

Description

Title of Invention: COMPOSITIONS AND METHODS FOR NON-INVASIVE RAPID TEST FOR DNA, RNA, AND PROTEIN MARKERS PRESENT IN NASOPHARYNGEAL CARCINOMA

[0001] Cross-reference to Related Applications:

[0002] This application claims the benefit of and priorities to U.S. Provisional Application No. 63/597,510 filed November 09, 2023, U.S. Provisional Application No. 63/554,695 filed February 16, 2024 and U.S. Provisional Application No. 63/706,419 filed October 11, 2024, which are hereby incorporated by reference in their entirety.

Field of the Invention:

[0003] This invention is generally related to compositions methods and kits for diagnosing and relapse prediction of nasopharyngeal carcinoma.

Background of the Invention:

[0004] Nasopharyngeal carcinoma (NPC) is a malignancy highly associated with Epstein-Barr virus (EBV). The gold standard of NPC diagnosis (NPCD) is nasopharyngeal biopsy coupled with in-situ hybridization (ISH) for Epstein-Barr encoded ribonucleic acid (EBER) to confirm the presence of undifferentiated carcinoma or non-keratinizing carcinoma. This method also require nasoendoscopy sampling and is therefore, invasive. Although plasma EBV DNA has been regarded as relatively sensitive in NPC screening, diagnosis and treatment response evaluation, it can still miss up to 15% of NPC if used as the only screening or diagnosis tool, especially early-stage small NPC and those NPC which are defective in producing plasma EBV DNA. Positive predictive value (PPV) of plasma EBV DNA in NPCD is just around 11%. In addition, plasma EBV DNA is currently used for disease monitoring after treatment and still remains unable to accurately predict relapse before treatment at time of initial diagnosis. About 50% local relapse and up to 20% metastatic disease are negative for plasma EBV DNA. PPV of plasma EBV DNA in NPC relapse prediction (NPCR) is just around 6-30%. Currently, about 30% of patients still relapse within 5 years despite intensive treatment.

[0005] RNA and protein (expression) biomarkers for NPCD or NPCR has also been studied. The most commonly mutated genes in NPC are TP53 and PIK3CA. TP53 is a tumor suppressor gene that is involved in cell cycle regulation and apoptosis. However, only around 6% NPC patients and 15% relapse cases were found to carry TP53 mutations. In terms of relapse prediction, detectable EBV DNA during post-treatment follow-up is associated with tumor recurrence. However, a cut-off value of 0 copy/mL for EBV

DNA during post-treatment follow-up has a sensitivity of around 50-80% and a positive predictive value of around 6-30%. There is neither consensus nor gold standard of using any DNA mutations for non-invasive NPCD or NPCR.

- [0006] The viral oncoprotein, Epstein–Barr virus latent membrane protein 1 (LMP1) is up-regulated in NPC. However, only 63% of NPC cases were found to have LMP1 expression at nasopharyngeal tumor by immunohistochemistry.
- [0007] There is currently no consensus or gold standard in using RNA and protein biomarkers for non-invasive NPCD or NPCR before treatment.
- [0008] It is an object of the present invention to provide compositions, methods and kits for non-invasive nasopharyngeal carcinoma diagnosis (NPCD) in a subject.
- [0009] It is an object of the present invention to provide compositions, methods, kits for non-invasive nasopharyngeal carcinoma relapse prediction (NPCR) before treatment and suggested follow-ups in a subject.
- [0010] Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present disclosure as it existed before the priority date of each claim of this application.
- [0011] Throughout this specification the word “comprise,” or variations such as “comprises” or “comprising,” will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

Summary of the Invention:

- [0012] Disclosed are compositions, kits and methods of diagnosing NPC or predicting NPC relapse thereof, in a subject, with suggested follow-ups. The methods, compositions and kits are based on the discovery of genetic mutations, RNA and/or protein expression biomarkers, and the epithelial cell types reflected by DNA/RNA/protein biomarkers, the presence of which allow the diagnosis of a subject as having NPC or a determination of having NPC relapse in the future, with high, sensitivity, specificity, and accuracy. Thus, the methods detect the presence of indicators of NPCD or NPCR, which can be: (i) genetic mutations, (ii) biomarker expression (RNA and/or protein expression), and/or (iii) cytometric indicators (i.e. cell type) reflected by expression of cell type-specific DNA/RNA/protein biomarkers.
- [0013] The methods include use: (a) genetic marker detection, by determining the presence of one or more genetic mutations as disclosed herein, in a sample obtained from the subject, the presence of which means in some forms that the subject will have an NPC relapse in the future (i.e. NPCR) or in other forms, that the subject has NPC, producing

a diagnosis result (i.e., NPCD) , respectively; (b) biomarker detection, by measuring the levels of at least one biomarker, in a sample obtained from the subject, wherein the presence and/or increased levels of each of the measured biomarkers relative to a control without NPC relapse or without NPC means in some forms that the subject will have an NPC relapse in the future (i.e. NPCR) or in other forms, that the subject has NPC, producing a diagnosis result (i.e., NPCD) , respectively; and/or (c) cytometric marker detection, by detecting the presence of cell types (as disclosed herein) , in a sample obtained from the subject, the presence of which means in some forms that the subject will have an NPC relapse in the future (i.e. NPCR) or in other forms, that the subject has NPC, producing a diagnosis result (i.e., NPCD) , respectively.

- [0014] A) . Methods for detecting the presence of NPC relapse (i.e. “NPCR”) in a biological sample obtained from a subject, use genetic marker, biomarker and/or cytometric marker detection. The cytometric marker detection includes genetic mutation detection and/or biomarker expression in the cytometric markers, as discussed further below.
- [0015] Genetic markers for detecting NPCR include detecting genetic mutations in cells in a biological sample obtained from the subject. In these forms, the methods detect: single-nucleotide variations (SNVs) and/or Insertions/Deletions (InDels) which include, but are not limited to a SNV in (i) HMGN2P3 (high mobility group nucleosomal binding domain 2 pseudogene 3) , (ii) ARL5A (ADP Ribosylation Factor Like GTPase 5A) , DHX57 (DEXH-box helicase 57) , (iii) IL32 (interleukin 32) and /or (iv) ATPAF1 (ATP Synthase Mitochondrial F1 Complex Assembly Factor 1) ; and/or an InDel in (i) DNAJC11 (DnaJ Heat Shock Protein Family (Hsp40) Member C11) , (ii) EIF2AK1 (Eukaryotic Translation Initiation Factor 2 Alpha Kinase 1) , (iii) FAM234A (Family With Sequence Similarity 234 Member A)) ; PARPBP (PARP1 Binding Protein)) and/or (iv) HDAC2 (Histone deacetylase 2) . The presence of one or more of these mutations is indicative of a relapse. In some forms, the presence two or more of the identified mutations are assessed, for example, IL32 and DHX57.
- [0016] In some forms, the method detects a relapse with a specificity of at least 80%. In some forms, the presence of IL32 (interleukin 32) (SNV) and DHX57 (DEXH-box helicase 57) (SNV) indicates a relapse, with a 100% sensitivity;
- [0017] Cytometric markers for detecting NPCR include genetic mutation detection and/or biomarker expression in specific cell types as disclosed below, and in these forms, the methods detect
- [0018] (i) . cytometric (cell type) mutations such as: (1) HDAC2 (chr6: 113, 970, 875) [CT>C, CTT, CTTT, CTTTT, CTTTTT] InDel or ATPAF1 (such as ATPAF1 chr1: 46668177 [A>C]) mutant LY6D⁺ neoplastic secretory-primed basal cluster 1 cells (SPB1) . SPB1 is the major or majority form of secretory-primed basal cells (SPB) ,

which is just before Goblet or secretory states in epithelial cell development; (2) chr 1: 5857099 [CCAGC] in frame deletion at TACSRD2 present in LY6D+ neoplastic SPB1 cells present in the sample; and/or (3) chr 1: 87328973 (Ato T) 5'UTR single-nucleotide mutation at LMO4 present in LY6D+ neoplastic SPB1 cells present in the sample; or

[0019] (ii) . cytometric (cell type) biomarkers expression in a biological sample obtained from the subject, specifically, markers of the LY6D-positive (+/-ATPAF1-mutated) neoplastic SPB1/SPB subpopulations, to detect NPC relapse in the subject. Exemplary biomarkers (predictive of relapse) are cytometric (cell type) biomarkers, specifically, markers of the LY6D+ neoplastic SPB1/SPB epithelial subpopulations, and they include mRNA or protein, peptide, or fragment thereof, encoded by one or more of the following genes, the upregulation or downregulation of which detect relapse, as discussed herein: (A) LY6D; KRT16 (Keratin 16); CEBPD (CCAAT enhancer binding protein delta); CDKN1A (cyclin-dependent kinase inhibitor 1A); PGM2 (Phosphoglucomutase 2); MEG3 (Maternally Expressed 3); CTNNBIP1 (Catenin Beta Interacting Protein 1); IGF2BP3 (Insulin-like growth factor 2 mRNA-binding protein 3); CLDND1 (claudin domain containing 1); DUSP11 (Dual-specificity phosphatase 11); FAF1 (Fas Associated Factor 1); BAG4 (BAG Cochaperone 4); SI-PA1L2 (Signal Induced Proliferation Associated 1 Like 2); AP3M2 (Adaptor Related Protein Complex 3 Subunit Mu 2); SERPINB12 (Serpin Family B Member 12); CALML3 (Calmodulin Like 3); CLCA4 (Chloride Channel Accessory 4); GPX2 (glutathione peroxidase 2), and LSP1 (Lymphocyte Specific Protein 1) for detecting the presence of relapse of NPC in a subject; the presence/high expression of one or more of these biomarkers in a biological sample obtained from the subject is indicative of the presence of NPC relapse in the future; and (B) for non-neoplastic LYD+cells, CLCA4 (Chloride Channel Accessory 4); SYT8 (Synaptotagmin 8); FGFR3 (Fibroblast growth factor receptor 3); SUSD4 (Sushi Domain Containing 4); TNNT3 (Troponin T3, Fast Skeletal Type); NSG1 (Neuronal Vesicle Trafficking Associated 1) for detecting the presence of upcoming relapse of NPC in a subject; the absence/low level of one or more of these biomarkers in a biological sample obtained from the subject is indicative of the presence of NPC relapse in the future.

[0020] Biomarker detection methods for detecting NPCR include in some forms, detecting the expression, in a biological sample obtained from the subject, of one or more of the following genes: NEDD8; CALML3; NDUFA13; BEX3; HNRNPA0; SLIRP; ADH5; GNG5; UBE2D2; PSMA4; SLC2A1; SNX3; LY6D; PSMA3; PSMB1; YBX1; PSMB5; PSMA6; DHTKD1, PRMT9, SEH1L, MROH1, MED25, NCAN, GLS, DENND5B, COX8A, AUP1, GAN. The expression can be determined by measuring mRNA or protein, peptide, or fragment thereof, encoded by these genes. In some

forms, prediction of NPC relapse is detected by a combination of at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, or at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, or at least seventeen of the aforementioned genes. In some forms, mutations in eighteen of the aforementioned genes, for example, NEDD8; CALML3; NDUFA13; BEX3; HNRNPA0; SLIRP; ADH5; GNG5; UBE2D2; PSMA4; SLC2A1; SNX3; LY6D; PSMA3; PSMB1; YBX1; PSMB5; PSMA6, are used to detect NPC relapse. In some forms, detection of higher RNA or protein levels of relapse-predicting biomarkers selected from NEDD8, CALML3, NDUFA13, BEX3, HNRNPA0, SLIRP, ADH5, GNG5, UBE2D2, PSMA4, SLC2A1, SNX3, LY6D, PSMA3, PSMB1, YBX1, PSMB5, PSMA6, DHTKD1, PRMT9, SEH1L, MROH1, MED25, NCAN, GLS, DENND5B, COX8A, AUP1, GAN predict NPC relapse.

- [0021] In some forms, increased expression of GAN, CYB561D2, DLST, PRMT9, OGA, PFDN4, CNDP1, PPP6R1, DPP3, ESD, CDH1, DAG1, AUP1, PGAM5, DAD1, FCN1, FGL1, SRGN, PGA3, PGA4, PGA5, FMOD predicts local relapse.
- [0022] In some forms, increased expression of GAS2L1, HDAC1, MARCHF2, TRIM28, TRAPPC11, ADAMTS7, HLA-DRB3, PKM, CXCL6, PRMT9, MROH1, MED25, NCAN, AMDHD1, PCSK1, ARFGAP1, ZNHIT1, ZNF326, CBX1, LRP5, SYNPO2, HIVEP1, DEFA3, HLA-DPB1, DUSP7, STOML3, GLS, DENND5B, AUP1, SLC17A5, SLC2A2 predicts distant relapse.
- [0023] Increased expression of NID2, APP, XYLT2, GNPAT, MRPL17, CONDMKN, CTSA, TNFSF12, SRGN, PROZ, CALU, C1S, FGL1, SOD3 predicts neck relapse.
- [0024] In some forms, prediction of NPC relapse is determined when the level of RNA or protein for any one or more of the aforementioned genes is higher in a subject by about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 92%, about 94%, about 96%, about 98%, about 100%, or more, compared to a control subject or pre-determined control value. Decreased expression of HMGB2, SNRPF, SRP14, RANBP3L, SAMHD1, SART3, IFI16, CROCC, NCF1B, SLIT1, SH3BGR, IGHV3-53, ANTXR1, EPB42, and/or SLC4A1 predicts relapse.
- [0025] Decreased expression of ZNF428, HDGFL2, GALE, HNRNPL, LUC7L2, HNRNPA1, THYN1, ANK3, LAMTOR5, MFGE8, CD69, REEP5, CHST7, ATP6V0A1, RDX, SLIT1, HLA-A, SH3BGR, SYNPO2, CD2BP2, SET, SELENOH, CCL19, RHAG, CLEC3B, EZR, SLC4A1, SLC29A1, FRZB, and/or SEPTIN11 predicts local relapse.
- [0026] Decreased expression of GASK1A, ANKRD30A, SMIM15, RIPOR1, B4GALT5, ZNF846, PRF1, MINPP1, FBL, RALYL, EPB41, TMEM132D, NCF1B, JAG1,

CLEC14A, GALC, SNRPB2, SLFN5, CALU, SLC12A4, CROCC, DSCAML1, IGHV3-53, TTN, IGHV4-39, CYP2C9, HSD11B1, COL2A1, and/or STAU2 predicts distant relapse.

- [0027] Decreased expression of AP1B1, CNBP, EXOC3L4, LSM6, ABCB6, ANTXR2, CLEC3B, LTF, F11, C1QTNF3, ANTXR1, and/or PIP4K2A predicts neck relapse.
- [0028] In some forms, the presence of EBV DNA, RNA or protein of RPMS1, LMP-1, or LMP-2B can detect NPC. High level EBV DNA, RNA or protein of RPMS1, BALF4, BALF5, or BALF0 can indicate NPC relapse prediction.
- [0029] B) . Methods for detecting the presence of indicators of NPC ("NPCD") , in a biological sample obtained from a subject, use genetic markers and/or biomarker detection. Thus, methods for detecting the presence of indicators of NPC (i.e. "NPCD") in a biological sample obtained from a subject, include:
- [0030] Genetic markers for detecting NPCD includes detecting one or more genetic mutations in cells in a biological sample obtained from the subject. In these forms the methods detect SNVs and/or InDels, which include but are not limited to a SNV in EEF2KMT; SOCS1; TESMIN; and IGFBP7s and/or InDels in EMP2; IL32; , and/or CSTA; , optionally in combination with EBV biomarkers. In these forms the disclosed methods diagnose NPC in a subject with at least about 80%sensitivity, preferably, at least 85, 90 or up to 95%sensitivity.
- [0031] Biomarkers for detecting NPCD includes detecting expression (and thus, increased/decreased expression) , in a biological sample obtained from the subject The expression of a gene can be determined by measuring mRNA or protein, peptide, or fragment thereof, encoded by the gene.
- [0032] In some forms, increased expression of one or more of the following genes: CKAP4, SYNGR2, MARCKSL1, CFL1, PDCD5, IL32, RAN, NME1, VCAM1, GAPDH, TUBB, HSPD1, LGALS1, TNFAIP3, ITGAV, RPLP0, STMN1, PPIA, CCL20, FSCN1, LGALS9, RPS19, YBX1, TPI1, ICAM1, ENO1, C1QBP, CXCL3, MIF, TAGLN2, CXCL10, UBD, CSTA, NFKBIA, SOCS1, MYBPC1, NTRK2, FKBP1A, SOX4, TUBA1C, PYCARD, TCIM and RPMS1 is used for detecting the presence of NPC in a subject . The presence/increased levels of one or more of these biomarkers in a biological sample obtained from the subject is indicative of presence of NPC. In some forms the disclosed methods diagnose NPC in a subject with at least about 80%sensitivity, preferably, at least 85, and up to 95%sensitivity. In some forms, co-detection of increased levels of expression of CCL20 (CCL20high) + IL32 (IL32high) + low levels of expression of LCN2 (LCN2low) attain 100%in sensitivity.
- [0033] In some forms, decreased expression any one or more of SLPI, WFDC2, AQP3, LCN2, BPIFB1, SERPINB3, CLU, SCGB1A1, CRIP1, KRT14, LYPD2, TFF3, TSPAN1, SCGB3A1, PIGR, MUC16, C19orf33, PRSS23, MSMB, CAPS, CXCL17,

ANXA1, AGR2, GSTA1, BPIFA1, MT1X, ALDH3A1, C20orf85, LGALS3, MUC4, and TACSTD2, is used to detect NPC.

- [0034] In some forms, a low RNA and/or protein level of any one or more of SLPI, WFDC2, AQP3, LCN2, BPIFB1, SERPINB3, CLU, SCGB1A1, CRIP1, KRT14, LYPD2, TFF3, TSPAN1, SCGB3A1, PIGR, MUC16, C19orf33, PRSS23, MSMB, CAPS, CXCL17, ANXA1, AGR2, GSTA1, BPIFA1, MT1X, ALDH3A1, C20orf85, LGALS3, MUC4, and TACSTD2, is used to detect NPC. Lower expression of ERAP2, COL3A1, RAVR1, EPHA1, MANEAL, MAP2K2, SIGIRR, BRD4, HEMGN, BDH2, AKAP12, TM4SF1, LCMT1, CLEC14A, IGHV3-13, TBC1D20, TRAPPC2, TRAPPC2B, WDR41, OXA1L, TTN, DST, TMEM132A, CKM, KMT2B, JMY, IGHV3-21, DNAH6, COCH, NSL1, SPTB, WWC3, KRT9, IGKV2-29, IGKV3D-20, KIFC3 detects de novo Metastasis of NPC.
- [0035] In some forms, NPC is detected when the level of RNA or protein for any one or more of the aforementioned genes is lower in a subject by about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 92%, about 94%, about 96%, about 98%, about 100%, compared to a control subject or pre-determined control value.
- [0036] Nasopharyngeal sample collection swabs, sample collection tubes and NPC assay kits are also provided. The kits can include a lateral flow device and/or a nasopharyngeal swab and a sample collection tube. In some forms the lateral flow assay is a form of probe-based assay in which the test sample flows along a solid substrate via capillary action with DNA/RNA/protein biomarker targets detected by probes including but not limited to antibodies or aptamers. The lateral flow device includes a solid substrate, such as a membrane strip, having an application point, an optional conjugate zone, a capture zone, and an absorbent. Binding agents are present in the conjugate zone, to bind one or more of the biomarkers markers disclosed herein. Capture agents are immobilized in the capture zone, which preferably contains a plurality of capture lines for detecting captured analyte (capture complex) .
- [0037] The binding agent in the capture zone can be an antibody or biomarker binding fragment thereof, or an aptamer. In some forms, the binding agent in the capture zone include antibodies that bind to a biomarker selected from the group consisting of PSMA4, CALML3, SLC2A1, SNX3, LY6D, YBX1, and RPMS1 and the device can be used to detect NPCR. In some forms, the binding agent in the capture zone include antibodies that bind to a biomarker selected from the group consisting of CKAP4, SYNGR2, CFL1, and RPMS1 and the device can be used for NPCD. In some forms, the device includes multiple capture zones for simultaneous detection of more than one biomarker and/or simultaneous detection of NPCR and NPCD.

[0038] Also described are computer-implemented systems (CISs) and/or methods (CIMs) containing one or more discriminative artificial intelligence (AI) platforms for analysing biological data using a signature matrix and outputting the occurrence of NPCR/NPCD based on the expression levels of certain biomarkers or combinations of biomarkers in the biological data, which are indicative of certain cell types. Columns in the signature matrix track cell types and rows track genes, and the number at each i th-row and j th-column represents the relative level of gene expression of a specific gene in the cell type. The AI platform assesses a subject's test results and makes a prediction about the occurrence of NPCR/NPCD based on comparisons with levels of gene expression levels or combinations thereof in the signature matrix. In clinical settings, if the gene expression data in the subject's test results contain more neoplastic SPB1 features when compared to the signature matrix, i.e., the gene expression is similar to the signature matrix, it indicates more or presence of neoplastic SPB1 in the subject's test sample, and/or if the gene expression data contain less non-neoplastic SPB1, it is likely the subject would have NPCR. The assessment is performed by uploading the signature matrix and subject's test results to the AI platform that analyzes the data and makes a prediction. The prediction is provided via a visual format (e.g., graphical user interface), an audio-format (e.g., via an audio signal that reports the prediction), or a combination thereof. The one or more AI platforms have been trained and validated using data involving gene expression levels of these biomarkers, associated cell types, and/or the occurrence of NPCR/NPCD.

Brief Description of the Drawings:

[0039] FIGs. 1A-1B are the schematic diagrams showing steps in discovering novel biomarkers from a single-cell RNA sequencing (scRNA-seq) NPC study largest in scale to date for disease diagnosis and relapse prediction using our proposed tests. 74 patients were recruited for the scRNA-seq study. Matched biopsies were harvested from normal adjacent tumor, primary tumor at nasopharynx and tumor from neck lymph nodes. Patients were followed for 3 years in median. 5 patients relapsed. Bioinformatic analyses were conducted to identify human-based DNA mutations, RNA/protein expression, and cytometric biomarkers that were associated with NPC malignancy and NPC relapse at the time of initial diagnosis (FIG. 1A). Microbe EBV RNA transcripts were also detected and analyzed. Biomarkers with highest neoplasticity, or NPCD, and/or NPCR accuracy were identified (FIG. 1B). FIG. 1C is a schematic diagram showing exemplary sampling methods and detection strategies utilizing the disclosed (1) human-based DNA, RNA/protein biomarker panel, (2) EBV-based panel and (3) cytometric panel for both (1) non-invasive rapid test with biological samples from nasopharyngeal swab or plasma, or (2) conventional

Rule 91,
06.01.2025

invasive endoscopy biopsy. FIG. 1D is a schematic diagram showing how the epithelial subpopulation (i.e. cytometric) biomarker combo can be applied to AI-powered platform to analyze expression data (e.g. transcriptomics data including RNA sequencing and microarray, proteomics data including mass spectrometry) to calculate the two risk scores (i.e. risk score 1 = neoplastic SPB1 abundance, and risk score 2 = difference between neoplastic SPB1 and non-neoplastic SPB1) to predict relapse at initial diagnosis before treatment and provide information for follow-ups early. FIGs. 1E-1G are diagrams showing a summary of biomarkers discovered for NPCD and/or NPCR.

- [0040] FIGs. 2A and 2B show mutations (i.e. SNVs and InDels) uniquely found in neoplastic cells from all possible malignant sites that predict relapse at initial diagnosis before treatment (i.e. NPCR) . FIG. 2A shows top 8 somatic mutations uniquely identified in neoplastic cells that were found strongly associated with relapse even at time of initial diagnosis. FIG. 2B shows performance metrics of the top 8 relapse-predicting somatic mutations.
- [0041] FIGs. 3A-3C shows novel mutations (i.e. SNVs and InDels) with high occurrence in NPC patients at initial diagnosis. FIG. 3A shows top 3 somatic mutations of neoplastic cells identified in NPC patients. EMP2 mutation was also found to have high coverage in EBV-negative NPC patients. FIG. 3B shows performance metrics of EMP2 mutation alone in NPC diagnosis. FIG. 3C shows that the disclosed mutation combo panel alone could achieve NPC diagnosis with sensitivity of 93.0%, which is significantly higher than golden standard plasma EBV DNA copy number alone. If mutation combo was used together with EBV plasma EBV DNA copy number, sensitivity could be further boosted up to 95.3%.
- [0042] FIGs. 4A-4E shows Lymphocyte Antigen 6 Family Member D (LY6D) -positive neoplastic secretory-primed basal (SPB) cells, in particular SPB1 (secretory primed basal cell type cluster 1, referred to hereinafter as SPB1) as an unique and novel neoplastic subpopulation that is consistently associated with and predicts relapse with high accuracy. FIG. 4A is a UMAP showing a total of 30 different epithelial subpopulations discovered by unsupervised clustering from the scRNA-seq data with known epithelial canonical markers. FIG. 4B is a violin plot showing that SPB1 is a novel epithelial subpopulation originating from SPB discovered by our scRNA-seq analysis. LY6D was found to be the biomarkers of SPB and in particular SPB1. Neoplastic SPB1 was found to have the highest LY6D expression while the remaining non-SPB1 subpopulations were found to have much lower LY6D expression. FIG. 4C is a table of statistical analysis showing difference in abundance of some neoplastic subpopulations between relapse and no relapse, without consideration of relapse time, establishing that neoplastic SPB1, SPB2 and SPB5 are the top neoplastic subpopulation that are

consistently associated with relapse, as revealed by Wilcoxon rank sum test. FIG. 4D shows log rank test that supported neoplastic SPB1, SPB2 and SPB5 as relapse-related. In particular, Cox regression analysis showed neoplastic SPB1 as the only neoplastic subpopulation that can predict relapse. KM plot also showed presence of SPB1 is associated with lower relapse-free survival probability. FIG. 4E is a performance analysis showed AUC of relapse prediction by neoplastic SPB1 to be 0.773 with an accuracy reaching 81%.

