# COMPOSITIONS AND METHODS FOR DETECTION OF LUNG CANCER

## CROSS-REFERENCE TO RELATED APPLICATIONS

[1] This application claims benefit of U.S. Provisional Application No. 63/297684 filed January 7, 2022, and U.S. Provisional Application No. 63/417299 filed October 18, 2022, the contents of which are hereby incorporated by reference in their entirety.

## BACKGROUND

[2] Early detection of cancer greatly increases the chance of successful treatment. However, many cancers including lung cancer still lack effective screening recommendations. Typical challenges for cancer-screening tests include limited sensitivity and specificity. A high rate of false-positive results can be of particular concern, as it can create difficult management decisions for clinicians and patients who would not want to unnecessarily administer (or receive) anti-cancer therapy that may potentially have undesirable side effects. Conversely, a high rate of false-negative results fails to satisfy the purpose of the screening test, as patients who need therapy are missed, resulting in a treatment delay and consequently a reduced possibility of success.

## SUMMARY

[3] The present disclosure, among other things, provides insights and technologies for achieving effective lung cancer screening from a biological sample. In some embodiments, such a biological sample is or comprises a bodily fluid-derived sample, e.g., in some embodiments a blood-derived sample. In some embodiments, the present disclosure, among other things, provides insights and technologies that are particularly useful for non-small cell lung cancer screening. In some embodiments, provided technologies are effective for detection of early stage lung cancers (*e.g.*, in some embodiments non-small cell lung cancer). In some embodiments, provided technologies are effective even when applied to populations comprising or consisting of asymptomatic individuals (*e.g.*, due to sufficiently high sensitivity and/or low rates of false positive and/or false negative results). In some embodiments, provided technologies are effective when applied to populations comprising or consisting of individuals (*e.g.*, asymptomatic individuals) without hereditary risk in developing lung cancer (*e.g.*, in some

embodiments non-small cell lung cancer). In some embodiments, provided technologies are effective when applied to populations comprising or consisting of symptomatic individuals (*e.g.*, individuals suffering from one or more symptoms of lung cancer). In some embodiments, provided technologies are effective when applied to populations comprising or consisting of individuals at risk for lung cancer (*e.g.*, individuals with hereditary and/or life-history associated risk factors for lung cancer). In some embodiments, provided technologies may be or include one or more compositions (*e.g.*, molecular entities or complexes, systems, cells, collections, combinations, kits, *etc.*) and/or methods (*e.g.*, of making, using, assessing, *etc.*), as will be clear to one skilled in the art reading the disclosure provided herein.

**[4]** In some embodiments, the present disclosure identifies the source of a problem with certain prior technologies including, for example, certain conventional approaches to detection and diagnosis of lung cancer. For example, the present disclosure appreciates that many conventional diagnostic assays, *e.g.*, X-ray imaging, sputum testing, low-dose CT scanning, and/ or molecular tests based on cell-free nucleic acids, serum biomarkers (*e.g.*, carcinoembryonic antigen (CEA), cytokeratin 19 fragment (CYFRA 21-1), neuron-specific enolase (NSE), progastrin-releasing peptide (ProGRP), and/or squamous cell carcinoma antigen (SCCA)), and/or bulk analysis of extracellular vesicles, can be time-consuming, costly, and/or lacking sensitivity and/or specificity sufficient to provide a reliable and comprehensive diagnostic assessment. In some embodiments, the present disclosure provides technologies (including systems, compositions, and methods) that solve such problems, among other things, by detecting co-localization of a target biomarker signature of lung cancer in individual nanoparticles having a size range of interest that includes extracellular vesicles, which comprises at least one extracellular vesicle-associated surface biomarker (including, e.g., membrane-bound and/or membrane-associated polypeptides) and at least one target biomarker selected from the group consisting of surface biomarkers, internal biomarkers, and RNA biomarkers. In some embodiments, the present disclosure provides technologies (including systems, compositions, and methods) that solve such problems, among other things, by detecting such target biomarker signature of lung cancer using a target entity detection approach that was developed by Applicant and described in US2020/0299780 and WO2020180741, which are based on interaction and/or co-localization of at least two or more target entities (*e.g.*, a target biomarker signature) in individual nanoparticles, e.g., extracellular vesicles.

**[5]** In some embodiments, extracellular vesicles for detection as described herein can be isolated from a bodily fluid of a subject by a size exclusion-based method. As will be understood by a skilled artisan, in some embodiments, a size exclusion-based method may provide a sample comprising nanoparticles having a size range of interest that includes nanoparticles. Accordingly, in some embodiments, provided technologies of the present disclosure encompass detection, in individual nanoparticles having a size range of interest (e.g., in some embodiments about 30 nm to about 1000 nm) that includes nanoparticles, of co-localization of at least two or more surface biomarkers (e.g., as described herein) that forms a target biomarker signature of lung cancer. A skilled artisan reading the present disclosure will understand that various embodiments described herein in the context of "nanoparticle(s)" can be also applicable in the context of "nanoparticles" as described herein.

**[6]** In some embodiments, the present disclosure, among other things, provides insights that screening of asymptotic individuals, *e.g.*, regular screening prior to or otherwise in absence of developed symptom(s), can be beneficial, and even important for effective management (*e.g.*, successful treatment) of lung cancer (*e.g.*, in some embodiments non-small cell lung cancer). In some embodiments, the present disclosure provides lung cancer screening systems that can be implemented to detect lung cancer (*e.g.*, in some embodiments non-small cell lung cancer), including early-stage cancer, in some embodiments in asymptomatic individuals. In some embodiments, provided technologies are implemented to achieve regular screening of asymptomatic individuals. The present disclosure provides, for example, compositions (*e.g.*, reagents, kits, components, *etc.*), and methods of providing and/or using them, including strategies that involve regular testing of one or more individuals (*e.g.*, symptomatic or asymptomatic individuals). The present disclosure defines usefulness of such systems, and provides compositions and methods for implementing them.

**[7]** In some embodiments, provided technologies achieve detection (*e.g.*, early detection, *e.g.*, in asymptomatic individual(s) and/or population(s)) of one or more features (*e.g.*, incidence, progression, responsiveness to therapy, recurrence, *etc.*) of lung cancer, with sensitivity and/or specificity (*e.g.*, rate of false positive and/or false negative results) appropriate to permit useful application of provided technologies to single-time and/or regular (*e.g.*, periodic) assessment. In some embodiments, provided technologies are useful in conjunction with regular medical examinations, such as but not limited to: physicals, general practitioner

visits, cholesterol/lipid blood tests, diabetes (type 2) screening, colonoscopies, blood pressure screening, thyroid function tests, prostate cancer screening, mammograms, HPV/Pap smears, and/or vaccinations. In some embodiments, provided technologies are useful in conjunction with treatment regimen(s); in some embodiments, provided technologies may improve one or more characteristics (*e.g.*, rate of success according to an accepted parameter) of such treatment regimen(s).

**[8]** In some aspects, provided are technologies for use in classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to lung cancer (*e.g.*, in some embodiments non-small cell lung cancer). In some embodiments, the present disclosure provides methods or assays for classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to lung cancer (*e.g.*, in some embodiments non-small cell lung cancer). In some embodiments, a provided method or assay comprises (a) detecting, in a biological sample (e.g., in some embodiments a bodily fluid-derived sample such as, e.g., but not limited to a blood-derived sample) from a subject in need thereof, nanoparticles (having a size range of interest that includes extracellular vesicles) expressing a target biomarker signature of lung cancer, the target biomarker signature comprising: at least one extracellular vesicle-associated surface biomarker (including, e.g., membrane-bound and/or membrane-associated polypeptides) and at least one target biomarker selected from the group consisting of: surface biomarkers (as described herein), intravesicular biomarkers (as described herein), and intravesicular RNA biomarkers (as described herein), wherein the surface biomarkers are selected from ADGRF1, ALCAM, ABCC3, ARSL, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, GPC4, HAS3, HS6ST2, KDELR3, KRTCAP3, IG1FR, LAMB3, LAPTM4B, LARGE2, LFNG, LSR, MAL2, MANEAL, MET, MSLN, MUC1, MUC21, NRCAM, PIGT, PODXL2, PRRG4, PRSS21, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, SLC44A4, SLC6A14, SLC7A7, SLC7A11, SMIM22, SMPDL3B, ST14, TACSTD2, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, UCHL1 and combinations thereof; the intravesicular biomarkers are selected from AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2,

SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and the intravesicular RNA (e.g., but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) biomarkers are selected from ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof; (b) comparing sample information indicative of level of the target biomarker signature-expressing nanoparticles in the biological sample to reference information including a reference threshold level; and (c) classifying the subject as having or being susceptible to lung cancer when the biological sample shows an elevated level of target biomarker signature-expressing nanoparticles relative to a classification cutoff referencing the reference threshold level.

**[9]** In some aspects, provided are technologies for use in classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to lung cancer (*e.g.*, in some embodiments non-small cell lung cancer). In some embodiments, the present disclosure provides methods or assays for classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to lung cancer (*e.g.*, in some embodiments non-small cell lung cancer). In some embodiments, a provided method or assay comprises (a) detecting, in a biological sample from a subject in need thereof, nanoparticles expressing a target biomarker signature of lung cancer (*e.g.*, in some embodiments non-small cell lung cancer), the target biomarker signature comprising: at least one nanoparticle-associated surface biomarker and at least one target biomarker selected from the group consisting of: surface biomarkers (as described herein), intravesicular biomarkers (as described herein), and intravesicular RNA biomarkers (as described herein); (b) comparing sample information indicative of level of the target biomarker signature-expressing nanoparticles in the biological sample to reference information including a reference threshold level; and (c) classifying the subject as having or being susceptible to lung

cancer when the biological sample shows an elevated level of target biomarker signature-expressing nanoparticles relative to a classification cutoff referencing the reference threshold level. In some embodiments, at least one such target biomarker is or comprises a surface biomarker selected from: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, Lewis X, Lewis Y/B antigen, LY6E, NOTCH2, NOTCH3, Phosphatidylserine, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA, Tn antigen, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof. In some embodiments, at least one such target biomarker is or comprises a surface biomarker selected from: ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3,

IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof. In some embodiments, at least one such target biomarker is or comprises an intravesicular biomarker selected from ABRACL, ACP5, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, AOC1, AP1M2, APOBEC3B, APOBEC3C, ARNTL2, ASF1B, AURKB, BAIAP2L1, BIRC5, C12orf45, C15orf48, C19orf33, C1S, C8orf4, CA9, CALML3, CAPNS2, CBLC, CCL19, CCNB2, CDC20, CDC45, CDCA4, CDCA5, CDK1, CDKN2A, CDKN2B, CENPW, CEP55, CES1, CHMP4C, CNN2, CPA3, CRABP2, CST1, CSTA, CTSC, CTSE, CYP2S1, DPYSL3, EFS, EGLN3, EHF, ELF3, ELF4, ENAH, ESRP1, ETV4, EVPL, FAM129B, FAM60A, FAM83A, FAM83D, FAM83H, FBP1, FERMT1, FOXA2, FOXE1, FOXM1, GBP6, GNA15, GPX2, GRHL2, GSTA1, HCK, HMGB3, HOXB7, ID1, IGF2BP2, IMPA2, IRF6, IVL, JUP, KIAA1522, KIF2C, KIFC1, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, LGALS3BP, LGALS7B, LSP1, MAGEA4, MAGEA6, MCM2, MDFI, MIF, MYBL2, MYH14, MZB1, NAPSA, NCF2, NNMT, NRARP, NUP210, NUSAP1, OSGIN1, PALLD, PITX1, PKP1, PKP3, PLEK, PLEK2, POSTN, PPP1R14C, PPP1R14D, PRAME, PTPN6, RBP1, RIN2, RIPK4, RPS4Y1, RRM2, S100A11, S100A14, S100A16, S100A2, S100P, SBK1, SCGB3A2, SERPINB13, SERPINB3, SERPINB5, SFTA2, SFTPA1, SFTPA2, SFTPB, SH3BP4, SNAI2, SOX2, SPI1, SPINK1, SPINT1, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, SULF1, SYK, SYTL1, TBC1D2, TEAD2, TEAD3, TFAP2C, TGFA, THBS2, TK1, TOP2A, TP63, TPD52, TPX2, TRIM29, TRIP13, UBE2C, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof. In some embodiments, at least one such target biomarker is or comprises an intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) selected from ABCA3,

ABCC1, ABCC3, ABRACL, ACP5, ADAM23, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, ANTXR1, AOC1, AP1M2, APOBEC3B, APOBEC3C, AQP3, AREG, ARNTL2, ARSL, ASF1B, ATP8B1, AURKB, B3GNT3, B3GNT5, BAIAP2L1, BCAM, BIK, BIRC5, C12orf45, C15orf48, C19orf33, C1S, C8orf4, CA12, CA9, CALML3, CAPNS2, CBLC, CCL19, CCL5, CCNB2, CD109, CD24, CD53, CD74, CD9, CDC20, CDC42EP1, CDC45, CDCA4, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CEACAM5, CEACAM6, CELSR1, CENPW, CEP55, CES1, CHMP4C, CLCA2, CLDN1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CNN2, COL17A1, CPA3, CRABP2, CST1, CSTA, CTSC, CTSE, CX3CL1, CXADR, CXCR4, CYBB, CYP2S1, CYP4F11, DAPL1, DMBT1, DPYSL3, DSC2, DSC3, DSG2, DSG3, DSP, EFNA1, EFS, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHB3, EPHX1, EPHX3, ESRP1, ETV4, EVA1A, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83A, FAM83D, FAM83H, FAT1, FAT2, FBLIM1, FBP1, FCER1G, FERMT1, FGFR2, FGFR3, FOLR1, FOXA2, FOXE1, FOXM1, FXYD3, GALNT3, GBP6, GJA1, GJB1, GJB2, GJB3, GJB5, GJB6, GNA15, GPC1, GPC3, GPC4, GPNMB, GPR87, GPRC5A, GPX2, GRHL2, GSTA1, HAS3, HCK, HMGB3, HOXB7, HS6ST2, ID1, IGF2BP2, IGSF9, IL2RG, IMPA2, IRF6, ITGA2, ITGA6, ITGB4, ITGB6, IVL, JAG2, JUP, KCNS3, KDELR3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, KRTCAP3, LAMB3, LAMP3, LAPTM5, LFNG, LGALS3BP, LGALS7B, LRP11, LRRC4, LSP1, LSR, LYPD3, MAGEA4, MAGEA6, MAL2, MANEAL, MAOA, MARCO, MCM2, MDFI, MET, MIF, MMP14, MPZL2, MSLN, MUC1, MUC21, MYBL2, MYH14, MYOF, MZB1, NAPSA, NCF2, NKG7, NNMT, NOTCH3, NRARP, NTRK2, NUP210, NUSAP1, OSGIN1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PERP, PIGR, PIGT, PITX1, PKP1, PKP3, PLEK, PLEK2, PLVAP, PMP22, PODXL2, POSTN, PPL, PPP1R14C, PPP1R14D, PRAME, PROM2, PRRG4, PRSS8, PTGES, PTGFRN, PTPN6, PTPRF, PTPRZ1, RAB25, RAB38, RAET1L, RARRES1, RBP1, RGS1, RHCG, RHOV, RIN2, RIPK4, ROS1, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, S100A2, S100P, SBK1, SCGB3A2, SCNN1A, SDC1, SERINC2, SERPINB13, SERPINB3, SERPINB5, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SH3BP4, SHISA2, SLC1A5, SLC2A1, SLC34A2, SLC40A1, SLC44A4, SLC6A14, SLC6A8, SLC7A7, SLC7A8, SMIM22, SMPDL3B, SNAI2, SOX2, SPI1, SPINK1, SPINT1, SPINT2, SPRR1A,

SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, TGFA, THBD, THBS2, TK1, TM4SF1, TMC4, TMC5, TMEM30B, TMEM45B, TMEM54, TMPRSS11D, TMPRSS11E, TMPRSS2, TMPRSS4, TNFRSF18, TNS4, TOP2A, TP53I11, TP63, TPD52, TPX2, TREM2, TRIM29, TRIP13, TSPAN1, TSPAN13, TSPAN6, TSPAN7, TSPAN8, TUSC3, TYROBP, UBE2C, UPK1B, VAMP8, VANGL2, WLS, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof.

**[10]** In some aspects, provided are technologies for use in classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to lung cancer (*e.g.*, in some embodiments non-small cell lung cancer such as, *e.g.*, lung adenocarcinoma (LUAD) and/or lung squamous cell carcinoma (LUSC)). In some embodiments, the present disclosure provides methods or assays for classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to lung cancer (*e.g.*, in some embodiments non-small cell lung cancer such as, *e.g.*, LUAD and/or LUSC)). In some embodiments, a provided method or assay comprises (a) detecting, in a biological sample from a subject in need thereof, nanoparticles expressing a target biomarker signature of lung cancer (*e.g.*, in some embodiments non-small cell lung cancer such as, e.g., LUAD and/or LUSC), the target biomarker signature comprising: at least one nanoparticle-associated surface biomarker and at least one target biomarker selected from the group consisting of: surface biomarkers (as described herein), intravesicular biomarkers (as described herein), and intravesicular RNA biomarkers (as described herein); (b) comparing sample information indicative of level of the target biomarker signature-expressing nanoparticles in the biological sample to reference information including a reference threshold level; and (c) classifying the subject as having or being susceptible to lung cancer (*e.g.*, in some embodiments non-small cell lung cancer such as, *e.g.,* LUAD and/or LUSC) when the biological sample shows an elevated level of target biomarker signature-expressing nanoparticles relative to a classification cutoff referencing the reference threshold level. In some embodiments, at least one such target biomarker is or comprises a surface biomarker selected from: ABCA3, ACBD3, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP1B1, ATP6AP2, BCAP31, B3GNT5, BSPRY, CDC42, CDH1, CDH3, CDKAL1, CELSR1, CIP2A, CISD2, CKAP4, CLPTM1L, CLSTN1, CPD, DPY19L1, DSG2, EGFR, EPCAM, FAM241B, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLPH3L, GRHL2, HACD3,

IER3IP1, IL1RAP, ITGA2, ITGB6, HAS3, KPNA2, KRTCAP3, LAD1, LAMB3, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LSR, MAL2, MAGT1, MARCKSL1, MET, MGAT1, MUC4, NCSTN, NECTIN4, NRAS, NRCAM, NUP210, PEX13, PIGN, PIGT, PLEC, PTPRF, PRSS21, SLC7A11, QSOX1, RAB25, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC35B2, SLC39A11, SMIM22, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TOMM22, TOR1AIP2, TRAM1, TTC33, VAMP8, VMA21, VRK2, VWA1, XPR1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, FLT4, GM3, HGF, IGF1R, IL6, LAG3, UCHL1, Lewis X, Lewis Y/B antigen, LY6E, NOTCH2, NOTCH3, Phosphatidylserine, TIGIT, TNFRSF10A, TNFRSF10B, TPBG, VEGFA, Tn antigen, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof. In some embodiments, at least one such target biomarker is or comprises a surface biomarker selected from: ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof. In some embodiments, at least one such target biomarker is or comprises an intravesicular biomarker selected from ABRACL, ACP5, AIF1, ALDH1A1, ALG1L, AP1M2, APOBEC3C, ASF1B, AURKB, BAIAP2L1, BIRC5, C12orf45, C15orf48, C19orf33, C1S, C8orf4, CBLC, CCL19, CCNB2, CDC20, CDCA5, CDK1, CDKN2A, CDKN2B, CEP55, CHMP4C, CNN2, CPA3, CRABP2, CST1, CTSC, DPYSL3, EGLN3, EHF, ELF3, ELF4, ENAH, ESRP1, ETV4, EVPL, FAM129B, FAM60A, FAM83A, FAM83H, GNA15, GRHL2, HCK, HMGB3, HOXB7, ID1, IMPA2, KIAA1522, KIF2C, KIFC1, KLF4, KLF5, KRT18, KRT19, KRT8, LGALS3BP, LSP1, MDFI, MIF, MYBL2, MYH14, MZB1, NCF2, NNMT, NUP210, NUSAP1, PALLD, PKP3, PLEK, PLEK2, POSTN, PTPN6, RIN2, RIPK4, RPS4Y1,

RRM2, S100A11, S100A14, S100A16, SBK1, SH3BP4, SPI1, SPINT1, SULF1, SYK, SYTL1, TBC1D2, TEAD2, TEAD3, TFAP2C, TGFA, THBS2, TK1, TOP2A, TPD52, TPX2, TRIP13, UBE2C, YAP1, ZC3H11A, ZNF217, and combinations thereof. In some embodiments, at least one such target biomarker is or comprises an intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) selected from ABCA3, ABRACL, ACP5, AIF1, ALDH1A1, ALG1L, ANTXR1, AP1M2, APOBEC3C, AQP3, AREG, ARSL, ASF1B, ATP8B1, AURKB, BAIAP2L1, BCAM, BIK, BIRC5, C12orf45, C15orf48, C19orf33, C1S, C8orf4, CBLC, CCL19, CCL5, CCNB2, CD24, CD53, CD74, CDC20, CDC42EP1, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CELSR1, CEP55, CHMP4C, CLDN4, CLDN7, CNN2, CPA3, CRABP2, CST1, CTSC, CX3CL1, CXADR, CXCR4, CYBB, DMBT1, DPYSL3, DSG2, EFNA1, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHX1, EPHX3, ESRP1, ETV4, EVA1A, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83A, FAM83H, FAT1, FBLIM1, FCER1G, GALNT3, GNA15, GPC4, GRHL2, HCK, HMGB3, HOXB7, HS6ST2, ID1, IL2RG, IMPA2, ITGA2, ITGB4, ITGB6, JAG2, KCNS3, KDELR3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT18, KRT19, KRT8, KRTCAP3, LAMB3, LAMP3, LAPTM5, LFNG, LGALS3BP, LRP11, LSP1, LSR, MAL2, MANEAL, MAOA, MARCO, MDFI, MET, MIF, MMP14, MPZL2, MYBL2, MYH14, MYOF, MZB1, NCF2, NKG7, NNMT, NOTCH3, NUP210, NUSAP1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PIGT, PKP3, PLEK, PLEK2, PLVAP, PMP22, PODXL2, POSTN, PPL, PROM2, PRSS8, PTGES, PTPN6, PTPRF, RAB25, RARRES1, RGS1, RHOV, RIN2, RIPK4, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, SBK1, SCNN1A, SDC1, SERINC2, SEZ6L2, SH3BP4, SHISA2, SLC1A5, SLC40A1, SLC7A7, SPI1, SPINT1, SPINT2, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, TGFA, THBS2, TK1, TM4SF1, TMC4, TMEM30B, TMEM54, TMPRSS11E, TMPRSS4, TNFRSF18, TOP2A, TP53I11, TPD52, TPX2, TREM2, TRIP13, TSPAN1, TSPAN13, TSPAN6, TSPAN7, TUSC3, TYROBP, UBE2C, VAMP8, WLS, YAP1, ZC3H11A, ZNF217, and combinations thereof.

**[11]** In some aspects, provided are technologies for use in classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to a certain type of lung cancer (*e.g.*, in particular embodiments lung adenocarcinoma (LUAD)). In some embodiments, the present

disclosure provides methods or assays for classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to a certain type of lung cancer (*e.g.*, in particular embodiments LUAD). In some embodiments, a provided method or assay comprises (a) detecting, in a biological sample from a subject in need thereof, nanoparticles expressing a target biomarker signature of a certain type of lung cancer (*e.g.*, in particular embodiments LUAD), the target biomarker signature comprising: at least one nanoparticle-associated surface biomarker and at least one target biomarker selected from the group consisting of: surface biomarkers (as described herein), intravesicular biomarkers (as described herein), and intravesicular RNA biomarkers (as described herein); (b) comparing sample information indicative of level of the target biomarker signature-expressing nanoparticles in the biological sample to reference information including a reference threshold level; and (c) classifying the subject as having or being susceptible to a certain type of lung cancer (e.g., in particular embodiments LUAD) when the biological sample shows an elevated level of target biomarker signature-expressing nanoparticles relative to a classification cutoff referencing the reference threshold level. In some embodiments, at least one such target biomarker that is particularly useful for classifying a subject as having or being susceptible to LUAD is or comprises a surface biomarker selected from: ABCC3, ACSL5, ATP11A, CD55, CEACAM5, CEACAM6, CLIC6, FOLR1, GOLM1, LPCAT1, MSLN, MUC1, NT5E, PLA2G4A, PLCH1, SLC34A2, SMPDL3B, TESC, TMC5, ERBB3, KDR, and combinations thereof. In some embodiments, at least one such target biomarker that is particularly useful for classifying a subject as having or being susceptible to LUAD is or comprises an intravesicular biomarker selected from AGR2, AOC1, CTSE, FBP1, FOXA2, KRT7, NAPSA, PPP1R14D, S100P, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, and combinations thereof. In some embodiments, at least one such target biomarker that is particularly useful for classifying a subject as having or being susceptible to LUAD is or comprises an intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) selected from ABCC3, AGR2, AOC1, B3GNT3, CEACAM5, CEACAM6, CLDN18, CLDN3, CLIC6, CTSE, FBP1, FOLR1, FOXA2, GJB1, GPRC5A, KRT7, MSLN, MUC1, MUC21, NAPSA, PIGR, PPP1R14D, ROS1, S100P, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SMIM22, SMPDL3B, SPINK1, TMC5, TMEM45B,

TMPRSS2, TSPAN8, and combinations thereof. In some embodiments, certain aforementioned target biomarkers are particularly useful in differentiating LUAD from LUSC.

**[12]** In some aspects, provided are technologies for use in classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to a certain type of lung cancer (*e.g.*, in particular embodiments lung squamous cell carcinoma (LUSC)). In some embodiments, the present disclosure provides methods or assays for classifying a subject (*e.g.*, an asymptomatic subject) as having or being susceptible to a certain type of lung cancer (*e.g.*, in particular embodiments LUSC). In some embodiments, a provided method or assay comprises (a) detecting, in a biological sample from a subject in need thereof, nanoparticles expressing a target biomarker signature of a certain type of lung cancer (*e.g.*, in particular embodiments LUSC), the target biomarker signature comprising: at least one nanoparticle-associated surface biomarker and at least one target biomarker selected from the group consisting of: surface biomarkers (as described herein), intravesicular biomarkers (as described herein), and intravesicular RNA biomarkers (as described herein); (b) comparing sample information indicative of level of the target biomarker signature-expressing nanoparticles in the biological sample to reference information including a reference threshold level; and (c) classifying the subject as having or being susceptible to a certain type of lung cancer (*e.g.*, in particular embodiments LUSC) when the biological sample shows an elevated level of target biomarker signature-expressing nanoparticles relative to a classification cutoff referencing the reference threshold level. In some embodiments, at least one such target biomarker that is particularly useful for classifying a subject as having or being susceptible to LUSC is or comprises a surface biomarker selected from: ABCC1, ATP11B, B4GALT4, CD109, CD9, CLCA2, CLDN1, CNTN1, CYP2S1, CYP4F11, CYP4F3, DSC2, DSC3, DSG3, EPHB3, FAT2, FBXO45, FERMT1, IGSF3, KLRG2, LAMC2, NECTIN1, PARL, PSMD2, PTDSS1, PTGFRN, RAB38, RAB6B, RAP2B, SLC2A1, TFRC, TMTC3, TRPV4, UGT1A6, UPK1B, XXYLT1, TNFSF18, and combinations thereof. In some embodiments, at least one such target biomarker that is particularly useful for classifying a subject as having or being susceptible to LUSC is or comprises an intravesicular biomarker selected from ADH7, AKR1C1, AKR1C2, AKR1C3, ALDH3A1, ALDH3B2, APOBEC3B, ARNTL2, CA9, CALML3, CAPNS2, CDC45, CDCA4, CENPW, CES1, CSTA, CYP2S1, EFS, FAM83D, FERMT1, FOXE1, FOXM1, GBP6, GPX2, GSTA1, IGF2BP2, IRF6, IVL, JUP, KRT13, KRT14, KRT15, KRT16, KRT17, KRT5, KRT6A, KRT6B, KRT6C, LGALS7B,

MAGEA4, MAGEA6, MCM2, NRARP, OSGIN1, PITX1, PKP1, PPP1R14C, PRAME, RBP1, S100A2, SERPINB13, SERPINB3, SERPINB5, SNAI2, SOX2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, TP63, TRIM29, ZNF750, and combinations thereof. In some embodiments, at least one such target biomarker that is particularly useful for classifying a subject as having or being susceptible to LUSC is or comprises an intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) selected from ABCC1, ADAM23, ADH7, AKR1C1, AKR1C2, AKR1C3, ALDH3A1, ALDH3B2, APOBEC3B, ARNTL2, B3GNT5, CA12, CA9, CALML3, CAPNS2, CD109, CD9, CDC45, CDCA4, CENPW, CES1, CLCA2, CLDN1, COL17A1, CSTA, CYP2S1, CYP4F11, DAPL1, DSC2, DSC3, DSG3, DSP, EFS, EPHB3, FAM83D, FAT2, FERMT1, FGFR2, FGFR3, FOXE1, FOXM1, FXYD3, GBP6, GJA1, GJB2, GJB3, GJB5, GJB6, GPC1, GPC3, GPNMB, GPR87, GPX2, GSTA1, HAS3, IGF2BP2, IGSF9, IRF6, ITGA6, IVL, JUP, KRT13, KRT14, KRT15, KRT16, KRT17, KRT5, KRT6A, KRT6B, KRT6C, LGALS7B, LRRC4, LYPD3, MAGEA4, MAGEA6, MCM2, NRARP, NTRK2, OSGIN1, PERP, PITX1, PKP1, PPP1R14C, PRAME, PRRG4, PTGFRN, PTPRZ1, RAB38, RAET1L, RBP1, RHCG, S100A2, SERPINB13, SERPINB3, SERPINB5, SLC2A1, SLC6A8, SLC7A8, SNAI2, SOX2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, THBD, TMPRSS11D, TNS4, TP63, TRIM29, UPK1B, VANGL2, ZNF750, and combinations thereof. In some embodiments, certain aforementioned target biomarkers are particularly useful in differentiating LUSC from LUAD.

**[13]** In some embodiments, methods or assays described herein may be performed for at least one more additional target biomarker signature (including, *e.g.*, at least one, at least two, at least three, or more additional target biomarker signatures). In some such embodiments, a classification cutoff may reference additional reference threshold level(s) corresponding to each additional target biomarker signature.

**[14]** In some embodiments, a nanoparticle-associated surface biomarker for use in a target biomarker signature of lung cancer used and/or described herein may be or comprise a tumor-specific biomarker and/or a tissue-specific biomarker (*e.g.*, a lung tissue-specific biomarker). In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise a non-specific marker, *e.g.*, it is present in one or more non-target tumors, and/or in one or more non-target tissues. In some embodiments, such a nanoparticle-associated surface

biomarker may be or comprise one or more of biomarkers selected from: ALCAM, B3GNT3, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CLDN3, CLDN4, DSG2, EGFR, EPCAM, FOLR1, GJB1, GJB2, IG1FR, LAMB3, MET, MSLN, MUC1, PIGT, PODXL2, ROS1, SDC1, SLC34A2, sTn antigen, Tn antigen, T antigen, SMPDL3B, ST14, TACSTD2, TMPRSS4, TNFRSF10B, TSPAN8, and combinations thereof. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise SLC34A2, CEACAM5, CEACAM6, and/or EPCAM. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise ALCAM, CD55, CDH1, CDH3, CD274 (PD-L1), CEACAM5, CEACAM6, DSG2, EGFR, EPCAM, FOLR1, IG1FR, MET, MSLN, MUC1, SLC34A2, sTn antigen, Tn antigen, T antigen, TACSTD2, TNFRSF10B, or combinations thereof. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise a SLC34A2 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise a CEACAM5 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise a CEACAM6 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise an EPCAM polypeptide. In some embodiments such a nanoparticle-associated surface biomarker may be or comprise a sTn antigen and/or a MUC1 polypeptide.

**[15]** In some embodiments, a target biomarker signature of lung cancer comprises a nanoparticle-associated surface biomarker (*e.g.*, ones described herein) and at least one (including, *e.g.*, 1, 2, 3, or more) additional target surface biomarker, which, in some embodiments, may be or comprise ALCAM, ABCC3, ARSL, B3GNT3, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FOLR1, GJB1, GJB2, GPC4, HS6ST2, IG1FR, KDELR3, KRTCAP3, LAMB3, LFNG, LSR, MANEAL, MET, MSLN, MUC1, MUC21, PIGT, PODXL2, PRRG4, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, ST14, sTn antigen, Tn antigen, T antigen, TACSTD2, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, or combinations thereof.

**[16]** In some embodiments, a nanoparticle-associated surface biomarker for use in a target biomarker signature of lung cancer used and/or described herein may be or comprise a tumor-specific biomarker and/or a tissue-specific biomarker (*e.g.*, a lung tissue-specific

biomarker). In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise a non-specific marker, *e.g.*, it is present in one or more non-target tumors, and/or in one or more non-target tissues. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise one or more of surface biomarkers described herein. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise one or more of biomarkers selected from: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, Lewis X, Lewis Y/B antigen, LY6E, NOTCH2, NOTCH3, Phosphatidylserine, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA, Tn antigen, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof.

**[17]** In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise a surface biomarker selected from: ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof.

**[18]** In some embodiments, a target biomarker signature of lung cancer comprises a nanoparticle-associated surface biomarker (*e.g.*, ones described herein) and at least one (including, *e.g.*, 1, 2, 3, or more) additional target surface biomarker, which, in some embodiments, may be or comprise Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC,

PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, Lewis X, Lewis Y/B antigen, LY6E, NOTCH2, NOTCH3, Phosphatidylserine, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA, Tn antigen, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof. In some embodiments, at least one (including, *e.g.*, 1, 2, 3, or more) additional target surface biomarker may be or comprise a surface biomarker selected from: ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof.

**[19]** In some embodiments, a target biomarker signature of lung cancer comprises a nanoparticle-associated surface biomarker (*e.g.*, ones described herein) and at least one target intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) biomarker, which, in some embodiments, may be or comprise at least one RNA transcript (e.g., mRNA transcript) encoded by a human gene as follows: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3,

LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, or combinations thereof.

**[20]** In some embodiments, a target biomarker signature of lung cancer comprises a nanoparticle-associated surface biomarker (*e.g.*, ones described herein) and at least one target intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) biomarker, which, in some embodiments, may be or comprise at least one RNA transcript (e.g., mRNA transcript) encoded by a human gene as follows: ABCA3, ABCC1, ABRACL, ACP5, ADAM23, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, ANTXR1, AP1M2, APOBEC3B, APOBEC3C, AQP3, AREG, ARNTL2, ASF1B, ATP8B1, AURKB, B3GNT5, BAIAP2L1, BCAM, BIK, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA12, CA9, CALML3, CAPNS2, CBLC, CCL19, CCL5, CCNB2, CD109, CD24, CD53, CD74, CD9, CDC20, CDC42EP1, CDC45, CDCA4, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CEACAM5, CEACAM6, CELSR1, CENPW, CEP55, CES1, CHMP4C, CLCA2, CLDN1, CLDN4, CLDN7, CNN2, COL17A1, CPA3, CRABP2, CSTA, CTSC, CTSE, CX3CL1, CXADR, CXCR4, CYBB, CYP2S1, CYP4F11, DAPL1, DPYSL3, DSC2, DSC3, DSG2, DSG3, DSP, EFNA1, EFS, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHB3, EPHX1, ESRP1, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83D, FAM83H, FAT1, FAT2, FBLIM1, FBP1, FCER1G, FERMT1, FGFR2, FGFR3, FOXE1, FOXM1, FXYD3, GALNT3, GBP6, GJA1, GJB2, GJB3, GJB5, GJB6, GNA15, GPC1, GPC3, GPNMB, GPR87, GPRC5A, GPX2, GRHL2, GSTA1, HAS3, HCK, HOXB7, ID1, IGF2BP2, IGSF9, IL2RG, IMPA2, IRF6, ITGA2, ITGA6, ITGB4, ITGB6, IVL, JAG2, JUP, KCNS3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, KRTCAP3, LAMP3, LAPTM5, LGALS7B, LRP11, LRRC4, LSP1, LSR, LYPD3, MAGEA4, MAGEA6, MAL2, MAOA, MARCO, MCM2, MDFI, MET, MMP14, MPZL2, MUC1, MYBL2, MYH14, MYOF, MZB1, NCF2, NKG7, NNMT, NOTCH3, NRARP, NTRK2, NUP210, NUSAP1, OSGIN1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PERP, PIGR,

PIGT, PITX1, PKP1, PKP3, PLEK, PLEK2, PLVAP, PMP22, POSTN, PPL, PPP1R14C, PRAME, PROM2, PRRG4, PRSS8, PTGES, PTGFRN, PTPN6, PTPRF, PTPRZ1, RAB25, RAB38, RAET1L, RARRES1, RBP1, RGS1, RHCG, RHOV, RIN2, RIPK4, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, S100A2, S100P, SCNN1A, SDC1, SERINC2, SERPINB13, SERPINB3, SERPINB5, SEZ6L2, SH3BP4, SHISA2, SLC1A5, SLC2A1, SLC34A2, SLC40A1, SLC6A8, SLC7A8, SNAI2, SOX2, SPI1, SPINT1, SPINT2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, THBD, THBS2, TK1, TM4SF1, TMC4, TMEM30B, TMEM54, TMPRSS11D, TMPRSS11E, TMPRSS4, TNFRSF18, TNS4, TOP2A, TP53I11, TP63, TPD52, TPX2, TREM2, TRIM29, TRIP13, TSPAN1, TSPAN13, TSPAN6, TSPAN7, TUSC3, TYROBP, UBE2C, UPK1B, VAMP8, VANGL2, WLS, YAP1, ZC3H11A, ZNF217, ZNF750, or combinations thereof.

**[21]** In some embodiments, a target biomarker signature of lung cancer comprises a nanoparticle-associated surface biomarker (*e.g.*, ones described herein) and at least one additional target intravesicular biomarker, which, in some embodiments, may be or comprise at least one polypeptide encoded by a human gene as follows: AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, or combinations thereof. In some embodiments, a target biomarker signature comprises at least one nanoparticle-associated surface biomarker, which is or comprises a SCL34A2 polypeptide and/or a CEACAM5 polypeptide; and at least one target biomarker SLC34A2, CEACAM5, CEACAM6 and/or EPCAM.

**[22]** In some embodiments, a target biomarker signature of lung cancer comprises a nanoparticle-associated surface biomarker (*e.g.*, ones described herein) and at least one additional target intravesicular biomarker, which, in some embodiments, may be or comprise at least one polypeptide encoded by a human gene as follows: ABRACL, ACP5, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, AP1M2, APOBEC3B, APOBEC3C, ARNTL2, ASF1B, AURKB, BAIAP2L1, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA9, CALML3, CAPNS2, CBLC, CCL19, CCNB2, CDC20, CDC45, CDCA4, CDCA5, CDK1, CDKN2A, CDKN2B, CENPW, CEP55, CES1, CHMP4C, CNN2, CPA3, CRABP2, CSTA, CTSC, CTSE, CYP2S1, DPYSL3, EFS, EGLN3, EHF, ELF3, ELF4,

ENAH, ESRP1, EVPL, FAM129B, FAM60A, FAM83D, FAM83H, FBP1, FERMT1, FOXE1, FOXM1, GBP6, GNA15, GPX2, GRHL2, GSTA1, HCK, HOXB7, ID1, IGF2BP2, IMPA2, IRF6, IVL, JUP, KIAA1522, KIF2C, KIFC1, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, LGALS7B, LSP1, MAGEA4, MAGEA6, MCM2, MDFI, MYBL2, MYH14, MZB1, NCF2, NNMT, NRARP, NUP210, NUSAP1, OSGIN1, PALLD, PITX1, PKP1, PKP3, PLEK, PLEK2, POSTN, PPP1R14C, PRAME, PTPN6, RBP1, RIN2, RIPK4, RPS4Y1, RRM2, S100A11, S100A14, S100A16, S100A2, S100P, SERPINB13, SERPINB3, SERPINB5, SH3BP4, SNAI2, SOX2, SPI1, SPINT1, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, SULF1, SYK, SYTL1, TBC1D2, TEAD2, TEAD3, TFAP2C, THBS2, TK1, TOP2A, TP63, TPD52, TPX2, TRIM29, TRIP13, UBE2C, YAP1, ZC3H11A, ZNF217, ZNF750, or combinations thereof.

**[23]** In some embodiments, a reference threshold level for use in a provided method or assay described herein is determined by levels of target biomarker signature-expressing nanoparticles observed in comparable samples from a population of non-lung cancer subjects.

**[24]** In some embodiments, a nanoparticle-associated surface biomarker included in a target biomarker signature may be detected using affinity agents (e.g., but not limited to antibody-based agents). In some embodiments, a nanoparticle-associated surface biomarker may be detected using a capture assay comprising an antibody-based agent. For example, in some embodiments, a capture assay for detecting the presence of a nanoparticle-associated surface biomarker in a nanoparticle may involve contacting a biological sample comprising nanoparticles with a capture agent directed to such a nanoparticle-associated surface biomarker. In some embodiments, such a capture agent may comprise a binding moiety directed to a nanoparticle-associated surface biomarker (*e.g.*, ones described herein), which may be optionally conjugated to a solid substrate. Without limitations, an exemplary capture agent for a nanoparticle-associated surface biomarker may be or comprising a solid substrate (*e.g.*, a magnetic bead) and a binding moiety (*e.g.*, an antibody agent) directed to a nanoparticle-associated surface biomarker.

**[25]** In some embodiments, a target biomarker included in a target biomarker signature may be detected using appropriate methods known in the art, which may vary with types of analytes to be detected (*e.g.*, surface analytes vs. intravesicular analytes; and/or polypeptides and/or glycoforms vs. carbohydrates vs. RNAs). For example, a person skilled in the art, reading

the present disclosure, will appreciate that a surface biomarker and/or an intravesicular biomarker may be detected using affinity agents (e.g., antibody-based agents)in some embodiments, while in some embodiments, an intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) biomarker may be detected using nucleic acid-based agents, *e.g.*, using quantitative reverse transcription PCR.

**[26]** For example, in some embodiments where a target biomarker is or comprises a surface biomarker and/or an intravesicular marker, such a target biomarker may be detected involving a proximity ligation assay, *e.g.*, following a capture assay (*e.g.*, ones as described herein) to capture nanoparticles that display a nanoparticle-associated surface biomarker (*e.g.*, ones as used and/or described herein). In some embodiments, such a proximity ligation assay may comprise contacting a biological sample comprising nanoparticles with a set of detection probes, each directed to a target biomarker, which set comprises at least two distinct detection probes, so that a combination comprising the nanoparticles and the set of detection probes is generated, wherein the two detection probes each comprise: (i) a binding moiety directed to a surface biomarker and/or an intravesicular biomarker; and (ii) an oligonucleotide domain coupled to the binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain. Such single-stranded overhang portions of the detection probes are characterized in that they can hybridize with each other when the detection probes are bound to the same nanoparticle. Such a combination comprising the nanoparticles and the set of detection probes is then maintained under conditions that permit binding of the set of detection probes to their respective targets on the nanoparticles such that their oligonucleotide domains are in close enough proximity to anneal to form a double-stranded complex. Such a double-stranded complex can be detected by contacting the double-stranded complex with a nucleic acid ligase to generate a ligated template; and detecting the ligated template. In some embodiments, a ligated template can be detected using quantitative PCR. The presence of such a ligated template is indicative of presence of nanoparticles that are positive for a target biomarker signature of lung cancer. While such a proximity ligation assay may perform better, *e.g.*, with higher specificity and/or sensitivity, than other existing proximity ligation assays, a person skilled in the art reading the present disclosure

will appreciate that other forms of proximity ligation assays that are known in the art may be used instead.

**[27]** In some embodiments where a target biomarker is or comprises an intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) marker, such a target biomarker may be detected involving a nucleic acid detection assay. In some embodiments, an exemplary nucleic acid detection assay may be or comprise reverse-transcription PCR.

**[28]** In some embodiments where a target biomarker is or comprises an intravesicular biomarker (*e.g.*, an intravesicular biomarker and/or an intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, etc.) biomarker), such a target biomarker may be detected involving, prior to a detection assay (*e.g.*, a proximity ligation assay as described herein), a sample treatment (*e.g.*, fixation and/or permeabilization) to expose such biomarker(s) within nanoparticles for subsequent detection.

**[29]** The present disclosure, among other things, recognizes that detection of a single lung cancer-associated serum protein or a plurality of lung cancer-associated biomarkers based on a bulk sample (*e.g.*, a bulk sample of extracellular vesicles), rather than at a resolution of a single extracellular vesicle, typically does not provide sufficient specificity and/or sensitivity in determination of whether a subject from whom the sample is obtained is likely to be suffering from or susceptible to lung cancer. The present disclosure, among other things, provides technologies, including systems, compositions, and/or methods, that solve such problems, including for example by specifically requiring that individual nanoparticles having a size range of interest that includes extracellular vesicles for detection be characterized by presence of a target biomarker signature comprising a combination of at least one or more extracellular vesicle-associated surface biomarkers and at least one or more target biomarkers. In particular embodiments, the present disclosure teaches technologies that require such individual nanoparticles be characterized by presence (*e.g.,* by expression) of such a target biomarker signature of lung cancer, while nanoparticles that do not comprise the target biomarker signature do not produce a detectable signal (*e.g.*, a level that is above a reference level, *e.g.*, by at least

10% or more, where in some embodiments, a reference level may be a level observed in a negative control sample, such as a sample in which individual nanoparticles comprising such a target biomarker signature are absent).

**[30]** As will be understood by a skilled artisan, in some embodiments, a sample comprising extracellular vesicles may also comprise nanoparticles having a size range of interest that includes extracellular vesicles. Thus, in some embodiments, provided technologies of the present disclosure in the context of extracellular vesicles are also applicable to detection of nanoparticles having a size range interest that includes extracellular vesicles. Accordingly, in some embodiments, the present disclosure, among other things, provides technologies for detection, in individual nanoparticles having a size range of interest (e.g., in some embodiments about 30 nm to about 1000 nm) that includes extracellular vesicles, of co-localization of at least two or more surface biomarkers (e.g., as described herein) that forms a target biomarker signature of lung cancer.

**[31]** In some embodiments, the present disclosure describes a method comprising steps of: (a) providing or obtaining a sample comprising nanoparticles having a size within the range of about 30 nm to about 1000 nm, which are isolated from a bodily fluid-derived sample (e.g., a blood-derived sample) of a subject; (b) detecting on surfaces of the nanoparticles co-localization of at least two surface biomarkers whose combined expression level has been determined to be associated with lung cancer, wherein the surface biomarkers are selected from (i) polypeptides encoded by human genes as follows: *PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1,*

*MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, NOTCH2, NOTCH3, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA*, and combinations thereof; and/or (ii) carbohydrate-dependent markers as follows: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, Phosphatidylserine, Tn antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, Lewis X, Lewis Y/B antigen, LY6E, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof; (c) comparing the detected co-localization level with the determined level; and (d) classifying the subject as having or being susceptible to lung cancer when the detected co-localization level is at or above the determined level.

**[32]** In some embodiments, the first surface biomarker and the second surface biomarker(s) are each independently selected from: (i) polypeptides encoded by human genes as follows: *PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1,*

*MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, NOTCH2, NOTCH3, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA*, and combinations thereof; and/or (ii) carbohydrate-dependent markers as follows: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, Phosphatidylserine, Tn antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, Lewis X, Lewis Y/B antigen, LY6E, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof.

**[33]** In some embodiments, the first surface biomarker and the second surface biomarker(s) are each independently selected from: ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof.

**[34]** Accordingly, in some embodiments, technologies provided herein can be useful for detection of incidence or recurrence of lung cancer in a subject and/or across a population of subjects. In some embodiments, a target biomarker signature may be selected for detection of lung cancer. In some embodiments, a target biomarker signature may be selected for detection

of a specific category of lung cancer, including, *e.g.*, but not limited to lung adenocarcinoma, small cell lung cancer, squamous and transitional cell lung cancer, large cell lung cancer, non-small cell carcinoma, other specified carcinomas, sarcomas, and other specified types of lung cancer as known in the art (see, *e.g.*, SEER Cancer Statistics Review 1975-2017). In some embodiments, technologies provided herein can be used periodically (*e.g.*, every year) to screen a human subject or across a population of human subjects for early-stage lung cancer or lung cancer recurrence.

**[35]** In some embodiments, a subject that is amenable to technologies provided herein for detection of incidence or recurrence of lung cancer may be an asymptomatic human subject and/or across an asymptomatic population. Such an asymptomatic subject may be a subject who has a family history of lung cancer, who has been previously treated for lung cancer, who is at risk of lung cancer recurrence after cancer treatment, who is in remission after lung cancer treatment, and/or who has been previously or periodically screened for the presence of lung cancer by chest X-ray, sputum analysis, low dose CT, and/or the presence of at least one lung cancer serum biomarker, *e.g.*, but not limited to CEA, CYFRA 21-1, NSE, ProGRP, and/or SCCA serum proteins. In some embodiments, such an asymptomatic subject may be a subject who is determined to have a normal medical diagnosis result from, *e.g.*, chest X-ray, sputum analysis, low dose CT analysis, or serum CEA, CYFRA 21-1, NSE, ProGRP, and/or SCCA levels. In some embodiments, such an asymptomatic subject may be a subject who is determined to have an abnormal medical diagnosis result from, *e.g.*, chest X-ray, sputum analysis, low dose CT analysis, and/or a serum level of CEA, CYFRA 21-1, NSE, ProGRP, and/or SCCA level when compared to results as typically observed in non-lung cancer subjects and/or normal healthy subjects. Alternatively, in some embodiments, an asymptomatic subject may be a subject who has not been previously screened for lung cancer, who has not been diagnosed for lung cancer, and/or who has not previously received lung cancer therapy.

**[36]** In some embodiments, a subject or population of subjects may be selected based on one or more characteristics such as age, race, geographic location, genetic history, medical history, personal and/or medical history (*e.g.*, smoking, alcohol, drugs, carcinogenic agents, diet, obesity, diabetes, physical activity, sun exposure, radiation exposure, exposure to infectious agents such as viruses, and/or occupational hazard).

**[37]** In some embodiments, technologies provided herein can be useful for selecting surgery or therapy for a subject who is suffering from or susceptible to lung cancer. In some embodiments, a lung cancer therapy and/or an adjunct therapy can be selected in light of findings based on technologies provided herein.

**[38]** In some embodiments, technologies provided herein can be useful for monitoring and/or evaluating efficacy of therapy administered to a subject (*e.g.*, a lung cancer subject).

**[39]** In some embodiments, the present disclosure provides technologies for managing patient care, *e.g.,* for one or more individual subjects and/or across a population of subjects. To give but a few examples, in some embodiments, the present disclosure provides technologies that may be utilized in screening (*e.g.,* temporally or incidentally motivated screening and/or non-temporally or incidentally motivated screening, *e.g.,* periodic screening such as annual, semi-annual, bi-annual, or with some other frequency). For example, in some embodiments, provided technologies for use in temporally motivated screening can be useful for screening one or more individual subjects or across a population of subjects (*e.g.*, asymptomatic subjects) who are older than a certain age (*e.g.*, over 30, 35, 40, 45, 50, 55, 60, 65, 70, or older). For example, in some embodiments, provided technologies for use in temporally motivated screening can be useful for screening one or more individual subjects or across a population of subjects (*e.g.*, asymptomatic subjects) who have a cigarette pack-year history greater than a certain number (*e.g.*, 5 pack years, 10 pack years, 15 pack years, 20 pack years, 25 pack years, 30 pack years, and/or greater than 35 pack years; 1 pack year is equal to 1 pack of cigarettes smoked per day for one year, while 2 packs smoked per day for one year would equal 2 pack years, or ½ pack smoked per day for two years would equal 1 pack years, etc.). In some embodiments, provided technologies for use in incidentally motivated screening can be useful for screening individual subjects who may have experienced an incident or event that motivates screening for lung cancer as described herein. For example, in some embodiments, an incidental motivation relating to determination of one or more indicators of cancer or susceptibility thereto may be or comprise, *e.g.*, an incident based on their family history (*e.g.*, a close relative such as blood-related relative was previously diagnosed for lung cancer), identification of one or more risk factors associated with lung cancer (*e.g.*, life history risk factors including but not limited to, *e.g.,* smoking, alcohol, diet, obesity, occupational hazard, *etc.*) and/or prior incidental findings from genetic tests (*e.g.*, genome sequencing), and/or imaging diagnostic tests (*e.g.*, X-ray, ultrasound, computerized tomography

(CT), low dose CT, and/or magnetic resonance imaging (MRI) scans), development of one or more signs or symptoms characteristic of lung cancer (*e.g.,* abnormal imaging results, and/or symptoms potentially indicative of lung cancer, *etc*.).

**[40]** In some embodiments, provided technologies for managing patient care can inform treatment and/or payment (*e.g.,* reimbursement for treatment) decisions and/or actions. For example, in some embodiments, provided technologies can provide determination of whether individual subjects have one or more indicators of incidence or recurrence of lung cancer, thereby informing physicians and/or patients when to initiate therapy in light of such findings. Additionally or alternatively, in some embodiments, provided technologies can inform physicians and/or patients of treatment selection, *e.g.*, based on findings of specific responsiveness biomarkers (*e.g.*, lung cancer responsiveness biomarkers). In some embodiments, provided technologies can provide determination of whether individual subjects are responsive to current treatment, *e.g.*, based on findings of changes in one or more levels of molecular targets associated with lung cancer, thereby informing physicians and/or patients of efficacy of such therapy and/or decisions to maintain or alter therapy in light of such findings.

**[41]** In some embodiments, provided technologies can inform decision making relating to whether health insurance providers reimburse (or not), *e.g.*, for (1) screening itself (*e.g.*, reimbursement available only for periodic/regular screening or available only for temporally and/or incidentally motivated screening); and/or for (2) initiating, maintaining, and/or altering therapy in light of findings by provided technologies. For example, in some embodiments, the present disclosure provides methods relating to (a) receiving results of a screening as described herein and also receiving a request for reimbursement of the screening and/or of a particular therapeutic regimen; (b) approving reimbursement of the screening if it was performed on a subject according to an appropriate schedule or response to a relevant incident and/or approving reimbursement of the therapeutic regimen if it represents appropriate treatment in light of the received screening results; and, optionally (c) implementing the reimbursement or providing notification that reimbursement is refused. In some embodiments, a therapeutic regimen is appropriate in light of received screening results if the received screening results detect a biomarker that represents an approved biomarker for the relevant therapeutic regimen (*e.g.,* as may be noted in a prescribing information label and/or via an approved companion diagnostic). Alternatively or additionally, the present disclosure contemplates reporting systems (*e.g.,*

implemented via appropriate electronic device(s) and/or communications system(s)) that permit or facilitate reporting and/or processing of screening results, and/or of reimbursement decisions as described herein.

**[42]** Some aspects provided herein relate to systems and kits for use in provided technologies. In some embodiments, a system or kit may comprise detection agents for a tumor biomarker signature of lung cancer (*e.g.*, ones described herein). In some embodiments, such a system or kit may comprise a capture agent for an extracellular vesicle-associated surface biomarker present in nanoparticles associated with lung cancer (*e.g.*, ones used and/or described herein); and (b) at least one or more detection agents directed to one or more target biomarkers of a target biomarker signature of lung cancer, which may be or comprise additional surface biomarker(s) (*e.g.*, ones as used and/or described herein), intravesicular biomarker(s) (*e.g.*, ones as used and/or described herein), and/or intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) biomarker(s)(*e.g.*, ones as used and/or described herein).

**[43]** In some embodiments, a capture agent included in a system and/or kit may comprise a binding moiety directed to an extracellular vesicle-associated surface biomarker (*e.g.*, ones described herein). In some embodiments, such a binding moiety may be conjugated to a solid substrate, which in some embodiments may be or comprise a solid substrate. In some embodiments, such a solid substrate may be or comprise a magnetic bead. In some embodiments, an exemplary capture agent included in a provided system and/or kit may be or comprise a solid substrate (*e.g.*, a magnetic bead) and an affinity reagent (e.g., but not limited to an antibody agent) directed to an extracellular vesicle-associated surface biomarker conjugated thereto.

**[44]** In some embodiments where a target biomarker includes a surface biomarker and/or an intravesicular biomarker, a system and/or kit may include detection agents for performing a proximity ligation assay (*e.g.*, ones as described herein). In some embodiments, such detection agents for performing a proximity ligation assay may comprise a set of detection probes, each directed to a target biomarker of a target biomarker signature, which set comprises at least two detection probes, wherein the two detection probes each comprise: (i) a polypeptide-binding moiety directed to a target biomarker; and (ii) an oligonucleotide domain coupled to the binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the

single-stranded overhang portions of the detection probes are characterized in that they can hybridize to each other when the detection probes are bound to the same nanoparticle.

**[45]** In some embodiments, a provided system and/or kit may comprise a plurality (*e.g.*, 2, 3, 4, 5, or more) of sets of detection probes, each set of which comprises two or more (*e.g.*, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more) detection probes. In some embodiments, at least one set of detection probes may be directed to detection for lung cancer. For example, in some embodiments, a provided system and/kit may comprise at least one set for detection probes for detection of lung cancer and at least one set of detection probes for detection of a different cancer (*e.g.*, pancreatic cancer). In some embodiments, two or more detection probes may be directed to different categories of lung cancer, including, *e.g.*, but not limited to lung adenocarcinoma lung cancer, small cell lung cancer, squamous and transitional cell lung cancer, Large cell lung cancer, non-small cell carcinoma lung cancer, other specified carcinoma lung cancer, sarcoma lung cancer, and other specified types of lung cancer as known in the art (see, *e.g.*, SEER Cancer Statistics Review 1975-2017). In some embodiments, two or more sets may be directed to detection of lung cancer of different categories. In some embodiments, two or more sets may be directed to detection of lung cancer of the same category. In some embodiments, two or more sets may be directed to detection of lung cancer of different stages. In some embodiments, two or more sets may be directed to detection of lung cancer of the same stage.

**[46]** In some embodiments, detection probes in a provided kit may be provided as a single mixture in a container. In some embodiments, multiple sets of detection probes may be provided as individual mixtures in separate containers. In some embodiments, each detection probe is provided individually in a separate container.

**[47]** In some embodiments where a target biomarker includes an intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) biomarker, such a system and/or kit may include detection agents for performing a nucleic acid detection assay. In some embodiments, such a system and/or kit may include detection agents for performing a quantitative reverse-transcription PCR, for example, which may comprise primers directed to intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan

noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) target(s).

**[48]** A skilled artisan reading the present disclosure will understand that a system or kit for detection of extracellular vesicles can also be employed to detect nanoparticles having a size range of interest that includes extracellular vesicles. Accordingly, in some embodiments, a system or kit may comprise (i) a capture agent for a first surface biomarker of a lung cancer-associated biomarker signature (e.g., as described herein) present on the surface of nanoparticles having a size range of interest that includes extracellular vesicles; and (ii) at least one or more detection agents directed to a second surface biomarker of the lung cancer-specific biomarker signature. In some embodiments, such nanoparticles have a size within the range of about 30 nm to about 1000 nm.

**[49]** In some embodiments, the present disclosure describes a kit for detection of lung cancer comprising: (a) a capture agent comprising a target-capture moiety directed to a first surface biomarker; and (b) at least one set of detection probes, which set comprises at least two detection probes each directed to a second surface biomarker, wherein the detection probes each comprise: (i) a target binding moiety directed at the second surface biomarker; and (ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same nanoparticle having a size within the range of about 30 nm to about 1000 nm; wherein at least the first surface biomarker and the second surface biomarker form a target biomarker signature determined to be associated with lung cancer, and wherein the first and second surface biomarkers are each independently selected from: (i) polypeptides encoded by human genes as follows: *PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1,*

*CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, NOTCH2, NOTCH3, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA,* and combinations thereof; and/or (ii) carbohydrate-dependent markers as follows: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, Phosphatidylserine, Tn antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, Lewis X, Lewis Y/B antigen, LY6E, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof.

**[50]** In some embodiments, the first surface biomarker and the second surface biomarker(s) are each independently selected from: (i) polypeptides encoded by human genes as follows: *PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX,*

*GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, NOTCH2, NOTCH3, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA,* and combinations thereof; and/or (ii) carbohydrate-dependent markers as follows: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, Phosphatidylserine, Tn antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, Lewis X, Lewis Y/B antigen, LY6E, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof.

**[51]** In some embodiments, the first surface biomarker and the second surface biomarker(s) are each independently selected from: ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof.

**[52]** In some embodiments, a provided system and/or kit may comprise at least one chemical reagent, *e.g.*, to process a sample and/or nanoparticles (including, e.g., in some embodiments extracellular vesicles) therein. In some embodiments, a provided system and/or kit may comprise at least one chemical reagent to process nanoparticles (including, e.g., in some embodiments extracellular vesicles) in a sample, including, *e.g.*, but not limited to a fixation agent, a permeabilization agent, and/or a blocking agent. In some embodiments, a provided system and/or kit may comprise a nucleic acid ligase and/or a nucleic acid polymerase. In some embodiments, a provided system and/or kit may comprise one or more primers and/or probes. In some embodiments, a provided system and/or kit may comprise one or more pairs of primers, for example for PCR, *e.g.*, quantitative PCR (qPCR) reactions. In some embodiments, a provided system and/or kit may comprise one or more probes such as, for example, hydrolysis probes which may in some embodiments be designed to increase the specificity of qPCR (*e.g.*, TaqMan probes). In some embodiments, a provided system and/or kit may comprise one or more multiplexing probes, for example as may be useful when simultaneous or parallel qPCR reactions are employed (*e.g.*, to facilitate or improve readout).

**[53]** In some embodiments, a provided system and/or kit can be used for screening (*e.g.*, regular screening) and/or other assessment of individuals (*e.g.*, asymptomatic or symptomatic subjects) for detection (*e.g.*, early detection) of lung cancer. In some embodiments, a provided system and/or kit can be used for screening and/or other assessment of individuals susceptible to lung cancer (*e.g.*, individuals with a known genetic, environmental, or experiential risk, *etc.*). In some embodiments, provided system and/or kits can be used for monitoring recurrence of lung cancer in a subject who has been previously treated. In some embodiments, provided systems and/or kits can be used as a companion diagnostic in combination with a therapy for a subject who is suffering from lung cancer. In some embodiments, provided systems and/or kits can be used for monitoring or evaluating efficacy of a therapy administered to a subject who is suffering from lung cancer. In some embodiments, provided systems and/or kits can be used for selecting a therapy for a subject who is suffering from lung cancer. In some embodiments, provided systems and/or kits can be used for making a therapy decision and/or selecting a therapy for a subject with one or more symptoms (*e.g.*, non-specific symptoms) associated with lung cancer.

**[54]** Complexes formed by performing methods described herein and/or using systems and/or kits described herein are also within the scope of disclosure. For example, in some embodiments, a complex comprising: (a) a nanoparticle expressing a target biomarker signature, at least two of which include at least one nanoparticle-associated surface biomarker and at least one target biomarker selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers, wherein the surface biomarkers are selected from ADGRF1, ALCAM, ABCC3, ARSL, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, GPC4, HAS3, HS6ST2, KDELR3, KRTCAP3, IG1FR, LAMB3, LAPTM4B, LARGE2, Lewis X antigen, LFNG, LSR, MAL2, MANEAL, MET, MSLN, MUC1, MUC21, NRCAM, PIGT, PODXL2, PRRG4, PRSS21, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, sTn antigen, Tn antigen, T antigen, SLC44A4, SLC6A14, SLC7A7, SLC7A11, SMIM22, SMPDL3B, ST14, TACSTD2, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, UCHL1, and combinations thereof; intravesicular biomarkers are selected from: AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and the intravesicular RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) biomarkers are selected from ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof, wherein the nanoparticle is immobilized onto a solid substrate comprising a binding moiety directed to such a nanoparticle-associated surface biomarker. In some embodiments, such a complex further comprises at least two detection probes directed to at least one target biomarker

of a target biomarker signature present in the nanoparticle, wherein each detection probe is bound to a respective target biomarker and each comprises: (i) a binding moiety directed to the target biomarker; and (ii) an oligonucleotide domain coupled to the binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the detection probes are hybridized to each other.

**[55]** In some embodiments, a nanoparticle-associated surface biomarker present in a nanoparticle that forms a complex comprises one or more of surface biomarkers described herein. In some embodiments, such a nanoparticle-associated surface biomarker may be or comprise one or more of biomarkers selected from: ALCAM, B3GNT3, CDCP1, CDHA1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CLDN3, CLDN4, DSG2, EGFR, EPCAM, FOLR1, GJB1, GJB2, IG1FR, LAMB3, MET, MSLN, MUC1, PIGT, PODXL2, ROS1, SDC1, SLC34A2, SMPDL3B, ST14, sTn antigen, Tn antigen, T antigen, TACSTD2, TMPRSS4, TSPAN8, TNFRSF10B, and combinations thereof. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a SLC34A2 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a CEACAM5 polypeptide. In some embodiments, such an nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a CEACAM6 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise an EPCAM polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise an ALCAM polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a CD55 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a CDH1 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a CDH3 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target

biomarker signature may be or comprise a CD274 (PD-L1) polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a DSG2 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a EGFR polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a FOLR1 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a IG1FR polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a MET polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a MSLN polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a MUC1 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a sTn antigen polypeptide glycosylation. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a Tn antigen polypeptide glycosylation. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a T antigen polypeptide glycosylation. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a TACSTD2 polypeptide. In some embodiments, such a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a TNFRSF10B polypeptide.

**[56]** In some embodiments, a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a MUC1 polypeptide. In some embodiments, a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target biomarker signature may be or comprise a Lewis Y antigen. In some embodiments, a nanoparticle-associated surface biomarker and/or a surface

biomarker included in a target biomarker signature may be or comprise a sLex antigen. In some embodiments, a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target-biomarker signature may be or comprise a T antigen. In some embodiments, a nanoparticle-associated surface biomarker and/or a surface biomarker included in a target-biomarker signature may be or comprise a Tn antigen.

**[57]** Also within the scope of the present disclosure is a complex comprising: a nanoparticle having a size range of interest that includes nanoparticles, and comprising a lung cancer-specific biomarker signature, which includes at least two surface biomarkers described herein, wherein the nanoparticle is immobilized onto a solid substrate comprising a binding moiety directed to a first surface biomarker of a lung cancer-specific biomarker signature. In some embodiments, such a complex is also bound to at least two detection probes each directed to a surface biomarker (which can be the same or different surface biomarker(s)) of the lung cancer-specific biomarker signature, wherein each detection probe is bound to a respective surface biomarker and each comprises: (i) a binding moiety directed to the surface biomarker; and (ii) an oligonucleotide domain coupled to the binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the detection probes are hybridized to each other.

**[58]** In some embodiments, the present disclosure describes a complex comprising: (a) a nanoparticle having a size within the range of about 30 nm to about 1000 nm and comprising at least a first surface biomarker and a second surface biomarker on its surface, which combination is determined to be a target biomarker signature for lung cancer, wherein the first surface biomarker and the second surface biomarker are each independently selected from: (i) polypeptides encoded by human genes as follows: *PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3,*

*FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, NOTCH2, NOTCH3, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA,* and combinations thereof; and/or (ii) carbohydrate-dependent markers as follows: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, Phosphatidylserine, Tn antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, Lewis X, Lewis Y/B antigen, LY6E, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof; (b) a solid substrate comprising a target-capture moiety directed to the first surface biomarker; wherein the target-capture moiety binds to the first surface biomarker of the nanoparticle such that the nanoparticle is immobilized on the solid substrate; and (c) at least a first detection probe and a second detection probe each bound to the nanoparticle, wherein each detection probe comprises: (i) a target binding moiety directed to the second surface biomarker; and (ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the first and second detection probes are hybridized to each other.

**[59]** In some embodiments, the first surface biomarker and the second surface biomarker(s) are each independently selected from: (i) polypeptides encoded by human genes as follows: *PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB,*

*SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, NOTCH2, NOTCH3, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA,* and combinations thereof; and/or (ii) carbohydrate-dependent markers as follows: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, Phosphatidylserine, Tn antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, Lewis X, Lewis Y/B antigen, LY6E, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof.

**[60]** In some embodiments, the first surface biomarker and the second surface biomarker(s) are each independently selected from: ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2,

CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof.

**[61]** These, and other aspects encompassed by the present disclosure, are described in more detail below and in the claims.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[62]** ***FIG. 1*** is a schematic diagram illustrating an exemplary workflow of profiling individual nanoparticles, e.g., extracellular vesicles. The figure shows purification of EVs from plasma using size exclusion chromatography (SEC) and immunoaffinity capture of EVs displaying a specific EV-associated surface marker (***Panel A***); detection of co-localized target markers (*e.g.*, intravesicular biomarkers or surface biomarkers) on captured EVs using a target entity detection assay according to some embodiments described herein (***Panel B***).

**[63]** ***FIG. 2*** is a schematic diagram illustrating a target entity detection assay according to some embodiments described herein. In some embodiments, a target entity detection assay uses a combination of detection probes, which combination is specific for detection of cancer. In some embodiments, a duplex system includes a first detection probe for a target biomarker 1 and a second detection probe for a target biomarker 2 are added to a sample comprising a biological entity (*e.g.*, extracellular vesicle). In some embodiments, detection probes each comprise a target binding moiety (*e.g.*, an affinity agent such as, e.g., an antibody agent against a target biomarker) coupled to an oligonucleotide domain, which comprises a double-stranded portion and a single-stranded overhang extended from one end of the oligonucleotide domain. A detection signal is generated when distinct target binding moieties (*e.g.*, affinity agents such as, e.g., antibody agents against target biomarker 1 and target biomarker 2, respectively) of the first and second detection probes are localized to the same biological entity (*e.g.*, an extracellular vesicle) in close proximity such that the corresponding single-stranded overhangs hybridize to each other, thus allowing ligation of their oligonucleotide domains to occur. For example, a control entity (*e.g.*, a biological entity from a healthy subject sample) does not express one or both of target biomarker 1 and target biomarker 2, so no

detection of signal can be generated. However, when a biological entity from a cancer sample (*e.g.*, a lung cancer sample) expresses target biomarker 1 and target biomarker 2, and the target biomarkers are present within a short enough distance of each other in the same biological entity (*e.g.*, extracellular vesicle), a detection signal is generated.

**[64]** ***FIG. 3*** shows a graphical representation of the prevalence of different lung and bronchus cancer histological subtypes. Data obtained from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program Cancer Statistics Review 1975-2017 the entire contents of which are incorporated herein by reference. Category "Other" includes large cell carcinoma, non-small cell carcinoma, other specified carcinomas, carcinoma NOS, sarcoma, and other specified types. Determination and grouping of histological classifications are noted within the review.

**[65]** ***FIG. 4*** is a graphical representation of the population demographics of a pilot patient cohort. A total of 39 patient plasma samples obtained from this cohort were analyzed, an overview of age, sex, and cohort size for the patient lung cancer diagnostic stage are shown. Cohort samples were subsequently evaluated by exemplary assays described herein.

**[66]** ***FIG. 5*** is a graphical representation of an exemplary lung adenocarcinoma (LUAD) diagnostic assay as described herein. Normalized signals of healthy controls and LUAD patient cohorts using SLC34A2 antibody based EV capture with CEACAM6 + CEACAM6 detection probes. The horizontal cutoff line represents a 100% specificity threshold. Sensitivities of 16.7% for stage I LUAD, 60% for stage II LUAD, 50% for stage III LUAD, and 100% for stage IV LUAD were achieved.

**[67]** ***FIG. 6*** is a graphical representation of an exemplary lung adenocarcinoma diagnostic assay as described herein. Normalized signals of healthy controls and LUAD patient cohorts using SLC34A2 antibody based EV capture with CEACAM6 + EPCAM detection probes. The horizontal cutoff line represents a 100% specificity threshold. Sensitivities of 20% for stage II LUAD, 50% for stage III LUAD, and 75% for stage IV LUAD were achieved.

**[68]** ***FIG. 7*** is a graphical representation of an exemplary lung adenocarcinoma diagnostic assay as described herein. Normalized signals of healthy controls and LUAD patient cohorts using CEACAM5 antibody based EV capture with CEACAM6 + SLC34A2 detection probes. The horizontal cutoff line represents a 100% specificity threshold. Sensitivities of 16.7%

for stage I LUAD, 20% for stage II LUAD, 50% for stage III LUAD, and 75% for stage IV LUAD were achieved.

**[69]** ***FIG. 8*** is a graphical representation of the correlation between exemplary lung adenocarcinoma diagnostic assays as described herein. Signal from SLC34A2 antibody based capture with CEACAM6 + CEACAM6 detection probes is depicted along the x-axis, while signal from SLC34A2 antibody based capture with CEACAM6 + EPCAM detection probes is depicted along the y-axis. Correlations were determined using the Pearson product-moment correlation coefficient. Strong correlations were observed, especially for stage III and stage IV LUAD samples.

**[70]** ***FIG. 9*** is a graphical representation of the correlation between exemplary Lung Adenocarcinoma diagnostic assays as described herein. Signal from SLC34A2 antibody based capture with CEACAM6 + CEACAM6 detection probes is depicted along the x-axis, while signal from CEACAM5 antibody based capture with CEACAM6 + SLC34A2 detection probes is depicted along the y-axis. Correlations were determined using the Pearson product-moment correlation coefficient. Strong correlations were observed, especially for stage III and stage IV LUAD samples.

**[71]** ***FIG. 10*** is a graphical representation of the correlation between exemplary Lung Adenocarcinoma diagnostic assays as described herein. Signal from SLC34A2 antibody based capture with CEACAM6 + EPCAM detection probes is depicted along the x-axis, while signal from CAECAM5 antibody based capture with CEACAM6 + SLC34A2 detection probes is depicted along the y-axis. Correlations were determined using the Pearson product-moment correlation coefficient. Strong correlations were observed, especially for stage III and stage IV LUAD samples.

**[72]** ***FIG. 11*** is a graphical representation of the population demographics of an expanded patient cohort. A total of 138 patient plasma samples obtained from this cohort were analyzed, an overview of age, sex, and cohort size for the patient lung cancer diagnostic stage are shown. Cohort samples were subsequently evaluated by exemplary assays described herein.

**[73]** ***FIG. 12*** is a graphical representation of an exemplary lung adenocarcinoma diagnostic assay as described herein. Normalized signals of healthy controls and LUAD patient cohorts using SLC34A2 antibody based EV capture with CEACAM6 + CEACAM6 detection

probes. The horizontal cutoff line represents a 99.9% specificity threshold. Sensitivities of 50% for stage II LUAD, 55.5% for stage III LUAD, and 71.4% for stage IV LUAD were achieved.

**[74]** ***FIG. 13*** is a schematic diagram illustrating a target entity detection assay according to some embodiments described herein. The figure shows an exemplary triplex target entity detection system, in which in some embodiments, three or more detection probes, each for a target biomarker, can be added to a sample comprising a biological entity (*e.g.*, extracellular vesicle). In some embodiments, detection probes each comprise a target binding moiety (*e.g.*, an affinity agent such as, e.g., an antibody agent against a target biomarker) coupled to an oligonucleotide domain, which comprises a double-stranded portion and a single-stranded overhang extended from one end of the oligonucleotide domain. A detection signal is generated when the corresponding single-stranded overhangs of all three or more detection probes hybridize to each other to form a linear double-stranded complex, and ligation of at least one strand of the double-stranded complex occurs, thus allowing a resulting ligated product to be detected.

**[75]** ***FIG. 14*** is a non-limiting example of a double-stranded complex comprising four detection probes connected to each other in a linear arrangement through hybridization of their respective single-stranded overhangs.

**[76]** ***FIG. 15*** is a schematic diagram illustrating a target entity detection assay of an exemplary embodiment described herein. In some embodiments, a plurality of detection probes, each for a distinct target, are added to a sample comprising a biological entity (*e.g.*, extracellular vesicle). In some embodiments, detection probes each comprise a target binding moiety (*e.g.*, an antibody agent) coupled to an oligonucleotide domain, which comprises a double-stranded portion and a single-stranded overhang extended from one end of the oligonucleotide domain. A detection signal is generated when all detection probes are localized to the same biological entity (*e.g.*, an extracellular vesicle or analyte) in close proximity such that the corresponding single-stranded overhangs hybridize to form a linear double-stranded complex, and ligation of at least one strand of the resulting linear double-stranded complex occurs, thereby allowing a ligated product to be detected.

**[77]** ***FIG. 16A*** is a graphical representation of the discriminatory power of a biomarker combination for LUAD detection by simulating "healthy patients" and "cancer patients." Using bioinformatic analysis, plasma samples from 5000 simulated "healthy patients" and 5000

simulated "cancer patients" were randomly selected from normal and cancer tissue databases, respectively. Simulated "cancer patients" were modeled to have tumors of varying size, ranging from 1g to 1000g. Based on the two pools of "healthy patients" and "cancer patients," sensitivity of a biomarker combination to detect LUAD at 99% specificity was calculated.

**[78]** ***FIG. 16B*** is an exemplary heatmap illustrating the ability of each possible combination of biomarkers from the list in Table 3 to detect LUAD based on simulated sensitivities for a 100g tumor as described herein. Each row represents one biomarker combination and each column represents one LUAD cancer patient. Light grey indicates that a LUAD patient's cancer was not detected using a given biomarker combination, and dark grey indicates that a LUAD patient's cancer was detected using a given biomarker combination. The heatmap shows the sensitivity thresholds based on a 100g tumor.

**[79]** ***FIG. 16C*** is a histogram illustrating the AUC values for each possible LUAD biomarker combination based on the list in Table 3 against other cancer types. To compare against other cancers, EV scores for biomarker combinations (wherein an EV score is a multiplication of all TPM expression values in a given biomarker combination) were calculated for lung adenocarcinomas and all other tumor types in a cancer molecular database (*e.g.*, the Cancer Genome Atlas), except for lung squamous carcinoma, and then the AUC was calculated. The histogram shows the distribution of AUC values against all other cancers. In general, biomarker combination described herein are shown to be specific to LUAD over other cancer types.

**[80]** ***FIG. 17A*** is a graphical representation of certain biomarker combinations relative to a Healthy Smoker Sample Pool for detection of Early Stage lung adenocarcinoma (LUAD). Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Early Stage LUAD Sample Pool from the Healthy Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Smoker Pooled samples and Early Stage LUAD Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[81]** ***FIG. 17B*** is a graphical representation of certain biomarker combinations relative to a Healthy Smoker Sample Pool for detection of Late Stage LUAD. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Late Stage LUAD Sample Pool from the Healthy Smoker Sample Pool. The x-axis represents the difference

in Ct value obtained from Healthy Smoker Pooled samples and Late Stage LUAD Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[82]** ***FIG. 18A*** is a graphical representation of certain biomarker combinations relative to a Healthy Smoker Sample Pool for detection of Early Stage lung squamous cell cancer (LUSC). Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Early Stage LUSC Sample Pool from the Healthy Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Smoker Pooled samples and Early Stage LUSC Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[83]** ***FIG. 18B*** is a graphical representation of certain biomarker combinations relative to a Healthy Smoker Sample Pool for detection of Late Stage LUSC. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Late Stage LUSC Sample Pool from the Healthy Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Smoker Pooled samples and Late Stage LUSC Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[84]** ***FIG. 19A*** is a graphical representation of certain biomarker combinations relative to a Healthy Non-Smoker Sample Pool for detection of Early Stage LUAD. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Early Stage LUAD Sample Pool from the Healthy Non-Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Non-Smoker Pooled samples and Early Stage LUAD Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[85]** ***FIG. 19B*** is a graphical representation of certain biomarker combinations biomarker combinations relative to a Healthy Non-Smoker Sample Pool for detection of Late Stage LUAD. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Late Stage LUAD Sample Pool from the Healthy Non-Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Non-Smoker Pooled samples and Late Stage LUAD Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[86]** ***FIG. 20A*** is a graphical representation of certain biomarker combinations relative to a Healthy Non-Smoker Sample Pool for detection of Early Stage LUSC. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Early Stage LUSC Sample Pool from the Healthy Non-Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Non-Smoker Pooled samples and Early Stage LUSC Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[87]** ***FIG. 20B*** is a graphical representation of certain biomarker combinations relative to a Healthy Non-Smoker Sample Pool for detection of Late Stage LUSC. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Late Stage LUSC Sample Pool from the Healthy Non-Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Non-Smoker Pooled samples and Late Stage LUSC Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[88]** ***FIG. 21A*** is a graphical representation of certain biomarker combinations relative to a Healthy Smoker Sample Pool for detection of Early Stage LUAD. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Early Stage LUAD Sample Pool from the Healthy Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Smoker Pooled samples and Early Stage LUAD Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probes).

**[89]** ***FIG. 21B*** is a graphical representation of certain biomarker combinations relative to a Healthy Smoker Sample Pool for detection of Late Stage LUAD. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Late Stage LUAD Sample Pool from the Healthy Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Smoker Pooled samples and Late Stage LUAD Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probes).

**[90]** ***FIG. 22A*** is a graphical representation of certain biomarker combinations relative to a Healthy Smoker Sample Pool for detection of Early Stage LUSC. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Early Stage

LUSC Sample Pool from the Healthy Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Smoker Pooled samples and Early Stage LUSC Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probes).

**[91]** ***FIG. 22B*** is a graphical representation of certain biomarker combinations relative to a Healthy Smoker Sample Pool for detection of Late Stage LUSC. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Late Stage LUSC Sample Pool from the Healthy Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Smoker Pooled samples and Late Stage LUSC Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probes).

**[92]** ***FIG. 23A*** is a graphical representation of certain biomarker combinations relative to a Healthy Non-Smoker Sample Pool for detection of Early Stage LUAD. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Early Stage LUAD Sample Pool from the Healthy Non-Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Non-Smoker Pooled samples and Early Stage LUAD Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probes).

**[93]** ***FIG. 23B*** is a graphical representation of certain biomarker combinations relative to a Healthy Non-Smoker Sample Pool for detection of Late Stage LUAD. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Late Stage LUAD Sample Pool from the Healthy Non-Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Non-Smoker Pooled samples and Late Stage LUAD Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[94]** ***FIG. 24A*** is a graphical representation of certain biomarker combinations relative to a Healthy Non-Smoker Sample Pool for detection of Early Stage LUSC. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Early Stage LUSC Sample Pool from the Healthy Non-Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Non-Smoker Pooled samples and

Early Stage LUSC Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probes).

**[95]** ***FIG. 24B*** is a graphical representation of certain biomarker combinations relative to a Healthy Non-Smoker Sample Pool for detection of Late Stage LUSC. Biomarker combinations were ranked (highest rank at the top of the chart) by their ability to distinguish the Late Stage LUSC Sample Pool from the Healthy Non-Smoker Sample Pool. The x-axis represents the difference in Ct value obtained from Healthy Non-Smoker Pooled samples and Late Stage LUSC Pooled samples. The y-axis represents certain biomarker combinations (target of capture probe, targets of detection probe).

**[96]** ***FIG. 25*** shows that the lung cancer (LC) training and validation cohorts included individuals with lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSqC), healthy never-smokers, and healthy ever-smokers. All samples were $K_2$EDTA plasma except those sourced from Reprocell (ACD plasma). Cancer samples were collected at time of diagnosis and prior to treatment, and staging was pathologically determined. Healthy controls had no medical history of cancer at the time of blood draw and smoking status was self-reported. COPD ever-smokers were clinically diagnosed with chronic obstructive pulmonary disease. One mL aliquots of each sample were purified using size exclusion chromatography and split across the various biomarker combinations evaluated, and all samples were tested in duplicate. Biological materials were provided by a tumor bank.

**[97]** ***FIG. 26*** shows that an LC panel comprising 8 biomarker combinations were trained on a cohort of 118 individual patient samples, resulting in a locked classifier which was validated in an independent, blinded cohort of 456 samples.

**[98]** ***FIG. 27***: (A) Box plot summarizing the probability of lung cancer based on an exemplary LC Panel (e.g., ones described herein) and locked classifier. Each data point represents the median of assay replicates for a unique sample and the red dotted line denotes 80% specificity relative to the healthy never-smoker, healthy ever-smoker, and COPD cohorts. The exemplary LC Panel exhibits similar signal distributions across healthy never-smokers, healthy ever-smokers, and individuals with COPD, indicating that it is not detecting a smoking signature. (B) ROC curve of an exemplary LC Panel (e.g., ones described herein) comparing stage I/II lung cancer to healthy never-smoker, healthy ever-smoker, and COPD cohorts. The

disclosed locked model demonstrates promising discrimination of LUAD and LUSqC from healthy never-smokers, healthy ever-smokers, and individuals with COPD.

**[99]** ***FIG. 28*** shows the number of lung cancer samples and assay sensitivity sorted by stage.

**[100]** ***FIG. 29*** shows that the signal from an exemplary LC assay (e.g., ones described herein) is correlated with tumor size but is not correlated with smoking history.

**[101]** ***FIG. 30*** shows a subset analysis of screen-detected stage I, II, and III/IV lung cancer cases compared to healthy never-smokers and healthy ever-smokers. Each data point represents the median of assay replicates for a unique sample and the red dotted line denotes 80% specificity relative to the healthy never-smoker, healthy ever-smoker, and COPD cohorts. The performance of an exemplary LC assay (e.g., ones described herein) is similar in screen-detected lung cancer cases.

**[102]** ***FIG. 31*** shows the performance of an exemplary LC assay (e.g., ones described herein) in lung cancer cases detected by routine screening.

## CERTAIN DEFINITIONS

**[103]** ***Administering:*** As used herein, the term "administering" or "administration" typically refers to the administration of a composition to a subject to achieve delivery of an agent that is, or is included in, a composition to a target site or a site to be treated. Those of ordinary skill in the art will be aware of a variety of routes that may, in appropriate circumstances, be utilized for administration to a subject, for example a human. For example, in some embodiments, administration may be parenteral. In some embodiments, administration may be oral. In some embodiments, administration may involve only a single dose. In some embodiments, administration may involve application of a fixed number of doses. In some embodiments, administration may involve dosing that is intermittent (*e.g.*, a plurality of doses separated in time) and/or periodic (*e.g.*, individual doses separated by a common period of time) dosing. In some embodiments, administration may involve continuous dosing (*e.g.*, perfusion) for at least a selected period of time.

**[104]** ***Affinity Agent***: The term "affinity agent" as used herein refers to an entity that is or comprises a target-binding moiety as described herein, and therefore binds to a target of interest (e.g., molecular target of interest such as a biomarker or an epitope). In many

embodiments, an affinity agent in accordance with the present disclosure binds specifically with a biomarker as described herein. In many embodiments, an affinity agent in accordance with the present disclosure binds specifically with a surface biomarker as described herein. In some embodiments, an affinity agent in accordance with the present disclosure binds specifically with a carbohydrate-dependent marker as described herein. In some embodiments, an affinity agent may be or comprise an antibody agent (e.g., an antibody or other entity that is or includes an antigen-binding portion thereof). Alternatively or additionally, in some embodiments, an affinity agent may selected from the group consisting of affimers, aptamers, lectins, sialic acid-binding immunoglobulin-type lectins (siglecs), and combinations thereof, and/or another binding agent that may be considered a ligand. In some embodiments, a target (e.g., a biomarker target) of an affinity agent is or comprises one or more polypeptide, nucleic acid, carbohydrate, and/or lipid moieties and/or entities).

**[105]** ***Agent***: In general, the term "agent", as used herein, is used to refer to an entity (e.g., for example, a lipid, metal, nucleic acid, polypeptide, polysaccharide, small molecule, etc, or complex, combination, mixture or system [e.g., cell, tissue, organism] thereof), or phenomenon (e.g., heat, electric current or field, magnetic force or field, etc). In appropriate circumstances, as will be clear from context to those skilled in the art, the term may be utilized to refer to an entity that is or comprises a cell or organism, or a fraction, extract, or component thereof. Alternatively or additionally, as context will make clear, the term may be used to refer to a natural product in that it is found in and/or is obtained from nature. In some instances, again as will be clear from context, the term may be used to refer to one or more entities that is man-made in that it is designed, engineered, and/or produced through action of the hand of man and/or is not found in nature. In some embodiments, an agent may be utilized in isolated or pure form; in some embodiments, an agent may be utilized in crude form. In some embodiments, potential agents may be provided as collections or libraries, for example that may be screened to identify or characterize active agents within them. In some cases, the term "agent" may refer to a compound or entity that is or comprises a polymer; in some cases, the term may refer to a compound or entity that comprises one or more polymeric moieties. In some embodiments, the term "agent" may refer to a compound or entity that is not a polymer and/or is substantially free of any polymer and/or of one or more particular polymeric moieties. In some embodiments, the

term may refer to a compound or entity that lacks or is substantially free of any polymeric moiety.