- [0043] FIG. 5. shows the performance metrics of identifying relapse by detecting ATPAF1 mutation carried in neoplastic SPB1 and different epithelial subpopulations in nasopharyngeal or neck lymph node tumor in our scRNA-seq data.
- [0044] FIG. 6 shows ROC with AUCs of different pathohistological scorings of LY6D immunohistochemistry results in predicting relapse (i.e. NPCR) .
- [0045] FIG. 7A shows the table of performance metrics of top significant relapse-predicting expression biomarkers within extracellular space or cell surface categories in Cox regression analysis, Wilcoxon test and logistic regression analysis. FIG. 7B shows the table of performance metrics of top significant relapse-predicting expression biomarkers not necessarily within extracellular space or cell surface categories in Cox regression analysis, Wilcoxon test and logistic regression analysis.
- [0046] FIGs. 8A-8C shows newly identified diagnostic expression biomarkers with high occurrence in NPC patients at initial diagnosis. FIG. 8A shows performance metrics of diagnostic expression biomarkers within extracellular space or cell surface categories significant in Wilcoxon test and logistic regression with a median difference of at least 0.5 between positive and negative results. FIG. 8B shows diagnostic expression biomarkers not necessarily within extracellular space or cell surface categories but significant in Wilcoxon test, with expression higher in NPC than non-NPC. FIG. 8C shows diagnostic expression biomarkers not necessarily within extracellular space or cell surface categories but significant in Wilcoxon test, with expression higher in non-NPC than NPC. Expression biomarker combo of chemokine (C-C motif) ligand 20 (CCL20) , IL32 and lipocalin-2 (LCN2) was found to achieve 100%in sensitivity and 96.3%in PPV in NPC diagnosis (i.e. NPCD) .
- [0047] FIG. 9A shows that EBV biomarker detection at nasopharyngeal tumor is more sensitive than plasma EBV DNA. EBV transcripts detected in an scRNA-seq reached sensitivity of 100%in identifying NPC patients. Even patients with plasma EBV DNA copy number = 0 copies/ml could be identified. FIG. 9B shows EBV expression biomarkers, RPMS1 in particular, can predict NPC relapse (i.e. NPCR) .
- [0048] FIGs. 10A-10D show performance of artificial intelligence (AI) -powered SPB1-guided relapse prediction/diagnosis before treatment using an independent RNA sequencing dataset as demonstration. FIG. 10A is a table showing AI-powered sub-

population abundance estimation in an independent RNA-seq dataset. Wilcoxon rank sum test was used to identify significant difference between relapse and non-relapse group. FIGs. 10B and 10C show identification of AI-powered non-neoplastic SPB1 abundance estimation for false positive discovery and the optimized AI-powered SPB1 relapse prediction score calculated by difference between neoplastic and non-neoplastic SPB1 (i.e. risk score 2) could reach sensitivity of 87.5% and accuracy of 86.1%. FIG. 10D is a schematic diagram demonstrating the use of optimized AI-powered SPB1-guided relapse prediction scores for NPC relapse diagnosis before treatment to maximize treatment beneficial outcomes.

[0049] FIGs. 11A-11I illustrate non-invasive rapid test sampling using next-generation 3D-printed nasopharyngeal swab, sample processing tube and customized panels of probes at rapid antigen test (RAT) for NPC diagnosis and relapse prediction conducted at home, community clinic, hospital or medical laboratory. FIG. 11A shows a 3D-printed nasopharyngeal swab tailor-made for high yield and least discomforting cellular and interstitial sample collection. FIG. 11B is a magnified side view of the swab tip design with dimensions. FIG. 11C is a magnified top view of the swab tip design with dimensions. FIGs. 11D-G shows designs of next-generation rapid diagnostic test station/hub including sample collection/distribution structure which maximize detection of biomarkers with consistency and minimizes sample input for detection of more than one biomarker panel. Slices of detection strips can be customized for personalized disease and prognostic detection. For example, inflammation (inflam) and drug resistance (DR) can be simultaneously detected in one strip. FIG. 11D shows a view of all RAT cassettes inserted to sampling hub. FIG. 11E shows another view showing the assembly or structure of RAT cassettes and before and after insertion to sampling hub. FIG. 11F shows the top view (cut in the middle) of a RAT cassette inserted to sampling hub. FIG. 11G shows the side view (cut in the middle) of a RAT cassette inserted to sampling hub. FIG. 11H is a next-generation sample processing tube including brushes/bristles therein for more efficient release of biomaterials from swab head. FIG. 11I. Sample processing tube with short hairs inside specially designed for releasing materials from (nasopharyngeal) swab.

[0050] FIGs. 12A-12B are illustrations of lateral flow devices made from a membrane strip having an application point at the proximal end, followed by a conjugation zone, a capture zone, and an absorbent zone. The arrow shows the direction of lateral flow from the proximal to distal end. A plurality of capture lines is shown in the capture zone.

[0051] FIG. 13 is a flow chart of computer-implemented systems (CISs) and/or methods (CIMs) containing one or more discriminative artificial intelligence (AI) platforms for analyzing biological data using a signature matrix and outputting the occurrence of

NPCR/NPCD based on the expression levels of certain biomarkers or combinations of biomarkers in the biological data, which are indicative of certain cell types.

[0052] FIG. 14A includes line plots showing association of different epithelial cell types with relapse. At normal tissue adjacent to tumor, non-neoplastic ciliated cells tended to be more abundant in no relapse than that in relapse, while non-neoplastic multipotent basal cells (MPB) tended to be more abundant in relapse than that in no relapse. At nasopharyngeal primary tumor, non-neoplastic ciliated cells, and non-neoplastic cells undergoing transition to secretory goblet cells tended to be more abundant in no relapse, while non-neoplastic MPB and SPB cells tended to be more abundant in relapse. As more neoplastic cells, SPB and proliferating basal cells (PB) tended to be more abundant in relapse. At lymph node, activated basal cells (AB) tended to be much more abundant in no relapse while basal cells (other than AB, SPB, MPB, PB) tended to be more abundant in relapse. FIG. 14B shows the statistical analysis of abundance of different major epithelial cell states between relapse and no relapse at different tissue types. FIG. 14C shows the statistical analysis of abundance of different major epithelial subtypes between relapse and no relapse at different tissue types.

[0053] FIG. 15 shows the table about discovery of significant upregulation of neoplastic MPB1 subpopulation at lymph node in relapse when compared with no relapse.

Detailed Description:

[0054] It is to be understood that the disclosed methods and compositions are not limited to specific synthetic methods, specific analytical techniques, or to particular reagents unless otherwise specified, and, as such, can vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

[0055] Methods for detecting the occurrence of nasopharyngeal carcinoma (NPC) and/or predicting NPC relapse in a subject has been developed. The disclosed methods are based on the discovery of genetic mutations (SNV and inDel) and expression biomarkers as well as cytometric markers, which have greater accuracy, specificity, positive predictive value and sensitivity for diagnosing NPC and predicting NPC relapse compared to existing expression biomarkers such as plasma EBV DNA. The methods detect an analyte in a sample obtained from a subject.

[0056] Expression levels of biomarkers can be assessed by measuring RNA or protein encoded by the genes discussed in detail below. Thus, high/low levels of expression of a gene can be determined by measuring the levels of RNA or protein encoded by that gene.

- [0057] Expression levels of biomarkers (as disclosed herein) , are compared to a control subject or pre-determined control value, to determine “increased” or “decreased” expression or the biomarker.
- [0058] The analyte to be detected can be any one or more of RNA, DNA, protein, peptide, or a fragment thereof. For example, the analyte to be detected can be RNA. In some forms, the analyte to be detected is DNA. In some forms, the analyte to be detected is a protein, peptide, or a fragment thereof. In some forms, a combination of two or more analytes are detected. For example, the analytes to be detected can be a combination of a nucleic acid and a peptide or protein. For example, the combination of the analytes to be detected can be a protein biomarker e.g., NPC relapse-predicting protein biomarker, and a nucleic acid e.g., an NPC-associated single nucleotide variation (SNV) . In another example, the analytes to be detected can be two nucleic acids e.g., one nucleic acid can be a NPC-associated SNV and a second nucleic acid can be an InDel mutation.
- [0059] One rationale of using nucleic acid (e.g. DNA, RNA) and protein biomarker panels at nasopharynx or blood plasma for non-invasive NPC diagnosis and relapse prediction Is at least because nucleic acid such as DNA/RNA or protein with the disclosed features could be released by neoplastic cells from different sites into the nearby nasopharyngeal tissue or into blood plasma/lymph. Probes that detect or amplify these biomarkers could be used at nasopharyngeal secretion or blood plasma.
- [0060] These embodiments are described further in detail below. The disclosed methods can employ a combination of detection of mutations, RNA expression and analyte expression to detect NPC or NPC relapse prediction.
- [0061] I. DEFINITIONS
- [0062] The term “assay” refers to an in vitro procedure for analyzing a sample to determine the presence, absence, or quantity of one or more analytes of interest.
- [0063] The terms “control” and “calibration” as used in connection with analytes, are used interchangeably to refer to analytes used as internal standards.
- [0064] The term “analyte” refers to a chemical substance of interest that is a potential constituent of a biological sample and is to be analyzed by an assay.
- [0065] A “lateral flow” assay is a device intended to detect the presence (or absence) of a target analyte in sample in which the test sample flows along a solid substrate via capillary action.
- [0066] The term “membrane” as used herein refers to a solid substrate with sufficient porosity to allow movement of antibodies or aptamers bound to analyte by capillary action along its surface and through its interior.
- [0067] The term “membrane strip” or “test strip” refers to a length and width of membrane sufficient to allow separation and detection of analyte.

- [0068] The term “application point” is the position on the membrane where a fluid can be applied.
- [0069] The term “immobilized” refers to chemical or physical fixation of an agent or particle to a location on or in a substrate, such as a membrane. For example, capture agents may be chemically conjugated to a membrane, and particles coated with capture agents may be physically trapped within a membrane.
- [0070] The term “capture particle” refers to a particle coated with a plurality of capture agents. In preferred embodiments, the capture particle is immobilized in a defined capture zone.
- [0071] The term “capture zone” refers to a point on a membrane strip at which one or more capture agents are immobilized.
- [0072] The term “antibody” refers to intact immunoglobulin molecules, fragments or polymers of immunoglobulin molecules, single chain immunoglobulin molecules, human or humanized versions of immunoglobulin molecules, and recombinant immunoglobulin molecules, as long as they are chosen for their ability to bind an analyte.
- [0073] The term “aptamer” refers to an oligonucleic acid or peptide molecule that binds to a specific target molecule. Aptamers are generally selected from a random sequence pool. The selected aptamers are capable of adapting unique tertiary structures and recognizing target molecules with high affinity and specificity.
- [0074] A “nucleic acid aptamer” is an oligonucleic acid that binds to a target molecule via its conformation. A nucleic acid aptamer may be constituted by DNA, RNA, or a combination thereof. Nucleic acid aptamers are typically engineered using SELEX (systematic evolution of ligands by exponential enrichment) .
- [0075] A “peptide aptamer” is a combinatorial peptide molecule with a randomized amino acid sequence that is selected for its ability to bind a target molecule. Peptide aptamers are typically selected from combinatorial peptide libraries using yeast two-hybrid or phage display assays.
- [0076] The term “biological sample” refers to a tissue (e.g., tissue biopsy) , organ, cell, cell lysate, or body fluid from a subject. Non-limiting examples of body fluids include blood, urine, plasma, serum, tears, lymph, bile, cerebrospinal fluid, interstitial fluid, aqueous or vitreous humor, colostrum, sputum, amniotic fluid, saliva, anal and vaginal secretions, perspiration, semen, transudate, exudate, and synovial fluid.
- [0077] A “sample collection apparatus,” as used herein, refers to an apparatus that can be used for collection of a biological sample or into which a collected biological sample can be deposited or stored.
- [0078] The term “metatype” refers to the analyte-binding site of a binding agent when bound to analyte. The term “idiotype” refers to the analyte binding site of a binding agent free of its analyte.

- [0079] The term “anti-metatype” refers to a binding agent that selectively recognizes a binding agent-analyte complex (metatype) but lacks specificity for either free analyte or free binding agent. The term “anti-idiotypic” refers to a binding agent that selectively recognizes the analyte binding site of another binding agent.
- [0080] The term “specifically binds” or “selectively binds” refers to a binding reaction which is determinative of the presence of the analyte in a heterogeneous population. Generally, a first molecule that “specifically binds” a second molecule has an affinity constant (K_a) greater than about 10^5 M^{-1} (e.g., 10^6 M^{-1} , 10^7 M^{-1} , 10^8 M^{-1} , 10^9 M^{-1} , 10^{10} M^{-1} , 10^{11} M^{-1} , and 10^{12} M^{-1} or more) with that second molecule.
- [0081] The term “detectable label” refers to any moiety that can be selectively detected in a screening assay. Examples include radiolabels, (e.g., ^3H , ^{14}C , ^{35}S , ^{125}I , ^{131}I), affinity tags (e.g., biotin /avidin or streptavidin), binding sites for antibodies, metal binding domains, epitope tags, fluorescent or luminescent moieties (e.g., fluorescein and derivatives, green fluorescent protein (GFP), rhodamine and derivatives, lanthanides), colorimetric probe, and enzymatic moieties (e.g., horseradish peroxidase, β -galactosidase, β -lactamase, luciferase, alkaline phosphatase).
- [0082] As used herein, the term “sensitivity” refers to the ability of a test to correctly identify true positives, i.e., patients with NPC (or patients predicted to have relapse). For example, sensitivity can be expressed as a percentage, the proportion of actual positives which are correctly identified as such (e.g., the percentage of test subjects having BC correctly identified by the test as having NPC). A test with high sensitivity has a low rate of false negatives, i.e., the cases of NPC not identified as such.
- [0083] As used herein, the term “specificity” refers to the ability of a test to correctly identify true negatives, i.e., the individuals that have no BC. For example, specificity can be expressed as a percentage, the proportion of actual negatives which are correctly identified as such (e.g., the percentage of test subjects not having NPC correctly identified by the test as not having NPC). A test with high specificity has a low rate of false positives, i.e., the cases of individuals not having BC but suggested by the test as having NPC.
- [0084] Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein.
- [0085] Use of the term “about” is intended to describe values either above or below the stated value in a range of approx. $\pm 10\%$; in other embodiments the values may range in value either above or below the stated value in a range of approx. $\pm 5\%$; in other embodiments the values may range in value either above or below the stated value in

a range of approx. $\pm 2\%$; in other embodiments the values may range in value either above or below the stated value in a range of approx. $\pm 1\%$. The preceding ranges are intended to be made clear by context, and no further limitation is implied.

[0086] All methods described herein can be performed in any suitable order unless otherwise indicated or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the embodiments and does not pose a limitation on the scope of the embodiments unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

[0087] Disclosed are materials, compositions, and components that can be used for, can be used in conjunction with, can be used in preparation for, or are products of the disclosed method and compositions. These and other materials are disclosed herein, and it is understood that when combinations, subsets, interactions, groups, etc. of these materials are disclosed that while specific reference of each various individual and collective combinations and permutation of these compounds may not be explicitly disclosed, each is specifically contemplated and described herein. For example, if a ligand is disclosed and discussed and a number of modifications that can be made to a number of molecules including the ligand are discussed, each and every combination and permutation of ligand and the modifications that are possible are specifically contemplated unless specifically indicated to the contrary. Thus, if a class of molecules A, B, and C are disclosed as well as a class of molecules D, E, and F and an example of a combination molecule, A-D is disclosed, then even if each is not individually recited, each is individually and collectively contemplated. Thus, in this example, each of the combinations A-E, A-F, B-D, B-E, B-F, C-D, C-E, and C-F are specifically contemplated and should be considered disclosed from disclosure of A, B, and C; D, E, and F; and the example combination A-D. Likewise, any subset or combination of these is also specifically contemplated and disclosed. Thus, for example, the subgroup of A-E, B-F, and C-E are specifically contemplated and should be considered disclosed from disclosure of A, B, and C; D, E, and F; and the example combination A-D. Further, each of the materials, compositions, components, etc. contemplated and disclosed as above can also be specifically and independently included or excluded from any group, subgroup, list, set, etc. of such materials.

[0088] These concepts apply to all aspects of this application including, but not limited to, steps in methods of making and using the disclosed compositions. Thus, if there are a variety of additional steps that can be performed it is understood that each of these additional steps can be performed with any specific embodiment or combination of em-

bodiments of the disclosed methods, and that each such combination is specifically contemplated and should be considered disclosed.

[0089] All methods described herein can be performed in any suitable order unless otherwise indicated or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the embodiments and does not pose a limitation on the scope of the embodiments unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

[0090] II. NPC DIAGNOSIS AND RELAPSE DETECTION

[0091] The disclosed methods detect the presence of one or more analytes, in a biological sample obtained from a subject, which can be DNA, RNA, protein, or a peptide fragment thereof. The disclosed methods detect the presence/absence of specific genetic mutations as disclosed herein, the expression of specific biomarkers (RNA/protein expression) , the presence/absence of specific cell types, including the presence of specific mutations in these specific cell types and the expression of specific biomarkers in these specific cell types. Acronyms used in the present disclosure are known in the art. Tables 2 and 3 provide commercially available sources for antibodies that are specific for disclosed biomarkers.

[0092] In the disclosed methods, a biological sample is assessed for the presence, absence, or most preferably, the quantity of an analyte. In preferred forms, the biological sample includes plasma, cells and fluid from the nasopharynx, the upper part of the pharynx, connecting with the nasal cavity above the soft palate.

[0093] In another preferred form, the biological sample includes cells and tissue e.g., cells and/or tissue from the nasopharynx and lymph nodes. In a particular form, the biological sample includes nasopharynx tumor cells such as tumor cells extracted from a nasopharynx tumor. In other forms, the biological sample is a bodily fluid, such as whole blood, plasma, serum, saliva, or oral fluid.

[0094] A. Detecting NPC Relapse

[0095] The disclosed methods for detecting the presence of NPC relapse (i.e. “NPCR”) in a biological sample obtained from a subject, use genetic markers (detection of genetic mutations) , biomarkers (detection of biomarker expression) and/or cytometric (cell-type specific) analyte detection. The cytometric analyte detection includes cell-type specific genetic mutation detection and/or cell-type specific biomarker expression, as discussed further below.

[0096] Detection of Mutation: In some forms NPC relapse can be determined by detecting, in a sample obtained from a subject, the presence of one or more of the following mutations as shown in the Table below:

Item	Gene	Chromosomal location	Mutation
Mut#1	IL32	chr16:3,065,801	[T>C] SNV mutation
Mut#2	DHX57	chr2:38,868,300	[T>A] SNV mutation
Mut#3	HMG2P3	chr16:26,032,755	[G>A] SNV mutation
Mut#4	DNAJC11	chr1:6,667,742	[TTC>T,TTCTC] InDel mutation
Mut#5	EIF2AK1	chr7:6,046,108	[A>AT] InDel mutation
Mut#6	FAM234A	chr16:254,566	[G>GT] InDel mutation
Mut#7	PARBP	chr12:102,123,938	[G>A,GA] InDel mutation
Mut#8	ARL5A	chr2:151,828,247	[A>G] SNV mutation
Mut#9*	Mut#1+ Mut#2	/	/
Mut#10	ATPAF1	chr1:46,668,177	[A>C] SNV mutation OR p.Val49Gly
Mut#11	HDAC2	chr6:113,970,875	[CT>C,CTT,CTTT,CTTTT,CTTTTT] InDel mutation
Mut#12	TACSRD2	chr1:58,577,099	[CCAG>C] InDel mutation OR p.Leu19del
Mut#13	LMO4	1:87328973	[A>T] SNV mutation

[0097] FIG. 1E (NPC relapse-predicting SNVs -InDels DNA and RNA mutation biomarkers). In some forms, NPC relapse is determined by a combination of mutations in at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, or at least eleven of the aforementioned genes. In some forms, mutations in all twelve of the aforementioned genes is used to detect NPC relapse. In a preferred embodiment, the method detects IL32 (chr16: 3, 065, 801) [T>C] SNV mutation and DHX57 (chr2: 38, 868, 300) [T>A] SNV mutation.

[0098] Detection of Biomarker Expression: In some forms NPC relapse can be determined by detecting, in a sample obtained from the subject, expression of one or more genes selected from the group consisting of: NEDD8 (Neural precursor cell expressed developmentally down-regulated protein 8), CALML3 (Calmodulin-like protein 3), NDUFA13 (NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13), BEX3 (Brain-expressed X-linked protein 3), HNRNPA0 (Heterogeneous nuclear ribonucleoprotein A0), SLIRP (SRA stem-loop-interacting RNA-binding protein), ADH5 (Alcohol dehydrogenase 5), GNG5 (Guanine nucleotide-binding protein G (I) / G (S) / G (O) subunit gamma-5), UBE2D2 (E3 ubiquitin-protein ligase E3D), PSMA4 (Proteasome subunit alpha type-4), SLC2A1 (Solute carrier family 2, facilitated glucose transporter member 1), SNX3 (Sorting nexin-3), LY6D (Lymphocyte antigen 6D), PSMA3 (Proteasome subunit alpha type-3), PSMB1 (Proteasome subunit beta type-1), YBX1 (Y-box-binding protein 1), PSMB5 (Proteasome subunit beta type-5), PSMA6 (Proteasome subunit alpha type-6), DHTKD1, PRMT9, SEH1L, MROH1, MED25, NCAN, GLS, DENND5B, COX8A, AUP1, and GAN (FIG. 1G, (NPC relapse-predicting RNA and protein expression biomarkers). The expression of each gene can be detected by measuring in a biological sample obtained from a subject,

Rule 91,
06.01.2025

RNA or protein, peptide, or fragment thereof, encoded by one or more of the gene. Increased expression of one or more of these genes in a subject, compared to a control subject or pre-determined control value, indicates NPCR. Increased expression of GAN, CYB561D2, DLST, PRMT9, OGA, PFDN4, CNDP1, PPP6R1, DPP3, ESD, CDH1, DAG1, AUP1, PGAM5, DAD1, FCN1, FGL1, SRGN, PGA3, PGA4, PGA5, FMOD predicts local relapse. Increased expression of GAS2L1, HDAC1, MARCHF2, TRIM28, TRAPPC11, ADAMTS7, HLA-DRB3, PKM, CXCL6, PRMT9, MROH1, MED25, NCAN, AMDHD1, PCSK1, ARFGAP1, ZNHIT1, ZNF326, CBX1, LRP5, SYNPO2, HIVEP1, DEFA3, HLA-DPB1, DUSP7, STOML3, GLS, DENND5B, AUP1, SLC17A5, SLC2A2 predicts distant relapse. Increased expression of NID2, APP, XYLT2, GNPAT, MRPL17, CONDMKN, CTSA, TNFSF12, SRGN, PROZ, CALU, C1S, FGL1, SOD3 predicts neck relapse. In some forms, prediction of NPC relapse is determined when the level of RNA or protein for any one or more of the aforementioned genes is higher in a subject by about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 92%, about 94%, about 96%, about 98%, about 100%, or more, compared to a control subject or pre-determined control value.

[0099] Decreased expression of HMGB2, SNRPF, SRP14, RANBP3L, SAMHD1, SART3, IFI16, CROCC, NCF1B, SLIT1, SH3BGR, IGHV3-53, ANT XR1, EPB42, SLC4A1 predicts relapse. Decreased expression of ZNF428, HDGFL2, GALE, HNRNPL, LUC7L2, HNRNPA1, THYN1, ANK3, LAMTOR5, MFGE8, CD69, REEP5, CHST7, ATP6V0A1, RDX, SLIT1, HLA-A, SH3BGR, SYNPO2, CD2BP2, SET, SELENOH, CCL19, RHAG, CLEC3B, EZR, SLC4A1, SLC29A1, FRZB, SEPTIN11 predicts local relapse. Decreased expression of GASK1A, ANKRD30A, SMIM15, RIPOR1, B4GALT5, ZNF846, PRF1, MINPP1, FBL, RALYL, EPB41, TMEM132D, NCF1B, JAG1, CLEC14A, GALC, SNRPB2, SLFN5, CALU, SLC12A4, CROCC, DSCAML1, IGHV3-53, TTN, IGHV4-39, CYP2C9, HSD11B1, COL2A1, STAU2 predicts distant relapse. Decreased expression of AP1B1, CNBP, EXOC3L4, LSM6, ABCB6, ANT XR2, CLEC3B, LTF, F11, C1QTNF3, ANT XR1, PIP4K2A predicts neck relapse.

[0100] In some forms, prediction of NPC relapse is detected by a combination of at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, or at least eleven, at least twelve, at least thirteen, at least fourteen, at least fifteen, at least sixteen, or at least seventeen of the aforementioned genes. In some forms, mutations in eighteen of the aforementioned genes (NEDD8; CALML3; NDUFA13; BEX3; HNRNPA0; SLIRP; ADH5; GNG5; UBE2D2;

PSMA4; SLC2A1; SNX3; LY6D; PSMA3; PSMB1; YBX1; PSMB5; PSMA6) are used to detect NPC relapse.

[0101] Detection of Cytometric analyte: In some forms NPC relapse can be determined by detecting, in a sample obtained from the subject, (i) LY6D-positive (LY6D+) neoplastic cells, and/or (ii) specific mutations in LY6D+ neoplastic cells and/or (iii) expression of specific biomarkers by the LY6D+ neoplastic cells and/or (iv) expression of specific biomarkers by the LY6D+ non-neoplastic cells (FIG. 1F (NPC relapse-predicting cytometric biomarkers (identifiable by cell type-specific and protein biomarkers)). In each instance increased/decreased expressed is determined by comparing expression of the same biomarker epithelial cells.

Rule 91,
06.01.2025

[0102] LY6D+ neoplastic cells: In some forms, NPC relapse is indicated by the presence of LY6D+ neoplastic cells reflected by the presence of, or higher expression of LY6D. Increased expression of LY6D is indicative of NPCR. In some preferred forms, NPC relapse prediction using LY6D IHC as an example, an Allred score ≥ 6 indicates NPC relapse. The modified Allred scoring system is a well-known, successfully clinically validated scoring system as shown in Table 1. (described in Arihilo, et al., Am J Clin Pathol 2007; 127 (3) : 356-365) .

[0103] Table 1. Modified Allred Score

Proportion score	Intensity score (membrane and/or cytoplasmic staining)	Modified Allred score
0 no cells positively stained with IHC	0 $\rightarrow$ negative	0-1 $\rightarrow$ negative
1 $\leq 1\%$ of cells positively stained with IHC	1 $\rightarrow$ Weak	2-3 $\rightarrow$ weak
2 1-10% of cells positively stained with IHC	2 $\rightarrow$ intermediate	4-6 $\rightarrow$ moderate
3 11-33% of cells positively stained with IHC	3 $\rightarrow$ strong	7-8 $\rightarrow$ strong
4 34-66% of cells positively stained with IHC		
5 67-100% of cells positively stained with IHC		

[0104] Specific mutations in the LY6D+ cells: In some forms, NPC relapse is indicated by the presence, in a sample obtained from a subject, of HDAC2 (chr6: 113, 970, 875) [CT>C, CTT, CTTT, CTTTT, CTTTTT] InDel and/or or ATPAF1 (chr1: 46668177)

[A>C]) mutant LY6D+ neoplastic secretory-primed basal cluster 1 cells (SPB1) . In some forms, NPC relapse is indicated by the presence, in a sample obtained from a subject, in frame deletion at TACSRD2 (chr 1: 5857099) in LY6D+ neoplastic SPB1 cells present in the sample; and/or 5'UTR single-nucleotide mutation (Ato T) at LMO4 (chr 1: 87328973) present in LY6D+neoplastic SPB1 cells present in the sample.

- [0105] The presence of one or more mutations is indicative of a relapse, and the absence of the mutations disclosed herein is indicative of no relapse. Preferably, the method detects a relapse with a specificity of at least 80%. In a preferred embodiment, the presence of IL32 (interleukin 32) (SNV) and DHX57 (DExH-box helicase 57) (SNV) indicates a relapse with a 100% sensitivity.
- [0106] Expression of specific biomarkers by the LY6D+ neoplastic cells: In some forms, NPC relapse detection is indicated by the presence of LY6D+ neoplastic SPB cells and/or increased expression of KRT16, CEBPD, CDKN1A, PGM2, and LY6D compared to the expression of these biomarkers in epithelial cells.
- [0107] In some forms, NPC relapse prediction is indicated by the presence of LY6D+neoplastic SPB1 cells reflected by presence and/or increased expression of MEG3, CTNNBIP1, and LY6D compared to the expression of these biomarkers epithelial cells.
- [0108] In some forms, NPC relapse prediction is indicated by the presence of LY6D+neoplastic SPB2 cells reflected by the presence of, and/or increased expression of IGF2BP3, FAF1, DUSP11, CLDND1, and LY6D compared to the expression of these biomarkers in epithelial cells.
- [0109] In some forms, NPC relapse prediction is indicated by the presence of LY6D+neoplastic SPB5 cells reflected by the presence of and/or increased expression of BAG4, SERPINB12, AP3M2, SIPA1L2, and LY6D compared to the expression of these biomarkers in epithelial cells.
- [0110] In some forms, NPC relapse prediction is indicated by the presence of “high neoplasticity” neoplastic SPB cells, reflected by expression of CALML3, CLCA4, GPX2, LSP1, and LY6D.
- [0111] In some forms, NPC relapse prediction is indicated by detecting the presence of LY6D+ non-neoplastic SPB cells expressing specific combinations of biomarkers in a sample obtained from a subject. In these forms, low levels of (i) LY6D+ non-neoplastic SPB cells reflected by expression of CLCA4, SYT8, FGFR3, and LY6D and/or (ii) LY6D+ non-neoplastic SPB1 cells reflected by expression of SUS4, TNNT3, NSG1, and LY6D, indicate NPCR.
- [0112] The disclosed methods contemplate combinations of detection of cytometric analytes. For example, in some forms, NPC relapse prediction is indicated by (i) the presence of HDAC-mutated neoplastic SPB1 cells reflected by the presence of

neoplastic SPB1 cells with HDAC2 InDel mutation, (ii) ATPAF1-mutated neoplastic SPB1 cells reflected by the presence of neoplastic SPB1 cells with ATPAF1 SNV mutation, (iii) low levels of non-neoplastic SPB cells reflected by expression of CLCA4, SYT8, FGFR3, and LY6D, (iv) low levels of non-neoplastic SPB1 cells reflected by RNA/protein expression of SUSD4, TNNT3, NSG1, and LY6D, and (v) high neoplasticity of neoplastic SPB cells reflected by expression of CALML3, CLCA4, GPX2, LSP1, and LY6D.

[0113] Table A: Summary of additional Biomarker Types

Biomarker Type	Diagnosis/ Relapse	Indication	List of Biomarkers
HDAC-mutated neoplastic SPB1 cells	Relapse	"Presence" indicates "NPCR"	MEG3, CTNNBIP1, LY6D, HDAC mutation
ATPAF1-mutated neoplastic SPB1 cells	Relapse	"Presence" indicates "NPCR"	MEG3, CTNNBIP1, LY6D, ATPAF1 mutation
EBV (DNA/RNA/protein)	Relapse	"High level" indicates "NPCR"	RPMS1, BALF4, BALF5, BALF0
EBV (DNA/RNA/protein)	Diagnosis	"Presence" indicates "NPCD"	RPMS1, LMP-1, LMP-2B

[0114] B. Detecting the presence of NPC

[0115] Detection of Mutation: In some forms the presence of NPC (i.e., NPCD) can be determined by detecting, in a sample obtained from a subject, the presence of one or more of the following mutations as shown in the Table below:

Item	Gene	Chromosomal location	Mutation
Mut#1	EMP2	chr16:10,543,633	[C>CA] InDel mutation
Mut#2	IL32	chr16:3,065,653	[AAGGTGACT...CGCCAGC>A,AAGC] InDel mutation
Mut#3	EEF2KMT	chr16:5,091,852	[G>A] SNV mutation
Mut#4	CSTA	chr3:122,337,565	[G>GA] InDel mutation
Mut#5	SOCS1	chr16:11,255,404	[A>C] SNV mutation
Mut#6	TESMIN	chr11:68,750,666	[A>T] SNV mutation
Mut#7	IGFBP7	chr4:57,110,068	[T>C] SNV mutation

[0116] In some forms, a combination of mutations in at least two, at least three, at least four, at least five, at least six of the aforementioned genes is used to detect NPC. In some forms, mutations in all seven (EMP2 + IL32 + EEF2KMT + CSTA + SOCS1 + TESMIN + IGFBP7) of the aforementioned genes are used to detect NPC.

[0117] The presence of the indicated mutation in a biological sample obtained from a subject is indicative of NPC, and the absence of the mutation (s) is indicative of the absence of NPC. The disclosed methods diagnose NPC (using identification of the mutations described above) in a subject with at least about 80%sensitivity, preferably, at least 85, 90 or up to 95%sensitivity.

- [0118] Detection of Biomarker expression: In some forms the presence of NPC (i.e., NPCD) can be determined by detecting, in a sample obtained from a subject, the expression of one or more of the following genes, where in some forms, high levels of expression is indicative of NPC and in some forms, low levels of expression is indicative of NPC, as disclosed further below: In some forms, high expression of any one or more of CKAP4 (Cytoskeleton-associated protein 4) , SYNGR2 (Synaptogyrin 2) , MARCKSL1 (Macrophage myristoylated alanine-rich C kinase substrate) , CFL1 (Protein CURLY FLAG LEAF 1) , PDCD5 (Programmed cell death protein 5) , IL32, RAN (Ras-related nuclear protein) , NME1 (N-methyl D-aspartate receptor subtype 2A) , VCAM1 (Vascular cell adhesion protein 1) , GAPDH (Glyceraldehyde-3-phosphate dehydrogenase) , TUBB (Tubulin Beta Class I) , HSPD1 (60 kDa heat shock protein) , LGALS1 (Galectin-1) , TNFAIP3 (TNFAIP3-interacting protein 2) , ITGAV (Integrin Subunit Alpha V) , RPLP0 (Large ribosomal subunit protein uL10) , STMN1 (Stathmin 1) , PPIA (Peptidyl-prolyl cis-trans isomerase A) , CCL20, FSCN1 (Fascin 1) , LGALS9 (Galectin-9) , RPS19 (Small ribosomal subunit protein eS19) , YBX1 (Y-box binding protein 1) , TPI1 (Triosephosphate isomerase) , ICAM1 (Inter-cellular adhesion molecule 1) , ENO1 (Enolase 1) , C1QBP (Complement component 1 Q subcomponent-binding protein) , CXCL3 (C-X-C motif chemokine 3) , MIF (Macrophage migration inhibitory factor) , TAGLN2 (Transgelin-1) , CXCL10 (C-X-C motif chemokine 10) , UBD (Ubiquitin D) , CSTA (Peptide transporter CstA) , NFKBIA (NF-kappa-B inhibitor alpha) , SOCS1 (Suppressor of cytokine signaling 1) , MYBPC1 (Myosin-binding protein C) , NTRK2 (B DNF/NT-3 growth factors receptor) , FKBP1A (Peptidyl-prolyl cis-trans isomerase FKBP1A) , SOX4 (Transcription factor SOX-4) , TUBA1C (Tubulin alpha-1C chain) , PYCARD (Apoptosis-associated speck-like protein containing a CARD) , and TCIM (Transcriptional and immune response regulator) . Higher expression of PBX2, CDK13, NPEPL1, DUSP19, BPNT1, LRRC75A, USP47, MAGT1, TARDBP, NDUFA12, ATOX1, MOSPD2, HBM, VTI1A, NFE2, MTHFS, IGFBP1, GCG, PRAP1, SYNPO2, FNIP2, NCAN, CRP, WWC3, GAN, DPY19L1, DHCR7, CGNL1, SLC25A1, OMD, DPT, HLA-DRB3, NTS, TRIM41, CCL17 indicates De Novo Metastasis.
- [0119] NPC is detected when the level of RNA or protein expressed by any one or more of the aforementioned genes is higher in a sample obtained from a subject by about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 92%, about 94%, about 96%, about 98%, about 100%, or more, compared to a control subject or pre-determined control value.
- [0120] In some forms, low expression of any one or more of SLPI (WD repeat containing protein slp1) , WFDC2 (WAP four-disulfide core domain protein 2) , AQP3

(Aquaporin-3) , LCN2 (Lipocalin 2) , BPIFB1 (BPI Fold Containing Family B Member 1) , SERPINB3 (Serp Family B Member 3) , CLU (Clusterin) , SCGB1A1 (Secretoglobin Family 1A Member 1) , CRIP1 (Cysteine-Rich Protein 1) , KRT14 (Keratin 14) , LYPD2 (Lymphocyte Antigen 6 Family Member D2) , TFF3 (Trefol Factor 3) , TSPAN1 (Tetraspanin 1) , SCGB3A1 (Secretoglobin Family 3A Member 1) , PIGR (Polymeric Immunoglobulin Receptor) , MUC16 (Mucin 16) , C19orf33 (Chromosome 19 Open Reading Frame 33) , PRSS23 (Protease, Serine 23) , MSMB (Microseminoprotein Beta) , CAPS (Calcyphosine) , CXCL17 (C-X-C Motif Chemokine Ligand 17) , ANXA1 (Annexin A1) , AGR2 (Anterior Gradient 2) , GSTA1 (Glutathione S-Transferase Alpha 1) , BPIFA1 (BPI Fold Containing Family A Member 1) , MT1X (Metallothionein 1X) , ALDH3A1 (Aldehyde Dehydrogenase 3 Family Member A1) , C20orf85 (Chromosome 20 Open Reading Frame 85) , LGALS3 (Galectin 3) , MUC4 (Mucin 4) , and TACSTD2 (Tumor-Associated Calcium Signal Transducer 2) can detect NPC. Lower expression of ERAP2, COL3A1, RAVR1, EPHA1, MANEAL, MAP2K2, SIGIRR, BRD4, HEMGN, BDH2, AKAP12, TM4SF1, LCMT1, CLEC14A, IGHV3-13, TBC1D20, TRAPPC2; TRAPPC2B, WDR41, OX-A1L, TTN, DST, TMEM132A, CKM, KMT2B, JMY, IGHV3-21, DNAH6, COCH, NSL1, SPTB, WWC3, KRT9, IGKV2-29, IGKV3D-20, KIFC3 detects De Novo Metastasis of NPC.

- [0121] NPC is detected when the level of RNA or protein expressed by any one or more of the aforementioned genes is lower in a sample obtained from a subject by about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 92%, about 94%, about 96%, about 98%, about 100%, compared to a control subject or pre-determined control value.
- [0122] III, Detection of Biomarkers
- [0123] FIG. 1 is a schematic diagram showing proposed exemplary sampling methods and detection strategies utilizing our the disclosed (1) human-based DNA, RNA/protein biomarker panel, (2) EBV-based panel and (3) cytometric panel for both (1) non-invasive rapid test with biological samples from nasopharyngeal swab or plasma, or (2) conventional invasive endoscopy biopsy.
- [0124] Protein expression can be detected using antibodies that bind to the protein, in combination with methods of protein detection known in the art. mRNA can be detected using Real-Time Polymerase Chain Reaction (RT-PCR) (primers available at <https://pga.mgh.harvard.edu/primerbank/>) .
- [0125] RNA and protein expression can be determined using methods including but not limited to immunohistochemistry (IHC) /Enzyme-linked Immunosorbent Assay (ELISA) and a lateral flow assays. RNA and proteins can be detected using lateral flow

assays such as the Rapid Antigen Test (antibody/aptamer probe-based lateral flow test targeting DN, RNA and protein biomarkers) .

- [0126] The presence of one or more mutations can be assessed in a sample obtained from the subject, using methods such as DNA sequencing, Droplet Digital Polymerase Chain Reaction (ddPCR) (primers designed using Bio-Rad platform at <https://www.bio-rad.com/digital-assays/assays-create/mutation>) , RT-PCR, aptamers, and single-cell RNA sequencing.
- [0127] Flow cytometry can be used (single cell based test targeting cell type-specific, DNA, RNA and protein biomarkers) .
- [0128] Other methods such as FISH (Fluorescence in situ hybridization) and immunohistochemistry can be used to detect specific biomarkers as disclosed therein.
- [0129] SNV and InDels in one or more genes disclosed herein that detect the presence of NPC or of a relapse in NPC, can be detected using real-time PCR, microarrays, sc-RNA sequencing, next generation sequencing and RNA or DNA single cell sequencing. A single nucleotide variant (SNV) is a variation of a single nucleotide in a population's genome. RNA sequencing (RNA-seq) is a genomic approach for the detection and quantitative analysis of messenger RNA molecules in a biological sample. scRNA-seq permits comparison of the transcriptomes of individual cells. The first, and most important, step in conducting scRNA-seq has been the effective isolation of viable, single cells from the tissue of interest. Next, isolated individual cells are lysed to allow capture of as many RNA molecules as possible. In order to specifically analyze polyadenylated mRNA molecules, and to avoid capturing ribosomal RNAs, poly [T] -primers are commonly used. Next, poly [T] -primed mRNA is converted to complementary DNA (cDNA) by a reverse transcriptase. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications is provided by Haques, et al., Genome Medicine 9, 75, 2017.
- [0130] The lateral flow assay can be used to detect a one or more expression biomarkers that detect the presence NPC or the presence of a relapse of NPC, identified and disclosed herein. Relevant mutations and expression biomarkers are discussed above.
- [0131] The assay generally involves combining the biological sample with an assay fluid, an analyte binding agent that specifically binds a drug analyte e.g., an antibody or aptamer, a calibration/control analyte, and a calibration/control binding agent that specifically binds the calibration analyte. Contacted capture particles may or may not have analyte bound to the analyte binding agent, depending on whether analyte is present in the fluid sample and whether analyte has bound to the analyte binding agent on the binding particles. Because there are multiple binding sites for analyte on the capture particles, the presence and the concentration of analyte bound to particles varies; the concentration of analyte bound to the particles increases proportionally

with the amount of analyte present in the fluid sample, and the probability of a particle being arrested in the sample capture zone similarly increases with increasing amount of analyte bound to the drug binding agent on the particles. Thus, the population of contacted binding particles may contain particles having various amount of analyte bound to the analyte binding agent, as well as particles having no analyte bound to the drug binding agent.

- [0132] In some forms, the NPC analyte and the control analyte have similar physical properties. For example, the control analyte is preferably a molecule of similar size to the NPC analyte of interest. In some forms, the analyte binding agent and the control binding agent also have similar properties. For example, if the analyte binding agent is an antibody, the calibration binding agent is also preferably an antibody. Moreover, the affinity and/or avidity of the calibration/control binding agent for the calibration/control analyte is preferably comparable (e.g., within one order of magnitude) to the affinity and/or avidity of the analyte binding agent for the NPC analyte.
- [0133] Preparation of Sample buffer
- [0134] In one embodiment, an assay fluid e.g., a sample buffer is introduced to the biological sample, forming a mixed fluid sample. The sample buffer used in antigen tests typically contains a combination of ingredients that help stabilize the sample and enhance the performance of the test. These ingredients can vary depending on the specific test kit, but commonly include one or more detergents, salts, blocking Proteins, preservatives, and buffering agents.
- [0135] The sample buffer typically includes one or more detergents to help break open cells and release viral particles. Exemplary detergents include Triton X-100, Triton™, X-114, IGEPAL®CA-630, TERGITOL™ 15-S-9, and Deviron®C16. In preferred forms, the detergent included in the sample buffer is Triton X-100 in a concentration from about 0.1% to about 1.0%, preferably about 0.5%. The sample buffer includes one or more salts to maintain the pH of the buffer and prevent degradation of the sample. Exemplary salts include Sodium chloride, ammonium sulfate, potassium chloride, and calcium chloride. In preferred forms, the salt included in the sample buffer is Sodium chloride in a concentration from about 0.1% to about 1.0%, preferably about 0.9%.
- [0136] The sample buffer includes one or more blocking agents e.g., blocking agents to aid stabilization of the cellular and Epstein-Barr viral particles and prevent them from sticking to the test components. Exemplary blocking agents include Bovine Serum Albumin (BSA), Polyvinylpyrrolidone (PVP), and purified proteins (e.g., casein). In preferred forms, the blocking agent included in the sample buffer is BSA in a concentration from about 0.1% to about 2.0%, preferably about 1.0%.

- [0137] In some forms, the sample buffer includes one or more preservatives to prevent bacterial growth and maintain the stability of the buffer. Exemplary preservatives include Sodium azide, and 0.01% thimerosal (merthiolate). In preferred forms, the preservative included in the sample buffer is Sodium azide in a concentration from about 0.02% to about 0.1%, preferably about 0.02%.
- [0138] In some forms, the sample buffer includes one or more additional buffering agents to maintain the pH of the solution and prevent interference with the test results. Exemplary buffering agents include Sodium phosphate, Potassium phosphate, and Sodium citrate. In preferred forms, the buffering agent included in the sample buffer is Sodium phosphate in a concentration from about 10mM to about 100mM, preferably about 50mM. In one preferred embodiment, the sample buffer contains Triton X-100 at a concentration of 0.5%, Sodium chloride at a concentration of 0.9%, BSA at a concentration of 1.0%, Sodium azide at a concentration of 0.02%, and Sodium phosphate at a concentration of 50mM. Bovine serum albumin (BSA): a protein that helps to stabilize the cellular and Epstein-Barr viral particles and prevent them from sticking to the test components.
- [0139] A. Lateral Flow Device/Assay
- [0140] In some forms, a lateral flow assay device is used to detect the one or more expression biomarkers disclosed herein in Section II. This device allows for rapid and simultaneous detection of genes/expression markers detecting NPC relapse and detecting the presence of NPC. In some forms, the device is Nucleic Acid Lateral Flow (NALF) device. In some forms, the device is a Nucleic Acid Lateral Flow ImmunoAssay device (NALFIA).
- [0141] In preferred embodiments, the disclosed NPC assay is a lateral flow assay, which is a form of immunoassay in which the test sample flows along a solid substrate via capillary action. FIG. 12A. As illustrated in FIG. 12A, a lateral flow device 10 includes a solid substrate 12, such as a membrane strip, having an application point 14, an optional conjugate zone 16, a capture zone 18, and an absorbent zone 20 (e.g., a wicking pad). Binding agents are optionally present in the conjugate zone 16. Capture agents are immobilized in the capture zone 18, which preferably contains a plurality of capture lines 22 for detecting captured analyte (capture complex). The sample pad ensures the controlled flow of the test solution, which migrates to the conjugate pad where nanoparticles labelled with antibodies (or any binding partner for the analyte in the sample) are stored. For example, the binding agent in the capture zone can be an antibody or biomarker binding fragment thereof, or an aptamer. If the target analyte is present, the labelled antibodies will bind to it and continue to migrate to the detection pad, whereupon the materials are captured by immobilized antibodies at a test line (T-line) to form a coloured strip while a subsequent control line (C-line) is used to col-

orimetric ally indicate that the solution has sufficiently migrated. Finally, the absorbent pad absorbs excess sample.

[0142] In some forms, the lateral flow device can use a multiplex detection format for detection of more than one target analytes. The assay can be performed over a strip containing a number of test lines equal to the number of target analytes to be analyzed. It is desirable to analyse multiple analytes simultaneously under same set of conditions. The multiplex detection format is useful in clinical diagnosis where multiple analytes which are inter-dependent in deciding the stage of a disease are to be detected. Lateral flow strips for this purpose can be built in various ways, i.e., by increasing length and test lines on a conventional strip or by making other structures like stars or T-shapes. The shape of the strip for the lateral flow device can be chosen based on the number of target analytes.

[0143] FIG. 11D-I show different configurations for an exemplary multiplex lateral flow device. The exemplary next-generation rapid diagnostic test station/hub includes sample collection/distribution structure which maximize detection of biomarkers with consistency and minimizes sample input for detection. Slices of detection strips can be customized for personalized disease and prognostic detection. For example, inflammation (inflam) and drug resistance (DR) can be simultaneously detected. The test can be conducted at home, community clinic or medical laboratory setting. Probes for DNA and RNA and protein biomarkers shown in Tables 1, 2 and/or the expression biomarkers disclosed herein can result in different colors indicating premalignancy to high malignancy, from low chance to high chance of early relapse within upcoming 3.5 years if conventional treatment alone is adopted.

[0144] In some forms, the binding agent in the capture zone include antibodies that bind to a biomarker selected from the group consisting of PSMA4, CALML3, SLC2A1, SNX3, LY6D, YBX1, and RPMS1, for example and the device can be used to detect NPCR. In some forms, the binding agent in the capture zone include antibodies that bind to a biomarker selected from the group consisting of CKAP4, SYNGR2, CFL1 and RPMS1 and the device can be used for NPCD. In some forms, the device includes multiple capture zones for simultaneous detection of more than one biomarker and/or simultaneous detection of NPCR and NPCD. The disclosed device can be used to detect the presence of any of the biomarkers disclosed herein, using a suitable binding partner.

[0145] i. Solid Substrate

[0146] The solid substrate 12, such as a membrane strip, can be made of a substance of sufficient porosity to allow movement of antibodies and analyte by capillary action along its surface and through its interior. Examples of suitable membrane substances include but are not limited to cellulose, cellulose nitrate, cellulose acetate, glass

fiber, nylon, polyelectrolyte ion exchange membrane, acrylic copolymer/nylon, and polyether sulfone. In a one embodiment, the membrane strip is made of cellulose nitrate (e.g., a cellulose nitrate membrane with a Mylar backing) or of glass fiber. In preferred forms, the membrane strip can be made of nitrocellulose. Nitrocellulose is a common binding matrix because of its high affinity for RNA and proteins and nucleic acids, and compatibility with a variety of detection methods (e.g., western blotting, dot-blot assays, and other RNA and protein or nucleic acid methods) .

- [0147] In an exemplary embodiment, the membrane strip is FUSION 5TM material (Whatman) , which is a single layer matrix material that performs all of the functions of a lateral flow strip.
- [0148] In some forms, the membrane strip is a nitrocellulose membrane or a polyvinylidene difluoride (PVDF) membrane, preferably a nitrocellulose membrane. Nitrocellulose membranes are known in the art and are commercially available e.g., from ThermoFisher Scientific (Cat#77010) and BioRad (Cat Nos: 1620115, 1620113 and 1620114) . For nitrocellulose membranes, the optimal pore size is about 0.1 μ m to about 0.45 μ m.
- [0149] ii. Application Point
- [0150] The solid substrate 12 includes an application point 14, which can optionally include an application pad. For example, if the sample containing the analyte contains particles or components that should preferentially be excluded from the immunoassay, an application pad can be used. The application pad may be used to modify the biological sample, e.g., adjust pH, filtering out solid components, separate whole blood constituents, and adsorb out unwanted antibodies. If an application pad is used, it rests on the membrane, immediately adjacent to or covering the application point. The application pad can be made of an absorbent substance which can deliver a fluid sample, when applied to the pad, to the application point on the membrane. Representative substances include cellulose, cellulose nitrate, cellulose acetate, nylon, polyelectrolyte ion exchange membrane, acrylic copolymer/nylon, polyether sulfone, or glass fibers. In one embodiment, the pad is a HemasepTM-V pad (Pall Corporation) . In another embodiment, the pad is a PallTM 133, PallTM A/D, or glass fiber pad.
- [0151] iii. Conjugate Zone
- [0152] The solid substrate 12 optionally contains a conjugate zone 16, which includes a conjugate pad containing binding agents. In some embodiments, the conjugate zone contains binding agents which bind the analyte to be measured and a control analyte. When the sample migrates through the conjugate zone containing binding agents, the analytes in the sample interacts with the binding agents to form capture complexes. In some forms, the conjugate pad can be made of cellulose fibers, glass fibers, or

plastic such as polyester, polypropylene, or polyethylene. Exemplary conjugate pads include WhatmanTM conjugate release pads, SureWick® Cellulose Fibre Pads and SureWick® Glass Fibre Pads.

[0153] iv. Absorbent Zone

[0154] In some forms, the lateral flow device contains an absorbent zone. The absorbent zone 20 preferably contains an absorbent pad. As shown in FIG. 12A, absorbent pads, when used, are placed at the distal end of the lateral flow device. If an absorbent pad is present, it can similarly be made from such absorbent substances as are described for an application pad. In a preferred embodiment, an absorbent pad allows continuation of the flow of liquid by capillary action past the capture zones and facilitates the movement of non-bound agents away from the capture zones. In some forms, the absorbent pad increases the total volume of sample that enters the lateral flow device. This increased sample volume facilitates washing unbound agents away from the test and control lines, thereby lowering the background and enhancing assay sensitivity.

[0155] v. Capture Zone

[0156] The capture zone 18 contains capture agent immobilized (e.g., coated on and/or permeated through the membrane) to the membrane strip. In preferred embodiments, the capture agent is conjugated to a capture particle that is immobilized in the capture zone 18.

[0157] The capture zone 18 is preferably organized into one or more capture lines containing capture agents. In preferred embodiments, the capture zone contains a plurality of capture lines for multiplex analysis, i.e., detection of two or more analytes. In addition, the capture zone 18 may contain one or more control capture lines for detecting the presence of control analyte (i.e., control or calibration capture zone). The “lines” for interacting with the materials in the liquid flow can be in a variety of shapes, orientations, and relationships. Most commonly, the “lines” are linear strips of material perpendicular to liquid flow. Also most commonly, different “lines” with different components are separate and do not overlap. These features are most consistent with the mechanics and operation of lateral flow devices. However, the lines can be in shapes other than a strip, can be oriented other than perpendicular to the liquid flow, and can overlap. For example, some lateral flow devices have the test line and the control line perpendicular to each other and overlapping so as to form a + symbol when both lines show a detectable signal. Preferably, the control analyte capture reagent specifically binds the control analyte but does not interact with the sample analyte being measured.

[0158] The calibration capture zone is preferably positioned such that the sample capture zone is between the application point and the calibration capture zone. In a preferred embodiment, the calibration capture zone is closely adjacent to the sample capture

zone, so that the dynamics of the capillary action of the components of the assay are similar (e.g., essentially the same) at both the calibration capture zone and the sample capture zone. For example, the two capture zones are sufficiently close together such that the speed of the liquid flow is similar over both zones. Although they are closely adjacent, the calibration capture zone and the sample capture zone are also sufficiently spaced such that the particles arrested in each zone can be quantitated individually (e.g., without crosstalk). Furthermore, in some forms, the sample capture zone is separated from the application point by a space that is a large distance, relative to the small distance between the sample capture zone and the calibration capture zone. Because particle capture is a rate limiting step in the assay, the distance between the application point and the capture zones (where particles are captured) must be sufficient to retard the speed of the liquid flow to a rate that is slow enough to allow capture of particles when the liquid flow moves over the sample capture zone. The optimal distances between the components on the membrane strip can be determined and adjusted using routine experimentation.

- [0159] In some embodiments, the capture zone 18 contains at least one capture line 22 with capture agents for detecting a dilution control analyte, i.e., an analyte that is typically present in the biological sample at predictable concentrations. Creatine is a particularly preferred dilution control analyte when the biological sample is urine. The typical human reference ranges for serum creatinine are 0.5 to 1.0 mg/dL (about 45-90 $\mu\text{mol/L}$) for women and 0.7 to 1.2 mg/dL (60-110 $\mu\text{mol/L}$) for men. Control analyte for nasal mucus?
- [0160] In some embodiments, the capture zone 18 contains one or more capture lines with capture agents for detecting reference analytes. The reference analytes may be administered to the biological sample at known concentrations. These reference values can facilitate quantitative correlations between label detection and analyte amounts.
- [0161] vi. Capture Particles
- [0162] Capture particles are particles, such as polymeric particles, which can be coated with the capture agent and immobilized to the membrane in the capture zone 18. In preferred embodiments, the particles are physically trapped within the membrane. This allows for selection of optimal particle chemistry that is not influenced by the need for chemical immobilization. Suitable capture particles include liposomes, colloidal gold, organic polymer latex particles, inorganic fluorescent particles, and phosphorescent particles. In some embodiments, the particles are polystyrene latex beads, and most particularly, polystyrene latex beads that have been prepared in the absence of surfactant, such as surfactant-free Superactive Uniform Aldehyde/Sulfate Latexes (Interfacial Dynamics Corp., Portland, Oregon).