**[106]** ***Amplification***: The terms "amplification" and "amplify" refers to a template-dependent process that results in an increase in the amount and/or levels of a nucleic acid molecule relative to its initial amount and/or level. A template-dependent process is generally a process that involves template-dependent extension of a primer molecule, wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, J. D. et al., In: Molecular Biology of the Gene, 4th Ed., W. A. Benjamin, Inc., Menlo Park, Calif. (1987); which is incorporated herein by reference for the purpose described herein.

**[107]** ***Antibody agent*:** As used herein, the term "antibody agent" refers to an agent that specifically binds to a particular antigen. In some embodiments, an antibody agent refers to a polypeptide that includes canonical immunoglobulin sequence elements sufficient to confer specific binding to a particular target antigen. As is known in the art, intact antibodies as produced in nature are approximately 150 kD tetrameric agents comprised of two identical heavy chain polypeptides (about 50 kD each) and two identical light chain polypeptides (about 25 kD each) that associate with each other into what is commonly referred to as a "Y-shaped" structure. Each heavy chain is comprised of at least four domains (each about 110 amino acids long)– an amino-terminal variable (VH) domain (located at the tips of the Y structure), followed by three constant domains: CH1, CH2, and the carboxy-terminal CH3 (located at the base of the Y's stem). A short region, known as the "switch", connects the heavy chain variable and constant regions. The "hinge" connects CH2 and CH3 domains to the rest of the antibody. Two disulfide bonds in this hinge region connect the two heavy chain polypeptides to one another in an intact antibody. Each light chain is comprised of two domains – an amino-terminal variable (VL) domain, followed by a carboxy-terminal constant (CL) domain, separated from one another by another "switch". Intact antibody tetramers are comprised of two heavy chain-light chain dimers in which the heavy and light chains are linked to one another by a single disulfide bond; two other disulfide bonds connect the heavy chain hinge regions to one another, so that the dimers are connected to one another, and the tetramer is formed. Naturally-produced antibodies are also glycosylated, typically on the CH2 domain. Each domain in a natural antibody has a structure characterized by an "immunoglobulin fold" formed from two beta sheets (e.g., 3-, 4-, or 5-

stranded sheets) packed against each other in a compressed antiparallel beta barrel. Each variable domain contains three hypervariable loops known as "complement determining regions" (CDR1, CDR2, and CDR3) and four somewhat invariant "framework" regions (FR1, FR2, FR3, and FR4). When natural antibodies fold, the FR regions form the beta sheets that provide the structural framework for the domains, and the CDR loop regions from both the heavy and light chains are brought together in three-dimensional space so that they create a single hypervariable antigen binding site located at the tip of the Y structure. The Fc region of naturally-occurring antibodies binds to elements of the complement system, and also to receptors on effector cells, including for example effector cells that mediate cytotoxicity. As is known in the art, affinity and/or other binding attributes of Fc regions for Fc receptors can be modulated through glycosylation or other modification. In some embodiments, antibodies produced and/or utilized in accordance with the present invention include glycosylated Fc domains, including Fc domains with modified or engineered such glycosylation. For purposes of the present invention, in certain embodiments, any polypeptide or complex of polypeptides that includes sufficient immunoglobulin domain sequences as found in natural antibodies can be referred to and/or used as an "antibody", whether such polypeptide is naturally produced (e.g., generated by an organism reacting to an antigen), or produced by recombinant engineering, chemical synthesis, or other artificial system or methodology. In some embodiments, an antibody is polyclonal; in some embodiments, an antibody is monoclonal. In some embodiments, an antibody has constant region sequences that are characteristic of rabbit, rodent (e.g., mouse, rat, hamster, etc.), camelid (e.g., llama, alpaca), sheep, goat, bovine, horse, chicken, donkey, shark, primate, human, or in vitro-derived (e.g., yeast, phage) antibodies. In some embodiments, antibody sequence elements are humanized, primatized, chimeric, *etc.*, as is known in the art. Moreover, the term "antibody" as used herein, can refer in appropriate embodiments (unless otherwise stated or clear from context) to any of the art-known or developed constructs or formats for utilizing antibody structural and functional features in alternative presentation. For example, in some embodiments, an antibody utilized in accordance with the present invention is in a format selected from, but not limited to, IgA, IgG, IgE or IgM antibodies; bi- or multi- specific antibodies (*e.g.*, Zybodies®, *etc.*); antibody fragments such as Fab fragments, Fab' fragments, F(ab')2 fragments, Fd' fragments, Fd fragments, and isolated CDRs or sets thereof; single chain Fvs; polypeptide-Fc fusions; single domain antibodies, alternative scaffolds or antibody

mimetics (e.g., anticalins, FN3 monobodies, Affibodies, Affilins, Affimers, Affitins, Alphabodies, Avimers, Fynomers, Im7, VLR, VNAR, Trimab, CrossMab, Trident); nanobodies, binanobodies, F(ab')2, Fab', di-sdFv, single domain antibodies, trifunctional antibodies, diabodies, and minibodies. etc. In some embodiments, relevant formats may be or include: Adnectins®; Affibodies®; Affilins®; Anticalins®; Avimers®; BiTE®s; cameloid antibodies; Centyrins®; ankyrin repeat proteins or DARPINs®; dual-affinity re-targeting (DART) agents; Fynomers®; shark single domain antibodies such as IgNAR; immune mobilizing monoclonal T cell receptors against cancer (ImmTACs); KALBITOR®s; MicroProteins; Nanobodies® minibodies; masked antibodies (*e.g.*, Probodies®); Small Modular ImmunoPharmaceuticals ("SMIPs™"); single chain or Tandem diabodies (TandAb®); TCR-like antibodies; Trans-bodies®; TrimerX®; VHHs. In some embodiments, an antibody may lack a covalent modification (*e.g.*, attachment of a glycan) that it would have if produced naturally. In some embodiments, an antibody may contain a covalent modification (*e.g.*, attachment of a glycan, a payload [*e.g.*, a detectable moiety, a therapeutic moiety, a catalytic moiety, *etc.*], or other pendant group [*e.g.*, poly-ethylene glycol, *etc.*]).

**[108]** ***Antigen***: As used herein, the term "antigen" refers to an entity (e.g., a molecule or a molecular structure such as, e.g., a peptide or protein, carbohydrate, lipoparticle, oligonucleotide, chemical molecule, or combinations thereof) that includes one or more epitopes and therefore is recognized and bound by an affinity agent (e.g., an antibody, affimer, or aptamer).

**[109]** ***Approximately or about***: As used herein, the term "approximately" or "about," as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In general, those skilled in the art, familiar within the context, will appreciate the relevant degree of variance encompassed by "about" or "approximately" in that context. For example, in some embodiments, the term "approximately" or "about" may encompass a range of values that are within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less of the referred value.

**[110]** ***Aptamer***: As used herein, the term "aptamer" typically refers to a nucleic acid molecule or a peptide molecule that binds to a specific target molecule (*e.g.*, an epitope). In some embodiments, a nucleic acid aptamer may be described by a nucleotide sequence and is typically about 15-60 nucleotides in length. A nucleic acid aptamer may be or comprise a single stranded

and/or double-stranded structure. In some embodiments, a nucleic acid aptamer may be or comprise DNA. In some embodiments, a nucleic acid aptamer may be or comprise RNA. Without wishing to be bound by any theory, it is contemplated that the chain of nucleotides in an aptamer form intramolecular interactions that fold the molecule into a complex three-dimensional shape, and this three-dimensional shape allows the aptamer to bind tightly to the surface of its target molecule. In some embodiments, a peptide aptamer may be described to have one or more peptide loops of variable sequence displayed by a protein scaffold. Peptide aptamers can be isolated from combinatorial libraries and often subsequently improved by directed mutation or rounds of variable region mutagenesis and selection. Given the extraordinary diversity of molecular shapes that exist within the universe of all possible nucleotide and/or peptide sequences, aptamers may be obtained for a wide array of molecular targets, including proteins and small molecules. In addition to high specificity, aptamers typically have very high affinities for their targets (*e.g.*, affinities in the picomolar to low nanomolar range for proteins or polypeptides). Because aptamers are typically synthetic molecules, aptamers are amenable to a variety of modifications, which can optimize their function for particular applications.

**[111]** ***Associated with:*** Two events or entities are "associated" with one another, as that term is used herein, if the presence, level and/or form of one is correlated with that of the other. For example, a particular biological phenomenon (*e.g.*, expression of a specific biomarker) is considered to be associated with lung cancer (*e.g.*, a specific type of lung cancer and/or stage of lung cancer), if its presence correlates with incidence of and/or susceptibility of lung cancer (*e.g.*, across a relevant population).

**[112]** ***Biological entity:*** In appropriate circumstances, as will be clear from context to those skilled in the art, the term "biological entity" may be utilized to refer to an entity or component that is present in a biological sample, *e.g.*, in some embodiments derived or obtained from a subject, which, in some embodiments, may be or comprise a cell or an organism, such as an animal or human, or, in some embodiments, may be or comprise a biological tissue or fluid. In some embodiments, a biological entity is or comprises a cell or microorganism, or a fraction, extract, or component thereof (including, *e.g.*, intracellular components and/or molecules secreted by a cell or microorganism). In some embodiments, a biological entity is or comprises a nanoparticle having a size within the range of about 30 nm to about 1000 nm, which in some embodiments are obtained from a bodily fluid sample (e.g., but not limited to a blood-derived

sample) of a subject. In some embodiments, such a nanoparticle may be or comprise a protein aggregate, including, e.g., in some embodiments comprising a glycan, and/or an extracellular vesicle. In some embodiments, such a nanoparticle may have a size within the range of about 30 nm to about 1000 nm, about 50 nm to about 500 nm, or about 75 nm to about 500 nm. For example, in some embodiments, a biological entity is or comprises a cell. In some embodiments, a biological entity is or comprises an extracellular vesicle. In some embodiments, a biological entity is or comprises a biological analyte (*e.g.*, a metabolite, carbohydrate, protein or polypeptide, enzyme, lipid, organelle, cytokine, receptor, ligand, and any combinations thereof). In some embodiments, a biological entity present in a sample is in a native state (*e.g.*, proteins or polypeptides remain in a naturally occurring conformational structure). In some embodiments, a biological entity is processed, *e.g.*, by isolating from a sample or deriving from a naturally occurring biological entity. For example, a biological entity can be processed with one or more chemical agents such that it is more desirable for detection utilizing technologies provided herein. As an example only, a biological entity may be a cell or extracellular vesicle that is contacted with a fixative agent (*e.g.*, but not limited to methanol and/or formaldehyde) to cause proteins and/or peptides present in the cell or extracellular vesicle to form crosslinks. In some embodiments, a biological entity is in an isolated or pure form (*e.g.*, isolated from a bodily fluid sample such as, *e.g.*, a blood, serum, plasma sample, *etc.*). In some embodiments, a biological entity may be present in a complex matrix (*e.g.*, a bodily fluid sample such as, *e.g.*, a blood, serum, or plasma sample, *etc.*).

**[113]** ***Biomarker***: The term "biomarker" typically refers to an entity, event, or characteristic whose presence, level, degree, type, and/or form, correlates with a particular biological event or state of interest, so that it is considered to be a "marker" of that event or state. To give but a few examples, in some embodiments, a biomarker may be or comprise a marker for a particular disease state, or for likelihood that a particular disease, disorder or condition may develop, occur, or reoccur. In some embodiments, a biomarker may be or comprise a marker for a particular disease or therapeutic outcome, or likelihood thereof. In some embodiments, a biomarker may be or comprise a marker for a particular tissue (*e.g.*, but not limited to brain, breast, colon, ovary and/or other tissues associated with a female reproductive system, pancreas, prostate and/or other tissues associated with a male reproductive system, liver, lung, and skin). Such a marker for a particular tissue, in some embodiments, may be specific for a healthy tissue,

specific for a diseased tissue, or in some embodiments may be present in a normal healthy tissue and diseased tissue (*e.g.*, a tumor); those skilled in the art, reading the present disclosure, will appreciate appropriate contexts for each such type of biomarker. In some embodiments, a biomarker may be or comprise a cancer-specific marker (*e.g.*, a marker that is specific to a particular cancer). In some embodiments, a biomarker may be or comprise a non-specific cancer marker (*e.g.,* a marker that is present in at least two or more cancers). A non-specific cancer marker may be or comprise, in some embodiments, a generic marker for cancers (*e.g.*, a marker that is typically present in cancers, regardless of tissue types), or in some embodiments, a marker for cancers of a specific tissue (*e.g.*, but not limited to brain, breast, colon, ovary and/or other tissues associated with a female reproductive system, pancreas, prostate and/or other tissues associated with a male reproductive system, liver, lung, and skin). Thus, in some embodiments, a biomarker is predictive; in some embodiments, a biomarker is prognostic; in some embodiments, a biomarker is diagnostic of the relevant biological event or state of interest. A biomarker may be or comprise an entity of any chemical class and may be or comprise a combination of entities. For example, in some embodiments, a biomarker may be or comprise a nucleic acid, a polypeptide, a lipid, a carbohydrate, a small molecule, an inorganic agent (*e.g.*, a metal or ion), or a combination thereof. In some embodiments, a biomarker is or comprises a portion of a particular molecule, complex, or structure; *e.g.,* in some embodiments, a biomarker may be or comprise an epitope. In some embodiments, a biomarker is a surface marker (*e.g.*, a surface marker) of an extracellular vesicle associated with lung cancer. In some embodiments, a biomarker is intravesicular (*e.g.*, a protein or RNA marker that is present within an extracellular vesicle). In some embodiments, a biomarker may be or comprise a genetic or epigenetic signature. In some embodiments, a biomarker may be or comprise a gene expression signature. In some embodiments, a "biomarker" appropriate for use in accordance with the present disclosure may refer to presence, level, and/or form of a molecular entity (*e.g.*, epitope) present in a target marker. For example, in some embodiments, two or more "biomarkers" as molecular entities (*e.g.*, epitopes) may be present on the same target marker (*e.g.*, a biomarker such as a surface biomarker present in an extracellular vesicle).

**[114]** ***Blood-derived sample:*** The term "blood-derived sample," as used herein, refers to a sample derived from a blood sample (*i.e.*, a whole blood sample) of a subject in need thereof. Examples of blood-derived samples include, but are not limited to, blood plasma (including, *e.g.*,

fresh frozen plasma), blood serum, blood fractions, plasma fractions, serum fractions, blood fractions comprising red blood cells (RBC), platelets, leukocytes, *etc*., and cell lysates including fractions thereof (for example, cells, such as red blood cells, white blood cells, *etc.*, may be harvested and lysed to obtain a cell lysate). In some embodiments, a blood-derived sample that is used with methods, systems, and/or kits described herein is a plasma sample.

**[115]** ***Cancer***: The term "cancer" is used herein to generally refer to a disease or condition in which cells of a tissue of interest exhibit relatively abnormal, uncontrolled, and/or autonomous growth, so that they exhibit an aberrant growth phenotype characterized by a significant loss of control of cell proliferation. In some embodiments, cancer may comprise cells that are precancerous (*e.g.*, benign), malignant, pre-metastatic, metastatic, and/or non-metastatic. The present disclosure provides technologies for detection of lung cancer.

**[116]** ***Capture assay***: As used herein, the term "capture assay" refers to a process of isolating or separating a biological entity of interest from a sample (*e.g.*, in some embodiments a bodily fluid-derived sample). In some embodiments, a biological entity of interest is isolated or separated from a sample (*e.g.*, in some embodiments a bodily fluid-derived sample) using a capture probe described herein. In some embodiments, a biological entity of interest that binds to a capture probe described herein is subject to a detection assay described herein. In some embodiments, a biological entity of interest amenable to a capture assay described herein is or comprises nanoparticles having a size range of interest that includes extracellular vesicles. In some embodiments, such a nanoparticle may have a size within the range of about 30 nm to about 1000 nm, about 50 nm to about 500 nm, or about 75 nm to about 500 nm. In some embodiments, a biological entity of interest amenable to a capture assay described herein is or comprises extracellular vesicles (*e.g.*, in some embodiments exosomes) of interest.

**[117]** ***Capture probe***: As used herein, the term "capture probe" refers to a capture agent for capturing a biological entity of interest from a sample (*e.g.*, in some embodiments a bodily fluid-derived sample, e.g., but not limited to a blood-derived sample). In many embodiments described herein, a capture agent comprises at least one target-capture moiety that binds to a surface polypeptide of a biological entity of interest. In some embodiments, such a biological entity of interest is or comprises nanoparticles having a size range of interest that includes extracellular vesicles. In some embodiments, such nanoparticles may have a size within the range of about 30 nm to about 1000 nm, about 50 nm to about 500 nm, or about 75 nm to about

500 nm. In some embodiments, such a biological entity of interest comprises extracellular vesicles (*e.g.*, in some embodiments exosomes). In some embodiments, a capture agent comprises at least one target moiety that binds to a surface biomarker (*e.g.*, ones described herein) of nanoparticles having a size within the range of about 30 nm to about 1000 nm, including, e.g., extracellular vesicles (*e.g.*, in some embodiments exosomes). In some embodiments, a target-capture moiety of a capture agent is or comprises an affinity agent described herein. In some embodiments, a target-capture moiety of a capture agent is or comprises an antibody agent. In some embodiments, a target-capture moiety of a capture agent is or comprises a lectin or a sialic acid-binding immunoglobulin-type lectin. In some embodiments, a capture agent may comprise a solid substrate such that its target-capture moiety is immobilized thereonto. In some embodiments, an exemplary solid substrate is a bead (*e.g.*, a magnetic bead). In some embodiments, a capture probe is or comprises a population of magnetic beads comprising a target-capture moiety that specifically binds to a surface biomarker described herein.

**[118]** ***Classification cutoff***: As used herein, the term "classification cutoff" refers to a level, value, or score, or a set of values, or an indicator that is used to predict a subject's risk for a disease or condition (*e.g.*, lung cancer), for example, by defining one or more dividing lines among two or more subsets of a population (*e.g.*, normal healthy subjects and subjects with inflammatory conditions *vs.* lung cancer subjects). In some embodiments, a classification cutoff may be determined referencing at least one reference threshold level (*e.g.*, reference cutoff) for a target biomarker signature described herein, optionally in combination with other appropriate variables, *e.g.*, age, life-history-associated risk factors, hereditary factors, physical and/or medical conditions of a subject. In some embodiments where a classification is based on a single target biomarker signature (*e.g.*, as described herein), a classification cutoff may be the same as a reference threshold (*e.g.*, cutoff) pre-determined for the single target biomarker signature. In some embodiments where a classification is based on two or more (e.g., 2, 3, 4, or more) target biomarker signatures, a classification cutoff may reference two or more reference thresholds (*e.g.*, cutoffs) each individually pre-determined for the corresponding target biomarker signatures, and optionally incorporate one or more appropriate variables, *e.g.*, age, life-history-associated risk factors, hereditary factors, physical and/or medical conditions of a subject. In some embodiments, a classification cutoff may be determined via a computer algorithm-

mediated analysis that references at least one reference threshold level (*e.g.*, reference cutoff) for a target biomarker signature described herein, optionally in combination with other appropriate variables, *e.g.*, age, life-history-associated risk factors, hereditary factors, physical and/or medical conditions of a subject.

**[119]** ***Close proximity***: The term "close proximity" as used herein, refers to a distance between two detection probes (*e.g.* two detection probes in a pair) that is sufficiently close enough such that an interaction between the detection probes (*e.g.*, through respective oligonucleotide domains) is expected to likely occur. For example, in some embodiments, probability of two detection probes interacting with each other (*e.g.*, through respective oligonucleotide domains) over a period of time when they are in sufficiently close proximity to each other under a specified condition (*e.g.*, when detection probes are bound to respective targets in an extracellular vesicle is at least 50% or more, including, *e.g.*, at least 60%, at least 70%, at least 80%, at least 90% or more. In some embodiments, a distance between two detection probes when they are in sufficiently close proximity to each other may range between approximately 0.1-1000 nm, or 0.5-500 nm, or 1-250 nm. In some embodiments, a distance between two detection probes when they are in sufficiently close proximity to each other may range between approximately 0.1-10 nm or between approximately 0.5-5 nm. In some embodiments, a distance between two detection probes when they are in sufficiently close proximity to each other may be less than 100 nm or shorter, including, *e.g.*, less than 90 nm, less than 80 nm, less than 70 nm, less than 60 nm, less than 50 nm, less than 40 nm, less than 30 nm, less than 20 nm, less than 10 nm, less than 5 nm, less than 1 nm, or shorter. In some embodiments, a distance between two detection probes when they are in sufficiently close proximity to each other may range between approximately 40-1000 nm or 40 nm-500 nm.

**[120]** ***Comparable*:** As used herein, the term "comparable" refers to two or more agents, entities, situations, sets of conditions, *etc.*, that may not be identical to one another but that are sufficiently similar to permit comparison therebetween so that one skilled in the art will appreciate that conclusions may reasonably be drawn based on differences or similarities observed. In some embodiments, comparable sets of conditions, circumstances, individuals, or populations are characterized by a plurality of substantially identical features and one or a small number of varied features. Those of ordinary skill in the art will understand, in context, what degree of identity is required in any given circumstance for two or more such agents, entities,

situations, sets of conditions, *etc*. to be considered comparable. For example, those of ordinary skill in the art will appreciate that sets of circumstances, individuals, or populations are comparable to one another when characterized by a sufficient number and type of substantially identical features to warrant a reasonable conclusion that differences in results obtained or phenomena observed under or with different sets of circumstances, individuals, or populations are caused by or indicative of the variation in those features that are varied.

**[121]** ***Complementary:*** As used herein, the term "complementary" in the context of nucleic acid base-pairing refers to oligonucleotide hybridization related by base-pairing rules. For example, the sequence "C-A-G-T" is complementary to the sequence "G-T-C-A." Complementarity can be partial or total. Thus, any degree of partial complementarity is intended to be included within the scope of the term "complementary" provided that the partial complementarity permits oligonucleotide hybridization. Partial complementarity is where one or more nucleic acid bases is not matched according to the base pairing rules. Total or complete complementarity between nucleic acids is where each and every nucleic acid base is matched with another base under the base pairing rules. In the context of identifying biomarker combinations for detection of a particular cancer, the term "complementary" is used herein in reference to sets of biomarkers having different information content (e.g., ability to detect cancer in distinct, substantially non-overlapping subgroups of subjects). For example, two sets of biomarkers – set 1 and set 2 – are said to be "complementary" to each other if, for example, set 1 detects cancer in a group (e.g., group A) of subjects in a population, and set 2 detects cancer in a substantially separate and non-overlapping group of subjects in the same population (e.g., group B), but not in Group A. Similarly, set 1 does not detect cancer in a substantial number of subjects in Group B.

**[122]** ***Detecting:*** The term "detecting" is used broadly herein to include appropriate means of determining the presence or absence of an extracellular vesicle expressing a target biomarker signature of lung cancer or any form of measurement indicative of such an extracellular vesicle. Thus, "detecting" may include determining, measuring, assessing, or assaying the presence or absence, level, amount, and/or location of an entity of interest (*e.g.*, a surface biomarker, an intravesicular biomarker, or an intravesicular RNA biomarker) that corresponds to part of a target biomarker signature in any way. In some embodiments, "detecting" may include determining, measuring, assessing, or quantifying a form of

measurement indicative of an entity of interest (*e.g.*, a ligated template indicative of a protein biomarker and/or an intravesicular biomarker, or a PCR amplification product indicative of an intravesicular mRNA). Quantitative and qualitative determinations, measurements or assessments are included, including semi-quantitative. Such determinations, measurements or assessments may be relative, for example when an entity of interest (*e.g.*, a surface biomarker, an intravesicular biomarker, or an intravesicular RNA biomarker) or a form of measurement indicative thereof is being detected relative to a control reference, or absolute. As such, the term "quantifying" when used in the context of quantifying an entity of interest (*e.g.*, a surface biomarker, an intravesicular biomarker, or an intravesicular RNA biomarker) or a form of measurement indicative thereof can refer to absolute or to relative quantification. Absolute quantification may be accomplished by correlating a detected level of an entity of interest (*e.g.*, a surface biomarker, an intravesicular biomarker, or an intravesicular RNA biomarker) or a form of measurement indicative thereof to known control standards (*e.g.*, through generation of a standard curve). Alternatively, relative quantification can be accomplished by comparison of detected levels or amounts between two or more different entities of interest (*e.g.*, different surface biomarkers, intravesicular biomarkers, or intravesicular RNA biomarkers) to provide a relative quantification of each of the two or more different entities of interest, *i.e.*, relative to each other.

**[123]** ***Detection label:*** The term "detection label" as used herein refers to any element, molecule, functional group, compound, fragment or moiety that is detectable. In some embodiments, a detection label is provided or utilized alone. In some embodiments, a detection label is provided and/or utilized in association with (*e.g.*, joined to) another agent. Examples of detection labels include, but are not limited to: various ligands, radionuclides (*e.g.*, $^{3}H$, $^{14}C$, $^{18}F$, $^{19}F$, $^{32}P$, $^{35}S$, $^{135}I$, $^{125}I$, $^{123}I$, $^{64}Cu$, $^{187}Re$, $^{111}In$, $^{90}Y$, $^{99m}Tc$, $^{177}Lu$, $^{89}Zr$, *etc.*), fluorescent dyes, chemiluminescent agents (such as, for example, acridinium esters, stabilized dioxetanes, and the like), bioluminescent agents, spectrally resolvable inorganic fluorescent semiconductors nanocrystals (*i.e.*, quantum dots), metal nanoparticles (*e.g.*, gold, silver, copper, platinum, *etc.*) nanoclusters, paramagnetic metal ions, enzymes, colorimetric labels (such as, for example, dyes, colloidal gold, and the like), biotin, digoxigenin, haptens, and proteins for which antisera or monoclonal antibodies are available.

**[124]** ***Detection probe*:** The term "detection probe" typically refers to a probe directed to detection and/or quantification of a specific target. In some embodiments, a detection probe is a quantification probe, which provides an indicator representing level of a specific target. In accordance with the present disclosure, a detection probe refers to a composition comprising a target binding entity, directly or indirectly, coupled to an oligonucleotide domain, wherein the target binding entity specifically binds to a respective target (*e.g.*, molecular target), and wherein at least a portion of the oligonucleotide domain is designed to permit hybridization with a portion of an oligonucleotide domain of another detection probe for a distinct target. In many embodiments, an oligonucleotide domain appropriate for use in the accordance with the present disclosure comprises a double-stranded portion and at least one single-stranded overhang. In some embodiments, an oligonucleotide domain may comprise a double-stranded portion and a single-stranded overhang at each end of the double-stranded portion. In some embodiments, a target binding entity of a detection probe is or comprises an affinity agent described herein. In some embodiments, a target binding entity of a detection probe is or comprises an antibody agent. In some embodiments, a target binding entity of a detection probe is or comprises a lectin or a sialic acid-binding immunoglobulin-type lectin (siglec).

**[125]** ***Double-stranded:*** As used herein, the term "double-stranded" in the context of oligonucleotide domain is understood by those of skill in the art that a pair of oligonucleotides exist in a hydrogen-bonded, helical arrangement typically associated with, for example, nucleic acid such as DNA. In addition to the 100% complementary form of double-stranded oligonucleotides, the term "double-stranded" as used herein is also meant to refer to those forms which include mismatches (*e.g.*, partial complementarity) and/or structural features as bulges, loops, or hairpins.

**[126]** ***Double-stranded complex:*** As used herein, the term "double-stranded complex" typically refers to a complex comprising at least two or more (including, *e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more) detection probes (*e.g.*, as provided and/or utilized herein), each directed to a target (which can be the same target or a distinct target), connected or coupled to one another in a linear arrangement through hybridization of complementary single-stranded overhangs of the detection probes. In some embodiments, such a double-stranded complex may comprise an extracellular vesicle, wherein respective target binding moieties of the detection probes are simultaneously bound to the extracellular vesicle.

**[127]** ***Epitope:*** As used herein, the term "epitope" includes any moiety that is specifically recognized by an affinity agent (*e.g.*, but not limited to, an antibody, affimer, and/or aptamer). In some embodiments, an epitope is comprised of a plurality of chemical atoms or groups on an antigen. In some embodiments, such chemical atoms or groups are surface-exposed when the antigen adopts a relevant three-dimensional conformation. In some embodiments, such chemical atoms or groups are physically near to each other in space when the antigen adopts such a conformation. In some embodiments, at least some such chemical atoms are groups are physically separated from one another when the antigen adopts an alternative conformation (*e.g.*, is linearized).

**[128]** ***Extracellular vesicle:*** As used herein, the term "extracellular vesicle" typically refers to a vesicle outside of a cell, *e.g.*, secreted by a cell. Examples of secreted vesicles include, but are not limited to exosomes, microvesicles, microparticles, ectosomes, oncosomes, and apoptotic bodies. Without wishing to be bound by theory, exosomes are nanometer-sized vesicles (*e.g.*, between 40 nm and 120 nm) of endocytic origin that may form by inward budding of the limiting membrane of multivesicular endosomes (MVEs), while microvesicles typically bud from the cell surface and their size may vary between 50 nm and 1000 nm. In some embodiments, an extracellular vesicle is or comprises an exosome and/or a microvesicle. In some embodiments, a sample comprising an extracellular vesicle is substantially free of apoptotic bodies. In some embodiments, a sample comprising nanoparticles may comprise extracellular nanoparticles shed or derived from one or more tissues (*e.g.*, cancerous tissues and/or non-cancerous or healthy tissues). In some embodiments, an extracellular vesicle in a sample may be shed or derived from a lung cancer tumor; in some embodiments, an extracellular vesicle is shed or derived from a tumor of a non-lung cancer. In some embodiments, an extracellular vesicle is shed or derived from a healthy tissue. In some embodiments, an extracellular vesicle is shed or derived from a benign lung tumor. In some embodiments, an extracellular vesicle is shed or derived from a tissue of a subject with symptoms (*e.g.*, non-specific symptoms) associated with lung cancer.

**[129]** ***Extracellular vesicle-associated membrane-bound polypeptide***: As used herein, such a term refers to a polypeptide that is present in the membrane of an extracellular vesicle. In some embodiments, such a biomarker may be associated with the extracellular side of the membrane. In some embodiments, such a polypeptide may be tumor-specific. In some

embodiments, such a polypeptide may be tissue-specific (*e.g.*, lung tissue-specific). In some embodiments, such a polypeptide may be non-specific, *e.g.*, it is present in one or more non-target tumors, and/or in one or more non-target tissues.

**[130]** ***Hybridization:*** As used herein, the term "hybridizing", "hybridize", "hybridization", "annealing", or "anneal" are used interchangeably in reference to pairing of complementary nucleic acids using any process by which a strand of nucleic acid joins with a complementary strand through base pairing to form a hybridization complex. Hybridization and the strength of hybridization (*e.g.*, strength of the association between the nucleic acids) is impacted by various factors including, *e.g.*, the degree of complementarity between the nucleic acids, stringency of the conditions involved, the melting temperature (T) of the formed hybridization complex, and the G:C ratio within the nucleic acids.

**[131]** ***Intravesicular protein biomarker:*** As used herein, the term "intravesicular protein biomarker" refers to a marker indicative of the state (*e.g.*, presence, level, and/or activity) of a polypeptide that is present within a biological entity (*e.g.*, a cell or an extracellular vesicle). In many embodiments, an intravesicular biomarker is associated with or present within an extracellular vesicle. In some embodiments, an intravesicular protein biomarker may be or comprise a phosphorylated polypeptide. In some embodiments, an intravesicular protein biomarker may be or comprise a mutated polypeptide.

**[132]** ***Intravesicular RNA biomarker:*** As used herein, the term "intravesicular RNA biomarker" refers to a marker indicative of the state (*e.g.*, presence and/or level) of a RNA (*e.g.*, but not limited to mRNA and noncoding RNA such as, e.g., orphan noncoding RNA, long noncoding RNA, piwi-interacting RNA, microRNA, circular RNA, etc.) that is present within a biological entity (*e.g.*, a cell or an extracellular vesicle). In many embodiments, an intravesicular RNA biomarker is associated with or present within an extracellular vesicle. In some embodiments, an intravesicular RNA biomarker is associated or specific to cancer. In some embodiments, an intravesicular RNA biomarker is or comprises an mRNA transcript. In some embodiments, an intravesicular RNA biomarker is or comprises a noncoding RNA. Exemplary noncoding RNAs may include, but are not limited to small nuclear RNA, microRNA (miRNA), small nucleolar RNA (snoRNA), circular RNA (circRNA), long noncoding RNA (lncRNA), small noncoding RNA, piwi-interacting RNA, etc.). Certain RNA biomarkers for cancer are described in the art, e.g., as described in Xi *et al.* "RNA Biomarkers: Frontier of Precision

Medicine for Cancer" *Noncoding RNA* (2017) 3:9, the contents of which are incorporated herein by reference for purposes described herein. In some embodiments, an intravesicular RNA biomarker is or comprise an orphan noncoding RNA (oncRNA). Certain oncRNAs that are cancer-specific were identified and described in the art, e.g., as described in Teng et al. "Orphan noncoding RNAs: novel regulators and cancer biomarkers" *Ann Transl Med* (2019) 7:S21; Fish et al. "Cancer cells exploit an orphan RNA to drive metastatic progression" *Nature Medicine* (2018) 24: 1743-1751; International Patent Publication WO 2019/094780, each of which are incorporated herein by reference for purposes described herein. In some embodiments, an intravesicular RNA biomarker is or comprises a long non-coding RNA. Certain non-coding RNA biomarkers for cancer are described in the art, e.g., as described in Qian *et al.* "Long Non-coding RNAs in Cancer: Implications for Diagnosis, Prognosis, and Therapy" *Front. Med.* (2020) Volume 7, Article 612393, the contents of which are incorporated herein by reference for purposes described herein. In some embodiments, an intravesicular RNA biomarker is or comprises piwiRNA. In some embodiments, an intravesicular RNA biomarker is or comprises miRNA. In some embodiments, an intravesicular RNA biomarker is or comprises snoRNA. In some embodiments, an intravesicular RNA biomarker is or comprises circRNA.

**[133]** ***Ligase:*** As used herein, the term "ligase" or "nucleic acid ligase" refers to an enzyme for use in ligating nucleic acids. In some embodiments, a ligase is enzyme for use in ligating a 3′-end of a polynucleotide to a 5′-end of a polynucleotide. In some embodiments, a ligase is an enzyme for use to perform a sticky-end ligation. In some embodiments, a ligase is an enzyme for use to perform a blunt-end ligation. In some embodiments, a ligase is or comprises a DNA ligase.

**[134]** ***Life-history-associated risk factors****:* As used herein, the term "life-history risk factors" refers to individuals' actions, experiences, medical history, and/or exposures in their lives which may directly or indirectly increase such individuals' risk for a condition, *e.g.*, lung cancer, relative to individuals who do not have such actions, experiences, medical history, and/or exposures in their lives. In some embodiments, non-limiting examples of life-history-associated risk factors include smoking, alcohol, drugs, carcinogenic agents, diet, obesity, diabetes, chronic obstructive pulmonary disease (COPD), physical activity, sun exposure, radiation exposure, bituminous smoke exposure, exposure to infectious agents such as viruses and bacteria, and/or occupational hazard (see *e.g.,* Jyoti Malhotra *et al., "Risk Factors for Lung Cancer Worldwide"*

*European Respiratory journal* (2016) 48: 889-902; which is incorporated herein by reference for the purpose described herein). One skilled in the art recognizes that the above list of life-history-associated risk factors contributing to cancer (*e.g.* lung cancer) susceptibility is not exhaustive but constantly evolving.

**[135]** ***Ligation:*** As used herein, the term "ligate", "ligating or "ligation" refers to a method or composition known in the art for joining two oligonucleotides or polynucleotides. A ligation may be or comprise a sticky-end ligation or a blunt-end ligation. In some embodiments, ligation involved in provided technologies is or comprises a sticky-end ligation. In some embodiments, ligation refers to joining a 3′ end of a polynucleotide to a 5′ end of a polynucleotide. In some embodiments, ligation is facilitated by use of a nucleic acid ligase.

**[136]** ***Nanoparticles:*** The term "nanoparticles" as used in the context of a sample for a detection assay (e.g., as described herein) refers to particles having a size range of about 30 nm to about 1000 nm. In some embodiments, nanoparticles have a size range of about 30 nm to about 750 nm. In some embodiments, nanoparticles have a size range of about 50 nm to about 750 nm. In some embodiments, nanoparticles have a size range of about 30 nm to about 500 nm. In some embodiments, nanoparticles have a size range of about 50 nm to about 500 nm. In some embodiments, nanoparticles are obtained from a bodily fluid sample of a subject, for example, in some embodiments by a size exclusion-based method (e.g., in some embodiments size exclusion chromatography). In some embodiments, nanoparticles are or comprise analyte aggregates, which in some embodiments may be or comprise protein or mucin aggregates. In some embodiments, nanoparticles are or comprise protein multimers. In some embodiments, nanoparticles are or comprise extracellular vesicles.

**[137]** ***Non-cancer subjects:*** As used herein, the term "non-cancer subjects" generally refers to subjects who do not have non-benign lung cancer. For example, in some embodiments, a non-cancer subject is a healthy subject. In some embodiments, a non-cancer subject is a healthy subject below age 55. In some embodiments, a non-cancer subject is a healthy subject of age 55 or above. In some embodiments, a non-cancer subject is a subject with non-lung related health diseases, disorders, or conditions. In some embodiments, a non-cancer subject is a subject having a benign lung tumor (*e.g.*, a benign mass observed in the thoracic or pulmonary cavity).

**[138]** ***Nucleic acid/ Oligonucleotide*:** As used herein, the term "nucleic acid" refers to a polymer of at least 10 nucleotides or more. In some embodiments, a nucleic acid is or comprises

DNA. In some embodiments, a nucleic acid is or comprises RNA. In some embodiments, a nucleic acid is or comprises peptide nucleic acid (PNA). In some embodiments, a nucleic acid is or comprises a single stranded nucleic acid. In some embodiments, a nucleic acid is or comprises a double-stranded nucleic acid. In some embodiments, a nucleic acid comprises both single and double-stranded portions. In some embodiments, a nucleic acid comprises a backbone that comprises one or more phosphodiester linkages. In some embodiments, a nucleic acid comprises a backbone that comprises both phosphodiester and non-phosphodiester linkages. For example, in some embodiments, a nucleic acid may comprise a backbone that comprises one or more phosphorothioate or 5'-N-phosphoramidite linkages and/or one or more peptide bonds, *e.g.*, as in a "peptide nucleic acid". In some embodiments, a nucleic acid comprises one or more, or all, natural residues (*e.g.*, adenine, cytosine, deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine, guanine, thymine, uracil). In some embodiments, a nucleic acid comprises on or more, or all, non-natural residues. In some embodiments, a non-natural residue comprises a nucleoside analog (*e.g.*, 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3 - methyl adenosine, 5-methylcytidine, C-5 propynyl-cytidine, C-5 propynyl-uridine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5 - propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, 6-O-methylguanine, 2-thiocytidine, methylated bases, intercalated bases, and combinations thereof). In some embodiments, a non-natural residue comprises one or more modified sugars (*e.g.*, 2'-fluororibose, ribose, 2'-deoxyribose, arabinose, and hexose) as compared to those in natural residues. In some embodiments, a nucleic acid has a nucleotide sequence that encodes a functional gene product such as an RNA or polypeptide. In some embodiments, a nucleic acid has a nucleotide sequence that comprises one or more introns. In some embodiments, a nucleic acid may be prepared by isolation from a natural source, enzymatic synthesis (*e.g.*, by polymerization based on a complementary template, *e.g.*, *in vivo* or *in vitro*, reproduction in a recombinant cell or system, or chemical synthesis. In some embodiments, a nucleic acid is at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10,000, 10,500, 11,000,

11,500, 12,000, 12,500, 13,000, 13,500, 14,000, 14,500, 15,000, 15,500, 16,000, 16,500, 17,000, 17,500, 18,000, 18,500, 19,000, 19,500, or 20,000 or more residues or nucleotides long.

**[139]** ***Nucleotide:*** As used herein, the term "nucleotide" refers to its art-recognized meaning. When a number of nucleotides is used as an indication of size, *e.g.*, of an oligonucleotide, a certain number of nucleotides refers to the number of nucleotides on a single strand, *e.g.*, of an oligonucleotide.

**[140]** ***Patient:*** As used herein, the term "patient" refers to any organism who is suffering or at risk of a disease or disorder or condition. Typical patients include animals (*e.g.*, mammals such as mice, rats, rabbits, non-human primates, and/or humans). In some embodiments, a patient is a human. In some embodiments, a patient is suffering from or susceptible to one or more diseases or disorders or conditions. In some embodiments, a patient displays one or more symptoms of a disease or disorder or condition. In some embodiments, a patient has been diagnosed with one or more diseases or disorders or conditions. In some embodiments, a disease or disorder or condition that is amenable to provided technologies is or includes cancer, or presence of one or more tumors. In some embodiments, a patient is receiving or has received certain therapy to diagnose and/or to treat a disease, disorder, or condition.

**[141]** ***Polypeptide*:** The term "polypeptide", as used herein, typically has its art-recognized meaning of a polymer of at least three amino acids or more. Those of ordinary skill in the art will appreciate that the term "polypeptide" is intended to be sufficiently general as to encompass not only polypeptides having a complete sequence recited herein, but also to encompass polypeptides that represent functional, biologically active, or characteristic fragments, portions or domains (*e.g.*, fragments, portions, or domains retaining at least one activity) of such complete polypeptides. In some embodiments, polypeptides may contain L-amino acids, D-amino acids, or both and/or may contain any of a variety of amino acid modifications or analogs known in the art. Useful modifications include, *e.g.*, terminal acetylation, amidation, glycosylation, lipidation, methylation, *etc*. In some embodiments, polypeptides may comprise natural amino acids, non-natural amino acids, synthetic amino acids, and combinations thereof (*e.g.,* may be or comprise peptidomimetics).

**[142]** ***Prevent or prevention:*** As used herein, "prevent" or "prevention," when used in connection with the occurrence of a disease, disorder, and/or condition, refers to reducing the risk of developing the disease, disorder and/or condition and/or to delaying onset of one or more

characteristics or symptoms of the disease, disorder or condition. Prevention may be considered complete when onset of a disease, disorder or condition has been delayed for a predefined period of time.

**[143]** ***Primer:*** As used herein, the term "primer" refers to an oligonucleotide capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced (*e.g.*, in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH). A primer is preferably single stranded for maximum efficiency in amplification. A primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of a primer can depend on many factors, *e.g.*, desired annealing temperature, etc.

**[144]** ***Reference:*** As used herein, "reference" describes a standard or control relative to which a comparison is performed. For example, in some embodiments, an agent, animal, individual, population, sample, sequence or value of interest is compared with a reference or control agent, animal, individual, population, sample, sequence or value. In some embodiments, a reference or control is tested and/or determined substantially simultaneously with the testing or determination of interest. In some embodiments, a reference or control is a historical reference or control, optionally embodied in a tangible medium. In some embodiments, a reference or control in the context of a reference level of a target refers to a level of a target in a normal healthy subject or a population of normal healthy subjects. In some embodiments, a reference or control in the context of a reference level of a target refers to a level of a target in a subject prior to a treatment. Typically, as would be understood by those skilled in the art, a reference or control is determined or characterized under comparable conditions or circumstances to those under assessment. In some embodiments, cell-line-derived extracellular vesicles are used as a reference or control. Those skilled in the art will appreciate when sufficient similarities are present to justify reliance on and/or comparison to a particular possible reference or control.

**[145]** ***Risk:*** As will be understood from context, "risk" of a disease, disorder, and/or condition refers to a likelihood that a particular individual will develop the disease, disorder, and/or condition. In some embodiments, risk is expressed as a percentage. In some embodiments, risk is from 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90 up to 100%. In some embodiments risk is expressed as a risk relative to a risk associated with a reference

sample or group of reference samples. In some embodiments, a reference sample or group of reference samples have a known risk of a disease, disorder, condition and/or event. In some embodiments a reference sample or group of reference samples are from individuals comparable to a particular individual. In some embodiments, relative risk is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more.

**[146]** ***Sample:*** As used herein, the term "sample" typically refers to an aliquot of material obtained or derived from a source of interest. In some embodiments, a sample is obtained or derived from a biological source (*e.g.*, a tissue or organism or cell culture) of interest. In some embodiments, a source of interest may be or comprise a cell or an organism, such as an animal or human. In some embodiments, a source of interest is or comprises biological tissue or fluid. In some embodiments, a biological tissue or fluid may be or comprise amniotic fluid, aqueous humor, ascites, bile, bone marrow, blood, breast milk, cerebrospinal fluid, cerumen, chyle, chime, ejaculate, endolymph, exudate, feces, gastric acid, gastric juice, lymph, mucus, pericardial fluid, perilymph, peritoneal fluid, pleural fluid, pus, rheum, saliva, sebum, semen, serum, smegma, sputum, synovial fluid, sweat, tears, urine, vaginal secretions, vitreous humor, vomit, and/or combinations or component(s) thereof. In some embodiments, a biological fluid may be or comprise an intracellular fluid, an extracellular fluid, an intravesicular fluid (blood plasma), an interstitial fluid, a lymphatic fluid, and/or a transcellular fluid. In some embodiments, a biological tissue or sample may be obtained, for example, by aspirate, biopsy (*e.g.*, fine needle or tissue biopsy), swab (*e.g.*, oral, nasal, skin, or vaginal swab), scraping, surgery, washing or lavage (*e.g.*, bronchoalveolar, ductal, nasal, ocular, oral, uterine, vaginal, or other washing or lavage). In some embodiments, a biological sample is or comprises a bodily fluid sample or a bodily fluid-derived sample. Examples of a bodily fluid sample or a bodily fluid-derived sample include, but are not limited to an amniotic fluid, bile, blood, breast milk, bronchoalveolar lavage fluid (BAL), cerebrospinal fluid, dialysate, feces, saliva, semen, synovial fluid, tears, urine, etc. In some embodiments, a bodily fluid sample or a bodily fluid-derived sample that may be useful in accordance with the present disclosure is or comprises a blood-derived sample, a saliva-derived sample, a sputum-derived sample, or a pleural effusion-derived sample. In some embodiments, a biological sample is or comprises a liquid biopsy. In some embodiments, a biological sample is or comprises cells obtained from an individual. In some embodiments, a sample is a "primary sample" obtained directly from a source of interest by any

appropriate means. In some embodiments, as will be clear from context, the term "sample" refers to a preparation that is obtained by processing (*e.g.*, by removing one or more components of and/or by adding one or more agents to) a primary sample. For example, a sample is a preparation that is processed by using a semi-permeable membrane or an affinity-based method such antibody-based method to separate a biological entity of interest from other non-target entities. Such a "processed sample" may comprise, for example, in some embodiments nanoparticles (e.g., nanoparticles having a size range of interest that includes extracellular vesicles), while, in some embodiments, nucleic acids and/or proteins, *etc.*, extracted from a sample. In some embodiments, a processed sample can be obtained by subjecting a primary sample to one or more techniques such as amplification or reverse transcription of nucleic acid, isolation and/or purification of certain components, *etc.*

**[147]** ***Selective or specific:*** The term "selective" or "specific", when used herein with reference to an agent having an activity, is understood by those skilled in the art to mean that the agent discriminates between potential target entities, states, or cells. For example, in some embodiments, an agent is said to bind "specifically" to its target if it binds preferentially with that target in the presence of one or more competing alternative targets. In many embodiments, specific interaction is dependent upon the presence of a particular structural feature of the target entity (*e.g.*, an epitope, a cleft, a binding site). It is to be understood that specificity need not be absolute. In some embodiments, specificity may be evaluated relative to that of a target-binding moiety for one or more other potential target entities (*e.g.*, competitors). In some embodiments, specificity is evaluated relative to that of a reference specific binding moiety. In some embodiments, specificity is evaluated relative to that of a reference non-specific binding moiety. In some embodiments, a target-binding moiety does not detectably bind to the competing alternative target under conditions of binding to its target entity. In some embodiments, a target-binding moiety binds with higher on-rate, lower off-rate, increased affinity, decreased dissociation, and/or increased stability to its target entity as compared with the competing alternative target(s).

**[148]** ***Small molecule:*** As used herein, the term "small molecule" means a low molecular weight organic and/or inorganic compound. In general, a "small molecule" is a molecule that is less than about 5 kilodaltons (kD) in size. In some embodiments, a small molecule is less than about 4 kD, 3 kD, about 2 kD, or about 1 kD. In some embodiments, the

small molecule is less than about 800 Daltons (D), about 600 D, about 500 D, about 400 D, about 300 D, about 200 D, or about 100 D. In some embodiments, a small molecule is less than about 2000 g/mol, less than about 1500 g/mol, less than about 1000 g/mol, less than about 800 g/mol, or less than about 500 g/mol. In some embodiments, a small molecule is not a polymer. In some embodiments, a small molecule does not include a polymeric moiety. In some embodiments, a small molecule is not a protein or polypeptide (*e.g.*, is not an oligopeptide or peptide). In some embodiments, a small molecule is not a polynucleotide (*e.g.*, is not an oligonucleotide). In some embodiments, a small molecule is not a polysaccharide. In some embodiments, a small molecule does not comprise a polysaccharide (*e.g.*, is not a glycoprotein, proteoglycan, glycolipid, *etc.*). In some embodiments, a small molecule is not a lipid. In some embodiments, a small molecule is biologically active. In some embodiments, suitable small molecules may be identified by methods such as screening large libraries of compounds (Beck- Sickinger & Weber (2001) *Combinational Strategies in Biology and Chemistry* (John Wiley & Sons, Chichester, Sussex); by structure-activity relationship by nuclear magnetic resonance (Shuker *et al.* (1996) "Discovering high-affinity ligands for proteins: SAR by NMR." *Science* 274: 1531-1534); encoded self-assembling chemical libraries (Melkko *et al.* (2004) "Encoded self-assembling chemical libraries." *Nature Biotechnol.* 22: 568-574); DNA-templated chemistry (Gartner *et al.* (2004) "DNA-templated organic synthesis and selection of a library of macrocycles." *Science* 305: 1601-1605); dynamic combinatorial chemistry (Ramstrom & Lehn (2002) "Drug discovery by dynamic combinatorial libraries." *Nature Rev. Drug Discov.* 1: 26-36); tethering (Arkin & Wells (2004) "Small-molecule inhibitors of protein-protein interactions: progressing towards the dream." *Nature Rev. Drug Discov.* 3: 301-317); and speed screen (Muckenschnabel *et al.* (2004) "SpeedScreen: label-free liquid chromatography-mass spectrometry-based high- throughput screening for the discovery of orphan protein ligands." *Anal. Biochem.* 324: 241-249). In some embodiments, a small molecule may have a dissociation constant for a target in the nanomolar range.

**[149]** ***Specific binding:*** As used herein, the term "specific binding" refers to an ability to discriminate between possible binding partners in the environment in which binding is to occur. A target-binding moiety that interacts with one particular target when other potential targets are present is said to "bind specifically" to the target with which it interacts. In some embodiments, specific binding is assessed by detecting or determining degree of association

between a target-binding moiety and its partner; in some embodiments, specific binding is assessed by detecting or determining degree of dissociation of a target-binding moiety-partner complex; in some embodiments, specific binding is assessed by detecting or determining ability of a target-binding moiety to compete an alternative interaction between its partner and another entity. In some embodiments, specific binding is assessed by performing such detections or determinations across a range of concentrations.

**[150]** ***Stage of cancer:*** As used herein, the term "stage of cancer" refers to a qualitative or quantitative assessment of the level of advancement of a cancer (*e.g.*, lung cancer). In some embodiments, criteria used to determine the stage of a cancer may include, but are not limited to, one or more of where the cancer is located in a body, tumor size, whether the cancer has spread to lymph nodes, whether the cancer has spread to one or more different parts of the body, *etc.* In some embodiments, cancer may be staged using the AJCC staging system. The AJCC staging system is a classification system, developed by the American Joint Committee on Cancer for describing the extent of disease progress in cancer patients, which utilizes in part the TNM scoring system: **T**umor size, Lymph **N**odes affected, **M**etastases. In some embodiments, cancer may be staged using a classification system that in part involves the TNM scoring system, according to which T refers to the size and extent of the main tumor, usually called the primary tumor; N refers to the number of nearby lymph nodes that have cancer; and M refers to whether the cancer has metastasized. In some embodiments, a cancer may be referred to as Stage 0 (abnormal cells are present but have not spread to nearby tissue, also called carcinoma *in situ*, or CIS; CIS is not cancer, but it may become cancer), Stage I-III (cancer is present; the higher the number, the larger the tumor and the more it has spread into nearby tissues), or Stage IV (the cancer has spread to distant parts of the body). In some embodiments, a cancer may be assigned to a stage selected from the group consisting of: *in situ* (abnormal cells are present but have not spread to nearby tissue); localized (cancer is limited to the place where it started, with no sign that it has spread); regional (cancer has spread to nearby lymph nodes, tissues, or organs): distant (cancer has spread to distant parts of the body); and unknown (there is not enough information to figure out the stage).

**[151]** ***Subject*:** As used herein, the term "subject" refers to an organism from which a sample is obtained, *e.g.,* for experimental, diagnostic, prophylactic, and/or therapeutic purposes. Typical subjects include animals (*e.g.,* mammals such as mice, rats, rabbits, non-human

primates, domestic pets, *etc.*) and humans. In some embodiments, a subject is a human subject, *e.g.*, a human male or female subject. In some embodiments, a subject is suffering from lung cancer. In some embodiments, a subject is susceptible to lung cancer. In some embodiments, a subject displays one or more symptoms or characteristics of lung cancer. In some embodiments, a subject displays one or more non-specific symptoms of lung cancer. In some embodiments, a subject does not display any symptom or characteristic of lung cancer. In some embodiments, a subject is someone with one or more features characteristic of susceptibility to or risk of lung cancer. In some embodiments, a subject is a patient. In some embodiments, a subject is an individual to whom diagnosis and/or therapy is and/or has been administered. In some embodiments, a subject is a subject (*e.g.*, male or female subject) determined to have a thoracic or pulmonary mass(es). In some embodiments, a subject is an asymptotic subject. Such an symptomatic subject may be a subject (*e.g.*, male or female subject) at average population risk, with life-history associated risk, or with hereditary risk. For example, such an asymptomatic subject may be a subject who has a family history of cancer, who has been previously treated for cancer, who is at risk of cancer recurrence after cancer treatment, who is in remission after cancer treatment, and/or who has been previously or periodically screened for the presence of at least one cancer biomarker. Alternatively, in some embodiments, an asymptomatic subject may be a subject who has not been previously screened for cancer, who has not been diagnosed for cancer, and/or who has not previously received cancer therapy. In some embodiments, a subject amenable to provided technologies is an individual selected based on one or more characteristics such as age, race, geographical location, genetic history, medical history, personal history (*e.g.*, smoking, alcohol, drugs, carcinogenic agents, diet, obesity, physical activity, sun exposure, radiation exposure, exposure to infectious agents such as viruses, and/or occupational hazard).

**[152]** ***Suffering from*:** An individual who is "suffering from" a disease, disorder, and/or condition has been diagnosed with and/or displays one or more symptoms of a disease, disorder, and/or condition.

**[153]** ***Surface analyte:*** As used herein, a "surface analyte" refers to an analyte present on the surface of a biological entity (e.g., a cell or a nanoparticle from a biological sample). In some embodiments, a surface analyte is or comprises a surface polypeptide or surface protein. In some embodiments, a surface analyte is or comprises a glycan.

**[154]** ***Surface biomarker:*** As used herein, a "surface biomarker" refers to a marker indicative of the state (e.g., presence, level, and/or activity) of a surface analyte (e.g., as described herein) of a biological entity (e.g., a cell or a nanoparticle including, e.g., in some embodiments an analyte aggregate (e.g., a protein or mucin aggregate) and/or an extracellular vesicle). In some embodiments, a surface biomarker is or comprises a surface protein biomarker. In some embodiments, a surface biomarker is or comprises a carbohydrate-dependent marker.

**[155]** ***Surface polypeptide or surface protein:*** As used interchangeably herein, the terms "surface polypeptide," and "surface protein" refer to a polypeptide or protein present in and/or on the surface of a biological entity (e.g., a cell or a nanoparticle including, e.g., in some embodiments an analyte aggregate (e.g., a protein or mucin aggregate) and/or an extracellular vesicle, etc.) through direct or indirect interactions. As will be understood by a skilled artisan, a surface protein, in some embodiments, may comprise a post-translational modification, including, e.g., but not limited to glycosylation. In some embodiments, a surface polypeptide or protein may be or comprise a membrane-bound polypeptide. In some embodiments, a membrane-bound polypeptide refers to a polypeptide or protein with one or more domains or regions present in and/or on the surface of the membrane of a biological entity (e.g., a cell, an extracellular vesicle, etc.). In some embodiments, a membrane-bound polypeptide may comprise one or more domains or regions spanning and/or associated with the plasma membrane of a biological entity (e.g., a cell, an extracellular vesicle, etc.). In some embodiments, a -bound polypeptide may comprise one or more domains or regions spanning and/or associated with the plasma membrane of a biological entity (e.g., a cell, an extracellular vesicle, etc.) and also protruding into the intracellular and/or intravesicular space. In some embodiments, a membrane-bound polypeptide may comprise one or more domains or regions associated with the plasma membrane of a biological entity (e.g., a cell, an extracellular vesicle, etc.), for example, via one or more non-peptidic linkages (e.g., through a glycosylphosphatidylinositol (GPI) anchor or lipidification or through non-covalent interaction). In some embodiments, a membrane-bound polypeptide may comprise one or more domains or regions that is/are anchored into either side of plasma membrane of a biological entity (e.g., a cell, an extracellular vesicle, etc.). In some embodiments, a surface protein is associated with or present on the surface of a nanoparticle (e.g., as described herein). In some embodiments, a surface protein is associated with or present within an extracellular vesicle. In some embodiments, a surface protein may be associated with

or present within a lung cancer-associated extracellular vesicle (e.g., an extracellular vesicle obtained or derived from a bodily fluid-derived sample (e.g., but not limited to a blood-derived sample) of a subject suffering from or susceptible to lung cancer). As will be understood by a skilled artisan, detection of the presence of at least a portion of a surface polypeptide or surface protein on/within extracellular vesicles can facilitate separation and/or isolation of lung cancer-associated extracellular vesicles from a biological sample (e.g., in some embodiments a bodily fluid-derived sample such as, e.g., but not limited to a blood-derived sample) (e.g., a blood or blood-derived sample) from a subject. In some embodiments, detection of the presence of a surface polypeptide or surface protein may be or comprise detection of an intravesicular portion (e.g., an intravesicular epitope) of such a surface polypeptide or surface protein. In some embodiments, detection of the presence of a surface polypeptide or surface protein may be or comprise detection of a membrane-spanning portion of such a surface polypeptide or surface protein. In some embodiments, detection of the presence of a surface polypeptide or surface protein may be or comprise detection of an extravesicular portion of such a surface polypeptide or surface protein.

**[156]** ***Surface protein biomarker:*** As used herein, the term "surface protein biomarker" refers to a marker indicative of the state (e.g., presence, level, and/or activity) of a surface protein (e.g., as described herein) of a biological entity (e.g., a cell or a nanoparticle including, e.g., in some embodiments an analyte aggregate (e.g., a protein or mucin aggregate) and/or an extracellular vesicle). In some embodiments, a surface protein refers to a polypeptide or protein with one or more domains or regions located in or on the surface of the membrane of a biological entity (e.g., a cell or an extracellular vesicle). In some embodiments, a surface protein biomarker may be or comprise an epitope that is present on the interior side (intravesicular) or the exterior side (extravesicular) of the membrane. In some embodiments, a surface protein biomarker is associated with or present in an extracellular vesicle. In some embodiments, a surface protein biomarker may be or comprise a mutated polypeptide. In some embodiments, a surface protein biomarker may be post-translationally modified (e.g., but not limited to glycosylated, phosphorylated, etc.). In some embodiments, a surface protein biomarker may be post-translationally processed and present in the form of a truncated polypeptide, for example, as a result of proteolytic cleavage). In some embodiments, a surface-protein biomarker may be or comprise an epitope that is present on the exterior surface of a nanoparticle.

**[157]** ***Susceptible to:*** An individual who is "susceptible to" a disease, disorder, and/or condition is one who has a higher risk of developing the disease, disorder, and/or condition than does a member of the general public. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition may not have been diagnosed with the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition may exhibit symptoms of the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition may not exhibit symptoms of the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition will develop the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition will not develop the disease, disorder, and/or condition.

**[158]** ***Target-binding moiety***: In general, the terms "target-binding moiety" and "binding moiety" are used interchangeably herein to refer to any entity or moiety that binds to a target of interest (*e.g.*, molecular target of interest such as a biomarker or an epitope). In many embodiments, a target-binding moiety of interest is one that binds specifically with its target (*e.g.*, a target biomarker) in that it discriminates its target from other potential binding partners in a particular interaction context. In general, a target-binding moiety may be or comprise an entity or moiety of any chemical class (*e.g.*, polymer, non-polymer, small molecule, polypeptide, carbohydrate, lipid, nucleic acid, *etc.*). In some embodiments, a target-binding moiety is a single chemical entity. In some embodiments, a target-binding moiety is a complex of two or more discrete chemical entities associated with one another under relevant conditions by non-covalent interactions. For example, those skilled in the art will appreciate that in some embodiments, a target-binding moiety may comprise a "generic" binding moiety (*e.g.*, one of biotin/avidin/streptavidin and/or a class-specific antibody) and a "specific" binding moiety (*e.g.*, an antibody or aptamers with a particular molecular target) that is linked to the partner of the generic biding moiety. In some embodiments, such an approach can permit modular assembly of multiple target binding moieties through linkage of different specific binding moieties with a generic binding moiety partner.

**[159]** ***Target biomarker signature:*** The term "target biomarker signature", as used herein, refers to a combination of (*e.g.*, at least 2 or more, including, *e.g.*, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at

least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 25, at least 30, or more) biomarkers, which combination correlates with a particular biological event or state of interest, so that one skilled in the art will appreciate that it may appropriately be considered to be a "signature" of that event or state. To give but a few examples, in some embodiments, a target biomarker signature may correlate with a particular disease or disease state, and/or with likelihood that a particular disease, disorder or condition may develop, occur, or reoccur. In some embodiments, a target biomarker signature may correlate with a particular disease or therapeutic outcome, or likelihood thereof. In some embodiments, a target biomarker signature may correlate with a specific cancer and/or stage thereof. In some embodiments, a target biomarker signature may correlate with lung cancer and/or a stage and/or a subtype thereof. In some embodiments, a target biomarker signature comprises a combination of (*e.g.*, at least 2 or more, including, *e.g.*, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 25, at least 30, or more) biomarkers that together are specific for an lung cancer or a subtype and/or a disease stage thereof), though one or more biomarkers in such a combination may be directed to a target (*e.g.*, a surface biomarker, an intravesicular biomarker, and/or an intravesicular RNA) that is not specific to the lung cancer. For example, in some embodiments, a target biomarker signature may comprise at least one biomarker specific to an lung cancer or a stage and/or subtype thereof (*i.e.,* an lung cancer-specific target), and may further comprise a biomarker that is not necessarily or completely specific for the lung cancer (*e.g.,* that may also be found on some or all biological entities such as, *e.g.*, cells, nanoparticles, *etc.*, that are not cancerous, are not of the relevant cancer, and/or are not of the particular stage and/or subtype of interest). That is, as will be appreciated by those skilled in the art reading the present specification, so long as a combination of biomarkers utilized in a target biomarker signature is or comprises a plurality of biomarkers that together are specific for the relevant target biological entities of interest (*e.g.*, lung cancer cells of interest or nanoparticles secreted by lung cancer cells such as lung cancer-specific extracellular vesicles) (*i.e.,* sufficiently distinguish the relevant target biological entities (*e.g.*, lung cancer cells of interest or nanoparticles secreted by lung cancer cells such as lung cancer-specific extracellular vesicles) for detection from other biological entities not of interest for detection), such a combination of biomarkers is a useful target biomarker signature in accordance with certain embodiments of the present disclosure.

**[160]** ***Therapeutic agent:*** As used interchangeably herein, the phrase "therapeutic agent" or "therapy" refers to an agent or intervention that, when administered to a subject or a patient, has a therapeutic effect and/or elicits a desired biological and/or pharmacological effect. In some embodiments, a therapeutic agent or therapy is any substance that can be used to alleviate, ameliorate, relieve, inhibit, prevent, delay onset of, reduce severity of, and/or reduce incidence of one or more symptoms or features of a disease, disorder, and/or condition. In some embodiments, a therapeutic agent or therapy is a medical intervention (*e.g.*, surgery, radiation, phototherapy) that can be performed to alleviate, relieve, inhibit, present, delay onset of, reduce severity of, and/or reduce incidence of one or more symptoms or features of a disease, disorder, and/or condition.

**[161]** ***Threshold level (e.g., cutoff):*** As used herein, the term "threshold level" refers to a level that are used as a reference to attain information on and/or classify the results of a measurement, for example, the results of a measurement attained in an assay. For example, in some embodiments, a threshold level (*e.g.*, a cutoff) means a value measured in an assay that defines the dividing line between two subsets of a population (*e.g.*, normal and/or non-lung cancer vs. lung cancer). Thus, a value that is equal to or higher than the threshold level defines one subset of the population, and a value that is lower than the threshold level defines the other subset of the population. A threshold level can be determined based on one or more control samples or across a population of control samples. A threshold level can be determined prior to, concurrently with, or after the measurement of interest is taken. In some embodiments, a threshold level can be a range of values.

**[162]** ***Treat:*** As used herein, the term "treat," "treatment," or "treating" refers to any method used to partially or completely alleviate, ameliorate, relieve, inhibit, prevent, delay onset of, reduce severity of, and/or reduce incidence of one or more symptoms or features of a disease, disorder, and/or condition. Treatment may be administered to a subject who does not exhibit signs of a disease, disorder, and/or condition. In some embodiments, treatment may be administered to a subject who exhibits only early signs of the disease, disorder, and/or condition, for example for the purpose of decreasing the risk of developing pathology associated with the disease, disorder, and/or condition. In some embodiments, treatment may be administered to a subject at a later-stage of disease, disorder, and/or condition.

**[163]** Standard techniques may be used for recombinant DNA, oligonucleotide synthesis, and tissue culture and transformation (*e.g.*, electroporation, lipofection). Enzymatic reactions and purification techniques may be performed according to manufacturer's specifications or as commonly accomplished in the art or as described herein. The foregoing techniques and procedures may be generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification. See *e.g.*, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)), which is incorporated herein by reference for the purpose described herein.

## DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS

**[164]** Lung cancer was responsible for an estimated 148,869 deaths in 2016 in the United States (U.S. Cancer statistics working group, 2019; which is incorporated herein by reference for the purpose described herein). The majority of these deaths are attributable to late diagnosis; the 5-year total survival rate for lung cancer in the United States from 2001 to 2007 was 15.6%. Patients with localized disease at diagnosis had a 5-year survival rate of 52%; however, the majority of patients received initial diagnosis when distant metastasis had already formed and those patients have a dismal 5-year survival rate of approximately 3.6% (Cruz *et al.*, 2011; which is incorporated herein by reference for the purpose described herein).

**[165]** Unfortunately, there are no inexpensive, and widely available recommended lung cancer screening tests for average-risk individuals. While many individuals at hereditary or life-history associated risk are currently screened by low-dose CT scanning, these tests are suboptimal for screening, because they are relatively expensive and of limited accessibility. For example, the Prostate, Lung, Colorectal and Ovarian Cancer Screening Randomized Trial found relatively cost effective and widely available chest X-ray imaging and sputum testing to be ineffective in altering lung cancer mortality rates, while low-dose CT scanning does successfully lower mortality rates, it often increases the number of unnecessary surgeries and may be considered costly, and of limited availability.

**[166]** The present disclosure, among other things, identifies the source of a problem with certain prior technologies including, for example, certain conventional approaches to

detection and diagnosis of lung cancer. For example, the present disclosure appreciates that many conventional diagnostic assays, *e.g.*, X-ray imaging, sputum testing, low-dose CT scanning, and/ or molecular tests based on cell-free nucleic acids, serum proteins (*e.g.*, CEA, CYFRA 21-1, NSE, ProGRP, and/or SCCA) and/or bulk analysis of extracellular vesicles, can be time-consuming, costly, and/or lacking sensitivity and/or specificity sufficient to provide a reliable and comprehensive diagnostic assessment. In some embodiments, the present disclosure provides technologies (including systems, compositions, and methods) that solve such problems, among other things, by identification of biomarker combinations that are predicted to exhibit high sensitivity and specificity for lung cancer based on bioinformatics analysis. In some embodiments, the present disclosure provides technologies (including systems, compositions, and methods) that solve such problems, by detecting co-localization of a target biomarker signature of lung cancer (*e.g.*, identified by bioinformatics analysis) in individual nanoparticles having a size range of interest that includes extracellular vesicles, which comprises at least one extracellular vesicle-associated surface biomarker and at least one target biomarker selected from the group consisting of surface biomarkers, internal biomarkers, and RNA biomarkers present in nanoparticles associated with lung cancer. In some embodiments, the present disclosure provides technologies (including systems, compositions, and methods) that solve such problems, among other things, by detecting such target biomarker signature of lung cancer using a target entity detection approach that was developed by Applicant and described in U.S. Application No. US2020/0299780, and International Application WO2020180741), which are based on interaction and/or co-localization of a target biomarker signature in individual nanoparticles. The contents of each of the aforementioned disclosures is incorporated herein by reference in their entirety.

**[167]** In some embodiments, extracellular vesicles for detection as described herein can be isolated from a bodily fluid of a subject by a size exclusion-based method. As will be understood by a skilled artisan, in some embodiments, a size exclusion-based method may provide a sample comprising nanoparticles having a size range of interest that includes extracellular vesicles. Accordingly, in some embodiments, provided technologies of the present disclosure encompass detection, in individual nanoparticles having a size range of interest (e.g., in some embodiments about 30 nm to about 1000 nm) that includes extracellular vesicles, of co-localization of at least two or more surface biomarkers (e.g., as described herein) that forms a

target biomarker signature of lung cancer. A skilled artisan reading the present disclosure will understand that various embodiments described herein in the context of "extracellular vesicle(s)" (e.g., assays for detecting individual extracellular vesicles and/or provided "extracellular vesicle-associated surface biomarkers") can be also applicable in the context of "nanoparticles" as described herein.

**[168]** The present disclosure, among other things, provides insights and technologies for achieving effective lung cancer screening, *e.g.,* for early detection of lung cancer. In some embodiments, the present disclosure provides technologies for early detection of lung cancer in subjects who may be experiencing one more symptoms associated with lung cancer. In some embodiments, the present disclosure provides technologies for early detection of lung cancer in subjects who are at hereditary risks for lung cancer. In some embodiments, the present disclosure provides technologies for early detection of lung cancer in subjects who may be at hereditary risk and/or experiencing one or more symptoms associated with lung cancer. In some embodiments, the present disclosure provides technologies for early detection of lung cancer in subjects who may have life-history risk factors (*e.g.*, but not limited to smoking). In some embodiments, the present disclosure provides technologies for screening individuals, *e.g.*, individuals with certain risks (*e.g.*, hereditary risk, life-history associated risk, or average risk), for early-stage non-small cell lung cancers such as, *e.g.*, lung adenocarcinoma (LUAD), and lung squamous cell carcinoma (LUSC). Non-small cell lung cancers are the most common subtype of lung cancer, in which 54% of cases are detected at an advanced stage (SEER Cancer Statistics Review 1975-2017). In some embodiments, provided technologies are effective for detection of early-stage lung cancers. In some embodiments, provided technologies are effective when applied to populations comprising or consisting of individuals having one or more symptoms that may be associated with lung cancer. In some embodiments, provided technologies are effective even when applied to populations comprising or consisting of asymptomatic or symptomatic individuals (*e.g.*, due to sufficiently high sensitivity and/or low rates of false positive and/or false negative results). In some embodiments, provided technologies are effective when applied to populations comprising or consisting of individuals (*e.g.*, asymptomatic, or symptomatic individuals) without hereditary risk, and/or life-history related risk of developing lung cancer. In some embodiments, provided technologies are effective when applied to populations comprising or consisting of individuals (*e.g.*, asymptomatic, or symptomatic individuals) with hereditary risk, and/or life-history related

risk of developing lung cancer. In some embodiments, provided technologies are effective when applied to populations comprising or consisting of individuals susceptible to lung cancer (*e.g.*, individuals with a known genetic, environmental, or experiential risk, *etc.*). In some embodiments, provided technologies may be or include one or more compositions (*e.g.*, molecular complexes, systems, collections, combinations, kits, *etc.*) and/or methods (*e.g.*, of making, using, assessing, *etc.*), as will be clear to one skilled in the art reading the disclosure provided herein.

**[169]** In some embodiments, provided technologies achieve detection (*e.g.*, early detection, *e.g.*, in asymptomatic individual(s) and/or population(s)) of one or more features (*e.g.*, incidence, progression, responsiveness to therapy, recurrence, *etc.*) of lung cancer, with sensitivity and/or specificity (*e.g.*, rate of false positive and/or false negative results) appropriate to permit useful application of provided technologies to single-time and/or regular (*e.g.*, periodic) assessment. In some embodiments, provided technologies are useful in conjunction with an individual's regular medical examinations, such as but not limited to: physicals, general practitioner visits, cholesterol/lipid blood tests, diabetes (type 2) screening, colonoscopies, blood pressure screening, thyroid function tests, prostate cancer screening, mammograms, HPV/Pap smears, and/or vaccinations. In some embodiments, provided technologies are useful in conjunction with treatment regimen(s); in some embodiments, provided technologies may improve one or more characteristics (*e.g.*, rate of success according to an accepted parameter) of such treatment regimen(s).

**[170]** In some embodiments, the present disclosure, among other things, provides insights that screening of asymptotic individuals, *e.g.*, regular screening prior to or otherwise in absence of developed symptom(s), can be beneficial, and even important for effective management (*e.g.*, successful treatment) of lung cancer. In some embodiments, the present disclosure provides lung cancer screening systems that can be implemented to detect lung cancer, including early-stage cancer, in some embodiments in asymptomatic individuals (*e.g.*, without hereditary, and/or life-history associated risks in lung cancer). In some embodiments, provided technologies are implemented to achieve regular screening of asymptomatic individuals (*e.g.*, with or without hereditary and/or life-history associated risk(s) in lung cancer). In some embodiments, provided technologies are implemented to achieve regular screening of symptomatic individuals (*e.g.*, with or without hereditary and/or life-history associated risk(s) in

lung cancer). The present disclosure provides, for example, compositions (*e.g.*, reagents, kits, components, *etc.*), and methods of providing and/or using them, including strategies that involve regular testing of one or more individuals (*e.g.*, asymptomatic individuals). The present disclosure defines usefulness of such systems and provides compositions and methods for implementing them.

***I. Lung Cancer Detection***

**[171]** Today there is no lung cancer screening test of any kind that is CDC recommended for screening asymptomatic individuals of average risk, while in the USA the age-adjusted incidence rate of lung cancer is 62 per 100,000 men and women per year. Almost as many Americans die from lung cancer every year as die from prostate, breast, and colon cancer combined. In 2010 alone, there was an estimated ~240,000 new cases of lung cancer and ~161,000 deaths from lung cancer in the US (Cruz *et al.*, 2011; which is incorporated herein by reference for the purpose described herein). While the total number of deaths from lung cancer in the USA have been declining since ~1985, the global rate of lung cancer is increasing steadily and has increased ~51% from 1985 to 2010 (Cruz *et al.*, 2011; which is incorporated herein by reference for the purpose described herein). Globally, lung cancer is the largest contributor to new cancer cases, and in 2010 there were approximately 1,350,000 new lung cancer cases worldwide (~12.4% of all new cancer cases) and ~1,180,000 deaths (~17.6% of total cancer related deaths). Approximately half of these new lung cancer cases are occurring in developing countries. While the 5-year lung cancer survival rate in Europe, China, and developing countries has been estimated at only 8.9% (Cruz *et al.*, 2011; which is incorporated herein by reference for the purpose described herein).

**[172]** The Surveillance, Epidemiology and End Results (SEER) data from 2013-2017 has reported extensively on the prevalence and epidemiology of lung cancer in the United States of America for the last 45 years. SEER reported the median age at diagnosis for cancer of the lung and bronchus as ~71 years. Lung cancer arises from the cells of the respiratory epithelium and can be divided into two broad categories. Small cell lung cancer (SCLC) is highly malignant and derived from cells exhibiting neuroendocrine characteristics, SCLC makes up ~12 % of all lung cancer cases (*FIG. 3*). Non-small cell lung cancer (NSCLC) accounts for the remaining 85% of cases, and is divided into three pathological subtypes: adenocarcinoma, squamous cell

carcinoma, and large cell carcinoma. Adenocarcinoma itself accounts for ~50% of all new lung cancer cases in the United States (*FIG. 3*), with squamous cell carcinoma accounting for ~23%, and large cell carcinoma and other lung and bronchial cancers accounting for the remaining ~15% (*FIG. 3*).

**[173]** The 5-year survival rates for all patients with invasive non-small cell cancer of the lung and bronchus is ~25%. Patients with localized lung cancer at initial diagnosis have ~63% 5 year survival rate, patients with regional lung cancer metastasis at initial diagnosis have ~35% 5 year survival rates, while those with distant lung cancer metastasis have a poor ~7% 5 year survival rates (SEER 1975-2017 review, Table 15.12). These data indicate that diagnoses for early-stage lung cancer is important as it can increase survival rates of lung cancer patients.

**[174]** Certain risk factors for lung cancer include age and a history of smoking. The seminal report by the US surgeon general in 1964 stated: (1) Cigarette smoking was associated with a 70% increase in the age-specific death rates of men and a lesser increase in the death rates of women; (2) Cigarette smoking was causally related to lung cancer in men, the magnitude of the effect far outweighed all other factors leading to lung cancer, and the risk increased directly with the duration of smoking and the number of cigarettes smoked per day; (3) Cigarette smoking was believed more important than occupational exposures in the causation of lung cancer in the general population; (4) Cigarette smoking was reported as the important cause of chronic bronchitis in the United States; and (5) Male cigarette smokers had a higher death rate from coronary artery disease than male nonsmokers.

**[175]** The International Agency for Research on Cancer (IARC) has identified at least 50 known carcinogens in tobacco smoke. Examples of such carcinogens include but are not limited to tobacco-specific *N*-nitrosamines (TSNAs) formed by nitrosation of nicotine during tobacco processing and during smoking. The chemical 4-(methylnitrosamino)-1(3-pyridyl)-1-butanone (NNK) is known to induce adenocarcinoma of the lung in experimental animals. NNK is known to bind to DNA and create DNA adducts, leading to DNA damage. Failure to repair this damage can lead to permanent mutations. NNK is associated with DNA mutations resulting in the activation of K-ras oncogenes, which is detected in 24% of human lung adenocarcinomas.

**[176]** It is estimated that one in nine smokers eventually develops lung cancer. The relative risk of lung cancer in long-term smokers has been estimated as 10-fold to 30-fold greater than that of lifetime nonsmokers. In the USA, greater than 80% of lung cancers occur in persons

with tobacco exposure, while globally, 15% of lung cancers in men, and up to 53% in women are not attributable to smoking, with never smokers accounting for ~25% of all lung cancer cases worldwide.

**[177]** High risk individuals and/or populations as defined by the CDC are 55-77 years of age, have a >30 cigarette pack-year history, are current smokers, or quit smoking within the last 15 years. For these individuals low-dose CT scanning is currently the recommended lung cancer screening tool. However, low-dose CT in high-risk (*e.g.,* patients as defined by the CDC guidelines) populations can be considered relatively expensive, of limited access, and to have unreasonably high levels of false positives (*e.g.*, the proportion of all positive tests that were falsely positive may be as great as 97.5%; Raghu *et al.*, 2020 and Kinsinger *et al.*, 2017 which are both incorporated herein by reference for the purpose described herein). The present disclosure, among other things, provides a cost-effective screening assay with sufficiently high specificity and/or sensitivity.

**[178]** Among other things, in certain embodiments the present disclosure provides an insight that there is a need for development of a lung cancer liquid biopsy assay for screening subjects with a hereditary and/or life-history associated risk for lung cancer and/or subjects who may be experiencing one or more symptoms associated with lung cancer. In certain embodiments, the present disclosure provides an insight that there is a need for development of a lung cancer liquid biopsy assay for screening symptomatic or asymptomatic subjects *e.g.,* prior to other screening methods, *e.g.,* imaging methods for lung cancer detection such as, *e.g.*, MRI, CT scan, *etc.*

**[179]** In some embodiments, the present disclosure provides technologies for effective screening of lung cancer in individuals at hereditary risk, or in individuals with life-history associated-risks. In some embodiments, the present disclosure provides technologies for effective screening of lung cancer in average-risk individuals. In some embodiments, the present disclosure provides technologies for effective screening of lung cancer in individuals with one or more symptoms associated with lung cancer. In some embodiments, the present disclosure provides technologies for effective screening of lung cancer in asymptomatic individuals. Despite being the largest killer of men and women among all cancers, there is currently no recommended lung cancer screening tool that is non-invasive, using bodily fluid (including, e.g., but not limited to blood) only and for asymptomatic and/or average-risk individuals (*e.g.,*

individuals under the age of 55 years, or individuals over the age of 55 years who have no history of smoking or have quit smoking for more than 15 years). This is due, in part, to the cost, limited available, potential side effects, and/or poor performance (*e.g.,* high false positive rate, or ineffectualness) of existing lung cancer screening technologies. Given the incidence of lung cancer in average-risk individuals, inadequate test specificities (<99.5%) can result in false positive results that outnumber true positives by more than an order of magnitude. This places a significant burden on the healthcare system and on the individuals being screened as false positive results lead to additional tests, unnecessary surgeries, and emotional/physical distress (Wu *et al.*, 2016).

**[180]** In some embodiments, the present disclosure provides an insight that a particularly useful lung cancer screening test would be characterized by: (1) ultrahigh specificity (>99.5%) to minimize the number of false positives, and (2) high sensitivity (>40%) for stage I and II lung cancer (*i.e.*, when prognosis is most favorable).

**[181]** For example, in some embodiments, a particularly useful lung cancer screening test may be characterized by a specificity of >98% and a sensitivity of >50%, for example, for stage I and II lung cancer. In some embodiments, a particularly useful lung cancer screening test may be characterized by a specificity of >98% and a sensitivity of >60%, for example, for stage I and II lung cancer. In some embodiments, a particularly useful lung cancer screening test may be characterized by a specificity of >98% and a sensitivity of >70%, for example, for stage I and II lung cancer. In some embodiments, a particularly useful lung cancer screening test may be characterized by a specificity of >99.5% and a sensitivity of >65%, for example, for stage I and II lung cancer. In some embodiments, a particularly useful lung cancer screening test may be characterized by a specificity of >99.5% and a sensitivity of >60%, for example, for stage I and II lung cancer. In some embodiments, a particularly useful lung cancer screening test may be characterized by a specificity of 99% or higher and a sensitivity of >10% or higher (including, *e.g.*, >15%, >20%, >25%. In some embodiments, a particularly useful lung cancer screening test may be characterized by a specificity of 99% or higher and a sensitivity of 50% or higher.

**[182]** In some embodiments, the present disclosure provides an insight that a lung cancer screening test involving more than one set of biomarker combinations (*e.g.,* at least two orthogonal biomarker combinations as described herein) can increase sensitivity of such an assay, as compared to that is achieved by one set of biomarker combination. For example, in

some embodiments, a lung cancer screening test involving at least two orthogonal biomarker combinations can achieve a specificity of at least 98% and a sensitivity of at least 50%. In some embodiments, a lung cancer screening test involving at least two orthogonal biomarker combinations can achieve a specificity of at least 98% and a sensitivity of at least 60%. In some embodiments, a lung cancer screening test involving at least two orthogonal biomarker combinations can achieve a specificity of 99% and a sensitivity of 50% or higher.

**[183]** In some embodiments, the present disclosure provides an insight that a particularly useful lung cancer screening test may be characterized by an acceptable positive predictive value (PPV) at an economically justifiable cost. PPV is the likelihood a patient has the disease following a positive test, and is influenced by sensitivity, specificity, and/or disease prevalence. One clinician consensus for the minimum PPV needed to screen for lung cancer is 10%. With a 10% PPV, there would be nine false positives for every one true positive (Lung Cancer Screening: Recommendation Statement., *Am Fam Physician*. 2005 Mar 15;71(6):1165-1168). These false positives place a significant burden on both the healthcare system and subjects being screened as they lead to additional tests, unnecessary surgeries, and emotional and physical distress. In some embodiments, assays described herein are particularly useful for early lung cancer detection that achieves a PPV of greater than 10% or higher, including, *e.g.,* greater than 15%, greater than 20%, or greater than 25% or higher, with a specificity cutoff of at least 85% or higher, including, e.g., at least 90%, at least 95%, or higher (e.g., a specificity cutoff of at least 98% for subjects at hereditary risk for lung cancer, or with a specificity cutoff of at least 99.5% for subjects experiencing one or more symptoms associated with lung cancer).

**[184]** In some embodiments, assays described herein are particularly useful as a first screening test for early lung cancer detection. In some embodiments, subjects who have received a positive test result from assays described herein are recommended to receive a follow-up test. In some such embodiments, assays described herein can be useful for early lung cancer detection that achieves a PPV of greater than 2% or higher, including, *e.g.,* greater than 3%, greater than 4%, greater than 5%, greater than 6% greater than 7%, greater than 8%, greater than 9%, greater than 10%, greater than 15%, greater than 20%, or greater than 25% or higher. In some such embodiments, assays described herein can achieve a specificity cutoff of at least 85% or higher, including, e.g., at least 90%, at least 95% or higher (*e.g.*, a specificity cutoff of at least 98% for

subjects at hereditary risk for lung cancer, or with a specificity cutoff of at least 99.5% for subjects experiencing one or more symptoms associated with lung cancer).

**[185]** Several different biomarker classes have been studied for a lung cancer liquid biopsy assay including circulating tumor DNA (ctDNA), circulating tumor cells (CTCs), bulk proteins, and extracellular vesicles (EVs). EVs are particularly promising due to their abundance and stability in the bloodstream relative to ctDNA and CTCs, suggesting improved sensitivity for early-stage cancers. Moreover, EVs contain cargo (*i.e.*, proteins, RNA, metabolites) that originated from the same cell, providing superior specificity over bulk protein measurements. While the diagnostic utility EVs has been studied, much of this work has pertained to bulk EV measurements or low-throughput single-EV analyses.

***II. Provided Biomarkers and/or Target Biomarker Signatures for Detection of Lung Cancer***

**[186]** The present disclosure, among other things, provides various target biomarkers or combinations thereof (*e.g.*, target biomarker signatures) for lung cancer. Such target biomarker signatures that are predicted to exhibit high sensitivity and specificity for lung cancer were discovered by a multi-pronged bioinformatics analysis and biological approach, which for example, in some embodiments involve computational analysis of a diverse set of data, *e.g.*, in some embodiments comprising one or more of sequencing data, expression data, mass spectrometry, histology, post-translational modification data, and/or *in vitro* and/or *in vivo* experimental data through machine learning and/or computational modeling.