- [0163] In preferred embodiments, the particles are monodispersed polymer microspheres based on melamine resin (MF) (e.g., available from Sigma-Aldrich). Melamine resin microspheres are manufactured by acid-catalysed hydrothermal polycondensation of methylol melamine in the temperature range of 70-100 °C without any surfactants. Unmodified MF particles have a hydrophilic, charged surface due to the high density of polar triazine-amino and -imino groups. The surface functional groups (methylol groups, amino groups, etc.) allow covalent attachment of other ligands. For special applications, the MF particles can be modified by incorporation of other functionalities such as carboxyl groups. This increases possible surface derivatization such as chromophore or fluorophore labelling.
- [0164] The particles can be labelled to facilitate detection by a means which does not significantly affect the physical properties of the particles. For example, the particles can be labelled internally (that is, the label is included within the particle, such as within the liposome or inside the polystyrene latex bead). Representative labels include luminescent labels; chemiluminescent labels; phosphorescent labels; fluorescent labels; phosphorescent labels; enzyme-linked labels; chemical labels, such as electroactive agents (e.g., ferrocyanide); and colorimetric labels, such as dyes. In one embodiment, a fluorescent label is used. In another embodiment, phosphorescent particles are used, particularly up-converting phosphorescent particles, such as those described in U.S. Patent No. 5,043,265.
- [0165] The particles are preferably coated with capture agent, such as a sample analyte capture agent and control analyte capture agent. They can be prepared by mixing the capture agent in a conjugation buffer. A covalent coupling onto the particles is then performed, resulting in random binding of the capture agents onto the particle.
- [0166] b. Binding Agents and Capture Agents
- [0167] Binding agents for use in the disclosed assays include any molecule that selectively binds NPC analytes or calibration analytes. In preferred embodiments, the binding agents are antibodies, such as monoclonal antibodies, or aptamers, such as nucleic acid or peptide aptamers.
- [0168] i. Antibodies
- [0169] Antibodies that can be used in the compositions and methods include whole immunoglobulin (i.e., an intact antibody) of any class, fragments thereof, and synthetic RNA and proteins containing at least the antigen binding variable domain of an antibody. The variable domains differ in sequence among antibodies and are used in the binding and specificity of each antibody for its specific antigen. However, the variability is not usually evenly distributed through the variable domains of antibodies. It is typically concentrated in three segments called complementarity determining regions (CDRs) or hypervariable regions both in the light chain and the heavy chain

variable domains. The more highly conserved portions of the variable domains are called the framework (FR). The variable domains of native heavy and light chains each comprise four FR regions, largely adopting a beta-sheet configuration, connected by three CDRs, which form loops connecting, and in some cases forming part of, the beta-sheet structure. The CDRs in each chain are held together in proximity by the FR regions and, with the CDRs from the other chain, contribute to the formation of the antigen binding site of antibodies. Therefore, the disclosed antibodies contain at least the CDRs necessary to maintain DNA binding and/or interfere with DNA repair.

- [0170] Fragments of antibodies which have bioactivity can also be used. The fragments, whether attached to other sequences or not, include insertions, deletions, substitutions, or other selected modifications of specific regions or amino acids residues, provided the activity of the fragment is not significantly altered or impaired compared to the non-modified antibody or antibody fragment.
- [0171] Techniques can also be adapted for the production of single-chain antibodies specific to an antigenic RNA and protein of the present disclosure. Methods for the production of single-chain antibodies are well known to those of skill in the art. A single chain antibody can be created by fusing together the variable domains of the heavy and light chains using a short peptide linker, thereby reconstituting an antigen binding site on a single molecule. Single-chain antibody variable fragments (ScFvs) in which the C-terminus of one variable domain is tethered to the N-terminus of the other variable domain via a 15 to 25 amino acid peptide or linker have been developed without significantly disrupting antigen binding or specificity of the binding. The linker is chosen to permit the heavy chain and light chain to bind together in their proper conformational orientation.
- [0172] Divalent single-chain variable fragments (di-ScFvs) can be engineered by linking two ScFvs. This can be done by producing a single peptide chain with two VH and two VL regions, yielding tandem ScFvs. ScFvs can also be designed with linker peptides that are too short for the two variable regions to fold together (about five amino acids), forcing ScFvs to dimerize. This type is known as diabodies. Diabodies have been shown to have dissociation constants up to 40-fold lower than corresponding ScFvs, meaning that they have a much higher affinity to their target. Still shorter linkers (one or two amino acids) lead to the formation of trimers (triabodies or tribodies). Tetrabodies have also been produced. They exhibit an even higher affinity to their targets than diabodies.
- [0173] Suitable antibodies for use in the disclosed NPC assay are known in the art and/or commercially available.
- [0174] Anti-PSMA4 antibodies are available as follows: (RNA and Protein tech (polyclonal-Cat#11943-2-AP; monoclonal-Cat#68203-1-Ig); Kondo et al., J. Biol Chem., 295 (6) :

- 1658-1672 (2020) ; Benvenuto et al., Sci Rep., 11 (1) : 19051 (2021) ; Chadchankar et al., PLoS ONE, 14 (11) : e0225145 (2019)) ; PSMA4 Monoclonal antibody (RNA and protein tech,) ; Abcam, Cat#ab191403) ; ThermoFisher Scientific, monoclonal-Cat #MA5-25812; polyclonal-Cat #PA5-76658) ; Novus Biologicals, Cat#NBP2-38754) ; and G Biosciences (ITA7053-100u) .
- [0175] Anti-CALML3 antibodies are available as follows: Millipore Sigma (Polyclonal-Cat#SAB1400036) ; ThermoFisher Scientific, Cat #PA5-30232; referenced in Bun-banjerdsuk et al., Mod. Pathol., 32 (7) : 943-956 (2019)) ; ThermoFisher Scientific, monoclonal-Cat #MA5-29079; polyclonal-Cat #PA5-118992) ; Novus Biologicals (Cat#NBP2-15667) ; RNA and Protein Tech (Cat#17275-1-AP and Novus Biologicals (cat #-NBP2-90114B) .
- [0176] Anti-SLC2A1 antibodies include, but are not limited to: Glut1 Polyclonal Antibody (Novus Biologicals, Cat#NB110-39113; as referenced in Pham, et al., Cancers (Basel) , 14 (5) : 1311 (2022) ; Balukoff et al., Nat Commun., 11 (1) : 5755 (2020) .
- [0177] Anti-GLUT1 antibodies are available as follows: GLUT1 Polyclonal Antibody (ThermoFisher Scientific, Cat #PA5-16793 and Cat #PA5-32428; Liu et al., Nat Cell Biol., 22 (4) : 476-486 (2020)) ; GLUT1 Recombinant Rabbit Monoclonal Antibody (SA0377) (ThermoFisher Scientific, Cat #MA5-31960; Wang et al., Cell Reports, 28 (5) : 1323-1334. e4 (2019) and Risha et al., Sci. Reports, 10 (1) : 13572 (2020)) ; Recombinant Anti-Glucose Transporter GLUT1 antibody (Abcam Cat. #ab115730; Sato T et al. Sci Rep 12: 74 (2022) ; Wu F et al. Cell Death Discov 8: 3 (2022) ; and Zhou X et al., Bioengineered 13: 2471-2485 (2022)) ; and RNA and protein tech, Cat#21829-1-AP, Li et al., Cancer Cell, 33 (3) : 368-385. e7 (2018) ; Wang et al., Int J Oral Sci. 10 (3) : 27 (2018)) .
- [0178] Anti-SNX3 antibodies are available as follows: SNX3 Polyclonal Antibody (Cusabio Biotech, Cat#CSB-PA589999) ; Rabbit Polyclonal Anti-SNX3 antibody (Abcam, Cat#ab56078; Cicek E et al., Oncogene 41: 220-232 (2022) ; Yang et al., Cells 11 (21) : 3358 (2022) ; and Cui Y et al., Traffic 22: 123-136 (2021)) ; SNX3 Polyclonal antibody (RNA and protein tech; Cat#10772-1-AP; O'Farrell, et al., Nat Cell Biol, 19 (12) : 1412-1423 (2017) ; McGough et al., Nature Commun., 9 (1) : 3737 (2018) ; and Lu et al., Cell Death Differentiation, 28 (10) : 2871-2887 (2021)) ; SNX3 Polyclonal Antibody (ThermoFisher Scientific, Cat #PA5-101906) ; and SNX3 Monoclonal Antibody (G-7) (Santa Cruz, Cat#sc-376667; Xu et al., Neurodegenerative Diseases, 18 (1) : 26-37 (2018)) .
- [0179] Anti-LY6D antibodies are available as follows: : Anti-LY6D Polyclonal Antibody (Sigma-Aldrich, Cat#HPA024755) ; LY6D Polyclonal Antibody (ThermoFisher Scientific, Cat #PA5-64167; DePianto et al., JCI Insight, 6 (8) : e143626 (2021)) ; Ly-6D Monoclonal Antibody (49-H4) , PE, eBioscience™ (ThermoFisher Scientific,

- Cat#12-5974-80; Barros-Silva, et al., Cell Reports, 25 (12) : 3504-3518. e6 (2018)) ; and LY6D Polyclonal Antibody (RNA and protein tech, Cat#17361-1-AP; Yao et al., Nat Commun, 11 (1) : 5079 (2020) ; Steiner et al., Cell Reports, 42 (4) : 112377 (2023) ; and Zhang et al., Diabetes, 66 (6) : 1535-1547 (2017)) .
- [0180] Anti-YBX1 antibodies are available as follows: YBX1 Polyclonal antibody (RNA and protein tech, Cat#20339-1-AP, referenced in An et al., Nature, 583 (7815) : 303-309 (2020) ; and Zhang et al., Sci Adv, 8 (5) : eabj3967 (2022)) ; YBX1 Recombinant Rabbit Monoclonal Antibody (10H29L41) (ThermoFisher Scientific, Cat #702245) ; YBX1 Polyclonal Antibody (ThermoFisher Scientific, Cat #PA5-83493, referenced in Lin et al., Stem Cell Res Ther., 10 (1) : 263 (2019)) ; and Anti-YB1 Rabbit Polyclonal antibody (Abcam, Cat#ab12148, referenced in Feng et al., JCI Insight 7 (6) : e150091 (2022) and Gao et al., Oncogenesis 11: 13 (2022)) .
- [0181] Anti-RPMS1 antibodies are available as follows: : RPMS Polyclonal antibody (Zhang et al., J Virol., 75 (6) : 2946-2956 (2001) ; and Anti-RPMS1 peptide (SGQPRWWPWG) antibody; Smith et al., J Virol., 74 (7) : 3082-3092, (2000)) .
- [0182] Anti-HMGN2P3 antibodies are available as follows: Novus Biologicals, Cat# NBP3-12793) .
- [0183] Anti-DNAJC11 antibodies are available as follow: DNAJC11 Polyclonal antibody (RNA and protein tech, Cat#17331-1-AP, Violitzi et al, J Proteome Res., 18 (11) : 3896-3912 (2019)) ; Recombinant Anti-DNAJC11 antibody [EPR15065 (B)] -C-terminal (Abcam, Cat# ab183518) ; DNAJC11 Polyclonal Antibody (ThermoFisher Scientific, Cat #PA5-85470; Cat #PA5-100479; and Cat #PA5-55956) .
- [0184] Anti-EIF2AK1 antibodies are available as follows: EIF2AK1 Polyclonal Antibody (Novus Biologicals, Cat#NBP1-83210; Cat#NBP1-56484; Cat#NBP3-05000; and Cat# NBP3-04999) and EIF2AK1 Monoclonal Antibody (2H1F3) (RNA and protein tech, Cat# 20499-1-AP, referenced in Fessler, et al., Nature, 579 (7799) : 433-437 (2020)) .
- [0185] Anti-FAM234A antibodies are available as follows: FAM234A Mouse Monoclonal Antibody [A5-A10] (HUABIO, Cat#M1010-1) ; and Anti-FAM234A Polyclonal Antibody (Atlas Antibodies, Cat#HPA071871) .
- [0186] An anti-PARPBP antibodies are available as follows: Novus Biologicals, Cat# NBP1-93969) . and ThermoFisher Scientific (Cat#PA5-58877) .
- [0187] Anti-ARL5A antibodies are available as follows: ARL5A Polyclonal Antibody (ThermoFisher Scientific, Cat #PA5-30509, Cat #PA5-67615; Cat #PA5-114063; and Cat #PA5-55479) and Anti-ARL5A Rabbit Polyclonal antibody (Abcam, Cat#ab104008) .
- [0188] Anti-DHX57 antibodies include, but are not limited to: DHX57 Rabbit Polyclonal Antibody (Thermofisher Scientific, Cat #PA5-57548 and Cat #PA5-95809) ; and Polyclonal Anti-DHX57 Antibody (Atlas Antibodies, Product#HPA036160) .

[0189] Other exemplary antibodies that can be used in the disclosed NPC assay are provided in Table 2 and Table 3.

[0190] Table 2: List of Exemplary Antibodies That Can Be Used in the disclosed assays

Gene	Antibody Name	Catalog Number	Company
NEDD8	NEDD8 Polyclonal Antibody	bs-3812R-Biotin	Bioss Inc.
CALML3	CALML3 Antibody	NBP2-90114B	Novus Biologicals
NDUFA13	Rabbit Anti-NDUFA13	038858-Biotin	United States Biological
BEX3	Rabbit Anti-NGFRAP1	038955-Biotin	United States Biological
HNRNPA0	Immunotag™ ROA0 Polyclonal Antibody	ITN0047-50u	G Biosciences
SLIRP	SLIRP/C14orf156 Polyclonal Antibody	bs-21033R	Bioss Inc.
ADH5	Immunotag™ ADH5 Antibody	ITA6855-100u	G Biosciences
GNG5	GNG5 antibody	orb24361	Biorbyt
UBE2D2	UBE2D2 antibody	orb415016	Biorbyt
PSMA4	Immunotag™ PSMA4 Antibody	ITA7053-100u	G Biosciences
SLC2A1	Slc2A1 Recombinant Antibody	CAC12206	Biomatik

SNX3	Immunotag™ SNX3 Antibody	ITA5415-100u	G Biosciences
LY6D	LY6D antibody	orb29008	Biorbyt
PSMA3	PSMA3 antibody	orb50969	Biorbyt
PSMB1	Proteasome subunit beta type-1 antibody	orb239675	Biorbyt
YBX1	Mouse Anti-YBX1 Antibody	135476-Biotin	United States Biological
PSMB5	PSMB5 antibody	orb46795	Biorbyt
PSMA6	Immunotag™ PSMA6 Antibody	ITA6988-100u	G Biosciences
CKAP4	CKAP4 antibody	orb414676	Biorbyt
SYNGR2	Mouse Anti-SYNGR2 Antibody	134136-Biotin	United States Biological
MARCKSL1	Mouse Anti-MARCKSL1 Antibody	129435-Biotin	United States Biological
CFL1	CFL1 antibody	orb44566	Biorbyt
PDCD5	PDCD5 Antibody	NBP1-76988B	Novus Biologicals
IL32	Mouse anti Human Interleukin-32:Biotin Antibody	MBS226187	MyBioSource.com
RAN	Ran Polyclonal Antibody	bs-4282R-Biotin	Bioss Inc.
NME1	NME1 Polyclonal Antibody	bs-1066R-Biotin	Bioss Inc.
VCAM1	VCAM1 Rabbit Polyclonal Antibody	AP22317BT-N	OriGene Technologies
GAPDH	Goat Anti-GAPDH Antibody	MBS423610	MyBioSource.com
TUBB	tubulin Beta Polyclonal Antibody	bs-4511R-Biotin	Bioss Inc.
HSPD1	HSPD1 antibody	orb687431	Biorbyt
LGALS1	LGALS1 antibody	orb352068	Biorbyt
TNFAIP3	TNFAIP3 antibody	orb47522	Biorbyt
ITGAV	ITGAV antibody	orb414808	Biorbyt
RPLP0	RPLP0 antibody	orb352268	Biorbyt
STMN1	STMN1 antibody	orb46146	Biorbyt
PPIA	Ppia Recombinant Antibody	CAC12118	Biomatik
CCL20	Biotin anti-human CCL20	534604	BioLegend
FSCN1	FSCN1 antibody	orb401039	Biorbyt
LGALS9	Galactin-9 antibody	orb240263	Biorbyt
RPS19	Rabbit Anti-RPS19, C.T Antibody	041238-Biotin	United States Biological
TPI1	TPI1 antibody	orb46998	Biorbyt
ICAM1	ICAM1 antibody	orb46974	Biorbyt
ENO1	Alpha-enolase antibody	orb239777	Biorbyt
C1QBP	Mouse Anti-Human C1QBP Monoclonal Antibody	CTMM-0322-JL145	Creative Biolabs
CXCL3	CXCL3 antibody	orb27053	Biorbyt
MIF	Goat anti-MIF, Biotinylated Antibody	MBS423623	MyBioSource.com
TAGLN2	Rabbit Anti-TAGLN2, CT Antibody	042589-Biotin	United States Biological

CXCL10	Mouse anti Human CXCL10:Biotin AB	MBS226131	MyBioSource.com
UBD	UBD antibody	orb350791	Biorbyt
CSTA	CSTA antibody	orb350388	Biorbyt
NFKBIA	NFKBIA antibody	orb688111	Biorbyt
SOCS1	anti-SOCS1 Antibody	ABIN5539590	antibodies-online
MYBPC1	MYBPC1 antibody	orb415924	Biorbyt
NTRK2	Mouse Anti-NTRK2 Antibody	130601-Biotin	United States Biological
FKBP1A	Mouse Anti-FKBP1A Antibody	126829-Biotin	United States Biological
SOX4	SOX4 antibody	orb51073	Biorbyt
TUBA1C	Rabbit Anti-TUBA1C, CT Antibody	043449-Biotin	United States Biological
PYCARD	anti-PYCARD Antibody	ABIN5539792	antibodies-online
TCIM	Anti-C8orf4 TCIM Antibody	A31995	BosterBio
SLPI	Mouse Anti-SLPI Antibody	133526-Biotin	United States Biological
WFDC2	Mouse Anti-WFDC2 Antibody	135376-Biotin	United States Biological
AQP3	Aquaporin 3 Antibody BIOTIN-Conjugated	MBS540993	MyBioSource.com
LCN2	NGAL (LCN2) Biotinylated Mouse Monoclonal Detection Antibody	TA700555	OriGene Technologies
BPIFB1	BPIFB1 Antibody	NBP2-90203B	Novus Biologicals
SERPINB3	Serpinb3 Polyclonal Antibody	CAC07323	Biomatik
CLU	CLU antibody	orb238210	Biorbyt
SCGB1A1	SCGB1A1 antibody	orb46074	Biorbyt
CRIP1	Rabbit Anti-CRIP1, CT Antibody	034224-Biotin	United States Biological
KRT14	Keratin, type I cytoskeletal 14 antibody	orb242952	Biorbyt
LYPD2	LYPD2 Antibody	MBS969993	MyBioSource.com
TFF3	Anti-Trefoil factor 3 (TFF3) Antibody	130-10147B-50	RayBiotech
TSPAN1	TSPAN1 antibody	orb357352	Biorbyt
SCGB3A1	Mouse Anti-SCGB3A1 Antibody	133004-Biotin	United States Biological
PIGR	Rabbit Anti-PIGR, CT Antibody	P4165-50-Biotin	United States Biological
MUC16	MUC16 / CA125 Monoclonal Antibody	LS-C86749	LSBio
C19orf33	C19ORF33 Antibody	orb1520525	Biorbyt
PRSS23	Serine protease 23 antibody	orb242376	Biorbyt
MSMB	MSMB antibody	orb400511	Biorbyt
CAPS	Caps Polyclonal Antibody	CAC11552	Biomatik
CXCL17	Immunotag™ CXCL17 Antibody	ITA2451-100u	G Biosciences
ANXA1	Annexin A1 antibody	orb241488	Biorbyt
AGR2	AGR2 antibody	orb686795	Biorbyt
GSTA1	GSTA1 antibody	orb24485	Biorbyt
BPIFA1	Anti-Plunc/BPIFA1 Antibody	A03162-2	BosterBio
MT1X	Metallothionein-IX antibody	orb241394	Biorbyt

ALDH3A1	ALDH3A1 Antibody	MBS1499519	MyBioSource.com
C20orf85	C20orf85 Antibody	PA5-64228	ThermoFisher Scientific
LGALS3	Galectin 3 antibody	orb238943	Biorbyt
MUC4	Biotin-Linked Antibody to Mucin 4	MBS2005469	MyBioSource.com
TACSTD2	TROP2 / TACSTD2 Antibody	10428-MM01-B	Sino Biological, Inc.
KRT16	KRT16 antibody	orb517911	Biorbyt
CEBPD	CEBP Delta Antibody	NBP2-41192B	Novus Biologicals
CDKN1A	Mouse Anti-CDKN1A Antibody	124806-Biotin	United States Biological
PGM2	Mouse Anti-PGM2 Antibody	131218-Biotin	United States Biological
MEG3	MEG3 antibody	ABIN7138336	Antibodies-online
CTNNBIP1	Mouse Anti-CTNNBIP1 Antibody	125456-Biotin	United States Biological
IGF2BP3	IGF2BP3 antibody	orb46811	Biorbyt
FAF1	FAF1 antibody	orb47258	Biorbyt
DUSP11	DUSP11 antibody	orb686647	Biorbyt
CLDND1	CLDND1 antibody	orb687667	Biorbyt
BAG4	BAG4 antibody	orb47874	Biorbyt
SERPINB12	SERPINB12 antibody	orb34842	Biorbyt
AP3M2	AP3M2 antibody	orb21258	Biorbyt
SIPA1L2	SIPA1L2 antibody	orb356716	Biorbyt
CLCA4	CLCA4 Antibody	MBS8561978	MyBioSource.com
SYT8	Rabbit Anti-SYT8, ID Antibody	042538-Biotin	United States Biological
FGFR3	FGFR3 Antibody BIOTIN-Conjugated	MBS541609	MyBioSource.com
SUSD4	SUSD4 antibody	orb357008	Biorbyt
TNNT3	anti-TNNT3 Antibody	ABIN5539524	antibodies-online
NSG1	Anti-NSG1 Antibody	A12107-1	BosterBio
GPX2	GPX2 antibody	orb52630	Biorbyt
LSP1	LSP1 Antibody	NBP3-00254B	Novus Biologicals
RPMS1	Zhang, et al., <i>J. Virology</i> , 75(6):2946-2956		
BALF4	InVivoMAb Anti-EBV/HHV4 gB/BALF4 Antibody	VVV22002	Antibody System
BALF5	BALF5 antibody	orb351876	Biorbyt
BALF0	Zhao et al., <i>Viruses</i> 11(12): 1099); Shao et al., <i>Protein Expr. Purif.</i> 2019; 162:44–50.		
LMP-1	Anti-EBV/HHV-4 LMP1/BNLF1 Antibody	RVV07003	Antibody System
LMP-2B	EBV LMP2 Polyclonal Antibody	BS-4700R	ThermoFisher Scientific

[0191] Table 3. Additional List of Exemplary Antibodies That Can Be Used in the disclosed assays

Gene	Antibody Name	Catalog Number	Company
------	---------------	----------------	---------

ABCB6	Mouse Anti-ABCB6 Antibody	122795-Biotin	United States Biological
ADAMTS7	ADAMTS7 antibody	orb353824	Biorbyt
AKAP12	Mouse Anti-AKAP12 Antibody	123132-Biotin	United States Biological
AMDHD1	AMDHD1 Antibody	CSB-PA850402LD01HU	CUSABIO Technology LLC
ANK3	Ank3 Polyclonal Antibody	CAC11390	Biomatik
ANKRD30A	Immunotag™ ANKRD30A Polyclonal Antibody	ITT0232-50u	G Biosciences
ANTXR2	ANTXR2 Polyclonal Antibody	MBS2541669	MyBioSource.com
AP1B1	AP1B1 Antibody	MBS3220159	MyBioSource.com
APP	App Recombinant Antibody	CAC11986	Biomatik
ARFGAP1	ARFGAP1 Polyclonal Antibody	CAF50460	Biomatik
ATOX1	ATOX1 antibody	orb517155	Biorbyt
ATP6V0A1	Atp6V0A1 Polyclonal Antibody	CAC09593	Biomatik
B4GALT5	B4Galt5 Polyclonal Antibody	CAC09235	Biomatik
BDH2	BDH2 antibody	orb354056	Biorbyt
BPNT1	Mouse Anti-BPNT1 Antibody	243865-Biotin	United States Biological
BRD4	Anti-BRD4 Antibody Picoband™	A00123-2	BosterBio
CBX1	Mouse Anti-CBX1 Antibody	124402-Biotin	United States Biological
CD69	Anti-CD69 Antibody Picoband™	A00529-2	BosterBio
CDH1	CDH1 Monoclonal Antibody	CAC10270	Biomatik
CDK13	Anti-CDK13 Antibody Picoband™	A05292-1	BosterBio
CHST7	Anti-CHST7 antibody	GTX120778	GeneTex
CLEC14A	CLEC14A antibody	orb517775	Biorbyt
CLEC3B	CLEC3B/Tetranectin Antibody	NBP3-00044B	Novus Biologicals
CNBP	CNBP antibody	orb350368	Biorbyt
CNDP1	Rabbit Anti-CNDP1, ID Antibody	034035-Biotin	United States Biological
COL3A1	COL3A1 Antibody	CSB-PA11459D0Rb	CUSABIO Technology LLC
CROCC	Immunotag™ CROCC Polyclonal Antibody	ITN1244-50u	G Biosciences
CTSA	CTSA antibody	orb23460	Biorbyt
CXCL6	CXCL6 antibody	orb354512	Biorbyt
CYB561D2	Immunotag™ CYB561D2 Polyclonal Antibody	ITT1164-50u	G Biosciences
DAG1	DAG1 antibody	orb686839	Biorbyt
DHTKD1	Rabbit Anti-DHTKD1, ID Antibody	034629-Biotin	United States Biological
DLST	Anti-DLST Antibody Picoband™	A05097-1	BosterBio
DPP3	Rabbit Anti-DPP3, CT Antibody	034781-Biotin	United States Biological
DUSP19	DUSP19 Antibody	MBS7002143	MyBioSource.com
EPB41	EPB41 Polyclonal Antibody	CAF50672	Biomatik
EPHA1	EPHA1 antibody	orb45833	Biorbyt
ERAP2	Anti-ERAP2 Antibody Picoband™	A04269-1	BosterBio
ESD	Esd Polyclonal Antibody	CAC11845	Biomatik
EXOC3L4	EXOC3L4 Antibody	PA5-60455	Thermo Fisher Scientific