**[187]** In some embodiments, a target biomarker signature of lung cancer comprises at least one or more (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, or more) surface biomarkers (*e.g.*, in some embodiments surface polypeptide present in extracellular vesicles associated with lung cancer; "extracellular vesicle-associated surface biomarker") and at least one or more (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, or more) target biomarkers selected from the group consisting of surface biomarker(s), intravesicular biomarker(s), and intravesicular RNA biomarker(s), such that the combination of such surface biomarker(s) and such target biomarker(s) present a target biomarker signature of lung cancer that provides (a) high specificity (*e.g.*, greater than 98% or higher such as greater than 99%, or greater than 99.5%) to minimize the number of false positives, and (b) high sensitivity (*e.g.*, greater than 40%, greater than 50%, greater than 60%, greater than 70%, greater than 80%) for stage I and II lung cancer when prognosis is most favorable.

**[188]** In some embodiments, the present disclosure recognizes that in certain embodiments, sensitivity and specificity rates for subjects with different lung cancer risk levels may vary depending upon the risk tolerance of the attending physician and/or the guidelines set forth by interested medical consortia. In certain embodiments, subjects at risk of lung cancer may be served with an 85% specificity rate or higher (including, e.g., at least 90%, at least 95% or higher specificity rate) with 50% sensitivity or higher (including, e.g., at least 60%, at least 70%, at least 80%, or higher sensitivity). In certain embodiments, at risk subjects with life-history-associated risk factors may be served with an 85% specificity rate or higher (including, e.g., at least 90%, at least 95% or higher specificity rate) with 50% sensitivity or higher (including, e.g., at least 60%, at least 70%, at least 80%, or higher sensitivity). In certain embodiments, symptomatic subjects may be served with an 85% specificity rate or higher (including, e.g., at least 90%, at least 95% or higher specificity rate) with 50% sensitivity or higher (including, e.g., at least 60%, at least 70%, at least 80%, or higher sensitivity). In certain embodiments, non-symptomatic subjects may be served with an 85% specificity rate or higher (including, e.g., at least 90%, at least 95% or higher specificity rate) with 50% sensitivity or higher (including, e.g., at least 60%, at least 70%, at least 80%, or higher sensitivity). In certain embodiments, subjects at risk of lung cancer may be served with a 99.5% specificity rate with 70% sensitivity or a 98% specificity rate with 80% sensitivity. In certain embodiments, at risk subjects with life-history-associated risk factors may be served with a 99.5% specificity rate with 70% sensitivity or a 98% specificity rate with 80% sensitivity. In some embodiments, an assay described herein for detection of lung cancer in at-risk subjects (e.g., with life-history-associated risk factors) may have a set sensitivity rate that is lower than 80% sensitivity, including *e.g.*, less than 70%, less than 60%, less than 50% or lower sensitivity rate. In certain embodiments, non-symptomatic subjects may be served with a 99.5% specificity rate with 70% sensitivity or a 98% specificity rate with 80% sensitivity. In some embodiments, an assay described herein for detection of lung cancer in non-symptomatic subjects may have a set sensitivity rate that is lower than 80% sensitivity, including *e.g.*, less than 70%, less than 60%, less than 50% or lower sensitivity rate. In some embodiments, technologies and/or assays described herein for detection of lung cancer in a symptomatic subject may have a lower sensitivity and/or specificity requirement than those for detection of lung cancer in an asymptomatic subject. In some embodiments, an assay described herein for detection of lung cancer in a symptomatic subject may have a set specificity

rate that is lower than 99.5% specificity, including *e.g.*, less than 99% sensitivity, less than 95%, less than 90%, or less than 85% specificity rate. In some embodiments, an assay described herein for detection of lung cancer in a symptomatic subject may have a set sensitivity rate that is lower than 80% sensitivity, including *e.g.*, less than 70%, or less than 60% sensitivity rate.

**[189]** The present disclosure, among other things, observes that the gold standard for screening high-risk smokers is a chest CT, which had a reported positive predictive value of 3.8% in such a high-risk population in a National Lung Screening Trial study (National Lung Screening Trial Research Team (2013) "Results of initial low-dose computed tomographic screening for lung cancer. New England Journal of Medicine," 368(21): 1980-1991). In some embodiments, the present disclosure, among other things, appreciates that a biomarker signature of lung cancer that provides a positive predictive value (PPV) of 3.8% or higher is particularly useful for screening individuals at risk for lung cancer. In some embodiments, a target biomarker signature of lung cancer comprises at least one surface biomarker (*e.g.*, surface biomarker present on the surfaces of extracellular vesicles associated with lung cancer) and at least one target biomarker selected from the group consisting of surface biomarker(s), intravesicular biomarker(s), and intravesicular RNA biomarker(s), such that the combination of such surface biomarker(s) and such target biomarker(s) present a target biomarker signature of lung cancer that provides a positive predictive value (PPV) of at least 3.8% or higher, including, *e.g.*, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10% or higher, at least 15% or higher, at least 20% or higher, at least 25% or higher, and/or at least 30% or higher, in high-risk population.

**[190]** In general, gene identifiers used herein refer to the Gene Identification catalogued by the UniProt Consortium (UniProt.org); one skilled in the art will understand that certain genes can be known by multiple names and will also readily recognize such multiple names.

**[191]** In general, carbohydrate identifiers used herein refer to Kegg Cancer-associated Carbohydrates database (genome.jp/kegg/disease/br08441.html); one skilled in the art will understand that certain carbohydrates can be known by multiple names and will also readily recognize such multiple names.

**[192]** In some embodiments, a target biomarker signature of lung cancer comprises at least one extracellular vesicle-associated surface biomarker (*e.g.,* surface polypeptide and/or carbohydrate-dependent marker present in nanoparticles associated with lung cancer) and at least

one target biomarker selected from the group consisting of surface biomarker(s), intravesicular biomarker(s), and intravesicular RNA biomarker(s), such that the combination of such extracellular vesicle-associated surface biomarker(s) and such target biomarker(s) is specific for lung cancer. In some embodiments, a target biomarker signature of lung cancer is or comprises: CD166 antigen (ALCAM) polypeptide, N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase 3 polypeptide encoded by the UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3 (B3GNT3) gene, CUB domain containing protein 1 polypeptide encoded by the CUB domain-containing protein 1 (CDCP1) gene, Cadherin-1 (CDH1) polypeptide, cadherin 3 polypeptide encoded by the cadherin 3 (CDH3) gene, Complement decay-accelerating factor (CD55) polypeptide, Programmed cell death 1 ligand 1 (CD274; also known as PD-L1) polypeptide, carcinoembryonic antigen cell adhesion molecule 5 polypeptide encoded by the carcinoembryonic antigen cell adhesion molecule 5 (CEACAM5) gene, carcinoembryonic antigen cell adhesion molecule 6 polypeptide encoded by the carcinoembryonic antigen cell adhesion molecule 6 (CEACAM6) gene, claudin 3 polypeptide encoded by the claudin 3 (CLDN3) gene, claudin 4 polypeptide encoded by the (CLDN4) gene, desmoglein 2 polypeptide encoded by the desmoglein 2 (DSG2) gene, Epidermal growth factor receptor (EGFR) polypeptide, epithelial cell adhesion molecule polypeptide encoded by the epithelial cell adhesion molecule (EPCAM) gene, folate receptor alpha polypeptide encoded by the folate receptor alpha (FOLR1) gene, gap junction beta-1 protein polypeptide encoded by the gap junction protein beta 1 (GJB1) gene, gap junction beta-2 protein polypeptide encoded by the gap junction protein beta 2 (GJB2) gene, Hepatocyte growth factor receptor (MET) polypeptide, Insulin-like growth factor 1 receptor (IG1FR) polypeptide, laminin subunit beta-3 polypeptide encoded by the laminin subunit beta 3 (LAMB3) gene, Mesothelin (MSLN) polypeptide, Mucin-1 (MUC1) polypeptide, GPI transamidase component PIG-T polypeptide encoded by the phosphatidylinositol glycan anchor biosynthesis class T (PIGT) gene, podocalyxin-like protein 2 polypeptide encoded by the podocalyxin like 2 (PODXL2) gene, proto-oncogene tyrosine-protein kinase ROS polypeptide encoded by the ROS proto-oncogene 1, receptor tyrosine kinase (ROS1) gene, syndecan 1 polypeptide encoded by the syndecan 1 (SDC1) gene, sodium-dependent phosphate transport protein 2B polypeptide encoded by the solute carrier family 34 (sodium phosphate) member 2 (SLC34A2) gene, acid sphingomyelinase-like phosphodiesterase 3b polypeptide encoded by the sphingomyelin phosphodiesterase acid like 3B (SMPDL3B) gene,

suppressor of tumorigenicity 14 protein polypeptide encoded by the ST14 transmembrane serine protease matriptase (ST14) gene, Tumor-associated calcium signal transducer 2 (TACSTD2) polypeptide, transmembrane protease serine 2 polypeptide encoded by the transmembrane serine protease 2 (TMPRSS2) gene, Tumor necrosis factor receptor superfamily member 10B (TNFRSF10B) polypeptide, tetraspanin-8 polypeptide encoded by the tetraspanin 8 (TSPAN8) gene, sTn polypeptide glycosylation, Tn polypeptide glycosylation, T polypeptide glycosylation, or combinations thereof.

**[193]** In some embodiments, a target biomarker signature of lung cancer is or comprises a surface biomarker selected from the group consisting of: protein tyrosine kinase 7 (PTK7), tetraspanin-8 (TSPAN8), cadherin EGF LAG seven-pass G-type receptor 2 (CELSR2), glypican 1 (GPC1), suppressor of tumorigenicity 14 protein (ST14), protein tyrosine phosphatase receptor type Z1 (PTPRZ1), G-protein coupled receptor 87 (GPR87), gap junction beta-5 (GJB5), gap junction beta-2 (GJB2), Ras homolog family member V (RHOV), Ly6/PLAUR domain-containing protein 3 (LYPD3), claudin 7 (CLDN7), desmoplakin (DSP), serine incorporator 2 (SERINC2), abhydrolase domain containing 17C (ABHD17C), p53 apoptosis effector (PERP), myelin protein zero like 2 (MPZL2), integrin beta 4 subunit (ITGB4), mesoderm specific transcript (MEST), glycoprotein Nmb (GPNMB), solute carrier family 35 member A2 (SLC35A2), alpha 1,4-galactosyltransferase (Gb3; CD77), integrin subunit alpha 6 (ITGA6), multidrug resistance-associated protein 5 (ABCC5), ATPase Na+/K+ transporting subunit beta 3 (ATP1B3), jagged1 (JAG1), transmembrane serine protease 11D (TMPRSS11D), Adhesion G-protein coupled receptor F1 (ADGRF1) polypeptide, Phospholipid-transporting ATPase ABCA3 (ABCA3) polypeptide, Multidrug resistance-associated protein 1 (ABCC1) polypeptide, ATP-binding cassette sub-family C member 3 (ABCC3) polypeptide, Golgi resident protein GCP60 (ACBD3) polypeptide, Long-chain-fatty-acid--CoA ligase 5 (ACSL5) polypeptide, Advanced glycosylation end product-specific receptor (AGER) polypeptide, CD166 antigen (ALCAM) polypeptide, AP-1 complex subunit mu-2 (AP1M2) polypeptide, Gamma-secretase subunit APH-1A (APH1A) polypeptide, MICOS complex subunit MIC26 (APOO) polypeptide, Phospholipid-transporting ATPase IH (ATP11A) polypeptide, Phospholipid-transporting ATPase IF (ATP11B) polypeptide, Sodium/potassium-transporting ATPase subunit beta-1 (ATP1B1) polypeptide, Renin receptor (ATP6AP2) polypeptide, Lactosylceramide 1,3-N-acetyl-beta-D-glucosaminyltransferase (also known as Beta-1,3-N-acetylglucosaminyltransferase 5)

(B3GNT5) polypeptide, Beta-1,4-galactosyltransferase 4 (B4GALT4) polypeptide, B-cell receptor-associated protein 31 (BCAP31) polypeptide, B box and SPRY domain-containing protein (BSPRY) polypeptide, CD109 antigen (CD109) polypeptide, Complement decay-accelerating factor (CD55) polypeptide, CD9 antigen (CD9) polypeptide, Cell division control protein 42 homolog (CDC42) polypeptide, Cadherin-1 (CDH1) polypeptide, Cadherin-3 (CDH3) polypeptide, Threonylcarbamoyladenosine tRNA methylthiotransferase (CDKAL1) polypeptide, Carcinoembryonic antigen-related cell adhesion molecule 5 (CEACAM5) polypeptide, Carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) polypeptide, Cadherin EGF LAG seven-pass G-type receptor 1 (CELSR1) polypeptide, Protein CIP2A (CIP2A) polypeptide, CDGSH iron-sulfur domain-containing protein 2 (CISD2) polypeptide, Cytoskeleton-associated protein 4 (CKAP4) polypeptide, Calcium-activated chloride channel regulator 2 (CLCA2) polypeptide, Claudin-1 (CLDN1) polypeptide, Chloride intracellular channel protein 6 (CLIC6) polypeptide, Cleft lip and palate transmembrane protein 1-like protein (CLPTM1L) polypeptide, Calsyntenin-1 (CLSTN1) polypeptide, Contactin-1 (CNTN1) polypeptide, Carboxypeptidase D (CPD) polypeptide, Cytochrome P450 2S1 (CYP2S1) polypeptide, Cytochrome P450 4F11 (CYP4F11) polypeptide, Cytochrome P450 4F3 (CYP4F3) polypeptide, Probable C-mannosyltransferase DPY19L1 (DPY19L1) polypeptide, Desmocollin-2 (DSC2) polypeptide, Desmocollin-3 (DSC3) polypeptide, Desmoglein-2 (DSG2) polypeptide, Desmoglein-3 (DSG3) polypeptide, Epidermal growth factor receptor (EGFR) polypeptide, Epithelial cell adhesion molecule (EPCAM) polypeptide, Ephrin type-B receptor 3 (EPHB3) polypeptide, FAM241B polypeptide, Protocadherin Fat 2 (FAT2) polypeptide, F-box/SPRY domain-containing protein 1 (FBXO45) polypeptide, Fermitin family homolog 1 (FERMT1) polypeptide, Folate receptor alpha (FOLR1) polypeptide, FXYD domain-containing ion transport regulator 3 (FXYD3) polypeptide, Frizzled-6 (FZD6) polypeptide, Polypeptide N-acetylgalactosaminyltransferase 1 (GALNT1) polypeptide, Polypeptide N-acetylgalactosaminyltransferase 3 (GALNT3) polypeptide, Polypeptide N-acetylgalactosaminyltransferase 5 (GALNT5) polypeptide, Polypeptide N-acetylgalactosaminyltransferase 6 (GALNT6) polypeptide, Polypeptide N-acetylgalactosaminyltransferase 14 (GALNT14) polypeptide, Vitamin K-dependent gamma-carboxylase (GGCX) polypeptide, Golgi membrane protein 1 (GOLM1) polypeptide, Golgi phosphoprotein 3-like (GOLPH3L) polypeptide, Grainyhead-like protein 2 homolog (GRHL2)

polypeptide, Very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase 3 (HACD3) polypeptide, Hyaluronan synthase 3 (HAS3) polypeptide, Immediate early response 3-interacting protein 1 (IER3IP1) polypeptide, Immunoglobulin superfamily member 3 (IGSF3) polypeptide, Interleukin-1 receptor accessory protein (IL1RAP) polypeptide, Integrin alpha-2 (ITGA2) polypeptide, Integrin beta-6 (ITGB6) polypeptide, Killer cell lectin-like receptor subfamily G member 2 (KLRG2) polypeptide, Importin subunit alpha-1 (KPNA2) polypeptide, Keratinocyte-associated protein 3 (KRTCAP3) polypeptide, Ladinin-1 (LAD1) polypeptide, Laminin subunit beta-3 (LAMB3) polypeptide, Laminin subunit gamma-2 (LAMC2) polypeptide, Lysosome-associated membrane glycoprotein 3 (LAMP3) polypeptide, Ragulator complex protein LAMTOR2 (LAMTOR2) polypeptide, Lysocardiolipin acyltransferase 1 (LCLAT1) polypeptide, Lysosomal-associated transmembrane protein 4B (LAPTM4B) polypeptide, LARGE xylosyl- and glucuronyltransferase 2 (LARGE2) polypeptide, Lysophosphatidylcholine acyltransferase 1 (LPCAT1) polypeptide, Lipolysis-stimulated lipoprotein receptor (LSR) polypeptide, MAL2 polypeptide, Magnesium transporter protein 1 (MAGT1) polypeptide, MARCKS-related protein (MARCKSL1) polypeptide, Hepatocyte growth factor receptor (MET) polypeptide, Alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase (MGAT1) polypeptide, Mesothelin (MSLN) polypeptide, Mucin-1 (MUC1) polypeptide, Mucin-4 (MUC4) polypeptide, Nicastrin (NCSTN) polypeptide, Nectin-1 (NECTIN1) polypeptide, Nectin-4 (NECTIN4) polypeptide, GTPase NRas (NRAS) polypeptide, Neuronal cell adhesion molecule (Nr-CAM) polypeptide, 5'-nucleotidase (NT5E) polypeptide, Nuclear pore membrane glycoprotein 210 (NUP210) polypeptide, Presenilins-associated rhomboid-like protein, mitochondrial (PARL) polypeptide, Peroxisomal membrane protein PEX13 (PEX13) polypeptide, GPI ethanolamine phosphate transferase 1 (PIGN) polypeptide, GPI transamidase component PIG-T (PIGT) polypeptide, Cytosolic phospholipase A2 (PLA2G4A) polypeptide, 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase eta-1 (PLCH1) polypeptide, Plectin (PLEC) polypeptide, 26S proteasome non-ATPase regulatory subunit 2 (PSMD2) polypeptide, Phosphatidylserine synthase 1 (PTDSS1) polypeptide, Prostaglandin F2 receptor negative regulator (PTGFRN) polypeptide, Receptor-type tyrosine-protein phosphatase F (PTPRF) polypeptide, Sulfhydryl oxidase 1 (QSOX1) polypeptide, Ras-related protein Rab-25 (RAB25) polypeptide, Ras-related protein Rab-38 (RAB38) polypeptide, Ras-related protein Rab-6B; (RAB6B) polypeptide, Ras-related protein Rap-2b (RAP2B) polypeptide, Protein RCC2 (RCC2)

polypeptide, GTP-binding protein Rit1 (RIT1) polypeptide, Secretory carrier-associated membrane protein 3 (SCAMP3) polypeptide, Syndecan-1 (SDC1) polypeptide, Protein sel-1 homolog 3 (SEL1L3) polypeptide, Protein Shroom2 (SHROOM2) polypeptide, Solute carrier family 2, facilitated glucose transporter member 1 (SLC2A1) polypeptide, Cystine/glutamate transporter (also known as solute carrier family 7 member 11) (SLC7A11) polypeptide, Sodium-dependent phosphate transport protein 2B (SLC34A2) polypeptide, Adenosine 3'-phospho 5'-phosphosulfate transporter 1 (SLC35B2) polypeptide, Zinc transporter ZIP11 (SLC39A11) polypeptide, Small integral membrane protein 22 (SMIM22) polypeptide, Testisin (also known as eosinophil serine protease 1 (ESP-1) or serine protease 21) polypeptide, Acid sphingomyelinase-like phosphodiesterase 3b (SMPDL3B) polypeptide, Sterol O-acyltransferase 1 (SOAT1) polypeptide, Spastin (SPAST) polypeptide, Translocon-associated protein subunit alpha (SSR1) polypeptide, Translocon-associated protein subunit delta (SSR4) polypeptide, Surfeit locus protein 4; (SURF4) polypeptide, Synaptogyrin-2 (SYNGR2) polypeptide, Tumor-associated calcium signal transducer 2 (TACSTD2) polypeptide, Calcineurin B homologous protein 3 (TESC) polypeptide, Transferrin receptor protein 1 (TFRC) polypeptide, Transmembrane channel-like protein 5 (TMC5) polypeptide, Calcium load-activated calcium channel (TMCO1) polypeptide, Transmembrane emp24 domain-containing protein 2 (TMED2) polypeptide, Transmembrane emp24 domain-containing protein 3 (TMED3) polypeptide, Transmembrane protein 132A (TMEM132A) polypeptide, Transmembrane protein 33 (TMEM33) polypeptide, Transmembrane protease serine 4 (TMPRSS4) polypeptide, Protein O-mannosyl-transferase TMTC3 (TMTC3) polypeptide, Mitochondrial import receptor subunit TOM22 homolog (TOMM22) polypeptide, Torsin-1A-interacting protein 2, isoform IFRG15 (TOR1AIP2) polypeptide, Translocating chain-associated membrane protein 1 (TRAM1) polypeptide, Transient receptor potential cation channel subfamily V member 4 (TRPV4) polypeptide, Tetratricopeptide repeat protein 33 (TTC33) polypeptide, UDP-glucuronosyltransferase 1-6 (UGT1A6) polypeptide, Uroplakin-1b (UPK1B) polypeptide, Vesicle-associated membrane protein 8 (VAMP8) polypeptide, Vacuolar ATPase assembly integral membrane protein VMA21 (VMA21) polypeptide, Serine/threonine-protein kinase VRK2 (VRK2) polypeptide, von Willebrand factor A domain-containing protein 1 (VWA1) polypeptide, Xenotropic and polytropic retrovirus receptor 1 (XPR1) polypeptide, Xyloside xylosyltransferase 1 (XXYLT1) polypeptide, Disintegrin and metalloproteinase domain-

containing protein 28 (ADAM28) polypeptide, Tyrosine-protein kinase receptor UFO (AXL) polypeptide, Basigin (BSG) polypeptide, Programmed cell death 1 ligand 1 (CD274; also known as PD-L1) polypeptide, Leukocyte surface antigen CD47 (CD47) polypeptide, Clusterin (CLU) polypeptide, Dickkopf-related protein 1 (DKK1) polypeptide, Receptor tyrosine-protein kinase erbB-3 (ERBB3) polypeptide, Vascular endothelial growth factor receptor 3 (FLT4) polypeptide, (N-glycolylneuraminic acid (NeuGc, NGNA)-gangliosides GM3) (GM3) polypeptide, Hepatocyte growth factor (HGF) polypeptide, Insulin-like growth factor 1 receptor (IGF1R) polypeptide, Interleukin-6 (IL6) polypeptide, Vascular endothelial growth factor receptor 2 (KDR) polypeptide, Lymphocyte activation gene 3 protein (LAG3) polypeptide, Ubiquitin carboxyl-terminal hydrolase isozyme L1 (UCH-L1) polypeptide, Lewis X antigen, Lewis Y/B antigen, Lymphocyte antigen 6E (LY6E) polypeptide, Neurogenic locus notch homolog protein 2 (NOTCH2) polypeptide, Neurogenic locus notch homolog protein 3 (NOTCH3) polypeptide, Phosphatidylserine presenting polypeptide, T-cell immunoreceptor with Ig and ITIM domains (TIGIT) polypeptide, Tumor necrosis factor receptor superfamily member 10A (TNFRSF10A) polypeptide, Tumor necrosis factor receptor superfamily member 10B (TNFRSF10B) polypeptide, Tumor necrosis factor ligand superfamily member 18 (TNFSF18) polypeptide, Trophoblast glycoprotein (TPBG) polypeptide, Vascular endothelial growth factor A (VEGFA) polypeptide, Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, Phosphatidylserine, αGalNAc-Ser/Thr (Tn) antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, Lewis X, Lewis Y/B antigen, LY6E, Lewis Y/CD174 antigen, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)) antigen, N-glycolyl GM3 ganglioside (NeuGcGM3), and combinations thereof.

**[194]** In some embodiments, extracellular vesicle-associated surface biomarker(s) included in a target biomarker signature of lung cancer is or comprises: SLC34A2 polypeptide, CEACAM5 polypeptide, CEACAM6 polypeptide, EpCAM polypeptide, and/or combinations thereof. SLC34A2 polypeptide is a multi-pass membrane transporter than has been studied as a therapeutic target for non-small cell lung cancer (Lin *et al.*, 2015; which is incorporated herein by reference for the purpose described herein). CEACAM5 polypeptide, a member of the carcinoembryonic antigen (CEA) family of cell adhesion molecules (CAM), is a cell surface glycoprotein that has been implicated in gastrointestinal cancers and is thought to be involved in cellular differentiation, apoptosis, and polarity. CEACAM6 polypeptide is a member of the same

protein family as CEACAM5, and has been implicated in Crohn's disease and pancreatic adenocarcinoma, and is thought to be involved in the innate immune system and cell surface interactions. EpCAM polypeptide is implicated in gastrointestinal carcinomas and is thought to function as a homotypic calcium-independent cell adhesion molecule. In some embodiments, SLC34A2 polypeptide, CEACAM5 polypeptide, CEACAM6 polypeptide, and/or EpCAM polypeptide are detected as intact EV associated trans-membrane proteins. In some embodiments of the present disclosure, SLC34A2 polypeptide, CEACAM5 polypeptide, CEACAM6 polypeptide, and/or EPCAM polypeptide are detected as EV associated trans-membrane polypeptides.

**[195]** In some embodiments, a target biomarker included in a target biomarker signature of lung cancer is or comprises a surface biomarker selected from the group consisting of: Adhesion G-protein coupled receptor F1 polypeptide encoded by ADGRF1 gene, CD166 antigen (ALCAM) polypeptide, canalicular multispecific organic anion transporter 2 polypeptide encoded by the ATP binding cassette subfamily C member 3 (ABCC3) gene, arylsulfatase L polypeptide encoded by the arylsulfatase L (ARSL) gene, N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase 3 polypeptide encoded by the UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3 (B3GNT3) gene, Lactosylceramide 1,3-N-acetyl-beta-D-glucosaminyltransferase polypeptide encoded by UDP-GlcNAc:beta-Gal beta-1,3-N-acetylglucosaminyltransferase 5 (B3GNT5) gene, CUB domain containing protein 1 polypeptide encoded by the CUB domain-containing protein 1 (CDCP1) gene, cadherin 1 polypeptide encoded by the cadherin 1 (CDH1) gene, cadherin 3 polypeptide encoded by the cadherin 3 (CDH3) gene, Complement decay-accelerating factor (CD55) polypeptide, Programmed cell death 1 ligand 1 (CD274; also known as PD-L1) polypeptide, carcinoembryonic antigen cell adhesion molecule 5 polypeptide encoded by the carcinoembryonic antigen cell adhesion molecule 5 (CEACAM5) gene, carcinoembryonic antigen cell adhesion molecule 6 polypeptide encoded by the carcinoembryonic antigen cell adhesion molecule 6 (CEACAM6) gene, cadherin EGF LAG seven-pass G-type receptor 1 polypeptide encoded by the cadherin EGF LAG seven-pass G-type receptor 1 (CELSR1) gene, claudin 18 polypeptide encoded by the claudin 18 (CLDN18) gene, claudin 3 polypeptide encoded by the claudin 3 (CLDN3) gene, claudin 4 polypeptide encoded by the (CLDN4) gene, claudin 7 polypeptide encoded by the claudin 7 (CLDN7) gene, chloride intracellular channel protein 6 polypeptide encoded by the chloride

intracellular channel 6 (CLIC6) gene, deleted in malignant brain tumors 1 protein polypeptide encoded by the deleted in malignant brain tumors 1 (DMBT1) gene, desmoglein 2 polypeptide encoded by the desmoglein 2 (DSG2) gene, Epidermal growth factor receptor (EGFR) polypeptide, epithelial cell adhesion molecule polypeptide encoded by the epithelial cell adhesion molecule (EPCAM) gene, epoxide hydrolase 3 polypeptide encoded by the epoxide hydrolase 3 (EPHX3) gene, protein eva-1 homolog A polypeptide encoded by the eva-1 homolog A, regulator of programmed cell death (EVA1A) gene, Protein FAM241B encoded by FAM241B gene, folate receptor alpha polypeptide encoded by the folate receptor alpha (FOLR1) gene, FXYD domain-containing ion transport regulator 3 polypeptide encoded by FXYD3 gene, Polypeptide N-acetylgalactosaminyltransferase 14 encoded by GALNT14 gene, gap junction beta-1 protein polypeptide encoded by the gap junction protein beta 1 (GJB1) gene, gap junction beta-2 protein polypeptide encoded by the gap junction protein beta 2 (GJB2) gene, Hepatocyte growth factor receptor (MET) polypeptide, Insulin-like growth factor 1 receptor (IG1FR) polypeptide, glypican 4 polypeptide encoded by the glypican 4 (GPC4) gene, Hyaluronan synthase 3 polypeptide encoded by HAS3 gene, heparan sulfate 6-O-sulfotransferase 2 polypeptide encoded by the heparin sulfate 6-O-sulfotransferase 2 (HS6ST2) gene, ER lumen protein-retaining receptor 3 polypeptide encoded by the ER lumen protein-retaining receptor 3 (KDELR3) gene, keratinocyte-associated protein 3 polypeptide encoded by the keratinocyte associated protein 3 (KRTCAP3) gene, laminin subunit beta-3 polypeptide encoded by the laminin subunit beta 3 (LAMB3) gene, Lysosomal-associated transmembrane protein 4B polypeptide encoded by LAPTM4B gene, LARGE xylosyl- and glucuronyltransferase 2 polypeptide encoded by LARGE2 gene, beta-1,3-N-acetylglucosaminyltransferase lunatic fringe polypeptide encoded by the LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase (LFNG) gene, lipolysis-stimulated lipoprotein receptor polypeptide encoded by the lipolysis stimulated lipoprotein receptor (LSR) gene, Protein MAL2 encoded by MAL2 gene, glycoprotein endo-alpha-1,2-mannosidase-like protein polypeptide encoded by the mannosidase endo-alpha like (MANEAL) gene, mesothelin polypeptide encoded by the mesothelin (MSLN) gene, mucin 1 polypeptide encoded by the mucin 1, cell surface associated (MUC1) gene, mucin 21 polypeptide encoded by the mucin 21, cell surface associated (MUC21) gene, Neuronal cell adhesion molecule (Nr-CAM) polypeptide encoded by NRCAM gene, GPI transamidase component PIG-T polypeptide encoded by the phosphatidylinositol glycan anchor biosynthesis

class T (PIGT) gene, podocalyxin-like protein 2 polypeptide encoded by the podocalyxin like 2 (PODXL2) gene, Testisin polypeptide encoded by PRSS21 gene, transmembrane gamma-carboxyglutamic acid protein 4 polypeptide encoded by the proline rich and Gla domain 4 (PRRG4) gene, proto-oncogene tyrosine-protein kinase ROS polypeptide encoded by the ROS proto-oncogene 1, receptor tyrosine kinase (ROS1) gene, syndecan 1 polypeptide encoded by the syndecan 1 (SDC1) gene, serine incorporator 2 polypeptide encoded by the serine incorporator 2 (SERINC2) gene, seizure 6-like protein 2 polypeptide encoded by the seizure related 6 homolog like 2 (SEZ6L2) gene, sodium-dependent phosphate transport protein 2B polypeptide encoded by the solute carrier family 34 (sodium phosphate) member 2 (SLC34A2) gene, choline transporter-like protein 4 polypeptide encoded by the solute carrier family 44 member 4 (SLC44A4) gene, sodium- and chloride-dependent neutral and basic amino acid transporter B(0+) polypeptide encoded by the solute carrier family 6 member 14 (SLC6A14) gene, Y+L amino acid transporter 1 polypeptide encoded by the solute carrier family 7 member 7 (SLC7A7) gene, Cystine/glutamate transporter polypeptide encoded by solute carrier family 7 member 11 (SLC7A11) gene, small integral membrane protein 22 polypeptide encoded by the small integral membrane protein 22 (SMIM22) gene, acid sphingomyelinase-like phosphodiesterase 3b polypeptide encoded by the sphingomyelin phosphodiesterase acid like 3B (SMPDL3B) gene, suppressor of tumorigenicity 14 protein polypeptide encoded by the ST14 transmembrane serine protease matriptase (ST14) gene, Tumor-associated calcium signal transducer 2 (TACSTD2) polypeptide, transmembrane channel-like protein 4 polypeptide encoded by the transmembrane channel like 4 (TMC4) gene, transmembrane channel-like protein 5 polypeptide encoded by the transmembrane channel like 5 (TMC5) gene, transmembrane protein 45B polypeptide encoded by the transmembrane protein 45B (TMEM45B) gene, transmembrane protease serine 2 polypeptide encoded by the transmembrane serine protease 2 (TMPRSS2) gene, transmembrane protease serine 4 polypeptide encoded by the transmembrane serine protease 4 (TMPRSS4) gene, Tumor necrosis factor receptor superfamily member 10B (TNFRSF10B) polypeptide, tetraspanin-1 polypeptide encoded by the tetraspanin 1 (TSPAN1) gene, tetraspanin-8 polypeptide encoded by the tetraspanin 8 (TSPAN8) gene, Ubiquitin carboxyl-terminal hydrolase isozyme L1 (UCH-L1) polypeptide encoded by UCHL1 gene, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, and combinations thereof.

**[196]** In some embodiments, a target biomarker included in a target biomarker signature of lung cancer is or comprises a surface biomarker selected from the group consisting of: SLC34A2 polypeptide, CEACAM5 polypeptide, CEACAM6 polypeptide, EpCAM polypeptide, and combinations thereof.

**[197]** In some embodiments, a target biomarker included in a target biomarker signature of lung cancer is or comprises a surface biomarker selected from the group consisting of: ALCAM polypeptide, CD55 polypeptide, CDH1 polypeptide, CDH3 polypeptide, CD274 (PD-L1) polypeptide, CEACAM5 polypeptide, CEACAM6 polypeptide, DSG2 polypeptide, EGFR polypeptide, EPCAM polypeptide, FOLR1 polypeptide, IG1FR polypeptide, MET polypeptide, MSLN polypeptide, MUC1 polypeptide, SLC34A2 polypeptide, sTn antigen, Tn antigen, T antigen, TACSTD2 polypeptide, TNFRSF10B polypeptide, and combinations thereof.

**[198]** In some embodiments, a target biomarker signature comprises one or more extracellular vesicle-associated surface biomarkers, selected from a list consisting of a ADGRF1 polypeptide, a ABCA3 polypeptide, a ABCC1 polypeptide, a ABCC3 polypeptide, a ACBD3 polypeptide, a ACSL5 polypeptide, a AGER polypeptide, a ALCAM polypeptide, a AP1M2 polypeptide, a APH1A polypeptide, a APOO polypeptide, a ATP11A polypeptide, a ATP11B polypeptide, a ATP1B1 polypeptide, a ATP6AP2 polypeptide, a B3GNT5 polypeptide, a B4GALT4 polypeptide, a BCAP31 polypeptide, a BSPRY polypeptide, a CD109 polypeptide, a CD55 polypeptide, a CD9 polypeptide, a CDC42 polypeptide, a CDH1 polypeptide, a CDH3 polypeptide, a CDKAL1 polypeptide, a CEACAM5 polypeptide, a CEACAM6 polypeptide, a CELSR1 polypeptide, a CIP2A polypeptide, a CISD2 polypeptide, a CKAP4 polypeptide, a CLCA2 polypeptide, a CLDN1 polypeptide, a CLIC6 polypeptide, a CLPTM1L polypeptide, a CLSTN1 polypeptide, a CNTN1 polypeptide, a CPD polypeptide, a CYP2S1 polypeptide, a CYP4F11 polypeptide, a CYP4F3 polypeptide, a DPY19L1 polypeptide, a DSC2 polypeptide, a DSC3 polypeptide, a DSG2 polypeptide, a DSG3 polypeptide, a EGFR polypeptide, a EPCAM polypeptide, a EPHB3 polypeptide, a FAM241B polypeptide, a FAT2 polypeptide, a FBXO45 polypeptide, a FERMT1 polypeptide, a FOLR1 polypeptide, a FXYD3 polypeptide, a FZD6 polypeptide, a GALNT1 polypeptide, a GALNT3 polypeptide, a GALNT5 polypeptide, a GALNT6 polypeptide, a GALNT14 polypeptide, a GGCX polypeptide, a GOLM1 polypeptide, a GOLPH3L polypeptide, a GRHL2 polypeptide, a HACD3 polypeptide, a HAS3 polypeptide, a IER3IP1 polypeptide, a IGSF3 polypeptide, a IL1RAP polypeptide, a ITGA2 polypeptide, a

ITGB6 polypeptide, a KLRG2 polypeptide, a KPNA2 polypeptide, a KRTCAP3 polypeptide, a LAD1 polypeptide, a LAMB3 polypeptide, a LAMC2 polypeptide, a LAMP3 polypeptide, a LAPTM4B polypeptide, a LAMTOR2 polypeptide, a LARGE2 polypeptide, a LCLAT1 polypeptide, a LPCAT1 polypeptide, a LSR polypeptide, a MAGT1 polypeptide, a MAL2 polypeptide, a MARCKSL1 polypeptide, a MET polypeptide, a MGAT1 polypeptide, a MSLN polypeptide, a MUC1 polypeptide, a MUC4 polypeptide, a NCSTN polypeptide, a NECTIN1 polypeptide, a NECTIN4 polypeptide, a NRAS polypeptide, a NRCAM polypeptide, a NT5E polypeptide, a NUP210 polypeptide, a PARL polypeptide, a PEX13 polypeptide, a PIGN polypeptide, a PIGT polypeptide, a PLA2G4A polypeptide, a PLCH1 polypeptide, a PLEC polypeptide, a PRSS21 polypeptide, a PSMD2 polypeptide, a PTDSS1 polypeptide, a PTGFRN polypeptide, a PTPRF polypeptide, a QSOX1 polypeptide, a RAB25 polypeptide, a RAB38 polypeptide, a RAB6B polypeptide, a RAP2B polypeptide, a RCC2 polypeptide, a RIT1 polypeptide, a SCAMP3 polypeptide, a SDC1 polypeptide, a SEL1L3 polypeptide, a SHROOM2 polypeptide, a SLC2A1 polypeptide, a SLC34A2 polypeptide, a SLC35B2 polypeptide, a SLC39A11 polypeptide, a SLC7A11 polypeptide, a SMIM22 polypeptide, a SMPDL3B polypeptide, a SOAT1 polypeptide, a SPAST polypeptide, a SSR1 polypeptide, a SSR4 polypeptide, a SURF4 polypeptide, a SYNGR2 polypeptide, a TACSTD2 polypeptide, a TESC polypeptide, a TFRC polypeptide, a TMC5 polypeptide, a TMCO1 polypeptide, a TMED2 polypeptide, a TMED3 polypeptide, a TMEM132A polypeptide, a TMEM33 polypeptide, a TMPRSS4 polypeptide, a TMTC3 polypeptide, a TOMM22 polypeptide, a TOR1AIP2 polypeptide, a TRAM1 polypeptide, a TRPV4 polypeptide, a TTC33 polypeptide, a UGT1A6 polypeptide, a UPK1B polypeptide, a VAMP8 polypeptide, a VMA21 polypeptide, a VRK2 polypeptide, a VWA1 polypeptide, a XPR1 polypeptide, a XXYLT1 polypeptide, a ADAM28 polypeptide, a AXL polypeptide, a BSG polypeptide, a CD274 polypeptide, a CD47 polypeptide, a CLU polypeptide, a DKK1 polypeptide, a ERBB3 polypeptide, a FLT4 polypeptide, a GM3 polypeptide, a HGF polypeptide, a IGF1R polypeptide, a IL6 polypeptide, a KDR polypeptide, a LAG3 polypeptide, a UCHL1 polypeptide, a Lewis X antigen, a Lewis Y/B antigen, a LY6E polypeptide, a NOTCH2 polypeptide, a NOTCH3 polypeptide, a Phosphatidylserine presenting polypeptide, a TIGIT polypeptide, a TNFRSF10A polypeptide, a TNFRSF10B polypeptide, a TNFSF18 polypeptide, a TPBG polypeptide, a VEGFA polypeptide, a Tn antigen, a Lewis

Y/CD174 antigen, a Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)) antigen, a NeuGcGM3 ganglioside, and combinations thereof.

**[199]** In some embodiments, a target biomarker signature comprises one or more extracellular vesicle-associated surface biomarkers, selected from a list consisting of a HS6ST2 polypeptide, a CYP2S1 polypeptide, a HAS3 polypeptide, a LAMC2 polypeptide, a ADAM23 polypeptide, a ABCA13 polypeptide, a TMPRSS4 polypeptide, a UGT1A6 polypeptide, a ILDR1 polypeptide, a CYP4F11 polypeptide, a PIGT polypeptide, a LAMB3 polypeptide, a PRSS21 polypeptide, a DSG3 polypeptide, a SDK2 polypeptide, and combinations thereof.

**[200]** In some embodiments, a target biomarker signature comprises one or more extracellular vesicle-associated surface biomarkers, selected from a list consisting of a HS6ST2 polypeptide, a CYP2S1 polypeptide, a HAS3 polypeptide, a LAMC2 polypeptide, a ADAM23 polypeptide, a ABCA13 polypeptide, a TMPRSS4 polypeptide, a UGT1A6 polypeptide, a ILDR1 polypeptide, a CYP4F11 polypeptide, a PIGT polypeptide, a LAMB3 polypeptide, a PRSS21 polypeptide, a DSG3 polypeptide, a SDK2 polypeptide, a FERMT1 polypeptide, a EPCAM polypeptide, a SDC1 polypeptide, a PANX2 polypeptide, a ULBP2 polypeptide, a ECE2 polypeptide, a KRTCAP3 polypeptide, a CLCA2 polypeptide, a KPNA2 polypeptide, a TMEM132A polypeptide, a ABCC1 polypeptide, a UPK1B polypeptide, a DSG2 polypeptide, a NECTIN1 polypeptide, a SHISA2 polypeptide, and combinations thereof.

**[201]** In some embodiments, a target biomarker included in a target biomarker signature of lung cancer is or comprises an intravesicular biomarker selected from the group consisting of: amiloride-sensitive amine oxidase [copper-containing] polypeptide encoded by the amine oxidase copper containing 1 (AOC1) gene, uncharacterized protein C12orf45 polypeptide encoded by the chromosome 12 open reading frame 45 (C12orf45) gene, cellular retinoic acid binding protein 2 polypeptide encoded by the cellular retinoic acid binding protein 2 (CRABP2) gene, cystatin SN polypeptide encoded by the cystatin SN (CST1) gene, ETS translocation variant 4 polypeptide encoded by the ETS variant transcription factor 4 (ETV4) gene, protein FAM83A polypeptide encoded by the family with sequence similarity 83 member A (FAM83A) gene, hepatocyte nuclear factor 3-beta polypeptide encoded by the forkhead box A2 (FOXA2) gene, high mobility group protein B3 polypeptide encoded by the high mobility group box 3 (HMGB3) gene, galectin-3-binding protein polypeptide encoded by the galectin 3 binding protein (LGALS3BP) gene, macrophage migration inhibitory factor polypeptide encoded by the

macrophage migration inhibitory factor (MIF) gene, napsin-A polypeptide encoded by the napsin A aspartic peptidase (NAPSA) gene, protein phosphatase 1 regulatory subunit 14D polypeptide encoded by the protein phosphatase 1 regulatory inhibitor subunit 14D (PPP1R14D) gene, protein S100-A14 polypeptide encoded by the S100 calcium binding protein A14 (S100A14) gene, serine/threonine-protein kinase SBK1 polypeptide encoded by the SH3 domain binding kinase 1 (SBK1) gene, secretoglobin family 3A member 2 polypeptide encoded by the secretoglobin family 3A member 2 (SCGB3A2) gene, surfactant-associated protein 2 polypeptide encoded by the surfactant associated 2 (SFTA2) gene, pulmonary surfactant-associated protein A1 polypeptide encoded by the surfactant protein A1 (SFTPA1) gene, pulmonary surfactant-associated protein A2 polypeptide encoded by the surfactant protein A2 (SFTPA2) gene, pulmonary surfactant-associated protein B polypeptide encoded by the surfactant protein B (SFTPB) gene, serine protease inhibitor Kazal-type 1 polypeptide encoded by the serine peptidase inhibitor Kazal type 1 (SPINK1) gene, protransforming growth factor alpha polypeptide encoded by the transforming growth factor alpha (TGFA) gene, zinc finger CCCH domain-containing protein 11A polypeptide encoded by the zinc finger CCCH-type containing 11A (ZC3H11A) gene, and combinations thereof.

**[202]** In some embodiments, a target biomarker in a target biomarker signature of lung cancer is or comprises an intravesicular biomarker selected from the group consisting of: a ABRACL polypeptide, a ACP5 polypeptide, a ADH7 polypeptide, a AGR2 polypeptide, a AIF1 polypeptide, a AKR1C1 polypeptide, a AKR1C2 polypeptide, a AKR1C3 polypeptide, a ALDH1A1 polypeptide, a ALDH3A1 polypeptide, a ALDH3B2 polypeptide, a ALG1L polypeptide, a AP1M2 polypeptide, a APOBEC3B polypeptide, a APOBEC3C polypeptide, a ARNTL2 polypeptide, a ASF1B polypeptide, a AURKB polypeptide, a BAIAP2L1 polypeptide, a BIRC5 polypeptide, a C15orf48 polypeptide, a C19orf33 polypeptide, a C1S polypeptide, a C8orf4 polypeptide, a CA9 polypeptide, a CALML3 polypeptide, a CAPNS2 polypeptide, a CBLC polypeptide, a CCL19 polypeptide, a CCNB2 polypeptide, a CDC20 polypeptide, a CDC45 polypeptide, a CDCA4 polypeptide, a CDCA5 polypeptide, a CDK1 polypeptide, a CDKN2A polypeptide, a CDKN2B polypeptide, a CENPW polypeptide, a CEP55 polypeptide, a CES1 polypeptide, a CHMP4C polypeptide, a CNN2 polypeptide, a CPA3 polypeptide, a CRABP2 polypeptide, a CSTA polypeptide, a CTSC polypeptide, a CTSE polypeptide, a CYP2S1 polypeptide, a DPYSL3 polypeptide, a EFS polypeptide, a EGLN3 polypeptide, a EHF

polypeptide, a ELF3 polypeptide, a ELF4 polypeptide, a ENAH polypeptide, a ESRP1 polypeptide, a EVPL polypeptide, a FAM129B polypeptide, a FAM60A polypeptide, a FAM83D polypeptide, a FAM83H polypeptide, a FBP1 polypeptide, a FERMT1 polypeptide, a FOXE1 polypeptide, a FOXM1 polypeptide, a GBP6 polypeptide, a GNA15 polypeptide, a GPX2 polypeptide, a GRHL2 polypeptide, a GSTA1 polypeptide, a HCK polypeptide, a HOXB7 polypeptide, a ID1 polypeptide, a IGF2BP2 polypeptide, a IMPA2 polypeptide, a IRF6 polypeptide, a IVL polypeptide, a JUP polypeptide, a KIAA1522 polypeptide, a KIF2C polypeptide, a KIFC1 polypeptide, a KLF4 polypeptide, a KLF5 polypeptide, a KRT13 polypeptide, a KRT14 polypeptide, a KRT15 polypeptide, a KRT16 polypeptide, a KRT17 polypeptide, a KRT18 polypeptide, a KRT19 polypeptide, a KRT5 polypeptide, a KRT6A polypeptide, a KRT6B polypeptide, a KRT6C polypeptide, a KRT7 polypeptide, a KRT8 polypeptide, a LGALS7B polypeptide, a LSP1 polypeptide, a MAGEA4 polypeptide, a MAGEA6 polypeptide, a MCM2 polypeptide, a MDFI polypeptide, a MYBL2 polypeptide, a MYH14 polypeptide, a MZB1 polypeptide, a NCF2 polypeptide, a NNMT polypeptide, a NRARP polypeptide, a NUP210 polypeptide, a NUSAP1 polypeptide, a OSGIN1 polypeptide, a PALLD polypeptide, a PITX1 polypeptide, a PKP1 polypeptide, a PKP3 polypeptide, a PLEK polypeptide, a PLEK2 polypeptide, a POSTN polypeptide, a PPP1R14C polypeptide, a PRAME polypeptide, a PTPN6 polypeptide, a RBP1 polypeptide, a RIN2 polypeptide, a RIPK4 polypeptide, a RPS4Y1 polypeptide, a RRM2 polypeptide, a S100A11 polypeptide, a S100A14 polypeptide, a S100A16 polypeptide, a S100A2 polypeptide, a S100P polypeptide, a SERPINB13 polypeptide, a SERPINB3 polypeptide, a SERPINB5 polypeptide, a SH3BP4 polypeptide, a SNAI2 polypeptide, a SOX2 polypeptide, a SPI1 polypeptide, a SPINT1 polypeptide, a SPRR1A polypeptide, a SPRR1B polypeptide, a SPRR2A polypeptide, a SPRR2D polypeptide, a SPRR2E polypeptide, a SPRR3 polypeptide, a SULF1 polypeptide, a SYK polypeptide, a SYTL1 polypeptide, a TBC1D2 polypeptide, a TEAD2 polypeptide, a TEAD3 polypeptide, a TFAP2C polypeptide, a THBS2 polypeptide, a TK1 polypeptide, a TOP2A polypeptide, a TP63 polypeptide, a TPD52 polypeptide, a TPX2 polypeptide, a TRIM29 polypeptide, a TRIP13 polypeptide, a UBE2C polypeptide, a YAP1 polypeptide, a ZC3H11A polypeptide, a ZNF217 polypeptide, a ZNF750 polypeptide, and combinations thereof.

**[203]** In some embodiments, a target biomarker included in a target biomarker signature of lung cancer is or comprises an intravesicular RNA (*e.g.*, mRNA) biomarker selected

from the group consisting of: ABCC3 RNA, AOC1 RNA, ARSL RNA, B3GNT3 RNA, C12orf45 RNA, CDCP1 RNA, CDH1 RNA, CDH3 RNA, CEACAM5 RNA, CEACAM6 RNA, CELSR1 RNA, CLDN18 RNA, CLDN3 RNA, CLDN4 RNA, CLDN7 RNA, CLIC6 RNA, CRABP2 RNA, CST1 RNA, DMBT1 RNA, DSG2 RNA, EPCAM RNA, EPHX3 RNA, ETV4 RNA, EVA1A RNA, FAM83A RNA, FOLR1 RNA, FOXA2 RNA, GJB1 RNA, GJB2 RNA, GPC4 RNA, HMGB3 RNA, HS6ST2 RNA, KDELR3 RNA, KRTCAP3 RNA, LAMB3 RNA, LFNG RNA, LGALS3BP RNA, LSR RNA, MANEAL RNA, MIF RNA, MSLN RNA, MUC1 RNA, MUC21 RNA, NAPSA RNA, PIGT RNA, PODXL2 RNA, PPP1R14D RNA, PRRG4 RNA, ROS1 RNA, S100A14 RNA, SBK1 RNA, SCGB3A2 RNA, SDC1 RNA, SERINC2 RNA, SEZ6L2 RNA, SFTA2 RNA, SFTPA1 RNA, SFTPA2 RNA, SFTPB RNA, SLC34A2 RNA, SLC44A4 RNA, SLC6A14 RNA, SLC7A7 RNA, SMIM22 RNA, SMPDL3B RNA, SPINK1 RNA, ST14 RNA, TGFA RNA, TMC4 RNA, TMC5 RNA, TMEM45B RNA, TMPRSS2 RNA, TMPRSS4 RNA, TSPAN1 RNA, TSPAN8 RNA, ZC3H11A RNA, and combinations thereof.

**[204]** In some embodiments, a target biomarker signature comprises one or more intravesicular RNA (*e.g.*, mRNA) biomarkers a list consisting of a ABCA3 RNA, a ABCC1 RNA, a ABRACL RNA, a ACP5 RNA, a ADAM23 RNA, a ADH7 RNA, a AGR2 RNA, a AIF1 RNA, a AKR1C1 RNA, a AKR1C2 RNA, a AKR1C3 RNA, a ALDH1A1 RNA, a ALDH3A1 RNA, a ALDH3B2 RNA, a ALG1L RNA, a ANTXR1 RNA, a AP1M2 RNA, a APOBEC3B RNA, a APOBEC3C RNA, a AQP3 RNA, a AREG RNA, a ARNTL2 RNA, a ASF1B RNA, a ATP8B1 RNA, a AURKB RNA, a B3GNT5 RNA, a BAIAP2L1 RNA, a BCAM RNA, a BIK RNA, a BIRC5 RNA, a C15orf48 RNA, a C19orf33 RNA, a C1S RNA, a C8orf4 RNA, a CA12 RNA, a CA9 RNA, a CALML3 RNA, a CAPNS2 RNA, a CBLC RNA, a CCL19 RNA, a CCL5 RNA, a CCNB2 RNA, a CD109 RNA, a CD24 RNA, a CD53 RNA, a CD74 RNA, a CD9 RNA, a CDC20 RNA, a CDC42EP1 RNA, a CDC45 RNA, a CDCA4 RNA, a CDCA5 RNA, a CDCP1 RNA, a CDH1 RNA, a CDH3 RNA, a CDK1 RNA, a CDKN2A RNA, a CDKN2B RNA, a CEACAM5 RNA, a CEACAM6 RNA, a CELSR1 RNA, a CENPW RNA, a CEP55 RNA, a CES1 RNA, a CHMP4C RNA, a CLCA2 RNA, a CLDN1 RNA, a CLDN4 RNA, a CLDN7 RNA, a CNN2 RNA, a COL17A1 RNA, a CPA3 RNA, a CRABP2 RNA, a CSTA RNA, a CTSC RNA, a CTSE RNA, a CX3CL1 RNA, a CXADR RNA, a CXCR4 RNA, a CYBB RNA, a CYP2S1 RNA, a CYP4F11 RNA, a DAPL1 RNA, a DPYSL3 RNA, a

DSC2 RNA, a DSC3 RNA, a DSG2 RNA, a DSG3 RNA, a DSP RNA, a EFNA1 RNA, a EFS RNA, a EGFR RNA, a EGLN3 RNA, a EHD2 RNA, a EHF RNA, a ELF3 RNA, a ELF4 RNA, a EMP1 RNA, a EMP2 RNA, a ENAH RNA, a EPCAM RNA, a EPHA2 RNA, a EPHB3 RNA, a EPHX1 RNA, a ESRP1 RNA, a EVPL RNA, a F11R RNA, a F2R RNA, a F2RL1 RNA, a F3 RNA, a FAM129B RNA, a FAM60A RNA, a FAM83D RNA, a FAM83H RNA, a FAT1 RNA, a FAT2 RNA, a FBLIM1 RNA, a FBP1 RNA, a FCER1G RNA, a FERMT1 RNA, a FGFR2 RNA, a FGFR3 RNA, a FOXE1 RNA, a FOXM1 RNA, a FXYD3 RNA, a GALNT3 RNA, a GBP6 RNA, a GJA1 RNA, a GJB2 RNA, a GJB3 RNA, a GJB5 RNA, a GJB6 RNA, a GNA15 RNA, a GPC1 RNA, a GPC3 RNA, a GPNMB RNA, a GPR87 RNA, a GPRC5A RNA, a GPX2 RNA, a GRHL2 RNA, a GSTA1 RNA, a HAS3 RNA, a HCK RNA, a HOXB7 RNA, a ID1 RNA, a IGF2BP2 RNA, a IGSF9 RNA, a IL2RG RNA, a IMPA2 RNA, a IRF6 RNA, a ITGA2 RNA, a ITGA6 RNA, a ITGB4 RNA, a ITGB6 RNA, a IVL RNA, a JAG2 RNA, a JUP RNA, a KCNS3 RNA, a KIAA1522 RNA, a KIF2C RNA, a KIFC1 RNA, a KITLG RNA, a KLF4 RNA, a KLF5 RNA, a KRT13 RNA, a KRT14 RNA, a KRT15 RNA, a KRT16 RNA, a KRT17 RNA, a KRT18 RNA, a KRT19 RNA, a KRT5 RNA, a KRT6A RNA, a KRT6B RNA, a KRT6C RNA, a KRT7 RNA, a KRT8 RNA, a KRTCAP3 RNA, a LAMP3 RNA, a LAPTM5 RNA, a LGALS7B RNA, a LRP11 RNA, a LRRC4 RNA, a LSP1 RNA, a LSR RNA, a LYPD3 RNA, a MAGEA4 RNA, a MAGEA6 RNA, a MAL2 RNA, a MAOA RNA, a MARCO RNA, a MCM2 RNA, a MDFI RNA, a MET RNA, a MMP14 RNA, a MPZL2 RNA, a MUC1 RNA, a MYBL2 RNA, a MYH14 RNA, a MYOF RNA, a MZB1 RNA, a NCF2 RNA, a NKG7 RNA, a NNMT RNA, a NOTCH3 RNA, a NRARP RNA, a NTRK2 RNA, a NUP210 RNA, a NUSAP1 RNA, a OSGIN1 RNA, a OSMR RNA, a PALLD RNA, a PDPN RNA, a PDZK1IP1 RNA, a PECAM1 RNA, a PERP RNA, a PIGR RNA, a PIGT RNA, a PITX1 RNA, a PKP1 RNA, a PKP3 RNA, a PLEK RNA, a PLEK2 RNA, a PLVAP RNA, a PMP22 RNA, a POSTN RNA, a PPL RNA, a PPP1R14C RNA, a PRAME RNA, a PROM2 RNA, a PRRG4 RNA, a PRSS8 RNA, a PTGES RNA, a PTGFRN RNA, a PTPN6 RNA, a PTPRF RNA, a PTPRZ1 RNA, a RAB25 RNA, a RAB38 RNA, a RAET1L RNA, a RARRES1 RNA, a RBP1 RNA, a RGS1 RNA, a RHCG RNA, a RHOV RNA, a RIN2 RNA, a RIPK4 RNA, a RPS4Y1 RNA, a RRM2 RNA, a S100A10 RNA, a S100A11 RNA, a S100A14 RNA, a S100A16 RNA, a S100A2 RNA, a S100P RNA, a SCNN1A RNA, a SDC1 RNA, a SERINC2 RNA, a SERPINB13 RNA, a SERPINB3 RNA, a SERPINB5 RNA, a SEZ6L2 RNA, a SH3BP4 RNA, a SHISA2 RNA, a SLC1A5 RNA, a

SLC2A1 RNA, a SLC34A2 RNA, a SLC40A1 RNA, a SLC6A8 RNA, a SLC7A8 RNA, a SNAI2 RNA, a SOX2 RNA, a SPI1 RNA, a SPINT1 RNA, a SPINT2 RNA, a SPRR1A RNA, a SPRR1B RNA, a SPRR2A RNA, a SPRR2D RNA, a SPRR2E RNA, a SPRR3 RNA, a ST14 RNA, a STEAP1 RNA, a SULF1 RNA, a SYK RNA, a SYTL1 RNA, a TACSTD2 RNA, a TBC1D2 RNA, a TEAD2 RNA, a TEAD3 RNA, a TFAP2C RNA, a THBD RNA, a THBS2 RNA, a TK1 RNA, a TM4SF1 RNA, a TMC4 RNA, a TMEM30B RNA, a TMEM54 RNA, a TMPRSS11D RNA, a TMPRSS11E RNA, a TMPRSS4 RNA, a TNFRSF18 RNA, a TNS4 RNA, a TOP2A RNA, a TP53I11 RNA, a TP63 RNA, a TPD52 RNA, a TPX2 RNA, a TREM2 RNA, a TRIM29 RNA, a TRIP13 RNA, a TSPAN1 RNA, a TSPAN13 RNA, a TSPAN6 RNA, a TSPAN7 RNA, a TUSC3 RNA, a TYROBP RNA, a UBE2C RNA, a UPK1B RNA, a VAMP8 RNA, a VANGL2 RNA, a WLS RNA, a YAP1 RNA, a ZC3H11A RNA, a ZNF217 RNA, a ZNF750 RNA, and combinations thereof.

**[205]** In some embodiments, a target biomarker signature for lung cancer comprises at least one or more (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, or more) extracellular vesicle-associated surface biomarkers (*e.g.*, ones described herein) and at least one or more (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, or more) surface biomarkers (*e.g.*, ones described herein). In some embodiments, at least one extracellular vesicle-associated surface biomarker and at least one surface biomarker are the same. In some embodiments, at least one extracellular vesicle-associated surface biomarker and at least one surface biomarker(s) of a target biomarker signature for lung cancer are distinct. For example, in some embodiments, a target biomarker signature for lung cancer comprises at least one extracellular vesicle-associated surface biomarker, which is or comprises a SLC34A2 polypeptide, and at least one surface biomarker, which is or comprises a CEACAM6 polypeptide, and/or an EpCAM polypeptide. In some embodiments, a target biomarker signature for lung cancer comprises at least one extracellular vesicle-associated surface biomarker, which is or comprises a CEACAM5 polypeptide, and at least one surface biomarker, which is or comprises a CEACAM6 polypeptide, and/or a SLC34A2 polypeptide.

**[206]** In some embodiments, a target biomarker signature for lung cancer comprises at least one or more (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, or more) extracellular vesicle-associated surface biomarkers (*e.g.*, ones described herein) and at least one or more (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, or more) intravesicular biomarkers (*e.g.*, ones described herein). In some such embodiments, the extracellular vesicle-associated surface biomarker(s) and the intravesicular biomarker(s) can be

encoded by the same gene, while the former is expressed in the membrane of the extracellular vesicle and the latter is expressed within the extracellular vesicle. In some such embodiments, the extracellular vesicle-associated surface biomarker(s) and the intravesicular biomarker(s) can be encoded by different genes.

**[207]** In some embodiments, a target biomarker signature for lung cancer comprises at least one or more (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, or more) extracellular vesicle-associated surface biomarkers (*e.g.*, ones described herein) and at least one or more (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, or more) intravesicular RNA (*e.g.*, mRNA) biomarkers (*e.g.*, ones described herein). In some such embodiments, the extracellular vesicle-associated surface biomarker(s) and the intravesicular RNA (*e.g.*, mRNA) biomarker(s) can be encoded by the same gene. In some such embodiments, the extracellular vesicle-associated surface biomarker(s) and the intravesicular RNA (*e.g.*, mRNA) biomarker(s) can be encoded by different genes.

**[208]** In some embodiments, a target biomarker signature for lung cancer comprises a combination of biomarkers as depicted in Table 4A and/or Table 4B. In some embodiments, a biomarker in such a combination is utilized as a capture probe polypeptide target (as an extracellular vesicle -associated surface biomarker), for example, as depicted in Table 4A and/or Table 4B. In some embodiments, a biomarker in such a combination is utilized as a detection probe polypeptide target (as a target surface biomarker); for example, as depicted in Table 4A and/or Table 4B.

**[209]** In some embodiments, a target biomarker signature for lung cancer comprises at least a SLC34A2 polypeptide (as an extracellular vesicle -associated surface biomarker) and a CEACAM6 polypeptide (as a target surface biomarker).

**[210]** In some embodiments, a target biomarker signature for lung cancer comprises at least a SLC34A2 polypeptide (as an extracellular vesicle-associated surface biomarker) and an EpCAM polypeptide (as a target surface biomarker).

**[211]** In some embodiments, a target biomarker signature for lung cancer comprises at least a CEACAM5 polypeptide (as an extracellular vesicle-associated surface biomarker) and a CEACAM6 polypeptide (as a target surface biomarker).

**[212]** In some embodiments, a target biomarker signature for lung cancer comprises at least a CEACAM5 polypeptide (as an extracellular vesicle-associated surface biomarker) and a SLC34A2 polypeptide (as a target surface biomarker).

**[213]** In some embodiments, a target biomarker signature for lung cancer comprises at least a SLC34A2 polypeptide (as an extracellular vesicle-associated surface biomarker) and at least two target surface biomarkers, which may be or comprise a CEACAM6 polypeptide and an EpCAM polypeptide.

**[214]** In some embodiments, a target biomarker signature for lung cancer comprises at least a CEACAM5 polypeptide (as an extracellular vesicle-associated surface biomarker) and at least two target surface biomarkers, which may be or comprise a CEACAM6 polypeptide and a SLC34A2 polypeptide.

**[215]** In some embodiments, a target biomarker signature for lung cancer comprises at least one of the following: (i) sTn antigen in combination with CEACAM5, CEACAM6, and/or MUC1; and (ii) MUC1 in combination with CEACAM5, Sialyl Lewis X, and/or Lewis Y.

**[216]** In some embodiments, a target biomarker signature for lung cancer comprises at least one of the following combinations: (i) (sTn antigen, CEACAM5); (ii) (sTn antigen, CEACAM5, MUC1); (iii) (sTn antigen, MUC1); (iv) (sTn antigen, CEACAM6); (v) (sTn antigen, CEACAM5, CEACAM6); (vi) (sTn antigen, MUC1, CEACAM6); (vii) (MUC1, CEACAM5); (viii) (MUC1, CEACAM5, Sialyl Lewis X antigen); and (ix) (MUC1, CEACAM5, Lewis Y antigen).

**[217]** In some embodiments, any one of the provided biomarkers can be detected and/or measured by protein and/or RNA (*e.g.*, mRNA) expression levels in wild-type form.

**[218]** In some embodiments, any one of the provided biomarkers can be detected and/or measured by protein and/or RNA (*e.g.*, mRNA) expression levels in mutant form. Thus, in some embodiments, mutant-specific detection of provided biomarkers (*e.g.*, proteins and/or RNA such as, *e.g.*, mRNAs) can be included.

**[219]** As noted herein, in some embodiments, a biomarker is or comprises a particular form of one or more polypeptides or proteins (*e.g.*, a pro- form, a truncated form, a modified form such as a glycosylated, phosphorylated, phosphatidylated, lipidated form, *etc.*). In some embodiments, detection of such form detects a plurality (and, in some embodiments, substantially all) polypeptides present in that form (*e.g.*, containing a particular modification such as, for example, a particular glycosylation, *e.g.*, sialyl-Tn (sTn) glycosylation, *e.g.*, a truncated O-glycan containing a sialic acid α-2,6 linked to GalNAc α-O-Ser/Thr. In some

embodiments, surface biomarkers described herein may comprise a glycosylated form of surface biomarkers. For example, in some embodiments, surface biomarkers ADGRF1, B3GNT3, B3GNT5, HAS3, LARGE2, MAL2, NRCAM, PODXL2, PRSS21, and SCL7A11 may comprise glycosylation. In some embodiments, surface biomarkers described herein may comprise a lipidated form of surface biomarkers. For example, in some embodiments, surface biomarkers PRSS21 and UCHL1 may comprise lipidation.

**[220]** Accordingly, in some embodiments, a surface biomarker can be or comprise a glycosylation moiety (*e.g.*, an sTn antigen moiety, a Tn antigen moiety, or a T antigen moiety). Thompsen-nouvelle (Tn) antigen is an O-linked glycan that is thought to be associated with a broad array of tumors. Tn is a single alpha-linked GalNAc added to Ser or Thr as the first step of a major O-linked glycosylation pathway. A skilled artisan will understand that in certain embodiments, T antigen typically refers to an O-linked glycan with the structure Galβ1-3GalNAc-.

**[221]** In some embodiments, a surface biomarker can be or comprise a tumor-associated post-translational modification. In some embodiments, such a post-translational modification can be or comprise tumor-specific glycosylation patterns such as mucins with glycans aberrantly truncated at the initial GalNAc (*e.g.*, Tn), or combinations thereof.

**[222]** In some embodiments, a target biomarker signature comprises targets of a combination as depicted in Table 4A and/or Table 4B, wherein a target may be used in a capture probe and/or detection probe. In some embodiments, a target biomarker signature comprises a target of capture probe as depicted in Table 4A and/or Table 4B and at least one or more (including, *e.g.*, at least two or more) targets of detection probes (*e.g.*, detection probe 1 and/or detection probe 2).

**[223]** In some embodiments, certain biomarker combinations as depicted in Table 4A and/or Table 4B that may be particularly useful (*e.g.*, with higher sensitivity, specificity and/or PPV) for lung cancer detection can undergo an initial round of screening using an advanced stage (*e.g.,* late stage, *e.g.,* stage III and/or IV) lung cancer sample pool and the healthy control sample pool as a reference. In some embodiments, select combinations can be further tested using early-stage lung cancer sample pools (*e.g.,* stage I and/or II, optionally differentiated as appropriate), benign lung tumor plasma sample pools (*e.g.*, as described herein), non-lung cancer sample pools (*e.g.*, as described herein), and/or any combination thereof. In some embodiments,

biomarker combination performance can be determined by calculating the difference in assay signal (*e.g.*, on a Ct basis) between the healthy sample pools and lung cancer sample pools.

**[224]** In some embodiments, certain biomarker combinations for lung cancer detection can be selected with a delta Ct greater than inter-assay variability. For example, in some embodiments, biomarker combinations with a delta Ct greater than 2.0 (corresponding to a fourfold difference) or 1.0 (corresponding to a twofold difference) are considered to provide particularly effective diagnostic utility (*e.g.*, providing a signal greater than inter-assay variability). See, *e.g.*, Example 10, which provides an exemplary analysis of certain combinations described herein.

**[225]** In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination as described in Table 4A and/or Table 4B. In certain embodiments, a target biomarker signature for lung cancer comprises a set of markers that differentiates late stage lung cancer samples from a control sample (*e.g.*, compared to healthy smoker samples, and/or compared to healthy nonsmoker samples; see *e.g.*, Table 4A and/or Table 4B). In certain embodiments, a target biomarker signature for lung cancer comprises a set of markers that differentiates early stage lung cancer samples from a control sample (*e.g.*, compared to healthy smoker samples, and/or compared to healthy nonsmoker samples; see *e.g.*, Table 4A and/or Table 4B). In some embodiments, an assay directed to detection of a target biomarker signature for lung cancer can comprise a combination of capture and detection probes as described in Table 4A and/or Table 4B.

**[226]** In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination as described in Example 10, Table 4A and/or Table 4B that differentiates subjects suffering from early-stage lung cancer (*e.g.*, early stage non-small cell lung cancer such as LUAD and/or LUSC) from subjects who do not have lung cancer (*e.g.*, healthy subjects or subjects with a condition that is not lung cancer or is not associated with a lung condition). In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination as described in Example 10, Table 4A and/or Table 4B that differentiates subjects suffering from late-stage lung cancer (*e.g.*, late stage non-small cell lung cancer such as LUAD and/or LUSC) from subjects who do not have lung cancer (*e.g.*, healthy subjects or subjects with a condition that is not lung cancer or is not associated with a lung condition). In some embodiments, a target biomarker signature for lung cancer can be or

comprise targets of a combination as described in Example 10, Table 4A and/or Table 4B that differentiates subjects suffering from early-stage lung cancer (*e.g.*, early stage non-small cell lung cancer such as LUAD and/or LUSC) from subjects who are suffering from late stage lung cancer (*e.g.*, late stage non-small cell lung cancer such as LUAD and/or LUSC).

**[227]** In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination that differentiate lung cancer from healthy samples in at least 8 out of 8 conditions tested as described in Example 10, Table 4A and/or Table 4B. In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination that differentiate lung cancer from healthy samples in at least 7 out of 8 conditions tested as described in Example 10, Table 4A and/or Table 4B. In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination that differentiate lung cancer from healthy samples in at least 6 out of 8 conditions tested as described in Example 10, Table 4A and/or Table 4B. In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination that differentiate lung cancer from healthy samples in at least 5 out of 8 conditions tested as described in Example 10, Table 4A and/or Table 4B. In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination that differentiate lung cancer from healthy samples in at least 4 out of 8 conditions tested as described in Example 10, Table 4A and/or Table 4B. In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination that differentiate lung cancer from healthy samples in at least 3 out of 8 conditions tested as described in Example 10, Table 4A and/or Table 4B. In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination that differentiate lung cancer from healthy samples in at least 2 out of 8 conditions tested as described in Example 10, Table 4A and/or Table 4B. In some embodiments, a target biomarker signature for lung cancer can be or comprise targets of a combination that differentiate lung cancer from healthy samples in at least 1 out of 8 conditions tested as described in Example 10, Table 4A and/or Table 4B.

**[228]** In certain embodiments, a target biomarker signature for lung cancer detection comprises: a TNFRSF10B biomarker and a PD-L1 biomarker; or a TNFRSF10B biomarker and a CEACAM6 biomarker; or a TNFRSF10B biomarker and a EGFR biomarker; or a TNFRSF10B biomarker and a IGF1R biomarker; or a ALCAM biomarker and a EPCAM biomarker; or a CEACAM6 biomarker and a MUC1 biomarker; or a EGFR biomarker and a T

antigen biomarker; or a EPCAM biomarker and a T biomarker; or a FOLR1 biomarker and a T antigen biomarker; or a Tn antigen biomarker and a TACSTD2 biomarker; or a TNFRSF10B biomarker and a FOLR1 biomarker; or a TNFRSF10B biomarker and a sTn antigen biomarker; or a ALCAM biomarker and a PD-L1 biomarker; or a EPCAM biomarker and a MUC1 biomarker; or a TNFRSF10B biomarker and a CD55 biomarker; or a TNFRSF10B biomarker and a MUC1 biomarker; or a FOLR1 biomarker and a TACSTD2 biomarker; or a MET biomarker and a MUC1 biomarker; or a MET biomarker and a sTn antigen biomarker; or a MUC1 biomarker and a TACSTD2 biomarker; or a PD-L1 biomarker and a MUC1 biomarker; or a PD-L1 biomarker and a Tn antigen biomarker; or a SLC34A2 biomarker and a MET biomarker; or a SLC34A2 biomarker and a T antigen biomarker; or a TNFRSF10B biomarker and a CEACAM5 biomarker; or a TNFRSF10B biomarker and a MSLN biomarker; or a TNFRSF10B biomarker and a Tn biomarker; or combinations thereof.

**[229]** In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a PD-L1 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a CEACAM6 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and an EGFR biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and an IGF1R biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises an ALCAM biomarker and an EPCAM biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a CEACAM6 biomarker and a MUC1 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises an EGFR biomarker and a T antigen biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises an EPCAM biomarker and a T antigen biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a FOLR1 biomarker and a T antigen biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a Tn antigen biomarker and a TACSTD2 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a FOLR1 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a sTn antigen biomarker. In certain embodiments, a target biomarker signature for lung cancer

detection comprises a ALCAM biomarker and a PD-L1 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a EPCAM biomarker and a MUC1 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a CD55 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a MUC1 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a FOLR1 biomarker and a TACSTD2 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a MET biomarker and a MUC1 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a MET biomarker and a sTn antigen biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a MUC1 biomarker and a TACSTD2 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a PD-L1 biomarker and a MUC1 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a PD-L1 biomarker and a Tn antigen biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a SLC34A2 biomarker and a MET biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a SLC34A2 biomarker and a T antigen biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a CEACAM5 biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a MSLN biomarker. In certain embodiments, a target biomarker signature for lung cancer detection comprises a TNFRSF10B biomarker and a Tn antigen biomarker.

**[230]** In some embodiments, a target biomarker signature comprises a combination of at least two target biomarkers, which combination can be selected from the following: a CYP2S1 polypeptide and a HS6ST2 polypeptide; or a ADAM23 polypeptide and a CYP2S1 polypeptide; or a ADAM23 polypeptide and a CYP4F11 polypeptide; or a ADAM23 polypeptide and a UGT1A6 polypeptide; or a ADAM23 polypeptide and a TMPRSS4 polypeptide; or a ADAM23 polypeptide and a ILDR1 polypeptide; or a DSG3 polypeptide and a UPK1B polypeptide; or a ABCA13 polypeptide and a CYP2S1 polypeptide; or a ABCA13 polypeptide and a ADAM23 polypeptide; or a ADAM23 polypeptide and a LAMC2 polypeptide; or a ADAM23 polypeptide and a HAS3 polypeptide; or a HS6ST2 polypeptide and a LAMC2 polypeptide; or a CYP4F11

polypeptide and a HS6ST2 polypeptide; or a ADAM23 polypeptide and a ULBP2 polypeptide; or a ABCA13 polypeptide and a HS6ST2 polypeptide; or a CYP4F11 polypeptide and a LAMC2 polypeptide; or a ABCA13 polypeptide and a UPK1B polypeptide; or a ABCA13 polypeptide and a CYP4F11 polypeptide; or a CYP2S1 polypeptide and a LAMC2 polypeptide; or a CYP2S1 polypeptide and a CYP4F11 polypeptide; or a HS6ST2 polypeptide and a PIGT polypeptide; or a CYP4F11 polypeptide and a SHISA2 polypeptide; or a ABCA13 polypeptide and a DSG2 polypeptide; or a ADAM23 polypeptide and a LAMB3 polypeptide; or a CYP2S1 polypeptide and a VTCN1 polypeptide; or a CYP4F11 polypeptide and a HAS3 polypeptide; or a CYP4F11 polypeptide and a SLC7A11 polypeptide; or a ADAM23 polypeptide and a DSG3 polypeptide; or a ADAM23 polypeptide and a FERMT1 polypeptide; or a ADAM23 polypeptide and a HS6ST2 polypeptide; or a HS6ST2 polypeptide and a ULBP2 polypeptide; or a CYP4F11 polypeptide and a RAP2B polypeptide; or a RACGAP1 polypeptide and a TFRC polypeptide; or a CYP2S1 polypeptide and a ULBP2 polypeptide; or a KLRG2 polypeptide and a UPK1B polypeptide; or a CLCA2 polypeptide and a CYP2S1 polypeptide; or a ADAM23 polypeptide and a PANX2 polypeptide; or a ABCA13 polypeptide and a HAS3 polypeptide; or a ADAM23 polypeptide and a SLC7A11 polypeptide; or a CYP4F11 polypeptide and a DSG3 polypeptide; or a ADAM23 polypeptide and a CYP4F3 polypeptide; or a CYP2S1 polypeptide and a KLRG2 polypeptide; or a HS6ST2 polypeptide and a SLC7A11 polypeptide; or a ADAM23 polypeptide and a RAP2B polypeptide; or a CYP4F11 polypeptide and a PIGT polypeptide; or a ADAM23 polypeptide and a SDC1 polypeptide; or a ADAM23 polypeptide and a ECE2 polypeptide; or a DSG2 polypeptide and a MARVELD2 polypeptide; or a CYP2S1 polypeptide and a SHISA2 polypeptide; or a ABCA13 polypeptide and a UGT1A6 polypeptide; or a ABCC1 polypeptide and a CYP4F11 polypeptide; or a ADAM23 polypeptide and a PIGT polypeptide; or a CYP2S1 polypeptide and a PIGT polypeptide; or a CYP2S1 polypeptide and a ILDR1 polypeptide; or a CYP2S1 polypeptide and a NECTIN1 polypeptide; or a CYP4F11 polypeptide and a RAB6B polypeptide; or a CYP4F11 polypeptide and a KLRG2 polypeptide; or a ADAM23 polypeptide and a FXYD3 polypeptide; or a CLCA2 polypeptide and a CYP4F11 polypeptide; or a ABCA13 polypeptide and a DSG3 polypeptide; or a CYP4F11 polypeptide and a XXYLT1 polypeptide; or a ADAM23 polypeptide and a PRSS21 polypeptide; or a FGFBP1 polypeptide and a MARVELD3 polypeptide; or a CYP2S1 polypeptide and a HAS3 polypeptide; or a ADAM23 polypeptide and a CLCA2 polypeptide; or a CYP4F11 polypeptide and a NECTIN1 polypeptide;

or a CLCA2 polypeptide and a HS6ST2 polypeptide; or a HS6ST2 polypeptide and a PRSS21 polypeptide; or a CYP2S1 polypeptide and a DSG2 polypeptide; or a CD9 polypeptide and a CYP4F11 polypeptide; or a CYP4F11 polypeptide and a FERMT1 polypeptide; or a CYP4F11 polypeptide and a TFRC polypeptide; or a ADAM23 polypeptide and a EPCAM polypeptide; or a DSG2 polypeptide and a FBXO45 polypeptide; or a CYP2S1 polypeptide and a PANX2 polypeptide; or a CYP2S1 polypeptide and a PRSS21 polypeptide; or a CYP4F11 polypeptide and a SDK1 polypeptide; or a CYP4F11 polypeptide and a ULBP2 polypeptide; or a ABCA13 polypeptide and a PIGT polypeptide; or a HS6ST2 polypeptide and a RAP2B polypeptide; or a CYP2S1 polypeptide and a ECE2 polypeptide; or a ADAM23 polypeptide and a NECTIN1 polypeptide; or a CYP2S1 polypeptide and a SLC7A11 polypeptide; or a ECE2 polypeptide and a HS6ST2 polypeptide; or a CYP4F11 polypeptide and a SDK2 polypeptide; or a PACC1 polypeptide and a TFRC polypeptide; or a CYP4F11 polypeptide and a KPNA2 polypeptide; or a ADAM23 polypeptide and a SLC12A8 polypeptide; or a ADAM23 polypeptide and a APOO polypeptide; or a APOO polypeptide and a MARVELD2 polypeptide; or a CD9 polypeptide and a PACC1 polypeptide; or a ECE2 polypeptide and a UPK1B polypeptide; or a ILDR1 polypeptide and a MARVELD2 polypeptide; or a ITGA2 polypeptide and a TMEM158 polypeptide; or a ADAM23 polypeptide and a UCHL1 polypeptide; or a CYP4F11 polypeptide and a PANX2 polypeptide; or a CYP4F11 polypeptide and a ILDR1 polypeptide; or a CYP4F11 polypeptide and a NRCAM polypeptide; or a ADAM23 polypeptide and a CDH1 polypeptide; or a CYP4F11 polypeptide and a ECE2 polypeptide; or combinations thereof. In some embodiments, a target biomarker in the foregoing combinations may be used as a target of a capture probe and/or a target of a detection probe of assays described herein.

**[231]** In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a CYP2S1 polypeptide and a HS6ST2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a CYP2S1 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a TMPRSS4 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a ILDR1 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ABCA13 polypeptide and a CYP2S1 polypeptide. In

some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ABCA13 polypeptide and a ADAM23 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a HS6ST2 polypeptide and a LAMC2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a CYP4F11 polypeptide and a HS6ST2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ABCA13 polypeptide and a HS6ST2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a CYP4F11 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a UGT1A6 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a LAMC2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a HAS3 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a LAMB3 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a ULBP2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a DSG3 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a CYP4F11 polypeptide and a LAMC2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a CYP4F11 polypeptide and a SLC7A11 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a FERMT1 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ADAM23 polypeptide and a CYP4F3 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a RACGAP1 polypeptide and a TFRC polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a CYP4F11 polypeptide and a HAS3 polypeptide. In some embodiments, a target biomarker signature

comprises at least two target biomarkers, which is or comprises a CD9 polypeptide and a DSG3 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a DSG3 polypeptide and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a DSG3 polypeptide and a RACGAP1 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a EPCAM polypeptide and a LAMP3 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a DSG3 polypeptide and a ITGA2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a CDH1 polypeptide and a DSG3 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a CDH3 polypeptide and a DSG3 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a KPNA2 polypeptide and a ULBP2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a TFRC polypeptide and a ULBP2 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a DSG3 polypeptide and a VTCN1 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a ULBP2 polypeptide and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a DSG3 polypeptide and a NECTIN1 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a DSG3 polypeptide and a PTPRZ1 polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a KPNA2 polypeptide and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises a RACGAP1 polypeptide and a ULBP2 polypeptide. In some embodiments, a target biomarker in the foregoing combinations may be used as a target of a capture probe and/or a target of a detection probe of assays described herein.

**[232]** In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and PD-L1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises

TNFRSF10B and CEACAM6 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and EGFR biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and IGF1R biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and EPCAM biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises EGFR and T antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises EPCAM and T biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises FOLR1 and T biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises Tn antigen and TACSTD2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and FOLR1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and PD-L1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises EPCAM and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and CD55 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises FOLR1 and TACSTD2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MET and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MET and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MUC1 and TACSTD2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or

comprises PD-L1 and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises PD-L1 and Tn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises SLC34A2 and MET biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises SLC34A2 and T antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and CEACAM5 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and MSLN biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and Tn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and T antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and TACSTD2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CD55 and PD-L1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CDH3 and CDH1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM5 and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and EGFR biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and EPCAM biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and FOLR1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and MSLN biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and TACSTD2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and CEACAM6 biomarkers. In some embodiments, a target biomarker signature comprises at least two target

biomarkers, which is or comprises EGFR and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises FOLR1 and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MUC1 and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises SLC34A2 and MSLN biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises T and CDH1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises Tn antigen and IGF1R biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and ALCAM biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CD55 and EPCAM biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and PD-L1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and SLC34A2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and CEACAM5 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and EPCAM biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and MET biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and T biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises EGFR and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises EGFR and TACSTD2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises EPCAM and TACSTD2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises IGF1R and T biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MET and Tn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MSLN and T antigen

biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises PD-L1 and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises PD-L1 and T antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises SLC34A2 and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises SLC34A2 and PD-L1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises SLC34A2 and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises SLC34A2 and Tn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TACSTD2 and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and CDH3 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and EPCAM biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and MET biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and CDH3 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and MET biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and SLC34A2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises ALCAM and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CD55 and CDH1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CD55 and MET biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CDH3 and T biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM5 and MSLN biomarkers. In some embodiments, a

target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM5 and TACSTD2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises CEACAM6 and Tn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and CDH1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and CDH3 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and MUC1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and SLC34A2 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises DSG2 and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises EGFR and MSLN biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises FOLR1 and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises FOLR1 and Tn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MET and T antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MSLN and sTn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MUC1 and T antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises MUC1 and Tn antigen biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises SLC34A2 and EGFR biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and CDH1 biomarkers. In some embodiments, a target biomarker signature comprises at least two target biomarkers, which is or comprises TNFRSF10B and SLC34A2 biomarkers. In some embodiments, a target biomarker in the foregoing combinations may be used as a target of a capture probe and/or a target of a detection probe of assays described herein.