FBL	Anti-Fibrillarin/FBL Antibody Picoband™	A03178-1	BosterBio
GALE	Anti-GALE Antibody Picoband™	A00551-1	BosterBio
GAN	Mouse Anti-GAN Antibody	127153-Biotin	United States Biological
GAS2L1	Gas2L1 Polyclonal Antibody	CAC10792	Biomatik
GASK1A	GASK1A Rabbit Polyclonal Antibody	TA372310	OriGene Technologies
GNPAT	GNPAT antibody	orb414764	Biorbyt
HBM	Immunotag™ HBM Polyclonal Antibody	ITN2128-50u	G Biosciences
HDAC1	HDAC1 antibody	orb355160	Biorbyt
HDGFL2	Anti-HDGFL2 Antibody Picoband™	A32401-2	BosterBio
HEMGN	Rabbit Anti-HEMGN, CT Antibody	036505-Biotin	United States Biological
HLA-DRB3	HLA-DRB3 antibody	orb400275	Biorbyt
HMGB2	HMGB2 antibody	orb52041	Biorbyt
HNRNPA1	HNRNPA1 Antibody	CSB-PA010600HD01HU	CUSABIO Technology LLC
HNRNPL	HNRNPL antibody	orb355208	Biorbyt
IFI16	Anti-IFI16 Antibody Picoband™	A00848-3	BosterBio
JAG1	Anti-Jagged1/JAG1 Antibody Picoband™	A00640-2	BosterBio
LAMTOR5	LAMTOR5 Antibody	MBS9416937	MyBioSource.com
LCMT1	LCMT1 antibody	orb752613	Biorbyt
LRP5	Mouse Anti-LRP5 Antibody	129208-Biotin	United States Biological
LRRC75A	Anti-LRRC75A Antibody Picoband™	A18461-1	BosterBio
LSM6	LSM6 antibody	orb355492	Biorbyt
LUC7L2	Immunotag™ LUC7L2 Antibody	ITA2149-100u	G Biosciences
MAGT1	Rabbit Anti-MAGT1, NT Antibody	038074-Biotin	United States Biological
MANEAL	MANEAL Antibody	MBS3209665	MyBioSource.com
MAP2K2	MAP2K2 antibody	orb46962	Biorbyt
MARCHF2	Mouse Anti-MARCH2 Antibody	129428-Biotin	United States Biological
MED25	MED25 Antibody	MBS7043801	MyBioSource.com
MFGE8	MFGE8 Antibody	CSB-PA07974D0Rb	CUSABIO Technology LLC
MINPP1	Minpp1 Polyclonal Antibody	CAC09946	Biomatik
MOSPD2	Anti-MOSPD2 Antibody Picoband™	A16695-1	BosterBio
MROH1	MROH1 Antibody	MBS3219094	MyBioSource.com
MRPL17	Immunotag™ MRPL17 Antibody	ITA8843-100u	G Biosciences
MTHFS	MTHFS Antibody	MBS1493889	MyBioSource.com
NCAN	Immunotag™ NCAN Antibody	ITA5114-100u	G Biosciences
NDUFA12	Rabbit Anti-NDUFA12, CT Antibody	038857-Biotin	United States Biological
NFE2	Anti-NFE2 Antibody Picoband™	A03765-1	BosterBio
NID2	Mouse Anti-NID2 Antibody	130371-Biotin	United States Biological
NPEPL1	Mouse Anti-NPEPL1 Antibody	130495-Biotin	United States Biological
OGA	Mouse Anti-MGEA5 Antibody	129653-Biotin	United States Biological

PBX2	Mouse Anti-PBX2 Antibody	130921-Biotin	United States Biological
PCSK1	Mouse Anti-PCSK1 Antibody	249885-Biotin	United States Biological
PFDN4	PFDN4 Antibody	CSB-PA017814LD01HU	CUSABIO Technology LLC
PKM	PKM Antibody	MBS7006426	MyBioSource.com
PPP6R1	PPP6R1 Antibody	MBS9215757	MyBioSource.com
PRF1	PRF1 Antibody	MBS7004334	MyBioSource.com
PRMT9	PRMT9 Antibody	MBS7045457	MyBioSource.com
RALYL	RALYL antibody	orb356316	Biorbyt
RANBP3L	RANBP3L antibody	orb517839	Biorbyt
RAVER1	Anti-RAVER1 Antibody Picoband™	A09359-2	BosterBio
RDX	Immunotag™ RDX Antibody	ITA5254-100u	G Biosciences
REEP5	Mouse Anti-REEP5 Antibody	132462-Biotin	United States Biological
RIPOR1	RIPOR1 Antibody	MBS3218431	MyBioSource.com
SAMHD1	Anti-SAMHD1 Antibody Picoband™	A00592-4	BosterBio
SAR13	SAR13 antibody	orb356580	Biorbyt
SEH1L	SEH1L Polyclonal Antibody	bs-8721R-Biotin	Bioss Inc
SIGIRR	Sigirr Polyclonal Antibody	CAC10868	Biomatik
SMIM15	smim15 antibody	orb861403	Biorbyt
SNRPF	SNRPF Polyclonal Antibody	MBS9431508	MyBioSource.com
SRP14	Anti-SRP14 antibody	GTX101867	GeneTex
TARDBP	TARDBP antibody	orb515603	Biorbyt
TBC1D20	TBC1D20 Antibody	MBS969043	MyBioSource.com
THYN1	THYN1 Polyclonal Antibody	bs-7104R-Biotin	Bioss Inc.
TM4SF1	TM4SF1 antibody	orb357132	Biorbyt
TMEM132D	TMEM132D antibody	orb515821	Biorbyt
TNFSF12	TNFSF12 Antibody	MBS1492223	MyBioSource.com
TRAPPC11	TRAPPC11 antibody	orb51793	Biorbyt
TRAPPC2	Mouse Anti-TRAPPC2 Antibody	134692-Biotin	United States Biological
TRAPPC2B	Anti-ZNF547 Antibody	orb1880124	Biorbyt
TRIM28	Trim28 Polyclonal Antibody	CAC07189	Biomatik
USP47	USP47 antibody	orb357500	Biorbyt
VTI1A	VTI1A antibody	orb51897	Biorbyt
WDR41	WDR41 antibody	orb687647	Biorbyt
XYLT2	XYLT2 Antibody	MBS8505767	MyBioSource.com
ZNF326	Rabbit Anti-ZNF326, NT Antibody	044194-Biotin	United States Biological
ZNF428	ZNF428 Polyclonal Antibody	PA5-71584	Thermo Fisher Scientific
ZNF846	ZNF846 Antibody	NBP3-20736B	Novus Biologicals
ZNHIT1	Immunotag™ ZNHIT1 Antibody	ITA2300-100u	G Biosciences
GLS	GLS Antibody	MBS7003190	MyBioSource.com
DENND5B	Rabbit Anti-DENND5B, CT Antibody	034585-Biotin	United States Biological
COX8A	Immunotag™ COX82 Antibody	ITA2380-100u	G Biosciences
AUP1	Anti-AUP1 Antibody Picoband™	A08937-1	BosterBio

IGFBP1	IGFBP1 Monoclonal Antibody	CAC10267	Biomatik
GCG	GCG Antibody	MBS7136090	MyBioSource.com
PRAP1	Anti-PRAP1 Antibody	MBS8307774	MyBioSource.com
SYNPO2	SYNPO2 Antibody	NBP2-82037B	Novus Biologicals
FNIP2	FNIP2 Antibody	NBP2-81941B	Novus Biologicals
PGAM5	PGAM5 antibody	orb52329	Biorbyt
DAD1	DAD1 Antibody	CSB-PA03124D0Rb	CUSABIO Technology LLC
FCN1	Mouse Anti-FCN1 Antibody	126725-Biotin	United States Biological
FGL1	Fgl1 Polyclonal Antibody	CAC11231	Biomatik
SRGN	SRGN antibody	orb517305	Biorbyt
PGA3	PGA3 Antibody	CSB-PA11289D0Rb	CUSABIO Technology LLC
FMOD	FMOD antibody	orb52562	Biorbyt
HIVEP1	HIVEP1 Antibody	MBS3219261	MyBioSource.com
DEFA3	Mouse Anti-DEFA3 Antibody	245281-Biotin	United States Biological
HLA-DPB1	Hla-Drb3 Polyclonal Antibody	CAC11113	Biomatik
DUSP7	DUSP7 Polyclonal Antibody	bs-7928R-Biotin	Bioss Inc.
STOML3	STOML3 antibody	orb53702	Biorbyt
SLC17A5	Anti-SLC17A5 Antibody Picoband™	A06657-2	BosterBio
SLC2A2	SLC2A2 antibody	orb46262	Biorbyt
PROZ	Immunotag™ Protein Z Monoclonal Antibody	ITM0538-50u	G Biosciences
CALU	Mouse Anti-CALU Antibody	244176-Biotin	United States Biological
C1S	C1S antibody	orb414640	Biorbyt
SOD3	SOD3 antibody	orb356900	Biorbyt
CRP	Crp Recombinant Antibody	CAC12107	Biomatik
WWC3	WWC3 antibody	orb687019	Biorbyt
DPY19L1	DPY19L1 antibody	orb418046	Biorbyt
DHCR7	DHCR7 antibody	orb516422	Biorbyt
CGNL1	Rabbit Anti-CGNL1, NT Antibody	033821-Biotin	United States Biological
SLC25A1	Anti-Slc25a1 Picoband™ Antibody	A05995-1	BosterBio
OMD	OMD antibody	orb516500	Biorbyt
DPT	Mouse Anti-DPT Antibody	126005-Biotin	United States Biological
NTS	NTS antibody	orb351808	Biorbyt
TRIM41	Rabbit Anti-TRIM41, CT Antibody	043239-Biotin	United States Biological
CCL17	Biotin anti-human CCL17 (TARC)	523003	BioLegend
SLIT1	Rabbit Anti-SLIT1, NT Antibody	042009-Biotin	United States Biological
SH3BGR	SH3BGR antibody	orb356692	Biorbyt
ANTXR1	TEM8/ANTXR1 Antibody	NBP3-21501B	Novus Biologicals
EPB42	Mouse Anti-EPB42 Antibody+B282:B289	126370-Biotin	United States Biological
SLC4A1	Mouse Anti-SLC4A1 Antibody	133495-Biotin	United States Biological
OXA1L	Rabbit Anti-OXA1L, CT Antibody	039550-Biotin	United States Biological
TTN	Ttn Polyclonal Antibody	CAC07287	Biomatik

DST	Dst Polyclonal Antibody	CAC08167	Biomatik
TMEM132A	Anti-TMEM132A Antibody Picoband™	A12054-2	BosterBio
CKM	Ckm Polyclonal Antibody	CAC07384	Biomatik
KMT2B	KMT2B antibody	orb686407	Biorbyt
HLA-A	Immunotag™ HLA-A Antibody	ITA6967-100u	G Biosciences
CD2BP2	Rabbit Anti-CD2BP2, NT Antibody	033489-Biotin	United States Biological
SET	SET Polyclonal Antibody	CAF50321	Biomatik
SELENOH	SELENOH Antibody	MBS9630600	MyBioSource.com
CCL19	Ccl19 Recombinant Antibody	CAC12049	Biomatik
RHAG	Immunotag™ CD241 Polyclonal Antibody	ITT5284-50u	G Biosciences
EZR	EZR Antibody	CSB-PA02159D0Rb	CUSABIO Technology LLC
SLC29A1	SLC29A1 antibody	orb46518	Biorbyt
FRZB	Anti-FRZB Antibody Picoband™	PB9372	BosterBio
SEPTIN11	SEPT11 Antibody	MBS7134993	MyBioSource.com
GALC	Mouse Anti-GALC Antibody	127114-Biotin	United States Biological
SNRPB2	SNRPB2 antibody	orb351044	Biorbyt
SLFN5	SLFN5 Antibody	MBS9430160	MyBioSource.com
SLC12A4	SLC12A4 antibody	orb39051	Biorbyt
DSCAML1	Mouse Anti-DSCAML1 Antibody	245499-Biotin	United States Biological
CYP2C9	Immunotag™ CP2C9 Polyclonal Antibody	ITN1835-50u	G Biosciences
HSD11B1	Anti-HSD11B1 Antibody Picoband™	A01565	BosterBio
COL2A1	COL2A1 Antibody	MBS7136224	MyBioSource.com
STAU2	Mouse Anti-STAU2 Antibody	252178-Biotin	United States Biological
LTF	iotin-Linked Antibody to Lactoferrin	MBS2003498	MyBioSource.com
F11	F11 Polyclonal Antibody	CAC11850	Biomatik
C1QTNF3	C1Qtnf3 Polyclonal Antibody	CAC08270	Biomatik
PIP4K2A	PIP4K2A Polyclonal Antibody	CAF50266	Biomatik
JMY	JMY antibody	orb416048	Biorbyt
IGHV3-21	https://www.sciencedirect.com/science/article/pii/S00064971139869884		
DNAH6	DNAH6 Antibody	MBS3212343	MyBioSource.com
COCH	Rabbit Anti-COCH, CT Antibody	034070-Biotin	United States Biological
NSL1	NSL1 antibody	orb518019	Biorbyt
SPTB	Anti-beta 1 Spectrin/SPTB Antibody Picoband™	A02392	BosterBio
KRT9	Rabbit Anti-KRT9, ID Antibody	037637-Biotin	United States Biological
IGKV2-29	Recombinant Anti-Human IGKV2-29 Antibody	MOB-2718z	Creative Biolabs
IGKV3D-20	Recombinant Anti-Human IGKV3D-20 Antibody	MOB-1384z	Creative Biolabs
KIFC3	Anti-KIFC3 antibody	GTX114507	GeneTex
PGA4	Rabbit Anti-PGA4, CT Antibody	039908-Biotin	United States Biological
PGA5	Mouse Anti-PGA5 Antibody	131197-Biotin	United States Biological

[0192] Antibodies that specifically bind an analyte can also be made using routine methods. For example, antibodies can be purified from animals immunized with analyte.

Monoclonal antibodies can be produced by fusing myeloma cells with the spleen cells from a mouse that has been immunized with the opioid analyte or with lymphocytes that were immunized in vitro. Antibodies can also be produced using recombinant technology.

[0193] The capture agent of the disclosed compositions and methods may be an antibody, such as an anti-metatype antibody. Anti-metatype antibodies are immunological reagents specific for the conformation of the liganded antibody active site which do not interact with bound ligand or unliganded antibody. An antibody that selectively binds a capture complex but not to free analyte may be obtained using standard methods known in the art. For example, a naive scFv antibody fragment phage display library may be used to select antibodies that bind to an immunocomplex of analyte and Fab fragments of antibodies that specifically bind the analyte. First the phages are preincubated to sort out those binding to Fab fragments as such. The unbound phages are separated and incubated with a mixture of analyte and immobilized Fab to select the phages that bind to the immunocomplex formed between the immobilized Fab and analyte. Unbound phages are washed away, and then those bound to the complex are eluted. The background is monitored by checking the binding to Fab in the absence of analyte. After several panning rounds a number of clones are picked up, sequenced and expressed resulting in an scFv fr.

[0194] ii. Nucleic Acid Aptamers

[0195] Nucleic acid aptamers are typically oligonucleotides ranging from 15-50 bases in length that fold into defined secondary and tertiary structures, such as stem-loops or G-quartets. The oligonucleotide may be DNA or RNA and may be modified for stability. A nucleic acid aptamer generally has higher specificity and affinity to a target molecule than an antibody. Nucleic acid aptamers preferably bind the target molecule with a K_d less than 10^{-6} , 10^{-8} , 10^{-10} , or 10^{-12} . Nucleic acid aptamers can also bind the target molecule with a very high degree of specificity. It is preferred that the nucleic acid aptamers have a K_d with the target molecule at least 10, 100, 1000, 10,000, or 100,000-fold lower than the K_d with other molecules. In addition, the number of target amino acid residues necessary for aptamer binding may be smaller than that of an antibody.

[0196] Nucleic acid aptamers are typically isolated from complex libraries of synthetic oligonucleotides by an iterative process of adsorption, recovery and reamplification. For example, nucleic acid aptamers may be prepared using the SELEX (Systematic Evolution of Ligands by Exponential Enrichment) method. The SELEX method involves selecting an RNA molecule bound to a target molecule from an RNA pool composed of RNA molecules each having random sequence regions and primer-binding regions at both ends thereof, amplifying the recovered RNA molecule via RT-

PCR, performing transcription using the obtained cDNA molecule as a template, and using the resultant as an RNA pool for the subsequent procedure. Such procedure is repeated several times to several tens of times to select RNA with a stronger ability to bind to a target molecule. The base sequence lengths of the random sequence region and the primer binding region are not particularly limited. In general, the random sequence region contains about 20 to 80 bases and the primer binding region contains about 15 to 40 bases. Specificity to a target molecule may be enhanced by prospectively mixing molecules similar to the target molecule with RNA pools and using a pool containing RNA molecules that did not bind to the molecule of interest. An RNA molecule that was obtained as a final product by such technique is used as an RNA aptamer. Representative examples of how to make and use aptamers to bind a variety of different target molecules can be found in U.S. Patent No. 5,476,766, U.S. Patent No 5,503,978, U.S. Patent No 5,631,146, U.S. Patent No 5,731,424, U.S. Patent No 5,780,228, U.S. Patent No 5,792,613, U.S. Patent No 5,795,721, U.S. Patent No 5,846,713, U.S. Patent No 5,858,660, U.S. Patent No 5,861,254, U.S. Patent No 5,864,026, U.S. Patent No 5,869,641, U.S. Patent No 5,958,691, U.S. Patent No 6,001,988, U.S. Patent No 6,011,020, U.S. Patent No 6,013,443, U.S. Patent No 6,020,130, U.S. Patent No 6,028,186, U.S. Patent No 6,030,776, and U.S. Patent No 6,051,698. An aptamer database containing comprehensive sequence information on aptamers and unnatural ribozymes that have been generated by in vitro selection methods is available at aptamer. icmb. utexas. edu.

[0197] In some embodiments, the nucleic acid aptamer may contain one or more modified nucleic acids (also referred to as xeno nucleic acids, or XNAs) for added chemical functionalities that may increase binding affinity of the nucleic acid aptamer to the immuno-complex. Non-limiting modified nucleic acids include but are not limited to unnatural base pairs (UBPs), base modifications such as for example, C7-modified deaza-adenine, C7-modified deaza-guanosine, C7-modified deaza-cytosine, C7-modified deaza-uridine; and sugar modifications such as for example, ribulonucleic acid, α -L-threose nucleic acid (TNA), 3'-2'phosphonomethyl-threosyl nucleic acid (tPhoNA) and 2'-deoxyxylonucleic acid (dXNA). In some forms, the modified nucleic acid may be introduced in the nucleic acid aptamer by in vitro evolution using an alternative for the phosphodiester backbone such as for example, phosphorothioates, boranophosphate, phosphonate, alkyl phosphonate nucleic acid, and peptide nucleic acid. In some forms, the modified nucleic acid may be introduced in the nucleic acid aptamer via a mutant T7 RNA polymerase that is tolerant of substitutions at the 2' position of the furanose ring. Substitutions that may be attached to C2' include but are not limited to a fluorine, an amine, or a methoxy group. In some forms, the modified nucleic acid may be introduced in the nucleic acid aptamer via R-group mod-

ifications at the 5th position of uracil. In these forms, the R-group can be one of many different sidechains known to those of skill in the art, ranging from hydrophobic to hydrophilic. Incorporation of synthetic nucleotides into nucleic acid aptamers using phosphodiester replacements and modified bases are known to those of skill in the art (See for example, Mayer G. *Angew Chem Int Ed Engl.* (2009) 48: pages 2672–2689; Keefe, A.D. and Cload, S.T. *Curr Opin Chem Biol.* (2008) ; 12: pages 448–456; Appella, D.H. *Curr. Opin. Chem. Biol.* (2009) 13 (5-6) : pages 687-696) .

[0198] iii. Peptide Aptamers

[0199] Peptide aptamers are small peptides with a randomized amino acid sequence that are selected for their ability to bind a target molecule. Peptide aptamer selection can be made using different systems, but the most used is currently the yeast two-hybrid system. Peptide aptamer can also be selected from combinatorial peptide libraries constructed by phage display and other surface display technologies such as mRNA display, ribosome display, bacterial display and yeast display. These experimental procedures are also known as biopannings. Among peptides obtained from biopannings, mimotopes can be considered as a kind of peptide aptamers. All the peptides panned from combinatorial peptide libraries have been stored in a special database with the name MimoDB.

[0200] Assay Fluid

[0201] An aqueous assay fluid can also be introduced to the biological sample, forming a mixed fluid sample. The assay fluid supports a reaction between the analyte and the labelled binding agent (e.g., does not interfere with binding) and has a viscosity that is sufficiently low to allow movement of the assay fluid by capillary action. In some embodiments, the assay fluid contains one or more of the following components: a buffering agent (e.g., phosphate) ; a salt (e.g., NaCl) ; a Protein stabilizer (e.g., bovine serum albumin “BSA” , casein, serum) ; and a detergent such as a non-ionic detergent or a surfactant (e.g., NINATE® 411, ZONYL® FSN 100, AEROSOL OT 100%, GEROPON® T-77, BIO-TERGE® AS-40, STANDAPOL® ES-1, TETRONIC® 1307, SURFNYOL® 465, SURFYNOL® 485, SURFYNOL® 104PG-50, IGEPAL® CA210, TRITON™ X-45, TRITON™ X-100, TRITON™ X305, SILWET® L7600, RHODASURF® ON-870, CREMOPHOR® EL, TWEEN® 20, TWEEN® 80, BRIJ 35, CHEMAL LA-9, PLURONIC® L64, SURFACTANT 10G, SPAN™ 60) . Optionally, if desired, the assay fluid can contain a thickening agent. Representative assay fluids include saline, or 50 mM Tris-HCl, pH 7.2. In some embodiments, the assay fluid is water.

[0202] Preparation of the Lateral Flow Device

- [0203] Disclosed are methods of preparing the lateral flow device to detect one or more analytes from a nasal swab sample for NPC diagnosis and relapse prediction. In one preferred embodiment, the lateral flow device is a multiplex test strip for the detection of NPC analytes e.g., peptide and nucleic acid antigens.
- [0204] Exemplary Test Strip Preparation Method
- [0205] Materials for preparing the disclosed test strips include one or more of a membrane strip e.g., nitrocellulose membrane, a conjugate pad, a sample pad, an absorbent, absorbent pad, backing card, aptamers or antibodies for one or more target antigens, capture particles e.g., colloidal gold particles, blocking agents, buffering solutions, dispenser and laminator.
- [0206] Procedure for Preparing the strip
- [0207] An exemplary procedure for preparing the disclosed lateral flow device such as a test strip includes one or more of the following steps:
- [0208] 1. Preparation of a nitrocellulose membrane, conjugate pad, sample pad, and absorbent pad according to the dimensions required for the lateral flow strip (approximately 5mm (width) × 40mm (length) .
- [0209] 2. Coating the nitrocellulose membrane with one or more capture aptamers or antibodies for one or more desired target antigens. The antibodies used in the NPC assay are generally diluted to an appropriate level, which is generally based on the concentration and affinity of the antibody. For example, the antibodies can be used at a dilution of 1: 50, 1: 100, 1: 200, 1: 300, 1: 400, 1: 500, 1: 750, 1: 1000, 1: 2000, 1: 3000, 1: 4000, 1: 5000, 1: 6000, 1: 7000, 1: 8000, 1: 9000, or 1: 10,000. Preferred dilution for antibodies specific for NPC analytes is 1: 200. Preferred dilutions for secondary antibodies are 1: 7000, 1: 8000, 1: 9000, or 1: 10,000 (range from 1: 7000 to 1: 10000) .
- [0210] 3. Conjugating capture particles such as colloidal gold particles (their amount depending on the amount of antibodies or aptamers on the conjugation pad) with one or more detection aptamers or antibodies for one or more desired target antigens.
- [0211] 4. Applying the conjugated gold particles to the conjugate pad and allow them to dry.
- [0212] 5. Applying a blocking agent e.g., BSA or casein, to the nitrocellulose membrane and/or conjugate pad to prevent non-specific binding.
- [0213] 6. Assembling the lateral flow strip with the sample pad, conjugate pad, nitro-cellulose membrane, and absorbent pad in the stated order.
- [0214] 7. Cutting the assembled test strip to the desired size and shape.
- [0215] Application of Sample to the Lateral Flow Device
- [0216] In one form, prior to the application of the fluid sample to the lateral flow device, a buffer solution, for example, is applied to the sample pad to wet it and to the conjugate pad to active the capture particles e.g., colloidal gold particles.

- [0217] Referring to FIG. 12A, the sample is applied to the application point 14 of the membrane strip, or to the application pad, if present. After the membrane strip is contacted with the sample, the membrane strip is maintained under conditions (e.g., sufficient time and fluid volume) which allow the labeled binding agents to move by capillary action along the membrane to and through the capture zone 18 and subsequently beyond the capture zones 18 (e.g., into a wicking pad), thereby removing any non-bound labeled binding agents from the capture zones. In some embodiments, the sample migrates through the conjugate zone containing binding agents. The analyte in the sample interacts with the binding agents to form capture complexes.
- [0218] As the applied sample passed through the membrane strip, analyte bound (sample/control analyte) to binding agent (capture complex) are immobilized by capture agents in the capture zone 18, which are preferably conjugated to immobilized capture particles. The capture zone 18 is preferably organized into one or more capture lines in specific areas of the capture zone where they serve to capture the capture complexes as they migrate by the capture lines. The capture zone 18 preferably contains a plurality of capture lines 22 for multiplex analysis and quantification.
- [0219] Capillary action subsequently moves any binding agents that have not been arrested onwards beyond the capture zone 18, for example, into a wicking pad which follows the capture 18 zone. If desired, a secondary wash step can be used. Assay fluid can be applied at the application point after the mixed fluid sample has soaked into the membrane or into the application pad, if present. The secondary wash step can be used at any time thereafter, provided that it does not dilute the mixed fluid sample. A secondary wash step can contribute to reduction of background signal when the capture particles are detected.
- [0220] Detection of Analyte and Interpretation of Results
- [0221] The amount of analyte bound by binding agents arrested in the capture zone (sandwich complex) may then be detected. The labeled binding or capture agents are preferably detected using an appropriate means for the type of label used. In some forms, the appearance of lines and/or color changes on the nitrocellulose membrane indicates the presence or absence of the desired analyte.
- [0222] The amount of analyte in the sample is directly related to the level of detection agent detected in a capture line. This value is preferably normalized by the amount of another detectable label immobilized within the membrane (e.g., capture zone) to account for variations in detection device and parameters (e.g., light intensity). This normalized value may then be plotted against a standard curve or response surface that correlates these normalized values to analyte concentration. For example, a standard curve or response surface may be prepared in advance using analyte standards. In

addition, three or more internal standard analytes may be detected in the assay and used to adjust or select the standard curve or surface from reference curves or surfaces.

[0223] In one preferred form, the test results, indicated by the intensity and/or contrast of visible colors, reflects the strength of detection of different NPC analytes. For example, in some forms, the test result outputs on the test strip are “No NPC” , “pre-malignant NPC” and/or “malignant NPC” . In some forms, the test results can indicate the chance of relapse upon NPC diagnosis, with output options e.g., “low” and/or “high” .

[0224] IV. Sample Collection Apparatus

[0225] The quantitative point-of-care assay may involve the use of a sample collection apparatus that is not in fluid contact with the solid phase apparatus. The sample collection apparatus can be any apparatus which can contain binding agents and to which a measured volume of fluid sample can be added. Representative sample collection apparatus includes a sample tube, a test tube, a vial, a pipette or pipette tip, or a syringe.

[0226] Sample Collection Swab

[0227] In one preferred embodiment, the sample collection apparatus is a nasopharyngeal swab. FIGs. 11A-11I illustrate non-invasive rapid test sampling using next-generation 3D-printed nasopharyngeal swab, sample processing tube and customized panels of probes at rapid antigen test (RAT) for NPC diagnosis and relapse prediction conducted at home, community clinic, hospital or medical laboratory. FIG. 11A shows a 3D-printed nasopharyngeal swab tailor-made for high yield and least discomforting cellular and interstitial sample collection. FIG. 11B is a magnified side view of the swab tip design with dimensions. FIG. 11C is a magnified top view of the swab tip design with dimensions. FIGs. 11D-G shows a design of next-generation rapid diagnostic test station/hub including sample collection/distribution structure which maximize detection of biomarkers with consistency and minimizes sample input for detection of more than one biomarker panel. Slices of detection strips can be customized for personalized disease and prognostic detection. For example, inflammation (inflam) and drug resistance (DR) can be simultaneously detected in one strip. FIG. 11D is showing a view of all RAT cassettes inserted to sampling hub. FIG. 11E is just another view showing the assembly or structure of RAT cassettes and before and after insertion to sampling hub. FIG. 11F is showing the top view (cut in the middle) of a RAT cassette inserted to sampling hub. FIG. 11G shows the side view (cut in the middle) of a RAT cassette inserted to sampling hub. FIG. 11H is a next-generation sample processing tube including brushes/bristles therein for more efficient release of biomaterials from swab head. FIG. 11I. Sample processing tube with short hairs inside specially designed for releasing materials from (nasopharyngeal) swab.

- [0228] The disclosed dimensions in FIGs. 11B, 11C, 11F and 11G are not limiting as one can readily adjust the various dimensions either above or below the stated value in a range of approx. +/-10%.
- [0229] As shown in FIG. 11B, the device includes a “head” for insertion into the nasal cavity, which is connected to a handle via a series of preferably cylindrical “poles” , of varying diameters. 3D printing is an additive manufacturing technique that creates three-dimensional objects by building successive layers of raw material such as metals, plastics, and ceramics. The objects are produced from a digital file, rendered from a magnetic resonance image (MRI) or a computer-aided design (CAD) drawing, which allows the manufacturer to easily make changes or adapt the product as desired. 3D printing approaches can differ in terms of how the layers are deposited and in the type of materials used. A variety of 3D printers are available on the market, ranging from inexpensive models aimed at consumers and capable of printing small, simple parts, to commercial grade printers that produce significantly larger and more complex products.
- [0230] The device is preferably made from an autoclavable medical grade high modulus photo polymer resin, which are known in the art and are reviewed for example, in Gutteridge, et al., Annals of 3D Printed Medicine Volume 5 , 100044 (<https://doi.org/10.1016/j.stlm.2021.100044>) , Tables, 3 and 4, therein.
- [0231] The head portion of the device is preferably generally cylindrical in shape it includes “hairs” –hollow cylindrical projections from the body of the head, for collecting cells and secretions. The head portion is hollow to collect cells scratched off and secretion, and allow materials collected to be released later for detection at a test strip. The head portion can measure about 3mm in diameter and about 12-15 mm in length (without taking into account the hairs) or if the hairs are taken to account, from about 5mm in outer diameter/width and 14-17 mm in length (See FIGs. 11B and 11C) .
- [0232] Hairs on the head of the swab hollow cylindrical projections, are used to collect cells scratched off and nasopharyngeal secretion, and allow materials collected to be released later for detection, for example, at a test strip. The hollow cylindrical hairs are located throughout the head and at the tip of swab to maximize materials (nasopharyngeal secretion and cells) and collection. Each hair is a hollow cylinder with an outer radius (R) , an inner radius (r) , a height and a thickness (R-r) . The hollow cylindrical hairs are about 0.3 mm thick, about 2 mm in outer diameter and about and about 1.3 mm high. The hollow cylindrical hairs can be located throughout the outer surface of the head of the device, about 0.75 mm apart.
- [0233] Dome-shaped cross wires (about 0.1 mm thick) on the cylindrical projections are used to gently scratch cells off, press against tissue and squeeze secretion out. Dome-shaped cross wires project about 0.2 mm from the plane of the hairs, thus, the length

from the top of the dome to the base of the hollow cylindrical hair (i.e., the portion resting on the outer surface of the head) is about 1.5 mm.

- [0234] Connecting the head and handle of the device are a first pole (connected to the head) of about 1 mm in diameter, a second/middle pole of about 2mm in diameter and which is hollow and is configured for extension or retraction of the device pushing the second pole over the first pole (retraction) or away from the second pole (extension) and a third pole, which is about 1 mm in diameter. The first pole can be about 20-50 mm in length, the second/middle pole can be about 20-50 mm in length and the third pole can be about 10-30 mm in length.
- [0235] The handle of the device is in some forms, turned to a hollow bulb. Upon squeezing, materials can be forced out of the swab. By contrast, upon releasing, materials can be sucked into the swab. In some forms, the hollow bulb is made from a suitable plastic material.
- [0236] Sample Collection tube
- [0237] A special tube for harvesting biomaterials collected by nasopharyngeal swab is provided. In some forms the tube is a microcentrifuge tube, with brushes/bristles as disclosed herein. Microcentrifuge tubes are small conical tubes. The top opening of the tube generally has a lid which is connected to the top of the tube by means of a plastic strip or hinge. These tubes are widely used by molecular biologists and biochemists. A microcentrifuge tube is described in U.S. Patent No. 5254314, the contents of which are herein incorporated by reference.
- [0238] The sample collection tube includes a tube having a sealed bottom end, an open top end and a lid which seals the top opening of the tube, with improvements to include flexible hair/bristle brushes.
- [0239] This specially designed plastic tube with flexible hair/bristle brushes on the interior wall can enhance the efficiency and sensitivity of diagnostic tests for conditions like nasopharyngeal carcinoma, where optimal sample collection is crucial for early detection and effective treatment.
- [0240] Design:
- [0241] 1. The tube is made of plastic material, typically polypropylene or a similar durable plastic such as listed above for the swab.
- [0242] 2. The interior wall of the tube is lined with multiple rows of soft, flexible hair-like bristles or brushes. In some forms, the tube has an inner diameter of about 9mm (exclusive of brushes) and an outer diameter of about 11 mm. Long hair (3.5mm in length) results in an open inner diameter of 2mm for insertion of swab head, whereas short hair (1.5mm in length) results in an open inner diameter of 6mm for insertion of swab head. Depth of the tube is around 40mm.

- [0243] 3. The bristles are arranged in a spiral or helical pattern along the length of the tube's interior, creating a textured surface. In some forms, the bristles are cylindrical in shape with a spherical tip long. In some forms, the bristles are "long bristles" ranging in length from about 1.5-5 mm, preferably from about 2 to about 3.5 mm. In this form the bristles tips can be separated by a distance of about 2-6 mm. Thus for example, when the bristle lengths are about 3.5 mm, the tips of the bristles are separated by a distance of about 2 mm. (FIG. 11H) . The bristles can have a diameter of about 0.5 mm and the space between bristles is about 1-2 mm. In some forms, the bristles are "short" bristles ranging in length from about 0.5-2.5 mm, preferably from about 1 to about 1.5 mm. In this form the bristles tips can be separated by a distance of about 2-6 mm (FIG. 11I) .
- [0244] 4. The bristles are made of a synthetic, medical-grade material (such as polypropylene (PP)) that is gentle yet effective in dislodging and capturing cellular and other biological materials.
- [0245] Functionality:
- [0246] 1. When the nasopharyngeal swab is inserted into the tube, the bristles on the interior wall make contact with the swab tip.
- [0247] 2. As the swab is rotated or gently moved back and forth within the tube, the bristles lightly scratch and dislodge the collected biomaterials, including cells, mucus, and other relevant analytes, from the swab surface.
- [0248] 3. The dislodged biomaterials are then suspended in the liquid medium within the tube, creating a more concentrated sample for downstream diagnostic testing.
- [0249] 4. The spiral or helical arrangement of the bristles helps to ensure that the entire surface area of the swab tip is effectively scratched and sampled.
- [0250] Benefits:
- [0251] 1. Increased sample yield: The flexible bristles effectively harvest a higher amount of biomaterials from the nasopharyngeal swab, leading to improved sensitivity and accuracy in diagnostic tests, such as nucleic acid amplification tests or cytological analysis.
- [0252] 2. Consistent sample collection: The standardized tube design and the controlled brushing action ensure more reproducible and reliable sample collection compared to traditional swab-only methods.
- [0253] 3. Reduced sample loss: The contained nature of the tube and the capturing of bio-materials by the bristles minimize the loss of sample material during the collection and transfer process.
- [0254] 4. Ease of use: The simple tube design allows for straightforward sample collection and processing, requiring minimal specialized training for healthcare providers.
- [0255] An exemplary procedure for collecting biological samples using nasopharyngeal swabs and processing the samples for NPC diagnosis and relapse prediction using

a lateral flow assay as disclosed herein. In this example, the lateral flow assay has two lines for NPC diagnosis and relapse prediction respectively. One or more of the following materials can be used to collect and process the biological samples: a nasopharyngeal swab e.g., a 3D-printed nasopharyngeal swab as disclosed above for tumor sample collection, a lateral flow device e.g., a rapid test (with two lines for NPC diagnosis and relapse prediction, sample buffer, and a dropper) .

- [0256] Collection and processing of the biological samples includes one or more of the following steps:
- [0257] 1. Collecting a nasopharyngeal sample from a patient. Methods of collecting nasopharyngeal samples are known and may include one or more of the following steps: (i) tilting the patient's back about 70 degrees, (ii) gently rotating the nasopharyngeal swab and inserting it less than one inch (about 2 cm) into nostril parallel to the palate (not upwards) until resistance is met at turbinates, (iii) rotating the swab several times against the nasal wall and (iv) repeating (i) to (iii) in the other nostril using the same swab.
- [0258] 2. Inserting the nasopharyngeal swab into a tube of sample buffer, and swirl the swab to ensure that it is fully immersed in the buffer.
- [0259] 3. Placing a few drops of the sample buffer onto the sample well of the lateral flow device.
- [0260] 4. Waiting for the control line to appear, which confirms that the device is working properly.
- [0261] 5. After the control line appears, waiting for a few minutes for the test analyte to be detected, if present. In some forms, the wait time can be up to 15 minutes.
- [0262] 6. Assessing the intensity of the colour on the two lines to interpret the test results. The results are presented in terms of intensity and contrast of visible colours will reflect the strength of detection of different biomarkers, indicating the degree of likelihood of suffering from NPC disease. In some forms, a positive result for NPC diagnosis and relapse prediction can be concluded if both lines appear with high intensity, i.e., a heavier or a darker line. When the line indicator for NPC diagnosis appears with high intensity, a positive result for NPC diagnosis only can be concluded. A negative result can be interpreted if neither line appears with high intensity. In some forms, the intensity of each line may vary depending on the amount and types of biomarkers detected, with higher intensity indicating a higher likelihood of NPC diagnosis or relapse.
- [0263] V. Kits
- [0264] Kits for diagnosis of nasopharyngeal carcinoma and/or monitoring the potential for nasopharyngeal carcinoma relapse are described. In particular, kits for detecting one or more target analytes that are biomarkers for NPC occurrence and relapse are provided.

In one embodiment, the kit includes the lateral flow device as disclosed herein. The kit optionally contains a sample collection apparatus such as a nasopharyngeal swab, sample buffer, and a dropper.

[0265] Kit components additionally can include analytes at known concentrations for generating a standard curve, capture particles, particles, and conjugation buffer for coating particles with binding agents, disposal apparatus (e.g., biohazard waste bags), and/or other information or instructions regarding the sample collection apparatus (e.g., lot information, expiration date, etc.). The kits contain some or all of the materials needed to measure one or more of the following analytes: HMGN2P3 (high mobility group nucleosomal binding domain 2 pseudogene 3); DNAJC11 (DnaJ Heat Shock Protein Family (Hsp40) Member C11); EIF2AK1 (Eukaryotic Translation Initiation Factor 2 Alpha Kinase 1); FAM234A (Family With Sequence Similarity 234 Member A); PARPBP (PARP1 Binding Protein); ARL5A (ADP Ribosylation Factor Like GTPase 5A); IL32 (interleukin 32); and DHX57 (DEXH-box helicase 57), PSMA4 (proteasome 20S subunit alpha 4), CALML3 (calmodulin like 3), SLC2A1 (solute carrier family 2 member 1), SNX3 (sorting nexin 3), LY6D (lymphocyte antigen 6 family member D), YBX1 (Y box binding Protein 1), RPMS1 (an open reading frame of BamHI-A rightward transcripts of Epstein-Barr virus).

[0266] In some forms, the kit contains a test strip that gives a positive reading only when the one or more target analytes are detected. Readout of the test strip would allow the clinician to have a sensitivity and specificity to determine whether the patient has nasopharyngeal carcinoma and/or the likelihood of relapse. In some forms, the kits can give a positive reading for a single biomarker or a positive reading for multiple biomarkers, e.g., a multiplex test strip.

[0267] VI. Computer-implemented systems and methods

[0268] Also described are CISs and/or CIMs containing one or more AI platforms for analyzing biological data and outputting the occurrence of a cancer, such as NPCR/NPCD. The analysis is based on the expression levels of certain biomarkers or combinations of biomarkers in the biological data. The biomarkers or combinations thereof are indicative of certain cell types. Preferably, the AI platform utilizes a “signature matrix” that has numerical entries of expression levels of these biomarkers amongst the cell types. In some forms, the biomarkers are genes. For example, columns in the signature matrix track cell types and rows track biomarkers (e.g., genes). Hence, the number at each *i*th-row and *j*th-column represents the relative level of the gene expression in the cell type. The AI platform assesses biological data (preferably from a subject’s test results) and makes a prediction about the occurrence of NPCR/NPCD based on comparisons with levels of biomarkers or combinations thereof in the signature matrix. Exemplary biomarkers (predictive of relapse) are cytometric (cell

type) biomarkers, specifically, markers of the LY6D+ neoplastic SPB1/SPB epithelial subpopulations. As an example, in clinical settings, if the gene expression data in the biological data contain more neoplastic SPB1 features when compared to the signature matrix, i.e., the gene expression is similar to the signature matrix, it indicates more or presence of neoplastic SPB1 in the biological data, and/or if the gene expression data contain less non-neoplastic SPB1, it is likely the subject would have NPCR. This procedure can be done by uploading the signature matrix and biological data (e.g., subject sample gene expression data) to the AI platform. The AI platform analyzes the data and provides a prediction. The prediction is provided via a visual format (e.g., graphical user interface), an audio-format (e.g., via an audio signal that reports the prediction), or a combination thereof. The one or more AI platforms have been trained and validated using data involving gene expression levels of these biomarkers, associated cell types, and/or the occurrence of NPCR/NPCD.

[0269] Also described is an AI platform that uses a signature matrix and biological data (e.g., a subject's test results) to predict the occurrence of a cancer, such as NPCR/NPCD, wherein the AI platform is operably linked to a computer processor, wherein the AI platform is configured to analyze data in the signature matrix and the biological data (e.g., a subject's test results) to make the prediction. The analysis is based on the expression levels of certain biomarkers or combinations of biomarkers in the biological data. The biomarkers or combinations thereof are indicative of certain cell types. Preferably, the AI platform utilizes a "signature matrix" that has numerical entries of expression levels of these biomarkers amongst the cell types. In some forms, the biomarkers are genes. For example, columns in the signature matrix track cell types and rows track biomarkers (e.g., genes). Hence, the number at each *i*th-row and *j*th-column represents the relative level of the gene expression in the cell type. The AI platform assesses biological data (preferably from a subject's test results) and makes a prediction about the occurrence of NPCR/NPCD based on comparisons with levels of biomarkers or combinations thereof in the signature matrix. Exemplary biomarkers (predictive of relapse) are cytometric (cell type) biomarkers, specifically, markers of the LY6D+neoplastic SPB1/SPB epithelial subpopulations.

[0270] Also described is a non-transitory computer-readable medium with computer executable instructions stored thereon executed by a processor to perform a method of predicting the occurrence of a cancer, such as NPCR/NPCD. The method involves (i) uploading a signature matrix and biological data (e.g., a subject's test results) to a computing device, and (ii) using an AI platform to analyze the data and transmit the results to a human, wherein the analysis is based on the expression levels of certain biomarkers or combinations of biomarkers in the biological data. The biomarkers or combinations thereof are indicative of certain cell types. Preferably, the AI platform

utilizes a “signature matrix” that has numerical entries of expression levels of these biomarkers amongst the cell types. In some forms, the biomarkers are genes. For example, columns in the signature matrix track cell types and rows track biomarkers (e.g., genes) . Hence, the number at each i th-row and j th-column represents the relative level of the gene expression in the cell type. The AI platform assesses biological data (preferably from a subject’s test results) and makes a prediction about the occurrence of NPCR/NPCD based on comparisons with levels of biomarkers or combinations thereof in the signature matrix. Exemplary biomarkers (predictive of relapse) are cytometric (cell type) biomarkers, specifically, markers of the LY6D+ neoplastic SPB1/SPB epithelial subpopulations.

[0271] The machine learning procedures may involve various supervised machine learning techniques, various semi-supervised machine learning techniques, and/or various unsupervised machine learning techniques. For instance, the machine learning procedures may utilize Logistic Regression, Gaussian Naïve Bayes, Random Forest, Gradient boosting, Adaptive Boosting, LPBoost, TotalBoost, BrownBoost, MadaBoost, LogitBoost, Extra Trees, Linear Discriminant Analysis, Support Vector Machines, Decision Tree, k-nearest neighbor, alternating decision trees (ADTree) , Decision Stumps, functional trees (FT) , logistic model trees (LMT) , linear classifiers, factor analysis, principal component analysis, neighborhood component analysis, sparse filtering, stochastic neighbor embedding, autoencoders, stacked autoencoders, neural networks, convolutional neural networks, feed forward neural network, Tabular Attention Network, or any other machine learning algorithm or statistical algorithm. In some forms, the machine learning procedures include, but not limited to, the use of support vector regression (SVR) , linear least-square regression (LLSR) , microarray microdissection with analysis of differences (MMAD) and digital sorting algorithm (DSA) . Machine learning analyses may be performed using one or more of various programming languages and platforms, such as R, Weka, Python, and/or Matlab, for example. Machine learning analyses may be performed using a machine learning platform, such as BigML.

[0272] Several computational tools, including linear least-square regression (LLSR) , microarray microdissection with analysis of differences (MMAD) , and digital sorting algorithm (DSA) , have been applied to the deconvolution of complex gene expression programming (GEP) mixtures to infer cellular composition. Although these approaches are effective for enumerating highly distinct cell types in mixtures with minimal unknown content (e.g., lymphocytes, monocytes, and neutrophils in whole blood) , they are sensitive to experimental noise, high unknown mixture content, and closely related cell types, limiting their utility for TIL assessment. CIBERSORT, a computational approach, aims to address these challenges. Like other methods,

CIBERSORT requires a specialized knowledgebase of gene expression signatures, termed a “signature matrix,” for the deconvolution of cell types of interest. However, in contrast to previous efforts, CIBERSORT implements a machine learning approach, called support vector regression (SVR), that improves deconvolution performance through a combination of feature selection and robust mathematical optimization techniques. In benchmarking experiments, CIBERSORT was more accurate than other methods in resolving closely related cell subsets and in mixtures with unknown cell types (e.g., solid tissues). Thus, CIBERSORT is a useful approach for high throughput characterization of diverse cell types, such as TILs, from complex tissues. The instant disclosure provides users with a practical roadmap for dissecting leukocyte content in tumor gene expression datasets with CIBERSORT.

[0273] Bar charts were generated for relapse prediction performance of matched biomarkers from both more costly single-cell RNA sequencing data (and data obtained from AI-powered platform that carries out deconvolution and cell fraction estimation from less costly bulk-sample expression data (data not shown)). Performance of both methods were found to be comparable and outperform existing golden standard plasma EBV DNA test especially in terms of accuracy and positive predictive value.

[0274] FIG. 13 is a flow chart of computer-implemented systems (CISs) and/or methods (CIMs) containing one or more discriminative artificial intelligence (AI) platforms for analyzing biological data using a signature matrix and outputting the occurrence of NPCR/NPCD based on the expression levels of certain biomarkers or combinations of biomarkers in the biological data, which are indicative of certain cell types.

[0275] The disclosed methods can further include providing the diagnosis result and prescribing one or more treatments for the subject if the presence and/or increased levels of one or more of the biomarkers listed above is detected. The methods can include further testing, such as a biopsy.

[0276] The presence of these specific biomarkers (DNA, RNA and proteins) in nasopharyngeal carcinoma which correlate with its diagnosis and disease relapse shall lead to an alternative treatment strategy including radical radiotherapy of not less than 70 Gy delivered over 6 to 7 weeks, with or without, chemotherapy in induction, concurrent and/or adjuvant setting with reference to radiotherapy, immunotherapy including but not limited to use of immune checkpoint inhibitors in induction, concurrent and/or adjuvant setting with reference to radiotherapy, and vaccination and/or targeted therapy against the disease relapse biomarkers in induction, concurrent and/or adjuvant setting with reference to radiotherapy. The absence of biomarkers correlating with disease relapse but with the presence of biomarkers correlating with diagnosis of nasopharyngeal carcinoma shall (1), for treatment, follow the current standard treatment with radical radiotherapy, with or without, chemotherapy

in induction, concurrent and/or adjuvant setting, and immunotherapy including but not limited to use of immune checkpoint inhibitors in induction, concurrent and/or adjuvant setting, or (2) , for prevention, lead to preventive measures including regular monitoring of the level of diagnosis biomarkers, and with or without vaccination and/or targeted therapy against the diagnosis biomarkers.

[0277] Treatments are known in the art, and include, but are not limited to radiation therapy or chemotherapy. Chemotherapeutic agents commonly used to treat NPC include, but are not limited to Carboplatin (Paraplatin) , Epirubicin (Ellence) , Paclitaxel (Taxol) , Docetaxel (Taxotere) , Gemcitabine (Gemzar) , Capecitabine (Xeloda) and Methotrexate. A chemotherapeutic drug may be used alone or combined with other drugs. Thus, the disclosed methods include determining the presence of one or more mutations or expression biomarkers as disclosed therein and administering one or more agents for treating cancer, to the subject. The method can further include guiding treatment regimens as to whether to initiate, continue, or discontinue treatment of nasopharyngeal carcinoma in a subject.

[0278] The disclosed methods, compositions and devices can be further understood by way of the following non-limiting examples.

[0279] EXAMPLES

[0280] Methods

[0281] Patient recruitment and biopsy collection

[0282] The study was approved by the institutional review board of The University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW 19-157) . All mentioned datasets/studies were carried out in accordance with ethical guidelines and all patients provided written informed consent. Seventy-four patients diagnosed histologically with NPC underwent nasoendoscopy and biopsy of the primary tumor and ultrasonography-guided fine needle aspiration of neck lymph nodes (if clinically palpable and locatable by ultrasonography) . A total of 155 biopsies from multiple sites including normal adjacent tumor, primary tumor at nasopharynx and tumor at cervical lymph node (if any) were collected and immediately processed for single-cell isolation and transcriptome sequencing.

[0283] Single-cell RNA sequencing

[0284] 84187 epithelial cells had been successfully profiled. After tissue digestion, single cell suspension went through cell counting and viability assessment using Countess II FL Automated Cell Counter. Single cell encapsulation and cDNA libraries were prepared by Chromium Next GEM Single Cell V (D) J Reagent Kit v1.1 and Chromium Next GEM Chip G Single Cell Kit according to 10X Genomics protocol. Cells in suspension are counted and loaded into individual wells of 10X Chromium Single Cell chip. Single cells are then encapsulated into Gel Beads-in-emulsion (GEM)

by 10X Chromium Single Cell Controller. Single Cell V (D) J Reagent Kit v1 or v1.1 is used to perform downstream steps according to manufacturer's protocol. Library size and concentration are determined by Qubit, quantitative PCR and Bioanalyzer assays. Illumina sequencing (Pair-End sequencing of 151bp) were carried out at Centre for PanorOmic Sciences (CPOS), Genomics Core, LKS Faculty of Medicine, The University of Hong Kong.

[0285] Bioinformatic Analysis

[0286] Quality check (QC) filters ($200 \leq \text{nFeature_RNA} \leq 9000$ & $\text{percent. mt} \leq 50$) were applied. Scrublet doublet analysis of cells from each sample was adopted for removing doublets. Here, we followed the suggested workflow written in the author's tutorial pages, with doublet rates set at 5% of recovered cells. FastMNN was used for data integration and batch effect removal. Unsupervised clustering was adopted for cell type classification. CNVs were inferred by inferCNV by using macrophages as reference cells while SNVs were called by our modified Mutect2 pipeline. Kaplan-Meier (KM) curve, log rank test, Cox regression were adopted for progression-free survival analysis to identify abundance association with relapse. Deconvolution analysis of GSE102349, a publicly available bulk RNA sequencing dataset with relapse information using CIBERSORTX, was adopted to verify relapse prediction of our selected biomarkers from relapse-associated neoplastic subcluster.

[0287] Immunohistochemistry

[0288] NPC patient tissue was formalin-fixed, paraffin-embedded, and sectioned at 3 μm . Paraffin sections were deparaffinized in xylene and rehydrated through a gradient of ethanol. Before antibody staining, antigen retrieval was performed in 10mM citrate buffer (pH 6.0) using Target Retrieval Solution (Dako). Sections were mounted using DPX Mountant (Sigma). Stained sections were imaged with a Nikon Model Eclipse Ni-U Microscope (Nikon).

[0289] In-situ hybridization for Epstein-Barr virus encoded RNA

[0290] The formalin-fixed paraffin-embedded tumor slides were all tested with in-situ hybridization and all patients' tumors tested positive with Epstein-Barr virus encoded RNA (EBER). The relapsed tumor specimens (if obtained) of 3 patients with local recurrence also tested positive with EBER.

[0291] Droplet Digital Polymerase Chain Reaction

[0292] DNA was extracted from brush biopsy from nasopharynx using Qiagen DNeasy Blood and Tissue Kit. The concentration of DNA was measured by the SMA4000 and the purity was evaluated through the measurement of the OD260/OD280 ratio. Extracted DNA was stored at -20°C until it was used. Digital droplet PCR (ddPCR) assays were performed on QX200 AutoDG Droplet Digital PCR system (Bio-Rad). Genomic mutations were detected by customized ddPCR probe kit according to the

manufacturer's protocol. The PCR program of ddPCR after droplet generation was as follows: 95°C for 10 min; 40 cycles of 94°C for 15 s and 58°C for 60 s; 98°C for 10 min; 4°C for 5 min. The reaction temperature was changed at a rate of 2°C/s.