**[233]** In some embodiments, a target biomarker signature comprises a combination of at least three target biomarkers, which combination can be selected from the following: a ADAM23

polypeptide, a CYP2S1 polypeptide, and a LAMC2 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a LAMC2 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a PIGT polypeptide; or a ADAM23 polypeptide, a ILDR1 polypeptide, and a LAMC2 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a DSG2 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a KPNA2 polypeptide; or a ABCA13 polypeptide, a ADAM23 polypeptide, and a UGT1A6 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a ULBP2 polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a NECTIN1 polypeptide; or a ADAM23 polypeptide, a UGT1A6 polypeptide, and a ULBP2 polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a VTCN1 polypeptide; or a ABCA13 polypeptide, a ADAM23 polypeptide, and a CYP2S1 polypeptide; or a CLCA2 polypeptide, a CYP2S1 polypeptide, and a HS6ST2 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a UGT1A6 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a NECTIN1 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a FERMT1 polypeptide; or a ABCA13 polypeptide, a DSG2 polypeptide, and a HAS3 polypeptide; or a ABCA13 polypeptide, a ADAM23 polypeptide, and a DSG3 polypeptide; or a ADAM23 polypeptide, a RAP2B polypeptide, and a TMPRSS4 polypeptide; or a ADAM23 polypeptide, a PIGT polypeptide, and a TMPRSS4 polypeptide; or a ADAM23 polypeptide, a TMPRSS4 polypeptide, and a ULBP2 polypeptide; or a CYP2S1 polypeptide, a HAS3 polypeptide, and a PIGT polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a XXYLT1 polypeptide; or a ADAM23 polypeptide, a ILDR1 polypeptide, and a ULBP2 polypeptide; or a ADAM23 polypeptide, a LAMC2 polypeptide, and a UGT1A6 polypeptide; or a ADAM23 polypeptide, a CLCA2 polypeptide, and a CYP2S1 polypeptide; or a ABCA13 polypeptide, a ADAM23 polypeptide, and a FERMT1 polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a SDC1 polypeptide; or a ABCA13 polypeptide, a ADAM23 polypeptide, and a HAS3 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a HAS3 polypeptide; or a ABCA13 polypeptide, a DSG2 polypeptide, and a HS6ST2 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a SDK2 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a KLRG2 polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a FERMT1 polypeptide; or a ABCA13 polypeptide, a ADAM23 polypeptide, and a CDH3 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a ECE2 polypeptide; or a ADAM23 polypeptide, a CDH3

polypeptide, and a EPCAM polypeptide; or a CELSR1 polypeptide, a CYP2S1 polypeptide, and a HS6ST2 polypeptide; or a ADAM23 polypeptide, a CDH3 polypeptide, and a ILDR1 polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a ILDR1 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a NECTIN1 polypeptide; or a ADAM23 polypeptide, a LAMC2 polypeptide, and a TMPRSS4 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a DSG3 polypeptide; or a ADAM23 polypeptide, a LAMP3 polypeptide, and a UGT1A6 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a RAP2B polypeptide; or a CYP2S1 polypeptide, a ECE2 polypeptide, and a HS6ST2 polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a HAS3 polypeptide; or a ADAM23 polypeptide, a ILDR1 polypeptide, and a LAMB3 polypeptide; or a ADAM23 polypeptide, a ILDR1 polypeptide, and a RAP2B polypeptide; or a CYP2S1 polypeptide, a FBXO45 polypeptide, and a HS6ST2 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a HS6ST2 polypeptide; or a CYP2S1 polypeptide, a LAMC2 polypeptide, and a PIGT polypeptide; or a ADAM23 polypeptide, a ECE2 polypeptide, and a ILDR1 polypeptide; or a ADAM23 polypeptide, a KLRG2 polypeptide, and a TMPRSS4 polypeptide; or a ABCA13 polypeptide, a FERMT1 polypeptide, and a HS6ST2 polypeptide; or a ABCA13 polypeptide, a DSG2 polypeptide, and a PIGT polypeptide; or a CDH3 polypeptide, a CYP2S1 polypeptide, and a ILDR1 polypeptide; or a ADAM23 polypeptide, a ILDR1 polypeptide, and a PIGT polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a ECE2 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a PRSS21 polypeptide; or a HS6ST2 polypeptide, a LAMC2 polypeptide, and a LSR polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a HS6ST2 polypeptide; or a ADAM23 polypeptide, a ECE2 polypeptide, and a TMPRSS4 polypeptide; or a HS6ST2 polypeptide, a LAMC2 polypeptide, and a RAP2B polypeptide; or a ADAM23 polypeptide, a HAS3 polypeptide, and a ILDR1 polypeptide; or a ABCA13 polypeptide, a ADAM23 polypeptide, and a NECTIN1 polypeptide; or a CYP2S1 polypeptide, a DSG3 polypeptide, and a UPK1B polypeptide; or a ADAM23 polypeptide, a SHISA2 polypeptide, and a TMPRSS4 polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a ITGA2 polypeptide; or a ADAM23 polypeptide, a KLRG2 polypeptide, and a UGT1A6 polypeptide; or a CYP2S1 polypeptide, a ILDR1 polypeptide, and a ULBP2 polypeptide; or a ADAM23 polypeptide, a SHISA2 polypeptide, and a UGT1A6 polypeptide; or a CLCA2 polypeptide, a CYP2S1 polypeptide, and a DSG2 polypeptide; or a

ADAM23 polypeptide, a CYP2S1 polypeptide, and a LAMB3 polypeptide; or a ADAM23 polypeptide, a NRCAM polypeptide, and a UGT1A6 polypeptide; or a CYP2S1 polypeptide, a FERMT1 polypeptide, and a ULBP2 polypeptide; or a CYP2S1 polypeptide, a PIGT polypeptide, and a VTCN1 polypeptide; or a ADAM23 polypeptide, a PIGT polypeptide, and a UGT1A6 polypeptide; or a ADAM23 polypeptide, a ECE2 polypeptide, and a UGT1A6 polypeptide; or a ABCA13 polypeptide, a ADAM23 polypeptide, and a SDC1 polypeptide; or a ABCA13 polypeptide, a CYP2S1 polypeptide, and a SDC1 polypeptide; or a ADAM23 polypeptide, a TMPRSS4 polypeptide, and a XXYLT1 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a LAMB3 polypeptide; or a CYP2S1 polypeptide, a LAMB3 polypeptide, and a ULBP2 polypeptide; or a CYP2S1 polypeptide, a UGT1A6 polypeptide, and a ULBP2 polypeptide; or a ADAM23 polypeptide, a KRTCAP3 polypeptide, and a LAMC2 polypeptide; or a CLCA2 polypeptide, a HS6ST2 polypeptide, and a PIGT polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a LSR polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a RACGAP1 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a LAPTM4B polypeptide; or a ADAM23 polypeptide, a LAMC2 polypeptide, and a PRSS21 polypeptide; or a CLCA2 polypeptide, a CYP2S1 polypeptide, and a ULBP2 polypeptide; or a ADAM23 polypeptide, a LAMB3 polypeptide, and a LAMC2 polypeptide; or a ADAM23 polypeptide, a ILDR1 polypeptide, and a XXYLT1 polypeptide; or a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a LARGE2 polypeptide; or a CYP2S1 polypeptide, a LAMC2 polypeptide, and a ULBP2 polypeptide; or a ADAM23 polypeptide, a TMEM132A polypeptide, and a TMPRSS4 polypeptide; or a CYP2S1 polypeptide, a LAMC2 polypeptide, and a SHISA2 polypeptide; or a ADAM23 polypeptide, a CYP2S1 polypeptide, and a FAT2 polypeptide; or a CYP2S1 polypeptide, a CYP4F11 polypeptide, and a KLRG2 polypeptide; or combinations thereof. In some embodiments, at least one target biomarker in the foregoing combinations may be used as a target of a capture probe, and at least two target biomarkers may be used as targets of detection probes.

**[234]** In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a CYP2S1 polypeptide, and a LAMC2 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a LAMC2 polypeptide. In some embodiments, a target biomarker signature comprises at

least three target biomarkers, which is or comprises a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a PIGT polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a ILDR1 polypeptide, and a LAMC2 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ABCA13 polypeptide, a CYP2S1 polypeptide, and a DSG2 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a KPNA2 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ABCA13 polypeptide, a ADAM23 polypeptide, and a UGT1A6 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a CYP2S1 polypeptide, a HS6ST2 polypeptide, and a ULBP2 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a CYP2S1 polypeptide, and a NECTIN1 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a UGT1A6 polypeptide, and a ULBP2 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a LAMC2 polypeptide, and a UGT1A6 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a CDH3 polypeptide, and a EPCAM polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a LAMP3 polypeptide, and a UGT1A6 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a NRCAM polypeptide, and a UGT1A6 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a PIGT polypeptide, and a UGT1A6 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a ECE2 polypeptide, and a UGT1A6 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a LAMC2 polypeptide, and a PRSS21 polypeptide. In some embodiments, a target biomarker signature comprises at least three

target biomarkers, which is or comprises a ADAM23 polypeptide, a LAMB3 polypeptide, and a LAMC2 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a FXYD3 polypeptide, and a ULBP2 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ADAM23 polypeptide, a RAP2B polypeptide, and a UGT1A6 polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a DSG3 polypeptide, a EPCAM polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a DSG3 polypeptide, a KPNA2 polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a DSG3 polypeptide, a LAMP3 polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a DSG3 polypeptide, a ULBP2 polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a DSG3 polypeptide, a RAP2B polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a DSG3 polypeptide, a LMNB2 polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a ABCC1 polypeptide, a DSG3 polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a DSG3 polypeptide, a TFRC polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a CD9 polypeptide, a DSG3 polypeptide, and a UPK1B polypeptide. In some embodiments, a target biomarker signature comprises at least three target biomarkers, which is or comprises a DSG3 polypeptide, a LAMB3 polypeptide, and a UPK1B polypeptide. In some embodiments, at least one target biomarker in the foregoing combinations may be used as a target of a capture probe, and at least two target biomarkers may be used as targets of detection probes.

**[235]** In certain embodiments, a target biomarker signature for lung cancer detection comprises a combination of at least three target biomarkers, which combination can be selected from the following: a CEACAM6 polypeptide, a MUC1 polypeptide, and a sTn antigen

polypeptide; or a TNFRSF10B polypeptide, a MUC1 polypeptide, and a CEACAM6 polypeptide; or a SLC34A2 polypeptide, a MUC1 polypeptide, and a CEACAM6 polypeptide; or a CEACAM6 polypeptide, a MUC1 polypeptide, and a MSLN polypeptide; or a FOLR1 polypeptide, a T antigen polypeptide, and a EGFR polypeptide; or a DSG2 polypeptide, a MUC1 polypeptide, and a CEACAM6 polypeptide; or combinations thereof. In some embodiments, at least one target biomarker in the foregoing combinations may be used as a target of a capture probe, and at least two target biomarkers may be used as targets of detection probes.

**[236]** In certain embodiments, a target biomarker signature for lung cancer detection comprises a CEACAM6 polypeptide, a MUC1 and a sTn antigen polypeptide. In certain embodiments, a target polypeptide signature for lung cancer detection comprises a TNFRSF10B polypeptide, a MUC1 polypeptide, and a CEACAM6 polypeptide. In certain embodiments, a target polypeptide signature for lung cancer detection comprises a SLC34A2 polypeptide, a MUC1 polypeptide, and a CEACAM6 polypeptide. In certain embodiments, a target polypeptide signature for lung cancer detection comprises a CEACAM6 polypeptide, a MUC1 polypeptide, and a MSLN polypeptide. In certain embodiments, a target polypeptide signature for lung cancer detection comprises a FOLR1 polypeptide, a T antigen polypeptide, and a EGFR polypeptide. In certain embodiments, a target polypeptide signature for lung cancer detection comprises a DSG2 polypeptide, a MUC1 polypeptide, and a CEACAM6 polypeptide. In some embodiments, at least one target biomarker in the foregoing combinations may be used as a target of a capture probe, and at least two target biomarkers may be used as targets of detection probes.

**[237]** In some embodiments, a target biomarker signature for lung cancer detection comprises at least one of the following biomarker combinations: (i) sTn antigen in combination with CEACAM5, CEACAM6, and/or MUC1; and (ii) MUC1 in combination with CEACAM5, Sialyl Lewis X, and/or Lewis Y. In some embodiments, at least one target biomarker in the foregoing combinations may be used as a target of a capture probe, and at least two target biomarkers may be used as targets of detection probes.

**[238]** In some embodiments, a target biomarker signature for lung cancer detection comprises at least one of the following biomarker combinations: (i) (sTn antigen, CEACAM5); (ii) (sTn antigen, CEACAM5, MUC1); (iii) (sTn antigen, MUC1); (iv) (sTn antigen, CEACAM6); (v) (sTn antigen, CEACAM5, CEACAM6); (vi) (sTn antigen, MUC1, CEACAM6); (vii) (MUC1, CEACAM5); (viii) (MUC1, CEACAM5, Sialyl Lewis X antigen);

and (ix) (MUC1, CEACAM5, Lewis Y antigen). In some embodiments, at least one target biomarker in the foregoing combinations may be used as a target of a capture probe, and at least two target biomarkers may be used as targets of detection probes.

### *III. Exemplary Methods of Detecting Provided Markers and/or Target Biomarker Signatures for Lung Cancer*

**[239]** In general, the present disclosure provides technologies according to which a target biomarker signature is analyzed and/or assessed in a biological sample (e.g., in some embodiments a bodily fluid-derived sample such as, e.g., but not limited to a blood-derived sample) comprising nanoparticles (including, e.g., extracellular vesicles) from a subject in need thereof; in some embodiments, a diagnosis or therapeutic decision is made based on such analysis and/or assessment.

**[240]** In some embodiments, methods of detecting a target biomarker signature include methods for detecting one or more provided markers of a target biomarker signature as proteins, glycans, or proteoglycans (including, e.g., but not limited to a protein with a carbohydrate or glycan moiety). Exemplary protein-based methods of detecting one or more provided markers include, but are not limited to, proximity ligation assay, mass spectrometry (MS) and immunoassays, such as immunoprecipitation; Western blot; ELISA; immunohistochemistry; immunocytochemistry; flow cytometry; and immuno-PCR. In some embodiments, an immunoassay can be a chemiluminescent immunoassay. In some embodiments, an immunoassay can be a high-throughput and/or automated immunoassay platform.

**[241]** In some embodiments, methods of detecting one or more provided markers as proteins, glycans, or proteoglycans (including, e.g., but not limited to a protein with a carbohydrate or glycan moiety) in a sample comprise contacting a sample with one or more antibody agents directed to the provided markers of interest. In some embodiments, such methods also comprise contacting the sample with one or more detection labels. In some embodiments, antibody agents are labeled with one or more detection labels.

**[242]** In some embodiments, detecting binding between a biomarker of interest and an antibody agent for the biomarker of interest includes determining absorbance values or emission values for one or more detection agents. For example, the absorbance values or emission values are indicative of amount and/or concentration of biomarker of interest expressed by nanoparticles

(*e.g.*, higher absorbance is indicative of higher level of biomarker of interest expressed by nanoparticles). In some embodiments, absorbance values or emission values for detection agents are above a threshold value. In some embodiments, absorbance values or emission values for detection agents is at least 1.3, at least 1.4, at least 1.5, at least 1.6, at least 1.7, at least 1.8, at least 1.9, at least 2.0, at least 2.5, at least 3.0, at least 3.5-fold or greater than a threshold value. In some embodiments, the threshold value is determined across a population of a control or reference group (*e.g.*, non-cancer subjects).

**[243]** In some embodiments, methods of detecting one or more provided markers include methods for detecting one or more provided markers as nucleic acids. Exemplary nucleic acid-based methods of detecting one or more provided markers include, but are not limited to, performing nucleic acid amplification methods, such as polymerase chain reaction (PCR), reverse transcription polymerase chain reaction (RT-PCR), transcription-mediated amplification (TMA), ligase chain reaction (LCR), strand displacement amplification (SDA), and nucleic acid sequence-based amplification (NASBA). In some embodiments, a nucleic acid-based method of detecting one or more provided markers includes detecting hybridization between one or more nucleic acid probes and one or more nucleotide sequences that encode a biomarker of interest. In some embodiments, the nucleic acid probes are each complementary to at least a portion of one of the one or more nucleotide sequences that encode the biomarker of interest. In some embodiments, the nucleotide sequences that encode the biomarker of interest include DNA (*e.g.*, cDNA). In some embodiments, the nucleotides that encode the biomarker of interest include RNA (*e.g.*, mRNA, snRNA, miRNA, siRNA, orphan noncoding RNA, long noncoding RNA, or piwi-interacting RNA).

**[244]** In some embodiments, methods of detecting one or more provided markers involve proximity-ligation-immuno quantitative polymerase chain reaction (pliq-PCR). Pliq-PCR can have certain advantages over other technologies to profile EVs. For example, pliq-PCR can have a sensitivity three orders of magnitude greater than other standard immunoassays, such as ELISAs (Darmanis *et al.*, 2010; which is incorporated herein by reference for the purpose described herein). In some embodiments, a pliq-PCR reaction can be designed to have an ultra-low LOD, which enables to detect trace levels of tumor-derived EVs, for example, down to a thousand EVs per mL.

**[245]** In some embodiments, methods for detecting one or more provided markers may involve other technologies for detecting EVs, including, *e.g.*, Nanoplasmic Exosome (nPLEX) Sensor (Im *et al.*, 2014; which is incorporated herein by reference for the purpose described herein) and the Integrated Magnetic-Electrochemical Exosome (iMEX) Sensor (Jeong *et al.*, 2016; which is incorporated herein by reference for the purpose described herein), which have reported LODs of ~$10^3$ and ~$10^4$ EVs, respectively (Shao *et al.*, 2018; which is incorporated herein by reference for the purpose described herein).

**[246]** In some embodiments, methods for detecting one or more provided biomarkers in nanoparticles can be based on bulk EV sample analysis.

**[247]** In some embodiments, methods for detecting one or more provided biomarkers in nanoparticles can be based on profiling individual EVs (*e.g.*, single-EV profiling assays), which is further discussed in the section entitled "*Exemplary Methods for Profiling Individual Nanoparticles*" below.

**[248]** In some embodiments, nanoparticles in a sample may be captured or immobilized on a solid substrate prior to detecting one or more provided biomarkers in accordance with the present disclosure. In some embodiments, nanoparticles may be captured on a solid substrate surface by non-specific interaction, including, *e.g.*, adsorption. In some embodiments, nanoparticles may be selectively captured on a solid substrate surface. For example, in some embodiments, a solid substrate surface may be coated with an agent that specifically binds to nanoparticles (*e.g.*, an antibody agent specifically targeting such nanoparticles, *e.g.*, associated with lung cancer). In some embodiments, a solid substrate surface may be coated with a member of an affinity binding pair and an entity of interest (*e.g.*, nanoparticles including, e.g., extracellular vesicles) to be captured may be conjugated to a complementary member of the affinity binding pair. In some embodiments, an exemplary affinity binding pair includes, *e.g.*, but is not limited to biotin and avidin-like molecules such as streptavidin. As will be understood by those of skilled in the art, other appropriate affinity binding pairs can also be used to facilitate capture of an entity of interest to a solid substrate surface. In some embodiments, an entity of interest may be captured on a solid substrate surface by application of a current, *e.g.*, as described in Ibsen *et al.* ACS Nano., 11: 6641-6651 (2017) and Lewis *et al.* ACS Nano., 12: 3311-3320 (2018), both of which are incorporated herein by reference for the purpose described

herein, and both of which describe use of an alternating current electrokinetic microarray chip device to isolate extracellular vesicles from an undiluted human blood or plasma sample.

**[249]** A solid substrate may be provided in a form that is suitable for capturing nanoparticles and does not interfere with downstream handling, processing, and/or detection. For example, in some embodiments, a solid substrate may be or comprise a bead (*e.g.*, a magnetic bead). In some embodiments, a solid substrate may be or comprise a surface. For example, in some embodiments, such a surface may be a capture surface of an assay chamber (including, *e.g.*, a tube, a well, a microwell, a plate, a filter, a membrane, a matrix, *etc.*). Accordingly, in some embodiments, a method described herein comprises, prior to detecting provided biomarkers in a sample, capturing or immobilizing nanoparticles on a solid substrate.

**[250]** In some embodiments, a sample may be processed, *e.g.*, to remove undesirable entities such as cell debris or cells, prior to capturing nanoparticles on a solid substrate surface. For example, in some embodiments, such a sample may be subjected to centrifugation, *e.g.*, to remove cell debris, cells, and/or other particulates. Additionally or alternatively, in some embodiments, such a sample may be subjected to size-exclusion-based purification or filtration. Various size-exclusion-based purification or filtration are known in the art and those skilled in the art will appreciate that in some cases, a sample may be subjected to a spin column purification based on specific molecular weight or particle size cutoff. Those skilled in the art will also appreciate that appropriate molecular weight or particle size cutoff for purification purposes can be selected, *e.g.*, based on the size of nanoparticles. For example, in some embodiments, size-exclusion separation methods may be applied to samples comprising nanoparticles to isolate a fraction of nanoparticles that include extracellular vesicles of a certain size (*e.g.*, greater than 30 nm and no more than 1000 nm, or greater than 70 nm and no more than 200 nm). Typically, nanoparticles may range from 30 nm to several micrometers in diameter. See, *e.g.*, Chuo *et al.*, "Imaging extracellular vesicles: current and emerging methods" *Journal of Biomedical Sciences* 25: 91 (2018) which is incorporated herein by reference for the purpose described herein, which provides information of sizes for different extracellular vesicle (EV) subtypes: migrasomes (0.5-3 μm), microvesicles (0.1-1 μm), oncosomes (1-10 μm), exomeres (<50 nm), small exosomes (60-80 nm), and large exosomes (90-120 nm). In some embodiments, nanoparticles having a size range of about 30 nm to 1000 nm may be isolated, for example, in some embodiments by one or more size-exclusion separation methods, for detection assay. In

some embodiments, specific EV subtype(s) may be isolated, for example, in some embodiments by one or more size-exclusion separation methods, for detection assay.

**[251]** In some embodiments, nanoparticles in a sample may be processed prior to detecting one or more provided biomarkers of a target biomarker signature for lung cancer. Different sample processing and/or preparation can be performed, *e.g.*, to stabilize targets (*e.g.*, target biomarkers) in nanoparticles to be detected, and/or to facilitate exposure of targets (*e.g.*, intravesicular proteins and/or RNA such as mRNA, snRNA, miRNA, siRNA, orphan noncoding RNA, long noncoding RNA, or piwi-interacting RNA) to a detection assay (*e.g.*, as described herein), and/or to reduce non-specific binding. Examples of such sample processing and/or preparation are known in the art and include, but are not limited to, crosslinking molecular targets (*e.g.*, fixation), permeabilization of biological entities (*e.g.*, cells or nanoparticles having a size range of interest that includes extracellular vesicles), and/or blocking non-specific binding sites.

**[252]** In one aspect, the present disclosure provides a method for detecting whether a target biomarker signature of lung cancer is present or absent in a biological sample from a subject in need thereof, which may be in some embodiments a biological sample (e.g., in some embodiments a bodily fluid-derived sample such as, e.g., but not limited to a blood-derived sample) comprising nanoparticles. In some embodiments, such a method comprises (a) detecting, in a biological sample such as a blood-derived sample (*e.g.*, a plasma sample) from a subject, biological entities of interest (including, *e.g.*, nanoparticles) expressing a target biomarker signature of lung cancer; and (b) comparing sample information indicative of the level of the target biomarker signature-expressing biological entities of interest (*e.g.*, nanoparticles) in the biological sample (*e.g.*, blood-derived sample) to reference information including a reference threshold level. In some embodiments, a reference threshold level corresponds to a level of biological entities of interest (*e.g.*, nanoparticles) that express such a target biomarker signature in comparable samples from a population of reference subjects, *e.g.*, non-cancer subjects. In some embodiments, exemplary non-cancer subjects include healthy subjects (*e.g.*, healthy subjects of specified age ranges, such as *e.g.*, below age 55 or above age 55), subjects with non-lung related health diseases, disorders, or conditions (including, *e.g.*, subjects having non-lung cancer such as ovarian cancer, colorectal cancer, *etc.*, or subjects having symptoms of chronic obstructive pulmonary disease or chronic lung infections), subjects having benign lung tumors

(*e.g.*, a benign mass observed in the lung through imaging such as chest X-ray or low-dose CT scan), and combinations thereof.

**[253]** In some embodiments, a sample is pre-screened for certain characteristics prior to utilization in an assay as described herein. In some embodiments, a sample meeting certain pre-screening criteria is more suitable for diagnostic applications than a sample failing pre-screening criteria. For example, in some embodiments samples are visually inspected for appearance using known standards, *e.g.*, is the sample normal, hemolyzed (red), icteric (yellow), and/or lipemic (turbid). In some embodiments, samples can then be rated on a known standard scale (*e.g.*, 1, 2, 3, 4, 5) and the results are recorded. In some embodiments, samples are visually inspected for hemolysis (*e.g.*, heme) and rated on a scale from 1-5, where the visual inspection correlates with a known concentration, *e.g.*, where 1 denotes approximately 0 mg/dL, 2 denotes approximately 50 mg/dL, 3 denotes approximately 150 mg/dL, 4 denotes approximately 250 mg/dL, and 5 denotes approximately 525 mg/dL. In some embodiments, samples are visually inspected icteric levels (*e.g.*, bilirubin) and rated on a scale from 1-5, where the visual inspection correlates with a known concentration, *e.g.*, where 1 denotes approximately 0 mg/dL, 2 denotes approximately 1.7 mg/dL, 3 denotes approximately 6.6 mg/dL, 4 denotes approximately 16 mg/dL, and 5 denotes approximately 30 mg/dL. In some embodiments, samples are visually inspected for turbidity (*e.g.* lipids) and rated on a scale from 1-5, where the visual inspection correlates with a known concentration, *e.g.*, where 1 denotes approximately 0 mg/dL, 2 denotes approximately 125 mg/dL, 3 denotes approximately 250 mg/dL, 4 denotes approximately 500 mg/dL, and 5 denotes approximately 1000 mg/dL.

**[254]** In some embodiments, samples scoring lower than a certain level on one or more metrics, *e.g.*, equal to or lower than a score of 4, may be utilized in an assay as described herein. In some embodiments, samples scoring lower than a certain level on one or more metrics, *e.g.*, equal to or lower than a score of 3, may be utilized in an assay as described herein. In some embodiments, samples scoring lower than a certain level on one or more metrics, *e.g.*, equal to or lower than a score of 2, may be utilized in an assay as described herein. In some embodiments, samples scoring lower than a certain level on all three metrics (*e.g.*, hemolyzed, icteric, and lipemic) *e.g.*, equal to or lower than a score of 2, may be utilized in an assay as described herein. In some embodiments, low visual inspection scores on pre-screening criteria such as hemolysis, bilirubin, and/or lipemia (*e.g.*, equal to or lower than a score of 2) may have no appreciable

effect (*e.g.*, not be correlated with) on diagnostic properties (*e.g.*, Ct values) produced in an assay as described herein.

**[255]** In some embodiments, a sample is determined to be positive for the presence of a target biomarker signature (*e.g.*, ones described herein) when it shows an elevated level of target biomarker signature-expressing nanoparticles relative to a reference threshold level (*e.g.*, ones described herein). In some embodiments, a sample is determined to be positive for the presence of a target biomarker signature (e.g., as reflected by the level of target biomarker signature-expressing nanoparticles) if its level is at least 30% or higher, including, *e.g.*, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or higher, as compared to a reference threshold level. In some embodiments, a sample is determined to be positive for the presence of a target biomarker signature (e.g., as reflected by the level of target biomarker signature-expressing nanoparticles) if its level is at least 2-fold or higher, including, *e.g.*, at least 3-fold, at least 4-fold, at least 5-fold, at least 6-fold, at least 7-fold, at least 8-fold, at least 9-fold, at least 10-fold, at least 50-fold, at least 100-fold, at least 250-fold, at least 500-fold, at least 750-fold, at least 1000-fold, at least 2500-fold, at least 5000-fold, or higher, as compared to a reference threshold level.

**[256]** In some embodiments, a binary classification system may be used to determine whether a sample is positive for the presence of a target biomarker signature. For example, in some embodiments, a sample is determined to be positive for the presence of a target biomarker signature (e.g., as reflected by the level of target biomarker signature-expressing nanoparticles) if its level is at or above a reference threshold level, *e.g.*, a cutoff value. In some embodiments, such a reference threshold level (*e.g.*, a cutoff value) may be determined by selecting a certain number of standard deviations away from an average value obtained from control subjects such that a desired sensitivity and/or specificity of a lung cancer detection assay (*e.g.*, ones described herein) can be achieved. In some embodiments, such a reference threshold level (*e.g.*, a cutoff value) may be determined by selecting a certain number of standard deviations away from a maximum assay signal obtained from control subjects such that a desired sensitivity and/or specificity of a lung cancer detection assay (*e.g.*, ones described herein) can be achieved. In some embodiments, such a reference threshold level (*e.g.*, a cutoff value) may be determined by selecting the less restrictive of either (i) a certain number of standard deviations away from an average value obtained from control subjects, or (ii) a certain number of standard deviations

away from a maximum assay signal obtained from control subjects, such that a desired sensitivity and/or specificity of a lung cancer detection assay (*e.g.*, ones described herein) can be achieved. In some embodiments, control subjects for determination of a reference threshold level (*e.g.*, a cutoff value) may include, but are not limited to healthy subjects, subjects with inflammatory conditions (*e.g.*, chronic obstructive pulmonary disease (COPD)), subjects with benign lung tumors, and combinations thereof. In some embodiments, healthy subjects and subjects with inflammatory conditions (*e.g.*, COPD) are included in determination of a reference threshold level (*e.g.*, a cutoff value). In some embodiments, subjects with benign lung tumors are not included in determination of a reference threshold level (*e.g.*, a cutoff value). In some embodiments, a reference threshold level (*e.g.*, a cutoff value) may be determined by selecting at least 1.5 standard deviations (SDs) or higher (including, *e.g.*, at least 1.6, at least 1.7, at least 1.8, at least 1.9, at least 2, at least 2.1, at least 2.2, at least 2.3, at least 2.4, at least 2.5, at least 2.6, at least 2.7, at least 2.8, at least 2.9, at least 3, at least 3.1, at least 3.2, at least 3.3, at least 3.4, at least 3.5, at least 3.6 or higher SDs) away from (i) an average value obtained from control subjects, or (ii) a maximum assay signal obtained from control subjects, such that a desired specificity (*e.g.*, at least 95% or higher specificity [including, *e.g.*, at least 96%, at least 97%, at least 98%, at least 99%, or higher specificity] such as in some embodiments at least 99.8% specificity) of a lung cancer detection assay (*e.g.*, ones described herein) can be achieved. In some embodiments, a reference threshold level (*e.g.*, a cutoff value) may be determined by selecting at least 2.9 SDs (*e.g.*, at least 2.93 SDs) away from (i) an average value obtained from control subjects, or (ii) a maximum assay signal obtained from control subjects, such that a desired specificity (*e.g.*, at least 99%, or higher specificity) of a lung cancer detection assay (*e.g.*, ones described herein) can be achieved. In some embodiments, a reference threshold level (*e.g.*, a cutoff value) may be determined by selecting at least 2.9 SDs (*e.g.*, at least 2.93 SDs) away from the less restrictive of (i) an average value obtained from control subjects, or (ii) a maximum assay signal obtained from control subjects, such that a desired specificity (*e.g.*, at least 99%, or higher specificity) of a lung cancer detection assay (*e.g.*, ones described herein) can be achieved. In some embodiments, such a reference threshold level (*e.g.*, a cutoff value) may be determined based on expression level (*e.g.*, transcript level) of a target biomarker in normal healthy tissues *vs.* in lung cancer samples such that the specificity and/or sensitivity of interest (*e.g.*, as described herein) can be achieved. In some embodiments, a reference threshold level (*e.g.*, a

cutoff value) may vary dependent on, for example, lung cancer stages and/or subtypes and/or patient characteristics, for example, patient age, risks factors for lung cancer (*e.g.*, hereditary risk *vs.* average risk, life-history-associated risk factors), symptomatic/asymptomatic status, and combinations thereof.

**[257]** In some embodiments, such a reference threshold level (*e.g.*, a cutoff value) may be determined based on a log-normal distribution around healthy subjects (*e.g.*, of specified age ranges), and optionally subjects with inflammatory conditions (*e.g.*, COPD) and selection of a level that is necessary to achieve the specificity of interest, *e.g.*, based on prevalence of lung cancer or a subtype thereof. In some embodiments, specificity of interest may be at least 70%, including, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, at least 99.5% or higher.

**[258]** The present disclosure, among other things, also provides technologies for determining whether a subject as having or being susceptible to lung cancer, for example, from a sample comprising nanoparticles with a size range of interest that includes extracellular vesicles. For example, in some embodiments, when a biological sample (e.g., in some embodiments a bodily fluid-derived sample such as, e.g., but not limited to a blood-derived sample) from a subject in need thereof shows a level of target biomarker signature-expressing nanoparticles that is at or above a reference threshold level, *e.g.*, cutoff value (*e.g.*, as determined in accordance with the present disclosure), then the subject is classified as having or being susceptible to lung cancer. In some such embodiments, a reference threshold level (*e.g.*, cutoff value) may be determined based on a log-normal distribution around healthy subjects (*e.g.*, of specified age ranges), and optionally subjects with inflammatory conditions (*e.g.*, COPD) and selection of a level that is necessary to achieve the specificity of interest, *e.g.*, based on prevalence of lung cancer or a subtype thereof. In some embodiments, specificity of interest may be at least 70%, including, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, at least 99.5% or higher.

**[259]** In some embodiments, a reference threshold level (*e.g.*, a cutoff value) may be determined based on expression level (*e.g.*, transcript level) of individual target biomarker(s) of a target biomarker signature in normal healthy tissues *vs.* in lung cancer samples such that the specificity and/or sensitivity of interest (*e.g.*, as described herein) can be achieved. In some embodiments, a reference threshold level (*e.g.*, a cutoff value) may vary dependent on, for

example, lung cancer stages and/or subtypes and/or patient characteristics, for example, patient age, risks factors for lung cancer (*e.g.*, hereditary risk *vs.* average risk, life-history-associated risk factors), symptomatic/asymptomatic status, and combinations thereof.

**[260]** In some embodiments, when a biological sample (e.g., in some embodiments a bodily-fluid derived sample such as, e.g., but not limited to a blood-derived sample) from a subject in need thereof shows a level of biomarker combination that satisfies a reference threshold level, then the subject is classified as having or being susceptible to lung cancer. For example, in some embodiments, when a bodily fluid-derived sample (e.g., but not limited to a blood-derived sample) from a subject in need thereof shows an elevated level of target biomarker signature-expressing extracellular vesicles relative to a reference threshold level, then the subject is classified as having or being susceptible to lung cancer. In some embodiments, a subject in need thereof is classified as having or being susceptible to lung cancer when the subject's biological sample (e.g., in some embodiments a bodily fluid-derived sample such as, e.g., but not limited to a blood-derived sample) shows a level of target biomarker signature-expressing nanoparticles that is at least 30% or higher, including, *e.g.*, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or higher, as compared to a reference threshold level. In some embodiments, a subject in need thereof is classified as having or being susceptible to lung cancer when the subject's biological sample (e.g., in some embodiments a bodily fluid-derived sample such as, e.g., but not limited to a blood-derived sample) shows a level of target biomarker signature-expressing nanoparticles that is at least 2-fold or higher, including, *e.g.*, at least 3-fold, at least 4-fold, at least 5-fold, at least 6-fold, at least 7-fold, at least 8-fold, at least 9-fold, at least 10-fold, at least 20-fold, at least 30-fold, at least 40-fold, at least 50-fold, at least 60-fold, at least 70-fold, at least 80-fold, at least 90-fold, at least 100-fold, at least 250-fold, at least 500-fold, at least 750-fold, at least 1000-fold, or higher, as compared to a reference threshold level.

**[261]** When a biological sample (e.g., in some embodiments a bodily fluid-derived sample such as, e.g., but not limited to a blood-derived sample) from a subject in need thereof shows a comparable level to a reference threshold level, then the subject is classified as not likely to have or as not likely to be susceptible to lung cancer. In some such embodiments, a reference threshold level corresponds to a level of nanoparticles that express a target biomarker signature in comparable samples from a population of reference subjects, *e.g.*, non-cancer

subjects. In some embodiments, exemplary non-cancer subjects include healthy subjects (*e.g.*, healthy subjects of specified age ranges, such as *e.g.*, below age 55 or above age 55), subjects with non-lung related health diseases, disorders, or conditions (including, *e.g.*, subjects having non-lung cancer such as pancreatic cancer, colorectal cancer, *etc.*, or subjects having symptoms of chronic obstructive pulmonary disease or chronic lung infections), subjects having benign lung tumors (*e.g.*, a benign mass observed in the lung through imaging such as chest X-ray or low-dose CT scan), and combinations thereof.

***IV. Exemplary Methods for Profiling Individual Nanoparticles***

**[262]** In some embodiments, assays for profiling individual extracellular vesicles (*e.g.*, single EV profiling assays) can be used to detect one or more provided biomarkers of one or more target biomarker signatures for lung cancer. For example, in some embodiments, such an assay may involve (i) a capture assay through targeting one or more provided markers of a target biomarker signature for lung cancer and (ii) a detection assay for at least one or more additional provided markers of such a target biomarker signature for lung cancer, wherein such a capture assay is performed prior to such a detection assay.

**[263]** A skilled artisan reading the present disclosure will understand that assays described herein for detecting or profiling individual extracellular vesicles can also detect surface biomarkers present on the surfaces of nanoparticles having a size of interest (e.g., in some embodiments a size within the range of about 30 nm to about 1000 nm) that includes extracellular vesicles.

**[264]** In some embodiments, a capture assay is performed to selectively capture tumor-associated nanoparticles (*e.g.*, lung tumor-associated nanoparticles) from a biological sample (*e.g.*, a bodily fluid-derived sample such as a blood-derived sample) of a subject in need thereof. In some embodiments, a capture assay is performed to selectively capture nanoparticles of a certain size range, and/or certain characteristic(s), for example, nanoparticles associated with lung cancer. In some such embodiments, prior to a capture assay, a biological sample (e.g., a bodily fluid-derived sample such as a blood-derived sample) may be pre-processed to remove contaminants, including, *e.g.*, but not limited to soluble proteins and interfering entities such as, *e.g.*, cell debris. For example, in some embodiments, nanoparticles are purified from a biological sample (e.g., a bodily fluid-derived sample such as a blood-derived sample) of a subject using

size exclusion chromatography. In some such embodiments, nanoparticles can be directly purified from a biological sample (e.g., a bodily fluid-derived sample such as a blood-derived sample) using size exclusion chromatography, which in some embodiments may remove at least 90% or higher (including, *e.g.*, at least 93%, 95%, 97%, 99% or higher) of soluble proteins and other interfering agents such as, *e.g.*, cell debris.

**[265]** In some embodiments, a capture assay comprises a step of contacting a biological sample (e.g., a bodily fluid-derived sample such as a blood-derived sample) with at least one capture agent comprising a target-capture moiety that binds to at least one or more provided biomarkers of a target biomarker signature for lung cancer. In some embodiments, a capture assay may be multiplexed, which comprises a step of contacting a biological sample (e.g., a bodily fluid-derived sample such as a blood-derived sample) with a set of capture agents, each capture agent comprising a target-capture moiety that binds to a distinct provided biomarker of a target biomarker signature for lung cancer. In some embodiments, a target-capture moiety is directed to an extracellular vesicle-associated surface biomarker or surface biomarker (*e.g.*, ones as described and/or utilized herein).

**[266]** In some embodiments, such a target-capture moiety may be immobilized on a solid substrate. Accordingly, in some embodiments, a capture agent employed in a capture assay is or comprises a solid substrate comprising at least one or more (*e.g.*, 1, 2, 3, 4, 5, or more) target-capture moiety conjugated thereto, each target-capture moiety directed to an extracellular vesicle-associated surface biomarker or surface biomarker (*e.g.*, ones as described and/or utilized herein). A solid substrate may be provided in a form that is suitable for capturing nanoparticles having a size range of interest that includes extracellular vesicles and does not interfere with downstream handling, processing, and/or detection. For example, in some embodiments, a solid substrate may be or comprise a bead (*e.g.*, a magnetic bead). In some embodiments, a solid substrate may be or comprise a surface. For example, in some embodiments, such a surface may be a capture surface of an assay chamber (including, *e.g.*, a tube, a well, a microwell, a plate, a filter, a membrane, a matrix, *etc.*). In some embodiments, a capture agent is or comprises a magnetic bead comprising a target-capture moiety conjugated thereto.

**[267]** In some embodiments, a detection assay is performed to detect one or more provided biomarkers of a target biomarker signature for lung cancer (*e.g.*, ones that are different from ones targeted in a capture assay) in nanoparticles that are captured by a capture assay (*e.g.*,

as described above). In some embodiments, a detection assay may comprise immuno-PCR. In some embodiments, an immuno-PCR may involve at least one probe targeting a single provided biomarker (*e.g.*, ones described herein) of a target biomarker signature for lung cancer. In some embodiments, an immuno-PCR may involve a plurality of (*e.g.*, at least two, at least three, at least four, or more) probes directed to different epitopes of the same biomarker (*e.g.*, ones described herein) of a target biomarker signature. In some embodiments, an immuno-PCR may involve a plurality of (*e.g.*, at least two, at least three, at least four, or more) probes, each directed to a different provided biomarker described herein.

**[268]** In some embodiments, a detection assay may comprise reverse transcription polymerase chain reaction (RT-PCR). In some embodiments, an RT-PCR may involve at least one primer/probe set targeting a single provided biomarker described herein. In some embodiments, an RT-PCR may involve a plurality of (*e.g.*, at least two, at least three, at least four, or more) primer/probe sets, each set directed to a different provided biomarker described herein.

**[269]** In some embodiments, a detection assay may comprise a proximity-ligation-immuno quantitative polymerase chain reaction (pliq-PCR), for example, to determine co-localization of one or more provided biomarkers of a target biomarker signature for lung cancer within nanoparticles having a size range of interest that includes extracellular vesicles (*e.g.*, captured extracellular vesicles that express at least one extracellular vesicle-associated surface biomarker).

**[270]** In some embodiments, a detection assay employs a target entity detection system that was developed by Applicant and described in U.S. Application No. 16/805,637 (issued as US11,085,089), and International Application PCT/US2020/020529 (published as WO2020180741), both filed February 28, 2020 and entitled "Systems, Compositions, and Methods for Target Entity Detection" (the "'089 patent" and the "'529 application"; both of which are incorporated herein by reference in their entirety) which are, in part, based on interaction and/or co-localization of a target biomarker signature in individual nanoparticles. For example, such a target entity detection system (as described in the '089 patent and'529 application and also further described below in the section entitled "*Provided Target Entity Detection Systems and Methods Involving the Same*") can detect in a sample (*e.g.,* in a biological, environmental, or other sample), in some embodiments at a single entity level,

entities of interest (*e.g.*, biological or chemical entities of interest, such as nanoparticles or analytes) comprising at least one or more (*e.g.*, at least two or more) targets (*e.g.*, molecular targets). Those skilled in the art, reading the present disclosure, will recognize that provided target entity detection systems are useful for a wide variety of applications and/or purposes, including, *e.g.*, for detection of lung cancer. For example, in some embodiments, provided target entity detection systems may be useful for medical applications and/or purposes. In some embodiments, provided target entity detection systems may be useful to screen (*e.g.*, regularly screen) individuals (*e.g.*, in some embodiments which may be asymptomatic individuals, or in some embodiments which may be individuals experiencing one or more symptoms associated with lung cancer, or in some embodiments which may be individuals at risk for lung cancer such as, *e.g.*, individuals with a hereditary risk for lung cancer and/or life-history-associated risk factor, or smoking individuals) for a disease or condition (*e.g.*, lung cancer). In some embodiments, provided target entity detection systems may be useful to screen (*e.g.*, regularly screen) individuals (*e.g.*, in some embodiments which may be asymptomatic individuals, or in some embodiments which may be individuals experiencing one or more symptoms associated with lung cancer, or in some embodiments which may be individuals at risk for lung cancer such as, *e.g.*, individuals with a hereditary risk for lung cancer and/or life-history-associated risk factor, or smoking individuals) for different types of cancer (*e.g.*, for a plurality of different cancers, one of which may be lung cancer). In some embodiments, provided target entity detection systems are effective even when applied to populations comprising or consisting of asymptomatic individuals (*e.g.*, due to sufficiently high sensitivity and/or low rates of false positive and/or false negative results). In some embodiments, provided target entity detection systems may be useful as a companion diagnostic in conjunction with a disease treatment (*e.g.*, treatment of lung cancer).

**[271]** In some embodiments, a plurality of (*e.g.*, at least two or more) detection assays may be performed to detect a plurality of biomarkers (*e.g.*, at least two or more) of one or more target biomarker signatures for lung cancer (*e.g.*, ones that are different from ones targeted in a capture assay) in nanoparticles, *e.g.*, ones that are captured by a capture assay (*e.g.*, as described above). For example, in some embodiments, a plurality of detection assays may comprise (i) a provided target entity detection system or a system described in the '089 patent and '529 application and/or described herein; and (ii) immuno-PCR. In some embodiments, a plurality of

detection assays may comprise (i) a provided target entity detection system or a system described in the '089 patent and '529 application; and (ii) RT-PCR.

**[272]** For example, in some embodiments, a subject's sample comprising extracellular vesicles may be first subjected to detection of surface biomarkers (e.g., as described herein) using a target entity detection system or a system described in the '089 patent and '529 application and/or described herein and then subjected to a lysis buffer to release intravesicular analytes, followed by a nucleic acid assay (e.g., in some embodiments RT-qPCR) for detection of one or more intravesicular RNA biomarkers. In some embodiments, one or more intravesicular RNA biomarkers may be or comprise an mRNA transcript encoded by a biomarker gene described herein. In some embodiments, one or more intravesicular RNA biomarkers may be or comprise a microRNA. In some embodiments, one or more intravesicular RNA biomarkers may be or comprise an orphan noncoding RNA. In some embodiments, one or more intravesicular RNA biomarkers may be or comprise a long noncoding RNA. In some embodiments, one or more intravesicular RNA biomarkers may be or comprise a piwi-interacting RNA. In some embodiments, one or more intravesicular RNA biomarkers may be or comprise a circular RNA. In some embodiments, one or more intravesicular RNA biomarkers may be or comprise a small nucleolar RNA.

***V. Provided Target Entity Detection Systems and Methods Involving the Same***

**[273]** In some embodiments, a target entity detection system that can be useful in a detection assay for one or more provided biomarkers of one or more target biomarker signatures for lung cancer includes a plurality of detection probes each for a specific target (*e.g.*, a provided biomarker of a target biomarker signature). In some embodiments, such a system may comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, or more detection probes each for a specific target (*e.g.*, a provided biomarker of a target biomarker signature). In some embodiments, such a system may comprise 2-50 detection probes each for a specific target (*e.g.*, a provided biomarker of a target biomarker signature). In some embodiments, such a system may comprise 2-30 detection probes each for a specific target (*e.g.*, a provided biomarker of a target biomarker signature). In some embodiments, such a system may comprise 2-25 detection probes each for a specific target (*e.g.*, a provided biomarker of a target biomarker signature). In some

embodiments, such a system may comprise 5-30 detection probes each for a specific target (*e.g.*, a provided biomarker of a target biomarker signature). In some embodiments, such a system may comprise 5-25 detection probes each for a specific target (*e.g.*, a provided biomarker of a target biomarker signature). In some embodiments, at least two of such detection probes in a set may be directed to the same biomarker of a target biomarker signature. In some embodiments, at least two of such detection probes in a set may be directed to the same epitope of the same biomarker of a target biomarker signature. In some embodiments, at least two of such detection probes in a set may be directed to different epitopes of the same biomarker of a target biomarker signature.

**[274]** In some embodiments, detection probes appropriate for use in a target entity detection system provided herein may be used for detection of a single disease or condition, *e.g.*, lung cancer. In some embodiments, detection probes appropriate for use in a target entity detection system provided herein may permit detection of at least two or more diseases or conditions, *e.g.*, one of which is lung cancer. In some embodiments, detection probes appropriate for use in a target entity detection system provided herein may permit detection of lung cancer of certain subtypes including, *e.g.*, lung adenocarcinoma lung cancer, small cell lung cancer, squamous and transitional cell lung cancer, large cell lung cancer, non-small cell carcinoma lung cancer, other specified carcinoma lung cancer, sarcoma lung cancer, and other specified types of lung cancer as known in the art (SEER Cancer Statistics Review 1975-2017). In some embodiments, detection probes appropriate for use in a target entity detection system provided herein may permit detection of lung cancer of certain stages, including, *e.g.*, stage I, stage II, stage III, and/or stage IV. Accordingly, in some embodiments, detection probes appropriate for use in a target entity detection system provided herein may comprise a plurality (*e.g.*, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, or more) of sets of detection probes, wherein each set is directed to detection of a different disease or a different type of disease or condition. For example, in some embodiments, detection probes appropriate for use in a target entity detection system provided herein comprises a plurality (*e.g.*, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, or more) of sets of detection probes, wherein in some embodiments, each set is directed to detection of a different type of cancer, one of which is lung cancer, or in some embodiments, each set is directed to detection of lung cancer of various subtypes and/or stages.

*Detection probes*

**[275]** In some embodiments, a detection probe as provided and/or utilized herein comprises a target-binding moiety and an oligonucleotide domain coupled to the target-binding moiety. In some embodiments, an oligonucleotide domain coupled to a target-binding moiety may comprise a double-stranded portion and a single-stranded overhang extended from at least one end of the oligonucleotide domain. In some embodiments, an oligonucleotide domain coupled to a target-binding moiety may comprise a double-stranded portion and a single-stranded overhang extended from each end of the oligonucleotide domain. In some embodiments, detection probes may be suitable for proximity-ligation-immuno quantitative polymerase chain reaction (pliq-PCR) and be referred to as pliq-PCR detection probes.

*A. Target-binding moieties*

**[276]** A target-binding moiety that is coupled to an oligonucleotide domain is an entity or an agent that specifically binds to a target (*e.g.*, a provided biomarker of a target biomarker signature; those skilled in the art will appreciate that, where the target biomarker is a particular form or moiety/component, the target-binding moiety specifically binds to that form or moiety/component). In some embodiments, a target-binding moiety may have a binding affinity (*e.g.*, as measured by a dissociation constant) for a target (*e.g.*, molecular target) of at least about $10^{-4}$M, at least about $10^{-5}$M, at least about $10^{-6}$M, at least about $10^{-7}$M, at least about $10^{-8}$M, at least about $10^{-9}$M, or lower. Those skilled in the art will appreciate that, in some cases, binding affinity (*e.g.*, as measured by a dissociation constant) may be influenced by non-covalent intermolecular interactions such as hydrogen bonding, electrostatic interactions, hydrophobic and Van der Waals forces between the two molecules. Alternatively or additionally, binding affinity between a ligand and its target molecule may be affected by the presence of other molecules. Those skilled in the art will be familiar with a variety of technologies for measuring binding affinity and/or dissociation constants in accordance with the present disclosure, including, *e.g.*, but not limited to ELISA, surface plasmon resonance (SPR) assays, Luminex Single Antigen (LSA) assays, bio-layer interferometry (BLI) (e.g., Octet) assays, grating-coupled interferometry, and spectroscopic assays.

**[277]** In some embodiments, a target-binding moiety may be or comprise an agent of any chemical class such as, for example, a carbohydrate, a nucleic acid, a lipid, a metal, a

polypeptide, a small molecule, *etc.*, and/or a combination thereof. In some embodiments, a target-binding moiety may be or comprise an affinity agent such as an antibody, affimer, aptamer, lectin, siglec, etc. In some embodiments, a target-binding moiety is or comprises an antibody agent, *e.g.*, an antibody agent that specifically binds to a target or an epitope thereof, *e.g.*, a provided biomarker of a target biomarker signature for lung cancer or an epitope thereof. In some embodiments, a target-binding moiety for a provided biomarker may be a commercially available. In some embodiments, a target-binding moiety for a provided biomarker may be designed and created for the purpose of use in assays as described herein. In some embodiments, a target-binding moiety is or comprises an aptamer, *e.g.*, an aptamer that specifically binds to a target or an epitope thereof, *e.g.*, a provided biomarker of a target biomarker signature for lung cancer or an epitope thereof. In some embodiments, a target-binding moiety is or comprises an affimer molecule that specifically binds to a target or an epitope thereof, *e.g.*, a provided biomarker of a target biomarker signature for lung cancer or an epitope thereof. In some embodiments, a target-binding moiety is or comprises a lectin or siglec that specifically binds to a carbohydrate-dependent marker as provided herein. In some embodiments, such an affimer molecule can be or comprise a peptide or polypeptide that binds to a target or an epitope thereof (*e.g.*, as described herein) with similar specificity and affinity to that of a corresponding antibody. In some embodiments, a target may be or comprise a target that is associated with lung cancer. For example, in some such embodiments, a cancer-associated target can be or comprise a target that is associated with more than one cancer (*i.e.*, at least two or more cancers). In some embodiments, a cancer-associated target can be or comprise a target that is typically associated with cancers. In some embodiments, a cancer-associated target can be or comprise a target that is associated with cancers of a specific tissue, *e.g.*, lung cancer. In some embodiments, a cancer-associated target can be or comprise a target that is specific to a particular cancer, *e.g.*, a particular lung cancer.

**[278]** In some embodiments, a target-binding moiety recognizes and specifically binds to a target present in a biological entity (including, *e.g.*, but not limited to cells and/or nanoparticles). For example, in some embodiments, a target-binding moiety may recognize and specifically bind to a tumor-associated antigen or epitope thereof. In some embodiments, a tumor-associated antigen may be or comprise an antigen that is associated with a cancer such as, for example, skin cancer, brain cancer (including, *e.g.*, glioblastoma), breast cancer, colorectal

cancer, liver cancer, lung cancer, ovarian cancer, pancreatic cancer, prostate cancer, and skin cancer. In some embodiments, a target-binding moiety may recognize a tumor antigen associated with lung cancer (*e.g.*, lung adenocarcinoma, small cell lung cancer, non-small cell lung cancer, *etc.*). In some embodiments, a target-binding moiety may recognize a tumor antigen associated with lung adenocarcinoma.

**[279]** In some embodiments, a target-binding moiety may specifically bind to an intravesicular target, *e.g.*, a provided intravesicular protein or RNA (*e.g.*, mRNA, snRNA, miRNA, siRNA, orphan noncoding RNA, long noncoding RNA, or piwi-interacting RNA). In some embodiments, a target-binding moiety may specifically bind to a surface target that is present on/within nanoparticles, *e.g.*, a surface biomarker present on lung cancer-associated nanoparticles.

**[280]** In some embodiments, a target-binding moiety is assessed for off-target interactions. In some embodiments, a target-binding moiety is assessed using immunocapture followed by mass spectrometry (e.g., to reveal off target binding events in a complex sample). In some embodiments, a target-binding moiety is assessed using protein or glycan arrays, e.g., where many thousands of human proteins or glycans are arrayed on a chip and an antibody's binding is profiled across all available targets (e.g., a specific antibody will only bind to a target of interest). In some embodiments, a target-binding moiety is assessed using traditional immunoassays such as Western blot. In some embodiments, a target-binding moiety is assessed for generic off-target non-specific binding (e.g., binding to other antibodies, DNA, lipids, etc.). In some embodiments, such generic off-target non-specific binding may be measured and identified using a negative control to identify a false positive signal (e.g., suggesting that one or more antibodies bind non-specifically, and not to a target).

**[281]** In some embodiments, a target-binding moiety is directed to a biomarker for a specific condition or disease (*e.g.*, cancer), which biomarker is or has been determined, for example, by analyzing a population or library (*e.g.,* tens, hundreds, thousands, tens of thousands, hundreds of thousands, or more) of patient biopsies and/or patient data to identify such a biomarker (*e.g.*, a predictive biomarker).

**[282]** In some embodiments, a relevant biomarker may be one identified and/or characterized, for example, via data analysis. In some embodiments, for example, a diverse set of data (*e.g.*, in some embodiments comprising one or more of bulk RNA sequencing, single-cell

RNA (scRNA) sequencing, mass spectrometry, histology, post-translational modification data, *in vitro* and/or *in vivo* experimental data) can be analyzed through machine learning and/or computational modeling to identify biomarkers (*e.g.*, predictive markers) that are highly specific to a disease or condition (*e.g.*, cancer).

**[283]** In some embodiments, a target-binding moiety is directed to a tissue-specific target, for example, a target that is associated with a specific tissue such as, for example, brain, breast, colon, ovary and/or other tissues associated with a female reproductive system, pancreas, prostate and/or other tissues associated with a male reproductive system, liver, lung, and skin. In some embodiments, such a tissue-specific target may be associated with a normal healthy tissue and/or a diseased tissue, such as a tumor. In some embodiments, a target-binding moiety is directed to a target that is specifically associated with a normal healthy condition of a subject.

**[284]** In some embodiments, individual target binding entities utilized in a plurality of detection probes (*e.g.*, as described and/or utilized herein) are directed to different targets. In some embodiments, such different targets may represent different marker proteins or polypeptides. In some embodiments, such different targets may represent different epitopes of the same marker proteins or polypeptides. In some embodiments, two or more individual target binding entities utilized in a plurality of detection probes (*e.g.*, as described and/or utilized herein) may be directed to the same target.

**[285]** In some embodiments, individual target binding entities utilized in a plurality of detection probes for detection of lung cancer may be directed to different target biomarkers of a target biomarker signature for lung cancer (*e.g.*, ones as described in the section entitled "*Provided Biomarkers and/or Target Biomarker Signatures for Detection of Lung Cancer*" above). For example, in some embodiments, at least two detection probes in a plurality may have their target binding entities directed to CEACAM6 and EpCAM, respectively. In some embodiments, at least two detection probes in a plurality may have their target binding entities directed to CEACAM6 and SLC34A2, respectively. In some embodiments, at least two detection probes in a plurality may have their target binding entities directed to MUC1 and CEACAM6 respectively.

**[286]** In some embodiments, individual target binding entities utilized in a plurality of detection probes for detection of lung cancer may be directed to the same target biomarker of a target biomarker signature for lung cancer (*e.g.*, ones as described in the section entitled

"*Provided Biomarkers and/or Target Biomarker Signatures for Detection of Lung Cancer*" above). In some embodiments, such target binding entities may be directed to the same or different epitopes of the same target biomarker of such a target biomarker signature for lung cancer. For example, in some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to CEACAM6 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to ALCAM (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to CD55 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to CD1 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to CDH3 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to CD274 (PD-L1) (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to CEACAM5 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to DSG2 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to EGFR (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to EPCAM (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an

extracellular domain thereof, and/or at the same epitope or at different epitopes).In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to FOLR1 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to IG1FR (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to MET (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to MSLN (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to MUC1 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to SLC34A2 (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to sTn antigen (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to Tn antigen (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to T antigen (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities each directed to TACSTD (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes). In some embodiments, at least two detection probes in a plurality may have their target binding entities

each directed to TNFRSF10B (*e.g.*, in its intact protein form or a fragment thereof, *e.g.*, an extracellular domain thereof, and/or at the same epitope or at different epitopes).

*B. Oligonucleotide domains*

**[287]** In some embodiments, an oligonucleotide domain for use in accordance with the present disclosure (*e.g.,* that may be coupled to a target-binding moiety) may comprise a double-stranded portion and a single-stranded overhang extended from one or both ends of the oligonucleotide domain. In some embodiments where an oligonucleotide domain comprises a single-stranded overhang extended from each end, a single-stranded overhang is extended from a different strand of a double-stranded portion. In some embodiments where an oligonucleotide domain comprises a single-stranded overhang extended from one end of the oligonucleotide domain, the other end of the oligonucleotide domain may be a blunt end.

**[288]** In some embodiments, an oligonucleotide domain may comprise ribonucleotides, deoxyribonucleotides, synthetic nucleotide residues that are capable of participating in Watson-Crick type or analogous base pair interactions, and any combinations thereof. In some embodiments, an oligonucleotide domain is or comprises DNA. In some embodiments, an oligonucleotide domain is or comprises peptide nucleic acid (PNA).

**[289]** In some embodiments, an oligonucleotide may have a length that is determined, at least in part, for example, by, *e.g.*, the physical characteristics of an entity of interest (*e.g.*, biological entity such as extracellular vesicles) to be detected, and/or selection and localization of molecular targets in an entity of interest (*e.g.*, biological entity such as extracellular vesicles) to be detected. In some embodiments, an oligonucleotide domain of a detection probe is configured to have a length such that when a first detection probe and a second detection probe bind to an entity of interest (*e.g.*, biological entity such as extracellular vesicles), the first single-stranded overhang and the second single-stranded overhang are in sufficiently close proximity to permit interaction (*e.g.*, hybridization) between the single-stranded overhangs. For example, when an entity of interest (*e.g.*, biological entity) is an extracellular vesicle (*e.g.*, an exosome), oligonucleotide domains of detection probes can each independently have a length such that their respective single-stranded overhangs are in sufficiently close proximity to anneal or interact with each other when the corresponding detection probes are bound to the same extracellular vesicle. For example, in some embodiments, oligonucleotide domains of detection probes for use in

detecting extracellular vesicles (*e.g.*, an exosome) may each independently have a length of about 20 nm to about 200 nm, about 40 nm to about 500 nm, about 40 nm to about 300 nm, or about 50 nm to about 150 nm. In some embodiments, oligonucleotide domains of detection probes for use in detecting extracellular vesicles (*e.g.*, an exosome) may each independently have a length of about 20 nm to about 200 nm. In some embodiments, lengths of oligonucleotide domains of detection probes in a set can each independently vary to increase and/or maximize the probability of them finding each other when they simultaneously bind to the same entity of interest. Such oligonucleotide domains designed for use in detection probes for detecting extracellular vesicles can also be used in detection probes for detecting nanoparticles having a size range of interest that includes extracellular vesicles.

**[290]** Accordingly, in some embodiments, an oligonucleotide domain for use in technologies provided herein may have a length in the range of about 20 up to about 1000 nucleotides. In some embodiments, an oligonucleotide domain may have a length in the range of about 30 up to about 1000 nucleotides. In some embodiments, an oligonucleotide domain may have a length in the range of about 30 to about 500 nucleotides, from about 30 to about 250 nucleotides, from about 30 to about 200 nucleotides, from about 30 to about 150 nucleotides, from about 40 to about 150 nucleotides, from about 40 to about 125 nucleotides, from about 40 to about 100 nucleotides, from about 40 to about 60 nucleotides, from about 50 to about 90 nucleotides, from about 50 to about 80 nucleotides. In some embodiments, an oligonucleotide domain may have a length of at least 20 or more nucleotides, including, *e.g.*, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 250, at least 500, at least 750, at least 1000 nucleotides or more. In some embodiments, an oligonucleotide domain may have a length of no more than 1000 nucleotides or lower, including, *e.g.*, no more than 900, no more than 800, no more than 700, no more than 600, no more than 500, no more than 400, no more than 300, no more than 200, no more than 100, no more than 90, no more than 80, no more than 70, no more than 60, no more than 50, no more than 40 nucleotides, no more than 30 nucleotides, no more than 20 nucleotides, or lower.

**[291]** In some embodiments, an oligonucleotide domain may have a length of about 20 nm to about 500 nm. In some embodiments, an oligonucleotide domain may have a length of about 20 nm to about 400 nm, about 30 nm to about 200 nm, about 50 nm to about 100 nm, about 30 nm to about 70 nm, or about 40 nm to about 60 nm. In some embodiments, an

oligonucleotide domain may have a length of at least about 20 nm or more, including, *e.g.*, at least about 30 nm, at least about 40 nm, at least about 50 nm, at least about 60 nm, at least about 70 nm, at least about 80 nm, at least about 90 nm, at least about 100 nm, at least about 200 nm, at least about 300 nm, at least about 400 nm or more. In some embodiments, an oligonucleotide domain may have a length of no more than 1000 nm or lower, including, *e.g.*, no more than 900 nm, no more than 800 nm, no more than 700 nm, no more than 600 nm, no more than 500 nm, no more than 400 nm, no more than 300 nm, no more than 200 nm, no more than 100 nm or lower.

**[292]** In some embodiments, a double-stranded portion of an oligonucleotide domain for use in technologies provided herein may have a length in the range of about 30 up to about 1000 nucleotides. In some embodiments, a double-stranded portion of an oligonucleotide domain may have a length in the range of about 30 to about 500 nucleotides, from about 30 to about 250 nucleotides, from about 30 to about 200 nucleotides, from about 30 to about 150 nucleotides, from about 40 to about 150 nucleotides, from about 40 to about 125 nucleotides, from about 40 to about 100 nucleotides, from about 50 to about 90 nucleotides, from about 50 to about 80 nucleotides. In some embodiments, a double-stranded portion of an oligonucleotide domain may have a length of at least 30 or more nucleotides, including, *e.g.*, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 250, at least 500, at least 750, at least 1000 nucleotides or more. In some embodiments, a double-stranded portion of an oligonucleotide domain may have a length of no more than 1000 nucleotides or lower, including, *e.g.*, no more than 900, no more than 800, no more than 700, no more than 600, no more than 500, no more than 400, no more than 300, no more than 200, no more than 100, no more than 90, no more than 80, no more than 70, no more than 60, no more than 50, no more than 40 nucleotides or lower.

**[293]** In some embodiments, a double-stranded portion of an oligonucleotide domain may have a length of about 20 nm to about 500 nm. In some embodiments, a double-stranded portion of an oligonucleotide domain may have a length of about 20 nm to about 400 nm, about 30 nm to about 200 nm, about 50 nm to about 100 nm, about 30 nm to about 70 nm, or about 40 nm to about 60 nm. In some embodiments, a double-stranded portion of an oligonucleotide domain may have a length of at least about 20 nm or more, including, *e.g.*, at least about 30 nm, at least about 40 nm, at least about 50 nm, at least about 60 nm, at least about 70 nm, at least about 80 nm, at least about 90 nm, at least about 100 nm, at least about 200 nm, at least about

300 nm, at least about 400 nm or more. In some embodiments, a double-stranded portion of an oligonucleotide domain may have a length of no more than 1000 nm or lower, including, *e.g.*, no more than 900 nm, no more than 800 nm, no more than 700 nm, no more than 600 nm, no more than 500 nm, no more than 400 nm, no more than 300 nm, no more than 200 nm, no more than 100 nm or lower.

**[294]** In some embodiments, a double-stranded portion of an oligonucleotide domain is characterized in that when detection probes are connected to each other through hybridization of respective complementary single-stranded overhangs (*e.g.*, as described and/or utilized herein), the combined length of the respective oligonucleotide domains (including, if any, a linker that links a target-binding moiety to an oligonucleotide domain) is long enough to allow respective target binding entities to substantially span the full characteristic length (*e.g.*, diameter) of an entity of interest (*e.g.*, an extracellular vesicle). For example, in some embodiments where nanoparticles (e.g., extracellular vesicles) are entities of interest, a combined length of oligonucleotide domains (including, if any, a linker that links a target-binding moiety to an oligonucleotide domain) of detection probes may be approximately 50 to 200 nm, when the detection probes are fully connected to each other.

**[295]** In some embodiments, a double-stranded portion of an oligonucleotide domain may comprise a binding site for a primer. In some embodiments, such a binding site for a primer may comprise a nucleotide sequence that is designed to reduce or minimize the likelihood for miss-priming or primer dimers. Such a feature, in some embodiments, can decrease the lower limit of detection and thus increase the sensitivity of systems provided herein. In some embodiments, a binding site for a primer may comprise a nucleotide sequence that is designed to have a similar annealing temperature as another primer binding site.

**[296]** In some embodiments, a double-stranded portion of an oligonucleotide domain may comprise a nucleotide sequence designed to reduce or minimize overlap with nucleic acid sequences (*e.g.*, DNA and/or RNA sequences) typically associated with genome and/or gene transcripts (*e.g.*, genomic DNA and/or RNA, such as mRNA of genes) of a subject (*e.g.*, a human subject). Such a feature, in some embodiments, may reduce or minimize interference of any genomic DNA and/or mRNA transcripts of a subject that may be present (*e.g.*, as contaminants) in a sample during detection.

**[297]** In some embodiments, a double-stranded portion of an oligonucleotide domain may have a nucleotide sequence designed to reduce or minimize formation of self-dimers, homo-dimers, or hetero-dimers.

**[298]** In some embodiments, a single-stranded overhang of an oligonucleotide domain for use in technologies provided herein may have a length of about 2 to about 20 nucleotides. In some embodiments, a single-stranded overhang of an oligonucleotide domain may have a length of about 2 to about 15 nucleotides, from about 2 to about 10 nucleotides, from about 3 to about 20 nucleotides, from about 3 to about 15 nucleotides, from about 3 to about 10 nucleotides. In some embodiments, a single-stranded overhang can have at least 1 to 5 nucleotides in length. In some embodiments, a single-stranded overhang of an oligonucleotide domain may have a length of at least 2 or more nucleotides, including, *e.g.*, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20 nucleotides, or more. In some embodiments, a single-stranded overhang of an oligonucleotide domain may have a length of no more than 20 nucleotides or lower, including, *e.g.*, no more than 15, no more than 14, no more than 13, no more than 12, no more than 11, no more than 10, no more than 9, no more than 8, no more than 7, no more than 6, no more than 5, no more than 4 nucleotides or lower.

**[299]** In some embodiments, a single-stranded overhang of an oligonucleotide domain may have a length of about 1 nm to about 10 nm. In some embodiments, a single-stranded overhang of an oligonucleotide domain may have a length of about 1 nm to about 5 nm. In some embodiments, a single-stranded overhang of an oligonucleotide domain may have a length of at least about 0.5 nm or more, including, *e.g.*, at least about 1 nm, at least about 1.5 nm, at least about 2 nm, at least about 3 nm, at least about 4 nm, at least about 5 nm, at least about 6 nm, at least about 7 nm, at least about 8 nm, at least about 9 nm, at least about 10 nm or more. In some embodiments, a single-stranded overhang of an oligonucleotide domain may have a length of no more than 10 nm or lower, including, *e.g.*, no more than 9 nm, no more than 8 nm, no more than 7 nm, no more than 6 nm, no more than 5 nm, no more than 4 nm, no more than 3 nm, no more than 2 nm, no more than 1 nm or lower.

**[300]** A single-stranded overhang of an oligonucleotide domain is designed to comprise a nucleotide sequence that is complementary to at least a portion of a single-stranded overhang of a second detection probe such that a double-stranded complex comprising a first detection

probe and a second detection probe can be formed through hybridization of the complementary single-stranded overhangs. In some embodiments, nucleotide sequences of complementary single-stranded overhangs are selected for optimal ligation efficiency in the presence of an appropriate nucleic acid ligase. In some embodiments, a single-stranded overhang has a nucleotide sequence preferentially selected for efficient ligation by a specific nucleic acid ligase of interest (*e.g.*, a DNA ligase such as a T4 or T7 ligase). For example, such a single-stranded overhang may have a nucleotide sequence of GAGT, *e.g.*, as described in Song *et al.*, "Enzyme-guided DNA sewing architecture" *Scientific Reports* 5: 17722 (2015), which is incorporated herein by reference for the purpose described herein.

**[301]** When two detection probes couple together through hybridization of respective complementary single-stranded overhangs, their respective oligonucleotide domains comprising the hybridized single-stranded overhangs can, in some embodiments, have a combined length of about 90%-110% or about 95%-105% of a characteristic length (*e.g.*, diameter) of an entity of interest (*e.g.*, a biological entity). For example, in some embodiments when a biological entity is an exosome, the combined length can be about 50 nm to about 200 nm, or about 75 nm to about 150 nm, or about 80 nm to about 120 nm.

*C. Coupling between a target-binding moiety and an oligonucleotide domain*

**[302]** An oligonucleotide domain and a target-binding moiety can be coupled together in a detection probe by a covalent linkage, and/or by a non-covalent association (such as, *e.g.*, a protein-protein interaction such as streptavidin-biotin interaction and/or an ionic interaction). In some embodiments, a detection probe appropriate for use in accordance with the present disclosure is a conjugate molecule comprising a target-binding moiety and an oligonucleotide domain, where the two components are typically covalently coupled to each other, *e.g.*, directly through a bond, or indirectly through one or more linkers. In some embodiments, a target-binding moiety is coupled to one of two strands of an oligonucleotide domain by a covalent linkage (*e.g.*, directly through a bond or indirectly through one or more linkers) and/or by a non-covalent association (such as, *e.g.*, a protein-protein interaction such as streptavidin-biotin interaction and/or ionic interaction).

**[303]** Where linkers are employed, in some embodiments, linkers are chosen to provide for covalent attachment of a target-binding moiety to one or both strands of an oligonucleotide

domain through selected linkers. In some embodiments, linkers are chosen such that the resulting covalent attachment of a target-binding moiety to one or both strands of an oligonucleotide domain maintains the desired binding affinity of the target-binding moiety for its target. In some embodiments, linkers are chosen to enhance binding specificity of a target-binding moiety for its target. Linkers and/or conjugation methods of interest may vary widely depending on a target-binding moiety, *e.g.*, its size and/or charges. In some embodiments, linkers are biologically inert.

**[304]** A variety of linkers and/or methods for coupling a target-binding moiety to an oligonucleotide is known to one of ordinary skill in the art and can be used in accordance with the present disclosure. In some embodiments, a linker can comprise a spacer group at either end with a reactive functional group capable of covalent attachment to a target-binding moiety. Examples of spacer groups that can be used in linkers include, but are not limited to, aliphatic and unsaturated hydrocarbon chains (including, *e.g.*, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, or longer), spacers containing heteroatoms such as oxygen (*e.g.*, ethers such as polyethylene glycol) or nitrogen (polyamines), peptides, carbohydrates, cyclic or acyclic systems that may contain heteroatoms. Non-limiting examples of a reactive functional group to facilitate covalent attachment include nucleophilic functional groups (*e.g.*, amines, alcohols, thiols, and/or hydrazides), electrophilic functional groups (*e.g.*, aldehydes, esters, vinyl ketones, epoxides, isocyanates, and/or maleimides), functional groups capable of cycloaddition reactions, forming disulfide bonds, or binding to metals. In some embodiments, exemplary reactive functional groups, but are not limited to, primary and secondary amines, hydroxamic acids, N- hydroxysuccinimidyl (NHS) esters, dibenzocyclooctyne (DBCO)-NHS esters, azido- NHS esters, azidoacetic acid NHS ester, propargyl-NHS ester, trans-cyclooctene-NHS esters, N-hydroxysuccinimidyl carbonates, oxycarbonylimidazoles, nitrophenylesters, trifluoroethyl esters, glycidyl ethers, vinylsulfones, maleimides, azidobenzoyl hydrazide, N-[4-(p-azidosalicylamino)butyl]-3'-[2'-pyridyldithio]propionamid), bis-sulfosuccinimidyl suberate, dimethyladipimidate, disuccinimidyltartrate, N- maleimidobutyryloxysuccinimide ester, N-hydroxy sulfosuccinimidyl-4- azidobenzoate, N-succinimidyl [4-azidophenyl]-1,3'-dithiopropionate, N- succinimidyl [4-iodoacetyl]aminobenzoate, glutaraldehyde, and succinimidyl 4-[N-maleimidomethyl]cyclohexane-1-carboxylate, 3-(2-pyridyldithio)propionic acid N-

hydroxysuccinimide ester (SPDP), 4-(N-maleimidomethyl)- cyclohexane-1-carboxylic acid N-hydroxysuccinimide ester (SMCC), and any combinations thereof.

**[305]** In some embodiments, a target-binding moiety (*e.g.*, a target binding antibody agent) is coupled or conjugated to one or both strands of an oligonucleotide domain using N-hydrosysuccinimide (NHS) ester chemistry. NHS esters react with free primary amines and result in stable covalent attachment. In some embodiments, a primary amino group can be positioned at a terminal end with a spacer group, *e.g.*, but not limited to an aliphatic and unsaturated hydrocarbon chain (*e.g.*, a C6 or C12 spacer group).

**[306]** In some embodiments, a target-binding moiety (*e.g.*, a target-binding affinity agent) can be coupled or conjugated to one or both strands of an oligonucleotide domain using a site-specific conjugation method known in the art, *e.g.*, to enhance the binding specificity of conjugated target-binding moiety (*e.g.*, conjugated targetbinding affinity agent). Examples of a site-specific conjugation method include, but are not limited to coupling or conjugation through a disulfide bond, C-terminus, carbohydrate residue or glycan, and/or unnatural amino acid labeling. In some embodiments where a target-binding moiety is or comprises an affinity agent, an oligonucleotide can be coupled or conjugated to the target-binding moiety via at least one or more free amine groups present in the target-binding moiety. In some embodiments, an oligonucleotide can be coupled or conjugated to a target-binding moiety that is or comprises an affinity agent via at least one or more reactive thiol groups present in the target-binding moiety. In some embodiments, an oligonucleotide can be coupled or conjugated to a target-binding moiety that is or comprises an antibody agent or a peptide aptamer via at least one or more carbohydrate residues present in the target-binding moiety.

**[307]** In some embodiments, a plurality of oligonucleotides (*e.g.*, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least ten, or more) can be coupled or conjugated to a target-binding moiety (*e.g.*, a target binding antibody agent).

*Exemplary duplex target entity detection system*

**[308]** In some embodiments, a target entity detection system as provided by the present disclosure (and useful, for example, for detecting, *e.g.,* at a single entity level, nanoparticles associated with lung cancer) may comprise a first population of first detection probes (*e.g.*, as described and/or utilized herein) for a provided target biomarker (*e.g.*, ones described herein) and

a second population of second detection probes (*e.g.*, as described and/or utilized herein) for a provided target biomarker (*e.g.*, ones described herein). In some embodiments, the first detection probes and the second detection probes are directed to the same provided target biomarker. In some embodiments, the first detection probes and the second detection probes are directed to different provided target biomarkers.

**[309]** ***FIG. 2*** illustrates an exemplary duplex target entity detection system for detecting, at a single entity level, an entity of interest (*e.g.*, biological entity such as an extracellular vesicle) comprising (i) at least one target (*e.g.*, a provided biomarker of a target biomarker signature for lung cancer) which expression level is high enough such that two molecules of the same target (*e.g.*, a provided biomarker of a target biomarker signature for lung cancer) are found in close proximity, or (ii) at least two or more distinct targets (*e.g.*,. provided biomarkers of a target biomarker signature for lung cancer). A first detection probe comprises a first target-binding moiety (*e.g.*, directed to a target cancer marker 1) and a first oligonucleotide domain coupled to the first target-binding moiety, the first oligonucleotide domain comprising a first double-stranded portion and a first single-stranded overhang extended from one end of the first oligonucleotide domain. As shown in ***FIG. 2***, a first oligonucleotide domain may be resulted from hybridization of a longer strand (strand 3) and a shorter strand (strand 1), thereby forming a double-stranded portion and a single-stranded overhang at one end. In some embodiments, a first target-binding moiety (*e.g.*, directed to target cancer marker 1) is coupled (*e.g.*, covalently coupled) to a5′end or 3′ end of a strand of a first oligonucleotide domain (*e.g.*, strand 1). In some embodiments, a 5′ end or 3′ end of a strand that is coupled to a first target-binding moiety may be modified with a linker (*e.g.*, as described and/or utilized herein with or without a spacer group). In some embodiments, a 5′ end of another strand of a first oligonucleotide domain (*e.g.*, strand 3) has a free phosphate group.

**[310]** In the embodiment depicted in ***FIG. 2,*** a second detection probe comprises a second target-binding moiety (*e.g.*, directed to a target cancer marker 2) and a second oligonucleotide domain coupled to the second target-binding moiety, the second oligonucleotide domain comprising a second double-stranded portion and a second single-stranded overhang extended from one end of the second oligonucleotide domain. As shown in ***FIG. 2***, a second oligonucleotide domain may be resulted from hybridization of a longer strand (strand 4) and a shorter strand (strand 2), thereby forming a double-stranded portion and a single-stranded

overhang at one end. In some embodiments, a second target-binding moiety (*e.g.*, directed to a target cancer marker 2) is coupled (*e.g.*, covalently coupled) to a 5′ end of a strand of a second oligonucleotide domain (*e.g.*, strand 2). In some embodiments, a 5′ end of a strand that is coupled to a second target-binding moiety may be modified with a linker (*e.g.*, as described and/or utilized herein with or without a spacer group). In some embodiments, a 5′ end of another strand of a second oligonucleotide domain (*e.g.*, strand 4) has a free phosphate group.

**[311]** At least portions of a first single-stranded overhang and a second single-stranded overhang are complementary to each other such that they can hybridize to form a double-stranded complex when they are in sufficiently close proximity, *e.g.*, when a first detection probe and a second detection probe simultaneously bind to the same entity of interest (*e.g.*, biological entity such as extracellular vesicle). In some embodiments, a first single-stranded overhang and a second single-stranded overhang have equal lengths such that when they hybridize to form a double-stranded complex, there is no gap (other than a nick to be ligated) between their respective oligonucleotide domains and each respective target-binding moiety is located at an opposing end of the double-stranded complex. For example, in some embodiments, a double-stranded complex forms before ligation occurs, wherein the double-stranded complex comprises a first detection probe and a second detection probe coupled to each other through direct hybridization of their respective single-stranded overhangs (*e.g.*, having 4 nucleotides in length), wherein each respective target-binding moiety (*e.g.*, directed to a target cancer marker 1 and a target cancer marker 2, respectively) is present at opposing ends of the double-stranded complex. In such embodiments, both strands of the double-stranded complex (comprising a nick between respective oligonucleotide domains) are ligatable, *e.g.*, for amplification and detection. In some embodiments, a double-stranded complex (*e.g.*, before ligation occurs) can comprise an entity of interest (*e.g.*, a biological entity such as an extracellular vesicle), wherein a first target-binding moiety (*e.g.*, directed to a target cancer marker 1) and a second target-binding moiety (*e.g.*, directed to a target cancer marker 2) are simultaneously bound to the entity of interest.

**[312]** In some embodiments of a duplex target entity detection system for detection of lung cancer, a first target-binding moiety of a first detection probe may be directed to a first target surface biomarker (*e.g.*, ones provided in the section entitled "*Provided Biomarkers and/or Target Biomarker Signatures for Detection of Lung Cancer*"), while a second target-binding moiety of a second detection probe may be directed to a second target surface biomarker (*e.g.*,

ones provided in the section entitled "*Provided Biomarkers and/or Target Biomarker Signatures for Detection of Lung Cancer*"). In some embodiments, a first target-binding moiety of a first detection probe may be directed to a first target intravesicular biomarker (*e.g.*, ones provided in the section entitled "*Provided Biomarkers and/or Target Biomarker Signatures for Detection of Lung Cancer*"), while a second target-binding moiety of a second detection probe may be directed to a second target intravesicular biomarker (*e.g.*, ones provided in the section entitled "*Provided Biomarkers and/or Target Biomarker Signatures for Detection of Lung Cancer*"). In some embodiments, the first target-binding moiety and the second target-binding moiety may be directed to the same or different epitopes of the same target surface biomarker or of the same target intravesicular biomarker. In some embodiments, the first target-binding moiety and the second target-binding moiety may be directed to the different target surface biomarkers or different target intravesicular biomarkers.

**[313]** In some embodiments of a duplex target entity detection system for detection of lung cancer, a first detection probe comprises a first target-binding moiety directed to CEACAM6 (*e.g.*, in intact protein form or a fragment thereof, such as an extracellular domain thereof) conjugated to a first oligonucleotide domain; whereas a second detection probe comprises a second target-binding moiety detected to CEACAM6 (*e.g.*, in intact protein form or a fragment thereof, such as an extracellular domain thereof) conjugated to a second oligonucleotide domain. In some such embodiments, the first target-binding moiety and the second target-binding moiety can be directed to the same or different epitope(s) of CEACAM6 (*e.g.*, in intact protein form or a fragment thereof, such as an extracellular domain thereof). In some embodiments, the double stranded portion of a first oligonucleotide domain and a second oligonucleotide domain may be the same. In some embodiments, the double stranded portion of a first oligonucleotide domain and a second oligonucleotide domain may be different.

**[314]** In some embodiments of a duplex target entity detection system for detection of lung cancer, a first detection probe comprises a first target-binding moiety directed to CEACAM6 polypeptide conjugated to a first oligonucleotide domain; whereas a second detection probe comprises a second target-binding moiety directed to EpCAM polypeptide conjugated to a second oligonucleotide domain. In some embodiments, the double stranded portion of a first oligonucleotide domain and a second oligonucleotide domain may be the same.

In some embodiments, the double stranded portion of a first oligonucleotide domain and a second oligonucleotide domain may be different.

**[315]** In some embodiments of a duplex target entity detection system for detection of lung cancer, a first detection probe comprises a first target-binding moiety directed to CEACAM6 polypeptide conjugated to a first oligonucleotide domain; whereas a second detection probe comprises a second target-binding moiety directed to SLC34A2 polypeptide conjugated to a second oligonucleotide domain. In some embodiments, the double stranded portion of a first oligonucleotide domain and a second oligonucleotide domain may be the same. In some embodiments, the double stranded portion of a first oligonucleotide domain and a second oligonucleotide domain may be different.

**[316]** In some embodiments, a duplex target entity detection system for detection of lung cancer may comprise at least two distinct sets of detection probes. For example, in some embodiments, each set may be directed to a distinct target biomarker signature comprising one or more target biomarkers (*e.g.*, ones described herein).

**[317]** In some embodiments, a duplex target entity detection system comprising at least two distinct sets of detection probes may also comprise a capture assay comprising a capture agent directed to an extracellular vesicle-associated surface biomarker.

**[318]** In some embodiments, any combination of biomarker probes (*e.g.*, a biomarker signature) including capture probes or detection probes as described herein may be utilized in combination with any other set of biomarker probes (*e.g.*, a biomarker signature) including capture probes or detection probes as described herein.

*Exemplary triplex or multiplex (n≥3) target entity detection system*

**[319]** In some embodiments, a target entity detection system as provided by the present disclosure (and useful, for example, for detecting, *e.g.,* at a single entity level, nanoparticles associated with lung cancer) may comprise n populations of distinct detection probes (*e.g.*, as described and/or utilized herein), wherein n ≥3. For example, in some embodiments when n =3, a target entity detection system may comprise a first detection probe (*e.g.*, as described and/or utilized herein) for a first target, a population of a second detection probe (*e.g.*, as described and/or utilized herein) for a second target, and a population of a third detection probe (*e.g.*, as described and/or utilized herein) for a third target.

**[320]** ***FIG. 13*** illustrates an exemplary triplex target entity detection system for detecting, at a single entity level, an entity of interest (*e.g.*, a biological entity such as a nanoparticle) comprising three distinct molecular targets. A first detection probe comprises a first target-binding moiety (*e.g.*, anti-cancer marker 1 antibody agent) and a first oligonucleotide domain coupled to the first target-binding moiety, the first oligonucleotide domain comprising a first double-stranded portion and a first single-stranded overhang extended from one end of the first oligonucleotide domain. As shown in ***FIG. 13***, a first oligonucleotide domain may be resulted from hybridization of a longer strand (strand 8) and a shorter strand (strand 1), thereby forming a double-stranded portion and a single-stranded overhang at one end. In some embodiments, a first target-binding moiety (*e.g.*, anti-cancer marker 1 antibody agent) is coupled (*e.g.*, covalently coupled) to a 5′ end of a strand of a first oligonucleotide domain (*e.g.*, strand 1). In some embodiments, a 5′ end of a strand that is coupled to a first target-binding moiety may be modified with a linker (*e.g.*, as described and/or utilized herein with or without a spacer group). In some embodiments, a 5′ end of another strand of a first oligonucleotide domain (*e.g.*, strand 8) has a free phosphate group.

**[321]** In the embodiment depicted in ***FIG. 13***, a second detection probe comprises a second target-binding moiety (*e.g.*, anti-cancer marker 3 antibody agent) and a second oligonucleotide domain coupled to the second target-binding moiety, the second oligonucleotide domain comprising a second double-stranded portion and a second single-stranded overhang extended from one end of the second oligonucleotide domain. As shown in ***FIG. 13***, a second oligonucleotide domain may be resulted from hybridization of a longer strand (strand 4) and a shorter strand (strand 2), thereby forming a double-stranded portion and a single-stranded overhang at one end. In some embodiments, a second target-binding moiety (*e.g.*, anti-cancer marker 3 antibody agent) is coupled (*e.g.*, covalently coupled) to a 5′ end of a strand of a second oligonucleotide domain (*e.g.*, strand 2). In some embodiments, a 5′ end of a strand that is coupled to a second target-binding moiety may be modified with a linker (*e.g.*, as described and/or utilized herein with or without a spacer group). In some embodiments, a 5′ end of another strand of a second oligonucleotide domain (*e.g.*, strand 4) has no free phosphate group.

**[322]** A third detection probe comprises a third target-binding moiety (*e.g.*, anti-cancer marker 2 antibody agent) and a third oligonucleotide domain coupled to the third target-binding moiety, the third oligonucleotide domain comprising a third double-stranded portion and a

single-stranded overhang extended from each end of the third oligonucleotide domain. For example, a single-stranded overhang is extended from one end of a strand of a third oligonucleotide domain while another single-stranded overhang is extended from an opposing end of a different strand of the third oligonucleotide domain. As shown in ***FIG. 13***, a third oligonucleotide domain may be resulted from hybridization of portions of two strands (*e.g.*, strands 9 and 10), thereby forming a double-stranded portion and a single-stranded overhang at each end. For example, a single-stranded overhang (3A) is formed at a 5′ end of strand 9 of a third detection probe, wherein the 5′ end of strand 9 has a free phosphate group. Additionally, a single-stranded overhang (3B) is formed at a 5′ end of strand 10 of the same third detection probe and a third target-binding moiety (*e.g.*, anti-target 2 antibody agent) is also coupled (*e.g.*, covalently coupled) to the 5′ end of strand 10. In some embodiments, a 5′ end of a strand (*e.g.*, strand 10) that is coupled to a third target-binding moiety may be modified with a linker (*e.g.*, as described and/or utilized herein with or without a spacer group).

**[323]** When all three detection probes are in sufficiently close proximity, *e.g.*, when all three detection probes simultaneously bind to the same entity of interest (*e.g.*, biological entity), (i) at least a portion of a single-stranded overhang (*e.g.*, 3A) of a third detection probe is hybridized to a corresponding complementary portion of a single-stranded overhang of a second detection probe, and (ii) at least a portion of another single-stranded overhang (*e.g.*, 3B) of the third detection probe is hybridized to a corresponding complementary portion of a single-stranded overhang of a first detection probe. As a result, a double-stranded complex comprising all three detection probes coupled to each other in a linear arrangement is formed by direct hybridization of corresponding single-stranded overhangs. See, *e.g.*, ***FIG. 13***.

**[324]** In some embodiments involving use of at least three or more ($n \geq 3$) detection probes in provided technologies, when single-stranded overhangs of detection probes anneal to each respective partner(s) to form a double-stranded complex, at least (n-2) target-binding moiety/moieties is/are present at internal position(s) of the double-stranded complex. In such embodiments, it is desirable to have internal target binding moieties present in a single strand of the double-stranded complex such that another strand of the double-stranded complex is free of any internal target binding moieties and is thus ligatable to form a ligated template. *e.g.*, for amplification and detection. See, *e.g.*, ***FIG. 13*** (using three detection probes), ***FIG. 14*** (using four detection probes), and ***FIG. 15*** (using n detection probes).

**[325]** In some embodiments where a strand of a double-stranded complex comprises at least one or more internal target binding moieties, the strand comprises a gap between an end of an oligonucleotide strand of a detection probe to which the internal target-binding moiety is coupled and an end of an oligonucleotide strand of another detection probe. The size of the gap is large enough such that the strand becomes non-ligatable in the presence of a nucleic acid ligase. In some embodiments, the gap may be 2-8 nucleotides in size or 2-6 nucleotides in size. In some embodiments, the gap is 6 nucleotides in size. In some embodiments, the overlap (hybridization region between single-stranded overhangs) can be 2-15 nucleotides in length or 4-10 nucleotides in length. In some embodiments, the overlap (hybridization region between single-stranded overhangs) is 8 nucleotides in length. The size of the gap and/or hybridization region are selected to provide an optimum signal separation from a ligated template (comprising no internal target binding moieties) and non-ligated template (comprising at least one internal target-binding moiety). It should be noted that while ***FIGS. 13-15*** do not show binding of detection probes to an entity of interest (*e.g.*, a biological entity), a double-stranded complex (*e.g.*, before ligation occurs) can comprise an entity of interest (*e.g.*, a biological entity such as extracellular vesicles), wherein at least three or more target binding moieties are simultaneously bound to the entity of interest.

**[326]** In some embodiments, selection of a combination (*e.g.,* a set) of detection probes (*e.g.*, number of detection probes and/or specific biomarkers) for use in a target entity detection system provided herein (*e.g.*, a duplex, triplex or multiplex target entity detection system described herein) is based on, for example, a desired specificity and/or a desired sensitivity that is deemed to be optimal for a particular application. For example, in some embodiments, a combination of detection probes is selected for detection of lung cancer (*e.g.*, for stage I, II, III, or IV) such that it provides a specificity of at least 95% or higher, including, *e.g.*, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, at least 99.8% or higher. In some embodiments, a combination of detection probes is selected for detection of lung cancer (*e.g.*, for stage I, II, III, or IV) such that it provides a sensitivity of at least 30% or higher, including, *e.g.*, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or higher. In some embodiments, a combination of detection probes is selected for detection of lung cancer (*e.g.*, for stage I, II, III, or IV) such that it provides a positive predictive value of at least 8% or higher, including, *e.g.*, at least 9%, at least 10%, at least 15%, at least

20%, at least 25%, at least 30%, at least 40%, at least 50%, or higher. In some embodiments, a combination of detection probes is selected for detection of lung cancer (*e.g.*, for stage I, II, III, or IV) such that it provides a positive predictive value of at least 2% or higher, including, *e.g.,* at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, or higher. In some embodiments, a combination of detection probes is selected for detection of lung cancer (*e.g.,* for stage I, II, III, or IV) such that it provides a limit of detection (LOD) below $1x10^7$ EV/mL sample or lower, including, *e.g.*, below $7x10^6$ EV/mL sample, below $6x10^6$ EV/mL sample, below $5x10^6$ EV/mL sample, below $4x10^6$ EV/mL sample, below $3x10^6$ EV/mL sample, below $2x10^6$ EV/mL sample, below $1x10^6$ EV/mL sample, or lower. In some embodiments, such lung cancer detection assay may be used to detect different subtypes of lung cancer including, *e.g.*, lung adenocarcinoma, small cell lung cancer, squamous and transitional cell lung cancer, large cell lung cancer, non-small cell carcinoma lung cancer, other specified carcinoma lung cancer, sarcoma lung cancer, and other specified types of lung cancer as known in the art (SEER Cancer Statistics Review 1975-2017). In some embodiments, such lung cancer detection assay may be used to detect lung cancer of an epithelial origin. In some embodiments, such lung cancer detection assay may be used to detect non-small cell lung cancer (*e.g.*, lung adenocarcinoma and/or lung squamous cell carcinoma). In some embodiments, such lung cancer detection assay may be used to detect lung adenocarcinoma. In some embodiments, such lung cancer detection assay may be used to detect lung squamous cell carcinoma.

**[327]** In some embodiments, a combination (*e.g.*, a set) of detection probes, rather than individual detection probes, confers specificity to detection of a disease, disorder, or condition (*e.g.*, a particular lung cancer and/or a stage of lung cancer as described herein), for example, one or more individual probes may be directed to a target that itself is not specific to lung cancer. For example, in some embodiments, a useful combination of detection probes in a target entity detection system provided herein (*e.g.*, a duplex, triplex or multiplex target entity detection system described herein) may comprise at least one detection probe directed to a target specific for the relevant disease, disorder, or condition (*i.e.,* a target that is specific to the relevant disease, disorder, or condition), and may further comprise at least one detection probe directed to a target that is not necessarily or completely specific for the relevant disease, disorder, or condition (*e.g.,* that may also be found on some or all cells that are healthy, are not of the

particular disease, disorder, or condition, and/or are not of the particular disease stage of interest). That is, as will be appreciated by those skilled in the art reading the present specification, so long as the set of detection probes utilized in accordance with the present invention is or comprises a plurality of individual detection probes that together are specific for detection of the relevant disease, disorder, or condition (*i.e.,* sufficiently distinguish biological entities for detection that are associated with the relevant disease, disorder, or condition from other biological entities not of interest for detection), the set is useful in accordance with certain embodiments of the present disclosure.

**[328]** In some embodiments, a target entity detection system provided herein (*e.g.*, a duplex, triplex or multiplex target entity detection system described herein) can comprise at least one or more (*e.g.*, at least 2 or more) control probes (in addition to target-specific detection probes, *e.g.*, as described and/or utilized herein, for example, in some embodiments to recognize disease-specific biomarkers such as cancer-specific biomarkers and/or tissue-specific biomarkers). In some embodiments, a control probe is designed such that its binding to an entity of interest (*e.g.*, a biological entity) inhibits (completely or partially) generation of a detection signal.

**[329]** In some embodiments, a control probe comprises a control binding moiety and an oligonucleotide domain (*e.g.*, as described and/or utilized herein) coupled to the control binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang extended from one end of the oligonucleotide domain. A control binding moiety is an entity or moiety that bind to a control reference. In some embodiments, a control reference can be or comprise a biomarker that is preferentially associated with a normal healthy cell. In some embodiments, a control reference can be or comprise a biomarker preferentially associated from a non-target tissue. In some embodiments, inclusion of a control probe can selectively remove or minimize detectable signals generated from false positives (*e.g.*, entities of interest comprising a control reference, optionally in combination with one or more targets to be detected). Other control probes described in U.S. Application No. 16/805,637 (published as US2020/0299780; issued as US11,085,089), and International Application PCT/US2020/020529 (published as WO2020180741), both filed February 28, 2020 and entitled "Systems, Compositions, and Methods for Target Entity Detection," the entire contents of each application are incorporated herein by reference in their entirety, can be useful in provided target entity detections systems.

**[330]** In some embodiments, the present disclosure provides insights, among other things, that detection probes as described or utilized herein may non-specifically bind to a solid substrate surface and some of them may remain in an assay sample even after multiple washes to remove any excess or unbound detection probes; and that such non-specifically bound detection probes may come off from the solid substrate surface and become free-floating in a ligation reaction, thus allowing them to interact with one another to generate a non-specific ligated template that produces an undesirable background signal. Accordingly, in some embodiments, a target entity detection system provided herein (*e.g.*, a duplex, triplex, or multiplex target entity detection described herein) can comprise at least one or more (*e.g.*, at least 2 or more) inhibitor oligonucleotides that are designed to capture residual detection probes that are not bound to an entity of interest but remain as free agents in a ligation reaction, thereby preventing such free-floating detection probes from interacting with other free-floating complementary detection probes to produce an undesirable background signal. In some embodiments, an inhibitor oligonucleotide may be or comprise a single-stranded or double-stranded oligonucleotide comprising a binding domain for a single-stranded overhang of a detection probe (*e.g.*, as described or utilized herein), wherein the inhibitor oligonucleotide does not comprise a primer binding site. The absence of such a primer binding site in an inhibitor oligonucleotide prevents a primer from binding to a non-specific ligated template resulting from ligation of a detectable probe to an inhibitor oligonucleotide, thereby reducing or inhibiting the non-specific ligated template from amplification and/or detection, *e.g.*, by polymerase chain reaction.

**[331]** In some embodiments, an inhibitor oligonucleotide comprises a binding domain for a single-stranded overhang of a detection probe (*e.g.*, as described or utilized herein), wherein the binding domain is or comprises a nucleotide sequence that is substantially complementary to the single-stranded overhang of the detection probe such that a free, unbound detection probe having a complementary single-stranded overhang can bind to the binding domain of the inhibitor oligonucleotide. In some embodiments, an inhibitor oligonucleotide may have a hairpin at one end. In some embodiments, an inhibitor oligonucleotide may be a single-stranded oligonucleotide comprising at one end a binding domain for a single-stranded overhang of a detection probe, wherein a portion of the single-stranded oligonucleotide can self-hybridize to form a hairpin at another end.

**[332]** In some embodiments, a target entity detection system provided herein (*e.g.*, a duplex, triplex or multiplex target entity detection system described herein) does not comprise a connector oligonucleotide that associates an oligonucleotide domain of a detection probe with an oligonucleotide domain of another detection probe. In some embodiments, a connector oligonucleotide is designed to bridge oligonucleotide domains of any two detection probes that would not otherwise interact with each other when they bind to an entity of interest. In some embodiments, a connector oligonucleotide is designed to hybridize with at least a portion of an oligonucleotide domain of a detection probe and at least a portion of an oligonucleotide domain of another detection probe. A connector oligonucleotide can be single-stranded, double-stranded, or a combination thereof. A connector oligonucleotide is free of any target-binding moiety (*e.g.*, as described and/or utilized herein) or control binding moiety. In at least some embodiments, no connector oligonucleotides are necessary to indirectly connect oligonucleotide domains of detection probes; in some embodiments, such connector oligonucleotides are not utilized, in part because detection probes as provided and/or utilized herein are designed such that their respective oligonucleotide domains have a sufficient length to reach and interact with each other when they are in sufficiently close proximity, *e.g.*, when the detection probes simultaneously bind to an entity of interest (*e.g.*, a biological entity such as an extracellular vesicle).

*Methods of using provided target entity detection systems*

**[333]** Provided target entity detection systems are useful in detecting an entity of interest (*e.g.*, a biological entity such as extracellular vesicles) in a sample (*e.g.*, in a biological, environmental, or other sample) for various applications and/or purposes associated with detection of lung cancer. Accordingly, some aspects provided herein relate to methods of using a plurality of (*e.g.*, at least 2, at least 3, or more) detection probes appropriate for use in accordance with the present disclosure. In some embodiments, a method comprises contacting an entity of interest (*e.g.*, a biological entity such as extracellular vesicles) in a sample (*e.g.*, a biological sample such as a blood-derived sample from a human subject) with a set of detection probes comprising at least 2 or more (including, *e.g.*, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20 or more) detection probes as described and/or utilized herein. In some embodiments, a method comprises subjecting a sample

comprising an entity of interest (*e.g.*, a biological entity such as extracellular vesicles) to a target entity detection system (*e.g.*, as provided herein). A plurality of detection probes (*e.g.*, at least two or more) can be added to a sample comprising an entity of interest (*e.g.*, a biological entity such as extracellular vesicles) at the same time or at different times (*e.g.*, sequentially). In some embodiments, a method may comprise, prior to contacting with a plurality of detection probes, contacting a sample comprising an entity of interest with at least one capture agent directed to an extracellular vesicle-associated surface biomarker.

**[334]** In certain embodiments, a provided target entity detection system for use in a method described herein may comprise a plurality of (*e.g.*, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20 or more) distinct sets (*e.g.*, combinations) of detection probes (*e.g.*, as described herein). In some embodiments, a method comprises contacting an entity of interest (*e.g.*, a biological entity such as extracellular vesicles) in a sample (*e.g.*, a biological sample such as a blood-derived sample from a human subject) with a plurality of sets of detection probes, wherein each set may comprise at least 2 or more (including, *e.g.*, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20 or more) detection probes as described and/or utilized herein. In some embodiments, a method comprises subjecting a sample comprising an entity of interest (*e.g.*, a biological entity such as extracellular vesicles) to a target entity detection system (*e.g.*, as provided herein). A plurality of detection probes and/or detection probe combinations (*e.g.*, at least two or more) can be added to a sample comprising an entity of interest (*e.g.*, a biological entity such as extracellular vesicles) at the same time or at different times (*e.g.*, sequentially). In some embodiments, a method may comprise, prior to contacting with a plurality of detection probes, contacting a sample comprising an entity of interest with at least one capture agent directed to an extracellular vesicle-associated surface biomarker.

**[335]** In some embodiments, the relationship between results (*e.g.*, Ct values and/or relative number of ligated nucleic acid templates (*e.g.*, ligated DNA templates)) from profiling one or more biomarker combinations in a sample can be combined with clinical information (including, *e.g.*, but not limited to patient age, past medical history, X-ray results, smoking history, *etc.*) and/or other information to better classify patients with or at risk for lung cancer.

Various classification algorithms can be used to interpret the relationship between multiple variables to increase an assay's sensitivity and/or specificity. In some embodiments, such algorithms include, but are not limited to, logistic regression models, support vector machines, gradient boosting machines, random forest algorithms, Naive Bayes algorithms, K-nearest neighborhood algorithms, and combinations thereof. In some embodiments, performance (*e.g.*, accuracy) of assays described herein can be improved, *e.g.*, by selection of biomarker combinations (*e.g.*, as described herein), selection of other factors or variables (*e.g.*, clinical information and/or life-style information, *e.g.,* smoking history) to include an algorithm, and/or selection of the type of algorithm itself.

**[336]** In certain embodiments, technologies described herein utilize a predictive algorithm that is trained and validated using data sets as described herein. In certain embodiments, technologies described herein are utilized to generate a risk score using an algorithm created from training samples which is designed to take into account results from at least two, *e.g.*, at least two, at least 3, at least 4, at least 5, or more than 5 separate assays comprising biomarker signatures (*e.g.*, as described herein). In certain embodiments, an algorithm-generated risk score can be generated at least in part using diagnostic data (*e.g.*, raw and/or normalized Ct values) from at least one individual assay (*e.g.*, individual biomarker signature). In certain embodiments, a reference threshold can be included within a risk score. In certain embodiments, multiple threshold levels denoting multiple different degrees of lung cancer risk may be included in a risk score. In some embodiments, separate target biomarker signature assays may be performed as individual assays in a series of assays, and individual assays may be weighted equally or differently in a predictive algorithm. In some embodiments, for example, weighting of individual assays combined in an algorithm (*e.g.*, a cohort of biomarker assays) may be determined by a number of factors including but not limited to the sensitivity of an individual assay, the specificity of an individual assay, the reproducibility of an individual assay, the variability of an individual assay, the positive predictive value of an individual assay, and/or the lowest limit of detection of a specific assay. In some embodiments, a cohort of biomarker assays may be ranked according to a characteristic (*e.g.*, sensitivity, specificity, lowest limit of detection *etc.*) and the biomarker assays may then be weighted based upon their relative rank.

**[337]** In some embodiments, a risk score generated by an algorithm (as described herein) can be presented in a suitable manner, *e.g.*, on a nominal scale, *e.g.*, on a scale of 0-100 reflecting a number of likelihoods, *e.g.*, including but not limited to the likelihood a subject has lung cancer, the likelihood a subject will develop lung cancer, and/or the likely stage of lung cancer. In some embodiments, a higher risk score can demonstrate that there is an increasing likelihood of disease pathology, *e.g.*, lower to higher values may reflect healthy controls, benign controls, stage I, stage II, stage III, and stage IV lung cancers. In some embodiments, a risk score can be utilized to reduce the potential of cross reactivity of technologies as described herein when compared with other cancer types.

**[338]** In some embodiments, a risk score may be generated from a combination of data derived from assays as described herein coupled with other applicable diagnostic data such as age, life history, X-ray results, MRI results, low-dose CT scanning, or any combination thereof. In some embodiments, a risk score provides predictive value above and beyond that of conventional standard of care diagnostic assay predictive values, *e.g.*, higher than predictive values provided low-dose CT scanning assays utilized in isolation or in combination with another diagnostic assay. In some embodiments, a risk score may be generated that has high specificity for lung cancers (*e.g.*, lung adenocarcinoma and/or lung squamous cell carcinoma) and has low sensitivity for other cancers.

**[339]** In some embodiments, a risk score may have an associated clinical cutoff for detection of lung cancer. For example, in some embodiments, a risk score's clinical cutoff for detection may require an assay that yields at least 40%, *e.g.*, at least 50%, at least 60%, or greater sensitivity for detection of both early and late stage lung cancer and has a minimum of 90% specificity, *e.g.*, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or greater specificity in a generally healthy population of subjects (e.g., aged 20 to 89 years of age) or in a population of subjects with hereditary risk. In some embodiments, sensitivity and specificity targets are the approximate lower bounds of the two-sided 95% confidence interval for the targeted 77% sensitivity and 99.5% specificity.

**[340]** In some embodiments, a training study is performed to provide the necessary data required to program a risk score algorithm. In some embodiments, such a training study may comprise a cohort of samples from a range of suppliers, including at least commercial suppliers, biobanks, purpose driven studies, and/or physicians. In some embodiments, a training study may

comprise positive samples from lung adenocarcinoma and/or lung squamous cell carcinoma cancer patients (*e.g.*, stage I, stage II, stage III, and/or stage IV), positive control samples from certain lung cancer cell lines, negative samples from benign lung tumor patients, negative samples from non-lung cancer patients (*e.g.*, brain cancer, breast cancer, colorectal cancer, endometrial cancer, ovarian cancer, melanoma, non-Hodgkin's lymphoma, pancreatic cancer, skin cancer, *etc.*), negative samples from inflammatory condition patients (*e.g.*, Crohn's disease, endometriosis, diabetes type II, lupus, pancreatitis, rheumatoid arthritis, ulcerative colitis, chronic obstructive pulmonary disease, *etc.*), negative samples from healthy patients, or any combination thereof. In some embodiments, a training study may comprise samples from patients of any appropriate age range, *e.g.*, <31 years old, 31-40 years old, 41-50 years old, 51-60 years old, 61-70 years old, 71-80 years old, or >80 years old. In some embodiments, a training study may comprise samples from patients of any race/ethnicity/descent, (*e.g.*, Caucasians, Africans, Asians *etc.*).

**[341]** In some embodiments, a validation study is performed to provide the necessary data required to confirm a risk score algorithm's utility. In some embodiments, such a validation study may comprise a cohort of samples from a range of suppliers, including at least commercial suppliers, biobanks, purpose driven studies, and/or physicians. In some embodiments, a validation study may comprise positive samples from lung adenocarcinoma and/or lung squamous cell carcinoma cancer patients (*e.g.*, stage I, stage II, stage III, and/or stage IV), positive control samples from certain lung cancer cell lines, negative samples from benign lung tumor patients, negative samples from non-lung cancer patients (*e.g.*, brain cancer, breast cancer, colorectal cancer, endometrial cancer, ovarian cancer, melanoma, non-Hodgkin's lymphoma, pancreatic cancer, skin cancer, *etc.*), negative samples from inflammatory condition patients (*e.g.*, Crohn's disease, endometriosis, diabetes type II, lupus, pancreatitis, rheumatoid arthritis, ulcerative colitis, chronic obstructive pulmonary disease, *etc.*), negative samples from healthy patients, or any combination thereof. In some embodiments, a validation study may comprise samples from patients of any appropriate age range, *e.g.*, <31 years old, 31-40 years old, 41-50 years old, 51-60 years old, 61-70 years old, 71-80 years old, or >80 years old. In some embodiments, a validation study may comprise samples from patients of any race/ethnicity/descent, (*e.g.*, Caucasians, Africans, Asians *etc.*).

**[342]** In certain embodiments, at least one target biomarker signature comprising at least one surface biomarker (e.g., extracellular vesicle-associated surface biomarker) and at least one (including, *e.g.*, at least two, or more) target biomarker (which may be selected from any of surface biomarkers described herein, intravesicular biomarkers described herein, and/or intravesicular RNA biomarkers described herein) may be embodied in a lung cancer detection assay. In some such embodiments, at least one capture agent is directed to the extracellular vesicle-associated surface biomarker, and at least one set of detection probes is directed to one or more of such target biomarkers described herein. For example, **FIGS. 5-7** disclose certain examples of target biomarker signatures, each of which may be embodied in a lung cancer detection assay (*e.g.*, ones described herein).

**[343]** In certain embodiments, at least two (including, *e.g.*, at least three or more) distinct target biomarker signatures each comprising at least one extracellular vesicle-associated surface biomarker and at least one (including, *e.g.*, at least two, or more) target biomarker (which may be selected from any of surface biomarkers described herein, intravesicular biomarkers described herein, and/or intravesicular RNA biomarkers described herein) may be embodied in a lung cancer detection assay. In some embodiments, each distinct target biomarker signature may have a different pre-determined cutoff value for individually determining whether a sample is positive for lung cancer. In some embodiments, a sample is determined to be positive for lung cancer if assay readout is above at least one of cutoff values for a plurality of (*e.g.*, at least 2 or more) target biomarker signatures. In some embodiments, a combination of cutoff values (*e.g.*, at least 2, at least 3, or more) can be utilized to create a diagnostic value with corollarily improved sensitivity and/or specificity.

**[344]** Accordingly, in some embodiments, a sample can be divided into aliquots such that a different capture agent and/or a different set of detection probes (*e.g.*, each directed to detection of a distinct disease or condition) can be added to a different aliquot. In such embodiments, provided technologies can be implemented with one aliquot at a time or multiple aliquots at a time (*e.g.*, for parallel assays to increase throughput).

**[345]** In some embodiments, amount of detection probes that is added to a sample provides a sufficiently low concentration of detection probes in a mixture to ensure that the detection probes will not randomly come into close proximity with one another in the absence of binding to an entity of interest (*e.g.*, biological entity), at least not to any great or substantial

degree. As such, in many embodiments, when detection probes simultaneously bind to the same entity of interest (*e.g.*, biological entity) through the binding interaction between respective targeting binding moieties of the detection probes and the binding sites of an entity of interest (*e.g.*, a biological entity), the detection probes come into sufficiently close proximity to one another to form double-stranded complex (*e.g.*, as described herein). In some embodiments, the concentration of detection probes in a mixture following combination with a sample may range from about 1 fM to 1 μM, such as from about 1pM to about 1 nM, including from about 1 pM to about 100 nM.

**[346]** In some embodiments, the concentration of an entity of interest (*e.g.*, a biological entity) in a sample is sufficiently low such that a detection probe binding to one entity of interest (*e.g.*, a biological entity) will not randomly come into close proximity with another detection probe binding to another entity of interest (*e.g.*, biological entity) in the absence of respective detection probes binding to the same entity of interest (*e.g.*, biological entity), at least not to any great or substantial degree. By way of example only, the concentration of an entity of interest (*e.g.*, biological entity) in a sample is sufficiently low such that a first target detection probe binding to a non-target entity of interest (*e.g.*, a non-cancerous biological entity such as an extracellular vesicle comprising a first target) will not randomly come into close proximity with another different target detection probe that is bound to another non-target entity of interest (*e.g.*, a non-cancerous biological entity such as an extracellular vesicle), at least not to any great or substantial degree, to generate a false positive detectable signal.

**[347]** Following contacting an entity of interest (*e.g.*, biological entity) in a sample with a set of detection probes, such a mixture may be incubated for a period of time sufficient for the detection probes to bind corresponding targets (*e.g.*, molecular targets), if present, in the entity of interest to form a double-stranded complex (*e.g.*, as described herein). In some embodiments, such a mixture is incubated for a period of time ranging from about 5 min to about 5 hours, including from about 30 min to about 2 hours, at a temperature ranging from about 10 to about 50 °C, including from about 20 °C to about 37 °C.

**[348]** A double-stranded complex (resulted from contacting an entity of interest such as a biological entity with detection probes) can then be subsequently contacted with a nucleic acid ligase to perform nucleic acid ligation of a free 3′ end hydroxyl and 5′ end phosphate end of oligonucleotide strands of detection probes, thereby generating a ligated template comprising

oligonucleotide strands of at least two or more detection probes. In some embodiments, prior to contacting an assay sample comprising a double-stranded complex with a nucleic acid ligase, at least one or more inhibitor oligonucleotide (*e.g.*, as described herein) can be added to the assay sample such that the inhibitor oligonucleotide can capture any residual free-floating detection probes that may otherwise interact with each other during a ligation reaction.

**[349]** As is known in the art, ligases catalyze the formation of a phosphodiester bond between juxtaposed 3'-hydroxyl and 5'-phosphate termini of two immediately adjacent nucleic acids when they are annealed or hybridized to a third nucleic acid sequence to which they are complementary. Any known nucleic acid ligase (*e.g.*, DNA ligases) may be employed, including but not limited to temperature sensitive and/or thermostable ligases. Non-limiting examples of temperature sensitive ligases include bacteriophage T4 DNA ligase, bacteriophage T7 ligase, and E. coli ligase. Non-limiting examples of thermostable ligases include Taq ligase, Tth ligase, and Pfu ligase. Thermostable ligase may be obtained from thermophilic or hyperthermophilic organisms, including but not limited to, prokaryotic, eukaryotic, or archael organisms. In some embodiments, a nucleic acid ligase is a DNA ligase. In some embodiments, a nucleic acid ligase can be a RNA ligase.

**[350]** In some embodiments, in a ligation step, a suitable nucleic acid ligase (*e.g.*, a DNA ligase) and any reagents that are necessary and/or desirable are combined with the reaction mixture and maintained under conditions sufficient for ligation of the hybridized ligation oligonucleotides to occur. Ligation reaction conditions are well known to those of skill in the art. During ligation, a reaction mixture, in some embodiments, may be maintained at a temperature ranging from about 20° C to about 45° C, such as from about 25° C to about 37° C for a period of time ranging from about 5 minutes to about 16 hours, such as from about 1 hour to about 4 hours. In yet other embodiments, a reaction mixture may be maintained at a temperature ranging from about 35° C to about 45° C, such as from about 37° C to about 42° C, *e.g.*, at or about 38 ° C, 39° C, 40° C or 41° C, for a period of time ranging from about 5 minutes to about 16 hours, such as from about 1 hour to about 10 hours, including from about 2 to about 8 hours.

**[351]** Detection of such a ligated template can provide information as to whether an entity of interest (*e.g.*, a biological entity) in a sample is positive or negative for targets to which detection probes are directed. For example, a detectable level of such a ligated template is indicative of a tested entity of interest (*e.g.*, a biological entity) comprising targets (*e.g.*,

molecular targets) of interest. In some embodiments, a detectable level is a level that is above a reference level, *e.g.*, by at least 10% or more, including, *e.g.*, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or more. In some embodiments, a reference level may be a level observed in a negative control sample, such as a sample in which an entity of interest comprising such targets is absent. Conversely, a non-detectable level (*e.g.*, a level that is below the threshold of a detectable level) of such a ligated template indicates that at least one of targets (*e.g.*, molecular targets) of interest is absent from a tested entity of interest (*e.g.*, a biological entity). Those of skill in the art will appreciate that a threshold that separates a detectable level from a non-detectable level may be determined based on, for example, a desired sensitivity level, and/or a desired specificity level that is deemed to be optimal for each application and/or purpose. For example, in some embodiments, a specificity of 99.7% may be achieved using a system provided herein, for example by setting a threshold that is three standard deviations above a reference level (*e.g.*, a level observed in a negative control sample, such as, *e.g.*, a sample derived from one or more normal healthy individuals). Additionally or alternatively, those of skill in the art will appreciate that a threshold of a detectable level (*e.g.*, as reflected by a detection signal intensity) may be 1 to 100-fold above a reference level.

**[352]** In some embodiments, a method provided herein comprises, following ligation, detecting a ligated template, *e.g.*, as a measure of the presence and/or amount of an entity of interest in a sample. In various embodiments, detection of a ligated template may be qualitative or quantitative. As such, in some embodiments where detection is qualitative, a method provides a reading or evaluation, *e.g.*, assessment, of whether or not an entity of interest (*e.g.*, a biological entity) comprising at least two or more targets (*e.g.*, molecular targets) is present in a sample being assayed. In other embodiments, a method provides a quantitative detection of whether an entity of interest (*e.g.*, a biological entity) comprising at least two or more targets (*e.g.*, molecular targets) is present in a sample being assayed, *e.g.*, an evaluation or assessment of the actual amount of an entity of interest (*e.g.*, a biological entity) comprising at least two or more targets (*e.g.*, molecular targets) in a sample being assayed. In some embodiments, such quantitative detection may be absolute or relative.

**[353]** A ligated template formed by using technologies provided herein may be detected by an appropriate method known in the art. Those of skill in the art will appreciate that

appropriate detection methods may be selected based on, for example, a desired sensitivity level and/or an application in which a method is being practiced. In some embodiments, a ligated template can be directly detected without any amplification, while in other embodiments, ligated template may be amplified such that the copy number of the ligated template is increased, *e.g.*, to enhance sensitivity of a particular assay. Where detection without amplification is practicable, a ligated template may be detected in a number of different ways. For example, oligonucleotide domains of detection probes (*e.g.*, as described and/or utilized herein) may have been directly labeled, *e.g.*, fluorescently or radioisotopically labeled, such that a ligated template is directly labeled. For example, in some embodiments, an oligonucleotide domain of a detection probe (*e.g.*, as provided and/or utilized herein) can comprise a detectable label. A detectable label may be a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Such labels include biotin for staining with labeled Streptavidin conjugate, magnetic beads (*e.g.*, Dynabeads®), fluorescent dyes (*e.g.*, fluorescein, Texas red, rhodamine, green fluorescent protein, and the like), radiolabels (*e.g.*, $^{3}H$, $^{125}I$, $^{34}S$, $^{14}C$, or $^{32}P$), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and calorimetric labels such as colloidal gold or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, *etc.*) beads. In some embodiments, a directly labeled ligated template may be size separated from the remainder of the reaction mixture, including unligated directly labeled ligation oligonucleotides, in order to detect the ligated template.

**[354]** In some embodiments, detection of a ligated template can include an amplification step, where the copy number of ligated nucleic acids is increased, *e.g.*, in order to enhance sensitivity of the assay. The amplification may be linear or exponential, as desired, where amplification can include, but is not limited to polymerase chain reaction (PCR); quantitative PCR, isothermal amplification, NASBA, digital droplet PCR, *etc.*

**[355]** Various technologies for achieving PCR amplification are known in the art; those skilled in the art will be well familiar with a variety of embodiments of PCR technologies, and will readily be able to select those suitable to amplify a ligated template generated using technologies provided herein. For example, in some embodiments, a reaction mixture that includes a ligated template is combined with one or more primers that are employed in the primer extension reaction, *e.g.*, PCR primers (such as forward and reverse primers employed in geometric (or exponential) amplification or a single primer employed in a linear amplification).

Oligonucleotide primers with which one or more ligated templates are contacted should be of sufficient length to provide for hybridization to complementary template DNA under appropriate annealing conditions. Primers are typically at least 10 bp in length, including, *e.g.*, at least 15 bp in length, at least 20 bp in length, at least 25 bp in length, at least 30 bp in length or longer. In some embodiments, the length of primers can typically range from about 15 to 50 bp in length, from about 18 to 30 bp, or about 20 to 35 bp in length. Ligated templates may be contacted with a single primer or a set of two primers (forward and reverse primers), depending on whether primer extension, linear, or exponential amplification of the template DNA is desired.

**[356]** In addition to the above components, a reaction mixture comprising a ligated template typically includes a polymerase and deoxyribonucleoside triphosphates (dNTPs). The desired polymerase activity may be provided by one or more distinct polymerase enzymes. In preparing a reaction mixture, *e.g.*, for amplification of a ligated template, various constituent components may be combined in any convenient order. For example, an appropriate buffer may be combined with one or more primers, one or more polymerases and a ligated template to be detected, or all of the various constituent components may be combined at the same time to produce the reaction mixture.

***VI. Uses***

**[357]** In some embodiments, one or more provided biomarkers of one or more target biomarker signatures for lung cancer can be detected in a sample comprising biological entities (including, *e.g.*, cells, circulating tumor cells, cell-free DNA, nanoparticles, *etc.*), for example, using methods of detecting and/or assays as described herein. In some embodiments, one or more provided biomarkers of one or more target biomarker signatures for lung cancer can be detected in a sample comprising nanoparticles, for example, using methods of detecting and/or assays as described herein.

**[358]** In some embodiments, a sample may be or comprise a biological sample. In some embodiments, a biological sample is a bodily fluid-derived sample of a subject (e.g., a human subject). In some embodiments, a biological sample can be derived from a blood or blood-derived sample of a subject (*e.g.*, a human subject) in need of such an assay. In some embodiments, a biological sample can be or comprise a primary sample (*e.g.*, a tissue or tumor sample) from a subject (*e.g.*, a human subject) in need of such an assay. In some embodiments, a

biological sample can be processed to separate one or more entities of interest (*e.g.*, biological entity) from non-target entities of interest, and/or to enrich one or more entities of interest (*e.g.*, biological entity). In some embodiments, an entity of interest present in a sample may be or comprise a biological entity, *e.g.*, a cell or nanoparticles having a size range that includes extracellular vesicles (*e.g.*, an exosome). In some embodiments, such a biological entity (*e.g.*, extracellular vesicle) may be processed or contacted with a chemical reagent, *e.g.*, to stabilize and/or crosslink targets (*e.g.*, provided target biomarkers) to be assayed in the biological entity and/or to reduce non-specific binding with detection probes. In some embodiments, a biological entity is or comprises a cell, which may be optionally processed, *e.g.*, with a chemical reagent for stabilizing and/or crosslinking targets (*e.g.*, molecular targets) and/or for reducing non-specific binding. In some embodiments, a biological entity is or comprises an extracellular vesicle (*e.g.*, an exosome), which may be optionally processed, *e.g.*, with a chemical reagent for stabilizing and/or crosslinking targets (*e.g.*, molecular targets) and/or for reducing non-specific binding.

**[359]** In some embodiments, technologies provided herein can be useful for managing patient care, *e.g.,* for one or more individual subjects and/or across a population of subjects. By way of example only, in some embodiments, provided technologies may be utilized in screening, which for example, may be performed periodically, such as annually, semi-annually, bi-annually, or with some other frequency as deemed to be appropriate by those skilled in the art. In some embodiments, such a screening may be temporally motivated or incidentally motivated. For example, in some embodiments, provided technologies may be utilized in temporally-motivated screening for one or more individual subjects or across a population of subjects (*e.g.*, asymptomatic subjects) who are older than a certain age (*e.g.*, over 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or older). As will be appreciated by those skilled in the art, in some embodiments, the screening age and/or frequency may be determined based on, for example, but not limited to prevalence of a disease, disorder, or condition (*e.g.*, cancer such as lung cancer). In some embodiments, provided technologies may be utilized in incidentally-motivated screening for individual subjects who may have experienced an incident or event that motivates screening for a particular disease, disorder, or condition (*e.g.*, cancer such as lung cancer). For example, in some embodiments, an incidental motivation relating to determination of one or more indicators of a disease, disorder, or condition (*e.g.*, cancer such as lung cancer) or susceptibility thereto may be or comprise, *e.g.*, an incident based on their family history (*e.g.*, a close relative such as

blood-related relative was previously diagnosed for such a disease, disorder, or condition such as lung cancer), identification of one or more life-history associated risk factors for a disease, disorder, or condition (*e.g.*, lung cancer) and/or prior incidental findings from genetic tests (*e.g.*, genome sequencing), and/or imaging diagnostic tests (*e.g.*, chest X-ray, or low-dose CT scanning), development of one or more signs or symptoms characteristic of a particular disease, disorder, or condition (*e.g.,* chronic obstructive pulmonary disease, and/or symptoms such as bloody sputum, persistent cough, shortness of breath, repeated and/or chronic respiratory infection, thoracic pain, unexplained weight loss, and/or fatigue), subjects having benign lung tumors (*e.g.*, a benign mass observed in the lung through imaging such as chest X-ray or low-dose CT scan), and combinations thereof.) and/or other incidents or events as will be appreciated by those skilled in the art.

**[360]** In some embodiments, provided technologies for managing patient care can inform treatment and/or payment (*e.g.,* reimbursement for treatment) decisions and/or actions. For example, in some embodiments, provided technologies can provide determination of whether individual subjects have one or more indicators of risk, incidence, or recurrence of a disease disorder, or condition (*e.g.*, cancer such as lung cancer), thereby informing physicians and/or patients when to provide/receive therapeutic or prophylactic recommendations and/or to initiate such therapy in light of such findings. In some embodiments, such individual subjects may be asymptomatic subjects, who may be temporally-motivated or incidentally-motivated to be screened at a regular frequency (*e.g.*, annually, semi-annually, bi-annually, or other frequency as deemed to be appropriate by those skilled in the art). In some embodiments, such individual subjects may be experiencing one or more symptoms that may be associated with lung cancer, who may be temporally-motivated or incidentally-motivated to be screened at a regular frequency (*e.g.*, annually, semi-annually, bi-annually, or other frequency as deemed to be appropriate by those skilled in the art). In some embodiments, such individual subjects may be subjects having a benign lung tumor and/or a chronic inflammatory condition, who may be temporally-motivated or incidentally-motivated to be screened at a regular frequency (*e.g.*, annually, semi-annually, bi-annually, or other frequency as deemed to be appropriate by those skilled in the art). In some embodiments, such individual subjects may be subjects at hereditary risk for lung cancer, who may be temporally-motivated or incidentally-motivated to be screened at a regular frequency (*e.g.*, annually, semi-annually, bi-annually, or other frequency as deemed

to be appropriate by those skilled in the art). In some embodiments, such individual subjects may be subjects with life-history associated risk, who may be temporally-motivated or incidentally-motivated to be screened at a regular frequency (*e.g.*, annually, semi-annually, bi-annually, or other frequency as deemed to be appropriate by those skilled in the art). In some embodiments, such individual subjects may be smoking subjects (*e.g.*, heavy smokers), who may be temporally-motivated or incidentally-motivated to be screened at a regular frequency (*e.g.*, annually, semi-annually, bi-annually, or other frequency as deemed to be appropriate by those skilled in the art).

**[361]** Additionally or alternatively, in some embodiments, provided technologies can inform physicians and/or patients of treatment selection, *e.g.*, based on findings of specific responsiveness biomarkers (*e.g.*, cancer responsiveness biomarkers). In some embodiments, provided technologies can provide determination of whether individual subjects are responsive to current treatment, *e.g.*, based on findings of changes in one or more levels of molecular targets associated with a disease, thereby informing physicians and/or patients of efficacy of such therapy and/or decisions to maintain or alter therapy in light of such findings. In some embodiments, provided technologies can provide determination of whether individual subjects are likely to be responsive to a recommended treatment, *e.g.*, based on findings of molecular targets (*e.g.*, provided biomarkers of one or more target biomarker signatures for lung cancer) that predict therapeutic effects of a recommended treatment on individual subjects, thereby informing physicians and/or patients of potential efficacy of such therapy and/or decisions to administer or alter therapy in light of such findings.

**[362]** In some embodiments, provided technologies can inform decision making relating to whether health insurance providers reimburse (or not), *e.g.*, for (1) screening itself (*e.g.*, reimbursement available only for periodic/regular screening or available only for temporally- and/or incidentally- motivated screening); and/or for (2) initiating, maintaining, and/or altering therapy in light of findings by provided technologies. For example, in some embodiments, the present disclosure provides methods relating to (a) receiving results of a screening that employs provided technologies and also receiving a request for reimbursement of the screening and/or of a particular therapeutic regimen; (b) approving reimbursement of the screening if it was performed on a subject according to an appropriate schedule (based on, *e.g.*, screening age such as older than a certain age, *e.g.*, over 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or older, and/or

screening frequency such as, *e.g.*, every 3 months, every 6 months, every year, every 2 years, every 3 years or at some other frequencies) or in response to a relevant incident and/or approving reimbursement of the therapeutic regimen if it represents appropriate treatment in light of the received screening results; and, optionally (c) implementing the reimbursement or providing notification that reimbursement is refused. In some embodiments, a therapeutic regimen is appropriate in light of received screening results if the received screening results detect a biomarker that represents an approved biomarker for the relevant therapeutic regimen (*e.g.,* as may be noted in a prescribing information label and/or via an approved companion diagnostic).

**[363]** Alternatively or additionally, the present disclosure contemplates reporting systems (*e.g.,* implemented via appropriate electronic device(s) and/or communications system(s)) that permit or facilitate reporting and/or processing of screening results (*e.g.*, as generated in accordance with the present disclosure), and/or of reimbursement decisions as described herein. Various reporting systems are known in the art; those skilled in the art will be well familiar with a variety of such embodiments, and will readily be able to select those suitable for implementation.

*Exemplary uses*

*A. Detection of lung cancer incidence or recurrence*

**[364]** The present disclosure, among other things, recognizes that detection of a single cancer-associated biomarker in a biological entity (*e.g.*, nanoparticles, including, e.g., extracellular vesicles) or a plurality of cancer-associated biomarkers based on a bulk sample, rather than at a resolution of a single biological entity (*e.g.*, individual extracellular vesicles), typically does not provide sufficient specificity and/or sensitivity in determination of whether a subject from whom the biological entity is obtained is likely to be suffering from or susceptible to cancer (*e.g.*, lung cancer). The present disclosure, among other things, provides technologies, including compositions and/or methods, that solve such problems, including for example by specifically requiring that an entity (*e.g.*, an extracellular vesicle) for detection be characterized by presence of a combination of at least two or more targets (*e.g.,* at least two or more provided biomarkers of a target biomarker signature for lung cancer). In particular embodiments, the present disclosure teaches technologies that require such an entity (*e.g.*, an extracellular vesicle) be characterized by presence (*e.g.,* by expression) of a combination of molecular targets that is

specific to cancer (*i.e.,* "target biomarker signature" of a relevant cancer, *e.g.*, lung cancer), while biological entities (*e.g.*, nanoparticles, including, e.g., extracellular vesicles) that do not comprise the targeted combination (*e.g.*, target biomarker signature) do not produce a detectable signal. Accordingly, in some embodiments, technologies provided herein can be useful for detection of risk, incidence, and/or recurrence of cancer in a subject. In some such embodiments, technologies provided herein are useful for detection of risk, incidence, and/or recurrence of lung cancer in a subject. For example, in some embodiments, a combination of two or more provided biomarkers are selected for detection of a specific cancer (*e.g.*, lung cancer) or various cancers (one of which includes lung cancer). In some embodiments, a specific combination of provided biomarkers for detection of lung cancer can be determined by analyzing a population or library (*e.g.* tens, hundreds, thousands, tens of thousands, hundreds of thousands, or more) of lung cancer patient biopsies and/or patient data to identify such a predictive combination. In some embodiments, a relevant combination of biomarkers may be one identified and/or characterized, for example, via data analysis. For example, in some embodiments, data analysis may comprise a bioinformatic analysis, for example, as described in Examples 7-9. In some embodiments, for example, a diverse set of lung cancer-associated data (*e.g.*, in some embodiments comprising one or more of bulk RNA sequencing, single-cell RNA (scRNA) sequencing, mass spectrometry, histology, post-translational modification data, *in vitro* and/or *in vivo* experimental data) can be analyzed through machine learning and/or computational modeling to identify a combination of predictive markers that is highly specific to lung cancer. In some embodiments, a combination of predictive markers to distinguish stages of cancer (*e.g.*, lung cancer) can be determined *in silico* based on comparing and analyzing diverse data (*e.g.*, in some embodiments comprising bulk RNA sequencing, scRNA sequencing, mass spectrometry, histology, post-translational modification data, *in vitro* and/or *in vivo* experimental data) relating to different stages of cancer (*e.g.*, lung cancer). For example, in some embodiments, technologies provided herein can be used to distinguish lung cancer subjects from non-lung cancer subjects, including, *e.g.*, healthy subjects, subjects diagnosed with benign tumors or thoracic masses, and subjects with non-lung-related diseases, disorders, and/or conditions (*e.g.*, subjects with non-lung cancer, or subjects with inflammatory conditions, *e.g.*, chronic obstructive pulmonary disease or chronic lung infections). In some embodiments, technologies provided herein can be useful for early detection of lung cancer, *e.g.*, detection of lung cancer of stage I or stage II. In some

embodiments, technologies provided herein can be useful for detection of one or more lung cancer subtypes, including, *e.g.*, lung adenocarcinoma, small cell lung cancer, squamous and transitional cell lung cancer, large cell lung cancer, non-small cell carcinoma, other specified lung carcinoma, lung sarcoma, and other specified types of lung cancer as known in the art (SEER Cancer Statistics Review 1975-2017). In some embodiments, technologies provided herein can be useful for screening individuals at hereditary risk, life-history associated risk, or average risk for early stage lung cancer (*e.g.*, lung adenocarcinoma and/or lung squamous cell carcinoma).

**[365]** In some embodiments, technologies provided herein can be useful for screening a subject for risk, incidence, or recurrence of a specific cancer in a single assay. For example, in some embodiments, technologies provided herein is useful for screening a subject for risk, incidence, or recurrence of lung cancer. In some embodiments, technologies provided herein can be used to screen a subject for risk or incidence of a specific cancer or a plurality of (*e.g.*, at least 2, at least 3, or more) cancers in a single assay. For example, in some embodiments, technologies provided herein can be used to screen a subject for a plurality of cancers in a single assay, one of which includes lung cancer and other cancers to be screened can be selected from the group consisting of brain cancer (including, *e.g.*, glioblastoma), breast cancer, colorectal cancer, pancreatic cancer, prostate cancer, liver cancer, ovarian cancer, and skin cancer.

**[366]** In some embodiments, provided technologies can be used periodically (*e.g.*, every year, every two years, every three years, *etc.*) to screen a human subject (*e.g.*, a human subject) for lung cancer (*e.g.*, early-stage lung cancer) or cancer recurrence. In some embodiments, a human subject amenable to such screening may be an adult or an elderly. In some embodiments, a human subject amenable to such screening may be older than a specified age, *e.g.,* age 45 and above, age 55 and above, age 65 and above, age 70 and above, age 75 and above, or age 80 and above. In some embodiments, a human subject amenable to such screening may have an age of about 50 or above. In some embodiments, a human subject amenable to such screening may have an age of 50 or less. In some embodiments, a human subject amenable to such screening may have an age over 35. In some embodiments, a human subject who is determined to have a genetic predisposition to lung cancer may be screened at a younger age than a human subject who has no family history risk.

**[367]** In some embodiments, a subject that is amenable to provided technologies for detection of incidence or recurrence of lung cancer may be a subject with a smoking history (*e.g.*, a heavy smoker), who in some embodiments may be experiencing one or more symptoms associated with lung cancer. In some embodiments, a subject that is amenable to provided technologies for detection of incidence or recurrence of lung cancer may be a human subject who is determined to have a benign lung tumor and/or one or more chronic inflammatory conditions (*e.g.*, COPD). In some embodiments, a subject that is amenable to provided technologies for detection of incidence or recurrence of lung cancer may be a subject, who has a family history of lung cancer (*e.g.*, subjects having one or more first-degree relatives with a history of lung cancer), who has been previously treated for cancer (*e.g.*, lung cancer), who is at risk of lung cancer recurrence after cancer treatment, who is in remission after lung cancer treatment, and/or who has been previously or periodically screened for lung cancer, *e.g.*, by screening for the presence of at least one lung cancer biomarker (*e.g.*, as described herein).

**[368]** In some embodiments, the present disclosure, among other things, provides insights that technologies described and/or utilized herein may be particularly useful for screening certain populations of subjects, e.g., subjects who are at higher susceptibility to developing lung cancer. In some embodiments, the present disclosure, among other things, recognizes that the resulting PPVs of technologies described and/or utilized herein for lung detection may be higher in lung cancer prone or susceptible populations. In some embodiments, the present disclosure, among other things, provides insights that screening of smoking individuals, *e.g.*, regular screening prior to or otherwise in absence of developed symptom(s), can be beneficial, and even important for effective management (*e.g.*, successful treatment) of lung cancer. In some embodiments, the present disclosure provides lung cancer screening systems that can be implemented to detect lung cancer, including early-stage cancer, in some embodiments in smoking individuals (*e.g.*, with or without hereditary and/or life-history risks in lung cancer and/or with or without symptoms associated with lung cancer). In some embodiments, provided technologies can be implemented to achieve regular screening of smoking individuals (*e.g.*, with or without hereditary and/or life-history risks in lung cancer and/or with or without symptoms associated with lung cancer). In some embodiments, provided technologies achieve detection (*e.g.*, early detection, *e.g.*, in symptomatic or asymptomatic individual(s) and/or population(s)) of one or more features (*e.g.*, incidence, progression,

responsiveness to therapy, recurrence, *etc.*) of lung cancer, with sensitivity and/or specificity (*e.g.*, rate of false positive and/or false negative results) appropriate to permit useful application of provided technologies to single-time and/or regular (*e.g.*, periodic) assessment. In some embodiments, provided technologies are useful in conjunction with a subject's periodic physical examination (*e.g.* every year, every other year, or at an interval approved by the attending physician). In some embodiments, provided technologies are useful in conjunction with treatment regimen(s); in some embodiments, provided technologies may improve one or more characteristics (*e.g.*, rate of success according to an accepted parameter) of such treatment regimen(s).

**[369]** In some embodiments, a subject that is amenable to provided technologies for detection of incidence or recurrence of lung cancer may be an asymptomatic human subject and/or across an asymptomatic population of subjects. Such an asymptomatic subject and/or across an asymptomatic population of subjects may be subject(s) who has/have a family history of cancers such as breast and/or ovarian cancer, leukemia, and/or lung cancer (*e.g.*, individuals having one or more first-degree relatives with a history of cancers known to be associated with genetic risk factors), who has been previously treated for cancer (*e.g.*, lung cancer), who is at risk of lung cancer recurrence after cancer treatment, who is in remission after lung cancer treatment, and/or who has been previously or periodically screened for lung cancer, *e.g.*, by screening for the presence of at least one lung cancer biomarker or through thoracic imaging (*e.g.,* X-ray imaging, sputum testing, low-dose CT scanning, and/ or molecular tests based on cell-free nucleic acids, serum proteins *e.g.*, CEA, CYFRA 21-1, NSE, ProGRP, and/or SCCA), Alternatively, in some embodiments, an asymptomatic subject may be a subject who has not been previously screened for lung cancer, who has not been diagnosed for lung cancer, and/or who has not previously received lung cancer therapy. In some embodiments, an asymptomatic subject may be a subject with a benign lung tumor. In some embodiments, an asymptomatic subject may be a subject who is susceptible to lung cancer (*e.g.*, at an average population risk, at an elevated life-history associated risk, or with hereditary risk for lung cancer).

**[370]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be selected based on one or more characteristics such as age, race, geographic location, genetic history, medical history, personal history (*e.g.*, smoking, alcohol, drugs, carcinogenic agents, diet, obesity, physical activity, sun

exposure, radiation exposure, exposure to infectious agents such as viruses, and/or occupational hazard). For example, in some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects determined to currently be or have been a smoker (*e.g.* cigarettes, cigars, pipe, and/or hookah).

**[371]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects determined to have worked in conditions known to expose the subject(s) to inhalable carcinogens, including but not limited to: copper ore smelting, lead ore smelting, zinc ore smelting, manufacture of insecticides, arsenic mining, asbestos mining, asbestos textile production, brake lining work, cement production, construction work, insulation work, shipyard work, ceramic manufacture, electronic and aerospace equipment manufacture, chemical manufacturing, chromate production, chromium electroplating, leather tanning, pigment production, nickel mining, nickel refining, nickel electroplating, production of stainless and heat-resistant steel, polycyclic aromatic production, aluminum production, hydrocarbon compound production, coke production, ferrochromium alloy production, nickel-containing ore smelting, roofing, radon mining, ceramics and glass production, granite working, metal ore smelting, silica mining and quarrying stone.

**[372]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects determined to have one or more germline mutations in lung cancer-associated genes (*e.g.,* genes associated with DNA repair pathways such as ATM or BRCA, and/or germline mutations in the potential oncogene epidermal growth factor receptor (EGFR)), and combinations thereof.

**[373]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects diagnosed with an imaging-confirmed thoracic mass or pulmonary mass.

**[374]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects at hereditary risk or life-history associated risk before undergoing a biopsy and/or a surgical procedure (e.g., lobectomy).

**[375]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or population of subjects determined to have COPD or pulmonary fibrosis. In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or population of subjects with a history of chronic bronchitis, tuberculosis, and/or pneumonia. In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or population of subjects determined to have HIV and/or AIDS. In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or population of subjects with high current or historical alcohol consumption. In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or population of subjects determined to have hereditary mutations in EGFR, cytochrome p450 enzymes, and/or DNA repair genes. In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or population of subjects exposed to radiation therapy and/or chemotherapy.

**[376]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects with one or more non-specific symptoms of lung cancer. In some embodiments, exemplary non-specific symptoms of lung cancer may include symptoms similar to those of chronic obstructive pulmonary disease, and/or symptoms such as bloody sputum, persistent cough, shortness of breath, repeated and/or chronic respiratory infection, thoracic pain, unexplained weight loss, hemoptysis, airway obstruction, and/or fatigue.

**[377]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects of diverse descent such as Asians, African Americans, Caucasians, Native Hawaiians or other Pacific Islanders, Hispanics or Latinos, American Indians or Alaska natives, non-Hispanic blacks, or non-Hispanic whites. In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects of diverse descent such as Asian Pacific Islanders, Hispanics, American Indian/Alaska natives, non-Hispanic black, or non-Hispanic white. In some embodiments, a

subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be a subject or a population of subjects of any race and/or any ethnicity.

**[378]** In some embodiments, a subject or population of subjects that are amenable to provided technologies for detection of lung cancer may be determined to have normal X-ray imaging, sputum testing, low-dose CT scanning, and/or molecular tests based on cell-free nucleic acids, serum proteins (*e.g.*, CEA, CYFRA 21-1, NSE, ProGRP, and/or SCCA). In some embodiments, such subjects may have received a negative indication of lung cancer from such diagnostic tests. In some embodiments, such subjects may have received a positive indication of lung cancer from such diagnostic tests. In some embodiments, technologies provided herein can be used in combination with other diagnostics assays including *e.g.,* but not limited to: (i) physicals, general practitioner visits, cholesterol/lipid blood tests, diabetes (type 2) screening, colonoscopies, blood pressure screening, thyroid function tests, prostate cancer screening, mammograms, HPV/Pap smears, and/or vaccinations; (ii) chest X-ray imaging, sputum testing, chest low-dose CT scanning, and/ or molecular tests based on cell-free nucleic acids, serum proteins (*e.g.*, CEA, CYFRA 21-1, NSE, ProGRP, and/or SCCA); (iii) a genetic assay to screen blood plasma for genetic mutations in circulating tumor DNA and/or protein biomarkers linked to cancer; (iv) an assay involving immunofluorescence staining to identify cell phenotype and marker expression, followed by amplification and analysis by next-generation sequencing; and/or (v) EGFR, KRAS, ALK, ROAS1, HER2, BRAF, and RET germline, somatic, and circulating cell mutation assays, or assays involving cell-free tumor DNA, liquid biopsy, serum protein and cell-free DNA, and/or circulating tumor cells.

*B. Selection of cancer therapy (e.g., lung cancer therapy)*

**[379]** In some embodiments, provided technologies can be used for selecting an appropriate treatment for a cancer patient (*e.g.*, a patient suffering from or susceptible to lung cancer). For example, some embodiments provided herein relate to a companion diagnostic assay for classification of patients for cancer therapy (*e.g.*, lung cancer and/or adjunct treatment) which comprises assessment in a patient sample (*e.g.*, a biological sample such as a blood-derived sample from a lung cancer patient) of a selected combination of provided biomarkers using technologies provided herein. Based on such an assay outcome, patients who are determined to be more likely to respond to a cancer therapy (*e.g.*, a lung cancer therapy and/or an adjunct

therapy, including, *e.g.*, Abraxane, Afatinib Dimaleate, Alectinib, Atezolizumab, Bevacizumab, Brigatinib, Capmatinib Hydrochloride, Carboplatin, Ceritinib, Crizotinib, Dabrafenib Mesylate, Dacomitinib, Docetaxel, Doxorubicin Hydrochloride, Durvalumab, Entrectinib, Erlotinib Hydrochloride, Everolimus, Gefitinib, Gemcitabine Hydrochloride, Ipilimumab, Lorlatinib, Mechlorethamine Hydrochloride, Methotrexate, Necitumumab, Nivolumab, Osimertinib Mesylate, Paclitaxel, Paclitaxel Albumin-stabilized Nanoparticle Formulation, Pembrolizumab, Pemetrexed Disodium, Ramucirumab, Selpercatinib, Trametinib, and/or Vinorelbine Tartrate) can be administered such a therapy, or patients who are determined to be non-responsive to a specific such therapy can be administered a different therapy.

*C. Evaluation of treatment efficacy (e.g., cancer treatment efficacy)*

**[380]** In some embodiments, technologies provided herein can be used for monitoring and/or evaluating efficacy of an anti-cancer therapy administered to a cancer patient (*e.g.*, lung cancer patient). For example, a biological sample (e.g., a bodily fluid sample including, but not limited to a blood-derived sample, etc.) can be collected from an lung cancer patient prior to or receiving an anti-cancer therapy (*e.g.*, Abraxane, Afatinib Dimaleate, Alectinib, Atezolizumab, Bevacizumab, Brigatinib, Capmatinib Hydrochloride, Carboplatin, Ceritinib, Crizotinib, Dabrafenib Mesylate, Dacomitinib, Docetaxel, Doxorubicin Hydrochloride, Durvalumab, Entrectinib, Erlotinib Hydrochloride, Everolimus, Gefitinib, Gemcitabine Hydrochloride, Ipilimumab, Lorlatinib, Mechlorethamine Hydrochloride, Methotrexate, Necitumumab, Nivolumab, Osimertinib Mesylate, Paclitaxel, Paclitaxel Albumin-stabilized Nanoparticle Formulation, Pembrolizumab, Pemetrexed Disodium, Ramucirumab, Selpercatinib, Trametinib, and/or Vinorelbine Tartrate.) at a first time point to detect or measure tumor burdens, *e.g.*, by detecting presence or amount of nanoparticles comprising a selected combination of biomarkers that is specific to detection of lung cancer. After a period of treatment, a second biological sample (e.g., a bodily fluid sample including, but not limited to a blood-derived sample, etc.) can be collected from the same lung cancer patient to detect changes in tumor burdens, *e.g.*, by detecting absence or reduction in amount of nanoparticles comprising a selected combination of biomarkers that is specific to detection of lung cancer. By monitoring levels and/or changes in tumor burdens over the course of treatment, appropriate course of action, *e.g.*, increasing or

decreasing the dose of a therapeutic agent, and/or administering a different therapeutic agent, can be taken.

***VII. Kits***

**[381]** Also provided are kits that find use in practicing technologies as described above. In some embodiments, a kit comprises a plurality of detection probes (*e.g.*, as described and/or utilized herein). In some embodiments, a provided kit may comprise two or more (*e.g.*, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more) detection probes. In some embodiments, individual detection probes may be directed at different targets. In some embodiments, two or more individual detection probes may be directed to the same target. In some embodiments, a provided kit comprises two or more different detection probes directed at different targets, and optionally may include at least one additional detection probe also directed at a target to which another detection probe is directed. In some embodiments, a provided kit comprises a plurality of subsets of detection probes, each of which comprises two or more detection probes directed at the same target. In some embodiments, a plurality of detection probes may be provided as a mixture in a container. In some embodiments, multiple subsets of detection probes may be provided as individual mixtures in separate containers. In some embodiments, each detection probe is provided individually in a separate container.

**[382]** In some embodiments, a kit for detection of lung cancer comprises: (a) a capture agent comprising a target-capture moiety directed to an extracellular vesicle-associated surface biomarker; and (b) a set of detection probes, which set comprises at least two detection probes each directed to a target biomarker of a target biomarker signature for lung cancer, wherein the detection probes each comprise:(i) a target binding moiety directed the target biomarker of the target biomarker signature for lung cancer; and (ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same extracellular vesicle. In these embodiments, such a target biomarker signature for lung cancer comprises:

at least one extracellular vesicle-associated surface biomarker and at least one target biomarker selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers, wherein:

- the surface biomarkers are selected from: ADGRF1, ALCAM, ABCC3, ARSL, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, GPC4, HAS3, HS6ST2, KDELR3, KRTCAP3, IG1FR, LAMB3, LAPTM4B, LARGE2, LFNG, LSR, MAL2, MANEAL, MET, MSLN, MUC1, MUC21, NRCAM, PIGT, PODXL2, PRRG4, PRSS21, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, SLC44A4, SLC6A14, SLC7A7, SLC7A11, SMIM22, SMPDL3B, ST14, TACSTD2, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, UCHL1 and combinations thereof; in some embodiments, a surface biomarker described herein may comprise at least one modification (e.g., in some embodiments glycosylation, and/or in some embodiments lipidation);
- the intravesicular biomarkers are selected from: AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; in some embodiments, an intravesicular biomarker described herein may comprise at least one post-translational modification;
- the intravesicular RNA biomarkers are selected from ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof.

**[383]** In some embodiments, a kit for detection of lung cancer comprises: (a) a capture agent comprising a target-capture moiety directed to an extracellular vesicle -associated surface biomarker; and (b) a set of detection probes, which set comprises at least two detection probes each directed to a target biomarker of a target biomarker signature for lung cancer, wherein the detection probes each comprise:(i) a target binding moiety directed the target biomarker of the target biomarker signature for lung cancer; and (ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same extracellular vesicle. In these embodiments, such a target biomarker signature for lung cancer comprises at least one extracellular vesicle-associated surface biomarker (*e.g.*, as described herein) and at least one target biomarker selected from the group consisting of: surface biomarkers (*e.g.*, as described herein), intravesicular biomarkers (*e.g.*, as described herein), and intravesicular RNA biomarkers (*e.g.*, as described herein). In some embodiments, one or more surface biomarkers utilized in a provided kit are selected from: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN,

PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, Lewis X antigen, Lewis Y/B antigen, LY6E, NOTCH2, NOTCH3, Phosphatidylserine, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA, Tn antigen, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof. In some embodiments, one or more intravesicular biomarkers utilized in a provided kit are polypeptides encoded by human genes as follows: ABRACL, ACP5, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, AOC1, AP1M2, APOBEC3B, APOBEC3C, ARNTL2, ASF1B, AURKB, BAIAP2L1, BIRC5, C12orf45, C15orf48, C19orf33, C1S, C8orf4, CA9, CALML3, CAPNS2, CBLC, CCL19, CCNB2, CDC20, CDC45, CDCA4, CDCA5, CDK1, CDKN2A, CDKN2B, CENPW, CEP55, CES1, CHMP4C, CNN2, CPA3, CRABP2, CST1, CSTA, CTSC, CTSE, CYP2S1, DPYSL3, EFS, EGLN3, EHF, ELF3, ELF4, ENAH, ESRP1, ETV4, EVPL, FAM129B, FAM60A, FAM83A, FAM83D, FAM83H, FBP1, FERMT1, FOXA2, FOXE1, FOXM1, GBP6, GNA15, GPX2, GRHL2, GSTA1, HCK, HMGB3, HOXB7, ID1, IGF2BP2, IMPA2, IRF6, IVL, JUP, KIAA1522, KIF2C, KIFC1, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, LGALS3BP, LGALS7B, LSP1, MAGEA4, MAGEA6, MCM2, MDFI, MIF, MYBL2, MYH14, MZB1, NAPSA, NCF2, NNMT, NRARP, NUP210, NUSAP1, OSGIN1, PALLD, PITX1, PKP1, PKP3, PLEK, PLEK2, POSTN, PPP1R14C, PPP1R14D, PRAME, PTPN6, RBP1, RIN2, RIPK4, RPS4Y1, RRM2, S100A11, S100A14, S100A16, S100A2, S100P, SBK1, SCGB3A2, SERPINB13, SERPINB3, SERPINB5, SFTA2, SFTPA1, SFTPA2, SFTPB, SH3BP4, SNAI2, SOX2, SPI1, SPINK1, SPINT1, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, SULF1, SYK, SYTL1, TBC1D2, TEAD2, TEAD3, TFAP2C, TGFA, THBS2, TK1, TOP2A, TP63, TPD52, TPX2, TRIM29, TRIP13, UBE2C, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof. In some embodiments, an intravesicular biomarker described herein may comprise at least one post-

translational modification. In some embodiments, one or more intravesicular RNA biomarkers utilized in a provided kit are RNA transcripts (e.g., mRNA transcripts) encoded by human genes as follows: ABCA3, ABCC1, ABCC3, ABRACL, ACP5, ADAM23, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, ANTXR1, AOC1, AP1M2, APOBEC3B, APOBEC3C, AQP3, AREG, ARNTL2, ARSL, ASF1B, ATP8B1, AURKB, B3GNT3, B3GNT5, BAIAP2L1, BCAM, BIK, BIRC5, C12orf45, C15orf48, C19orf33, C1S, C8orf4, CA12, CA9, CALML3, CAPNS2, CBLC, CCL19, CCL5, CCNB2, CD109, CD24, CD53, CD74, CD9, CDC20, CDC42EP1, CDC45, CDCA4, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CEACAM5, CEACAM6, CELSR1, CENPW, CEP55, CES1, CHMP4C, CLCA2, CLDN1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CNN2, COL17A1, CPA3, CRABP2, CST1, CSTA, CTSC, CTSE, CX3CL1, CXADR, CXCR4, CYBB, CYP2S1, CYP4F11, DAPL1, DMBT1, DPYSL3, DSC2, DSC3, DSG2, DSG3, DSP, EFNA1, EFS, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHB3, EPHX1, EPHX3, ESRP1, ETV4, EVA1A, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83A, FAM83D, FAM83H, FAT1, FAT2, FBLIM1, FBP1, FCER1G, FERMT1, FGFR2, FGFR3, FOLR1, FOXA2, FOXE1, FOXM1, FXYD3, GALNT3, GBP6, GJA1, GJB1, GJB2, GJB3, GJB5, GJB6, GNA15, GPC1, GPC3, GPC4, GPNMB, GPR87, GPRC5A, GPX2, GRHL2, GSTA1, HAS3, HCK, HMGB3, HOXB7, HS6ST2, ID1, IGF2BP2, IGSF9, IL2RG, IMPA2, IRF6, ITGA2, ITGA6, ITGB4, ITGB6, IVL, JAG2, JUP, KCNS3, KDELR3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, KRTCAP3, LAMB3, LAMP3, LAPTM5, LFNG, LGALS3BP, LGALS7B, LRP11, LRRC4, LSP1, LSR, LYPD3, MAGEA4, MAGEA6, MAL2, MANEAL, MAOA, MARCO, MCM2, MDFI, MET, MIF, MMP14, MPZL2, MSLN, MUC1, MUC21, MYBL2, MYH14, MYOF, MZB1, NAPSA, NCF2, NKG7, NNMT, NOTCH3, NRARP, NTRK2, NUP210, NUSAP1, OSGIN1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PERP, PIGR, PIGT, PITX1, PKP1, PKP3, PLEK, PLEK2, PLVAP, PMP22, PODXL2, POSTN, PPL, PPP1R14C, PPP1R14D, PRAME, PROM2, PRRG4, PRSS8, PTGES, PTGFRN, PTPN6, PTPRF, PTPRZ1, RAB25, RAB38, RAET1L, RARRES1, RBP1, RGS1, RHCG, RHOV, RIN2, RIPK4, ROS1, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, S100A2, S100P, SBK1, SCGB3A2, SCNN1A, SDC1, SERINC2, SERPINB13, SERPINB3, SERPINB5, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB,

SH3BP4, SHISA2, SLC1A5, SLC2A1, SLC34A2, SLC40A1, SLC44A4, SLC6A14, SLC6A8, SLC7A7, SLC7A8, SMIM22, SMPDL3B, SNAI2, SOX2, SPI1, SPINK1, SPINT1, SPINT2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, TGFA, THBD, THBS2, TK1, TM4SF1, TMC4, TMC5, TMEM30B, TMEM45B, TMEM54, TMPRSS11D, TMPRSS11E, TMPRSS2, TMPRSS4, TNFRSF18, TNS4, TOP2A, TP53I11, TP63, TPD52, TPX2, TREM2, TRIM29, TRIP13, TSPAN1, TSPAN13, TSPAN6, TSPAN7, TSPAN8, TUSC3, TYROBP, UBE2C, UPK1B, VAMP8, VANGL2, WLS, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof.

**[384]** In some embodiments, when at least one target biomarker is selected from one or more of the provided surface biomarkers, the selected surface biomarker(s) and the at least one extracellular vesicle-associated surface biomarker are different. In some embodiments, when at least one target biomarker is selected from one or more of the provided surface biomarkers, the selected surface biomarker(s) and the at least one extracellular vesicle-associated surface biomarker are the same (with the same or different epitopes).

**[385]** In some embodiments, a capture agent provided in a kit comprises a target-capture moiety directed to an extracellular vesicle-associated surface biomarker, which is or comprises one or more of: *PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1,*

*RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, NOTCH2, NOTCH3, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA,* and combinations thereof; and/or (ii) carbohydrate-dependent markers as follows: Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, Lewis X, Lewis Y/B antigen, LY6E, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof. In some embodiments, a capture agent provided in a kit comprises a target-capture moiety directed to an extracellular vesicle-associated surface biomarker, which is or comprises one or more of: ALCAM, B3GNT3, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CLDN3, CLDN4, DSG2, EGFR, EPCAM, FOLR1, GJB1, GJB2, IG1FR, LAMB3, MET, MSLN, MUC1, PIGT, PODXL2, ROS1, SDC1, SLC34A2, SMPDL3B, ST14, sTn antigen, Tn antigen, T antigen, TACSTD2, TMPRSS4, TSPAN8, TNFRSF10B, and combinations thereof. In some embodiments, a capture agent provided in a kit comprises a target-capture moiety directed to an extracellular vesicle-associated surface biomarker, which is or comprises one or more polypeptides selected from SLC34A2, CEACAM5, CEACAM6, EpCAM, and/or combinations thereof.

**[386]** In some embodiments, a target binding moiety of at least two detection probes provided in a kit is each directed to the same target biomarker of a target biomarker signature. In some embodiments, such the same target biomarker is or comprises CEACAM6. In some such embodiments, an oligonucleotide domain of such at least two detection probes are different.

**[387]** In some embodiments, a target binding moiety of at least two detection probes provided in a kit is each directed to a distinct target biomarker of a target biomarker signature. For example, in some embodiments, a kit comprises at least two detection probes directed to CEACAM6 and EpCAM, respectively. In some embodiments, a kit comprises at least two detection probes directed to CEACAM6 and SLC34A2, respectively.

**[388]** In some embodiments, a target binding moiety of a detection probe may be or comprise an affinity agent, which in some embodiments may be or comprise an antibody (*e.g.*, a monoclonal antibody). In some embodiments, a target binding moiety of a detection probe may be or comprise an affinity agent, which in some embodiments may be or comprise a lectin or siglec.

**[389]** In some embodiments, a kit is directed to detection of a target biomarker signature for lung cancer as described herein (including, *e.g.*, but not limited to ones depicted in Tables 4-5). In some embodiments, a kit is directed to detection of a target biomarker signature for lung cancer comprising at least a surface biomarker as described herein (including, *e.g.*, but not limited to ones provided in Table 3, as an extracellular vesicle-associated surface biomarker) and at least a surface biomarker as described herein (including, *e.g.*, but not limited to ones provided in Table 3, as a target surface biomarker).

**[390]** In some embodiments, a kit is directed to a target biomarker signature for lung cancer comprising at least a SLC34A2 polypeptide (as an extracellular vesicle-associated surface biomarker) and a CEACAM6 polypeptide (as a target surface biomarker).

**[391]** In some embodiments, a kit is directed to a target biomarker signature for lung cancer comprising at least a SLC34A2 polypeptide (as an extracellular vesicle-associated surface biomarker) and an EpCAM polypeptide (as a target surface biomarker).

**[392]** In some embodiments, a kit is directed to a target biomarker signature for lung cancer comprising at least a CEACAM5 polypeptide (as an extracellular vesicle-associated surface biomarker) and a CEACAM6 polypeptide (as a target surface biomarker).

**[393]** In some embodiments, a kit is directed to a target biomarker signature for lung cancer comprising at least a CEACAM5 polypeptide (as an extracellular vesicle-associated surface biomarker) and a SLC34A2 polypeptide (as a target surface biomarker).

**[394]** In some embodiments, a kit is directed to a target biomarker signature for lung cancer comprising at least a CEACAM5 polypeptide (as an extracellular vesicle-associated surface biomarker); and a CEACAM6 polypeptide and a SLC34A2 polypeptide (as two distinct target surface biomarkers).

**[395]** In some embodiments, a kit is directed to a target biomarker signature for lung cancer comprising at least a SLC34A2 polypeptide (as an extracellular vesicle-associated surface

biomarker); and a CEACAM6 polypeptide and an EpCAM polypeptide (as two distinct target surface biomarkers).

**[396]** In some embodiments, a kit may comprise at least one enzymatic or chemical reagent such as an enzyme, fixation agent, a permeabilization agent, and/or a blocking agent.

**[397]** In some embodiments, a kit may comprise one or more nucleic acid ligation reagents (*e.g.*, a nucleic acid ligase such as a DNA ligase and/or a buffer solution).

**[398]** In some embodiments, a kit may comprise at least one or more amplification reagents such as PCR amplification reagents. In some embodiments, a kit may comprise one or more nucleic acid polymerases (*e.g.*, DNA polymerases), one or more pairs of primers, nucleotides, and/or a buffered solution.

**[399]** In some embodiments, a kit may comprise a solid substrate for capturing an entity (*e.g.*, biological entity) of interest. For example, such a solid substrate may be or comprise a bead (*e.g.*, a magnetic bead). In some embodiments, such a solid substrate may be or comprise a surface. In some embodiments, a surface may be or comprise a capture surface (*e.g.*, an entity capture surface) of an assay chamber, such as, *e.g.*, a filter, a matrix, a membrane, a plate, a tube, a well (*e.g.*, but not limited to a microwell), *etc*. In some embodiments, a surface (*e.g.*, a capture surface) of a solid substrate can be coated with a capture agent (*e.g.*, affinity agent) for an entity (*e.g.*, biological entity) of interest.

**[400]** In some embodiments, a set of detection probes provided in a kit may be selected for diagnosis of lung cancer.

**[401]** In some embodiments, a kit may comprise a plurality of sets of detection probes, wherein each set of detection probes is directed for detection of a specific cancer and comprises at least 2 or more detection probes. For example, such a kit can be used to screen a subject for various cancers, one of which is lung cancer while other cancers may be selected from skin cancer, ovarian cancer, breast cancer, colorectal cancer, pancreatic cancer, prostate cancer, brain cancer, and liver cancer in a single assay.

**[402]** In some embodiments, kits provided herein may include instructions for practicing methods described herein. These instructions may be present in kits in a variety of forms, one or more of which may be present in the kits. One form in which these instructions may be present is as printed information on a suitable medium or substrate, *e.g.*, a piece or pieces of paper on which the information is printed, in the packaging of kits, in a package insert, *etc*.

Yet another means may be a computer readable medium, *e.g.*, diskette, CD, USB drive, *etc.*, on which instructional information has been recorded. Yet another means that may be present is a website address which may be used via the internet to access instructional information. Any convenient means may be present in the kits.

**[403]** In some embodiments where kits are for use as companion diagnostics, such kits can include instructions for identifying patients that are likely to respond to a therapeutic agent (*e.g.*, identification of biomarkers that are indicative of patient responsiveness to the therapeutic agent). In some embodiments, such kits can comprise a therapeutic agent for use in tandem with the companion diagnostic test.

**[404]** Exemplary embodiments below are also within the scope of the present disclosure:

**[405]** One aspect of the disclosure herein is a method comprising steps of:

(a) providing or obtaining a blood-derived sample from a subject;

(b) detecting, in the blood-derived sample, extracellular vesicles expressing a first target biomarker signature ("first target biomarker signature-expressing extracellular vesicles"), the first target biomarker signature comprising:

at least one extracellular vesicle-associated surface biomarker and

at least one target biomarker selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers, wherein:

the surface biomarkers are selected from ADGRF1, ABCC3, ALCAM, ARSL, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, GPC4, HAS3, HS6ST2, IG1FR, KDELR3, KRTCAP3, LAMB3, LAPTM4B, LARGE2, LFNG, LSR, MAL2, MANEAL, MET, MSLN, MUC1, MUC21, NRCAM, PIGT, PODXL2, PRRG4, PRSS21, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SLC7A11, SMIM22, SMPDL3B, ST14, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, TACSTD2, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, UCHL1 and combinations thereof;

the intravesicular biomarkers are selected from AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and

the intravesicular RNA biomarkers are selected from RNA transcripts (e.g., mRNA transcripts) encoded by human genes as follows: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof;

(c) comparing sample information indicative of level of the first target biomarker signature-expressing extracellular vesicles in the blood-derived sample to reference information including a first reference threshold level;

(d) classifying the subject as having or being susceptible to lung cancer when the blood-derived sample shows an elevated level of first target biomarker signature-expressing extracellular vesicles relative to a classification cutoff referencing the first reference threshold level.

**[406]** In some embodiments of the disclosed method, when the at least one target biomarker is selected form one or more of the surface biomarkers, the selected surface biomarker(s) and the at least one extracellular vesicle-associated surface biomarker are different.

**[407]** In some embodiments of the disclosed method, the steps of (b) and (c) are repeated for at least a second target biomarker signature, and wherein the classification cutoff references the first reference threshold level and at least a second reference threshold level corresponding to the at least a second target biomarker signature.

**[408]** In some embodiments of the disclosed method, the extracellular vesicle-associated surface biomarker is or comprises ADGRF1, ALCAM, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CLDN3, CLDN4, DSG2,

EGFR, EPCAM, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, HAS3, IG1FR, LAMB3, LAPTM4B, LARGE2, MAL2, MET, MSLN, MUC1, NRCAM, PIGT, PODXL2, PRSS21, ROS1, SDC1, SLC34A2, SLC7A11, SMIM22, SMPDL3B, ST14, UCHL1, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, TACSTD2, TMPRSS4, TSPAN8, TNFRSF10B, or combinations thereof.

**[409]** In some embodiments of the disclosed method, the first and/or second target biomarker signature comprises at least one extracellular vesicle-associated surface biomarker and at least two target biomarkers selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers.

**[410]** In some embodiments of the disclosed method, the at least two target biomarkers comprise one of the following combinations:

- at least two distinct surface biomarkers;
- at least two distinct intravesicular biomarkers;
- at least two distinct intravesicular RNA biomarkers;
- a surface biomarker and an intravesicular biomarker;
- a surface biomarker and an intravesicular RNA biomarker; and
- an intravesicular biomarker and an intravesicular RNA biomarker.

**[411]** One aspect of the disclosure herein is a method comprising steps of:

(a) providing or obtaining a blood-derived sample from a subject;

(b) detecting, in the blood-derived sample, extracellular vesicles expressing a first target biomarker signature ("first target biomarker signature-expressing extracellular vesicles"), the first target biomarker signature comprising:

at least one extracellular vesicle-associated surface biomarker and

at least one target biomarker, wherein the target biomarker is or comprises a surface biomarker, wherein the target biomarker signature comprises a biomarker combination as listed in Table 4;

(c) comparing sample information indicative of level of the first target biomarker signature-expressing extracellular vesicles in the blood-derived sample to reference information including a first reference threshold level;

(d) classifying the subject as having or being susceptible to lung cancer when the blood-derived sample shows an elevated level of first target biomarker signature-expressing extracellular vesicles relative to a classification cutoff referencing the first reference threshold level.

**[412]** In some embodiments of the disclosed method, the extracellular vesicle-associated surface biomarker is or comprises a CEACAM5 polypeptide and/or a SLC34A2 polypeptide.

**[413]** In some embodiments of the disclosed method, the first and/or second target biomarker signature comprises (i) at least one extracellular vesicle-associated surface biomarker, which is or comprises a CEACAM5 polypeptide and/or a SLC34A2 polypeptide; and (ii) at least one target biomarker CEACAM6, which may be a surface biomarker or an intravesicular RNA biomarker.

**[414]** In some embodiments of the disclosed method, the first and/or second target biomarker signature comprises (i) at least one extracellular vesicle-associated surface biomarker, which is or comprises a CEACAM5 polypeptide and/or a SLC34A2 polypeptide; and (ii) at least two target biomarkers CEACAM6 and EPCAM, each of which may be a surface biomarker or an intravesicular RNA biomarker.

**[415]** In some embodiments of the disclosed method, the first and/or second target biomarker signature comprises (i) at least one extracellular vesicle-associated surface biomarker, which is or comprises a CEACAM5 polypeptide; and (ii) at least two target biomarkers CEACAM6 and SLC34A2, each of which may be a surface biomarker or an intravesicular RNA biomarker.

**[416]** In some embodiments of the disclosed method, the first or second reference threshold level is determined by levels of target biomarker signature-expressing extracellular vesicles observed in comparable samples from a population of non-cancer subjects.

**[417]** In some embodiments of the disclosed method, the population of non-cancer subjects comprises one or more of the following subject populations: healthy subjects, subjects diagnosed with benign tumors, and subjects with non-lung-related diseases, disorders, and/or conditions.

**[418]** In some embodiments of the disclosed method, the blood-derived sample has been subjected to size exclusion chromatography to isolate (*e.g.*, directly from the blood-derived sample) nanoparticles having a size range of interest that includes extracellular vesicles.

**[419]** In some embodiments of the disclosed method, the step of detecting comprises a capture assay.

**[420]** In some embodiments of the disclosed method, the capture assay involves contacting the blood-derived sample with a capture agent comprising a target-capture moiety that binds to the at least one extracellular vesicle-associated surface biomarker.

**[421]** In some embodiments of the disclosed method, the capture agent is or comprises a solid substrate comprising the target-capture moiety conjugated thereto.

**[422]** In some embodiments of the disclosed method, the solid substrate comprises a magnetic bead.

**[423]** In some embodiments of the disclosed method, the target-capture moiety is or comprises an antibody agent.

**[424]** In some embodiments of the disclosed method, the step of detecting comprises a detection assay.

**[425]** In some embodiments of the disclosed method, the step of detecting comprises a capture assay and a detection assay, the capture assay being performed prior to the detection assay.

**[426]** In some embodiments of the disclosed method, when the first and/or second target biomarker signature comprises at least one intravesicular RNA biomarkers, the detection assay involves reverse transcription qPCR.

**[427]** In some embodiments of the disclosed method, when the first and/or second target biomarker signature comprises at least one intravesicular biomarker, the target biomarker signature-expressing extracellular vesicles are processed involving fixation and/or permeabilization prior to the detection assay.

**[428]** In some embodiments of the disclosed method, when the first and/or second target biomarker signature comprises at least one surface biomarker and/or intravesicular biomarker, the detection assay involves an immunoassay (including, *e.g.*, immuno-PCR, and/or proximity ligation assay).

**[429]** In some embodiments of the disclosed method, the detection assay involves a proximity ligation assay.

**[430]** In some embodiments of the disclosed method, the proximity ligation assay comprises the steps of:

(a) contacting the target biomarker signature-expressing extracellular vesicles that express the at least one extracellular vesicle-associated surface biomarker ("extracellular vesicle-associated surface biomarker-expressing extracellular vesicles") with a set of detection probes, each directed to a target biomarker of the target biomarker signature, which set comprises at least two detection probes, so that a combination comprising the extracellular vesicles and the set of detection probes is generated,

wherein the detection probes each comprise:

(i) a target binding moiety directed to the target biomarker of the target biomarker signature; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the detection probes are characterized in that they can hybridize to each other when the detection probes are bound to the same extracellular vesicle,

(b) maintaining the combination under conditions that permit binding of the set of detection probes to their respective targets on the extracellular vesicles such that the at least two detection probes can bind to the same extracellular vesicle that express the target biomarker signature to form a double-stranded complex;

(c) contacting the double-stranded complex with a nucleic acid ligase to generate a ligated template; and

(d) detecting the ligated template, wherein presence of the ligated template is indicative of presence in the blood-derived sample of the target biomarker signature-expressing extracellular vesicles; and

(e) optionally repeating steps a through d at least one additional time using an orthogonal target biomarker signature.

**[431]** In some embodiments of the disclosed method, the target binding moiety of the at least two detection probes is directed to the same target biomarker (*e.g.*, CEACAM6).

**[432]** In some embodiments of the disclosed method, the oligonucleotide domain of the at least two detection probes are different.

**[433]** In some embodiments of the disclosed method, the target-capture moiety of the capture assay is or comprises at least one antibody agent directed to the at least one extracellular vesicle-associated surface biomarker (*e.g.*, a CEACAM5 polypeptide, an EpCAM polypeptide, and/or a SLC34A2 polypeptide).

**[434]** In some embodiments of the disclosed method, the method is performed to screen for early-stage lung cancer, late-stage lung cancer, or recurrent lung cancer in the subject.

**[435]** In some embodiments of the disclosed method, the subject has at least one or more of the following characteristics:

(i) an asymptomatic subject who is susceptible to lung cancer (*e.g.*, at an average population risk (*i.e.*, without hereditary risk) or with hereditary risk for lung cancer);

(ii) a subject with a family history of lung cancer (*e.g.*, a subject having one or more first-degree relatives with a history of lung cancer);

(iii) a subject who is or was a smoker;

(iv) a subject who has exposure to secondhand smoke, radon gas, asbestos, bituminous "smoky coal", and/or other carcinogens (*e.g.*, arsenic, chromium, nickel, ionizing radiation, polycyclic aromatic hydrocarbons, nitric oxide, high levels of particulate matter < 2.5μm);

(v) a subject aged 40 or over;

(vi) a subject with one or more non-specific symptoms of lung cancer, optionally wherein at least one of the non-specific symptoms is similar to one or more common respiratory symptoms such as coughing, hemoptysis, airway obstruction, and shortness of breath, associated with a non-cancer disease, disorder, or condition;

(vii) a subject recommended for chest imaging such as X-ray, CT scan, or low-dose CT scan;

(viii) a subject diagnosed with an imaging-confirmed lung mass;

(ix) a subject with a benign lung tumor;

(x) a subject who has been previously treated for lung cancer;

(xi) a subject determined to have COPD;

(xii) a subject determined to have pulmonary fibrosis;

(xiii) a subject with a history of chronic bronchitis, tuberculosis, and/or pneumonia;

(xiv) a subject determined to have HIV and/or AIDS;

(xv) a subject with high current or historical alcohol consumption;

(xvi) a subject with hereditary mutations in EGFR, cytochrome p450 enzymes, and/or DNA repair genes; and

(xvii) a subject exposed to radiation therapy and/or chemotherapy.

**[436]** In some embodiments of the disclosed method, the method is used in combination with one or more of the following diagnostic assays:

(i) the subject's annual physical examination;

(ii) a chest imaging (*e.g.*, X-ray, CT scan, or low-dose CT scan);

(iii) sputum cytology;

(iv) a genetic assay to screen blood plasma for genetic mutations in circulating tumor DNA and/or protein biomarkers linked to cancer; and

(v) an assay involving immunofluorescent staining to identify cell phenotype and marker expression, followed by amplification and analysis by next-generation sequencing.

**[437]** In some embodiments of the disclosed method, the lung cancer is small cell lung cancer or non-small cell lung cancer.