[0293] Enzyme-Linked Immunosorbent Assay

[0294] 100 μ l of sample was added in duplicate to wells of a Dynex Immulon 4HBX 96-well plate. After overnight incubation at 4 °C, the plate was centrifuged for 10 min at 1400 rpm. The supernatant was discarded and the wells were dried for 5 min with a hair dryer and then fixed for 10 min at room temperature with methanol. After the methanol was discarded, plates were either used immediately or wrapped in aluminum foil and stored at -20 °C. For probing, plates were washed 4 $\times$ with PBS containing 0.05% Tween 20 and then blocked with PBS containing 1% human serum albumin (HSA) for 1–2 h at room temperature. After washing once, primary antibody was added. After incubating 1 h at room temperature, plates were washed 4 $\times$ with PBS/ Tween and rabbit anti-mouse immunoglobins conjugated to biotin (Dako, cat# #0354) was added diluted 1: 1000I PBS/HSA/Tween. Plates were incubated 1 h at room temperature. Then they were washed and ExtrAvidin peroxidase (Sigma, cat #E-2886) was added diluted 1: 1000 in PBS/HSA/Tween. After 45 min at room temperature, plates were washed and TMB (Sigma) was added. After 5–10 min, color development was stopped by the addition of HCl.

[0295] Statistical Analysis

[0296] Data were expressed as mean $\pm$ SEM. Independent sample t test/Wilcoxon rank sum test and ANOVA/Kruskal-Wallis were used to compare differences of parametric and non-parametric continuous variables. Pearson/Spearman correlation and Chi-square tests were used to evaluation association between variables. Binary logistic regression with univariate and multivariate analyses were performed to identify predictive factors of NPC recurrence. Kaplan-Meier methods were performed to evaluate survival endpoints of recurrence-free survival (RFS) and overall survival (OS) . Restricted mean survival time (RMST) was also performed to provide survival estimation as well. Log-rank tests were performed to compare survival difference in subgroup analyses. Cox proportional hazard models with univariable and multivariable analyses were performed to identify prognostic factors of RFS and OS. All statistical analyses were analyzed either by Statistical Package for Social Sciences (SPSS) version 25 (IBM, USA) or R programming. P values <0.05 (two-sided) were considered statistically significant.

[0297] Results

[0298] The disclosure herein is based on a large-scale NPC single-cell RNA sequencing (scRNA-seq) .

[0299] FIGs. 1A-1B are the schematic diagrams demonstrating steps in discovering novel biomarkers from a single-cell RNA sequencing (scRNA-seq) NPC study for disease diagnosis and relapse prediction using our proposed tests. 74 patients were recruited for the scRNA-seq study. Matched biopsies were harvested from normal adjacent tumor, primary tumor at nasopharynx and tumor from neck lymph nodes. Patients were followed for 3 years in median. 5 patients relapsed. Bioinformatic analyses were conducted to identify human-based DNA mutations, RNA/protein expression, and cytometric biomarkers that were associated with NPC malignancy and NPC relapse at the time of initial diagnosis. As shown in FIG. 1B, microbe EBV RNA transcripts were also detected and analyzed. Biomarkers with highest neoplasticity, or NPCD, and/or NPCR accuracy were identified. FIG. 1C is a schematic diagram showing proposed sampling methods and detection strategies utilizing our (1) human-based DNA, RNA/protein biomarker panel, (2) EBV-based panel and (3) cytometric panel for both (1) non-invasive rapid test with biological samples from nasopharyngeal swab or plasma, or (2) conventional invasive endoscopy biopsy. FIG. 1D is a schematic diagram showing that the epithelial subpopulation (i.e. cytometric) biomarker can applied to AI-powered platform to analyze expression data (e.g. transcriptomics data including RNA sequencing and microarray, proteomics data including mass spectrometry) to calculate the two risk scores (i.e. risk score 1 = neoplastic SPB1 abundance, and risk score 2 = difference between neoplastic SPB1 and non-neoplastic SPB1) to predict relapse at initial diagnosis before treatment and provide information for follow-ups early. FIGs. 1E-1G are diagrams showing a summary of all biomarkers discovered for NPCD and/or NPCR.

Rule 91,
06.01.2025

[0300] FIGs. 2A-2B show mutations (i.e. SNVs and InDels) uniquely found in neoplastic cells from all possible malignant sites that predict relapse at initial diagnosis before treatment (i.e. NPCR) . FIG. 2A shows top 8 somatic mutations uniquely identified in neoplastic cells that were found strongly associated with relapse even at time of initial diagnosis. FIG. 2B shows performance metrics of the top 8 relapse-predicting somatic mutations, Receiver operating characteristic curve (ROC) were generated with area under ROC curve (AUC) of the top 2 relapse-predicting somatic mutations showing an AUC of 82.9% (IL32) and 79.5% (DHX57) . Kaplan-Meier (KM) curve of the top 2 relapse-predicting somatic mutations were generated (data not shown) . Relapse prediction performance could be further boosted by combining IL32+DHX57 SNV, attaining a sensitivity of 100%, specificity of 84.2%and Accuracy of 86%; its Receiver operating characteristic curve (ROC) showed an area under ROC curve (AUC) of 92.1%.

- [0301] FIGs. 3A-3C shows novel mutations (i.e. SNVs and InDels) with high occurrence in NPC patients at initial diagnosis. FIG. 3A shows top 3 somatic mutations of neoplastic cells identified in NPC patients. EMP2 mutation was also found to have high coverage in EBV-negative NPC patients. FIG. 3B shows performance metrics of EMP2 mutation alone in NPC diagnosis. FIG. 3C shows that the disclosed mutation combo panel alone could achieve NPC diagnosis with sensitivity of 93.0%, which is significantly higher than golden standard plasma EBV DNA copy number alone. If mutation combo was used together with EBV plasma EBV DNA copy number, sensitivity could be further boosted up to 95.3%.
- [0302] FIGs. 4A-4E shows Lymphocyte Antigen 6 Family Member D (LY6D) -positive neoplastic secretory-primed basal cells (SPB) , in particular SPB1 (secretory primed basal cell type cluster 1, referred to hereinafter as SPB1) as a unique and novel neoplastic subpopulation that is consistently associated with and predicts relapse with high accuracy. FIG. 4A is a UMAP showing a total of 30 different epithelial subpopulations discovered by unsupervised clustering from the scRNA-seq data with known epithelial canonical markers. FIG. 4B is a violin plot showing that SPB1 is a novel epithelial subpopulation originating from SPB discovered by our scRNA-seq analysis. LY6D was found to be the biomarkers of SPB and in particular SPB1. Neoplastic SPB1 was found to have the highest LY6D expression while the remaining non-SPB1 subpopulations were found to have much lower LY6D expression. Statistical analysis showing difference in abundance of some neoplastic subpopulations between relapse and no relapse, without consideration of relapse time. FIG. 4C is a table showing that neoplastic SPB1, SPB2 and SPB5 were found to be the top neoplastic subpopulation that are consistently associated with relapse, as revealed by Wilcoxon rank sum test. Statistical analysis showing difference in abundance of some neoplastic subpopulations between relapse and no relapse, with consideration of relapse time. FIG. 4D shows log rank test that supported neoplastic SPB1, SPB2 and SPB5 as relapse-related. In particular, Cox regression analysis showed neoplastic SPB1 as the only neoplastic subpopulation that can predict relapse. KM plot also showed presence of SPB1 is associated with lower relapse-free survival probability. FIG. 4E is a performance analysis showed AUC of relapse prediction by neoplastic SPB1 to be 0.773 with an accuracy reaching 81%. The graph shows AUC, a way of showing performance without setting particular a threshold of detection, while the table is showing performance metrics including sensitivity, accuracy and specificity at a particular threshold of detection.
- [0303] Identification of gene mutations present in LY6D-positive neoplastic SPB1 that associate with relapse revealed specific mutations) carried by LY6D+ neoplastic SPB1, as well as epithelial cells in general, found to be more predominant in relapse

patients. The HDAC2 mutation (HDAC2 (chr6: 113, 970, 875) [CT>C, CTT, CTTT, CTTTT, CTTTTT] InDel) was detected in neoplastic SPB1 detected at NP tumor and lymph node in as shown by scRNA-seq data. HDAC2-mutated neoplastic SPB1 at nasopharyngeal tumor was found to be highly predictive of relapse with a specificity of 96.1%, sensitivity of 75% and accuracy of 94.5%. However, HDAC2-mutated neoplastic SPB1 at nasopharyngeal tumor or neck lymph node were found to be highly predictive of relapse with specificity of 96.1%, sensitivity reaching 80% and accuracy of 94.6% (Logistic Regression-odd ratio =9.8 and P Value = 0.001; Cox Regression-Hazard ratio =35.06, P Value = 0.001) . An [Ato C] missense SNV (chr1: 46668177) at ATP synthase mitochondrial F1 complex assembly factor 1 (ATPAF1) present in LY6D-positive neoplastic SPB1 was identified as predictor of early disease relapse. ATPAF1 mutation carried by LY6D-positive neoplastic SPB1 or epithelial cells were found to be more predominant in relapse patients. (FIG. 5) The table shows the performance metrics of identifying relapse by detecting ATPAF1 mutation carried in neoplastic SPB1 and different epithelial subpopulations in nasopharyngeal or neck lymph node tumor in the scRNA-seq data generated herein. KM curve data showed that ATPAF1-mutated neoplastic SPB1 at nasopharyngeal tumor or lymph node results in a lower relapse-free survival (RFS) (data not shown) . ATPAF1-mutated neoplastic SPB1 at nasopharyngeal tumor or lymph node together were found to be significantly predictive of relapse with an AUC of up to 0.890, outperforming plasma EBV DNA test (AUC = 0.736) .

- [0304] Probability of survival against time in LY6D-positive neoplastic subpopulation. LY6D immunohistochemistry of nasopharyngeal FFPE slides collected at initial diagnosis confirmed presence of LY6D-positive neoplastic subpopulation (with Allred score ≥ 6) mostly in relapse patients, as revealed by log rank test and Cox regression analysis. LY6D IHC Allred score < 6 , median survival (not reached) ; LY6D IHC Allred score ≥ 6 , median survival 27.9 months (data not shown) . FIG. 6 shows ROC with AUCs of different pathohistological scorings of LY6D immunohistochemistry results in predicting relapse (i.e. NPCR) . The graph shows AUC, a way of showing performance without setting particular a threshold of detection, while the table is showing performance metrics including sensitivity, accuracy and specificity at a particular threshold of detection.
- [0305] FIG. 7A shows the table of performance metrics of top significant relapse-predicting expression biomarkers within extracellular space or cell surface categories in Cox regression analysis, Wilcoxon test and logistic regression analysis. FIG. 7B shows the table of performance metrics of top significant relapse-predicting expression biomarkers not necessarily within extracellular space or cell surface categories in Cox regression analysis, Wilcoxon test and logistic regression analysis.

- [0306] FIGs. 8A-8C shows newly identified diagnostic expression biomarkers with high occurrence in NPC patients at initial diagnosis. FIG. 8A shows performance metrics of diagnostic expression biomarkers within extracellular space or cell surface categories significant in Wilcoxon test and logistic regression with a median difference of at least 0.5 between positive and negative results. Top expression biomarkers CKAP4, SYNGR2 and CFL1 has an AUC of 96.3%, 92.2% and 91.7%, respectively. FIG. 8B shows diagnostic expression biomarkers not necessarily within extracellular space or cell surface categories but significant in Wilcoxon test, with expression higher in NPC than non-NPC. FIG. 8C shows diagnostic expression biomarkers not necessarily within extracellular space or cell surface categories but significant in Wilcoxon test, with expression higher in non-NPC than NPC. Expression biomarker combo of chemokine (C-C motif) ligand 20 (CCL20), IL32 and lipocalin-2 (LCN2) was found to achieve 100% in sensitivity and 96.3% in PPV in NPC diagnosis (i.e. NPCD).
- [0307] FIG. 9A shows that EBV biomarker detection at nasopharyngeal tumor is more sensitive than plasma EBV DNA. EBV transcripts detected in our scRNA-seq reached sensitivity of 100% in identifying NPC patients. Even patients with plasma EBV DNA copy number = 0 copies/ml could be identified. FIG. 9B shows EBV expression biomarkers, RPMS1 in particular, can predict NPC relapse (i.e. NPCR).
- [0308] FIGs. 10A-10C show performance of artificial intelligence (AI)-powered SPB1-guided relapse prediction/diagnosis before treatment using an independent RNA sequencing dataset as demonstration. FIG. 10A is a table showing AI-powered sub-population abundance estimation in an independent RNA-seq dataset using biomarkers identified from each neoplastic and non-neoplastic epithelial sub. Wilcoxon rank sum test was used to identify significant difference between relapse and non-relapse group. A KM plot demonstrated association of AI-powered neoplastic SPB1 abundance estimation (i.e. risk score 1) with relapse. AUC of AI-powered neoplastic SPB1 abundance estimation in relapse identification reached 0.790. Sensitivity and accuracy of AI-powered neoplastic SPB1 abundance estimation in relapse prediction reached 75.0%. FIGs. 10B and 10C show identification of AI-powered non-neoplastic SPB1 abundance estimation for false positive discovery and the optimized AI-powered SPB1 relapse prediction score calculated by difference between neoplastic and non-neoplastic SPB1 (i.e. risk score 2) could reach sensitivity of 87.5% and accuracy of 86.1%. FIG. 10D is a schematic diagram demonstrating the use of optimized AI-powered SPB1-guided relapse prediction scores for NPC relapse diagnosis before treatment to maximize treatment beneficial outcomes.
- [0309] FIGs. 11A-11J illustrate non-invasive rapid test sampling using next-generation 3D-printed nasopharyngeal swab, sample processing tube and customized panels of probes at rapid antigen test (RAT) for NPC diagnosis and relapse prediction conducted at

home, community clinic, hospital or medical laboratory. FIG. 11A shows a 3D-printed nasopharyngeal swab tailor-made for high yield and least discomforting cellular and interstitial sample collection. FIG. 11B is a magnified side view of the swab tip design with dimensions. FIG. 11C is a magnified top view of the swab tip design with dimensions. FIGs. 11D-G show designs of next-generation rapid diagnostic test station/hub including sample collection/distribution structure which maximize detection of biomarkers with consistency and minimizes sample input for detection of more than one biomarker panel. Slices of detection strips can be customized for personalized disease and prognostic detection. For example, inflammation (inflam) and drug resistance (DR) can be simultaneously detected in one strip. FIG. 11D shows a view of all RAT cassettes inserted to sampling hub. FIG. 11E is another view showing the assembly or structure of RAT cassettes and before and after insertion to sampling hub. FIG. 11F shows the top view (cut in the middle) of a RAT cassette inserted to sampling hub. FIG. 11G shows the side view (cut in the middle) of a RAT cassette inserted to sampling hub. FIG. 11H is a next-generation sample processing tube including brushes/bristles therein for more efficient release of biomaterials from swab head. FIG. 11I. Sample processing tube with short hairs inside specially designed for releasing materials from (nasopharyngeal) swab.

- [0310] FIGs. 12A-12B are illustrations of lateral flow devices made from a membrane strip having an application point at the proximal end, followed by a conjugation zone, a capture zone, and an absorbent zone. The arrow shows the direction of lateral flow from the proximal to distal end. A plurality of capture lines is shown in the capture zone.
- [0311] Bar charts were generated to show relapse prediction performance of matched biomarkers from both more costly single-cell RNA sequencing data and data obtained from AI-powered platform that carries out deconvolution and cell fraction estimation from less costly bulk-sample expression data (data not shown) . Performance of both methods were found to be comparable and outperform existing golden standard plasma EBV DNA test especially in terms of accuracy and positive predictive value.
- [0312] FIG. 14A are line plots showing association of different epithelial cell types with relapse. At normal tissue adjacent to tumor, non-neoplastic ciliated cells tended to be more abundant in no relapse than that in relapse, while non-neoplastic multipotent basal cells (MPB) tended to be more abundant in relapse than that in no relapse. At nasopharyngeal primary tumor, non-neoplastic ciliated cells, and non-neoplastic cells undergoing transition to secretory goblet cells tended to be more abundant in no relapse, while non-neoplastic MPB and secretory-primed basal cells (SPB) tended to be more abundant in relapse. As more neoplastic cells, SPB and proliferating basal cells (PB) tended to be more abundant in relapse. At lymph node, activated basal cells (AB)

tended to be much more abundant in no relapse while basal cells (other than AB, SPB, MPB, PB) tended to be more abundant in relapse. FIG. 14B shows the statistical analysis of abundance of different major epithelial cell states between relapse and no relapse at different tissue types. Basal cell states were found significantly increased while ciliated and goblet cells tended to decrease at normal tissue adjacent to tumor and nasopharyngeal primary tumor in relapse. FIG. 14C shows the statistical analysis of abundance of different major epithelial subtypes between relapse and no relapse at different tissue types. Ciliated cell is the only subtype found significantly reduced at normal tissue adjacent to tumor in relapse, compared with no relapse. At normal tissue adjacent to tumor, MPB and SPB showed a trend to increase in relapse while transition cells to goblet states tended to decrease in relapse. At nasopharyngeal primary tumor, basal cells in general including SPB (excluding AB, MPB, transition cells to goblet cells) tended to increase in relapse while AB, ciliated, goblet and transition cells to goblet cells tended to decrease in relapse. At lymph node, AB tended to decrease in relapse while basal cells (excluding AB, MPB, SPB and transition cells to goblet cells) tended to increase in relapse while AB tended to decrease in relapse.

- [0313] FIG. 15 shows the table about discovery of significant upregulation of neoplastic MPB1 subpopulation at lymph node in relapse when compared with no relapse.
- [0314] Univariate and multivariate Cox regression analysis confirmed that Allred score ≥ 6 , presence of neoplastic SPB1 subpopulation at nasopharynx and nasopharynx/lymph node, presence of neoplastic ATPAF1-mutated SPB1 subpopulation at nasopharynx and nasopharynx/lymph node could be independent predictor of relapse. Univariate and multivariate Cox regression analysis confirmed that Allred score ≥ 6 could be independent predictor of overall survival.
- [0315] In summary, the present studies led to the discovery of a number of previously uncovered DNA mutations and expression biomarkers with performance metrics outperforming many biomarkers previously identified and implicated in NPC diagnosis and relapse prediction.
- [0316] 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 method and compositions described herein. Such equivalents are intended to be encompassed by the following claims.
- [0317] MIC SNT Signature expression of SPB and SPB1 subpopulations relative to other epithelial cell types for AI analysis:

Gene	Neoplastic AB1	Neoplastic AB2	Neoplastic Basal1	Neoplastic Basal2	Neoplastic Basal3	Neoplastic Basal4	Neoplastic Basal5	Neoplastic MPB1	Neoplastic MPB2
KRT5	30.86	25.202	53.657	27.017	6.668	5.887	41.699	43.218	58.477
KRT15	36.651	29.09	50.295	17.968	4.816	1.076	53.635	46.537	92.575
LY8D	0.77	6.286	1.343	1.292	0.071	1.211	1.33	0.912	0.742
SLMOA9	1.066	0.852	0.716	1.89	0.14	1.067	1.055	0.83	0.191
CD31	3.791	11.177	6.076	16.062	5.241	47.156	5.174	6.354	6.406
SLMOA2	1.707	2.808	8.796	2.831	0.638	0.839	10.398	13.907	12.437
KRT17	10.23	2.251	22.052	7.255	4.457	2.395	26.308	30.454	26.775
SLMOA8	0.991	0.603	0.902	1.346	0.553	0.566	1.108	0.638	0.423
SPN	3.114	2.119	5.773	3.344	0.557	0.813	7.707	6.613	6.709
KRT13	5.033	4.077	5.68	3.366	0.712	1.143	5.441	3.366	7.332
KRT16	2.407	0.369	2.147	1.49	0.126	0.481	3.296	0.259	0.289
KRT14	2.851	1.575	2.583	1.304	0.011	0.274	3.279	1.32	1.859
LY8D	6.791	13.595	13.271	13.027	9.106	19.166	10.465	13.617	14.62
CTNNT3	5.13	1.626	5.821	7.067	2.538	16.644	3.647	6.453	3.589
LCN2	0.622	0.53	0.163	0.857	0.855	1.043	0.999	0.091	0.073
R32	4.151	8.623	6.522	6.352	5.439	5.43	11.194	7.062	5.182
MEAT1	32.983	9.058	7.793	5.191	2.3	6.678	5.925	7.385	8.167
WDR72	1.419	1.301	1.353	2.65	1.255	2.556	1.096	2.267	1.962
ADP3	0.567	0.897	0.969	0.354	0.04	0.596	1.011	2.505	1.592
DSP	36.398	6.336	7.425	5.283	2.355	3.061	5.81	4.439	6.772
PRC4HD	2.249	2.078	6.4	5.603	3.519	4.651	5.256	5.114	5.839
CAHML3	2.531	2.187	5.296	2.638	0.193	1.439	3.213	2.524	5.199
DAP11	1.643	1.425	3.335	3.438	0.703	1.356	4.025	3.614	2.613
SOD4	11.215	5.451	10.35	8.929	11.568	4.3	7.632	6.282	10.401
SLP1	0.879	0.765	0.289	0.272	0.272	0.349	0.291	0.94	0.349
R189	9.804	8.321	12.079	7.594	3.951	5.126	12.125	8.316	32.836
LGALS1	4.768	11.081	10.928	10.495	2.452	7.642	10.069	8.295	8.922
SPNMBL	2.255	1.814	3.015	1.199	0.665	0.415	3.076	4.905	2.076
SLF3	6.195	1.907	2.967	1.991	0.256	2.127	2.503	1.097	2.121

Gene	Neoplastic MP83	Neoplastic PB1	Neoplastic PB2	Neoplastic PB81	Neoplastic PB82	Neoplastic PB83	Neoplastic PB84	Neoplastic SP85	Non-neoplastic AB1	Non-neoplastic AB2
KMT5	23.419	24.758	35.82	45.862	53.783	6.932	61.891	25.166	17.687	10.363
KMT15	24.978	11.536	16.19	30.963	42.406	10.187	15.613	9.432	17.77	10.822
KMT6	0.05	0.784	0.803	17.743	12.14	15.528	11.646	5.265	0.747	0.464
SLC6A8	0.061	1.182	0.515	17.536	5.402	29.087	145.599	15.796	8.511	4.522
CSF4	2.182	12.931	12.995	15.108	10.432	14.103	18.969	9.174	2.255	2.627
SLC6A2	2.787	4.321	3.565	15.047	11.055	2.083	34.398	7.901	5.713	8.422
KMT17	29.201	10.461	10.948	12.819	16.65	2.962	62.789	11.729	7.85	10.634
SLC6A8	0.298	0.769	0.636	12.811	6.503	12.154	109.981	7.513	3.79	1.571
SPN	3.299	2.979	3.423	12.378	14.187	2.899	20.788	6.319	2.686	2.931
KMT13	0.073	2.915	4.275	11.587	24.327	6.583	10.622	5.134	3.333	0.753
KMT16	2.777	0.941	1.411	11.368	6.586	0.194	27.023	4.275	4.055	0.86
KMT6A	0.267	0.51	0.792	11.361	10.132	1.667	47.035	7.507	2.144	3.974
USD	2.033	8.307	10.232	10.899	14.409	17.205	17.131	13.137	5.724	6.136
CCND1	4.058	7.86	7.143	10.636	5.24	7.385	4.302	13.952	2.127	2.745
LCN2	0.465	0.82	0.545	9.628	4.375	66.137	2.902	3.825	5.571	0.834
IL32	1.292	4.796	5.84	8.902	5.856	8.561	8.042	2.624	3.799	5.003
MEI11	2.662	3.802	5.211	8.637	12.815	13.421	11.921	6.059	52.885	7.253
MYO22	5.224	1.518	1.588	7.005	4.213	42.802	7.437	4.161	11.112	4.901
KMT5	0.167	0.48	0.418	6.97	2.935	13.968	2.661	2.835	7.425	9.056
ESR	3.483	4.72	6.173	6.95	9.23	4.532	8.871	16.223	4.678	2.456
PTCD2	4.736	5.43	5.474	6.636	6.403	7.183	3.911	4.387	1.529	2.273
PA14A18	0.161	2.791	3.681	6.319	7.739	0.888	4.513	2.197	2.886	0.708
DAP1	8.346	1.475	1.518	5.663	5.442	2.503	1.866	3.535	0.768	0.421
SOX4	9.627	6.184	7.008	5.432	9.133	7.297	5.447	6.644	6.652	1.647
SLP	0.11	0.215	0.131	5.358	1.888	31.562	7.189	8.138	8.613	4.463
KMT6	6.97	6.236	6.64	5.34	28.299	6.645	11.905	14.319	10.081	7.165
LSAL31	9.561	12.593	10.942	5.177	4.124	0.32	8.785	1.424	2.282	4.704
GPMB8	0.569	0.91	0.973	5.133	4.353	1.127	4.284	1.925	0.975	0.818
KMT	1.604	1.267	1.531	4.68	12.396	13.395	5.248	10.559	10.263	1.56

Non-metabolic Base1	Non-metabolic Base2	Non-metabolic Base3	Non-metabolic Base4	Non-metabolic Base5	Non-metabolic Chloride1	Non-metabolic Chloride2	Non-metabolic Chloride3	Non-metabolic Chloride4	Non-metabolic Chloride5
31.833	17.108	2.257	3.186	17.733	0.02	0.34	2.021	0.201	6.368
31.473	16.255	7.407	3.739	38.511	0.182	0.348	0.413	0.740	4.002
4.905	0.865	0	0.029	1.087	0.017	0.034	0.69	0.086	1.065
11.005	6.586	3.734	1.288	3.439	0.157	0.112	1.98	0.168	8.302
4.782	14.883	9.513	34.437	1.597	0.129	0.1	1.344	1.836	1.836
77.495	0.952	0.952	0.653	38.392	0.31	0.028	4.915	0.177	12.613
29.651	0.792	0.792	0.309	27.802	0.149	0.129	0.447	0.52	6.832
30.123	9.438	7.958	0.821	2.795	0.016	0.071	0.765	0.036	5.522
5.719	0.789	0.789	4.381	5.98	0.408	0.282	0.505	0.496	2.563
5.427	6.924	2.434	4.258	3.494	0.028	0.162	0.175	0.262	0.128
5.175	0.899	0	0.496	1.683	0.003	0.068	0.019	0.013	0.708
30.834	0.554	0.101	0.486	5.724	0.02	0.046	0.187	0.059	3.345
37.858	9.318	8.317	34.034	3.771	0.899	4.787	1.537	0.837	3.037
3.15	2.973	2.996	1.521	2.395	0.62	0.308	0.365	1.241	1.36
3.907	2.907	5.244	0.145	0.359	30.628	8.472	90.576	28.863	2.514
2.678	7.584	4.955	3.43	6.112	0.519	1.05	1.018	1.635	3.751
9.752	11.549	9.752	7.034	5.369	12.13	43.342	167.662	45.579	38.905
18.582	4.935	10.489	13.887	5.923	50.874	11.984	3.7	6.288	51.37
25.945	3.977	0.937	0.208	17.101	6.279	4.81	6.898	4.014	9.496
4.447	3.184	2.966	1.372	2.46	2.943	3.305	2.956	2.616	2.362
3.125	4.772	2.813	5.207	1.913	3.708	0.282	2.093	1.038	2.448
2.764	1.381	0.287	1.206	0.687	0.197	0.043	0.135	0.174	0.531
0.935	0.411	0.195	0.087	0.785	0.019	0.063	0.138	0.019	0.21
4.085	6.525	15.913	1.986	1.784	0.59	0.971	2.466	0.636	2.262
6.314	3.063	1.263	2.842	5.602	37.43	8.979	96.637	27.172	88.85
14.081	6.791	3.234	7.523	20.938	1.278	1.818	2.139	1.649	6.582
2.481	5.504	1.213	1.703	1.86	0.219	0.383	0.414	0.524	1.31
1.515	0.862	0	0.646	0.646	0.014	0.079	0.084	0.074	0.217
2.957	1.192	1.31	1.31	2.173	7.484	5.134	6.883	7.752	5.258