**[438]** In some embodiments of the disclosed method, the lung cancer is non-small cell lung cancer.

**[439]** In some embodiments of the disclosed method, the non-small cell lung cancer is lung adenocarcinoma.

**[440]** In some embodiments of the disclosed method, the method is performed to monitor a lung cancer patient for response to treatment of an anti-lung cancer therapy (*e.g.*, surgery, radiation therapy, chemotherapy, radiosurgery, targeted drug therapy, immunotherapy) and/or for cancer recurrence/metastasis.

**[441]** In some embodiments of the disclosed method for detecting cancer, the method comprising steps of: detecting on surfaces of intact extracellular vesicles from a human blood sample co-localization of at least two biomarkers whose combined expression level has been determined to be associated with cancer; comparing the detected co-localization level with the determined level; and detecting cancer when the detected co-localization level is at or above the determined level.

**[442]** In some embodiments of the disclosed method for detecting cancer, the method comprising steps of: contacting a sample comprising exosomes with a set of detection probes that specifically bind to surface biomarkers on the exosomes to detect cancer-associated exosomes in

the sample with a specificity within a range of 95% to 100% and sensitivity within a range of 30% to 100%.

**[443]** In some embodiments of the disclosed method, the method comprising steps of: capturing exosomes from a biological sample with a capture agent that selectively interacts with a cancer-specific surface biomarker on the exosomes; and contacting the captured exosomes with at least one set of at least two detection probes that each selectively interacts with a surface biomarker on the exosomes; and detecting a product formed when the at least two detection probes of the set are in sufficiently close proximity, such detection indicating co-localization of the surface biomarkers.

**[444]** In some embodiments of the disclosed method, the method comprising steps of: contacting a sample comprising exosomes with a set of probes that specifically bind to surface biomarkers on the exosomes to detect cancer-associated exosomes in the sample, wherein: (i) each probe in the set comprises a target binding moiety directed to a surface biomarker on the exosomes; and (ii) the set comprises at least one capture probe and at least two detection probes, wherein each detection probe further comprises a detection moiety.

**[445]** In some embodiments of the disclosed method, the method comprising steps of: performing a proximity assay that detects a surface biomarker signature on exosomes from a human subject, the step of performing being performed a period of time after a performance of a prior assay to detect the surface biomarker signature on exosomes from the human subject; and comparing results of the performed assay with those of the prior assay.

**[446]** In some embodiments of the disclosed method, the method comprising steps of: contacting exosomes with at least two detection probes, wherein each detection probe comprises (i) a binding moiety; and (ii) an oligonucleotide entity, wherein the binding moiety is the same and the oligonucleotide entities complement one another.

**[447]** In some embodiments of the disclosed method, the method comprising detecting marker proximity on exosome surfaces, including an improvement that comprises contacting the exosomes with at least a pair of binding agents that each comprise a binding moiety and a proximity moiety, wherein the binding moieties are the same and the proximity moieties complement one another; and detecting an interaction between the proximity moieties.

**[448]** One aspect of the disclosure herein is a kit for detection of lung cancer comprising:

(a) a capture agent comprising a target-capture moiety directed to an extracellular vesicle-associated surface biomarker; and

(b) at least one set of detection probes, which set comprises at least two detection probes each directed to a target biomarker of a target biomarker signature for lung cancer, wherein the detection probes each comprise:

(i) a target binding moiety directed at the target biomarker of the target biomarker signature for lung cancer; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same extracellular vesicle;

wherein the target biomarker signature for lung cancer comprises:

at least one extracellular vesicle-associated surface biomarker and

at least one target biomarker selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers, wherein:

the surface biomarkers are selected from ADGRF1, ABCC3, ALCAM, ARSL, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, GPC4, HAS3, HS6ST2, IG1FR, KDELR3, KRTCAP3, LAMB3, LAPTM4B, LARGE2, LFNG, LSR, MAL2, MANEAL, MET, MSLN, MUC1, MUC21, NRCAM, PIGT, PODXL2, PRRG4, PRSS21, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SLC7A11, SMIM22, SMPDL3B, ST14, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, TACSTD2, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, UCHL1 and combinations thereof;

the intravesicular biomarkers are selected from AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D,

S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and

the intravesicular RNA biomarkers are selected from RNA transcripts (e.g., mRNA transcripts) encoded by human genes as follows: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof.

**[449]** In some embodiments of the disclosed kit, when the at least one target biomarker is selected from one or more of the surface biomarkers, the selected surface biomarker(s) and the at least one extracellular vesicle-associated surface biomarker are different.

**[450]** In some embodiments of the disclosed kit, the extracellular vesicle-associated surface biomarker is or comprises ADGRF1, ALCAM, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CLDN3, CLDN4, DSG2, EGFR, EPCAM, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, HAS3, IG1FR, LAMB3, LAPTM4B, LARGE2, MAL2, MET, MSLN, MUC1, NRCAM, PIGT, PODXL2, PRSS21, ROS1, SDC1, SLC34A2, SLC7A11, SMIM22, SMPDL3B, ST14, UCHL1, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, TACSTD2, TMPRSS4, TSPAN8, TNFRSF10B, or combinations thereof.

**[451]** One aspect of the disclosure herein is a kit for detection of lung cancer comprising:

(a) a capture agent comprising a target-capture moiety directed to an extracellular vesicle-associated surface biomarker; and

(b) at least one set of detection probes, which set comprises at least two detection probes each directed to a target biomarker of a target biomarker signature for lung cancer, wherein the detection probes each comprise:

(i) a target binding moiety directed at the target biomarker of the target biomarker signature for lung cancer; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same extracellular vesicle;

wherein the target biomarker signature for lung cancer comprises:

at least one extracellular vesicle-associated surface biomarker and

at least one target biomarker, wherein the target biomarker is or comprises a surface biomarker, wherein the target biomarker signature comprises a biomarker combination as listed in Table 4.

**[452]** In some embodiments of the disclosed kit, the extracellular vesicle-associated surface biomarker is or comprises a CEACAM5 polypeptide and/or a SLC34A2 polypeptide.

**[453]** In some embodiments of the disclosed kit, the target binding moiety of the at least two detection probes is each directed to the same target biomarker of the target biomarker signature.

**[454]** In some embodiments of the disclosed kit, the same target biomarker is or comprises CEACAM6.

**[455]** In some embodiments of the disclosed kit, the oligonucleotide domain of the at least two detection probes are different.

**[456]** In some embodiments of the disclosed kit, the target binding moiety of the at least two detection probes is each directed to a distinct target biomarker of the target biomarker signature.

**[457]** In some embodiments of the disclosed kit, the extracellular vesicle-associated surface biomarker is or comprises a CEACAM5 polypeptide and/or a SLC34A2 polypeptide; and the at least two detection probes are directed to CEACAM6 and EPCAM, respectively.

**[458]** In some embodiments of the disclosed kit, the extracellular vesicle-associated surface biomarker is or comprises a CEACAM5 polypeptide; and the at least two detection probes are directed to CEACAM6 and SLC34A2, respectively.

**[459]** In some embodiments of the disclosed kit, the kit further comprising at least one additional regent (*e.g.*, a ligase, a fixation agent, and/or a permeabilization agent).

**[460]** In some embodiments of the disclosed kit, the kit comprising at least two sets (including, *e.g.*, at least three sets) of detection probes, which each set comprises at least two detection probes each directed to a target biomarker of a distinct target biomarker signature for lung cancer.

**[461]** One aspect of the disclosure herein is a kit of any one of claims 44-46, comprising:

(a) a first capture agent comprising a target-capture moiety;

(b) a second capture agent comprising a target-capture moiety;

(c) at least two sets of detection probes, wherein the detection probes each comprise:

(i) a target binding moiety directed at a target surface biomarker; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same extracellular vesicle.

**[462]** In some embodiments of the disclosed kit, the kit comprising:

(a) a first capture agent comprising a target-capture moiety;

(b) a second capture agent comprising a target-capture moiety;

(c) a third capture agent comprising a target-capture moiety;

(d) at least three sets of detection probes, wherein the detection probes each comprise:

(i) a target binding moiety directed at a target surface biomarker; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same extracellular vesicle.

**[463]** One aspect of the disclosure herein is a complex comprising:

(a) an extracellular vesicle expressing a target biomarker signature for lung cancer, wherein the target biomarker signature comprises:

at least one extracellular vesicle-associated surface biomarker and

at least one target biomarker selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers, wherein:

the surface biomarkers are selected from ADGRF1, ABCC3, ALCAM, ARSL, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, GPC4, HAS3, HS6ST2, IG1FR, KDELR3, KRTCAP3, LAMB3, LAPTM4B, LARGE2, LFNG, LSR, MAL2, MANEAL, MET, MSLN, MUC1, MUC21, NRCAM, PIGT, PODXL2, PRRG4, PRSS21, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SLC7A11, SMIM22, SMPDL3B, ST14, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, TACSTD2, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, UCHL1 and combinations thereof;

the intravesicular biomarkers are selected from AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and

the intravesicular RNA biomarkers are selected from RNA transcripts (e.g., mRNA transcripts) encoded by human genes as follows: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof;

wherein the extracellular vesicle is immobilized onto a solid substrate comprising a target-capture moiety directed to the extracellular vesicle-associated surface biomarker;

(b) a first detection probe and a second detection probe each bound to the extracellular vesicle, wherein each detection probe comprises:

(i) a target binding moiety directed to one of the target biomarker of the tumor target biomarker signature; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the first and second detection probes are hybridized to each other.

**[464]** In some embodiments of the disclosed complex, when the at least one target biomarker is selected from one or more of the surface biomarkers, the selected surface biomarker(s) and the at least one extracellular vesicle-associated surface biomarker are different;

**[465]** In some embodiments of the disclosed complex, the extracellular vesicle-associated surface biomarker is or comprises ADGRF1, ALCAM, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CLDN3, CLDN4, DSG2, EGFR, EPCAM, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, HAS3, IG1FR, LAMB3, LAPTM4B, LARGE2, MAL2, MET, MSLN, MUC1, NRCAM, PIGT, PODXL2, PRSS21, ROS1, SDC1, SLC34A2, SLC7A11, SMIM22, SMPDL3B, ST14, UCHL1, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, TACSTD2, TMPRSS4, TSPAN8, TNFRSF10B, or combinations thereof.

**[466]** One aspect of the disclosure herein is a complex comprising:

(a) an extracellular vesicle expressing a target biomarker signature for lung cancer, wherein the target biomarker signature comprises:

at least one extracellular vesicle-associated surface biomarker and

at least one target biomarker, wherein the target biomarker is or comprises a surface biomarker, wherein the target biomarker signature is or comprises a biomarker combination as listed in Table 4;

wherein the extracellular vesicle is immobilized onto a solid substrate comprising a target-capture moiety directed to the extracellular vesicle-associated surface biomarker;

(b) a first detection probe and a second detection probe each bound to the extracellular vesicle, wherein each detection probe comprises:

(i) a target binding moiety directed to one of the target biomarker of the tumor target biomarker signature; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the first and second detection probes are hybridized to each other.

**[467]** In some embodiments of the disclosed complex, the extracellular vesicle-associated surface biomarker is or comprises a CEACAM5 polypeptide and/or a SLC34A2 polypeptide and/or a CLDN6 polypeptide.

**[468]** In some embodiments of the disclosed complex, the target binding moiety of the at least two detection probes is each directed to the same target biomarker of the target biomarker signature.

**[469]** In some embodiments of the disclosed complex, the same target biomarker is or comprises CEACAM6.

**[470]** In some embodiments of the disclosed complex, the oligonucleotide domain of the at least two detection probes are different.

**[471]** In some embodiments of the disclosed complex, the target binding moiety of the at least two detection probes is each directed to a distinct target biomarker of the target biomarker signature.

**[472]** In some embodiments of the disclosed complex, the extracellular vesicle-associated surface biomarker is or comprises a CEACAM5 polypeptide and/or a SLC34A2 polypeptide; and the at least two detection probes are directed to CEACAM6 and EPCAM, respectively.

**[473]** In some embodiments of the disclosed complex, the extracellular vesicle-associated surface biomarker is or comprises a CEACAM5 polypeptide; and the at least two detection probes are directed to CEACAM6 and SLC34A2, respectively.

**[474]** In some embodiments of the disclosed complex, the solid substrate comprises a magnetic bead.

**[475]** In some embodiments of the disclosed complex, the target-capture moiety is or comprises an antibody agent.

**[476]** In some embodiments of the disclosed complex, the complex comprising: (a) an exosome having at least one target biomarker on its surface; and (b) a first detection probe and a second detection probe each bound to the exosome, wherein each of the first detection probe and the second detection probe comprises: (i) a target binding moiety directed to a target biomarker expressed by the exosome; and (ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the first and second detection probes are hybridized to each other.

**[477]** In some embodiments of the disclosed complex, the complex comprising extracellular vesicles from a human blood sample bound to a set of at least two probes, each of which comprises a biomarker binding moiety and an oligonucleotide domain, wherein two or more bound probes are in proximity to one another so that their oligonucleotide domains hybridize to each other to form a ligatable hybrid.

**[478]** In some embodiments of the disclosed complex, the complex comprising: (a) an exosome comprising a cancer-associated target biomarker signature; and (b) at least a first detection probe and a second detection probe each bound to the exosome, wherein each of the detection probes comprise: (i) a target binding moiety directed to the target biomarker signature; and (ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the detection probes are at least partially complementary.

**[479]** One aspect of the disclosure herein is a set of probes for use in a method, kit, or complex of any one of claims 1-71, wherein each set of probes comprises: (a) a biomarker binding moiety that specifically binds to a surface biomarker on extracellular vesicles from cancer cells; and (b) an oligonucleotide domain, wherein the oligonucleotide domains of probes within the set are arranged and constructed so that, when the probes are bound to their target biomarkers, their oligonucleotide domains hybridize to one another to form a ligatable hybrid only when the target biomarkers are in proximity to one another.

**[480]** Other features of the invention will become apparent in the course of the following description of exemplary embodiments, which are given for illustration of the invention and are not intended to be limiting thereof.

## EXEMPLIFICATION

### Example 1: Detection of an exemplary target biomarker signature in individual extracellular vesicles associated with lung cancer

**[481]** The present Example describes synthesis of detection probes for targets (*e.g.*, target biomarker(s)) each comprising a target-binding moiety and an oligonucleotide domain (comprising a double-stranded portion and a single stranded overhang) coupled to the target-binding moiety. The present Example further demonstrates that use of such detection probes to detect the presence or absence of biological entities (*e.g.*, extracellular vesicles) comprising two or more distinct targets.

**[482]** In some embodiments, a detection probe can comprise a double-stranded oligonucleotide with an antibody agent specific to a target cancer biomarker at one end and a single stranded overhang at another end. When two or more detection probes are bound to the same biological entity (*e.g.*, an extracellular vesicle), the single-stranded overhangs of the detection probes are in close proximity such that they can hybridize to each other to form a double-stranded complex, which can be subsequently ligated and amplified for detection.

**[483]** This study employed at least two detection probes in a set. In some embodiments, such at least two detection probes, which may be directed to the same target biomarker. In some embodiments, such at least two detection probes directed to the same target are directed to different epitopes of the same target or to the same epitope of the same target. In some embodiments, such at least two detection probes are directed to distinct targets. A a skilled artisan reading the present disclosure will understand that two detection probes can be directed to different target biomarkers, or that three or more detection probes, each directed towards a distinct target protein, may be used. Further, compositions and methods described in this Example can be extended to applications in different biological samples (*e.g.*, comprising extracellular vesicles).

**[484]** The present Example shows experimental data from certain experiments demonstrating technologies provided herein are capable of detecting lung cancer (*e.g.*, lung

adenocarcinoma and/or lung squamous cell carcinoma) in patient samples using an exemplary biomarker combinations as described herein (*e.g.*, SLC34A2, CEACAM5, CEACAM6, and EpCAM, *e.g.*, in some embodiments, using SLC34A2 capture with CEACAM6 + CEACAM6 extracellular vesicles in PBS using a duplex system assay described herein. See, for example, ***FIGS. 1 and 2***.

**[485]** With such a duplex system assay capable of detecting cell line-derived extracellular vesicles (CLD-EVs), a study containing patient samples from a primary cohort of patients (see *FIG. 4*) with lung cancer of each stage (Stage I-IV) and/or subtypes and from various control groups (*e.g.*, healthy subjects) were performed.

*Overview of an exemplary assay*

**[486]** In some embodiments, a target entity detection system described herein is a duplex system. In some embodiments, such a duplex system, *e.g.*, as illustrated in ***FIG. 2***, utilizes two antibodies that each recognize a different epitope. Paired double-stranded template DNAs are also utilized in qPCR, each of which has specific four-base 5′ overhangs complementary to the 5′ overhang on its partner. Each antibody is conjugated with one of the two double-stranded DNA templates. When the antibodies bind their target epitopes, the sticky ends of the respective templates can hybridize. These sticky ends are then ligated together by T7 ligase, prior to PCR amplification. For hybridization between the two DNA templates to occur, the two antibodies need to be bound close enough to each other (within 50 to 60 nm, the length of the DNA linker and antibody). Any templates that bind but remain unligated will not produce PCR product, as shown in ***FIG. 2.***

*Healthy controls versus stage I, II, III, and IV lung adenocarcinoma plasma:*

**[487]** Plasma samples from healthy controls and Lung Adenocarcinoma (LUAD) patients were processed to obtain purified extracellular vesicles, which were interrogated using an exemplary assay as described below.

**[488]** Purified EVs were captured using magnetic beads covalently conjugated with anti-SLC34A2 , or anti-CEACAM5 antibodies. The EVs captured by the beads were profiled using a set of two detection probes, each comprising an antibody directed to a target biomarker (*e.g.*, CEACAM6, EpCAM, or SLC34A2) and a distinct oligonucleotide domain (*e.g.*, ones as described herein).

**[489]** The biomarker combinations were carefully selected to minimize cross-reactivity with healthy-tissue-derived extracellular vesicles, which are (i) SLC34A2 and CEACAM6, (ii) SLC34A2, CEACAM6, and EPCAM, and (iii) CEACAM5, CEACAM6, and SLC34A2. The cross-reactivity of such a biomarker combination with healthy tissues was bioinformatically predicted, in part, by using a heatmap of differentially expressed mRNAs in lung adenocarcinoma. Thus, different combinations of markers can be predicted to be much more abundant on the surface of lung cancer extracellular vesicles than on the surface of extracellular vesicles from healthy tissues. In some embodiments, such a biomarker combination may be selected from: (i) SLC34A2 capture probe and CEACAM6 + CEACAM6 detection probes, (ii) SLC34A2 capture probe and CEACAM6 + EPCAM detection probes, and (iii) CEACAM5 capture probe and CEACAM6 + SLC34A2 detection probes. Table 1 represents transcript expression scores of indicated biomarkers, as expressed in a lung cancer cell line vs. negative control cell line (*e.g.,* non-lung cancer cell line).

*Table 1. The transcript expression scores for the following biomarker combination, as expressed in certain lung cancer cell lines vs. negative control cell line (e.g., non-lung cancer line)*

| Genes | Lung Cancer Cell Line 1 | Negative Non-Lung Cancer |
|---|---|---|
| CEACAM6 | ++++ | - |
| SLC34A2 | +++ | - |
| EPCAM | +++ | + |
| CEACAM5 | ++ | - |

**Exemplary Methods:**

*Oligonucleotides*

**[490]** In some embodiments, oligonucleotides can have the following sequence structure and modifications. It is noted that the strand numbers below correspond to the numerical values associated with strands shown in ***FIG. 2***.

***Strand 1 v1:***

`/5AzideN/CAGTCTGACACAGCAGTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGG CTAGACAGAGGTGT`, where /5AzideN/ refers to an azide group linked to the 5′ oligonucleotide terminus via a NHS ester linker, or

/5AmMC12/CAGTCTGACACAGCAGTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGG CTAGACAGAGGTGT, where /5AmMC12/ refers to an amine group (*e.g.*, a primary amino group) linked to the 5′ oligonucleotide terminus via a 12-carbon spacer, or /5ThiolMC6/CAGTCTGACACAGCAGTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACT GGCTAGACAGAGGTGT, where /5ThiolMC6/ refers to a thiol linked to the 5′ oligonucleotide terminus via a 6-carbon spacer.

***Strand 2 v1:***

/5AzideN/GACCTGACCTACAGTGACCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCT CCTGGTCTCACTAG, where /5AzideN/ refers to an azide group linked to the 5′ oligonucleotide terminus via a NHS ester linker, or /5AmMC12/GACCTGACCTACAGTGACCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCT CCTGGTCTCACTAG, where /5AmMC1/ refers to an amine group (*e.g.*, a primary amino group) linked to the 5′ oligonucleotide terminus via a 12-carbon spacer, or /5ThiolMC6/GACCTGACCTACAGTGACCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGG CTCCTGGTCTCACTAG, where /5ThiolMC6/ refers to a thiol linked to the 5′ oligonucleotide terminus via a 6-carbon spacer

***Strand 3 v1:***

/5Phos/GAGTACACCTCTGTCTAGCCAGTCACGGATGTCAAGGGTAGCAGCGACGATTAACGA CTGCTGTGTCAGACTG, wherein /5Phos/ refers to a phosphate group linked to the 5′ oligonucleotide terminus

***Strand 4 v1:***

/5Phos/ACTCCTAGTGAGACCAGGAGCCAAACTGGACAGTGGCTAATCAGGCAAGGCTATGGT CACTGTAGGTCAGGTC, wherein /5Phos/ refers to a phosphate group linked to the 5′ oligonucleotide terminus

***Strand 5 v1:***

CAGTCTGACACAGCAGTCGT

***Strand 6 v1:***

GACCTGACCTACAGTGACCA

***Strand 7 (Probe) v1:***

/56-FAM/TGGCTAGAC/ZEN/AGAGGTGTACTCCTAGTGAGA/3IABkFQ/, wherein /56-FAM/ refers to a fluorescein (*e.g.*, 6-FAM) at the 5′ oligonucleotide terminus; and /3IABkFQ/ refers to a fluorescein quencher at the 3′ oligonucleotide terminus

**[491]** In some embodiments, oligonucleotides can have the following sequence structure and modifications. It is noted that the strand numbers below correspond to the numerical values associated with strands shown in ***FIG. 2***.

***Strand 1 v2:***

/5AzideN/CAGTCTGACTCACCACTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGG CTAGACAGAGGTGT, where /5AzideN/ refers to an azide group linked to the 5′ oligonucleotide terminus via a NHS ester linker, or

/5AmMC12/CAGTCTGACTCACCACTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACTGG CTAGACAGAGGTGT, where /5AmMC12/ refers to an amine group (*e.g.*, a primary amino group) linked to the 5′ oligonucleotide terminus via a 12-carbon spacer, or

/5ThiolMC6/CAGTCTGACTCACCACTCGTTAATCGTCGCTGCTACCCTTGACATCCGTGACT GGCTAGACAGAGGTGT, where /5ThiolMC6/ refers to a thiol linked to the 5′ oligonucleotide terminus via a 6-carbon spacer

***Strand 2 v2:***

/5AzideN/CACCAGACCTACGAAGTCCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCT CCTGGTCTCACTAG, where /5AzideN/ refers to an azide group linked to the 5′ oligonucleotide terminus via a NHS ester linker, or

/5AmMC12/CACCAGACCTACGAAGTCCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGGCT CCTGGTCTCACTAG, where /5AmMC1/ refers to an amine group (*e.g.*, a primary amino group) linked to the 5′ oligonucleotide terminus via a 12-carbon spacer, or

/5ThiolMC6/CACCAGACCTACGAAGTCCATAGCCTTGCCTGATTAGCCACTGTCCAGTTTGG CTCCTGGTCTCACTAG, where /5ThiolMC6/ refers to a thiol linked to the 5′ oligonucleotide terminus via a 6-carbon spacer

***Strand 3 v2:***

/5Phos/GAGTACACCTCTGTCTAGCCAGTCACGGATGTCAAGGGTAGCAGCGACGATTAACGA GTGGTGAGTCAGACTG, wherein /5Phos/ refers to a phosphate group linked to the 5′ oligonucleotide terminus

***Strand 4 v2:***

/5Phos/ACTCCTAGTGAGACCAGGAGCCAAACTGGACAGTGGCTAATCAGGCAAGGCTATGGA CTTCGTAGGTCTGGTG, wherein /5Phos/ refers to a phosphate group linked to the 5′ oligonucleotide terminus

***Strand 5 v2:***

CAGTCTGACTCACCACTCGT

***Strand 6 v2:***

CACCAGACCTACGAAGTCCA

***Strand 7 (Probe) v2:***

/56-FAM/TGGCTAGAC/ZEN/AGAGGTGTACTCCTAGTGAGA/3IABkFQ/, wherein /56-FAM/ refers to a fluorescein (*e.g.*, 6-FAM) at the 5′ oligonucleotide terminus; and /3IABkFQ/ refers to a fluorescein quencher at the 3′ oligonucleotide terminus

**[492]** In some embodiments, oligonucleotides can have the following sequence structure and modifications. It is noted that the strand numbers below correspond to the numerical values associated with strands shown in ***FIG. 2***.

***Strand 1 v1-med:***

/5AzideN/CAGTCTGACACAGCAGTCGTGACTGGCTAGACAGAGGTGT, where /5AzideN/ refers to an azide group linked to the 5′ oligonucleotide terminus via a NHS ester linker, or /5AmMC12/CAGTCTGACACAGCAGTCGTGACTGGCTAGACAGAGGTGT, where /5AmMC12/ refers to an amine group (*e.g.,* a primary amino group) linked to the 5′ oligonucleotide terminus via a 12-carbon spacer, or /5ThiolMC6/CAGTCTGACACAGCAGTCGTGACTGGCTAGACAGAGGTGT, where /5ThiolMC6/ refers to a thiol linked to the 5′ oligonucleotide terminus via a 6-carbon spacer.

***Strand 2 v1-med:***

/5AzideN/GACCTGACCTACAGTGACCATTGGCTCCTGGTCTCACTAG, where /5AzideN/ refers to an azide group linked to the 5′ oligonucleotide terminus via a NHS ester linker, or /5AmMC12/GACCTGACCTACAGTGACCATTGGCTCCTGGTCTCACTAG, where /5AmMC1/ refers to an amine group (*e.g.,* a primary amino group) linked to the 5′ oligonucleotide terminus via a 12-carbon spacer, or /5ThiolMC6/GACCTGACCTACAGTGACCATTGGCTCCTGGTCTCACTAG, where /5ThiolMC6/ refers to a thiol linked to the 5′ oligonucleotide terminus via a 6-carbon spacer

***Strand 3 v1-med:***

/5Phos/GAGTACACCTCTGTCTAGCCAGTCACGACTGCTGTGTCAGACTG, wherein /5Phos/ refers to a phosphate group linked to the 5′ oligonucleotide terminus

***Strand 4 v1-med:***

/5Phos/ACTCCTAGTGAGACCAGGAGCCAATGGTCACTGTAGGTCAGGTC, wherein /5Phos/ refers to a phosphate group linked to the 5′ oligonucleotide terminus

***Strand 5 v1:***

CAGTCTGACACAGCAGTCGT

***Strand 6 v1:***

GACCTGACCTACAGTGACCA

***Strand 7 (Probe) v1:***

/56-FAM/TGGCTAGAC/ZEN/AGAGGTGTACTCCTAGTGAGA/3IABkFQ/, wherein /56-FAM/ refers to a fluorescein (*e.g.,* 6-FAM) at the 5′ oligonucleotide terminus; and /3IABkFQ/ refers to a fluorescein quencher at the 3′ oligonucleotide terminus.

*Antibody-oligonucleotide (e.g., antibody-DNA) conjugation:*

**[493]** Antibody aliquots ranging from 25-100 µg were conjugated with oligonucleotide strands. For example, 60 µg aliquots of antibody was conjugated with hybridized strands 1+3 and 2+4, for example, using copper-free click chemistry. The first step was to prepare DBCO-functionalized antibodies to participate in the conjugation reaction with azide-modified oligonucleotide domain (*e.g.*, DNA domain). This began with reacting the antibodies with the DBCO-PEG5-NHS heterobifunctional cross linker. The reaction between the NHS ester and available lysine groups was allowed to take place at room temperature for 2 hours, after which unreacted crosslinker was removed using centrifugal ultrafiltration. To complete the conjugation, azide-modified oligonucleotide domains (*e.g.*, DNA domain) and the DBCO-functionalized antibodies were allowed to react overnight at room temperature. The concentration of conjugated antibody was measured, for example, using the Qubit protein assay.

*Cell Culture*

**[494]** Negative control cells (e.g., non-lung cancer cells such as melanoma cells or healthy cells) were grown in Eagle's Minimum Essential Medium (EMEM) with 10% exosome-free FBS and 50 units of penicillin/streptomycin per mL. Lung adenocarcinoma cells were grown in Roswell Park Memorial Institute (RPMI 1640) with 10% exosome-free FBS and 50 units of

penicillin/streptomycin per mL. Exemplary lung cancer cell lines that may be useful to develop an assay for detection of lung cancer (e.g., ones as described herein) include, but are not limited to, HCC4006, PC9, LO68, LUDLU-1, COR-L105, SKLU1, SKMES1, NCI-H727, LC-2/AD, NCIH358, ChaGo-K-1, MOR/CPR, MOR/0.4R, MOR/0.2R, NCIH-322 and cells lines described and discussed in Gazdar et al., "Lung Cancer Cell Lines as Tools for Biomedical Discovery and Research" Journal of the National Cancer Institute: 2010 September 8; 102(17): 1310-1321, which is incorporated herein by reference for the purpose described herein. All cell lines were maintained at 5% CO2 and 37 °C and the passage number was below 20.

*Purification of extracellular vesicles from cell culture medium*

**[495]** In some embodiments, lung cancer cells and negative control cells may be grown in their respective media until they reach ~80% confluence. The cell culture medium may be collected and spun at 300 RCF for 5 minutes at room temperature (RT) to remove cells and debris. The supernatant may then be collected and used in assays as described herein or frozen at –80 °C.

*Thawing*

**[496]** If prior to use, samples were stored at -80 °C, they are thawed. In brief, 50 mL tubes containing frozen conditioned media placed in plastic racks, the racks are placed in an empty ice bucket. Room temperature (RT) water is added, and samples are allowed to thaw, with periodic inversion/shaking to facilitate thawing. Tubes are consolidated such that all the tubes for each cell line are the same volume. A typical purification volume is approximately 200 mLs of spent medium per cell line. If larger batches are desired, this volume can be increased.

*Clarification*

**[497]** In some embodiments, samples are clarified prior to use. Clarification of media serves to remove cells and debris. In brief, 1) spin at 1300 RCF for 10 mins; transfer supernatant to a new 50 mL conical tube using a pipette, leaving ~1 cm of medium (to avoid disturbing the pellet), the remaining media is not decanted; 2) spin at 2000 RCF for 30 mins; transfer supernatant to a new 50 mL conical tube using a pipette, leaving ~1 cm of medium (to avoid disturbing the pellet), the remaining media is not decanted.

*Concentrate Media*

**[498]** In some embodiments, samples are concentrated. In brief: 1) a single 15 mL 10 kDa MWCO filter is used for approximately 100 mLs of medium (for example, for a 200 mL

batch, two10 kDa MWCO ultrafiltration tubes will be needed). In some embodiments, the same ultrafiltration column can be sequentially added to and re-spun to enable the concentration of large volumes of medium. In general, columns were utilized according to the manufacturer's protocol. Columns are spun for 10-12 minutes each time, at maximum speed (2500 to 4,300 RCF). 2) When each of the two tubes containing the same spent medium reaches ~1500 uL, the two tubes are combined into one, the now empty tube may be utilized as a balance. 3) When removing the concentrated medium, the sides of the concentration chamber may be flushed to release as many entrapped EVs as possible, while avoiding frothing, the consolidated media may be concentrated until there is 1 mL left. 4) The media is transferred to a 1.5 mL protein LoBind tube, with the 1 mL line marked, if necessary, volume is corrected to 1 mL with 20 nm filtered 1x PBS.

*Final Clarification Spin*

**[499]** To remove any remaining debris, the concentrated media can be centrifuged at 10,000 RCF for 10 minutes at 21 °C in a tabletop Eppendorf centrifuge.

*Run Concentrated Media Through Prepared IZON Columns*

**[500]** Izon columns are washed as described by the manufacturer, 20 nm filtered 1X PBS can be used to both wash the columns and recover the samples. 1 mL of concentrated spent medium can be run through the column and fractions can be collected (e.g., fractions 7, 8, and 9) in 5-mL Eppendorf flip-cap tubes, following the manufacturer's protocol.

*Particle counts:*

**[501]** Particle counts were obtained, *e.g.*, using a SpectraDyne particle counting instrument using the TS400 chips, to measure nanoparticle range between 65 and 1000 nm. In some embodiments, a particle size that is smaller than 65 nm or larger than 1000 nm may be desirable.

*Generation of patient plasma pools:*

**[502]** In some embodiments, pooled patient plasma pools were utilized. In brief, 1 mL aliquots of patient plasma were thawed at room temperature for at least 30 minutes. The tubes were vortexed briefly and spun down to consolidate plasma to the bottom of each tube. Plasma samples from a given patient cohort were combined in an appropriately sized container and mixed thoroughly by end-over-end mixing. Each plasma pool was split into 1 mL aliquots in Protein Lo-bind 1.5 mL Eppendorf tubes and refrozen at -80°C.

*Whole-plasma clarification (optional):*

**[503]** In some embodiments, prior to EVs purification, samples were blinded by personnel who would not participate in sample-handling. The patient-identification information was only revealed after the experiment was completed to enable data analysis. 1 mL aliquots of whole plasma were removed from storage at -80°C and subjected to three clarification spins to remove cells, platelets, and debris.

*Size-exclusion chromatography purification of EVs from clarified plasma:*

**[504]** Each clarified plasma sample (individual samples or pooled samples) was run through a single-use, size-exclusion purification column to isolate the EVs. Nanoparticles having a size range of about 65 nm to about 1000 nm were collected for each sample. In some embodiments, particle size that is smaller than 65 nm or larger than 1000 nm may be desirable.

*Capture-antibody conjugation to magnetic-capture beads:*

**[505]** Antibodies were conjugated to magnetic beads (*e.g.,* epoxy-functionalized Dynabeads™). Briefly, beads were weighed in a sterile environment and resuspended in buffer. Antibodies, at approximately 8 µg of Ab per mg of bead, were mixed with the functionalized beads and the conjugation reaction took place overnight at 37°C with end-over-end mixing. The beads were washed several times using the wash buffer provided by the conjugation kit and were stored at 4°C in the provided storage buffer, or at -20°C in a glycerol-based storage buffer.

*Direct capture of purified plasma EVs using antibody-conjugated magnetic beads:*

**[506]** In certain embodiments, purified plasma EVs were directly captured from clarified plasma samples. For example, for SLC34A2 or CEACAM5 capture, a diluted sample of purified plasma EVs were incubated with magnetic beads conjugated with respective antibodies for an appropriate time period at an appropriate temperature, *e.g.*, at room temperature.

*Binding of antibody-oligonucleotide conjugates to EVs bound on magnetic capture beads:*

**[507]** Antibody-oligonucleotide conjugates (*e.g.*, anti-SLC34A2, CEACAM6, or EPCAM antibody-oligonucleotide conjugates; "antibody probes") were diluted in an appropriate buffer at their optimal concentrations. Antibody probes were allowed to interact with a sample comprising EVs bound on magnetic capture beads.

*Post-binding washes:*

**[508]** In some embodiments, samples were washed, *e.g.*, multiple times, in an appropriate buffer.

*Ligation:*

**[509]** After the wash to remove unbound antibody-oligonucleotide conjugates, the beads with bound extracellular vesicles and bound antibody-oligonucleotide conjugates were contacted with a ligation mix. The mixtures were incubated for 20 minutes at room temperature.

*PCR:*

**[510]** Following ligation, the beads with bound extracellular vesicles and bound antibody-oligonucleotide conjugates were contacted with a PCR mix. PCR was performed in a 96-well plate, *e.g.*, on the Quant Studio 3, with the following exemplary PCR protocol: hold at 95 °C for 1 minute, perform 50 cycles of 95 °C for 5 seconds and 62 °C for 15 seconds. The rate of temperature change was chosen to be standard (2 °C per second). A single qPCR reaction was perform for each experimental replicate and ROX was used as the passive reference to normalize the qPCR signals. Data was then downloaded from the Quant Studio 3 machine and analyzed and plotted in Python 3.7.

*Data analysis:*

**[511]** In some embodiments, a binary classification system can be used for data analysis. In some embodiments, signals from a detection assay may be normalized based on a reference signal. For example, in some embodiments, normalized signals for a single antibody duplex were calculated by choosing a reference sample. In some embodiments, the equations used to calculate the normalized signal for an arbitrary sample *i* are given below, where $\text{Signal}_{\text{max}}$ is the signal from the highest concentration cell-line EVs standard.

$$\Delta Ct_i = Ct_{ref} - Ct_i$$

$$Signal_i = 2^{\Delta Ct_i}$$

$$Norm\ Signal_i = \frac{Signal_i}{Signal_{max}}$$

**Representative results:**

*LUAD cell line experiments*

**[512]** Purified cell-line EVs were diluted to an optimal concentration in an appropriate buffer and captured using appropriate capture probes (*e.g.*, anti-SLC34A2-functionalized beads or anti-CEACAM5-functionalized beads (1 mL replicates)). Captured EVs were analyzed using

antibody probes (*e.g.*, as described herein). It was observed that biomarker combinations described herein (*e.g.*, in combination with an exemplary assay such as, *e.g.*, as described in the present Example and illustrated in ***FIGS. 1-2***) is capable of distinguishing LUAD-derived EVs from the negative control cell line, with a signal strength that is well-correlated with the expression of the two markers (see Table 1).

*LUAD pilot patient plasma study*

**[513]** The demographics of the patients included in the LUAD patient plasma sample pilot study are provided in ***FIG. 4***. Care was taken to match age and gender as closely as possible across the different sample cohorts.

**[514]** Replicates of one milliliter or less (*e.g.,* 500µl or less, 400 µl or less, 300 µl or less, 200 µl or less, or 100 µl or less) of patient sample plasma was clarified as described above and EVs were purified using size-exclusion chromatography. EVs were captured using anti-SLC34A2 magnetic beads or anti-CEACAM5 magnetic beads. EVs captured by the anti-SLC34A2 magnetic beads were profiled using CEACAM6 + CEACAM6 detection probes or CEACAM6 + EPCAM detection probes (see ***FIGS. 5*** and ***6***). EVs captured by the anti-CEACAM5 magnetic beads were profiled using CEACAM6 + SLC34A2 detection probes (see ***FIG. 7***).

**Discussion:**

**[515]** The present Example describes the use of biomarker combinations in the assay described herein (*e.g.*, combinations of SLC34A2 + CEACAM6, or SLC34A2 + CEACAM6 + EPCAM, or CEACAM5+CEACAM6 + SLC34A2) (*e.g.*, in combination with a duplex assay as described in the present Example and illustrated in ***FIGS. 1-2***) can be capable of detecting lung adenocarcinoma with >99% specificity. In some embodiments, a biomarker combination includes SLC34A2 capture and CEACAM6 + CEACAM6 detection probes. In some such embodiments, a biomarker combination includes SLC34A2 capture and CEACAM6 + EpCAM detection probes. In some such embodiments, a biomarker combination includes CEACAM5 capture and CEACAM6 + SLC34A2 detection probes. In some embodiments, use of two or more biomarker combinations in an assay may increase the specificity and/or sensitivity of the assay.

**[516]** In some embodiments, a dendron, which can add up to 16 strands of oligonucleotide domain (*e.g.*, DNA) per antibody, can be used instead of one or two strands of DNA per antibody, for example, to enhance signal-to-noise.

**Example 2: Embodiments of biomarker signatures for lung cancer detection**

**[517]** In some embodiments, a biomarker combination including SLC34A2, CEACAM5, CEACAM6, and/or EPCAM (*e.g.*, in some embodiments, SLC34A2 capture and CEACAM6 + CEACAM6 detection probes, or SLC34A2 capture and CEACAM6 + EPCAM detection probes, or CEACAM5 capture and CEACAM6 + SLC34A2 detection probes) were used for detection of lung cancer (*e.g.*, following an assay as described in Example 1) in various subject populations including, *e.g.*, healthy controls, Stage I LUAD; Stage II LUAD; Stage III LUAD; and Stage IV LUAD.

**[518]** In some embodiments, two or more biomarker combinations described herein may be used together for detection of lung cancer, *e.g.,* to increase sensitivity of an assay. For example, in some embodiments, an assay for lung cancer detection may involve at least two biomarker combinations, wherein such at least two biomarker combinations each may comprise a different biomarker combination described herein (*e.g.*, but not limited to ones included in Tables 4-5).

**[519]** In some embodiments, three or more biomarker combinations described herein may be used together for detection of lung cancer, *e.g.,* to increase sensitivity of an assay. For example, in some embodiments, an assay for lung cancer detection may involve at least three biomarker combinations, wherein such at least three biomarker combinations each may comprise a different biomarker combination described herein (*e.g.*, but not limited to ones included in Tables 4-5).

**Example 3: Assessment of extracellular vesicle (EV) surface biomarkers as lung cancer biomarkers**

**[520]** In some embodiments, lung cancer detection includes detection of at least EV surface biomarker(s) following immunoaffinity capture of extracellular vesicles.

**[521]** In some embodiments, one or more surface biomarkers or extracellular membrane biomarkers that are present on extracellular vesicles ("capture biomarkers") can be used for immunoaffinity capture of lung cancer-associated extracellular vesicles. Examples of such capture biomarkers may include, but are not limited to ALCAM, B3GNT3, CDCP1, CDH1,

CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CLDN3, CLDN4, DSG2, EGFR, EPCAM, FOLR1, IG1FR, GJB1, GJB2, LAMB3, MET, MSLN, MUC1, PIGT, PODXL2, ROS1, SDC1, SLC34A2, SMPDL3B, ST14, sTn antigen, Tn antigen, T antigen, TACSTD2, TMPRSS4, TSPAN8, TNFRSF10B, and/or combinations thereof. Additional examples of such capture biomarkers may include but are not limited to Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, ADGRF1, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAT2, FAM241B, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SLC7A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, Lewis X antigen, Lewis Y/B antigen, LY6E, NOTCH2, NOTCH3, Phosphatidylserine, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA, Tn antigen, Lewis Y/CD174 antigen, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)) antigen, NeuGcGM3, and combinations thereof.

**[522]** In some embodiments, EV immunoassay methodology (*e.g.*, ones described herein such as in Example 1) and biomarker-validation process (*e.g.*, ones described herein such as in Example 1) can be used to assess additional surface biomarkers for lung cancer. In some embodiments, an antibody directed to a capture biomarker (*e.g.*, a surface biomarker present in lung cancer-associated EVs) is conjugated to magnetic beads and evaluated, optionally first on cell-line EVs then on patient samples, for its ability to bind the specific target protein. The antibody-coated bead is assessed for its ability to capture lung cancer-associated EVs and the EVs captured by the antibody-coated bead are read out using a target entity detection system (*e.g.*, a duplex system as described herein involving a set of two detection probes each directed to a target marker (*e.g.*, as described herein).

**[523]** In some embodiments, captured EVs can be read out using at least one or more (*e.g.*, 1, 2, 3, or more) of the following surface biomarkers: ADGRF1, ALCAM, ABCC3, ARSL, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, GPC4, HAS3, HS6ST2, KDELR3, KRTCAP3, IG1FR, LAMB3, LAPTM4B, LARGE2, LFNG, LSR, MAL2, MANEAL, MET, MSLN, MUC1, MUC21, NRCAM, PIGT, PODXL2, PRRG4, PRSS21, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, SLC44A4, SLC6A14, SLC7A7, SLC7A11, SMIM22, SMPDL3B, ST14, TACSTD2, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, UCHL1, and combinations thereof. In some embodiments, captured EVs can be read out using a set of detection probes (*e.g.*, as utilized and/or described herein), at least two of which are directed to one or more (*e.g.*, 1, 2, 3, or more) of the following surface biomarkers: ADGRF1, ALCAM, ABCC3, ARSL, B3GNT3, B3GNT5, CDCP1, CDH1, CDH3, CD55, CD274 (PD-L1), CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, DMBT1, DSG2, EGFR, EPCAM, EPHX3, EVA1A, FAM241B, FOLR1, FXYD3, GALNT14, GJB1, GJB2, GPC4, HAS3, HS6ST2, KDELR3, KRTCAP3, IG1FR, LAMB3, LAPTM4B, LARGE2, LFNG, LSR, MAL2, MANEAL, MET, MSLN, MUC1, MUC21, NRCAM, PIGT, PODXL2, PRRG4, PRSS21, ROS1, SDC1, SERINC2, SEZ6L2, SLC34A2, Lewis X antigen, sialyl Lewis X antigen, sTn antigen, Tn antigen, T antigen, SLC44A4, SLC6A14, SLC7A7, SLC7A11, SMIM22, SMPDL3B, ST14, TACSTD2, TMC4,

TMC5, TMEM45B, TMPRSS2, TMPRSS4, TNFRSF10B, TSPAN1, TSPAN8, UCHL1, and combinations thereof. In some embodiments, a set of detection probes comprises two detection probes each directed to the same surface. In some embodiments, a set of detection probes comprises two detection probes each directed to a distinct surface biomarker.

**[524]** In some embodiments, captured EVs can be read out using at least one or more (*e.g.*, 1, 2, 3, or more) of the following surface biomarkers: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, Lewis X antigen, Lewis Y/B antigen, LY6E, NOTCH2, NOTCH3, Phosphatidylserine, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA, Tn antigen, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and

combinations thereof. In some embodiments, captured EVs can be read out using a set of detection probes (*e.g.*, as utilized and/or described herein), at least two of which are directed to one or more (*e.g.*, 1, 2, 3, or more) of the following surface biomarkers: Sialyl Tn (sTn) antigen, Sialyl-6T antigen (6-sialyl core 1), T antigen, PTK7, TSPAN8, CELSR2, GPC1, ST14, PTPRZ1, GPR87, GJB5, GJB2, RHOV, LYPD3, CLDN7, DSP, SERINC2, ABHD17C, PERP, MPZL2, ITGB4, MEST, GPNMB, SLC35A2, Gb3 (CD77), ITGA6, ABCC5, ATP1B3, JAG1, TMPRSS11D, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide, (Lc3), Forssman antigen, ABCA3, ABCC1, ABCC3, ACBD3, ACSL5, ADGRF1, AGER, ALCAM, AP1M2, APH1A, APOO, ATP11A, ATP11B, ATP1B1, ATP6AP2, B3GNT5, B4GALT4, BCAP31, BSPRY, CD109, CD55, CD9, CDC42, CDH1, CDH3, CDKAL1, CEACAM5, CEACAM6, CELSR1, CIP2A, CISD2, CKAP4, CLCA2, CLDN1, CLIC6, CLPTM1L, CLSTN1, CNTN1, CPD, CYP2S1, CYP4F11, CYP4F3, DPY19L1, DSC2, DSC3, DSG2, DSG3, EGFR, EPCAM, EPHB3, FAM241B, FAT2, FBXO45, FERMT1, FOLR1, FXYD3, FZD6, GALNT1, GALNT3, GALNT5, GALNT6, GALNT14, GGCX, GOLM1, GOLPH3L, GRHL2, HACD3, HAS3, IER3IP1, IGSF3, IL1RAP, ITGA2, ITGB6, KLRG2, KPNA2, KRTCAP3, LAD1, LAMB3, LAMC2, LAMP3, LAMTOR2, LAPTM4B, LARGE2, LCLAT1, LPCAT1, LSR, MAGT1, MAL2, MARCKSL1, MET, MGAT1, MSLN, MUC1, MUC4, NCSTN, NECTIN1, NECTIN4, NRAS, NRCAM, NT5E, NUP210, PARL, PEX13, PIGN, PIGT, PLA2G4A, PLCH1, PLEC, PRSS21, PSMD2, PTDSS1, PTGFRN, PTPRF, QSOX1, RAB25, RAB38, RAB6B, RAP2B, RCC2, RIT1, SCAMP3, SDC1, SEL1L3, SHROOM2, SLC7A11, SLC2A1, SLC34A2, SLC35B2, SLC39A11, SMIM22, SMPDL3B, SOAT1, SPAST, SSR1, SSR4, SURF4, SYNGR2, TACSTD2, TESC, TFRC, TMC5, TMCO1, TMED2, TMED3, TMEM132A, TMEM33, TMPRSS4, TMTC3, TOMM22, TOR1AIP2, TRAM1, TRPV4, TTC33, UGT1A6, UPK1B, VAMP8, VMA21, VRK2, VWA1, XPR1, XXYLT1, ADAM28, AXL, BSG, CD274, CD47, CLU, DKK1, ERBB3, FLT4, GM3, HGF, IGF1R, IL6, KDR, LAG3, UCHL1, Lewis X antigen, Lewis Y/B antigen, LY6E, NOTCH2, NOTCH3, Phosphatidylserine, TIGIT, TNFRSF10A, TNFRSF10B, TNFSF18, TPBG, VEGFA, Tn antigen, Lewis Y/CD174, Sialyl Lewis X (sLex) (also known as Sialyl SSEA-1 (SLX)), NeuGcGM3, and combinations thereof. In some embodiments, a set of detection probes comprises two detection probes each directed to the same surface biomarker. In some

embodiments, a set of detection probes comprises two detection probes each directed to a distinct surface biomarker.

**Example 4: Assessment of mRNA in extracellular vesicles (intravesicular mRNA) as lung cancer biomarkers**

**[525]** In some embodiments, lung cancer detection includes detection of at least intravesicular mRNA(s) following immunoaffinity capture of extracellular vesicles.

**[526]** In some embodiments, one or more surface biomarkers or extracellular membrane biomarkers that are present on extracellular vesicles ("capture biomarkers"), for example, capture biomarkers as described in Example 3, can be used for immunoaffinity capture of lung cancer-associated extracellular vesicles.

**[527]** In some embodiments, EV nucleic acid detection assay (*e.g.*, reverse transcription PCR using primer-probe sets) and biomarker-validation process (*e.g.*, ones described herein such as in Example 1) can be used to assess mRNA biomarker candidates for lung cancer. In some embodiments, an antibody directed to a capture biomarker (*e.g.*, a surface biomarker present in lung cancer-associated EVs) is conjugated to magnetic beads and evaluated, optionally first on cell-line EVs then on patient samples, for its ability to bind the specific target biomarker. The antibody-coated bead is assessed for its ability to capture lung cancer-associated EVs and the captured EVs by the antibody-coated bead is profiled for their mRNA contents, for example, using one-step quantitative reverse transcription PCR (RT-qPCR) master mix.

**[528]** In some embodiments, captured EVs can be read out by detection of at least one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof. In some embodiments, captured EVs can be read out by detection of at least one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs using RT-qPCR: ABCA3, ABCC1, ABRACL, ACP5, ADAM23, ADH7,

AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, ANTXR1, AP1M2, APOBEC3B, APOBEC3C, AQP3, AREG, ARNTL2, ASF1B, ATP8B1, AURKB, B3GNT5, BAIAP2L1, BCAM, BIK, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA12, CA9, CALML3, CAPNS2, CBLC, CCL19, CCL5, CCNB2, CD109, CD24, CD53, CD74, CD9, CDC20, CDC42EP1, CDC45, CDCA4, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CEACAM5, CEACAM6, CELSR1, CENPW, CEP55, CES1, CHMP4C, CLCA2, CLDN1, CLDN4, CLDN7, CNN2, COL17A1, CPA3, CRABP2, CSTA, CTSC, CTSE, CX3CL1, CXADR, CXCR4, CYBB, CYP2S1, CYP4F11, DAPL1, DPYSL3, DSC2, DSC3, DSG2, DSG3, DSP, EFNA1, EFS, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHB3, EPHX1, ESRP1, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83D, FAM83H, FAT1, FAT2, FBLIM1, FBP1, FCER1G, FERMT1, FGFR2, FGFR3, FOXE1, FOXM1, FXYD3, GALNT3, GBP6, GJA1, GJB2, GJB3, GJB5, GJB6, GNA15, GPC1, GPC3, GPNMB, GPR87, GPRC5A, GPX2, GRHL2, GSTA1, HAS3, HCK, HOXB7, ID1, IGF2BP2, IGSF9, IL2RG, IMPA2, IRF6, ITGA2, ITGA6, ITGB4, ITGB6, IVL, JAG2, JUP, KCNS3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, KRTCAP3, LAMP3, LAPTM5, LGALS7B, LRP11, LRRC4, LSP1, LSR, LYPD3, MAGEA4, MAGEA6, MAL2, MAOA, MARCO, MCM2, MDFI, MET, MMP14, MPZL2, MUC1, MYBL2, MYH14, MYOF, MZB1, NCF2, NKG7, NNMT, NOTCH3, NRARP, NTRK2, NUP210, NUSAP1, OSGIN1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PERP, PIGR, PIGT, PITX1, PKP1, PKP3, PLEK, PLEK2, PLVAP, PMP22, POSTN, PPL, PPP1R14C, PRAME, PROM2, PRRG4, PRSS8, PTGES, PTGFRN, PTPN6, PTPRF, PTPRZ1, RAB25, RAB38, RAET1L, RARRES1, RBP1, RGS1, RHCG, RHOV, RIN2, RIPK4, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, S100A2, S100P, SCNN1A, SDC1, SERINC2, SERPINB13, SERPINB3, SERPINB5, SEZ6L2, SH3BP4, SHISA2, SLC1A5, SLC2A1, SLC34A2, SLC40A1, SLC6A8, SLC7A8, SNAI2, SOX2, SPI1, SPINT1, SPINT2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, THBD, THBS2, TK1, TM4SF1, TMC4, TMEM30B, TMEM54, TMPRSS11D, TMPRSS11E, TMPRSS4, TNFRSF18, TNS4, TOP2A, TP53I11, TP63, TPD52, TPX2, TREM2, TRIM29, TRIP13, TSPAN1, TSPAN13, TSPAN6,

TSPAN7, TUSC3, TYROBP, UBE2C, UPK1B, VAMP8, VANGL2, WLS, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof.

**[529]** In some embodiments, captured EVs can be read out by detection of at least one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof; and at least one or more (*e.g.*, 1, 2, 3, or more) of EV surface biomarkers described in Example 3. In some embodiments, captured EVs can be read out by detection of at least one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs using RT-qPCR: ABCA3, ABCC1, ABRACL, ACP5, ADAM23, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, ANTXR1, AP1M2, APOBEC3B, APOBEC3C, AQP3, AREG, ARNTL2, ASF1B, ATP8B1, AURKB, B3GNT5, BAIAP2L1, BCAM, BIK, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA12, CA9, CALML3, CAPNS2, CBLC, CCL19, CCL5, CCNB2, CD109, CD24, CD53, CD74, CD9, CDC20, CDC42EP1, CDC45, CDCA4, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CEACAM5, CEACAM6, CELSR1, CENPW, CEP55, CES1, CHMP4C, CLCA2, CLDN1, CLDN4, CLDN7, CNN2, COL17A1, CPA3, CRABP2, CSTA, CTSC, CTSE, CX3CL1, CXADR, CXCR4, CYBB, CYP2S1, CYP4F11, DAPL1, DPYSL3, DSC2, DSC3, DSG2, DSG3, DSP, EFNA1, EFS, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHB3, EPHX1, ESRP1, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83D, FAM83H, FAT1, FAT2, FBLIM1, FBP1, FCER1G, FERMT1, FGFR2, FGFR3, FOXE1, FOXM1, FXYD3, GALNT3, GBP6, GJA1, GJB2, GJB3, GJB5, GJB6, GNA15, GPC1, GPC3, GPNMB, GPR87, GPRC5A, GPX2, GRHL2, GSTA1, HAS3, HCK, HOXB7, ID1, IGF2BP2, IGSF9, IL2RG, IMPA2, IRF6, ITGA2, ITGA6, ITGB4, ITGB6, IVL, JAG2, JUP, KCNS3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8,

KRTCAP3, LAMP3, LAPTM5, LGALS7B, LRP11, LRRC4, LSP1, LSR, LYPD3, MAGEA4, MAGEA6, MAL2, MAOA, MARCO, MCM2, MDFI, MET, MMP14, MPZL2, MUC1, MYBL2, MYH14, MYOF, MZB1, NCF2, NKG7, NNMT, NOTCH3, NRARP, NTRK2, NUP210, NUSAP1, OSGIN1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PERP, PIGR, PIGT, PITX1, PKP1, PKP3, PLEK, PLEK2, PLVAP, PMP22, POSTN, PPL, PPP1R14C, PRAME, PROM2, PRRG4, PRSS8, PTGES, PTGFRN, PTPN6, PTPRF, PTPRZ1, RAB25, RAB38, RAET1L, RARRES1, RBP1, RGS1, RHCG, RHOV, RIN2, RIPK4, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, S100A2, S100P, SCNN1A, SDC1, SERINC2, SERPINB13, SERPINB3, SERPINB5, SEZ6L2, SH3BP4, SHISA2, SLC1A5, SLC2A1, SLC34A2, SLC40A1, SLC6A8, SLC7A8, SNAI2, SOX2, SPI1, SPINT1, SPINT2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, THBD, THBS2, TK1, TM4SF1, TMC4, TMEM30B, TMEM54, TMPRSS11D, TMPRSS11E, TMPRSS4, TNFRSF18, TNS4, TOP2A, TP53I11, TP63, TPD52, TPX2, TREM2, TRIM29, TRIP13, TSPAN1, TSPAN13, TSPAN6, TSPAN7, TUSC3, TYROBP, UBE2C, UPK1B, VAMP8, VANGL2, WLS, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof; and at least one or more (*e.g.,* 1, 2, 3, or more) of EV surface biomarkers described in Example 3.

**[530]** In some such embodiments, captured EVs can be read out (i) by detection of one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs using RT-qPCR: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof; and (ii) by using a set of detection probes (*e.g.*, as utilized and/or described herein), at least one of which are directed to one or more (*e.g.*, 1, 2, 3, or more) of EV surface biomarkers described in Example 3. In some such embodiments, captured EVs can be read out (i) by detection of one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs using RT-qPCR: ABCA3, ABCC1, ABRACL, ACP5,

ADAM23, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, ANTXR1, AP1M2, APOBEC3B, APOBEC3C, AQP3, AREG, ARNTL2, ASF1B, ATP8B1, AURKB, B3GNT5, BAIAP2L1, BCAM, BIK, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA12, CA9, CALML3, CAPNS2, CBLC, CCL19, CCL5, CCNB2, CD109, CD24, CD53, CD74, CD9, CDC20, CDC42EP1, CDC45, CDCA4, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CEACAM5, CEACAM6, CELSR1, CENPW, CEP55, CES1, CHMP4C, CLCA2, CLDN1, CLDN4, CLDN7, CNN2, COL17A1, CPA3, CRABP2, CSTA, CTSC, CTSE, CX3CL1, CXADR, CXCR4, CYBB, CYP2S1, CYP4F11, DAPL1, DPYSL3, DSC2, DSC3, DSG2, DSG3, DSP, EFNA1, EFS, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHB3, EPHX1, ESRP1, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83D, FAM83H, FAT1, FAT2, FBLIM1, FBP1, FCER1G, FERMT1, FGFR2, FGFR3, FOXE1, FOXM1, FXYD3, GALNT3, GBP6, GJA1, GJB2, GJB3, GJB5, GJB6, GNA15, GPC1, GPC3, GPNMB, GPR87, GPRC5A, GPX2, GRHL2, GSTA1, HAS3, HCK, HOXB7, ID1, IGF2BP2, IGSF9, IL2RG, IMPA2, IRF6, ITGA2, ITGA6, ITGB4, ITGB6, IVL, JAG2, JUP, KCNS3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, KRTCAP3, LAMP3, LAPTM5, LGALS7B, LRP11, LRRC4, LSP1, LSR, LYPD3, MAGEA4, MAGEA6, MAL2, MAOA, MARCO, MCM2, MDFI, MET, MMP14, MPZL2, MUC1, MYBL2, MYH14, MYOF, MZB1, NCF2, NKG7, NNMT, NOTCH3, NRARP, NTRK2, NUP210, NUSAP1, OSGIN1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PERP, PIGR, PIGT, PITX1, PKP1, PKP3, PLEK, PLEK2, PLVAP, PMP22, POSTN, PPL, PPP1R14C, PRAME, PROM2, PRRG4, PRSS8, PTGES, PTGFRN, PTPN6, PTPRF, PTPRZ1, RAB25, RAB38, RAET1L, RARRES1, RBP1, RGS1, RHCG, RHOV, RIN2, RIPK4, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, S100A2, S100P, SCNN1A, SDC1, SERINC2, SERPINB13, SERPINB3, SERPINB5, SEZ6L2, SH3BP4, SHISA2, SLC1A5, SLC2A1, SLC34A2, SLC40A1, SLC6A8, SLC7A8, SNAI2, SOX2, SPI1, SPINT1, SPINT2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, THBD, THBS2, TK1, TM4SF1, TMC4, TMEM30B, TMEM54, TMPRSS11D, TMPRSS11E, TMPRSS4, TNFRSF18, TNS4, TOP2A, TP53I11, TP63, TPD52, TPX2, TREM2, TRIM29, TRIP13, TSPAN1, TSPAN13, TSPAN6, TSPAN7, TUSC3, TYROBP, UBE2C, UPK1B, VAMP8, VANGL2, WLS, YAP1, ZC3H11A,

ZNF217, ZNF750, and combinations thereof; and (ii) by using a set of detection probes (*e.g.*, as utilized and/or described herein), at least one of which are directed to one or more (*e.g.*, 1, 2, 3, or more) of EV surface biomarkers described in Example 3.

**[531]** In some such embodiments, a set of detection probes comprises at least one detection probe directed to an EV surface biomarker. In some such embodiments, a set of detection probes comprises at least two detection probes directed to the same EV surface biomarker (with the same or different epitopes). In some such embodiments, a set of detection probes comprises at least two detection probes directed to distinct EV surface biomarkers. In some embodiments, a sample comprising an EV surface biomarker and intravesicular mRNA can be contacted with an anti-EV surface biomarker affinity agent (*e.g.*, an antibody directed to EV surface biomarker as described in Example 3) conjugated to a single-stranded oligonucleotide (*e.g.*, DNA) that serves as one of two primers in a pair for an intravesicular mRNA biomarker (*e.g.*, described in Example 4) such that the anti-EV surface biomarker antibody is bound to the EV surface biomarker while the conjugated single-stranded oligonucleotide is hybridized with the intravesicular mRNA biomarker present in the same sample. A second primer of the pair and an RT-qPCR probe are then added to perform an RT-qPCR for detection of the presence of an intravesicular mRNA and an EV surface biomarker in a single sample.

**[532]** In some embodiments, captured EVs can be read out by detection of at least one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof; and at least one or more (*e.g.*, 1, 2, 3, or more) of EV intravesicular biomarkers described in Example 5. In some such embodiments, captured EVs can be read out (i) by detection of one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs using RT-qPCR: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4,

CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof; and (ii) by using a set of detection probes (*e.g.*, as utilized and/or described herein), at least one of which are directed to one or more (*e.g.*, 1, 2, 3, or more) of intravesicular biomarkers described in Example 5.

**[533]** In some embodiments, captured EVs can be read out by detection of at least one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs: ABCA3, ABCC1, ABRACL, ACP5, ADAM23, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, ANTXR1, AP1M2, APOBEC3B, APOBEC3C, AQP3, AREG, ARNTL2, ASF1B, ATP8B1, AURKB, B3GNT5, BAIAP2L1, BCAM, BIK, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA12, CA9, CALML3, CAPNS2, CBLC, CCL19, CCL5, CCNB2, CD109, CD24, CD53, CD74, CD9, CDC20, CDC42EP1, CDC45, CDCA4, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CEACAM5, CEACAM6, CELSR1, CENPW, CEP55, CES1, CHMP4C, CLCA2, CLDN1, CLDN4, CLDN7, CNN2, COL17A1, CPA3, CRABP2, CSTA, CTSC, CTSE, CX3CL1, CXADR, CXCR4, CYBB, CYP2S1, CYP4F11, DAPL1, DPYSL3, DSC2, DSC3, DSG2, DSG3, DSP, EFNA1, EFS, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHB3, EPHX1, ESRP1, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83D, FAM83H, FAT1, FAT2, FBLIM1, FBP1, FCER1G, FERMT1, FGFR2, FGFR3, FOXE1, FOXM1, FXYD3, GALNT3, GBP6, GJA1, GJB2, GJB3, GJB5, GJB6, GNA15, GPC1, GPC3, GPNMB, GPR87, GPRC5A, GPX2, GRHL2, GSTA1, HAS3, HCK, HOXB7, ID1, IGF2BP2, IGSF9, IL2RG, IMPA2, IRF6, ITGA2, ITGA6, ITGB4, ITGB6, IVL, JAG2, JUP, KCNS3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, KRTCAP3, LAMP3, LAPTM5, LGALS7B, LRP11, LRRC4, LSP1, LSR, LYPD3, MAGEA4, MAGEA6, MAL2, MAOA, MARCO, MCM2, MDFI, MET, MMP14, MPZL2, MUC1, MYBL2, MYH14, MYOF, MZB1, NCF2, NKG7, NNMT, NOTCH3, NRARP, NTRK2, NUP210, NUSAP1, OSGIN1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PERP, PIGR,

PIGT, PITX1, PKP1, PKP3, PLEK, PLEK2, PLVAP, PMP22, POSTN, PPL, PPP1R14C, PRAME, PROM2, PRRG4, PRSS8, PTGES, PTGFRN, PTPN6, PTPRF, PTPRZ1, RAB25, RAB38, RAET1L, RARRES1, RBP1, RGS1, RHCG, RHOV, RIN2, RIPK4, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, S100A2, S100P, SCNN1A, SDC1, SERINC2, SERPINB13, SERPINB3, SERPINB5, SEZ6L2, SH3BP4, SHISA2, SLC1A5, SLC2A1, SLC34A2, SLC40A1, SLC6A8, SLC7A8, SNAI2, SOX2, SPI1, SPINT1, SPINT2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, THBD, THBS2, TK1, TM4SF1, TMC4, TMEM30B, TMEM54, TMPRSS11D, TMPRSS11E, TMPRSS4, TNFRSF18, TNS4, TOP2A, TP53I11, TP63, TPD52, TPX2, TREM2, TRIM29, TRIP13, TSPAN1, TSPAN13, TSPAN6, TSPAN7, TUSC3, TYROBP, UBE2C, UPK1B, VAMP8, VANGL2, WLS, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof; and at least one or more (*e.g.*, 1, 2, 3, or more) of EV intravesicular biomarkers described in Example 5. In some such embodiments, captured EVs can be read out (i) by detection of one or more (*e.g.*, 1, 2, 3, or more) of the following mRNAs using RT-qPCR: ABCA3, ABCC1, ABRACL, ACP5, ADAM23, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, ANTXR1, AP1M2, APOBEC3B, APOBEC3C, AQP3, AREG, ARNTL2, ASF1B, ATP8B1, AURKB, B3GNT5, BAIAP2L1, BCAM, BIK, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA12, CA9, CALML3, CAPNS2, CBLC, CCL19, CCL5, CCNB2, CD109, CD24, CD53, CD74, CD9, CDC20, CDC42EP1, CDC45, CDCA4, CDCA5, CDCP1, CDH1, CDH3, CDK1, CDKN2A, CDKN2B, CEACAM5, CEACAM6, CELSR1, CENPW, CEP55, CES1, CHMP4C, CLCA2, CLDN1, CLDN4, CLDN7, CNN2, COL17A1, CPA3, CRABP2, CSTA, CTSC, CTSE, CX3CL1, CXADR, CXCR4, CYBB, CYP2S1, CYP4F11, DAPL1, DPYSL3, DSC2, DSC3, DSG2, DSG3, DSP, EFNA1, EFS, EGFR, EGLN3, EHD2, EHF, ELF3, ELF4, EMP1, EMP2, ENAH, EPCAM, EPHA2, EPHB3, EPHX1, ESRP1, EVPL, F11R, F2R, F2RL1, F3, FAM129B, FAM60A, FAM83D, FAM83H, FAT1, FAT2, FBLIM1, FBP1, FCER1G, FERMT1, FGFR2, FGFR3, FOXE1, FOXM1, FXYD3, GALNT3, GBP6, GJA1, GJB2, GJB3, GJB5, GJB6, GNA15, GPC1, GPC3, GPNMB, GPR87, GPRC5A, GPX2, GRHL2, GSTA1, HAS3, HCK, HOXB7, ID1, IGF2BP2, IGSF9, IL2RG, IMPA2, IRF6, ITGA2, ITGA6, ITGB4, ITGB6, IVL, JAG2, JUP, KCNS3, KIAA1522, KIF2C, KIFC1, KITLG, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8,

KRTCAP3, LAMP3, LAPTM5, LGALS7B, LRP11, LRRC4, LSP1, LSR, LYPD3, MAGEA4, MAGEA6, MAL2, MAOA, MARCO, MCM2, MDFI, MET, MMP14, MPZL2, MUC1, MYBL2, MYH14, MYOF, MZB1, NCF2, NKG7, NNMT, NOTCH3, NRARP, NTRK2, NUP210, NUSAP1, OSGIN1, OSMR, PALLD, PDPN, PDZK1IP1, PECAM1, PERP, PIGR, PIGT, PITX1, PKP1, PKP3, PLEK, PLEK2, PLVAP, PMP22, POSTN, PPL, PPP1R14C, PRAME, PROM2, PRRG4, PRSS8, PTGES, PTGFRN, PTPN6, PTPRF, PTPRZ1, RAB25, RAB38, RAET1L, RARRES1, RBP1, RGS1, RHCG, RHOV, RIN2, RIPK4, RPS4Y1, RRM2, S100A10, S100A11, S100A14, S100A16, S100A2, S100P, SCNN1A, SDC1, SERINC2, SERPINB13, SERPINB3, SERPINB5, SEZ6L2, SH3BP4, SHISA2, SLC1A5, SLC2A1, SLC34A2, SLC40A1, SLC6A8, SLC7A8, SNAI2, SOX2, SPI1, SPINT1, SPINT2, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, ST14, STEAP1, SULF1, SYK, SYTL1, TACSTD2, TBC1D2, TEAD2, TEAD3, TFAP2C, THBD, THBS2, TK1, TM4SF1, TMC4, TMEM30B, TMEM54, TMPRSS11D, TMPRSS11E, TMPRSS4, TNFRSF18, TNS4, TOP2A, TP53I11, TP63, TPD52, TPX2, TREM2, TRIM29, TRIP13, TSPAN1, TSPAN13, TSPAN6, TSPAN7, TUSC3, TYROBP, UBE2C, UPK1B, VAMP8, VANGL2, WLS, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof; and (ii) by using a set of detection probes (*e.g.*, as utilized and/or described herein), at least one of which are directed to one or more (*e.g.*, 1, 2, 3, or more) of intravesicular biomarkers described in Example 5.

**[534]** In some embodiments, captured EVs can be read out by detection of at least one (e.g., 1, 2, 3, or more) mRNA biomarker described above; and at least one (e.g., 1, 2, 3, or more) EV intravesicular biomarkers described in Example 5. In some such embodiments, captured EVs can be read out (i) by detection of one or more (e.g., 1, 2, 3, or more) mRNAs; and (ii) by using a set of detection probes (e.g., as utilized and/or described herein), at least one of which are directed to one or more (e.g., 1, 2, 3, or more) intravesicular biomarkers described in Example 5. In some embodiments, a set of detection probes comprises at least one detection probe directed to an intravesicular biomarker (*e.g.*, as described herein). In some embodiments, a set of detection probes comprises at least two detection probes each directed to the same intravesicular biomarker (*e.g.*, with the same epitope or different epitopes). In some embodiments, a set of detection probes comprises at least two detection probes each directed to a distinct intravesicular biomarker (*e.g.*, as described herein). In some embodiments, a sample comprising EV intravesicular biomarker and intravesicular mRNA) can be contacted with an anti-EV

intravesicular biomarker affinity agent (*e.g.*, an antibody directed to EV intravesicular biomarker as described in Example 5) conjugated to a single-stranded oligonucleotide (*e.g.*, DNA) that serves as one of two primers in a pair for an intravesicular mRNA biomarker (*e.g.*, described in Example 4) such that the anti-EV intravesicular biomarker affinity agent is bound to the EV intravesicular biomarker while the conjugated single-stranded oligonucleotide is hybridized with the intravesicular mRNA biomarker present in the same sample. A second primer of the pair and an RT-qPCR probe are then added to perform an RT-qPCR for detection of the presence of an intravesicular mRNA and an intravesicular biomarker in a single sample.

**Example 5: Assessment of intravesicular biomarkers as lung cancer biomarkers**

**[535]** In some embodiments, lung cancer detection includes detection of at least intravesicular biomarker(s) following immunoaffinity capture of extracellular vesicles.

**[536]** In some embodiments, one or more surface biomarkers or extracellular membrane biomarkers that are present on extracellular vesicles ("capture biomarkers"), for example, capture biomarkers as described in Example 3, can be used for immunoaffinity capture of lung cancer-associated extracellular vesicles.

**[537]** In some embodiments, EV immunoassay methodology (*e.g.*, ones described herein such as in Example 1) and biomarker-validation process (*e.g.*, ones described herein such as in Example 1) can be used to assess intravesicular biomarkers for lung cancer. In some embodiments, an antibody directed to a capture biomarker (*e.g.*, a surface biomarker present in lung cancer-associated EVs) is conjugated to magnetic beads and evaluated, first on cell-line EVs then on patient samples, for its ability to bind the specific target biomarker. The antibody-coated bead is assessed for its ability to capture lung cancer-associated EVs and the captured EVs by the antibody-coated beads are fixed and/or permeabilized prior to being profiled for their intravesicular biomarkers using a target entity detection system (*e.g.*, a duplex system as described herein involving a set of two detection probes, each directed to a target marker that is distinct from the capture protein).

**[538]** In some embodiments, captured EVs after fixation and/or permeabilization can be read out using at least one or more (*e.g.*, 1, 2, 3, or more) of the following intravesicular biomarkers: AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof. In some

embodiments, captured EVs after fixation and/or permeabilization can be read out using a set of detection probes (*e.g.*, as utilized and/or described herein), at least two of which are directed to one or more (*e.g.*, 1, 2, 3, or more) of the following intravesicular biomarkers: AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof. In some embodiments, a set of detection probes comprises two detection probes each directed to the same intravesicular biomarker. In some embodiments, a set of detection probes comprises two detection probes each directed to a distinct intravesicular biomarker.

**[539]** In some embodiments, captured EVs after fixation and/or permeabilization can be read out using at least one or more (*e.g.*, 1, 2, 3, or more) of the following intravesicular biomarkers: ABRACL, ACP5, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, AP1M2, APOBEC3B, APOBEC3C, ARNTL2, ASF1B, AURKB, BAIAP2L1, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA9, CALML3, CAPNS2, CBLC, CCL19, CCNB2, CDC20, CDC45, CDCA4, CDCA5, CDK1, CDKN2A, CDKN2B, CENPW, CEP55, CES1, CHMP4C, CNN2, CPA3, CRABP2, CSTA, CTSC, CTSE, CYP2S1, DPYSL3, EFS, EGLN3, EHF, ELF3, ELF4, ENAH, ESRP1, EVPL, FAM129B, FAM60A, FAM83D, FAM83H, FBP1, FERMT1, FOXE1, FOXM1, GBP6, GNA15, GPX2, GRHL2, GSTA1, HCK, HOXB7, ID1, IGF2BP2, IMPA2, IRF6, IVL, JUP, KIAA1522, KIF2C, KIFC1, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, LGALS7B, LSP1, MAGEA4, MAGEA6, MCM2, MDFI, MYBL2, MYH14, MZB1, NCF2, NNMT, NRARP, NUP210, NUSAP1, OSGIN1, PALLD, PITX1, PKP1, PKP3, PLEK, PLEK2, POSTN, PPP1R14C, PRAME, PTPN6, RBP1, RIN2, RIPK4, RPS4Y1, RRM2, S100A11, S100A14, S100A16, S100A2, S100P, SERPINB13, SERPINB3, SERPINB5, SH3BP4, SNAI2, SOX2, SPI1, SPINT1, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, SULF1, SYK, SYTL1, TBC1D2, TEAD2, TEAD3, TFAP2C, THBS2, TK1, TOP2A, TP63, TPD52, TPX2, TRIM29, TRIP13, UBE2C, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof. In some embodiments, captured EVs after fixation and/or permeabilization can be read out using a set of detection probes (*e.g.*, as utilized and/or described herein), at least two of which are directed to one or more (*e.g.*, 1, 2, 3, or more) of the following intravesicular biomarkers: ABRACL, ACP5, ADH7, AGR2, AIF1, AKR1C1,

AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, AP1M2, APOBEC3B, APOBEC3C, ARNTL2, ASF1B, AURKB, BAIAP2L1, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA9, CALML3, CAPNS2, CBLC, CCL19, CCNB2, CDC20, CDC45, CDCA4, CDCA5, CDK1, CDKN2A, CDKN2B, CENPW, CEP55, CES1, CHMP4C, CNN2, CPA3, CRABP2, CSTA, CTSC, CTSE, CYP2S1, DPYSL3, EFS, EGLN3, EHF, ELF3, ELF4, ENAH, ESRP1, EVPL, FAM129B, FAM60A, FAM83D, FAM83H, FBP1, FERMT1, FOXE1, FOXM1, GBP6, GNA15, GPX2, GRHL2, GSTA1, HCK, HOXB7, ID1, IGF2BP2, IMPA2, IRF6, IVL, JUP, KIAA1522, KIF2C, KIFC1, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, LGALS7B, LSP1, MAGEA4, MAGEA6, MCM2, MDFI, MYBL2, MYH14, MZB1, NCF2, NNMT, NRARP, NUP210, NUSAP1, OSGIN1, PALLD, PITX1, PKP1, PKP3, PLEK, PLEK2, POSTN, PPP1R14C, PRAME, PTPN6, RBP1, RIN2, RIPK4, RPS4Y1, RRM2, S100A11, S100A14, S100A16, S100A2, S100P, SERPINB13, SERPINB3, SERPINB5, SH3BP4, SNAI2, SOX2, SPI1, SPINT1, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, SULF1, SYK, SYTL1, TBC1D2, TEAD2, TEAD3, TFAP2C, THBS2, TK1, TOP2A, TP63, TPD52, TPX2, TRIM29, TRIP13, UBE2C, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof. In some embodiments, a set of detection probes comprises two detection probes each directed to the same intravesicular biomarker. In some embodiments, a set of detection probes comprises two detection probes each directed to a distinct intravesicular biomarkers.

**[540]** In some embodiments, captured EVs after fixation and/or permeabilization can be read out using (i) at least one or more (*e.g.*, 1, 2, 3, or more) of the following intravesicular biomarkers: AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and (ii) at least one or more (*e.g.*, 1, 2, 3, or more) of EV surface biomarkers described in Example 3. In some embodiments, captured EVs after fixation and/or permeabilization can be read out using a set of detection probes (*e.g.*, as utilized and/or described herein), which comprises (i) a first detection probe directed to one or more (*e.g.*, 1, 2, 3, or more) of the following intravesicular biomarkers: AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB,

SPINK1, TGFA, ZC3H11A, and combinations thereof; and (ii) a second detection probe directed to one or more (*e.g.*, 1, 2, 3, or more) of EV surface biomarkers described in Example 3.

**[541]** In some embodiments, captured EVs after fixation and/or permeabilization can be read out using (i) at least one or more (*e.g.*, 1, 2, 3, or more) of the following intravesicular biomarkers: ABRACL, ACP5, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, AP1M2, APOBEC3B, APOBEC3C, ARNTL2, ASF1B, AURKB, BAIAP2L1, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA9, CALML3, CAPNS2, CBLC, CCL19, CCNB2, CDC20, CDC45, CDCA4, CDCA5, CDK1, CDKN2A, CDKN2B, CENPW, CEP55, CES1, CHMP4C, CNN2, CPA3, CRABP2, CSTA, CTSC, CTSE, CYP2S1, DPYSL3, EFS, EGLN3, EHF, ELF3, ELF4, ENAH, ESRP1, EVPL, FAM129B, FAM60A, FAM83D, FAM83H, FBP1, FERMT1, FOXE1, FOXM1, GBP6, GNA15, GPX2, GRHL2, GSTA1, HCK, HOXB7, ID1, IGF2BP2, IMPA2, IRF6, IVL, JUP, KIAA1522, KIF2C, KIFC1, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, LGALS7B, LSP1, MAGEA4, MAGEA6, MCM2, MDFI, MYBL2, MYH14, MZB1, NCF2, NNMT, NRARP, NUP210, NUSAP1, OSGIN1, PALLD, PITX1, PKP1, PKP3, PLEK, PLEK2, POSTN, PPP1R14C, PRAME, PTPN6, RBP1, RIN2, RIPK4, RPS4Y1, RRM2, S100A11, S100A14, S100A16, S100A2, S100P, SERPINB13, SERPINB3, SERPINB5, SH3BP4, SNAI2, SOX2, SPI1, SPINT1, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, SULF1, SYK, SYTL1, TBC1D2, TEAD2, TEAD3, TFAP2C, THBS2, TK1, TOP2A, TP63, TPD52, TPX2, TRIM29, TRIP13, UBE2C, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof; and (ii) at least one or more (*e.g.*, 1, 2, 3, or more) of EV surface biomarkers described in Example 3. In some embodiments, captured EVs after fixation and/or permeabilization can be read out using a set of detection probes (*e.g.*, as utilized and/or described herein), which comprises (i) a first detection probe directed to one or more (*e.g.*, 1, 2, 3, or more) of the following intravesicular biomarkers: ABRACL, ACP5, ADH7, AGR2, AIF1, AKR1C1, AKR1C2, AKR1C3, ALDH1A1, ALDH3A1, ALDH3B2, ALG1L, AP1M2, APOBEC3B, APOBEC3C, ARNTL2, ASF1B, AURKB, BAIAP2L1, BIRC5, C15orf48, C19orf33, C1S, C8orf4, CA9, CALML3, CAPNS2, CBLC, CCL19, CCNB2, CDC20, CDC45, CDCA4, CDCA5, CDK1, CDKN2A, CDKN2B, CENPW, CEP55, CES1, CHMP4C, CNN2, CPA3, CRABP2, CSTA, CTSC, CTSE, CYP2S1, DPYSL3, EFS, EGLN3, EHF, ELF3, ELF4, ENAH, ESRP1, EVPL, FAM129B, FAM60A, FAM83D, FAM83H, FBP1, FERMT1,

FOXE1, FOXM1, GBP6, GNA15, GPX2, GRHL2, GSTA1, HCK, HOXB7, ID1, IGF2BP2, IMPA2, IRF6, IVL, JUP, KIAA1522, KIF2C, KIFC1, KLF4, KLF5, KRT13, KRT14, KRT15, KRT16, KRT17, KRT18, KRT19, KRT5, KRT6A, KRT6B, KRT6C, KRT7, KRT8, LGALS7B, LSP1, MAGEA4, MAGEA6, MCM2, MDFI, MYBL2, MYH14, MZB1, NCF2, NNMT, NRARP, NUP210, NUSAP1, OSGIN1, PALLD, PITX1, PKP1, PKP3, PLEK, PLEK2, POSTN, PPP1R14C, PRAME, PTPN6, RBP1, RIN2, RIPK4, RPS4Y1, RRM2, S100A11, S100A14, S100A16, S100A2, S100P, SERPINB13, SERPINB3, SERPINB5, SH3BP4, SNAI2, SOX2, SPI1, SPINT1, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, SPRR3, SULF1, SYK, SYTL1, TBC1D2, TEAD2, TEAD3, TFAP2C, THBS2, TK1, TOP2A, TP63, TPD52, TPX2, TRIM29, TRIP13, UBE2C, YAP1, ZC3H11A, ZNF217, ZNF750, and combinations thereof; and (ii) a second detection probe directed to one or more (*e.g.*, 1, 2, 3, or more) of EV surface biomarkers described in Example 3.

**[542]** In some embodiments, captured EVs after fixation and/or permeabilization can be read out by detecting an EV intravesicular biomarker and an EV intravesicular mRNA together in a single sample as described in Example 4 above.

**[543]** The present Example further demonstrates exemplary methods for detection of at least one (e.g., 1, 2, 3, or more) intravesicular RNA biomarker in extracellular vesicles derived from cancer cell lines. In some embodiments, such a method comprises immunoaffinity capture of extracellular vesicles as described herein (e.g., via a surface-bound protein such as a surface biomarker described herein), followed by detection of intravesicular RNA, for example, by reverse-transcription qPCR (RT-qPCR). In some embodiments, extracellular vesicles are captured by a cancer-associated surface biomarker, e.g., in some embodiments using antibody-functionalized solid substrate (e.g., magnetic beads). In some embodiments, captured extracellular vesicles are lysed to release their nucleic acid cargo prior to detection of intravesicular RNA. In some embodiments, intravesicular RNA is or comprises mRNA.

**[544]** In some embodiments, cell lines were selected that originate from or are associated with cancer (e.g., a particular cancer type). In some embodiments, such cell lines were selected that originate from or are associated with colon/colorectal cancer, leukemia, melanoma, ovarian cancer, or sarcoma (e.g., rhabdoid tumor). In some embodiments, G-401, K562, NIH:OVCAR-3, SK-MEL-1, or T84 cell lines were selected.

**[545]** In some embodiments, extracellular vesicles were purified from conditioned cell culture medium, counted, immunoaffinity captured, and washed via methods as described herein (e.g., as described in Example 1).

**[546]** Each RT-qPCR reaction mixture included a PCR reaction mixture (e.g., 50% (volume) Luna One-Step reaction mix, 5% (volume) Luna WarmStart RT enzyme mix, 5% (volume) primer-TaqMan probe mixture), and a variable combination of water, captured extracellular vesicles, and lysing agent. RT-qPCR was performed, for example, on the Quant Studio 3, with a suitable PCR protocol, e.g., hold at 55°C for 10 minutes, hold at 95°C for 1 minute, perform 50 cycles of 95°C for 5 seconds and 62°C for 15 seconds, and standard melt curve. The rate of temperature change was chosen to be standard (2°C per second). All qPCRs were performed in doublets or triplets and ROX was used as the passive reference to normalize the qPCR signals. Data was then downloaded from the Quant Studio 3 machine and analyzed and plotted in Python 3.7. Primers and TaqMan probes for each gene were purchased from Integrated DNA Technologies (IDT) as a 20X concentrate.

**[547]** As an initial experiment, *MIF* mRNA was found to be detected in 5e7 bulk extracellular vesicles that were lysed with 1% IGEPAL. Table I shows *MIF* expression in transcript per million (TPM) from different cell lines.

**[548]** **Table I: Summary of *MIF* mRNA expression for cancer cell lines.** MIF RT-PCR signal (45-Ct) for a panel of bulk cell line EVs with varying gene expression. Samples were tested in singlicate and 5e7 EVs were used per reaction.

| Cell Line | MIF Expression (TPM) |
|---|---|
| SK-MEL-1 | 1302.5 |
| NIH:OVCAR-3 | 1175.3 |
| T84 | 196.8 |
| G-401 | 0.8 |

**[549]** A similar experiment was performed to further demonstrate this approach across different intravesicular RNA biomarkers. Table II shows mRNA transcript expression levels in 5e7 bulk extracellular vesicles from different cell lines and shows that mRNA is detectable in cell-line EVs at levels that are dependent on cell gene expression.

**[550]** **Table II: Summary of expression of four different mRNA transcripts for cancer cell lines**

| mRNA transcript | Cell Line | Expression (TPM) | RT-PCR signal of 5e7 lysed EVs | |
|---|---|---|---|---|
| | | | Replicate 1 | Replicate 2 |
| CLDN6 | OVCAR-3 | 203.4 | 34.0 | 33.9 |
| | SK-MEL-1 | 0.1 | 40.5 | Undetected |
| | No template control | - | Undetected | Undetected |

| mRNA transcript | Cell Line | Expression (TPM) | RT-PCR signal of 5e7 lysed EVs | |
|---|---|---|---|---|
| | | | Replicate 1 | Replicate 2 |
| FAM83A | K-562 | 220.1 | 29.9 | 29.7 |
| | G-401 | 0.1 | Undetected | Undetected |
| | No template control | - | Undetected | Undetected |
| HMGB3 | NCI-H1781 | 759.5 | 27.4 | 27.3 |
| | OVCAR-3 | 293.6 | 37.4 | 37.5 |
| | No template control | - | Undetected | Undetected |
| B3GNT3 | NCI-H1781 | 134.9 | 33.6 | 33.9 |
| | G-401 | 0 | Undetected | Undetected |
| | No template control | - | Undetected | Undetected |

**[551]** Additionally, an experiment was performed to detect the colocalization of at least one intravesicular RNA biomarkers with at least one surface biomarker (e.g., a surface marker that is associated with extracellular vesicles). In some embodiments, extracellular vesicles are captured using antibody-functionalized beads directed to a surface biomarker that is present on the surface of the extracellular vesicles. For example, in the present Example, EPCAM-targeted beads were used to capture extracellular vesicles. Bound extracellular vesicles were lysed and *MIF* mRNA content was quantified via RT-qPCR. In some embodiments, a positive control cell line is selected that expresses a surface biomarker for capture and/or an intravesicular RNA biomarker for detection (e.g., EPCAM+, MIF+), while a negative control cell line is selected that does not express a target biomarker for capture and/or detection (e.g., EPCAM-, MIF+). NIH:OVCAR-3 was selected as a positive control cell line and SK-MEL-1 was selected as a negative control cell line. Multiple detergent conditions were also assessed in this experiment to assess the effect of detergent (e.g., Tween-20) concentration on assay performance. These data indicate that reducing detergent (e.g., Tween-20) concentration can improve assay performance. Without wishing to be bound by a particular theory, this effect may likely be due to preservation of membrane integrity during extracellular vesicle capture, as Tween-20 may permeabilize membranes.

**[552]** The present Example demonstrates that intravesicular RNA can be detected via RT-qPCR. In particular, the present Example demonstrates that colocalization of surface biomarkers and intravesicular RNA in extracellular vesicles can be detected by immunoaffinity capture via a surface biomarker followed by RT-qPCR analysis of intravascular RNA.

**Example 6: Development of a lung cancer liquid biopsy assay**

**[553]** The present Example describes development of a lung cancer liquid biopsy assay, for example, for screening hereditary-, environmental-, and average-risk individuals. Despite being the leading killer of men and women among all cancers (Barta *et al.*, 2019; which is incorporated herein by reference for the purpose described herein), while low-dose chest

tomography (LDCT) is endorsed by the United States Preventative Services Task Force (USPSTF), there is currently no low cost, recommended, and highly accessible lung cancer screening tool (Lung Cancer Screening, PDQ®, 2020). This is due, in part, to the poor performance of proposed lung cancer screening technologies such as X-ray and/or sputum analysis (Oken *et al.,* 2011; which is incorporated herein by reference for the purpose described herein). Given the incidence of lung cancer, inadequate test specificities (<99.5%) result in false positive results and may lead to identification of some lung cancers with low aggressiveness. This places a significant burden on the healthcare system and on individuals being screened as false positive results lead to additional tests, unnecessary surgeries, and emotional/physical distress. As a result, it may be desirable to develop a lung cancer screening test that may exhibit two features to provide clinical utility: (1) ultrahigh specificity (>99.5%) to minimize the number of false positives, and (2) high sensitivity (>40%) for stage I and II lung cancer when prognosis is most favorable. The development of such a test has the potential to save tens of thousands of lives each year.

**[554]** Several different biomarker classes have been studied for a lung cancer liquid biopsy assay including circulating tumor DNA (ctDNA), circulating tumor cells (CTCs), bulk proteins, and extracellular vesicles (EVs). EVs are particularly promising due to their abundance and stability in the bloodstream relative to ctDNA and CTCs, suggesting improved sensitivity for early-stage cancers. Moreover, EVs contain cargo (*e.g.*, proteins, RNA, metabolites) that originated from the same cell, providing superior specificity over bulk protein measurements. While the diagnostic utility EVs has been studied, much of this work has pertained to bulk EV measurements or low-throughput single-EV analyses.

**[555]** This present Example describes one aspect of an exemplary approach for early stage lung cancer detection through the profiling of individual extracellular vesicles (EVs) in human plasma. EVs, including exosomes and microvesicles, contain co-localized proteins, RNAs, metabolites, and other compounds representative of their cell of origin (Kosaka *et al.*, 2019; which is incorporated herein by reference for the purpose described herein). The detection of strategically chosen co-localized markers within a single EV can enable the identification of cell type with ultrahigh specificity, including the ability to distinguish cancer cells from normal tissues. As opposed to other cancer diagnostic approaches that rely on cell death for biomarkers to enter the blood (*i.e.*, cfDNA), EVs are released at a high rate by functioning cells. Single cells

have been shown to release as many as 10,000 EVs per day *in vitro* (Balaj *et al.*, 2011; which is incorporated herein by reference for the purpose described herein). In addition, it is widely accepted that cancer cells release EVs at a higher rate than healthy cells (Bebelman *et al.* 2018; which is incorporated herein by reference for the purpose described herein).

**[556]** In one aspect, the present disclosure provides insights and technologies involving identification of genes that are upregulated in lung cancer versus healthy tissues using Applicant's proprietary bioinformatic biomarker discovery process. From a list of upregulated biomarkers, biomarker combinations that are predicted to exhibit high sensitivity and specificity for lung cancer are designed. Using an exemplary individual EV assay (see, *e.g.*, illustrated in ***FIG. 1 or 2*** and/or described herein), co-localization of such biomarkers on an individual vesicle is detected, indicating that the grouping of biomarkers originated from the same cell. This provides superior specificity to bulk biomarker measurements, including bulk EV assays, given that many upregulated cancer biomarkers are expressed by one or more healthy tissues. Through the careful design of biomarker combinations, signals from competing tissues can be reduced or eliminated, including those closely related to lung cancer. In some embodiments, the present disclosure provides technologies with high or ultrahigh specificity that is particularly helpful as a lung cancer screening test for which a positive-predictive value of at least comparable to that of a chest CT, which is the gold standard for screening high risk smokers, is sought. See, *e.g.*, National Lung Screening Trial Research Team (2013) "Results of initial low-dose computed tomographic screening for lung cancer" *New England Journal of Medicine*, *368*(21):1980-1991. In some embodiments, the present disclosure provides technologies useful for lung cancer screening with a positive predictive value of at least 3.8 % or higher (including, *e.g.,* at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 15%, at least 20%, at least 25% or higher) in a high-risk population.

*Biomarker Discovery*

**[557]** In some embodiments, a biomarker discovery process leverages bioinformatic analysis of large databases and an understanding of the biology of lung cancer and extracellular vesicles.

*Individual Extracellular Vesicle Analysis*

**[558]** The detection of tumor-derived EVs in the blood requires an assay that has sufficient selectivity and sensitivity to detect relatively few tumor-derived EVs per milliliter of

plasma in a background of 10 billion EVs from a diverse range of healthy tissues. The present disclosure, among other things, provides technologies that address this challenge. For example, in some embodiments, an assay for individual extracellular vesicle analysis is illustrated in ***FIG. 1***, which is performed in three key steps as outlined below:

1. EVs are purified from patient plasma using size-exclusion chromatography (SEC), which removes greater than 99% of soluble proteins and other interfering compounds.
2. Tumor-specific EVs are captured using antibody-functionalized magnetic beads specific to a membrane-associated surface biomarker.
3. A modified version of proximity-ligation-immuno quantitative polymerase chain reaction (pliq-PCR) is performed to determine the co-localization of additional protein biomarkers contained on or within the captured EVs.

**[559]** In many embodiments of a modified version of a pliq-PCR assay, two or more different antibody-oligonucleotide conjugates are added to the EVs captured by the antibody-functionalized magnetic bead and the antibodies subsequently bind to their biomarker targets. The oligonucleotides are composed of dsDNA with single-stranded overhangs that are complementary, and thus, capable of hybridizing when in close proximity (*i.e.*, when the corresponding biomarker targets are located on the same EV). After washing away unbound antibody-oligonucleotide species, adjacently bound antibody-oligonucleotide species are ligated using a standard DNA ligase reaction. Subsequent qPCR of the ligated template strands enables the detection and relative quantification of co-localized biomarker species. In some embodiments, two to twenty distinct antibody-oligonucleotide probes can be incorporated into such an assay, *e.g.*, as described in U.S. Application No. 16/805,637 (issued as US11,085,089), and International Application PCT/US2020/020529 (published as WO2020180741), both filed February 28, 2020 and entitled "Systems, Compositions, and Methods for Target Entity Detection"; which are both incorporated herein by reference in their entirety for any purpose.

**[560]** pliq-PCR has numerous advantages over other technologies to profile EVs. For example, pliq-PCR has a sensitivity three orders of magnitude greater than other standard immunoassays, such as ELISAs (Darmanis *et al.*, 2010; which is incorporated herein by reference for the purpose described herein). The ultra-low LOD of a well-optimized pliq-PCR reaction enables detection of trace levels of tumor-derived EVs, down to a thousand EVs per mL. This compares favorably with other emerging EV analysis technologies, including the

Nanoplasmic Exosome (nPLEX) Sensor (Im *et al.*, 2014; which is incorporated herein by reference for the purpose described herein) and the Integrated Magnetic-Electrochemical Exosome (iMEX) Sensor (Jeong *et al.*, 2016; which is incorporated herein by reference for the purpose described herein), which have reported LODs of ~$10^3$ and ~$10^4$ EVs, respectively (Shao *et al.*, 2018; which is incorporated herein by reference for the purpose described herein). Moreover, in some embodiments, a modified version of pliq-PCR approach does not require complicated equipment and can uniquely detect the co-localization of multiple biomarkers on individual EVs.

**[561]** In some embodiments, to further improve the sensitivity and specificity of an individual EV profiling assay, other classes of EV biomarkers include mRNA and intravesicular biomarkers (in addition to EV surface biomarkers) can be identified and included in an assay.

*Preliminary Work*

**[562]** Through preliminary studies, a workflow was developed in which biomarker candidates are validated to be present in EVs and capable of being detected by commercially available antibodies or mRNA primer-probe sets. For a given biomarker of interest, one or more cell lines expressing (positive control) and not expressing the biomarker of interest (negative control) can be cultured to harvest their EVs through concentrating their cell culture media and performing purification to isolate nanoparticles having a size range of interest (*e.g.*, using SEC). Typically, extracellular vesicles may range from 30 nm to several micrometers in diameter. See, *e.g.*, Chuo *et al.*, "Imaging extracellular vesicles: current and emerging methods" *Journal of Biomedical Sciences* 25: 91 (2018), which is incorporated herein by reference for the purpose described herein, which provides information of sizes for different extracellular vesicle (EV) subtypes: migrasomes (0.5-3 μm), microvesicles (0.1-1 μm), oncosomes (1-10 μm), exomeres (<50 nm), small exosomes (60-80 nm), and large exosomes (90-120 nm). In some embodiments, nanoparticles having a size range of about 30 nm to 1000 nm may be isolated for detection assay. In some embodiments, specific EV subtype(s) may be isolated for detection assay.

**[563]** Through a proprietary biomarker discovery process, four membrane-bound biomarkers (SLC34A2, CEACAM5, CEACAM6, and EPCAM) that are upregulated in LUAD versus healthy tissues were identified and used in proof-of-concept experiments in cell-line EVs and lung cancer patient samples.

**[564]** To detect assay signal from EVs that contain co-localized SLC34A2, CEACAM5, CEACAM6, and/or EPCAM markers, in some embodiments, an assay configuration involving SLC34A2 immunoaffinity capture and two distinct antibody-oligonucleotide probes specific to CEACAM6 was developed. In some embodiments, an assay configuration involving SLC34A2 immunoaffinity capture and two distinct antibody-oligonucleotide probes specific to CEACAM6 and EPCAM was developed. In some embodiments, an assay configuration involving CEACAM5 immunoaffinity capture and two distinct antibody-oligonucleotide probes specific to CEACAM6 and SLC34A2 was developed.

**[565]** Purified cell-line EVs were captured using anti-SLC34A2-functionalized, or anti-EPCAM-functionalized magnetic beads. It was observed that higher expressing cell line exhibited a significant increase in signal relative to the lower expressing cell line. This observation demonstrates that in some embodiments, a single EV profiling assay (*e.g.*, ones described herein) is capable of detecting co-localized membrane-bound biomarkers on single EVs with very high sensitivity.

**[566]** Following the validation of SLC34A2, CEACAM5, CEACAM6, and EPCAM in lung cancer cell-line EVs, a pilot cohort study and expanded cohort study was performed on lung cancer patient plasma samples using an optimized and operator-blinded assay protocol. All plasma samples were purchased from the same source, were processed according to the same blood collection protocol, and patient samples were collected prior to the initiation of any treatment (*i.e.*, treatment naïve). The patient cohorts included in the study of this Example are described in ***FIGS. 4 and 11.*** The results of the clinical pilot study are provided in ***FIGS. 5-10,*** while results for the expanded cohort study are presented in ***FIG. 12***.

**[567]** During the clinical pilot study, specificity was determined by assuming a log-normal distribution around all healthy controls (n=20). To determine the cutoff, the mean of the healthy control distribution was determined, and the appropriate number of standard deviations was added, such that substantially 100% of the values would be expected to be below the cutoff. In some embodiments of an assay configuration involving SLC34A2 immunoaffinity capture and two distinct antibody-oligonucleotide probes specific to CEACAM6, sensitivity of 16.7% for stage I LUAD, 60% for stage II LUAD, 50% for stage III LUAD, and 100% for stage IV LUAD were achieved. In some embodiments of an assay configuration involving SLC34A2 immunoaffinity capture and two distinct antibody-oligonucleotide probes specific to CEACAM6

and EPCAM sensitivities of 20% for stage II LUAD, 50% for stage III LUAD, and 75% for stage IV LUAD were achieved. In some embodiments of an assay configuration involving CEACAM5 immunoaffinity capture and two distinct antibody-oligonucleotide probes specific to CEACAM6 and SLC34A2 sensitivities of 16.7% for stage I LUAD, 20% for stage II LUAD, 50% for stage III LUAD, and 75% for stage IV LUAD were achieved.

**[568]** The predictive capacity of these three exemplary LUAD diagnostic assays were compared to each other and correlations were determined using the Pearson product-moment correlation coefficient. Strong correlations were observed, especially for stage III and stage IV LUAD samples (***FIGS. 8-10***).

**[569]** ***FIG. 8*** is a graphical representation of the correlation between exemplary lung adenocarcinoma diagnostic assays as described herein. Signal from SLC34A2 antibody based capture with CEACAM6 + CEACAM6 detection probes is depicted along the x-axis, while signal from SLC34A2 antibody based capture with CEACAM6 + EpCAM detection probes is depicted along the y-axis. Correlations were determined using the Pearson product-moment correlation coefficient. Strong correlations were observed, especially for stage III and stage IV LUAD samples.

**[570]** ***FIG. 9*** is a graphical representation of the correlation between exemplary lung adenocarcinoma diagnostic assays as described herein. Signal from SLC34A2 antibody based capture with CEACAM6 + CEACAM6 detection probes is depicted along the x-axis, while signal from CEACAM5 antibody based capture with CEACAM6 + SLC34A2 detection probes is depicted along the y-axis. Correlations were determined using the Pearson product-moment correlation coefficient. Strong correlations were observed, especially for stage III and stage IV LUAD samples.

**[571]** ***FIG. 10*** is a graphical representation of the correlation between exemplary lung adenocarcinoma diagnostic assays as described herein. Signal from SLC34A2 antibody based capture with CEACAM6 + EPCAM detection probes is depicted along the x-axis, while signal from CAECAM5 antibody based capture with CEACAM6 + SLC34A2 detection probes is depicted along the y-axis. Correlations were determined using the Pearson product-moment correlation coefficient. Strong correlations were observed, especially for stage III and stage IV LUAD samples.

**[572]** During the clinical expanded cohort study, specificity was determined by assuming a log-normal distribution around all healthy controls (n=90). To determine the cutoff, the mean of the healthy control distribution was determined, and the appropriate number of standard deviations was added, such that 99.9% of the values would be expected to be below the cutoff. In some embodiments of an assay configuration involving SLC34A2 immunoaffinity capture and two distinct antibody-oligonucleotide probes specific to CEACAM6, sensitivity of 50% for stage II LUAD, 55.5% for stage III LUAD, and 71.4% for stage IV LUAD were achieved.

**[573]** A specificity of 98%, 98.5%, 99%, 99.5%, 99.9%, or 100% can be used to evaluate the PPV for screening different LUAD at risk populations, *e.g.*, where prevalence is at ~ per 100,000 for smokers, or ~ per 100,000 for non-smokers.

**[574]** To further improve the performance of an exemplary single EV profile assay (*e.g.*, ones described herein) for detection of lung cancer, in some embodiments, additional biomarker candidates including surface biomarkers and intravesicular mRNAs/biomarkers can be identified.

**[575]** In some embodiments, it was previously demonstrated by Applicant the feasibility of EV-mRNA detection using purified cell-line EVs in bulk. Through immunoaffinity capture of a membrane bound protein marker, this approach enables the detection of two co-localized biomarkers. Moreover, EV-mRNA detection requires a simpler protocol because RT-qPCR can be performed directly after immunoaffinity capture. In some embodiments, mRNA detection using EVs can be performed by capturing EVs using capture probes (*e.g.*, as described herein) and detecting a particular lung cancer mRNA biomarker. EVs that express both capture probe marker and lung cancer mRNA biomarker are selectively detected.

**Example 7: Bioinformatically-identified biomarkers and biomarker combinations**

**[576]** The present Example illustrates an exemplary bioinformatically driven approach for identification of certain biomarkers and biomarker combinations that can be useful for lung cancer diagnosis.

*Bioinformatic filtering*

**[577]** There are more than 55,000 transcripts captured in the Genotype-Tissue Expression (GTEx) database (*e.g.*, a primary data resource for normal tissue gene expression) and the Cancer Genome Atlas (TCGA) database (*e.g.*, a primary data resource for cancer tissue

gene expression). To identify biomarkers that are useful for detection of lung cancer, two filtering steps were applied to the data.

**[578]** In some embodiments, UniProt filter was used. Biomarkers that have a valid UniProt entry, which includes evidence that a biomarker protein was found to be associated with a membrane, were considered in the analysis (*e.g.*, proteins with no evidence of being membrane associated were optionally filtered out). Such a filtering step may optionally distinguish between different membranes of interest or level of confidence of the provided evidence.

**[579]** In some embodiments, Vesiclepedia filter was used. Vesiclepedia (a repository of extracellular vesicle publications) was used to filter the results. Vesiclepedia lists the number of EV related references published for each gene (*e.g.*, Entrez). These references were used as a proxy for presence of a given biomarker in or on EVs. If no EV-related publications exist for a given biomarker, it is less likely to be an actual EV biomarker, and was thus filtered from the list of biomarkers for further consideration.

*Minimum expression and differentiation filtering*

**[580]** In some embodiments, a minimum expression level of a biomarker is considered. Low biomarker expression may produce stochastic noise and make robust signal detection difficult and unreliable. To overcome this challenge, one or more (including all of) of the following expression filters were applied. In particular embodiments, four expression filters were applied.

*Minimal expression in the cancer of interest*

**[581]** In some embodiments, a minimum number of samples were used to show expression levels that were detectable in sample from the cancer of interest (e.g., a plasma sample from a cancer patient), while leaving room for discovery of subtypes that potentially have differential gene expression profiles. To achieve this filter, in some embodiments, the 80th percentile of gene expression in the TCGA cancer of interest (*e.g.*, lung cancer) was calculated, and in some embodiments, biomarkers that have a transcript per million (TPM) value of >15 at the $80^{th}$ percentile were considered.

*Minimal expression associated cell-lines*

**[582]** In some embodiments, positive control cell-lines were utilized for testing of antibodies directed towards bioinformatically-predicted biomarkers. In some embodiments, the Cancer Cell Line Encyclopedia (CCLE) gene expression set, which contains >1000 cell-line

profiles, was utilized to reduce biomarker lists to those for which cell-lines expressing a biomarker of interest exist. In some embodiments, the 90$^{th}$ percentile of expression for each biomarker across cancer-specific cell-lines was calculated, and in some embodiments, biomarkers with a TPM >15 at the 90$^{th}$ percentile were considered.

*Minimum expression in the cancer of interest on a protein level*

**[583]** One skilled in the art will understand that not all genes that are expressed are ultimately translated into proteins. Accordingly, in some embodiments, mass spectrometry data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) were utilized to filter for protein-expressing genes. In some embodiments, biomarkers with a spectral count greater than 10 were considered to be expressed.

*Minimum differentiation against normal tissue*

**[584]** In some embodiments, assays described herein achieved superior specificity by requiring co-expression of at least two biomarkers, and in some embodiments, at least three biomarkers, on the same extracellular vesicle. Simple differential gene expression of normal tissues yielded too many false negative values. Instead, in some embodiments, a biomarker signature comprises a combination of biomarkers that may include biomarkers that were highly expressed in multiple tissue types, but only when they were paired with other biomarkers that provided additional discriminatory power (*e.g.*, highly tissue specific and/or highly cancer specific). However, such an analysis could capture housekeeping genes, such as GAPDH, which were ubiquitously expressed, and accordingly were not necessarily useful as discriminatory biomarkers. To remove such markers, in some embodiments, a z-score comparing cancerous tissue (*e.g.*, lung cancer) and every tissue type in GTEx for a given biomarker was calculated. In some embodiments, a biomarker with a z-score of 5 at the 80th percentile, in at least one normal tissue type was selected (*e.g.*, at least one normal tissue was clearly excluded by a biomarker candidate).

*Simulation and stochastic sampling*

**[585]** Discriminatory power of a biomarker signature candidate or biomarker combination candidate comprising at least two or more (including, *e.g.*, at least three or more) biomarkers can be determined by simulating and comparing expression of such a biomarker signature candidate in normal subjects (*e.g.*, subjects who were determined not to have lung cancer) to that in cancer subjects. Combinations of at least 2 and at least 3 biomarkers were

generated based on filtered biomarker sets. An EV score, which estimated the number of EVs generated by a profiled tissue, was calculated for a given combination by multiplying TPM values of all markers in a given combination.

**[586]** To simulate a population of normal subjects, a cohort of 5000 plasma samples from 5000 "healthy individuals" was created. Individual samples were created by randomly selecting tissue samples from each of the 54 tissues in the GTEx database and multiplying the TPM values of expressed genes with the estimated weight in grams of each organ based on a healthy individual. EV scores were then summed for an individual across tissues to simulate an individual. EV scores were then summed across tissues for a simulated individual. In addition to a healthy cohort, 5000 samples from "cancer individuals" were created by repeating the "healthy" pool generation technique, but with an added step of adding EV scores of randomly selected lung cancer (*e.g.,* LUAD and/or LUSC) samples from TCGA, multiplying the sample by 1, 10, or 100, corresponding to a 1g, a 10g, or a 100g tumor. Using these two sample pools of "healthy" and "cancer" individuals, sensitivity for each biomarker combination candidate at 99% specificity was calculated. This metric was then used to rank biomarker combination candidates.

**[587]** For initial biomarker signature selection, in some embodiments, 1 million combinations of three biomarkers were randomly sampled, and in some embodiments simulations were conducted using a 100g tumor, and 1000 individuals in each of the cancer and the healthy pool. In some embodiments, biomarker combinations were then ranked based on their sensitivity value at 99% specificity. In some embodiments, single biomarkers were then ranked based on the top 0.5 percentile of their rank in the combination list.

**Example 8: Correlation of bioinformatically-identified biomarkers and biomarker combinations with pathways known in the art**

**[588]** The present Example describes a gene set enrichment analysis for determination of overlap between certain bioinformatically-predicted biomarkers and published gene pathways. One skilled in the art will recognize that in certain cases, lists of single genes can be challenging to appropriately interpret. Fortunately, there are resources that provide functional lists of genes, such as, for example, lists of genes that encode proteins that are components of the same biochemical pathway or phenomenon. Comparing a bioinformatically-identified list of biomarkers to known gene sets and biochemical pathways can impose structure on a list of biomarkers.

**[589]** Table 2 shows an enrichment analysis of certain bioinformatically-identified biomarkers when compared to all gene sets in the Molecular Signature Database Category 2 - Cannonical pathways (v.7.4.) from the Broad Institute. This database includes, among other resources, KEGG, Biocarta, and Reactome data. Each p-value is a result of a Chi-square test, comparing a particular gene set with a list of certain bioinformatically-identified biomarkers against the background of all genes in MSigDB C2-CP database. Biomarkers were ranked with the highest overlap first, and in some embodiments, overlaps with a nominal p-value of 0.05 were considered.

**[590]** Table 2 shows certain molecular pathways that are enriched in a list of bioinformatically identified biomarkers, following correction for multiple testing, several molecular pathways exhibited a false discovery rate (FDR) of less than 0.05. Such molecular pathways provide a biological theme for certain bioinformatically identified biomarkers.

**Table 2 - Enrichment analysis of certain bioinformatically identified biomarkers**

| Gene Ontology Pathway, Source and Description | Raw P value | FDR | Pathway Gene # | Exemplary Biomarkers Included |
|---|---|---|---|---|
| REACTOME_APOPTOTIC_CLEAVAGE_OF_CELL_ADHESION_PROTEINS | 0.0E+00 | 0.0E+00 | 11 | CDH1, DSG2, DSG3 |
| REACTOME_MISCELLANEOUS_SUBSTRATES | 0.0E+00 | 0.0E+00 | 12 | CYP2S1, CYP4F11, CYP4F3 |
| REACTOME_CELL_JUNCTION_ORGANIZATION | 0.0E+00 | 0.0E+00 | 92 | CDH1, CDH3, LAMB3, LAMC2, NECTIN1, SDK1, SDK2 |
| REACTOME_SYNTHESIS_OF_LEUKOTRIENES_LT_AND_EOXINS_EX | 0.0E+00 | 0.0E+00 | 21 | ABCC1, CYP4F11, CYP4F3 |
| REACTOME_CELL_CELL_JUNCTION_ORGANIZATION | 0.0E+00 | 0.0E+00 | 65 | CDH1, CDH3, NECTIN1, SDK1, SDK2 |
| REACTOME_CELL_CELL_COMMUNICATION | 0.0E+00 | 0.0E+00 | 130 | CDH1, CDH3, LAMB3, LAMC2, NECTIN1, SDK1, SDK2 |
| PID_A6B1_A6B4_INTEGRIN_PATHWAY | 0.0E+00 | 4.0E-07 | 46 | CD9, CDH1, LAMB3, LAMC2 |
| KEGG_ECM_RECEPTOR_INTERACTION | 0.0E+00 | 1.1E-05 | 84 | ITGA2, ITGB6, LAMB3, LAMC2, SDC1 |
| PID_INTEGRIN4_PATHWAY | 0.0E+00 | 1.9E-05 | 11 | LAMB3, LAMC2 |
| REACTOME_TYPE_I_HEMIDESMOSOME_ASSEMBLY | 0.0E+00 | 1.9E-05 | 11 | LAMB3, LAMC2 |
| REACTOME_LAMININ_INTERACTIONS | 0.0E+00 | 3.5E-05 | 30 | ITGA2, LAMB3, LAMC2 |
| REACTOME_MET_ACTIVATES_PTK2_SIGNALING | 0.0E+00 | 3.5E-05 | 30 | ITGA2, LAMB3, LAMC2 |
| REACTOME_EICOSANOIDS | 0.0E+00 | 9.0E-05 | 12 | CYP4F11, CYP4F3 |

| Gene Ontology Pathway, Source and Description | Raw P value | FDR | Pathway Gene # | Exemplary Biomarkers Included |
|---|---|---|---|---|
| REACTOME_NON_INTEGRIN_MEMBRANE_ECM_INTERACTIONS | 0.0E+00 | 9.2E-05 | 59 | ITGA2, LAMB3, LAMC2, SDC1 |
| REACTOME_ADHERENS_JUNCTIONS_INTERACTIONS | 1.0E-07 | 2.0E-04 | 33 | CDH1, CDH3, NECTIN1 |
| REACTOME_APOPTOTIC_CLEAVAGE_OF_CELLULAR_PROTEINS | 7.0E-07 | 1.9E-03 | 38 | CDH1, DSG2, DSG3 |
| REACTOME_FATTY_ACIDS | 1.0E-06 | 2.9E-03 | 15 | CYP4F11, CYP4F3 |
| REACTOME_ANCHORING_FIBRIL_FORMATION | 1.0E-06 | 2.9E-03 | 15 | LAMB3, LAMC2 |
| REACTOME_MET_PROMOTES_CELL_MOTILITY | 2.0E-06 | 5.9E-03 | 41 | ITGA2, LAMB3, LAMC2 |
| PID_INTEGRIN5_PATHWAY | 5.1E-06 | 1.5E-02 | 17 | ITGB6, SDC1 |
| KEGG_CELL_ADHESION_MOLECULES_CAMS | 1.2E-05 | 3.5E-02 | 133 | CDH1, CDH3, NECTIN1, NRCAM, SDC1 |
| REACTOME_APOPTOTIC_EXECUTION_PHASE | 4.1E-05 | 1.2E-01 | 52 | CDH1, DSG2, DSG3 |
| REACTOME_ARACHIDONIC_ACID_METABOLISM | 1.5E-04 | 4.4E-01 | 59 | ABCC1, CYP4F11, CYP4F3 |
| REACTOME_OTHER_INTERLEUKIN_SIGNALING | 1.9E-04 | 5.5E-01 | 24 | PTPRZ1, SDC1 |
| REACTOME_O_LINKED_GLYCOSYLATION | 2.4E-04 | 7.1E-01 | 111 | B3GNT5, GALNT14, GALNT3, LARGE2 |
| REACTOME_O_LINKED_GLYCOSYLATION_OF_MUCINS | 2.5E-04 | 7.2E-01 | 62 | B3GNT5, GALNT14, GALNT3 |
| WP_OXIDATION_BY_CYTOCHROME_P450 | 2.9E-04 | 8.3E-01 | 63 | CYP2S1, CYP4F11, CYP4F3 |
| PID_INTEGRIN_CS_PATHWAY | 3.8E-04 | 1.0E+00 | 26 | ITGA2, ITGB6 |
| PID_INTEGRIN1_PATHWAY | 4.4E-04 | 1.0E+00 | 66 | ITGA2, LAMB3, LAMC2 |
| REACTOME_CYTOCHROME_P450_ARRANGED_BY_SUBSTRATE_TYPE | 4.4E-04 | 1.0E+00 | 66 | CYP2S1, CYP4F11, CYP4F3 |
| REACTOME_SYNDECAN_INTERACTIONS | 5.1E-04 | 1.0E+00 | 27 | ITGA2, SDC1 |
| WP_METAPATHWAY_BIOTRANSFORMATION_PHASE_I_AND_II | 6.1E-04 | 1.0E+00 | 185 | CYP2S1, CYP4F11, CYP4F3, HS6ST2, UGT1A6 |
| WP_HIPPOMERLIN_SIGNALING_DYSREGULATION | 6.6E-04 | 1.0E+00 | 123 | CDH1, CDH3, ITGA2, ITGB6 |
| KEGG_O_GLYCAN_BIOSYNTHESIS | 1.1E-03 | 1.0E+00 | 30 | GALNT14, GALNT3 |
| PID_NECTIN_PATHWAY | 1.1E-03 | 1.0E+00 | 30 | CDH1, NECTIN1 |
| KEGG_ARRHYTHMOGENIC_RIGHT_VENTRICULAR_CARDIOMYOPATHY_ARVC | 1.2E-03 | 1.0E+00 | 74 | DSG2, ITGA2, ITGB6 |
| WP_ARRHYTHMOGENIC_RIGHT_VENTRICULAR_CARDIOMYOPATHY | 1.4E-03 | 1.0E+00 | 76 | DSG2, ITGA2, ITGB6 |
| REACTOME_HS_GAG_BIOSYNTHESIS | 1.4E-03 | 1.0E+00 | 31 | HS6ST2, SDC1 |
| REACTOME_SIGNALING_BY_MET | 1.9E-03 | 1.0E+00 | 79 | ITGA2, LAMB3, LAMC2 |
| REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION | 3.0E-03 | 1.0E+00 | 301 | CDH1, ITGA2, ITGB6, LAMB3, LAMC2, SDC1 |
| KEGG_SMALL_CELL_LUNG_CANCER | 3.0E-03 | 1.0E+00 | 84 | ITGA2, LAMB3, LAMC2 |
| WP_ALPHA_6_BETA_4_SIGNALING_PATHWAY | 3.2E-03 | 1.0E+00 | 35 | LAMB3, LAMC2 |

| Gene Ontology Pathway, Source and Description | Raw P value | FDR | Pathway Gene # | Exemplary Biomarkers Included |
|---|---|---|---|---|
| REACTOME_INTEGRIN_CELL_SURFACE_INTERACTIONS | 3.2E-03 | 1.0E+00 | 85 | CDH1, ITGA2, ITGB6 |
| KEGG_HEMATOPOIETIC_CELL_LINEAGE | 3.8E-03 | 1.0E+00 | 87 | CD9, ITGA2, TFRC |
| REACTOME_RAC2_GTPASE_CYCLE | 4.1E-03 | 1.0E+00 | 88 | DSG2, RACGAP1, TFRC |
| REACTOME_ELECTRIC_TRANSMISSION_ACROSS_GAP_JUNCTIONS | 4.7E-03 | 1.0E+00 | 5 | PANX2 |
| REACTOME_HYALURONAN_BIOSYNTHESIS_AND_EXPORT | 4.7E-03 | 1.0E+00 | 5 | HAS3 |
| REACTOME_RNA_POLYMERASE_II_TRANSCRIPTION | 5.9E-03 | 1.0E+00 | 1374 | |
| REACTOME_POST_TRANSLATIONAL_MODIFICATION_SYNTHESIS_OF_GPI_ANCHORED_PROTEINS | 6.3E-03 | 1.0E+00 | 94 | PIGT, PRSS21, ULBP2 |
| REACTOME_RAC3_GTPASE_CYCLE | 6.3E-03 | 1.0E+00 | 94 | DSG2, RACGAP1, TFRC |
| NABA_BASEMENT_MEMBRANES | 6.9E-03 | 1.0E+00 | 40 | LAMB3, LAMC2 |
| WP_FERROPTOSIS | 6.9E-03 | 1.0E+00 | 40 | SLC7A11, TFRC |
| WP_SMALL_CELL_LUNG_CANCER | 8.2E-03 | 1.0E+00 | 98 | ITGA2, LAMB3, LAMC2 |
| BIOCARTA_MRP_PATHWAY | 1.1E-02 | 1.0E+00 | 6 | ABCC1 |
| REACTOME_ACROSOME_REACTION_AND_SPERM_OOCYTE_MEMBRANE_BINDING | 1.1E-02 | 1.0E+00 | 6 | CD9 |
| REACTOME_UPTAKE_AND_FUNCTION_OF_DIPHTHERIA_TOXIN | 1.1E-02 | 1.0E+00 | 6 | CD9 |
| REACTOME_SENSING_OF_DNA_DOUBLE_STRAND_BREAKS | 1.1E-02 | 1.0E+00 | 6 | KPNA2 |
| REACTOME_VLDL_CLEARANCE | 1.1E-02 | 1.0E+00 | 6 | LSR |
| KEGG_ABC_TRANSPORTERS | 1.1E-02 | 1.0E+00 | 44 | ABCA13, ABCC1 |
| REACTOME_PHASE_I_FUNCTIONALIZATION_OF_COMPOUNDS | 1.3E-02 | 1.0E+00 | 106 | CYP2S1, CYP4F11, CYP4F3 |
| WP_MECHANOREGULATION_AND_PATHOLOGY_OF_YAPTAZ_VIA_HIPPO_AND_NONHIPPO_MECHANISMS | 1.7E-02 | 1.0E+00 | 48 | CDH1, ITGB6 |
| PID_ARF6_TRAFFICKING_PATHWAY | 1.9E-02 | 1.0E+00 | 49 | CDH1, ITGA2 |
| BIOCARTA_EEA1_PATHWAY | 2.0E-02 | 1.0E+00 | 7 | TFRC |
| REACTOME_ATTACHMENT_OF_GPI_ANCHOR_TO_UPAR | 2.0E-02 | 1.0E+00 | 7 | PIGT |
| REACTOME_NECTIN_NECL_TRANS_HETERODIMERIZATION | 2.0E-02 | 1.0E+00 | 7 | NECTIN1 |
| REACTOME_NRCAM_INTERACTIONS | 2.0E-02 | 1.0E+00 | 7 | NRCAM |
| REACTOME_NEUROFASCIN_INTERACTIONS | 2.0E-02 | 1.0E+00 | 7 | NRCAM |
| KEGG_FOCAL_ADHESION | 2.3E-02 | 1.0E+00 | 199 | ITGA2, ITGB6, LAMB3, LAMC2 |
| REACTOME_TRANSPORT_OF_SMALL_MOLECULES | 2.4E-02 | 1.0E+00 | 728 | ABCC1, CLCA2, FXYD3, GNG4, LSR, PARL, SLC15A2, SLC7A11, TFRC |
| WP_FOCAL_ADHESION | 2.5E-02 | 1.0E+00 | 202 | ITGA2, ITGB6, LAMB3, LAMC2 |
| PID_RAC1_PATHWAY | 2.9E-02 | 1.0E+00 | 54 | CDH1, RACGAP1 |

| Gene Ontology Pathway, Source and Description | Raw P value | FDR | Pathway Gene # | Exemplary Biomarkers Included |
|---|---|---|---|---|
| REACTOME_GLYCOSAMINOGLYCAN_METABOLISM | 2.9E-02 | 1.0E+00 | 124 | HAS3, HS6ST2, SDC1 |
| BIOCARTA_CTBP1_PATHWAY | 3.1E-02 | 1.0E+00 | 8 | CDH1 |
| REACTOME_CREB1_PHOSPHORYLATION_THROUGH_THE_ACTIVATION_OF_CAMKII_CAMKK_CAMKIV_CASCADE | 3.1E-02 | 1.0E+00 | 8 | KPNA2 |
| REACTOME_PHOSPHATE_BOND_HYDROLYSIS_BY_NTPDASE_PROTEINS | 3.1E-02 | 1.0E+00 | 8 | ENTPD3 |
| REACTOME_HEPARAN_SULFATE_HEPARIN_HS_GAG_METABOLISM | 3.1E-02 | 1.0E+00 | 55 | HS6ST2, SDC1 |
| WP_FOCAL_ADHESIONPI3KAKTMTORSIGNALING_PATHWAY | 3.3E-02 | 1.0E+00 | 309 | GNG4, ITGA2, ITGB6, LAMB3, LAMC2 |
| REACTOME_BIOLOGICAL_OXIDATIONS | 4.1E-02 | 1.0E+00 | 222 | CYP2S1, CYP4F11, CYP4F3, UGT1A6 |
| REACTOME_CHL1_INTERACTIONS | 4.5E-02 | 1.0E+00 | 9 | ITGA2 |
| REACTOME_INLA_MEDIATED_ENTRY_OF_LISTERIA_MONOCYTOGENES_INTO_HOST_CELLS | 4.5E-02 | 1.0E+00 | 9 | CDH1 |
| REACTOME_ASSEMBLY_OF_COLLAGEN_FIBRILS_AND_OTHER_MULTIMERIC_STRUCTURES | 4.7E-02 | 1.0E+00 | 61 | LAMB3, LAMC2 |

**Example 9: Correlation of bioinformatically-identified biomarkers and biomarker combinations with clinical covariates and known mutational drivers.**

**[591]** The present Example illustrates potential associations between known lung cancer clinical covariates and certain bioinformatically-predicted biomarkers; and potential associations between known lung cancer mutational drivers (e.g., somatic mutational drivers) and certain bioinformatically-predicted biomarkers.

**[592]** In some embodiments, a heatmap can be useful for such analyses. For example, a heatmap that shows the row-scaled gene expression of certain identified biomarker genes using The Cancer Genome Atlas (TCGA) LUSC samples was generated, where each row represented a biomarker candidate and each column represented gene expression in a LUSC sample. Pearson correlation metrics were used to cluster rows and Euclidean distance to cluster columns. Both dimensions were clustered using Ward's algorithm, and an optimal leaf ordering algorithm was employed.

**[593]** In some embodiments, one or more clinical covariates were considered in addition to gene expression of certain bioinformatically-identified biomarkers. Such analysis can be

useful to provide an indication on potential subgroups, including staging, lymph node involvement, microsatellite instability and others.

**[594]** In some embodiments, clinical covariates included nodal involvement (*e.g.*, n0, n1, n2, n3, or n4), cancer stage, tumor grade, histological type, fibrosis, vascular tumor cell type and/or smoking amount (*e.g.*, pack years). In some embodiments, cancer stage included stage I, stage II, stage III, or stage IV cancers. In some embodiments, smoking amount was quantified in terms of pack years (*e.g.,* as described herein, *e.g.,* wherein 1 pack year is equal to 1 pack of cigarettes smoked per day for one year).

**[595]** This clinical covariate analysis did not identify any strong enrichments within the TCGA sample, demonstrating that certain bioinformatically-identified biomarker combinations can be particularly useful to identify lung cancer samples (*e.g.*, LUSC samples) irrespective of a particular clinical covariate (data not shown).

**[596]** In some embodiments, one or more mutational drivers (including, *e.g.*, mutation and copy number of alteration profiles) were considered in addition to gene expression of certain bioinformatically-identified biomarkers. For example, certain major known mutational drivers of lung cancer include, but are not limited to mutations in TP53, RB1, PTEN, PIK3CA, NOTCH1, NFE2L2, KMT2D, KEAP1, or combinations thereof. For each of these drivers, cancer-associated mutations may include copy number alterations (CNAs; including, *e.g.*, but not limited to amplification and/or deletion) and/or mutations (including, *e.g.*, but not limited to inframe mutation, missense mutation, splice, and/or truncating mutation). A clustering analysis was performed to identify associations between bioinformatically-predicted biomarkers, biomarker combinations, and certain major mutational drivers of lung cancer.

**[597]** This mutational driver analysis did not identify any strong enrichments within the TCGA sample, demonstrating that certain bioinformatically-predicted biomarkers and/or biomarker combinations can be particularly useful to identify lung cancer samples (*e.g.*, LUSC samples) irrespective of a particular mutational driver (data not shown).

**Example 10: Development of exemplary lung cancer liquid biopsy assays**

**[598]** The present Example describes development of exemplary lung cancer liquid biopsy assays, for example, for screening subjects with or without symptoms who have average risk, hereditary risk, and/or life-history associated risk. Steps leading to biomarker discovery

were performed as described in Examples 7 to 9 of the present disclosure. Sample preparation and assays were performed as described in Example 1 and Example 2.

**[599]** Lung cancer biomarkers were identified using bioinformatic pathways as described herein; certain biomarkers (presented in Table 3) were tested using commercially-available monoclonal antibodies targeting cell-line EVs to identify antibody clones that exhibit high avidity and selectivity in exemplary assays (data not shown). At least one suitable antibody clone was identified for each biomarker listed in Table 3.

**Table 3 – Lung cancer biomarkers tested in exemplary duplex assays described herein**

| # | Biomarker | # | Biomarker |
|---|---|---|---|
| 1 | ALCAM | 12 | IG1FR |
| 2 | CD55 | 13 | MET |
| 3 | CDH1 | 14 | MSLN |
| 4 | CDH3 | 15 | MUC1 |
| 5 | CD274 (PD-L1) | 16 | SLC34A2 |
| 6 | CEACAM5 | 17 | sTn antigen |
| 7 | CEACAM6 | 18 | Tn antigen |
| 8 | DSG2 | 19 | T antigen |
| 9 | EGFR | 20 | TACSTD2 |
| 10 | EPCAM | 21 | TNFRSF10B |
| 11 | FOLR1 | | |

**[600]** Tables 4A, 4B, and 4C, below, show certain duplex biomarker combinations that were evaluated and findings thereof. In certain embodiments, a duplex biomarker combination comprises two of distinct biomarkers listed in Table 3 above. In some embodiments, a duplex biomarker combination comprises at least one of the biomarker combinations described in Table 4A or 4B.

**Table 4A – Certain biomarker combinations for exemplary duplex assays**

| Test | Capture Probe Target | Detection Probe Target | Test | Capture Probe Target | Detection Probe Target |
|---|---|---|---|---|---|
| 1 | ALCAM | CD274 | 104 | FOLR1 | MSLN |
| 2 | ALCAM | CD55 | 105 | FOLR1 | MUC1 |
| 3 | ALCAM | CDH1 | 106 | FOLR1 | SLC34A2 |
| 4 | ALCAM | CDH3 | 107 | FOLR1 | sTn antigen |
| 5 | ALCAM | CEACAM5 | 108 | FOLR1 | T antigen |

| Test | Capture Probe Target | Detection Probe Target | Test | Capture Probe Target | Detection Probe Target |
|---|---|---|---|---|---|
| 6 | ALCAM | CEACAM6 | 109 | FOLR1 | TACSTD2 |
| 7 | ALCAM | EGFR | 110 | FOLR1 | Tn antigen |
| 8 | ALCAM | EPCAM | 111 | IGF1R | CDH1 |
| 9 | ALCAM | FOLR1 | 112 | IGF1R | T antigen |
| 10 | ALCAM | IGF1R | 113 | MET | CDH1 |
| 11 | ALCAM | MET | 114 | MET | CDH3 |
| 12 | ALCAM | MSLN | 115 | MET | IGF1R |
| 13 | ALCAM | MUC1 | 116 | MET | MSLN |
| 14 | ALCAM | SLC34A2 | 117 | MET | MUC1 |
| 15 | ALCAM | sTn antigen | 118 | MET | sTn antigen |
| 16 | ALCAM | T antigen | 119 | MET | T antigen |
| 17 | ALCAM | TACSTD2 | 120 | MET | TACSTD2 |
| 18 | ALCAM | Tn antigen | 121 | MET | Tn antigen |
| 19 | CDH3 | CDH1 | 122 | MSLN | CDH1 |
| 20 | CDH3 | IGF1R | 123 | MSLN | IGF1R |
| 21 | CDH3 | MSLN | 124 | MSLN | sTn antigen |
| 22 | CDH3 | sTn antigen | 125 | MSLN | T antigen |
| 23 | CDH3 | T antigen | 126 | MUC1 | CDH1 |
| 24 | CDH3 | TACSTD2 | 127 | MUC1 | CDH3 |
| 25 | CEACAM5 | CDH1 | 128 | MUC1 | IGF1R |
| 26 | CEACAM5 | CDH3 | 129 | MUC1 | MSLN |
| 27 | CEACAM5 | IGF1R | 130 | MUC1 | sTn antigen |
| 28 | CEACAM5 | MET | 131 | MUC1 | T antigen |
| 29 | CEACAM5 | MSLN | 132 | MUC1 | TACSTD2 |
| 30 | CEACAM5 | MUC1 | 133 | MUC1 | Tn antigen |
| 31 | CEACAM5 | sTn antigen | 134 | SLC34A2 | CD274 |
| 32 | CEACAM5 | T antigen | 135 | SLC34A2 | CDH1 |
| 33 | CEACAM5 | TACSTD2 | 136 | SLC34A2 | CDH3 |
| 34 | CEACAM5 | Tn antigen | 137 | SLC34A2 | CEACAM5 |
| 35 | CEACAM6 | CD274 | 138 | SLC34A2 | EGFR |
| 36 | CEACAM6 | CDH1 | 139 | SLC34A2 | IGF1R |
| 37 | CEACAM6 | CDH3 | 140 | SLC34A2 | MET |
| 38 | CEACAM6 | CEACAM5 | 141 | SLC34A2 | MSLN |
| 39 | CEACAM6 | EGFR | 142 | SLC34A2 | MUC1 |
| 40 | CEACAM6 | EPCAM | 143 | SLC34A2 | sTn antigen |
| 41 | CEACAM6 | FOLR1 | 144 | SLC34A2 | T antigen |
| 42 | CEACAM6 | IGF1R | 145 | SLC34A2 | TACSTD2 |

| Test | Capture Probe Target | Detection Probe Target | Test | Capture Probe Target | Detection Probe Target |
|---|---|---|---|---|---|
| 43 | CEACAM6 | MET | 146 | SLC34A2 | Tn antigen |
| 44 | CEACAM6 | MSLN | 147 | sTn antigen | CDH1 |
| 45 | CEACAM6 | MUC1 | 148 | sTn antigen | IGF1R |
| 46 | CEACAM6 | SLC34A2 | 149 | T antigen | CDH1 |
| 47 | CEACAM6 | sTn antigen | 150 | TACSTD2 | CDH1 |
| 48 | CEACAM6 | T antigen | 151 | TACSTD2 | IGF1R |
| 49 | CEACAM6 | TACSTD2 | 152 | TACSTD2 | MSLN |
| 50 | CEACAM6 | Tn antigen | 153 | TACSTD2 | sTn antigen |
| 51 | DSG2 | ALCAM | 154 | TACSTD2 | T antigen |
| 52 | DSG2 | CD274 | 155 | Tn antigen | CDH1 |
| 53 | DSG2 | CD55 | 156 | Tn antigen | CDH3 |
| 54 | DSG2 | CDH1 | 157 | Tn antigen | IGF1R |
| 55 | DSG2 | CDH3 | 158 | Tn antigen | MSLN |
| 56 | DSG2 | CEACAM5 | 159 | Tn antigen | TACSTD2 |
| 57 | DSG2 | CEACAM6 | 160 | TNFRSF10B | ALCAM |
| 58 | DSG2 | EGFR | 161 | TNFRSF10B | CD274 |
| 59 | DSG2 | EPCAM | 162 | TNFRSF10B | CD55 |
| 60 | DSG2 | FOLR1 | 163 | TNFRSF10B | CDH1 |
| 61 | DSG2 | IGF1R | 164 | TNFRSF10B | CDH3 |
| 62 | DSG2 | MET | 165 | TNFRSF10B | CEACAM5 |
| 63 | DSG2 | MSLN | 166 | TNFRSF10B | CEACAM6 |
| 64 | DSG2 | MUC1 | 167 | TNFRSF10B | EGFR |
| 65 | DSG2 | SLC34A2 | 168 | TNFRSF10B | EPCAM |
| 66 | DSG2 | sTn antigen | 169 | TNFRSF10B | FOLR1 |
| 67 | DSG2 | T antigen | 170 | TNFRSF10B | IGF1R |
| 68 | DSG2 | TACSTD2 | 171 | TNFRSF10B | MET |
| 69 | DSG2 | Tn antigen | 172 | TNFRSF10B | MSLN |
| 70 | EGFR | CD274 | 173 | TNFRSF10B | MUC1 |
| 71 | EGFR | CDH1 | 174 | TNFRSF10B | SLC34A2 |
| 72 | EGFR | CDH3 | 175 | TNFRSF10B | sTn antigen |
| 73 | EGFR | CEACAM5 | 176 | TNFRSF10B | T antigen |
| 74 | EGFR | IGF1R | 177 | TNFRSF10B | TACSTD2 |
| 75 | EGFR | MET | 178 | TNFRSF10B | Tn antigen |
| 76 | EGFR | MSLN | 179 | CD274 | CDH1 |
| 77 | EGFR | MUC1 | 180 | CD274 | CDH3 |
| 78 | EGFR | sTn antigen | 181 | CD274 | CEACAM5 |
| 79 | EGFR | T antigen | 182 | CD274 | IGF1R |

| Test | Capture Probe Target | Detection Probe Target | Test | Capture Probe Target | Detection Probe Target |
|---|---|---|---|---|---|
| 80 | EGFR | TACSTD2 | 183 | CD274 | MET |
| 81 | EGFR | Tn antigen | 184 | CD274 | MSLN |
| 82 | EPCAM | CD274 | 185 | CD274 | MUC1 |
| 83 | EPCAM | CDH1 | 186 | CD274 | sTn antigen |
| 84 | EPCAM | CDH3 | 187 | CD274 | T antigen |
| 85 | EPCAM | CEACAM5 | 188 | CD274 | TACSTD2 |
| 86 | EPCAM | EGFR | 189 | CD274 | Tn antigen |
| 87 | EPCAM | FOLR1 | 190 | CD55 | CD274 |
| 88 | EPCAM | IGF1R | 191 | CD55 | CDH1 |
| 89 | EPCAM | MET | 192 | CD55 | CDH3 |
| 90 | EPCAM | MSLN | 193 | CD55 | CEACAM5 |
| 91 | EPCAM | MUC1 | 194 | CD55 | CEACAM6 |
| 92 | EPCAM | SLC34A2 | 195 | CD55 | EGFR |
| 93 | EPCAM | sTn antigen | 196 | CD55 | EPCAM |
| 94 | EPCAM | T antigen | 197 | CD55 | FOLR1 |
| 95 | EPCAM | TACSTD2 | 198 | CD55 | IGF1R |
| 96 | EPCAM | Tn antigen | 199 | CD55 | MET |
| 97 | FOLR1 | CD274 | 200 | CD55 | MSLN |
| 98 | FOLR1 | CDH1 | 201 | CD55 | MUC1 |
| 99 | FOLR1 | CDH3 | 202 | CD55 | SLC34A2 |
| 100 | FOLR1 | CEACAM5 | 203 | CD55 | sTn antigen |
| 101 | FOLR1 | EGFR | 204 | CD55 | T antigen |
| 102 | FOLR1 | IGF1R | 205 | CD55 | TACSTD2 |
| 103 | FOLR1 | MET | 206 | CD55 | Tn antigen |

*Additional Assessment of Certain Biomarker Combinations*

**[601]** Certain biomarker combinations were assayed in subjects' samples from lung adenocarcinoma, lung squamous cell carcinoma, or healthy never-smoker pools. Certain biomarker combinations, e.g., those as provided in Table 4B, were observed to be particularly useful for the detection of lung cancer, including, e.g., lung adenocarcinoma and lung squamous cell carcinoma (LUNG), or for the detection of lung adenocarcinoma (LUAD), or for the detection of lung squamous cell carcinoma (LUSqC), based on the exemplary cutoff criteria described below. The median of the assay replicates was used for this analysis.

- LUNG
    - o $\Delta Ct > 2$ between LUAD pool and healthy never-smoker pool

- o ΔCt > 2 between LUSqC pool and healthy never-smoker pool
- LUAD
  - o ΔCt > 2 between LUAD pool and healthy never-smoker pool
  - o ΔCt < 2 between LUSqC pool and healthy never-smoker pool
- LUSqC
  - o ΔCt > 2 between LUSqC pool and healthy never-smoker pool
  - o ΔCt < 2 between LUAD pool and healthy never-smoker pool

**Table 4B – Certain biomarker combinations for exemplary duplex assays.** Note that for each biomarker combination disclosed below, in some embodiments, a capture biomarker may be used as a detection biomarker, and a detection biomarker may be used as a capture biomarker.

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 1 | STn, MUC1 + CEACAM6 | LUNG |
| 2 | STn, FOLR1 + FOLR1 | LUAD |
| 3 | STn, CD9 + CD9 | |
| 4 | STn, TSPAN8 + TSPAN8 | |
| 5 | STn, TFRC + TFRC | |
| 6 | STn, MUC1 + MUC1 | LUNG |
| 7 | STn, CEACAM5 + CEACAM5 | LUNG |
| 8 | STn, sLeX + sLeX | LUNG |
| 9 | STn, CD55 + CD55 | |
| 10 | STn, EPCAM + EPCAM | LUNG |
| 11 | STn, MSLN + MSLN | LUAD |
| 12 | STn, LeX + LeX | |
| 13 | STn, BCAP31 + BCAP31 | |
| 14 | STn, ITGA2 + ITGA2 | |
| 15 | STn, NT5E + NT5E | LUAD |
| 16 | STn, MUC1 + MUC1 | LUAD |
| 17 | STn, TACSTD2 + TACSTD2 | LUNG |
| 18 | STn, ALCAM + ALCAM | |
| 19 | STn, MARCKSL1 + MARCKSL1 | LUNG |
| 20 | STn, MET + MET | |
| 21 | STn, VWA1 + VWA1 | |
| 22 | STn, PS + PS | |
| 23 | STn, CD274 + CD274 | |
| 24 | STn, T + T | LUNG |
| 25 | STn, MUC4 + MUC4 | LUNG |
| 26 | STn, CEACAM6 + CEACAM6 | LUNG |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 27 | STn, IL1RAP + IL1RAP | |
| 28 | STn, PTGFRN + PTGFRN | |
| 29 | STn, LEY + LEY | LUAD |
| 30 | STn, SDC1 + SDC1 | LUSqC |
| 31 | MUC1, CEACAM6 + STn | LUNG |
| 32 | MUC1, LeX + LeX | LUAD |
| 33 | MUC1, FOLR1 + FOLR1 | LUAD |
| 34 | MUC1, CD9 + CD9 | |
| 35 | MUC1, sLeX + sLeX | LUAD |
| 36 | MUC1, CD55 + CD55 | LUAD |
| 37 | MUC1, NT5E + NT5E | LUAD |
| 38 | MUC1, TFRC + TFRC | |
| 39 | MUC1, PTGFRN + PTGFRN | |
| 40 | MUC1, CEACAM6 + CEACAM6 | LUNG |
| 41 | MUC1, TACSTD2 + TACSTD2 | LUNG |
| 42 | MUC1, STn + STn | LUNG |
| 43 | MUC1, LEY + LEY | LUNG |
| 44 | MUC1, ALCAM + ALCAM | LUAD |
| 45 | MUC1, BCAP31 + BCAP31 | |
| 46 | MUC1, CD274 + CD274 | |
| 47 | MUC1, EPCAM + EPCAM | LUNG |
| 48 | MUC1, MUC1 + MUC1 | |
| 49 | MUC1, MUC1 + MUC1 | LUAD |
| 50 | MUC1, PS + PS | |
| 51 | MUC1, CEACAM5 + CEACAM5 | LUNG |
| 52 | MUC1, MSLN + MSLN | |
| 53 | MUC1, MUC4 + MUC4 | LUNG |
| 54 | MUC1, Tn + Tn | |
| 55 | MUC1, MARCKSL1 + MARCKSL1 | |
| 56 | MUC1, ITGA2 + ITGA2 | |
| 57 | MUC1, T + T | |
| 58 | MUC1, DSC2 + DSC2 | |
| 59 | MUC1, VWA1 + VWA1 | LUSqC |
| 60 | MUC1, SLC34A2 + SLC34A2 | LUAD |
| 61 | EPCAM, LeX + LeX | |
| 62 | EPCAM, FOLR1 + FOLR1 | |
| 63 | EPCAM, ALCAM + ALCAM | |
| 64 | EPCAM, TACSTD2 + TACSTD2 | |
| 65 | EPCAM, LEY + LEY | |
| 66 | EPCAM, BCAP31 + BCAP31 | |
| 67 | EPCAM, CD55 + CD55 | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 68 | EPCAM, T + T | |
| 69 | EPCAM, NT5E + NT5E | |
| 70 | EPCAM, sLeX + sLeX | LUAD |
| 71 | EPCAM, STn + STn | LUSqC |
| 72 | EPCAM, CD9 + CD9 | |
| 73 | EPCAM, CD274 + CD274 | |
| 74 | EPCAM, MUC1 + MUC1 | |
| 75 | EPCAM, CEACAM6 + CEACAM6 | |
| 76 | EPCAM, MUC1 + MUC1 | |
| 77 | EPCAM, PTGFRN + PTGFRN | |
| 78 | EPCAM, MUC4 + MUC4 | |
| 79 | EPCAM, MARCKSL1 + MARCKSL1 | |
| 80 | EPCAM, TFRC + TFRC | |
| 81 | EPCAM, MET + MET | |
| 82 | EPCAM, PS + PS | |
| 83 | EPCAM, TSPAN8 + TSPAN8 | |
| 84 | EPCAM, EGFR + EGFR | LUNG |
| 85 | FOLR1, TACSTD2 + TACSTD2 | |
| 86 | FOLR1, MUC1 + MUC1 | |
| 87 | FOLR1, LeX + LeX | |
| 88 | FOLR1, CD9 + CD9 | |
| 89 | FOLR1, MUC1 + MUC1 | |
| 90 | FOLR1, MUC1 + MUC1 | |
| 91 | FOLR1, Tn + Tn | |
| 92 | FOLR1, STn + STn | LUNG |
| 93 | FOLR1, BCAP31 + BCAP31 | |
| 94 | FOLR1, PS + PS | |
| 95 | FOLR1, sLeX + sLeX | |
| 96 | FOLR1, TFRC + TFRC | |
| 97 | FOLR1, LEY + LEY | |
| 98 | FOLR1, ITGA2 + ITGA2 | |
| 99 | FOLR1, NT5E + NT5E | |
| 100 | FOLR1, EGFR + EGFR | LUSqC |
| 101 | FOLR1, MARCKSL1 + MARCKSL1 | LUSqC |
| 102 | FOLR1, T + T | |
| 103 | MSLN, CD55 + CD55 | |
| 104 | MSLN, CD9 + CD9 | |
| 105 | MSLN, TACSTD2 + TACSTD2 | |
| 106 | MSLN, TSPAN8 + TSPAN8 | |
| 107 | MSLN, BCAP31 + BCAP31 | |
| 108 | MSLN, EPCAM + EPCAM | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 109 | MSLN, LeX + LeX | |
| 110 | MSLN, CEACAM6 + CEACAM6 | LUAD |
| 111 | MSLN, FOLR1 + FOLR1 | |
| 112 | MSLN, TFRC + TFRC | |
| 113 | MSLN, MUC1 + MUC1 | |
| 114 | MSLN, NT5E + NT5E | |
| 115 | SLC34A2, LeX + LeX | |
| 116 | SLC34A2, FOLR1 + FOLR1 | |
| 117 | SLC34A2, CD55 + CD55 | |
| 118 | SLC34A2, CEACAM6 + CEACAM6 | |
| 119 | SLC34A2, sLeX + sLeX | |
| 120 | SLC34A2, NT5E + NT5E | |
| 121 | SLC34A2, TACSTD2 + TACSTD2 | |
| 122 | SLC34A2, BCAP31 + BCAP31 | |
| 123 | SLC34A2, CD274 + CD274 | |
| 124 | SLC34A2, LEY + LEY | |
| 125 | SLC34A2, ITGA2 + ITGA2 | |
| 126 | SLC34A2, PTGFRN + PTGFRN | LUSqC |
| 127 | TACSTD2, FOLR1 + FOLR1 | |
| 128 | TACSTD2, LeX + LeX | |
| 129 | TACSTD2, CD9 + CD9 | |
| 130 | TACSTD2, EPCAM + EPCAM | LUNG |
| 131 | TACSTD2, MUC1 + MUC1 | |
| 132 | TACSTD2, LEY + LEY | |
| 133 | TACSTD2, MUC1 + MUC1 | |
| 134 | TACSTD2, sLeX + sLeX | LUAD |
| 135 | TACSTD2, CEACAM6 + CEACAM6 | |
| 136 | TACSTD2, MUC1 + MUC1 | |
| 137 | TACSTD2, CD55 + CD55 | |
| 138 | TACSTD2, TSPAN8 + TSPAN8 | |
| 139 | TACSTD2, NT5E + NT5E | |
| 140 | TACSTD2, MSLN + MSLN | |
| 141 | TACSTD2, Tn + Tn | |
| 142 | TACSTD2, TFRC + TFRC | |
| 143 | TACSTD2, CEACAM5 + CEACAM5 | |
| 144 | TACSTD2, ALCAM + ALCAM | |
| 145 | TACSTD2, PS + PS | |
| 146 | TACSTD2, BCAP31 + BCAP31 | |
| 147 | TACSTD2, MUC4 + MUC4 | |
| 148 | TACSTD2, CD274 + CD274 | |
| 149 | TACSTD2, T + T | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 150 | TACSTD2, MARCKSL1 + MARCKSL1 | LUAD |
| 151 | T, EPCAM + EPCAM | LUNG |
| 152 | T, CEACAM6 + CEACAM6 | LUNG |
| 153 | T, ITGA2 + ITGA2 | LUAD |
| 154 | T, CD55 + CD55 | |
| 155 | T, Tn + Tn | LUAD |
| 156 | T, NT5E + NT5E | |
| 157 | T, CD9 + CD9 | |
| 158 | T, PTGFRN + PTGFRN | LUSqC |
| 159 | T, EGFR + EGFR | |
| 160 | T, TFRC + TFRC | |
| 161 | T, FOLR1 + FOLR1 | |
| 162 | T, MUC1 + MUC1 | LUAD |
| 163 | T, PS + PS | |
| 164 | T, BCAP31 + BCAP31 | LUAD |
| 165 | T, LeX + LeX | LUAD |
| 166 | T, TSPAN8 + TSPAN8 | LUAD |
| 167 | T, ALCAM + ALCAM | |
| 168 | T, MUC1 + MUC1 | |
| 169 | T, MUC1 + MUC1 | LUAD |
| 170 | T, MSLN + MSLN | |
| 171 | T, MET + MET | LUAD |
| 172 | T, TACSTD2 + TACSTD2 | LUAD |
| 173 | T, CD274 + CD274 | |
| 174 | T, MUC4 + MUC4 | |
| 175 | LEY, CD9 + CD9 | |
| 176 | LEY, TACSTD2 + TACSTD2 | |
| 177 | LEY, ITGA2 + ITGA2 | |
| 178 | LEY, MUC1 + MUC1 | LUAD |
| 179 | LEY, CD55 + CD55 | |
| 180 | LEY, NT5E + NT5E | |
| 181 | LEY, ALCAM + ALCAM | |
| 182 | LEY, BCAP31 + BCAP31 | |
| 183 | LEY, MUC1 + MUC1 | LUAD |
| 184 | LEY, MUC1 + MUC1 | LUAD |
| 185 | LEY, EPCAM + EPCAM | |
| 186 | LEY, LeX + LeX | |
| 187 | LEY, TFRC + TFRC | |
| 188 | LEY, PS + PS | |
| 189 | LEY, FOLR1 + FOLR1 | |
| 190 | LEY, EGFR + EGFR | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 191 | LEY, TSPAN8 + TSPAN8 | |
| 192 | LEY, LEY + LEY | |
| 193 | LEY, MARCKSL1 + MARCKSL1 | LUNG |
| 194 | LEY, CEACAM5 + CEACAM5 | LUNG |
| 195 | LEY, MSLN + MSLN | |
| 196 | LEY, CD274 + CD274 | |
| 197 | LEY, MUC4 + MUC4 | LUNG |
| 198 | LEY, IL1RAP + IL1RAP | |
| 199 | PS, ALCAM + ALCAM | |
| 200 | PS, BCAP31 + BCAP31 | |
| 201 | PS, EPCAM + EPCAM | |
| 202 | PS, T + T | |
| 203 | PS, CEACAM5 + CEACAM5 | |
| 204 | PS, CD9 + CD9 | |
| 205 | PS, FOLR1 + FOLR1 | |
| 206 | PS, LeX + LeX | LUAD |
| 207 | PS, MARCKSL1 + MARCKSL1 | |
| 208 | PS, LEY + LEY | |
| 209 | PS, MUC1 + MUC1 | |
| 210 | PS, CD274 + CD274 | LUNG |
| 211 | LeX, FOLR1 + FOLR1 | LUAD |
| 212 | LeX, EPCAM + EPCAM | LUAD |
| 213 | LeX, CD55 + CD55 | |
| 214 | LeX, TACSTD2 + TACSTD2 | LUAD |
| 215 | LeX, TSPAN8 + TSPAN8 | |
| 216 | LeX, BCAP31 + BCAP31 | |
| 217 | LeX, PS + PS | LUAD |
| 218 | LeX, NT5E + NT5E | LUAD |
| 219 | LeX, MUC1 + MUC1 | LUNG |
| 220 | LeX, CD9 + CD9 | |
| 221 | LeX, MUC1 + MUC1 | LUAD |
| 222 | LeX, TFRC + TFRC | |
| 223 | LeX, SLC34A2 + SLC34A2 | |
| 224 | LeX, MSLN + MSLN | |
| 225 | LeX, ALCAM + ALCAM | |
| 226 | LeX, STn + STn | LUAD |
| 227 | LeX, VWA1 + VWA1 | |
| 228 | LeX, MUC1 + MUC1 | LUAD |
| 229 | LeX, CEACAM5 + CEACAM5 | LUAD |
| 230 | LeX, LEY + LEY | LUAD |
| 231 | LeX, IL1RAP + IL1RAP | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 232 | LeX, ITGA2 + ITGA2 | |
| 233 | LeX, MET + MET | |
| 234 | LeX, PTGFRN + PTGFRN | |
| 235 | LeX, CD274 + CD274 | |
| 236 | LeX, CEACAM6 + CEACAM6 | |
| 237 | LeX, MARCKSL1 + MARCKSL1 | |
| 238 | LeX, T + T | |
| 239 | LeX, EGFR + EGFR | LUAD |
| 240 | LeX, SDC1 + SDC1 | |
| 241 | sLeX, CD9 + CD9 | |
| 242 | sLeX, TACSTD2 + TACSTD2 | LUNG |
| 243 | sLeX, TSPAN8 + TSPAN8 | |
| 244 | sLeX, NT5E + NT5E | LUAD |
| 245 | sLeX, BCAP31 + BCAP31 | |
| 246 | sLeX, EPCAM + EPCAM | LUAD |
| 247 | sLeX, FOLR1 + FOLR1 | LUAD |
| 248 | sLeX, CD55 + CD55 | |
| 249 | sLeX, MUC1 + MUC1 | LUAD |
| 250 | sLeX, CEACAM5 + CEACAM5 | LUNG |
| 251 | sLeX, TFRC + TFRC | |
| 252 | sLeX, ALCAM + ALCAM | |
| 253 | sLeX, MUC1 + MUC1 | LUAD |
| 254 | sLeX, MET + MET | |
| 255 | sLeX, VWA1 + VWA1 | |
| 256 | sLeX, T + T | LUAD |
| 257 | sLeX, CD274 + CD274 | |
| 258 | sLeX, CEACAM6 + CEACAM6 | LUAD |
| 259 | sLeX, STn + STn | LUAD |
| 260 | sLeX, MUC1 + MUC1 | LUAD |
| 261 | sLeX, LeX + LeX | |
| 262 | sLeX, IL1RAP + IL1RAP | |
| 263 | sLeX, MUC4 + MUC4 | LUNG |
| 264 | sLeX, SDC1 + SDC1 | |
| 265 | MUC1, FOLR1 + FOLR1 | LUAD |
| 266 | MUC1, LeX + LeX | LUAD |
| 267 | MUC1, MUC1 + MUC1 | |
| 268 | MUC1, TACSTD2 + TACSTD2 | |
| 269 | MUC1, MUC1 + MUC1 | |
| 270 | MUC1, NT5E + NT5E | |
| 271 | MUC1, STn + STn | LUNG |
| 272 | MUC1, sLeX + sLeX | LUAD |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 273 | MUC1, CD55 + CD55 | |
| 274 | MUC1, CD9 + CD9 | |
| 275 | MUC1, CEACAM6 + CEACAM6 | LUAD |
| 276 | MUC1, TFRC + TFRC | |
| 277 | MUC1, CD9 + CD9 | |
| 278 | MUC1, TACSTD2 + TACSTD2 | LUNG |
| 279 | MUC1, FOLR1 + FOLR1 | LUAD |
| 280 | MUC1, MUC1 + MUC1 | LUAD |
| 281 | MUC1, CEACAM6 + CEACAM6 | LUNG |
| 282 | MUC1, LeX + LeX | |
| 283 | MUC1, sLeX + sLeX | LUAD |
| 284 | MUC1, CD55 + CD55 | |
| 285 | MUC1, TFRC + TFRC | |
| 286 | MUC1, NT5E + NT5E | |
| 287 | MUC1, EPCAM + EPCAM | LUNG |
| 288 | MUC1, MUC1 + MUC1 | LUAD |
| 289 | MUC1, BCAP31 + BCAP31 | |
| 290 | MUC1, Tn + Tn | |
| 291 | MUC1, STn + STn | LUAD |
| 292 | MUC1, ALCAM + ALCAM | |
| 293 | MUC1, LEY + LEY | |
| 294 | MUC1, ITGA2 + ITGA2 | |
| 295 | MUC1, PTGFRN + PTGFRN | |
| 296 | MUC1, PS + PS | |
| 297 | MUC1, EGFR + EGFR | |
| 298 | MUC1, MSLN + MSLN | |
| 299 | MUC1, MARCKSL1 + MARCKSL1 | |
| 300 | MUC1, TSPAN8 + TSPAN8 | |
| 301 | CEACAM5, LeX + LeX | LUAD |
| 302 | CEACAM5, sLeX + sLeX | LUAD |
| 303 | CEACAM5, TACSTD2 + TACSTD2 | |
| 304 | CEACAM5, CD55 + CD55 | |
| 305 | CEACAM5, MUC4 + MUC4 | |
| 306 | CEACAM5, CEACAM6 + CEACAM6 | LUAD |
| 307 | CEACAM5, LEY + LEY | |
| 308 | CEACAM5, PTGFRN + PTGFRN | |
| 309 | CEACAM5, NT5E + NT5E | |
| 310 | CEACAM5, CD9 + CD9 | |
| 311 | CEACAM5, ALCAM + ALCAM | |
| 312 | CEACAM5, STn + STn | LUNG |
| 313 | CEACAM5, PS + PS | LUSqC |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 314 | CEACAM5, Tn + Tn | |
| 315 | CEACAM5, EGFR + EGFR | LUNG |
| 316 | CEACAM5, T + T | |
| 317 | CEACAM5, FOLR1 + FOLR1 | |
| 318 | CEACAM5, IL1RAP + IL1RAP | |
| 319 | CEACAM6, CD55 + CD55 | |
| 320 | CEACAM6, CD9 + CD9 | |
| 321 | CEACAM6, TACSTD2 + TACSTD2 | LUSqC |
| 322 | CEACAM6, TSPAN8 + TSPAN8 | |
| 323 | CEACAM6, LeX + LeX | |
| 324 | CEACAM6, NT5E + NT5E | |
| 325 | CEACAM6, PTGFRN + PTGFRN | LUSqC |
| 326 | CEACAM6, EGFR + EGFR | |
| 327 | CEACAM6, FOLR1 + FOLR1 | LUAD |
| 328 | CEACAM6, CD274 + CD274 | |
| 329 | CEACAM6, EPCAM + EPCAM | |
| 330 | CEACAM6, MUC1 + MUC1 | LUNG |
| 331 | CEACAM6, BCAP31 + BCAP31 | |
| 332 | CEACAM6, MUC1 + MUC1 | LUNG |
| 333 | CEACAM6, sLeX + sLeX | |
| 334 | CEACAM6, ITGA2 + ITGA2 | |
| 335 | CEACAM6, CEACAM5 + CEACAM5 | LUNG |
| 336 | CEACAM6, T + T | |
| 337 | CEACAM6, LEY + LEY | |
| 338 | CEACAM6, PS + PS | |
| 339 | CEACAM6, TFRC + TFRC | |
| 340 | CEACAM6, MSLN + MSLN | |
| 341 | CEACAM6, SLC34A2 + SLC34A2 | |
| 342 | CEACAM6, STn + STn | LUNG |
| 343 | CEACAM6, ALCAM + ALCAM | |
| 344 | CEACAM6, MET + MET | |
| 345 | CEACAM6, Tn + Tn | |
| 346 | CEACAM6, MUC4 + MUC4 | |
| 347 | CEACAM6, IL1RAP + IL1RAP | LUSqC |
| 348 | CEACAM6, SDC1 + SDC1 | |
| 349 | TRPV4, CEACAM6 + CEACAM6 | LUSqC |
| 350 | TRPV4, TFRC + TFRC | |
| 351 | TRPV4, ITGA2 + ITGA2 | |
| 352 | TRPV4, CD9 + CD9 | |
| 353 | TRPV4, CEACAM5 + CEACAM5 | |
| 354 | TRPV4, MUC4 + MUC4 | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 355 | TRPV4, VWA1 + VWA1 | |
| 356 | TRPV4, BCAP31 + BCAP31 | |
| 357 | TRPV4, PTGFRN + PTGFRN | |
| 358 | TRPV4, CD55 + CD55 | |
| 359 | TRPV4, LEY + LEY | LUSqC |
| 360 | TRPV4, IL1RAP + IL1RAP | |
| 361 | TRPV4, MUC1 + MUC1 | LUAD |
| 362 | TRPV4, EGFR + EGFR | |
| 363 | TRPV4, TSPAN8 + TSPAN8 | |
| 364 | TRPV4, MET + MET | LUNG |
| 365 | TRPV4, STn + STn | LUAD |
| 366 | TRPV4, EPCAM + EPCAM | |
| 367 | TRPV4, NT5E + NT5E | |
| 368 | TRPV4, TACSTD2 + TACSTD2 | LUSqC |
| 369 | TRPV4, CD274 + CD274 | |
| 370 | TRPV4, Tn + Tn | |
| 371 | TRPV4, SLC34A2 + SLC34A2 | |
| 372 | TRPV4, FOLR1 + FOLR1 | |
| 373 | ALCAM, EPCAM + EPCAM | |
| 374 | ALCAM, ALCAM + ALCAM | |
| 375 | ALCAM, CD9 + CD9 | |
| 376 | ALCAM, LEY + LEY | |
| 377 | ALCAM, PS + PS | LUAD |
| 378 | ALCAM, T + T | |
| 379 | ALCAM, TACSTD2 + TACSTD2 | |
| 380 | ALCAM, LeX + LeX | |
| 381 | ALCAM, ITGA2 + ITGA2 | |
| 382 | ALCAM, sLeX + sLeX | LUSqC |
| 383 | ALCAM, CD55 + CD55 | |
| 384 | ALCAM, NT5E + NT5E | |
| 385 | ITGA2, NT5E + NT5E | |
| 386 | ITGA2, T + T | |
| 387 | ITGA2, CD55 + CD55 | |
| 388 | ITGA2, CD9 + CD9 | |
| 389 | ITGA2, PTGFRN + PTGFRN | |
| 390 | ITGA2, LEY + LEY | |
| 391 | ITGA2, MUC4 + MUC4 | |
| 392 | ITGA2, MUC1 + MUC1 | |
| 393 | ITGA2, EGFR + EGFR | |
| 394 | ITGA2, ALCAM + ALCAM | |
| 395 | ITGA2, TFRC + TFRC | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 396 | ITGA2, FOLR1 + FOLR1 | |
| 397 | ITGA2, CD274 + CD274 | |
| 398 | ITGA2, BCAP31 + BCAP31 | |
| 399 | ITGA2, PS + PS | |
| 400 | ITGA2, MUC1 + MUC1 | |
| 401 | ITGA2, MUC1 + MUC1 | |
| 402 | ITGA2, LeX + LeX | |
| 403 | PTGFRN, LeX + LeX | |
| 404 | PTGFRN, CEACAM6 + CEACAM6 | |
| 405 | PTGFRN, ITGA2 + ITGA2 | |
| 406 | PTGFRN, NT5E + NT5E | |
| 407 | PTGFRN, CD9 + CD9 | |
| 408 | PTGFRN, CEACAM5 + CEACAM5 | |
| 409 | PTGFRN, CD55 + CD55 | |
| 410 | PTGFRN, TACSTD2 + TACSTD2 | |
| 411 | PTGFRN, MUC4 + MUC4 | |
| 412 | PTGFRN, T + T | |
| 413 | PTGFRN, CD274 + CD274 | |
| 414 | PTGFRN, MUC1 + MUC1 | |
| 415 | IL1RAP, TFRC + TFRC | |
| 416 | IL1RAP, PS + PS | |
| 417 | IL1RAP, CD9 + CD9 | |
| 418 | IL1RAP, LeX + LeX | |
| 419 | IL1RAP, CD274 + CD274 | |
| 420 | IL1RAP, MSLN + MSLN | |
| 421 | IL1RAP, FOLR1 + FOLR1 | |
| 422 | IL1RAP, ALCAM + ALCAM | |
| 423 | IL1RAP, TACSTD2 + TACSTD2 | |
| 424 | IL1RAP, BCAP31 + BCAP31 | |
| 425 | IL1RAP, PTGFRN + PTGFRN | |
| 426 | IL1RAP, Tn + Tn | |
| 427 | IL1RAP, TSPAN8 + TSPAN8 | |
| 428 | IL1RAP, VWA1 + VWA1 | |
| 429 | IL1RAP, MUC4 + MUC4 | |
| 430 | IL1RAP, SDC1 + SDC1 | |
| 431 | IL1RAP, CEACAM5 + CEACAM5 | |
| 432 | IL1RAP, NT5E + NT5E | |
| 433 | NT5E, CD9 + CD9 | |
| 434 | NT5E, ITGA2 + ITGA2 | |
| 435 | NT5E, CEACAM6 + CEACAM6 | LUSqC |
| 436 | NT5E, TSPAN8 + TSPAN8 | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 437 | NT5E, CD55 + CD55 | |
| 438 | NT5E, LeX + LeX | |
| 439 | NT5E, TACSTD2 + TACSTD2 | LUSqC |
| 440 | NT5E, CD274 + CD274 | |
| 441 | NT5E, T + T | |
| 442 | NT5E, TFRC + TFRC | |
| 443 | NT5E, EGFR + EGFR | |
| 444 | NT5E, sLeX + sLeX | |
| 445 | NT5E, PTGFRN + PTGFRN | |
| 446 | NT5E, MUC1 + MUC1 | |
| 447 | NT5E, EPCAM + EPCAM | |
| 448 | NT5E, FOLR1 + FOLR1 | |
| 449 | NT5E, MUC1 + MUC1 | LUAD |
| 450 | NT5E, LEY + LEY | |
| 451 | NT5E, ALCAM + ALCAM | |
| 452 | NT5E, PS + PS | |
| 453 | NT5E, MET + MET | LUAD |
| 454 | NT5E, MUC1 + MUC1 | LUAD |
| 455 | NT5E, SLC34A2 + SLC34A2 | |
| 456 | NT5E, IL1RAP + IL1RAP | |
| 457 | CD9, EPCAM + EPCAM | |
| 458 | CD9, FOLR1 + FOLR1 | |
| 459 | CD9, TSPAN8 + TSPAN8 | |
| 460 | CD9, LeX + LeX | |
| 461 | CD9, TFRC + TFRC | |
| 462 | CD9, TACSTD2 + TACSTD2 | |
| 463 | CD9, CD55 + CD55 | |
| 464 | CD9, PS + PS | |
| 465 | CD9, CEACAM6 + CEACAM6 | |
| 466 | CD9, NT5E + NT5E | |
| 467 | CD9, MSLN + MSLN | |
| 468 | CD9, MUC1 + MUC1 | |
| 469 | CD9, STn + STn | |
| 470 | CD9, MUC1 + MUC1 | |
| 471 | CD9, MUC1 + MUC1 | |
| 472 | CD9, ITGA2 + ITGA2 | |
| 473 | CD9, LEY + LEY | |
| 474 | CD9, ALCAM + ALCAM | |
| 475 | CD9, sLeX + sLeX | LUAD |
| 476 | CD9, CD274 + CD274 | |
| 477 | CD9, BCAP31 + BCAP31 | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 478 | CD9, PTGFRN + PTGFRN | |
| 479 | CD9, IL1RAP + IL1RAP | |
| 480 | CD9, MET + MET | |
| 481 | CD9, MUC4 + MUC4 | |
| 482 | CD9, T + T | |
| 483 | CD9, MARCKSL1 + MARCKSL1 | |
| 484 | CD9, EGFR + EGFR | |
| 485 | CD9, CEACAM5 + CEACAM5 | |
| 486 | CD9, VWA1 + VWA1 | |
| 487 | EGFR, CD55 + CD55 | |
| 488 | EGFR, NT5E + NT5E | |
| 489 | EGFR, CEACAM6 + CEACAM6 | |
| 490 | EGFR, T + T | LUSqC |
| 491 | EGFR, ITGA2 + ITGA2 | |
| 492 | EGFR, PTGFRN + PTGFRN | LUSqC |
| 493 | EGFR, FOLR1 + FOLR1 | LUSqC |
| 494 | EGFR, PS + PS | LUSqC |
| 495 | EGFR, MUC4 + MUC4 | |
| 496 | EGFR, TSPAN8 + TSPAN8 | |
| 497 | EGFR, CD9 + CD9 | |
| 498 | EGFR, MET + MET | LUSqC |
| 499 | TSPAN8, EPCAM + EPCAM | |
| 500 | TSPAN8, CD9 + CD9 | |
| 501 | TSPAN8, LeX + LeX | |
| 502 | TSPAN8, CD55 + CD55 | |
| 503 | TSPAN8, BCAP31 + BCAP31 | |
| 504 | TSPAN8, CEACAM6 + CEACAM6 | |
| 505 | TSPAN8, NT5E + NT5E | |
| 506 | TSPAN8, sLeX + sLeX | |
| 507 | TSPAN8, TACSTD2 + TACSTD2 | |
| 508 | TSPAN8, EGFR + EGFR | LUNG |
| 509 | TSPAN8, MSLN + MSLN | |
| 510 | TSPAN8, PS + PS | |
| 511 | TSPAN8, STn + STn | LUAD |
| 512 | TSPAN8, LEY + LEY | |
| 513 | TSPAN8, ITGA2 + ITGA2 | |
| 514 | TSPAN8, T + T | |
| 515 | TSPAN8, VWA1 + VWA1 | |
| 516 | TSPAN8, MUC1 + MUC1 | |
| 517 | TSPAN8, PTGFRN + PTGFRN | |
| 518 | TSPAN8, CD274 + CD274 | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 519 | TSPAN8, TFRC + TFRC | |
| 520 | TSPAN8, MUC4 + MUC4 | |
| 521 | TSPAN8, MUC1 + MUC1 | |
| 522 | TSPAN8, MUC1 + MUC1 | |
| 523 | MUC4, TACSTD2 + TACSTD2 | |
| 524 | MUC4, CEACAM5 + CEACAM5 | LUNG |
| 525 | MUC4, EGFR + EGFR | LUSqC |
| 526 | MUC4, sLeX + sLeX | LUNG |
| 527 | MUC4, CD9 + CD9 | |
| 528 | MUC4, NT5E + NT5E | LUAD |
| 529 | MUC4, FOLR1 + FOLR1 | LUAD |
| 530 | MUC4, PTGFRN + PTGFRN | |
| 531 | MUC4, CD55 + CD55 | LUAD |
| 532 | MUC4, MUC1 + MUC1 | LUAD |
| 533 | MUC4, ITGA2 + ITGA2 | |
| 534 | MUC4, ALCAM + ALCAM | |
| 535 | MUC4, EPCAM + EPCAM | |
| 536 | MUC4, LeX + LeX | LUAD |
| 537 | MUC4, MUC1 + MUC1 | LUAD |
| 538 | MUC4, LEY + LEY | LUAD |
| 539 | MUC4, PS + PS | |
| 540 | MUC4, CD274 + CD274 | |
| 541 | MUC4, Tn + Tn | |
| 542 | MUC4, TSPAN8 + TSPAN8 | |
| 543 | MUC4, SDC1 + SDC1 | |
| 544 | MUC4, SLC34A2 + SLC34A2 | |
| 545 | MUC4, STn + STn | LUAD |
| 546 | MUC4, CEACAM6 + CEACAM6 | LUAD |
| 547 | TNFRSF10B, CD55 + CD55 | |
| 548 | TNFRSF10B, EGFR + EGFR | LUSqC |
| 549 | TNFRSF10B, TFRC + TFRC | |
| 550 | TNFRSF10B, NT5E + NT5E | |
| 551 | TNFRSF10B, LEY + LEY | |
| 552 | TNFRSF10B, CD9 + CD9 | |
| 553 | TNFRSF10B, BCAP31 + BCAP31 | |
| 554 | TNFRSF10B, CD274 + CD274 | |
| 555 | TNFRSF10B, PTGFRN + PTGFRN | |
| 556 | TNFRSF10B, T + T | |
| 557 | TNFRSF10B, FOLR1 + FOLR1 | |
| 558 | TNFRSF10B, LeX + LeX | |
| 559 | CD274, CEACAM6 + CEACAM6 | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 560 | CD274, CD9 + CD9 | |
| 561 | CD274, TFRC + TFRC | |
| 562 | CD274, CD55 + CD55 | |
| 563 | CD274, NT5E + NT5E | |
| 564 | CD274, PTGFRN + PTGFRN | |
| 565 | CD274, TACSTD2 + TACSTD2 | |
| 566 | CD274, ITGA2 + ITGA2 | |
| 567 | CD274, sLeX + sLeX | |
| 568 | CD274, FOLR1 + FOLR1 | |
| 569 | CD274, LEY + LEY | |
| 570 | CD274, LeX + LeX | |
| 571 | CD274, EGFR + EGFR | |
| 572 | CD274, PS + PS | |
| 573 | CD274, IL1RAP + IL1RAP | |
| 574 | CD274, MSLN + MSLN | |
| 575 | CD274, MUC4 + MUC4 | |
| 576 | CD274, Tn + Tn | |
| 577 | LAMB3, Tn + Tn | |
| 578 | LAMB3, CEACAM6 + CEACAM6 | |
| 579 | LAMB3, T + T | |
| 580 | LAMB3, LEY + LEY | |
| 581 | LAMB3, CD9 + CD9 | |
| 582 | LAMB3, TFRC + TFRC | |
| 583 | LAMB3, TSPAN8 + TSPAN8 | |
| 584 | LAMB3, ITGA2 + ITGA2 | |
| 585 | LAMB3, MUC4 + MUC4 | |
| 586 | LAMB3, BCAP31 + BCAP31 | |
| 587 | LAMB3, PTGFRN + PTGFRN | |
| 588 | LAMB3, TACSTD2 + TACSTD2 | |
| 589 | SDC1, Tn + Tn | |
| 590 | SDC1, LeX + LeX | |
| 591 | SDC1, sLeX + sLeX | |
| 592 | SDC1, CD9 + CD9 | |
| 593 | SDC1, T + T | |
| 594 | SDC1, TACSTD2 + TACSTD2 | |
| 595 | TFRC, CD9 + CD9 | |
| 596 | TFRC, FOLR1 + FOLR1 | |
| 597 | TFRC, CD274 + CD274 | |
| 598 | TFRC, NT5E + NT5E | |
| 599 | TFRC, TACSTD2 + TACSTD2 | LUSqC |
| 600 | TFRC, LEY + LEY | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 601 | TFRC, ITGA2 + ITGA2 | |
| 602 | TFRC, IL1RAP + IL1RAP | |
| 603 | TFRC, T + T | LUSqC |
| 604 | TFRC, MUC1 + MUC1 | |
| 605 | TFRC, STn + STn | LUNG |
| 606 | TFRC, MUC1 + MUC1 | |
| 607 | CD55, LeX + LeX | |
| 608 | CD55, CEACAM6 + CEACAM6 | |
| 609 | CD55, CD9 + CD9 | |
| 610 | CD55, TSPAN8 + TSPAN8 | |
| 611 | CD55, ITGA2 + ITGA2 | |
| 612 | CD55, MSLN + MSLN | |
| 613 | CD55, EPCAM + EPCAM | LUAD |
| 614 | CD55, EGFR + EGFR | |
| 615 | CD55, LEY + LEY | |
| 616 | CD55, NT5E + NT5E | |
| 617 | CD55, TACSTD2 + TACSTD2 | |
| 618 | CD55, T + T | |
| 619 | CD55, CD274 + CD274 | |
| 620 | CD55, BCAP31 + BCAP31 | |
| 621 | CD55, PS + PS | |
| 622 | CD55, MUC1 + MUC1 | LUAD |
| 623 | CD55, MUC1 + MUC1 | |
| 624 | CD55, PTGFRN + PTGFRN | |
| 625 | CD55, ALCAM + ALCAM | |
| 626 | CD55, sLeX + sLeX | |
| 627 | CD55, CEACAM5 + CEACAM5 | LUNG |
| 628 | CD55, MUC4 + MUC4 | |
| 629 | CD55, STn + STn | LUNG |
| 630 | CD55, MET + MET | |
| 631 | CD109, ITGA2 + ITGA2 | |
| 632 | CD109, CD9 + CD9 | |
| 633 | CD109, T + T | |
| 634 | CD109, CD55 + CD55 | |
| 635 | CD109, PTGFRN + PTGFRN | |
| 636 | CD109, EPCAM + EPCAM | |
| 637 | CD109, EGFR + EGFR | |
| 638 | CD109, LeX + LeX | |
| 639 | CD109, TSPAN8 + TSPAN8 | |
| 640 | CD109, IL1RAP + IL1RAP | |
| 641 | CD109, VWA1 + VWA1 | LUSqC |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 642 | CD109, CD274 + CD274 | |
| 643 | IGSF3, CD9 + CD9 | |
| 644 | IGSF3, LeX + LeX | |
| 645 | IGSF3, LEY + LEY | |
| 646 | IGSF3, EGFR + EGFR | |
| 647 | IGSF3, STn + STn | |
| 648 | IGSF3, MUC4 + MUC4 | |
| 649 | IGSF3, MSLN + MSLN | LUAD |
| 650 | IGSF3, IL1RAP + IL1RAP | |
| 651 | IGSF3, CEACAM5 + CEACAM5 | |
| 652 | IGSF3, NT5E + NT5E | |
| 653 | IGSF3, TSPAN8 + TSPAN8 | |
| 654 | IGSF3, Tn + Tn | |
| 655 | MET, CD9 + CD9 | |
| 656 | MET, T + T | LUSqC |
| 657 | MET, LeX + LeX | |
| 658 | MET, CD55 + CD55 | |
| 659 | MET, Tn + Tn | LUSqC |
| 660 | MET, NT5E + NT5E | |
| 661 | MET, EGFR + EGFR | LUSqC |
| 662 | MET, sLeX + sLeX | |
| 663 | MET, ALCAM + ALCAM | |
| 664 | MET, BCAP31 + BCAP31 | |
| 665 | MET, TFRC + TFRC | |
| 666 | MET, CEACAM6 + CEACAM6 | |
| 667 | DSC3, CD9 + CD9 | |
| 668 | DSC3, CEACAM6 + CEACAM6 | |
| 669 | DSC3, TACSTD2 + TACSTD2 | |
| 670 | DSC3, EGFR + EGFR | |
| 671 | DSC3, BCAP31 + BCAP31 | |
| 672 | DSC3, CEACAM5 + CEACAM5 | LUAD |
| 673 | DSC3, TFRC + TFRC | |
| 674 | DSC3, FOLR1 + FOLR1 | |
| 675 | DSC3, CD55 + CD55 | |
| 676 | DSC3, ITGA2 + ITGA2 | |
| 677 | DSC3, sLeX + sLeX | |
| 678 | DSC3, PS + PS | |
| 679 | DSC3, PTGFRN + PTGFRN | |
| 680 | DSC3, T + T | |
| 681 | DSC3, VWA1 + VWA1 | |
| 682 | DSC3, LeX + LeX | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 683 | DSC3, SLC34A2 + SLC34A2 | |
| 684 | DSC3, MARCKSL1 + MARCKSL1 | |
| 685 | DSC3, ALCAM + ALCAM | |
| 686 | DSC3, TSPAN8 + TSPAN8 | |
| 687 | DSC3, CD274 + CD274 | |
| 688 | DSC3, LEY + LEY | |
| 689 | DSC3, MSLN + MSLN | |
| 690 | DSC3, MUC1 + MUC1 | |
| 691 | IGF1R, CD9 + CD9 | |
| 692 | IGF1R, BCAP31 + BCAP31 | |
| 693 | IGF1R, CD55 + CD55 | |
| 694 | IGF1R, MUC1 + MUC1 | |
| 695 | IGF1R, TFRC + TFRC | |
| 696 | IGF1R, MUC1 + MUC1 | LUNG |
| 697 | IGF1R, LEY + LEY | |
| 698 | IGF1R, PS + PS | |
| 699 | IGF1R, PTGFRN + PTGFRN | |
| 700 | IGF1R, ALCAM + ALCAM | |
| 701 | IGF1R, FOLR1 + FOLR1 | |
| 702 | IGF1R, CEACAM5 + CEACAM5 | |
| 703 | IGF1R, NT5E + NT5E | |
| 704 | IGF1R, sLeX + sLeX | |
| 705 | IGF1R, Tn + Tn | |
| 706 | IGF1R, T + T | |
| 707 | IGF1R, MUC4 + MUC4 | |
| 708 | IGF1R, CEACAM6 + CEACAM6 | |
| 709 | MARCKSL1, ALCAM + ALCAM | |
| 710 | MARCKSL1, LEY + LEY | |
| 711 | MARCKSL1, EGFR + EGFR | LUNG |
| 712 | MARCKSL1, T + T | |
| 713 | MARCKSL1, NT5E + NT5E | |
| 714 | MARCKSL1, CD9 + CD9 | |
| 715 | BCAP31, TSPAN8 + TSPAN8 | |
| 716 | BCAP31, EPCAM + EPCAM | |
| 717 | BCAP31, LeX + LeX | |
| 718 | BCAP31, FOLR1 + FOLR1 | |
| 719 | BCAP31, CD9 + CD9 | |
| 720 | BCAP31, CEACAM6 + CEACAM6 | LUSqC |
| 721 | BCAP31, TACSTD2 + TACSTD2 | |
| 722 | BCAP31, LEY + LEY | |
| 723 | BCAP31, sLeX + sLeX | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 724 | BCAP31, NT5E + NT5E | |
| 725 | BCAP31, MUC1 + MUC1 | |
| 726 | BCAP31, MSLN + MSLN | |
| 727 | BCAP31, T + T | |
| 728 | BCAP31, MUC1 + MUC1 | LUAD |
| 729 | BCAP31, PS + PS | |
| 730 | BCAP31, MUC1 + MUC1 | |
| 731 | STn, EGFR + EGFR | LUAD |
| 732 | STn, SLC34A2 + SLC34A2 | LUAD |
| 733 | MUC1, SDC1 + SDC1 | LUAD |
| 734 | EPCAM, MSLN + MSLN | |
| 735 | EPCAM, Tn + Tn | |
| 736 | EPCAM, MUC1 + MUC1 | LUSqC |
| 737 | EPCAM, SLC34A2 + SLC34A2 | |
| 738 | FOLR1, SLC34A2 + SLC34A2 | |
| 739 | FOLR1, CEACAM5 + CEACAM5 | |
| 740 | SLC34A2, EPCAM + EPCAM | |
| 741 | TACSTD2, MET + MET | |
| 742 | TACSTD2, SLC34A2 + SLC34A2 | |
| 743 | TACSTD2, SDC1 + SDC1 | |
| 744 | T, CEACAM5 + CEACAM5 | LUAD |
| 745 | T, SLC34A2 + SLC34A2 | |
| 746 | LEY, T + T | |
| 747 | LEY, sLeX + sLeX | LUNG |
| 748 | LEY, Tn + Tn | LUAD |
| 749 | LEY, SLC34A2 + SLC34A2 | |
| 750 | LEY, CEACAM6 + CEACAM6 | LUAD |
| 751 | sLeX, LEY + LEY | |
| 752 | sLeX, SLC34A2 + SLC34A2 | |
| 753 | sLeX, PS + PS | |
| 754 | sLeX, EGFR + EGFR | LUAD |
| 755 | MUC1, PS + PS | |
| 756 | MUC1, MARCKSL1 + MARCKSL1 | |
| 757 | MUC1, T + T | LUSqC |
| 758 | MUC1, CEACAM5 + CEACAM5 | LUNG |
| 759 | MUC1, CD274 + CD274 | |
| 760 | MUC1, SLC34A2 + SLC34A2 | |
| 761 | CEACAM5, CD274 + CD274 | |
| 762 | TRPV4, LeX + LeX | |
| 763 | TRPV4, SDC1 + SDC1 | |
| 764 | PTGFRN, LEY + LEY | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 765 | PTGFRN, FOLR1 + FOLR1 | |
| 766 | PTGFRN, Tn + Tn | |
| 767 | EGFR, CD274 + CD274 | |
| 768 | EGFR, TACSTD2 + TACSTD2 | |
| 769 | EGFR, LEY + LEY | |
| 770 | EGFR, CEACAM5 + CEACAM5 | |
| 771 | EGFR, MUC1 + MUC1 | |
| 772 | TSPAN8, SLC34A2 + SLC34A2 | |
| 773 | MUC4, MSLN + MSLN | LUAD |
| 774 | MUC4, T + T | |
| 775 | TNFRSF10B, TACSTD2 + TACSTD2 | |
| 776 | TNFRSF10B, STn + STn | LUNG |
| 777 | TNFRSF10B, MUC1 + MUC1 | LUAD |
| 778 | TNFRSF10B, MUC1 + MUC1 | |
| 779 | CD274, CEACAM5 + CEACAM5 | |
| 780 | CD274, T + T | |
| 781 | CD274, STn + STn | |
| 782 | SDC1, MUC1 + MUC1 | LUAD |
| 783 | SDC1, CEACAM6 + CEACAM6 | LUAD |
| 784 | TFRC, MUC1 + MUC1 | |
| 785 | TFRC, CEACAM6 + CEACAM6 | LUSqC |
| 786 | TFRC, LeX + LeX | fail |
| 787 | TFRC, sLeX + sLeX | LUAD |
| 788 | TFRC, MUC4 + MUC4 | |
| 789 | CD55, SLC34A2 + SLC34A2 | |
| 790 | IGSF3, CD274 + CD274 | LUNG |
| 791 | MET, TACSTD2 + TACSTD2 | |
| 792 | DSC3, MUC1 + MUC1 | |
| 793 | DSC3, MET + MET | |
| 794 | DSC3, Tn + Tn | |
| 795 | DSC3, SDC1 + SDC1 | |
| 796 | IGF1R, SLC34A2 + SLC34A2 | |
| 797 | IGF1R, STn + STn | |
| 798 | IGF1R, CD274 + CD274 | |
| 799 | IGF1R, EGFR + EGFR | LUNG |
| 800 | IGF1R, SDC1 + SDC1 | |
| 801 | PTK7, T + T | |
| 802 | PTK7, MUC4 + MUC4 | |
| 803 | MARCKSL1, CD274 + CD274 | |
| 804 | MARCKSL1, FOLR1 + FOLR1 | |
| 805 | MARCKSL1, PTGFRN + PTGFRN | |

| # | Exemplary combination (Capture biomarker, Detection biomarker 1 + Detection biomarker 2) | Observation |
|---|---|---|
| 806 | MARCKSL1, LeX + LeX | |
| 807 | CEACAM5, KRT19 + KRT19 | |
| 808 | EGFR, KRT19 + KRT19 | |
| 809 | EPCAM, KRT19 + KRT19 | |
| 810 | LeX, KRT19 + KRT19 | |
| 811 | MUC1, KRT19 + KRT19 | |
| 812 | MUC4, KRT19 + KRT19 | |
| 813 | SDC1, KRT19 + KRT19 | |
| 814 | sLeX, KRT19 + KRT19 | |
| 815 | STn, KRT19 + KRT19 | |
| 816 | TACSTD2, KRT19 + KRT19 | |
| 817 | T, KRT19 + KRT19 | |
| 818 | TRPV4, KRT19 + KRT19 | |
| 819 | CEACAM6, KRT19 + KRT19 | |
| 820 | KRT19, CEACAM5 + CEACAM5 | |
| 821 | KRT19, EGFR + EGFR | |
| 822 | KRT19, LeX + LeX | |
| 823 | KRT19, MUC4 + MUC4 | |
| 824 | KRT19, PTGFRN + PTGFRN | |
| 825 | KRT19, SDC1 + SDC1 | |
| 826 | KRT19, sLeX + sLeX | |
| 827 | KRT19, STn + STn | |
| 828 | KRT19, TACSTD2 + TACSTD2 | |
| 829 | KRT19, T + T | |
| 830 | KRT19, CEACAM6 + CEACAM6 | LUAD |
| 831 | EPCAM, SDC1 + SDC1 | |
| 832 | T, SDC1 + SDC1 | |
| 833 | CEACAM5, SDC1 + SDC1 | |
| 834 | EGFR, SDC1 + SDC1 | |

**[602]** Additional cutoff criteria were evaluated for certain biomarker combinations disclosed in Table 4B including:

- A cutoff of 3 Ct for both LUAD and LUSqC pools
- A cutoff of 3 for the LUAD pool and a cutoff of 2 for the LUSqC pool

**Table 4C – Summary of biomarker combination assessments based on the results of Table 4B using various ΔCt cutoffs, as indicated above.**

| Classification | ΔCt Cutoff | | |
|---|---|---|---|
| | 2 (both pools) | 3 (both pools) | 3 (LUAD), 2 (LUSqC) |
| LUNG | 53 | 28 | 42 |
| LUAD | 98 | 78 | 64 |
| LUSqC | 35 | 17 | 46 |
| Unclassified | 646 | 702 | 780 |

**[603]** FIGS. 16A-20B and Table 5 below summarizes results of using certain duplex biomarker combinations as listed in Table 4A and/or Table 4B above to detect lung cancer, *e.g.*, different types and/or stages of lung cancer. Assays were performed as described in Example 1 and Example 2. The results shown in Table 5 compare healthy non-smoker pooled samples, healthy smoker pooled samples, early stage lung adenocarcinoma pooled samples, late stage lung adenocarcinoma pooled samples, early stage lung squamous cell carcinoma pooled samples, and late stage lung squamous cell carcinoma pooled samples. Suitable positive and negative controls were conducted for each biomarker combination to validate the assay worked appropriately (*e.g.*, positive control samples comprising extracellular vesicles from lung cancer cell lines and negative control samples comprising no extracellular vesicles). The results documented that certain biomarker combinations successfully differentiate lung cancer samples from noted reference samples (for example, when measured by at least a 2 fold difference in amplification performed as described in Example 1 with the biomarker combinations depicted).

**[604]** Healthy Nonsmokers Pool included serum samples from subjects (*e.g.*, female and male subjects) who have no medical history of lung cancer. In some instances, such subjects may have a medical condition that is not lung cancer or is not associated with a lung disorder, including, *e.g.*, chronic gastritis, coronary artery disease, esophageal ulcer, hypertension, osteoarthritis, chronic cholecystitis, varicose veins, obesity, chronic prostatitis, and/or hypertension. All samples were from Caucasians, aged 55 to 78 years, and with body mass indexes (BMIs) ranging from 22 to 31.

**[605]** Healthy Smokers Pool included serum samples from subjects (*e.g.*, female and male subjects) who have no medical history of lung cancer but has a history of smoking equating to 20 to 30 pack years. In some instances, such subjects may have a medical condition that is not lung cancer or is not associated with a lung disorder, including, *e.g.*, chronic bronchitis, hypertension, and/or gastroduodenitis. All samples were from Caucasians, aged 50 to 67 years, and with BMIs ranging from 23 to 29.

**[606]** Early Stage LUAD Pool included serum samples from subjects (*e.g.*, female and male subjects) who have been determined to have early stage LUAD. All samples were from Caucasians, aged 36 to 77 years, with a varied history of smoking ranging from non-smokers to 45 pack years, and with BMIs ranging from 19 to 41. In some instances, such subjects had at least one or more of the following medical conditions (in addition to early stage LUAD): hypertension, obesity, asthma, ischemic heart disease, kidney cancer, atherosclerosis, diabetes mellitus type 2, chronic bronchitis, uterine cancer, breast cancer, chronic pyelonephritis, cervical cancer, gastric ulcer, renal cancer, chronic gastritis, and/or osteoarthritis.

**[607]** Late Stage LUAD Pool included serum samples from subjects (*e.g.*, female and male subjects) who have been determined to have late stage LUAD. All samples were from Caucasians, aged 51 to 72 years, with a varied history of smoking ranging from non-smokers to 54 pack years, and with BMIs ranging from 19 to 40. In some instances, such subjects had at least one or more of the following medical conditions (in addition to late stage LUAD): cerebrovascular disease, ischemic heart disease, atherosclerosis, hypertension, coronary artery disease, obesity, nodular goiter, chronic cholecystitis, diabetes mellitus type 2, stroke, chronic bronchitis, chronic gastritis, angina, thyrotoxicosis, and/or chronic obstructive bronchitis.

**[608]** Early Stage LUSC Pool included serum samples from subjects (*e.g.*, female and male subjects) who have been determined to have early stage LUSC. All samples were from Caucasians, aged 49 to 82 years, with a varied history of smoking ranging from non-smokers to 50 pack years, and with BMIs ranging from 19 to 33. In some instances, such subjects had at least one or more of the following medical conditions (in addition to early stage LUSC): stroke, chronic bronchitis, diabetes mellitus type 2, atherosclerosis, chronic atrophic gastritis, coronary artery disease, hypertension, chronic gastritis, chronic venous insufficiency, chronic obstructive pulmonary disease, steatohepatitis, gastric cancer, paroxysmal atrial fibrillation, kidney cancer, sick sinus syndrome, irregular heart rhythm, sensory hearing loss, vibration disease, colitis, impaired glucose tolerance, peptic ulcer, duodenal ulcer, and/or obesity.

**[609]** Late Stage LUSC Pool included serum samples from subjects (*e.g.*, female and male subjects) who have been determined to have late stage LUSC. All samples were from Caucasians, aged 49 to 66 years, with history of smoking ranging from 20 to 52 pack years, and with BMIs ranging from 23 to 31. In some instances, such subjects had at least one or more of the following medical conditions (in addition to late stage LUSC): chronic bronchitis,

pneumosclerosis, pulmonary failure, chronic gastritis, chronic alcoholism, gastric ulcer, coronary artery disease, atherosclerosis, emphysema, duodenal ulcer, chronic obstructive pulmonary disease, varicose vein disease, pain syndrome, hypertension, and/or atherosclerotic cardiosclerosis.

**[610]** The present Example illustrates that assays described herein were capable of distinguishing lung cancer-derived EVs from those originating from healthy tissues using certain biomarker signatures provided herein. Thus, in some embodiments, certain biomarker signatures provided herein (*e.g.*, as listed in Table 4A and/or Table 4B) are particularly useful for detecting lung cancer, or particularly non-small cell lung cancer (*e.g.*, lung adenocarcinoma and/or lung squamous cell carcinoma). In certain embodiments, biomarker combinations listed in Table 4A and/or Table 4B in combination with EV assays as described herein were able to detect lung cancer, for example, in some embodiments with a signal at least two times greater than a pool of plasma samples from healthy individuals (*e.g.*, nonsmokers and/or smokers).

**[611]** As shown below, in some embodiments, certain biomarker combinations distinguished between the late-stage lung cancer pool and the healthy pool (non-smokers and/or smokers). Among other things, such findings demonstrate that provided biomarker combinations, used in accordance with the present disclosure, distinguish lung malignancies from non-cancerous conditions, thereby documenting a particularly advantageous feature of disclosed technologies. In some situations, such feature is particularly beneficial, as lung tumors can generate similar symptoms to other diseases associated with smoking, such as chronic bronchitis and/or chronic obstructive pulmonary disorders, and are often difficult to differentiate using conventional screening methods.

**[612]** As shown below, in some embodiments, certain biomarker combinations distinguished the early-stage lung cancer pool from the healthy pool (non-smokers and smokers). Among other things, such findings demonstrate that provided biomarker combinations, used in accordance with the present disclosure, distinguish lung malignancies over individuals without lung cancer, thereby documenting a particularly advantageous feature of disclosed technologies. In some situations, such feature is particularly beneficial, as correctly differentiating lung cancer samples from healthy individuals (smokers and/or non-smokers) at an early stage can increase the chance of successfully combating the disease while maintaining patient wellbeing.

**[613]** Among other things, results presented in the present Example demonstrate that provided technologies detect lung cancer (*e.g.*, non-small cell lung cancer such as, *e.g.*, lung adenocarcinoma and/or lung squamous cell carcinoma), and specifically demonstrate that provided technologies (*e.g.,* using biomarker combinations as described herein to detect biomarkers in and/or on EVs) distinguish lung cancer-derived EVs from EVs that are derived from non-lung cancer tissues.

**[614]** In some embodiments, a target biomarker signature provided herein for detection of lung cancer can comprise at least three target biomarkers, at least one of which is a surface biomarker as described herein. In some embodiments, at least two or at least three of such target biomarkers are distinct surface biomarkers described herein. In some embodiments, a target biomarker signature that can be useful for detection of lung cancer (*e.g.*, non-small cell lung cancer) comprises TNFRSF10B, MUC1, and CEACAM6. In some embodiments, a target biomarker signature that can be useful for detection of lung cancer (*e.g.*, non-small cell lung cancer) comprises DSG2, MUC1, and CEACAM6. In some embodiments, a target biomarker signature that can be useful for detection of lung cancer (*e.g.*, non-small cell lung cancer) comprises CEACAM6, MUC1, and sTn antigen. In some embodiments, a target biomarker signature that can be useful for detection of lung cancer (*e.g.*, non-small cell lung cancer) comprises SLC34A2, MUC1, and CEACAM6. In some embodiments, a target biomarker signature that can be useful for detection of lung cancer (*e.g.*, non-small cell lung cancer) comprises CEACAM6, MUC1, and MSLN. In some embodiments, a target biomarker signature that can be useful for detection of lung cancer (*e.g.*, non-small cell lung cancer) comprises FOLR1, T antigen, and EGFR.

**[615]** In some embodiments, target biomarker signatures described herein may be detected by an assay described herein (*e.g.*, a proximity-based ligation assay described herein). In some embodiments, a proximity-based ligation assay can comprise a capture probe directed to TNFRSF10B, a detection probe directed to MUC1, and a detection probe directed to CEACAM6. In certain embodiments, such a proximity-based ligation assay described herein was able to distinguish early stage lung cancer (*e.g.*, early stage non-small cell lung cancer such as, *e.g.*, LUAD) from healthy nonsmokers and healthy smokers (see *e.g.*, FIGS. 21A-24B; and Table 6. In some embodiments, a proximity-based ligation assay can comprise a capture probe directed to DSG2, a detection probe directed to MUC1, and a detection probe directed to CEACAM6. In

certain embodiments, such a proximity-based ligation assay described herein was able to distinguish early-stage lung cancer (*e.g.*, early stage non-small cell lung cancer such as, *e.g.*, LUAD) from healthy nonsmokers and healthy smokers (see *e.g.*, FIGS. 21A-24B; and Table 6). In some embodiments, a proximity-based ligation assay can comprise a capture probe directed to CEACAM6, a detection probe directed to MUC1, and a detection probe directed to sTn antigen. In certain embodiments, such a proximity-based ligation assay described herein was able to distinguish early-stage lung cancer (*e.g.*, early-stage non-small cell lung cancer such as, *e.g.*, LUAD) from healthy nonsmokers and healthy smokers (see *e.g.*, FIGS. 21A-24B; and Table 6). In some embodiments, a proximity-based ligation assay can comprise a capture probe directed to SLC34A2, a detection probe directed to MUC1, and a detection probe directed to CEACM6. In certain embodiments, such a proximity-based ligation assay described herein was able to distinguish early-stage lung cancer (*e.g.*, early-stage non-small cell lung cancer such as, *e.g.*, LUAD) from healthy nonsmokers and healthy smokers (see *e.g.*, FIGS. 21A-24B; and Table 6). In some embodiments, a proximity-based ligation assay can comprise a capture probe directed to CEACM6, a detection probe directed to MUC1, and a detection probe directed to MSLN. In certain embodiments, such a proximity-based ligation assay described herein was able to distinguish late-stage lung cancer (*e.g.*, late-stage non-small cell lung cancer such as, *e.g.*, LUAD) from healthy nonsmokers and healthy smokers (see *e.g.*, FIGS. 21A-24B; and Table 6). In some embodiments, a proximity-based ligation assay can comprise a capture probe directed to FOLR1, a detection probe directed to T antigen, and a detection probe directed to EGFR. In certain embodiments, such a proximity-based ligation assay described herein was able to distinguish late-stage lung cancer (*e.g.*, late-stage non-small cell lung cancer such as, *e.g.*, LUAD) from healthy nonsmokers (see *e.g.*, FIGS. 21A-24B; and Table 6).

**Table 5 - Detection of lung cancer using certain exemplary 2-biomarker combinations**

| Biomarker Combination | Differentiates Early Stage (ES) or Late Stage (LS) LUAD Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | | Differentiates Early Stage (ES) or Late Stage (LS) LUSC Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | |
|---|---|---|---|---|---|---|---|---|
| | ES vs S | LS vs S | ES vs NS | LS vs NS | ES vs S | LS vs S | ES vs NS | LS vs NS |
| TNFRSF10B + PD-L1 | * | * | * | * | * | * | * | * |
| TNFRSF10B + CEACAM6 | * | * | * | * | | * | * | * |
| TNFRSF10B + EGFR | * | * | | * | * | * | * | * |
| TNFRSF10B + IGF1R | * | * | * | * | * | | * | * |
| ALCAM + EPCAM | | | * | * | * | * | * | * |
| CEACAM6 + MUC1 | * | * | * | * | | * | | * |
| EGFR + T antigen | | * | | * | * | * | * | * |
| EPCAM + T antigen | | * | | * | * | * | * | * |
| FOLR1 + T antigen | | * | | * | * | * | * | * |
| Tn antigen + TACSTD2 | * | * | * | * | * | * | | |
| TNFRSF10B + FOLR1 | * | | * | * | | * | * | * |
| TNFRSF10B + sTn antigen | * | | * | * | | * | * | * |
| ALCAM + PD-L1 | | * | | * | * | * | * | |
| EPCAM + MUC1 | | * | | * | * | * | * | |

| Biomarker Combination | Differentiates Early Stage (ES) or Late Stage (LS) LUAD Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | | Differentiates Early Stage (ES) or Late Stage (LS) LUSC Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | |
|---|---|---|---|---|---|---|---|---|
| | ES vs S | LS vs S | ES vs NS | LS vs NS | ES vs S | LS vs S | ES vs NS | LS vs NS |
| TNFRSF10B + CD55 | | * | * | * | * | | * | |
| TNFRSF10B + MUC1 | * | * | * | * | * | | | |
| FOLR1 + TACSTD2 | | | | | * | * | * | * |
| MET + MUC1 | * | * | | * | * | | | |
| MET + sTn antigen | | * | | * | | * | | * |
| MUC1 + TACSTD2 | | * | | * | | * | | * |
| PD-L1 + MUC1 | * | * | | * | * | | | |
| PD-L1 + Tn antigen | * | * | | * | * | | | |
| SLC34A2 + MET | | * | | * | | | * | * |
| SLC34A2 + T antigen | | | | * | * | | * | * |
| TNFRSF10B + CEACAM5 | | | * | * | | | * | * |
| TNFRSF10B + MSLN | | * | | * | | | * | * |
| TNFRSF10B + Tn antigen | | | * | * | | | * | * |
| ALCAM + T antigen | | * | | * | | | | * |
| ALCAM + TACSTD2 | | * | | | | * | | * |

| Biomarker Combination | Differentiates Early Stage (ES) or Late Stage (LS) LUAD Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | | Differentiates Early Stage (ES) or Late Stage (LS) LUSC Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | |
|---|---|---|---|---|---|---|---|---|
| | ES vs S | LS vs S | ES vs NS | LS vs NS | ES vs S | LS vs S | ES vs NS | LS vs NS |
| CD55 + PD-L1 | * | * | | | * | | | |
| CDH3 + CDH1 | | | | | * | * | * | |
| CEACAM5 + MUC1 | | * | | * | | * | | |
| CEACAM6 + EGFR | | * | | | * | * | | |
| CEACAM6 + EPCAM | | * | | | | * | | * |
| CEACAM6 + FOLR1 | | * | | * | | * | | |
| CEACAM6 + MSLN | | * | | * | | | | * |
| CEACAM6 + sTn antigen | | | * | | | * | | * |
| CEACAM6 + TACSTD2 | | | | * | | * | | * |
| DSG2 + CEACAM6 | * | | * | | | | | * |
| EGFR + MUC1 | | * | | * | * | | | |
| FOLR1 + MUC1 | | * | | * | * | | | |
| MUC1 + sTn antigen | | * | | * | | * | | |
| SLC34A2 + MSLN | | * | | | | * | | * |
| T antigen + CDH1 | | | | * | | | * | * |

| Biomarker Combination | Differentiates Early Stage (ES) or Late Stage (LS) LUAD Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | | Differentiates Early Stage (ES) or Late Stage (LS) LUSC Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | |
|---|---|---|---|---|---|---|---|---|
| | ES vs S | LS vs S | ES vs NS | LS vs NS | ES vs S | LS vs S | ES vs NS | LS vs NS |
| Tn antigen + IGF1R | * | * | | | * | | | |
| TNFRSF10B + ALCAM | | | * | * | | | * | |
| CD55 + EPCAM | * | | * | | | | | |
| CEACAM6 + PD-L1 | | * | | | * | | | |
| CEACAM6 + SLC34A2 | | * | | * | | | | |
| DSG2 + CEACAM5 | | | * | | | | | * |
| DSG2 + EPCAM | | | * | | | | | * |
| DSG2 + MET | | | * | * | | | | |
| DSG2 + T antigen | | * | | * | | | | |
| EGFR + sTn antigen | | * | | | | * | | |
| EGFR + TACSTD2 | | | | * | | | | * |
| EPCAM + TACSTD2 | | | | | | * | | * |
| IGF1R + T antigen | | | | | * | | * | |
| MET + Tn antigen | | * | | * | | | | |
| MSLN + T antigen | | * | | * | | | | |

| Biomarker Combination | Differentiates Early Stage (ES) or Late Stage (LS) LUAD Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | | Differentiates Early Stage (ES) or Late Stage (LS) LUSC Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | |
|---|---|---|---|---|---|---|---|---|
| | ES vs S | LS vs S | ES vs NS | LS vs NS | ES vs S | LS vs S | ES vs NS | LS vs NS |
| PD-L1 + sTn antigen | | | | | | * | | * |
| PD-L1 + T antigen | * | | * | | | | | |
| SLC34A2 + MUC1 | | * | | * | | | | |
| SLC34A2 + PD-L1 | | * | | | * | | | |
| SLC34A2 + sTn antigen | | | | | | * | | * |
| SLC34A2 + Tn antigen | | * | | * | | | | |
| TACSTD2 + sTn antigen | | | | | | * | | * |
| TNFRSF10B + CDH3 | | | * | | | | | * |
| TNFRSF10B + EPCAM | | | | | | * | | * |
| TNFRSF10B + MET | * | | * | | | | | |
| ALCAM + CDH3 | | | | | | * | | |
| ALCAM + MET | | | | | | | | * |
| ALCAM + MUC1 | | * | | | | | | |
| ALCAM + SLC34A2 | | | | | | | * | |
| ALCAM + sTn antigen | | | | | | | | * |

| Biomarker Combination | Differentiates Early Stage (ES) or Late Stage (LS) LUAD Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | | Differentiates Early Stage (ES) or Late Stage (LS) LUSC Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | |
|---|---|---|---|---|---|---|---|---|
| | ES vs S | LS vs S | ES vs NS | LS vs NS | ES vs S | LS vs S | ES vs NS | LS vs NS |
| CD55 + CDH1 | | | | * | | | | |
| CD55 + MET | | | * | | | | | |
| CDH3 + T antigen | | | | | | * | | |
| CEACAM5 + MSLN | | | | * | | | | |
| CEACAM5 + TACSTD2 | | | | | | * | | |
| CEACAM6 + Tn antigen | * | | | | | | | |
| DSG2 + CDH1 | | | * | | | | | |
| DSG2 + CDH3 | | * | | | | | | |
| DSG2 + MUC1 | | | * | | | | | |
| DSG2 + SLC34A2 | | | * | | | | | |
| DSG2 + sTn antigen | | | * | | | | | |
| EGFR + MSLN | | | | * | | | | |
| FOLR1 + sTn antigen | | * | | | | | | |
| FOLR1 + Tn antigen | | | | * | | | | |
| MET + T antigen | | | | | | | * | |

| Biomarker Combination | Differentiates Early Stage (ES) or Late Stage (LS) LUAD Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | | Differentiates Early Stage (ES) or Late Stage (LS) LUSC Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | |
|---|---|---|---|---|---|---|---|---|
| | ES vs S | LS vs S | ES vs NS | LS vs NS | ES vs S | LS vs S | ES vs NS | LS vs NS |
| MSLN + sTn antigen | | | | | | | | * |
| MUC1 + T antigen | | | | | | | | * |
| MUC1 + Tn antigen | | * | | | | | | |
| SLC34A2 + EGFR | | * | | | | | | |
| TNFRSF10B + CDH1 | | * | | | | | | |
| TNFRSF10B + SLC34A2 | | | | | | | | * |

*Table 5* depicting evaluation of certain exemplary two target biomarker combinations for detection of lung cancer (*e.g.*, non-small cell lung cancer). Comparisons are of Early Stage (ES) or Late Stage (LS) Lung adenocarcinoma (LUAD) or Lung squamous cell carcinoma (LUSC) to healthy Smoker (S) sample pools or healthy Nonsmoker (NS) sample pools. Rows depict biomarker combinations that differentiate lung cancer samples from healthy smoker or nonsmoker samples. Detection/differentiation is denoted as * and corresponds to at least a 2 fold amplification difference (*e.g.*, at least 3 fold amplification difference, at least 4 fold amplification difference, *etc.*) between cancerous samples when compared to the noted reference samples. It should be noted that the two biomarkers of such certain exemplary 2-biomarker combinations can be utilized as a target of a capture probe and/or a detection probe for a provided proximity-based ligation assay.

**Table 6 – Detection of lung cancer using certain exemplary 3-biomarker combinations**

| | Differentiates Early Stage (ES) or Late Stage (LS) LUAD Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | | Differentiates Early Stage (ES) or Late Stage (LS) LUSC Samples From Healthy Smoker (S) or Healthy Nonsmoker (NS) Samples? | | | |
|---|---|---|---|---|---|---|---|---|
| **Biomarker Combination** | **ES vs S** | **LS vs S** | **ES vs NS** | **LS vs NS** | **ES vs S** | **LS vs S** | **ES vs NS** | **LS vs NS** |
| CEACAM6, MUC1 + sTn antigen | * | * | * | * | * | * | * | * |
| TNFRSF10B, MUC1 + CEACAM6 | * | * | * | * | | * | | * |
| SLC34A2, MUC1 + CEACAM6 | * | * | | * | * | * | | |
| CEACAM6, MUC1 + MSLN | | * | | * | | * | | * |
| FOLR1, T antigen + EGFR | | | | * | | | | |
| DSG2, MUC1 + CEACAM6 | * | * | | | | | | |

*Table 6* depicting evaluation of certain exemplary three target biomarker combinations for detection of lung cancer (*e.g.*, non-small cell lung cancer). Comparisons are of Early Stage (ES) or Late Stage (LS) Lung adenocarcinoma (LUAD) or Lung squamous cell carcinoma (LUSC) to healthy Smoker (S) sample pools or healthy Nonsmoker (NS) sample pools. Rows depict biomarker combinations that differentiate lung cancer samples from healthy smoker or nonsmoker samples. Detection/differentiation is denoted as * and corresponds to at least a 2 fold amplification difference (*e.g.*, at least 3 fold amplification difference, at least 4 fold amplification difference, *etc.*) between cancerous samples when compared to the noted reference samples. For each biomarker combination, the first biomarker listed is a target of a capture probe, while the following biomarkers listed are targets of detection probes. It should be noted that a biomarker may be suitable as a target for one or more detection probes and/or a capture for a provided proximity-based ligation assay.

**Example 11: Further characterization of exemplary lung cancer liquid biopsy assays**

**[616]** This example is directed towards further characterization of various biomarker combinations in exemplary lung cancer liquid biopsy assays using different cell populations and patient populations. Useful biomarker combinations for detection of lung cancer can be determined by screening various combinations of biomarkers disclosed herein (*e.g.,* surface biomarkers *e.g.,* as described in examples 1-6), in pooled control and patient plasma samples. Pooled samples can provide an estimation for the average assay signal among various patient cohorts, facilitating biomarker combination triage.

**[617]** In some embodiments, a biomarker combination for detection of lung cancer can comprise one, two, three, four, five, six, seven or more biomarkers (*e.g.,* ones described herein), wherein such a combination comprises at least one biomarker for capturing extracellular vesicles (*e.g.,* by immunoaffinity capture) and at least one biomarker (including, *e.g.,* one, two, three, four, five, six, seven, or more biomarkers) for detection by pliq-PCR assay. In some embodiments, a target biomarker for capturing extracellular vesicles (*e.g.,* immunoaffinity capture) can be same as at least one target biomarker for pliq-PCR based analysis. In some embodiments, target biomarkers for capturing extracellular vesicles and pliq-PCR based analysis can each be distinct.

**[618]** Various biomarker combinations can be validated in pooled patient cohorts. Patient cohorts can include appropriate patient and/or control populations. In certain embodiments, a patient cohort can comprise healthy non-smoker subjects (*e.g.,* healthy subjects aged between 45 and 85 years of age), healthy subjects with a history of smoking (*e.g.,* healthy subjects aged between 45 and 85 years of age), subjects with late stage lung cancer (*e.g.*, stage III and stage IV non-small cell lung cancer such as, *e.g.*, lung adenocarcinoma or lung squamous cell carcinoma ), subjects with early stage lung cancer (*e.g.*, stage I and II non-small cell lung cancer such as, *e.g.*, lung adenocarcinoma or lung squamous cell carcinoma), subjects with benign tumors, subjects with inflammatory diseases, disorders, or conditions, and/or subjects with other cancers.

**[619]** In some embodiments, a two-step screening procedure can characterize performance of various biomarker combinations. For example, various biomarker combinations (*e.g.,* various combinations of biomarkers for a capture probe (*e.g.,* for immunoaffinity capture) and detection probes) can initially be screened in healthy background pools (from various age

groups) and an advanced stage lung cancer (*e.g.*, LUAD and/or LUSC) pool. In some embodiments, biomarker combinations that exhibit marginal separation between the healthy and lung cancer pools (*e.g.,* a ΔCt less than 4, corresponding to less than a 16-fold difference) can be eliminated from further study as a biomarker combination to use in isolation, however, such biomarker combinations may be useful when combined with additional biomarker combinations as described herein. In some embodiments, biomarker combinations that exhibit marginal separation between the healthy and lung cancer pools (*e.g.,* a ΔCt less than 2, corresponding to less than a 4-fold difference) can be eliminated from further study as a biomarker combination to use in isolation, however, such biomarker combinations may be useful when combined with additional biomarker combinations as described herein. In some embodiments, biomarker combinations that exhibit marginal separation between the healthy and lung cancer pools (*e.g.,* a ΔCt less than 1, corresponding to less than a 2-fold difference) can be eliminated from further study as a biomarker combination to use in isolation, however, such biomarker combinations may be useful when combined with additional biomarker combinations as described herein.

**[620]** In some embodiments, a two-step screening procedure can characterize performance of various biomarker combinations. For example, various biomarker combinations (*e.g.,* various combinations of biomarkers for a capture probe (*e.g.,* for immunoaffinity capture) and detection probes) can initially be screened in healthy background pools (from various age groups) and an advanced stage lung cancer pool. In some embodiments, biomarker combinations that exhibit good diagnostic performance (*e.g.,* a ΔCt greater than 1, corresponding to greater than a 2-fold difference) can undergo a second round of screening with pooled samples from early stage lung cancer, benign tumors, other cancers, and/or inflammatory conditions. In some embodiments, biomarker combinations that exhibit good diagnostic performance (*e.g.,* a ΔCt greater than 2, corresponding to greater than a 4-fold difference) can undergo a second round of screening with pooled early stage lung cancer, benign tumors, other cancers, and/or inflammatory conditions. In some embodiments, biomarker combinations that exhibit good diagnostic performance (*e.g.,* a ΔCt greater than 4, corresponding to greater than a 16-fold difference) can undergo a second round of screening with pooled early stage lung cancer, benign tumors, other cancers, and/or inflammatory conditions. In some embodiments, top biomarker combinations (*e.g.,* the best performing approximately 20 biomarker combinations) that can distinguish early and late stage lung cancer from the control pools can be further evaluated.

**[621]** Incorporation of one or more additional biomarker combinations in an exemplary assay (*e.g.,* ones described herein) can improve its sensitivity. As noted above, by utilizing at least two biomarker combinations, improved performance of an assay (*e.g.,* at least approximately 10% sensitivity at 99% specificity; or at least approximately 80% sensitivity at 98% specificity; or at least approximately 70% sensitivity at 99.5% specificity) can be achieved. Pooled patient samples can approximate the assay signals across individual patient populations, a trait which can provide a realistic matrix to assess a large number of biomarker combinations in an efficient manner. The top biomarker combinations (*e.g.,* up to 20 biomarker combinations) identified can be further tested in individual patient plasma based pilot studies. In some embodiments, an individual patient plasma based pilot study can comprise control patients who are healthy smokers or nonsmokers, control patients who have conditions that induce symptoms similar to lung cancer, control patients who have other types of cancer, and/or control patients with benign tumors. In some embodiments, an individual patient plasma based pilot study can comprise test patients who are either healthy smokers or nonsmokers, test patients who are symptomatic, test patients who are asymptomatic, test patients with stage I or stage II lung cancer, and/or test patients with stage III or stage IV lung cancer. Non-lung cancer health conditions can be aged-matched to the lung cancer cohort. It can be expected that the control samples (*e.g.,* healthy controls, benign tumors, and/or other off-target conditions, *etc.*) can provide an estimate of a log-normal distribution to set signal cutoffs pertaining to a 98% specificity (with approximately 95%) for hereditary risk and/or lifestyle associated risk screening assays and 99.5% specificity (with approximately) for symptomatic triaging assays. Biomarker combination performance characteristics can be evaluated using bivariate associations between combinations to assess independence, and top biomarker combinations can be further tested using a three-variable logistic regression model for predicting lung cancer.

**[622]** The identification of novel biomarker combinations that can distinguish lung cancer cases from controls and identify orthogonality between combinations can be achieved. This orthogonality can aid in distinguishing lung cancer cases from the control cohorts, increasing the sensitivity of exemplary assays as described herein. To assess racial diversity in the patient cohort, samples can be obtained from vendors which source from diverse populations. In certain embodiments, exemplary biomarker combinations may be specific to particular racial and/or ethnic groups. In certain embodiments, testing exemplary biomarker combinations in

diverse racial and/or ethnic groups can identify biomarker combinations of appropriate diagnostic value for specific racial and/or ethnic groups.

**[623]** Validation of additional lung cancer biomarker combinations in primary patient samples can improve the diagnostic performance of exemplary assays. In some embodiments, through the incorporation of additional, orthogonal biomarker combinations, assay performance can improve beyond 50% sensitivity or higher at 99% specificity or higher. In some embodiments, through incorporation of additional, orthogonal biomarker combinations, assay performance can assay performance can improve beyond 70% sensitivity or higher at 98% specificity and 60% sensitivity or higher at 99.5% specificity.

**Example 12: Validation of exemplary lung cancer liquid biopsy assays**

**[624]** This example relates to the validation of exemplary lung cancer liquid biopsy assays in additional populations/cohorts. Independent validation studies to assess exemplary lung cancer diagnostic assays as described herein using additional cohorts of patient samples can be performed. Samples can be obtained from any appropriate source. Technical experts are blinded to sample designations prior to any result analysis, and/or assay results are analyzed by an outside independent technical expert. To ensure sampling from a population representative of the United States, at least approximately 20% of samples can be sourced from non-white ethnicities (United States Census Bureau, 2018), depending on sample availability.

**[625]** In certain embodiments, patient cohorts analyzed can include: patients at hereditary and/or lifestyle associated risk for lung cancer prior to undertaking any risk-reducing operations, and appropriate relative control cohorts. In certain embodiments, a biomarker combination (*e.g.,* ones described herein) can provide at least an approximately 80% sensitivity at approximately 98% specificity in subjects with hereditary and/or life-style associated risk. In certain embodiments, a biomarker combination (*e.g.,* ones described herein) can provide at least approximately 70% sensitivity at approximately 99.5% specificity in subjects with hereditary and/or life-style associated risk. Moreover, in certain embodiments, exemplary assays can differentiate between subjects with benign tumors and those with lung cancer, resulting in few false positives.

**[626]** In certain embodiments, additional validation studies of exemplary assays as described herein can be conducted utilizing longitudinally-collected samples from independent sources. In some embodiments, samples can be obtained or derived from longitudinally

collected blood draws (*e.g.,* blood draws collected at temporally distinct time points) from lung cancer cases. In addition, blood draws from age-matched controls can be analyzed. Using exemplary logistic regression models for a 99% specificity cutoff, assessments can be made of how many years prior to diagnosis (*e.g.,* however many years prior to lung cancer diagnosis blood draws are available from) exemplary assay are capable of detecting lung cancer while maintaining an annual specificity of 99%. Assay sensitivity can be calculated as a function of year prior to diagnosis while ensuring specificity is maintained in the control samples at 99%.

***References Cited***

Balaj, L., Lessard, R., Dai, L., Cho, Y.J., Pomeroy, S.L., Breakefield, X.O. and Skog, J., 2011. Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. *Nature communications*, *2*, p.180.

Bebelman, M.P., Smit, M.J., Pegtel, D.M., and Baglio, S.R., 2018. Biogenesis and function of extracellular vesicles in cancer. *Pharmacology & Therapeutics*, *188*, pp.1-11.

Geena X. Wu, Dan J. Raz, Laura Brown, and Virginia Sun, 2016. Psychological burden associated with lung cancer screening: a systematic review. *Clin Lung Cancer,* 2016 Sep; 17(5): 315-324.

Charles S. Dela Cruz, Lynn T. Tanoue, and Richard A. Matthay. 2011. Lung Cancer: Epidemiology, Etiology, and Prevention, Clin Chest Med 32 (2011) 605–644. doi:10.1016/j.ccm.2011.09.001

Darmanis, S., Nong, R.Y., Hammond, M., Gu, J., Alderborn, A., Vänelid, J., Siegbahn, A., Gustafsdottir, S., Ericsson, O., Landegren, U. and Kamali-Moghaddam, M., 2010. Sensitive plasma protein analysis by microparticle-based proximity ligation assays. *Molecular & cellular proteomics*, *9*(2), pp.327-335.

Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2016, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2016/, based on November 2018 SEER data submission, posted to the SEER web site, April 2019.

Im, H., Shao, H., Park, Y.I., Peterson, V.M., Castro, C.M., Weissleder, R. and Lee, H., 2014. Label-free detection and molecular profiling of exosomes with a nano-plasmonic sensor. *Nature biotechnology*, *32*(5), p.490.

Jeong, S., Park, J., Pathania, D., Castro, C.M., Weissleder, R. and Lee, H., 2016. Integrated magneto–electrochemical sensor for exosome analysis. *ACS nano*, *10*(2), pp.1802-1809.

Kinsinger LS, Anderson C, Kim J, *et al.* Implementation of Lung Cancer Screening in the Veterans Health Administration. JAMA Intern Med. 2017;177(3):399-406. doi:10.1001/jamainternmed.2016.9022

Martin M Oken, Willam G Hocking, Paul A Kvale, Gerald L Andriole, Saundra S Buys, Timothy R Church, E David Crawford, Mona N Fouad, Claudine Isaacs, Douglas J Reding, Joel L Weissfeld, Lance A Yokochi, Barbara O'Brien, Lawrence R Ragard, Joshua M Rathmell, Thomas L Riley, Patrick Wright, Neil Caparaso, Ping Hu, Grant Izmirlian, Paul F Pinsky, Philip C Prorok, Barnett S Kramer, Anthony B Miller, John K

Gohagan, Christine D Berg, PLCO Project Team., 2011. Screening by Chest Radiograph and Lung Cancer Mortality: The Prostate, Lung, Colorectal, and Ovarian (PLCO) Randomized Trial. *JAMA*, 2011 Nov 2; 306(17):1865-1873.

Raghu VK, Zhao W, Pu J, *et al.* 2019. Feasibility of lung cancer prediction from low-dose CT scan and smoking factors using causal models. *Thorax* 2019;74:643–649.

Shao, H., Im, H., Castro, C.M., Breakefield, X., Weissleder, R. and Lee, H., 2018. New technologies for analysis of extracellular vesicles. *Chemical reviews*, *118*(4), pp.1917-1950.

## EQUIVALENTS

**[627]** Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. It is to be understood that the invention encompasses all variations, combinations, and permutations in which one or more limitations, elements, clauses, descriptive terms, *etc*., from one or more of the listed claims is introduced into another claim dependent on the same base claim (or, as relevant, any other claim) unless otherwise indicated or unless it would be evident to one of ordinary skill in the art that a contradiction or inconsistency would arise. Further, it should also be understood that any embodiment or aspect of the invention can be explicitly excluded from the claims, regardless of whether the specific exclusion is recited in the specification. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the claims that follow.

## CLAIMS

What is claimed is:

1. A method comprising steps of:

(a) providing or obtaining a biological sample from a subject;

(b) detecting, in the biological sample, nanoparticles having a size within the range of about 30 nm to about 1000 nm and expressing a first target biomarker signature ("first target biomarker signature-expressing nanoparticles"), the first target biomarker signature comprising:

at least one nanoparticle-associated surface biomarker and

at least one target biomarker selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers, wherein:

the surface biomarkers are selected from ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof;

the intravesicular biomarkers are selected from AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and

the intravesicular RNA biomarkers are selected from RNA transcripts (e.g., mRNA transcripts) encoded by human genes as follows: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4,

HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof;

(c) comparing sample information indicative of level of the first target biomarker signature-expressing nanoparticles in the biological sample to reference information including a first reference threshold level;

(d) classifying the subject as having or being susceptible to lung cancer when the biological sample shows an elevated level of first target biomarker signature-expressing nanoparticles relative to a classification cutoff referencing the first reference threshold level.

2. The method of claim 1, wherein when the at least one target biomarker is selected form one or more of the surface biomarkers, the selected surface biomarker(s) and the at least one nanoparticle-associated surface biomarker are different.

3. The method of claim 1 or 2, wherein the steps of (b) and (c) are repeated for at least a second target biomarker signature, and wherein the classification cutoff references the first reference threshold level and at least a second reference threshold level corresponding to the at least a second target biomarker signature.

4. The method of any one of claims 1-3, wherein the nanoparticle-associated surface biomarker is or comprises ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen, Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-

6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, or combinations thereof.

5. The method of any one of claims 1-4, wherein the first and/or second target biomarker signature comprises at least one nanoparticle-associated surface biomarker and at least two target biomarkers selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers.

6. The method of any one of claims 1-5, wherein the at least two target biomarkers comprise one of the following combinations:

- at least two distinct surface biomarkers;
- at least two distinct intravesicular biomarkers;
- at least two distinct intravesicular RNA biomarkers;
- a surface biomarker and an intravesicular biomarker;
- a surface biomarker and an intravesicular RNA biomarker; and
- an intravesicular biomarker and an intravesicular RNA biomarker.

7. A method comprising steps of:

(a) providing or obtaining a biological sample from a subject;

(b) detecting, in the biological sample, nanoparticles having a size within the range of about 30 nm to about 1000 nm and expressing a first target biomarker signature ("first target biomarker signature-expressing nanoparticles"), the first target biomarker signature comprising:

at least one nanoparticle-associated surface biomarker and

at least one target biomarker, wherein the target biomarker is or comprises a surface biomarker, wherein the target biomarker signature comprises a biomarker combination as listed in Tables 4A or 4B;

(c) comparing sample information indicative of level of the first target biomarker signature-expressing nanoparticles in the biological sample to reference information including a first reference threshold level;

(d) classifying the subject as having or being susceptible to lung cancer when the biological sample shows an elevated level of first target biomarker signature-expressing nanoparticles relative to a classification cutoff referencing the first reference threshold level.

8. The method of any one of claims 1-6, wherein the nanoparticle-associated surface biomarker is or comprises sTn antigen and/or MUC1.

9. The method of any one of claims 1-8, wherein the first and/or second target biomarker signature comprises (i) at least one nanoparticle-associated surface biomarker, which is or comprises a sTn antigen or MUC1, and (ii) at least one target biomarker selected from CEACAM5, CEACAM6, MUC1, Sialyl Lewis X, Lewis Y, and combinations thereof , which may be a surface biomarker or an intravesicular RNA biomarker.

10. The method of any one of claims 1-8, wherein the first and/or second target biomarker signature comprises (i) at least one nanoparticle-associated surface biomarker, which is or comprises a sTn antigen or MUC1 and (ii) at least two target biomarkers selected from CEACAM5, CEACAM6, MUC1, Sialyl Lewis X, Lewis Y, and combinations thereof , each of which may be a surface biomarker or an intravesicular RNA biomarker.

11. The method of any one of claims 1-8, wherein the first and/or second target biomarker signature comprises (i) at least one nanoparticle-associated surface biomarker, which is or comprises a sTn antigen, and (ii) at least two target biomarkers selected from MUC1, CEACAM5, CEACAM6, and combinations thereof, each of which may be a surface biomarker or an intravesicular RNA biomarker, or (i) at least one nanoparticle-associated surface biomarker, which is or comprises MUC1, and (ii) at least two target biomarkers selected from CEACAM5, Sialyl Lewis X, Lewis Y, and combinations thereof.

12. The method of any one of claims 1-11, wherein the first or second reference threshold level is determined by levels of target biomarker signature-expressing nanoparticles observed in comparable samples from a population of non-cancer subjects.

13. The method of claim 11, wherein the population of non-cancer subjects comprises one or more of the following subject populations: healthy subjects, subjects diagnosed with benign tumors, and subjects with non-lung-related diseases, disorders, and/or conditions.

14. The method of any one of claims 1-13, wherein the biological sample has been subjected to size exclusion chromatography to isolate (*e.g.*, directly from the biological sample) nanoparticles having a size range of interest that includes nanoparticles.

15. The method of any one of claims 1-14, wherein the step of detecting comprises a capture assay.

16. The method of claim 15, wherein the capture assay involves contacting the biological sample with a capture agent comprising a target-capture moiety that binds to the at least one nanoparticle-associated surface biomarker.

17. The method of claim 15, wherein the capture agent is or comprises a solid substrate comprising the target-capture moiety conjugated thereto.

18. The method of claim 17, wherein the solid substrate comprises a magnetic bead.

19. The method of any one of claims 16-18, wherein the target-capture moiety is or comprises an antibody agent.

20. The method of any one of claims 1-19, wherein the step of detecting comprises a detection assay.

21. The method of any one of claims 1-19, wherein the step of detecting comprises a capture assay and a detection assay, the capture assay being performed prior to the detection assay.

22. The method of any one of claims 20-21, wherein when the first and/or second target biomarker signature comprises at least one intravesicular RNA biomarkers, the detection assay involves reverse transcription qPCR.

23. The method of any one of claims 20-22, wherein when the first and/or second target biomarker signature comprises at least one intravesicular biomarker, the target biomarker signature-expressing nanoparticles are processed involving fixation and/or permeabilization prior to the detection assay.

24. The method of any one of claims 20-23, wherein when the first and/or second target biomarker signature comprises at least one surface biomarker and/or intravesicular biomarker, the detection assay involves an immunoassay (including, *e.g.*, immuno-PCR, and/or proximity ligation assay).

25. The method of claim 24, wherein the detection assay involves a proximity ligation assay.

26. The method of claim 25, wherein the proximity ligation assay comprises the steps of:

(a) contacting the target biomarker signature-expressing nanoparticles that express the at least one nanoparticle-associated surface biomarker ("nanoparticle-associated surface biomarker-expressing nanoparticles") with a set of detection probes, each directed to a target biomarker of the target biomarker signature, which set comprises at least two detection probes, so that a combination comprising the nanoparticles and the set of detection probes is generated,

wherein the detection probes each comprise:

(i) a target binding moiety directed to the target biomarker of the target biomarker signature; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the detection probes are characterized in that they can hybridize to each other when the detection probes are bound to the same nanoparticle,

(b) maintaining the combination under conditions that permit binding of the set of detection probes to their respective targets on the nanoparticles such that the at least two

detection probes can bind to the same nanoparticle that express the target biomarker signature to form a double-stranded complex;

(c) contacting the double-stranded complex with a nucleic acid ligase to generate a ligated template; and

(d) detecting the ligated template, wherein presence of the ligated template is indicative of presence in the biological sample of the target biomarker signature-expressing nanoparticles; and

(e) optionally repeating steps a through d at least one additional time using an orthogonal target biomarker signature.

27. The method of claim 26, wherein the target binding moiety of the at least two detection probes is directed to the same target biomarker.

28. The method of claim 27, wherein the oligonucleotide domain of the at least two detection probes are different.

29. The method of any one of claims 19-28, wherein the target-capture moiety of the capture assay is or comprises at least one antibody agent directed to the at least one nanoparticle-associated surface biomarker.

30. The method of any one of claims 1-29, wherein the method is performed to screen for early-stage lung cancer, late-stage lung cancer, or recurrent lung cancer in the subject.

31. The method of any one of claims 1-30, wherein the subject has at least one or more of the following characteristics:

(i) an asymptomatic subject who is susceptible to lung cancer (*e.g.*, at an average population risk (*i.e.*, without hereditary risk) or with hereditary risk for lung cancer);

(ii) a subject with a family history of lung cancer (*e.g.*, a subject having one or more first-degree relatives with a history of lung cancer);

(iii) a subject who is or was a smoker;

(iv) a subject who has exposure to secondhand smoke, radon gas, asbestos, bituminous "smoky coal", and/or other carcinogens (*e.g.*, arsenic, chromium, nickel, ionizing radiation, polycyclic aromatic hydrocarbons, nitric oxide, high levels of particulate matter < 2.5μm);

(v) a subject aged 30 or over;

(vi) a subject with one or more non-specific symptoms of lung cancer, optionally wherein at least one of the non-specific symptoms is similar to one or more common respiratory symptoms such as coughing, hemoptysis, airway obstruction, and shortness of breath, associated with a non-cancer disease, disorder, or condition;

(vii) a subject recommended for chest imaging such as X-ray, CT scan, or low-dose CT scan;

(viii) a subject diagnosed with an imaging-confirmed lung mass;

(ix) a subject with a benign lung tumor;

(x) a subject who has been previously treated for lung cancer;

(xi) a subject determined to have COPD;

(xii) a subject determined to have pulmonary fibrosis;

(xiii) a subject with a history of chronic bronchitis, tuberculosis, and/or pneumonia;

(xiv) a subject determined to have HIV and/or AIDS;

(xv) a subject with high current or historical alcohol consumption;

(xvi) a subject with hereditary mutations in EGFR, cytochrome p450 enzymes, and/or DNA repair genes; and

(xvii) a subject exposed to radiation therapy and/or chemotherapy.

32. The method of any one of claims 1-31, wherein the method is used in combination with one or more of the following diagnostic assays:

(i) the subject's annual physical examination;

(ii) a chest imaging (*e.g.*, X-ray, CT scan, or low-dose CT scan);

(iii) sputum cytology;

(iv) a genetic assay to screen blood plasma for genetic mutations in circulating tumor DNA and/or protein biomarkers linked to cancer; and

(v) an assay involving immunofluorescent staining to identify cell phenotype and marker expression, followed by amplification and analysis by next-generation sequencing.

33. The method of any one of claims 1-32, wherein the lung cancer is small cell lung cancer or non-small cell lung cancer.

34. The method of any one of claims 1-33, wherein the lung cancer is non-small cell lung cancer.

35. The method of claim 34, wherein the non-small cell lung cancer is lung adenocarcinoma.

36. The method of any one of claims 1-35, wherein the method is performed to monitor a lung cancer patient for response to treatment of an anti-lung cancer therapy (*e.g.*, surgery, radiation therapy, chemotherapy, radiosurgery, targeted drug therapy, immunotherapy) and/or for detection of cancer recurrence/metastasis, and/or for differentiating a benign lung mass (e.g., nodule or lesion) from lung cancer.

37. The method of any one of claims 1-35 for detecting cancer, the method comprising steps of: detecting on surfaces of intact nanoparticles from a human blood sample co-localization of at least two biomarkers whose combined expression level has been determined to be associated with cancer; comparing the detected co-localization level with the determined level; and detecting cancer when the detected co-localization level is at or above the determined level.

38. The method of any one of claims 1-35 for detecting cancer, the method comprising steps of: contacting a sample comprising exosomes with a set of detection probes that specifically bind to surface biomarkers on the exosomes to detect cancer-associated exosomes in the sample with a specificity within a range of 95% to 100% and sensitivity within a range of 30% to 100%.

39. The method of any one of claims 1-35, comprising steps of: capturing exosomes from a biological sample with a capture agent that selectively interacts with a cancer-specific surface biomarker on the exosomes; and contacting the captured exosomes with at least one set of at least two detection probes that each selectively interacts with a surface biomarker on the exosomes; and detecting a product formed when the at least two detection probes of the set are in sufficiently close proximity, such detection indicating co-localization of the surface biomarkers.

40. The method of any one of claims 1-35, comprising steps of: contacting a sample comprising exosomes with a set of probes that specifically bind to surface biomarkers on the exosomes to detect cancer-associated exosomes in the sample, wherein: (i) each probe in the set comprises a target binding moiety directed to a surface biomarker on the exosomes; and (ii) the set comprises at least one capture probe and at least two detection probes, wherein each detection probe further comprises a detection moiety.

41. The method of any one of claims 1-35, comprising steps of: performing a proximity assay that detects a surface biomarker signature on exosomes from a human subject, the step of performing being performed a period of time after a performance of a prior assay to detect the surface biomarker signature on exosomes from the human subject; and comparing results of the performed assay with those of the prior assay.

42. The method of any one of claims 1-35, comprising steps of: contacting exosomes with at least two detection probes, wherein each detection probe comprises (i) a binding moiety; and (ii) an oligonucleotide entity, wherein the binding moiety is the same and the oligonucleotide entities complement one another.

43. The method of any one of claims 1-35, comprising detecting marker proximity on exosome surfaces, including an improvement that comprises contacting the exosomes with at least a pair of binding agents that each comprise a binding moiety and a proximity moiety, wherein the binding moieties are the same and the proximity moieties complement one another; and detecting an interaction between the proximity moieties.

44. A kit for detection of lung cancer comprising:

(a) a capture agent comprising a target-capture moiety directed to a nanoparticle-associated surface biomarker, wherein the nanoparticle has a size within the range of about 30 nm to about 1000 nm; and

(b) at least one set of detection probes, which set comprises at least two detection probes each directed to a target biomarker of a target biomarker signature for lung cancer, wherein the detection probes each comprise:

(i) a target binding moiety directed at the target biomarker of the target biomarker signature for lung cancer; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same nanoparticle;

wherein the target biomarker signature for lung cancer comprises:

at least one nanoparticle-associated surface biomarker and

at least one target biomarker selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers, wherein:

the surface biomarkers are selected from ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen, Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof;

the intravesicular biomarkers are selected from AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF,

NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and

the intravesicular RNA biomarkers are selected from RNA transcripts (e.g., mRNA transcripts) encoded by human genes as follows: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof.

45. The kit of claim 44, wherein when the at least one target biomarker is selected from one or more of the surface biomarkers, the selected surface biomarker(s) and the at least one nanoparticle-associated surface biomarker are different.

46. The kit of claim 44 or 45, wherein the nanoparticle-associated surface biomarker is or comprises ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen, Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, or combinations thereof.

47. A kit for detection of lung cancer comprising:

(a) a capture agent comprising a target-capture moiety directed to a nanoparticle-associated surface biomarker, wherein the nanoparticle has a size within the range of about 30 nm to about 1000 nm; and

(b) at least one set of detection probes, which set comprises at least two detection probes each directed to a target biomarker of a target biomarker signature for lung cancer, wherein the detection probes each comprise:

(i) a target binding moiety directed at the target biomarker of the target biomarker signature for lung cancer; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same nanoparticle;

wherein the target biomarker signature for lung cancer comprises:

at least one nanoparticle-associated surface biomarker and

at least one target biomarker, wherein the target biomarker is or comprises a surface biomarker, wherein the target biomarker signature comprises a biomarker combination as listed in Tables 4A or 4B.

48. The kit of any one of claims 44-46, wherein the nanoparticle-associated surface biomarker is or comprises sTn antigen and/or MUC1.

49. The kit of any one of claims 44-48, wherein the target binding moiety of the at least two detection probes is each directed to the same target biomarker of the target biomarker signature.

50. The kit of claim 49, wherein when the nanoparticle-associated surface biomarker is or comprises sTn antigen, the same target biomarker to which the detection probes are directed is or comprises CEACAM5, CEACAM6, or MUC1; or wherein when the nanoparticle-associated

surface biomarker is or comprises MUC1, the same target biomarker to which the detection probes are directed is or comprises CEACAM5, Sialyl Lewis X, or Lewis Y.

51. The kit of claim 50, wherein the oligonucleotide domain of the at least two detection probes are different.

52. The kit of any one of claims 44-48, wherein the target binding moiety of the at least two detection probes is each directed to a distinct target biomarker of the target biomarker signature.

53. The kit of claim 52, wherein the nanoparticle-associated surface biomarker is or comprises a sTn antigen and/or MUC1, and the at least two detection probes are directed to at least two target biomarkers selected from CEACAM5, CEACAM6, MUC1, Sialyl Lewis X, and Lewis Y, and combinations thereof.

54. The kit of claim 52, wherein the nanoparticle-associated surface biomarker is or comprises a sTn antigen, and the at least two detection probes are directed to at least two target biomarkers selected from MUC1, CEACAM5, CEACAM6, and combinations thereof; or wherein the nanoparticle-associated surface biomarker is or comprises MUC1, and the at least two detection probes are directed to the at least two target biomarkers selected from CEACAM5, Sialyl Lewis X, Lewis Y, and combinations thereof .

55. The kit of any one of claims 44-54, further comprising at least one additional regent (*e.g.*, a ligase, a fixation agent, and/or a permeabilization agent).

56. The kit of any one of claims 44-55, comprising at least two sets (including, *e.g.*, at least three sets) of detection probes, which each set comprises at least two detection probes each directed to a target biomarker of a distinct target biomarker signature for lung cancer.

57. The kit of any one of claims 44-46, comprising:

(a) a first capture agent comprising a target-capture moiety;

(b) a second capture agent comprising a target-capture moiety;

(c) at least two sets of detection probes, wherein the detection probes each comprise:

(i) a target binding moiety directed at a target surface biomarker; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same nanoparticle.

58. The kit of any one of claims 44-46, comprising:

(a) a first capture agent comprising a target-capture moiety;

(b) a second capture agent comprising a target-capture moiety;

(c) a third capture agent comprising a target-capture moiety;

(d) at least three sets of detection probes, wherein the detection probes each comprise:

(i) a target binding moiety directed at a target surface biomarker; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same nanoparticle.

59. A complex comprising:

(a) a nanoparticle having a size within the range of about 30 nm to about 1000 nm and expressing a target biomarker signature for lung cancer, wherein the target biomarker signature comprises:

at least one nanoparticle-associated surface biomarker and

at least one target biomarker selected from the group consisting of: surface biomarkers, intravesicular biomarkers, and intravesicular RNA biomarkers,

wherein:

the surface biomarkers are selected from ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen, Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof;

the intravesicular biomarkers are selected from AOC1, C12orf45, CRABP2, CST1, ETV4, FAM83A, FOXA2, HMGB3, LGALS3BP, MIF, NAPSA, PPP1R14D, S100A14, SBK1, SCGB3A2, SFTA2, SFTPA1, SFTPA2, SFTPB, SPINK1, TGFA, ZC3H11A, and combinations thereof; and

the intravesicular RNA biomarkers are selected from RNA transcripts (e.g., mRNA transcripts) encoded by human genes as follows: ABCC3, AOC1, ARSL, B3GNT3, C12orf45, CDCP1, CDH1, CDH3, CEACAM5, CEACAM6, CELSR1, CLDN18, CLDN3, CLDN4, CLDN7, CLIC6, CRABP2, CST1, DMBT1, DSG2, EPCAM, EPHX3, ETV4, EVA1A, FAM83A, FOLR1, FOXA2, GJB1, GJB2, GPC4, HMGB3, HS6ST2, KDELR3, KRTCAP3, LAMB3, LFNG, LGALS3BP, LSR, MANEAL, MIF, MSLN, MUC1, MUC21, NAPSA, PIGT, PODXL2, PPP1R14D, PRRG4, ROS1, S100A14, SBK1, SCGB3A2, SDC1, SERINC2, SEZ6L2, SFTA2, SFTPA1, SFTPA2, SFTPB, SLC34A2, SLC44A4, SLC6A14, SLC7A7, SMIM22, SMPDL3B, SPINK1, ST14, TGFA, TMC4, TMC5, TMEM45B, TMPRSS2, TMPRSS4, TSPAN1, TSPAN8, ZC3H11A, and combinations thereof;

wherein the nanoparticle is immobilized onto a solid substrate comprising a target-capture moiety directed to the nanoparticle-associated surface biomarker;

(b) a first detection probe and a second detection probe each bound to the nanoparticle, wherein each detection probe comprises:

(i) a target binding moiety directed to one of the target biomarker of the tumor target biomarker signature; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the first and second detection probes are hybridized to each other.

60. The complex of claim 59, wherein when the at least one target biomarker is selected from one or more of the surface biomarkers, the selected surface biomarker(s) and the at least one nanoparticle-associated surface biomarker are different;

61. The complex of claim 59 or 60, wherein the nanoparticle-associated surface biomarker is or comprises ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen, Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, or combinations thereof.

62. A complex comprising:

(a) a nanoparticle having a size within the range of about 30 nm to about 1000 nm and expressing a target biomarker signature for lung cancer, wherein the target biomarker signature comprises:

at least one nanoparticle-associated surface biomarker and

at least one target biomarker, wherein the target biomarker is or comprises a surface biomarker, wherein the target biomarker signature is or comprises a biomarker combination as listed in Tables 4A or 4B;

wherein the nanoparticle is immobilized onto a solid substrate comprising a target-capture moiety directed to the nanoparticle-associated surface biomarker;

(b) a first detection probe and a second detection probe each bound to the nanoparticle, wherein each detection probe comprises:

(i) a target binding moiety directed to one of the target biomarker of the tumor target biomarker signature; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the first and second detection probes are hybridized to each other.

63. The complex of any one of claims 59-61, wherein the nanoparticle-associated surface biomarker is or comprises a sTn antigen or MUC1.

64. The complex of any one of claims 59-62, wherein the target binding moiety of the at least two detection probes is each directed to the same target biomarker of the target biomarker signature.

65. The complex of claim 64, wherein when the nanoparticle-associated surface biomarker is or comprises sTn antigen, the same target biomarker to which the detection probes are directed is or comprises CEACAM5, CEACAM6, or MUC1; or wherein when the nanoparticle-associated surface biomarker is or comprises MUC1, the same target biomarker to which the detection probes are directed is or comprises CEACAM5, Sialyl Lewis X, or Lewis Y.

66. The complex of claim 64, wherein the oligonucleotide domain of the at least two detection probes are different.

67. The complex of any one of claims 59-62, wherein the target binding moiety of the at least two detection probes is each directed to a distinct target biomarker of the target biomarker signature.

68. The complex of claim 67, wherein the nanoparticle-associated surface biomarker is or comprises a sTn antigen or MUC1 and the at least two detection probes are directed to at least two target biomarkers selected from CEACAM5, CEACAM6, MUC1, Sialyl Lewis X, Lewis Y, and combinations thereof.

69. The complex of claim 67, wherein the nanoparticle-associated surface biomarker is or comprises a sTn antigen, and the at least two detection probes are directed to at least two target biomarkers selected from MUC1, CEACAM5, CEACAM6, and combinations thereof; or wherein the nanoparticle-associated surface biomarker is or comprises MUC1, and the at least two detection probes are directed to the at least two target biomarkers selected from CEACAM5, Sialyl Lewis X, Lewis Y, and combinations thereof.

70. The complex of any one of claims 59-69, wherein the solid substrate comprises a magnetic bead.

71. The complex of any one of claims 59-70, wherein the target-capture moiety is or comprises an antibody agent.

72. The complex of any one of claims 59-71, comprising: (a) an exosome having at least one target biomarker on its surface; and (b) a first detection probe and a second detection probe each bound to the exosome, wherein each of the first detection probe and the second detection probe comprises: (i) a target binding moiety directed to a target biomarker expressed by the exosome; and (ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the first and second detection probes are hybridized to each other.

73. The complex of any one of claims 59-71, comprising nanoparticles from a human blood sample bound to a set of at least two probes, each of which comprises a biomarker binding moiety and an oligonucleotide domain, wherein two or more bound probes are in proximity to one another so that their oligonucleotide domains hybridize to each other to form a ligatable hybrid.

74. The complex of any one of claims 59-71, comprising: (a) an exosome comprising a cancer-associated target biomarker signature; and (b) at least a first detection probe and a second detection probe each bound to the exosome, wherein each of the detection probes comprise: (i) a target binding moiety directed to the target biomarker signature; and (ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the detection probes are at least partially complementary.

75. A set of probes for use in a method, kit, or complex of any one of claims 1-71, wherein each set of probes comprises: (a) a biomarker binding moiety that specifically binds to a surface biomarker on nanoparticles from cancer cells, wherein the nanoparticles have a size within the range of about 30 nm to about 1000 nm; and (b) an oligonucleotide domain, wherein the oligonucleotide domains of probes within the set are arranged and constructed so that, when the probes are bound to their target biomarkers, their oligonucleotide domains hybridize to one another to form a ligatable hybrid only when the target biomarkers are in proximity to one another.

76. A method for differentiating benign lung mass (e.g., nodule and/or lesion) from lung cancer, wherein the method comprises:

(a) detecting, in a blood-derived sample from a subject determined to have a lung mass, on surfaces of nanoparticles having a size within the range of about 300 nm to about 1000 nm co-localization of at least one biomarker combination, which comprises at least one capture biomarker and at least one detection biomarker, where the at least one capture biomarker and the at least one detection biomarker are each independently selected from: ALCAM, BCAP31,

CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen, Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof;

(b) comparing the detected co-localization level with a reference level; and

(c) identifying the lung mass of the subject to be likely benign when the detected co-localization level is or comparable to the reference level; or identifying the lung mass to be cancerous when the detected co-localization level is above the reference level.

77. The method of claim 76, wherein the method for differentiating benign lung mass from lung cancer has a specificity within a range of 90% to 100% and sensitivity within a range of 65% to 95%.

78. A method for detection of early-stage lung cancer, wherein the method comprises:

(a) detecting, in a blood-derived sample from a subject, on surfaces of nanoparticles having a size within the range of about 300 nm to about 1000 nm co-localization of at least one biomarker combination, which comprises at least one capture biomarker and at least one detection biomarker, where the at least one capture biomarker and the at least one detection biomarker are each independently selected from:

ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen, Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2,

NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof;

(b) comparing the detected co-localization level with a reference level; and

(c) identifying the subject to be negative for lung cancer when the detected co-localization level is or comparable to the reference level; or identifying the subject as likely to have or be susceptible to lung cancer, when the detected co-localization level is above the reference level.

79. The method of claim 78, wherein the method for detection of early-stage lung cancer has a specificity within a range of 90% to 100% and sensitivity within a range of 80% to 95%.

80. The method of any one of claims 76-79, wherein the detecting comprises detecting on surfaces of the nanoparticles co-localization of the at least one biomarker combination, wherein the at least one biomarker combination is selected from one of the following:

(i) sTn antigen in combination with CEACAM5, CEACAM6, and/or MUC1; and

(ii) MUC1 in combination with CEACAM5, Sialyl Lewis X, and/or Lewis Y.

81. The method of any one of claims 76-79, wherein the detecting comprises detecting on surfaces of the nanoparticles co-localization of the at least one biomarker combination, wherein the at least one biomarker combination is selected from one of the following combinations:

(i) (sTn antigen, CEACAM5);

(ii) (sTn antigen, CEACAM5, MUC1);

(iii) (sTn antigen, MUC1);

(iv) (sTn antigen, CEACAM6);

(v) (sTn antigen, CEACAM5, CEACAM6);

(vi) (sTn antigen, MUC1, CEACAM6);

(vii) (MUC1, CEACAM5);

(viii) (MUC1, CEACAM5, Sialyl Lewis X antigen); and

(ix) (MUC1, CEACAM5, Lewis Y antigen).

82. The method of any one of claims 76-81, wherein the detecting comprises:

(a) capturing the nanoparticles from the blood-derived sample with a capture probe that selectively interacts with the at least one capture biomarker on the nanoparticles;

(b) contacting the captured nanoparticles with at least one set of at least two detection probes that each selectively interacts with the at least one detection biomarker on the nanoparticles; and

(c) detecting a product formed when the at least two detection probes of the set are in sufficiently close proximity on the individual nanoparticles.

83. The method of claim 82, wherein the capture probe comprises a target-capture moiety that binds to the capture biomarker.

84. The method of claim 83, wherein the target-capture moiety is or comprises an antibody agent directed to the capture biomarker.

85. The method of any one of claims 76-84, wherein the capture biomarker is or comprises a polypeptide encoded by human gene *MUC1*, or a sialyl Tn (sTn) antigen.

86. The method of any one of claims 82-85, wherein the capture probe is or comprises a solid substrate comprising the target-capture moiety conjugated thereto.

87. The method of claim 86, wherein the solid substrate comprises a magnetic bead.

88. The method of any one of claims 82-87, wherein the at least two detection probes each comprise:

(i) a target binding moiety directed to the at least one detection biomarker; and

(ii) an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain,

wherein the single-stranded overhang portions of the detection probes are characterized in that they can hybridize to each other when the detection probes are bound to the same nanoparticle.

89. The method of any one of claims 82-88, wherein the product was formed when the at least two detection probes of the set are in sufficiently close proximity on the individual nanoparticles such that the single-stranded overhang portions of the at least two detection probes of the set hybridize to each other to form a double-stranded complex.

90. The method of claim 89, wherein the product formed comprises a ligated template upon contacting the double-stranded complex with a nucleic acid ligase.

91. The method of any one of claims 82-90, wherein the target binding moieties of the at least two detection probes are each directed to the same detection biomarker.

92. The method of claim 91, wherein the oligonucleotide domain of the at least two detection probes are different.

93. The method of claim 91 or 92, wherein when the capture biomarker is or comprises a sTn antigen, the same detection biomarker to which the detection probes are directed is or comprises CEACAM5, CEACAM6, or MUC1; or wherein when the capture biomarker is or comprises MUC1, the same detection biomarker to which the detection probes are directed is or comprises CEACAM5, Sialyl Lewis X, or Lewis Y.

94. The method of claim 82-90, wherein the target binding moieties of the at least two detection probes are each directed to a distinct detection biomarker.

95. The method of claim 94, wherein when the capture biomarker is or comprises a sTn antigen and/or MUC1, the at least two detection probes are directed to at least two detection biomarkers selected from CEACAM5, CEACAM6, MUC1, Sialyl Lewis X, and Lewis Y, and combinations thereof.

96. The method of claim 94, wherein the capture biomarker is or comprises a sTn antigen, and the at least two detection probes are directed to at least two detection biomarkers selected

from MUC1, CEACAM5, CEACAM6, and combinations thereof; or wherein the capture biomarker is or comprises MUC1, and the at least two detection probes are directed to the at least two detection biomarkers selected from CEACAM5, Sialyl Lewis X, Lewis Y, and combinations thereof.

97. The method of any one of claims 76-96, wherein the nanoparticles are or comprise intact extracellular vesicles.

98. A kit comprising:

at least one set of probes for a biomarker combination specific for detection of lung cancer,

wherein the biomarker combination comprises at least one capture biomarker on nanoparticles having a size within the range of about 300 nm to about 1000 nm and at least one detection biomarker on the nanoparticles, and wherein the capture biomarker and the detection biomarker are each independently selected from:

ALCAM, BCAP31, CD109, CD274, CD55, CD9, CDH3, CEACAM5, CEACAM6, DSC2, DSC3, EGFR, EPCAM, FOLR1, HACD3, IL1RAP, IGF1R, IGSF3, ITGA2, LAMB3, Lewis Y antigen, Lewis X antigen, MARCKSL1, MET, MSLN, MUC1, MUC4, NT5E, Phosphatidylserine, PTGFRN, PTK7, SDC1, Sialyl Lewis X antigen, SLC34A2, Sialyl Tn antigen (sTn antigen), Tn antigen, T antigen, TACSTD2, TNFRSF10B, TFRC, TRPV4, TSPAN8, CD274, CDH1, CNTN1, GOLM1, VWA1, CELSR2, CLDN1, DSG2, DSG3, GPC1, LAMC2, NECTIN1, ST14, CLCA2, PTPRZ1, SLC2A1, LYPD3, Gb3 (CD77), GJB2, MPZL2, RAP2B, ITGB4, DSP, MEST, ITGA6, ABCC5, ATP1B3, JAG1, GPNMB, PERP, TMPRSS11D, Sialyl-6T antigen (6-sialyl core 1), Sialyl-T antigen, Sialyl Lewis A antigen (CA19-9), Globo H, Gb5 (SSEA-3), Lactotriaosylceramide (Lc3), Forssman antigen, and combinations thereof; and

wherein the at least one set of probes comprises:

a capture probe comprising a target-capture moiety directed to the capture biomarker; and

at least two detection probes each comprising a target binding moiety directed to the at least one detection biomarker.

99. The kit of claim 98, comprising a plurality of sets of probes, each set for a distinct biomarker combination specific for detection of lung cancer.

100. The kit of claim 98 or 99, wherein the biomarker combination(s) is/are selected from one of the following:

(i) sTn antigen in combination with CEACAM5, CEACAM6, and/or MUC1; and

(ii) MUC1 in combination with CEACAM5, Sialyl Lewis X, and/or Lewis Y.

101. The kit of claim 98 or 99, wherein the biomarker combination(s) is/are selected from one of the following:

(i) (sTn antigen, CEACAM5);

(ii) (sTn antigen, CEACAM5, MUC1);

(iii) (sTn antigen, MUC1);

(iv) (sTn antigen, CEACAM6);

(v) (sTn antigen, CEACAM5, CEACAM6);

(vi) (sTn antigen, MUC1, CEACAM6);

(vii) (MUC1, CEACAM5);

(viii) (MUC1, CEACAM5, Sialyl Lewis X antigen); and

(ix) (MUC1, CEACAM5, Lewis Y antigen).

102. The kit of any one of claims 98-101, wherein the capture probe and detection probes selectively bind to respective biomarkers on the nanoparticles with a specificity within a range of 90% to 100% and sensitivity within a range of 65% to 95%.

103. The kit of any one of claims 98-102, wherein the detection probes each further comprises an oligonucleotide domain coupled to the target binding moiety, the oligonucleotide domain comprising a double-stranded portion and a single-stranded overhang portion extended from one end of the oligonucleotide domain, wherein the single-stranded overhang portions of the at least two detection probes are characterized in that they can hybridize to each other when the at least two detection probes are bound to the same nanoparticle.

104. The kit of any one of claims 98-103, wherein the target binding moieties of the at least two detection probes are each directed to the same detection biomarker on the nanoparticles.

105. The kit of claim 104, wherein the oligonucleotide domain of the at least two detection probes are different.

106. The kit of claim 104 or 105, wherein when the capture biomarker is or comprises sTn antigen, the same detection biomarker to which the detection probes are directed is or comprises CEACAM5, CEACAM6, or MUC1; or wherein when the capture biomarker is or comprises MUC1, the same detection biomarker to which the detection probes are directed is or comprises CEACAM5, Sialyl Lewis X, or Lewis Y.

107. The kit of claim 98-103, wherein the target binding moieties of the at least two detection probes are each directed to a distinct detection biomarker on the nanoparticles.

108. The kit of claim 107, wherein the capture biomarker is or comprises a sTn antigen and/or MUC1, and the at least two detection probes are directed to at least two detection biomarkers selected from CEACAM5, CEACAM6, MUC1, Sialyl Lewis X, and Lewis Y, and combinations thereof.

109. The kit of claim 107, wherein the capture biomarker is or comprises a sTn antigen, and the at least two detection probes are directed to at least two detection biomarkers selected from MUC1, CEACAM5, CEACAM6, and combinations thereof; or wherein the capture biomarker is or comprises MUC1, and the at least two detection probes are directed to the at least two detection biomarkers selected from CEACAM5, Sialyl Lewis X, Lewis Y, and combinations thereof.

110. The kit of any one of claims 98-109, further comprising at least one additional reagent (*e.g.*, a ligase, a fixation agent, and/or a permeabilization agent).

111. The kit of any one of claims 98-110, wherein the nanoparticles are or comprise intact extracellular vesicles.