Gene	Non-neoplastic Gblet2	Non-neoplastic Gblet3	Non-neoplastic Gblet4	Non-neoplastic Gblet5	Non-neoplastic Gblet6	Non-neoplastic Gblet7	Non-neoplastic Gblet8	Non-neoplastic Gblet9	Non-neoplastic Gblet10	Non-neoplastic Gblet11	Non-neoplastic Gblet12	Non-neoplastic Gblet13
2073	0.077	0.019	1.379	0.132	14.51	0.276	20.013	18.42	17.613	24.792		
88713	0.241	0.017	0.824	0.254	38.449	3.392	45.683	52.84	36.663	11.139		
1960	0.017	0.012	0.101	0.085	0.069	0	0.283	0.135	0.106	9.281		
502049	0.283	0.037	7.525	0.493	1.245	0.103	5.315	1.987	1.797	35.593		
5313	0.548	0.263	0.283	0.849	0.964	0.409	1.599	0.871	0.371	21.516		
530942	0.112	0.112	2.352	0.179	22.99	1.123	47.785	31.889	22.993	78.582		
88712	0.098	0.122	3.135	0.122	20.993	0.023	23.471	31.055	23.08	48.905		
530946	0.081	0.082	3.046	0.046	0.346	0	3.106	0.511	0.246	36.281		
59	0.346	0.05	0.447	0.194	4.213	0.549	0.213	0.061	0.373	11.236		
88713	0.018	0.008	0.008	0.008	0.722	0.013	3.459	4.348	0.095	11.8		
88716	0.012	0.004	0.134	0.013	0	0	0.031	0.15	0.039	1.025		
88716	0.04	0.025	0.376	0.226	0.974	0	2.617	5.018	0.214	33.856		
88716	0.816	0.082	5.86	2.845	4.255	6.725	6.482	6.402	4.541	6.857		
88716	0.787	0.278	3.49	2.32	1.642	0.361	2.132	2.225	2.871	2.295		
530942	0.756	0.701	67.416	12.857	2.654	0.782	3.409	0.363	18.573	4.283		
530942	0.568	0.295	2.659	1.575	2.686	1.957	1.293	0.903	1.249	1.594		
530942	15.457	8.772	3.744	18.893	9.06	3.34	5.774	7.059	4.213	3.938		
530942	171.9	88.275	186.659	74.459	23.228	2.371	13.177	5.84	21.049	30.38		
530942	1.61	1.093	3.482	5.985	24.867	5.087	23.048	22.728	9.911	25.557		
530942	0.785	0.319	2.482	2.594	3.286	4.935	2.848	2.222	2.986	4.851		
530942	0.276	0.176	1.046	0.332	1.177	0.319	2.281	2.033	1.67	2.929		
530942	0.003	0	0.09	0.005	1.402	1.847	0.993	0.892	0.1	2.574		
530942	0.001	0.001	0.057	0.008	0.743	0	0.522	0.379	0.463	0.905		
530942	0.058	0.144	3.876	1.148	2.101	3.523	2.438	3.107	2.073	2.409		
530942	975.926	234.46	343.791	156.506	47.537	1.32	8.177	7.235	15.311	42.838		
530942	0.033	0.077	0.022	0.398	5.395	3.062	8.281	33.146	3.849	3.286		
530942	0.736	0.773	0.53	0.3	1.17	2.084	0.765	0.449	2.027	1.575		
530942	0.005	0.009	0.008	0.042	0.042	0.057	1.283	0.926	0.41	0.689		
530942	3.502	0.864	6.988	6.337	2.433	8.329	1.673	4.566	2.856	2.987		

Gene	Non-neoplastic: PB2	Non-neoplastic: SPB	Goblet transition1	Non-neoplastic: SPB	Goblet transition2	Non-neoplastic: SPB3	Non-neoplastic: SPB3	Non-neoplastic: SPB4	Non-neoplastic: SPB5
8875	28.075	0.682	0	23.686	26.412	4.29	47.599	34.652	34.652
8876	14.824	0.559	4.186	5.368	10.095	1.35	2.605	6.198	6.198
1181	1.686	2.375	0.621	9.928	4.991	4.821	41.443	23.554	23.554
310049	8.275	16.378	581.075	21.189	9.663	32.003	223.772	38.74	38.74
3374	9.183	3.071	17.482	7.146	4.592	7.573	33.228	7.2	7.2
310062	47.747	2.637	0.116	47.497	40.625	21.113	140.878	107.84	107.84
88717	16.663	3.666	67.812	6.277	13.948	5.496	84.137	21.563	21.563
310048	6.653	9.764	213.378	11.313	5.522	26.53	252.604	36.012	36.012
999	7.701	1.179	2.014	9.182	9.567	5.217	40.821	15.367	15.367
88715	5.674	1.065	113.458	7.261	12.531	3.146	83.69	27.151	27.151
88716	2.202	0.066	4.365	0.863	0.634	0.096	49.82	2.512	2.512
88766	10.171	1.764	19.665	9.147	11.045	3.727	128.834	30.672	30.672
999	8.527	6.566	0.766	3.454	2.515	4.129	3.942	2.665	2.665
25802	3.31	1.947	0.372	4.083	3.583	2.487	2.71	3.455	3.455
1062	1.907	102.666	204.8	9.15	4.078	43.085	19.367	8.07	8.07
833	4.566	1.305	1.137	1.182	0.884	1.194	1.568	0.789	0.789
88215	9.951	25.613	96.977	8.326	12.138	9.794	12.063	5.904	5.904
94722	6.7	212.093	80.891	43.844	39.684	108.132	24.256	32.69	32.69
88293	11.408	12.768	5.04	50.525	53.564	35.372	48.502	50.423	50.423
109	4.708	2.327	0.442	5.182	4.558	3.046	7.474	5.182	5.182
3552	3.552	1.35	0.255	2.976	2.305	2.799	2.652	3.218	3.218
816180	3.15	0.38	7.131	5	3.299	3.502	14.194	4.954	4.954
244603	0.408	0.234	0	1.074	0.899	0.793	0.457	0.584	0.584
94511	4.209	3.272	3.358	1.72	3.412	3.188	1.999	1.256	1.256
3094	4.188	134.65	53.259	48.182	40.336	98.214	34.342	24.827	24.827
8391	6.971	2.998	7.864	3.762	19.612	4.241	6.261	4.256	4.256
2058	4.661	0.207	0	0.317	0.211	0.062	0.464	0.937	0.937
166163	0.577	0.11	0	1.232	0.637	0.409	0.686	0.684	0.684
699048	1.402	10.93	29.058	5.97	11.961	9.366	4.776	3.248	3.248

Gene	Neoplastic AB1	Neoplastic AG2	Neoplastic Basal1	Neoplastic Basal2	Neoplastic Basal3	Neoplastic Basal4	Neoplastic Basal5	Neoplastic PRP61	Neoplastic PRP82
TUBA1B	2.847	7.506	4.655	6.009	5.656	6.488	4.743	2.796	4.389
IFI6	1.168	2.862	1.607	1.998	0.719	2.963	2.959	2.795	1.525
AGR2	0.334	0.407	0.526	0.421	0.168	0.594	0.672	1.217	0.449
GSLA1	0.319	0.388	0.729	0.992	0.159	0.936	0.458	0.313	0.556
CTC20	3.624	5.209	3.615	11.811	4.809	36.787	2.959	1.466	1.763
BSI2	2.676	4.58	4.991	4.398	1.957	3.682	4.926	6.28	5.798
H2AFZ	1.678	2.983	3.397	4.274	4.026	6.659	3.162	2.589	2.976
IFI1M1	2.692	7.625	4.879	6.854	1.756	3.868	9.454	13.106	9.479
SOD2	2.562	4.712	1.583	2.805	1.077	1.263	2.137	1.821	1.807
HLA-DQA1	4.287	2.671	5.677	7.176	6.23	14.639	4.322	4.736	3.382
CLDN1	8.404	3.397	5.489	2.913	1.059	0.921	3.823	4.866	7.616
FOX	10.195	9.077	12.702	6.599	3.511	6.129	12.537	5.265	34.989
CXCL10	2.563	2.592	2.984	12.055	0.93	11.732	10.435	3.874	1.28
IDO1	1.45	0.54	0.274	0.85	1.62	1.727	0.698	0.179	0.111
CEBPD	2.041	0.854	2.166	1.836	0.679	0.462	2.268	1.851	3.509
FNTA	1.384	0.668	1.197	1.403	1.14	0.998	1.134	1.075	1.109
MYBPC1	10.186	1.626	3.826	2.518	0.621	5.75	2.271	1.513	2.65
LYPD3	0.922	0.263	0.713	0.415	0.071	0.275	0.671	0.431	0.728
HLA-DRE5	1.852	1.51	2.327	2.531	1.244	3.67	2.228	2.689	1.286
IFR2	8.782	5.062	6.885	5.05	3.457	4.097	6.169	2.833	16.163
CTSC	3.11	1.522	1.361	1.972	2.537	0.604	1.865	1.071	1.086
DUSP1	4.946	4.482	5.693	3.216	1.308	4.447	6.154	3.218	13.315
MT2A	1.252	4.072	1.559	5.717	4.056	5.488	3.364	0.705	0.94
MSMB	0.147	0.191	0.255	0.24	0.016	0.107	0.22	0.115	0.142
SERPINF1	0.538	0.812	0.447	0.406	0.093	0.952	0.558	0.568	0.494
FABP5	0.371	0.364	0.387	0.308	0.457	1.202	0.423	0.409	0.592
TGM	7.122	2.566	2.572	2.755	1.293	1.997	1.826	1.365	2.991
MTOR2	1.989	1.627	2.516	3.322	0.764	2.74	2.048	2.823	1.923
RHOV	0.936	0.224	0.518	0.482	0.21	1.056	0.528	0.136	0.334
DS1	8.076	4.838	4.936	2.11	0.899	0.762	3.906	5.881	6.938

Gene	Neoplastic MPB3	Neoplastic PB1	Neoplastic PB2	Neoplastic PB3	Neoplastic PB4	Neoplastic PB5	Non-neoplastic AD1	Non-neoplastic PB2
TUBA1B	5.147	30.757	17.398	4.594	4.189	4.156	5.145	2.942
IFB6	0.544	1.199	1.077	4.433	1.903	7.878	2.084	1.852
AGR2	0.201	0.346	0.257	4.288	2.875	20.193	0.777	1.554
GSTA1	0.077	0.486	0.553	4.073	1.709	8.554	0.953	1.779
CC120	0.07	6.036	8.415	4.026	5.109	6.724	12.227	8.373
BST2	2.292	3.11	3.284	3.916	3.404	5.036	4.924	1.691
H2AFZ	1.866	15.573	10.484	3.481	3.29	3.258	2.472	4.493
IFITM1	4.09	4.243	4.139	3.444	3.115	10.804	5.233	1.795
SOD2	1.155	0.764	1.184	3.421	2.422	3.28	10.355	3.201
HLA-DQA1	1.349	7.075	7.098	3.332	4.725	5.49	5.303	1.062
CLDN1	0.51	2.549	3.131	3.341	5.193	1.569	4.546	2.712
FOS	5.017	6.224	6.709	3.006	33.215	4.001	14.225	11.822
CXCL10	1.102	4.488	5.309	2.892	3.23	3.087	11.071	3.452
HDG1	0.029	0.703	0.51	2.78	1.907	6.371	5.579	4.494
CEBPD	1.015	1.075	1.221	2.769	3.939	2.149	3.119	1.228
FN1A	0.918	1.159	1.126	2.56	1.479	1.685	1.288	4.229
MYBPC1	0.304	1.699	2.523	2.415	2.243	4.153	0.716	5.543
LYPD3	0.323	0.274	0.36	2.41	2.691	0.837	2.948	3.45
HLA-DH85	0.835	2.201	2.612	2.407	1.97	8.046	4.239	0.964
IER2	3.518	5.149	5.492	2.302	14.296	2.887	5.673	10.401
CTSC	1.194	1.266	1.442	2.301	1.803	3.689	3.96	13.026
DUSP1	0.528	2.975	3.472	2.09	20.218	1.926	7.307	3.724
MT2A	2.216	5.263	4.962	2.085	1.699	3.411	2.013	0.846
NSM8	0.003	0.091	0.18	1.978	2.272	6.077	0.819	0.498
SERPINB1	0.154	0.326	0.352	1.806	1.068	4.174	2.808	1.563
FABP5	0.036	0.765	0.553	1.792	0.767	0.728	1.88	2.425
TCM4	0.32	1.294	2.039	1.65	2.805	2.993	3.483	20.087
NTRK2	3.695	2.504	2.426	1.602	1.939	1.048	0.583	0.591
RHOV	0.056	0.385	0.474	1.593	2.53	2.594	1.768	0.893
DST	2.134	1.613	2.029	1.536	3.065	0.516	4.126	0.581

[illegible]

Gene	Non-neoplastic: F8L	Non-neoplastic: SFB	Globet transition2	Non-neoplastic: SFB1	Non-neoplastic: SFB2	Non-neoplastic: SFB3	Non-neoplastic: SFB4	Non-neoplastic: SFB5
TUBA1B	21.915	17.39	2.361	0	2.436	3.071	2.425	2.746
IF6	6.991	5.685	7.407	0.227	3.97	4.055	7.786	9.131
A382	5.403	1.211	32.231	2.18	9.98	7.523	20.291	5.576
GSTA1	0.957	0.444	11.243	0	5.604	6.849	14.45	1.744
CC120	3.974	4.851	0.935	0.349	0.52	0.199	0.792	1.752
86T3	4.882	5.365	4.172	0.233	3.529	3.221	4.76	6.286
H2AF2	20.267	11.486	2.521	1.879	3.255	3.195	3.089	2.134
BTM1	16.993	14.181	15.265	0.51	11.978	10.895	16.646	8.624
SOD2	1.277	2.608	1.918	12.972	1.593	1.675	1.38	4.948
HLA-DQA1	3.084	4.335	3.116	0.333	2.275	2.335	2.449	1.611
CLDN3	1.86	2.141	0.522	0.116	2.235	2.465	1.019	2.249
FOX	5.074	12.883	6.181	5.55	5.872	43.269	9.463	7.747
CXCL10	6.408	4.36	0.565	0.11	0.73	0.258	1.776	3.082
IDO1	0.758	0.453	5.965	6.686	0.611	0.697	1.153	1.62
TERPO	1.132	0.894	1.098	2.806	1.384	1.962	1.512	1.41
FN1A	0.524	0.809	0.478	0.372	0.571	0.733	0.486	0.366
MYBPC1	0.016	0.373	0.055	0	0.022	0.089	0.034	0
LYF53	0.86	0.243	0.241	2.019	1.313	1.13	0.804	8.306
HLA-DHB5	3.573	3.271	4.732	0.465	3.224	4.468	3.733	2.15
HER2	1.957	4.456	2.63	5.167	1.82	10.207	2.921	3.5
C7SC	2.521	2.353	5.175	0.227	4.239	4.308	5.627	6.049
DUSP1	1.649	6.023	3.519	8.586	2.337	11.673	5.485	5.972
M12A	31.98	19.999	9.629	0.255	9.603	9.608	9.201	2.679
MSMR	0.582	0.275	67.779	6.445	1.719	1.06	7.346	1.128
SERPINE1	3.238	0.995	4.291	9.992	9.19	6.05	8.413	21.631
FABP5	6.934	3.604	1.167	0	2.377	3.392	2.211	35.346
TCM4	0.087	0.247	0.993	0	0.367	0.864	0.317	0.605
NR1K3	0.36	0.681	0.051	0	0.161	0.198	0.051	0.077
PRCV	0.608	0.498	1.276	0	1.611	2.266	1.524	2.185
DST	1.459	1.667	0.225	0	1.041	1.151	0.287	0.306

Gene	Neoplastic AB1	Neoplastic AB2	Neoplastic Basal1	Neoplastic Basal2	Neoplastic Basal3	Neoplastic Basal4	Neoplastic Basal5	Neoplastic MPPI1	Neoplastic MPPI2
ZFP36	4.732	5.092	5.575	3.194	1.568	2.007	6.356	1.786	16.963
TSPAN1	0.222	0.31	0.264	0.198	0.065	0.553	0.281	0.66	0.233
ISG15	0.684	1.346	0.699	1.078	0.461	1.113	2.414	1.095	0.663
STMN1	0.97	2.51	2.68	2.955	4.209	1.751	1.976	1.774	1.806
EIF4EBP1	0.526	1.073	1.038	1.364	2.772	1.02	0.963	0.59	0.554
SERPINE3	0.149	0.85	0.005	0.01	0.009	0.004	0.051	0.076	0.003
LAMB3	2.902	1.657	1.922	1.473	0.651	1.2	1.962	3.069	3.177
ADRF	0.424	0.3	1.218	0.808	0.111	2.205	0.378	2.623	1.344
MTX	0.367	0.799	0.728	1.291	0.81	0.75	1.29	0.57	0.523
PTTG1	0.395	0.59	0.639	0.649	0.563	0.798	0.877	0.475	0.334
COL17A1	2.591	2.18	1.832	1.427	0.628	0.248	1.663	3.267	2.206
FABP7	0.642	0.091	0.517	1.377	3.005	0.769	0.934	0.73	1.771
NTS	0.196	0.035	0.373	0.595	0.164	0.268	0.382	0.034	0.016
RAC2	0.749	1.94	0.905	1.297	0.642	1.538	2.402	1.034	0.794
NRM	3.38	7.615	3.312	11.673	10.193	18.235	8.974	1.306	1.234
PDZK1P1	0.147	0.063	0.39	0.33	0.323	0.24	0.426	0.704	0.524
PI3	0.02	0.015	0.008	0.081	0.003	0.006	0.021	0.01	0
HES1	1.764	0.707	1.908	0.866	0.388	0.497	1.418	0.68	4.246
SOC31	1.792	1.045	3.052	4.192	0.869	4.745	1.872	2.514	3.549
NME1	0.582	1.729	1.168	1.505	3.049	1.067	1.271	1.041	0.578
HIT3	0.553	0.663	0.511	0.793	0.204	0.755	0.905	0.617	0.579
LPK1B	0.222	0.121	0.053	0.034	0.015	0.035	0.097	0.204	0.003
IGHG1	4.111	57.137	0.505	0.605	0.157	0.65	0.849	0.924	0.331
ITF	0.666	0.074	1.015	1.359	0.456	6.418	0.644	0.713	0.623
ALDH3A1	0.145	0.188	0.204	0.17	0.015	0.164	0.245	1.499	0.408
CLU	0.614	0.718	0.741	1.527	0.092	2.823	0.543	5.74	1.901
CCNI1	4.425	2.646	1.3	1.024	0.652	0.961	1.356	1.156	2.352
CLCA2	0.168	0.183	0.116	0.071	0.002	0.004	0.193	0.263	0.169
CRABP2	0.163	0.141	0.313	0.166	0.053	0.589	0.295	0.573	0.296
MTHFD2	0.805	0.628	0.553	1.059	2.128	1.412	0.665	0.27	0.323

Gene	Neoclastic MPB3	Neoclastic PB1	Neoclastic PB2	Neoclastic SPB1	Neoclastic SPB2	Neoclastic SPB3	Neoclastic SPB4	Neoclastic SPB5	Non-neoclastic AB1	Non-neoclastic AB2
ZFP36	1.029	2.179	2.899	1.515	13.638	6.205	6.247	6.091	6.951	5.584
TSPAN1	0.037	0.166	0.151	1.398	0.558	5.148	0.744	0.28	1.588	0.627
ISG15	0.154	0.61	0.726	1.396	0.915	2.951	1.258	0.623	1.98	1.591
STAM1	2.153	9.98	5.908	1.314	1.408	1.298	0.46	2.406	1.157	1.755
EIF4EBP1	1.325	2.113	1.627	1.292	0.732	0.734	1.09	1.958	0.294	1.046
SCRPNB3	0	0.019	0.01	1.26	0.047	7.175	1.073	1.396	6.944	3.064
LAMB3	2.119	0.735	0.922	1.221	1.895	1.333	3.713	0.202	1.528	0.679
ADRF	0.213	1.409	1.092	1.167	1.532	1.346	0.693	0.335	0.762	0.343
MTX	0.317	1.125	1.081	1.148	1.613	1.146	1.075	0.938	1.652	1.722
PTTG1	0.097	5.707	1.938	1.13	1.004	0.378	0.736	0.873	0.267	0.509
COL17A1	1.362	0.845	0.94	1.115	1.537	0.598	1.741	0.113	0.976	0.69
FABP7	1.395	0.707	0.639	1.09	0.239	0	0.987	0.256	0.06	0
MTS	0.007	0.747	0.548	1.017	0.279	1.04	0.096	0.371	0.385	0.016
RAC2	0.664	1.139	1.035	1.004	0.704	0.498	1.266	0.523	0.766	2.002
VIM	12.466	11.013	9.97	0.99	0.863	0.231	0.541	0.297	3.713	7.498
PDZK1IP1	0.228	0.123	0.155	0.984	0.326	2.021	3.852	0.765	0.399	0.558
PB3	0.013	0.008	0.005	0.937	0.065	1.14	5.36	0.75	0.113	0.009
HES1	0.48	0.969	1.016	0.921	7.27	1.358	1.664	1.688	1.286	0.445
SOC3	0.788	3.745	3.842	0.917	1.883	0.362	2.182	0.378	1.07	0.785
NMIF4	1.854	1.607	1.624	0.894	0.835	0.738	0.813	1.102	0.321	0.451
IFT3	0.098	0.335	0.472	0.882	0.592	1.109	0.843	0.38	0.794	0.571
UPK1B	0.012	0.052	0.024	0.861	0.199	2.75	0.268	3.532	0.647	0.307
IGHG1	0.062	0.4	1.381	0.828	0.176	0.585	0.304	0.481	5.208	66.322
LIF	0	0.589	0.359	0.815	2.445	0.206	20.416	0.607	0.717	0.392
ALDH3A1	0.071	0.212	0.201	0.81	0.79	2.267	0.168	0.201	2.226	1.418
CLU	0.667	0.478	0.493	0.798	0.666	1.707	2.275	0.186	2.258	2.988
CCN1	0.739	0.931	1.182	0.77	1.743	1.288	1.165	3.542	3.511	1.217
CLCA2	0.172	0.082	0.061	0.721	0.409	0.248	0.704	0.492	0.323	0.259
CRABP2	0.203	0.252	0.230	0.711	1.058	1.973	1.638	0.62	0.272	0.234
MTHFD2	0.502	1.786	1.448	0.706	0.323	0.59	0.66	0.798	0.961	0.856

Source	Non-metabolic Band1	Non-metabolic Band2	Non-metabolic Band3	Non-metabolic Band4	Non-metabolic Band5	Non-metabolic Band6	Non-metabolic Band7	Non-metabolic Band8	Non-metabolic Band9	Non-metabolic Band10	Non-metabolic Band11	Non-metabolic Band12	Non-metabolic Band13	Non-metabolic Band14	Non-metabolic Band15
ZFP36	4.502	2.068	2.135	2.541	15.431	1.585	2.439	2.115	2.538	2.238					
150A01	0.328	0.206	1.074	0.305	0.305	25.101	12.025	15.328	1.834						
16C15	3.274	3.942	3.009	3.265	2.64	3.99	2.356	2.508	6.487	0.698					
370A01	1.508	3.674	0.587	0.617	0.652	0.297	0.155	1.021	0.564	0.855					
1744B01	0.553	0.927	0.843	0.649	0.278	0.269	0.089	0.208	0.738	0.738					
380A03	5.221	0.897	0.053	0	1.747	0.137	0.368	13.743	1.254	12.663					
130A03	2.29	0.897	0	0.844	2.406	0.056	0.084	0.229	0.117	0.488					
130A04	2.723	0.897	0.359	0.447	2.892	0.06	0.187	0.245	0.229	0.833					
1471X	15.755	1.136	1.136	8.794	8.794	0.187	0.147	0.245	0.586	7.968					
171C1	0.558	0.795	0.795	0.422	0.483	0.177	0.177	0.433	0.182	3.851					
COL17A1	2.037	0.599	0.842	0.821	2.123	0.015	0.015	0.041	0.007	0.227					
17A002	0.431	0.183	0	0	0	0.087	0.115	0.071	0.043	4.147					
17X	0.431	0.098	0.098	0	0.168	0.02	0.063	0.473	0.069	0.272					
18A12	0.619	0.117	0.153	0.627	1.561	0.181	0.858	0.731	1.892	1.306					
171X	1.719	1.27	9.661	1.066	0.096	2.881	1.529	1.529	3.827	4.869					
18C24102	0.335	0.72	0.109	0.791	0.543	0.331	0.085	1.519	0.43						
173	0.159	0.103	0.046	0.647	0.008	0.098	0.037	1.57	0.105	0.317					
16C31	2.284	1.84	1.51	0.894	1.008	0.066	0.839	0.856	0.314	1.09					
16C31	0.599	2.587	0.23	3.078	0.759	0.072	0.124	0.203	0.172	0.37					
16A4	1.129	1.727	1.727	1.123	0.839	0.183	0.403	1.063	0.549	0.776					
1611	1.159	1.918	0.916	1.485	1.139	0.697	0.494	1.265	1.025	0.255					
16A18	0.772	0.11	0.053	0.677	0.112	0.096	0.107	1.125	0.104	1.859					
16103	1.433	0.841	0.337	0.532	0.532	0.661	2.046	1.955	0.592	3.331					
171	0.825	1.028	0.362	7.312	0.065	0.143	0.708	1.624	3.645	83.293					
161013A1	0.552	0.659	0.023	0.026	0.176	5.598	2.048	3.514	6.471	1.658					
171	0.853	3.984	0.16	1.407	1.806	9.422	5.193	6.44	4.8	5.183					
17A1	1.065	0.907	1.044	0	0.244	0.575	1.384	0.723	0.683	3.098					
17A2	0.678	0.023	0	0	0.001	0.001	0.005	0.065	0.044	0.211					
17A002	0.921	0.295	0.295	0.017	0.24	0.025	0.005	0.586	0.011	0.525					
171102	0.233	0.753	1.075	2.388	0.434	0.154	0.116	0.267	0.201	0.722					

[illegible]

Gene	Non-neoplastic PB2	Non-neoplastic SFB	Goblet transition1	Non-neoplastic SFB-Goblet transition2	Non-neoplastic SPB1	Non-neoplastic SPB2	Non-neoplastic SPB3	Non-neoplastic SPB4	Non-neoplastic SPB5
ZFP36	4.422	2.758	24.033	2.362	14.776	3.641	3.811	2.014	2.014
TPST4L1	0.395	13.935	2.297	7.454	5.906	12.36	2.747	2.91	2.91
ISG15	2.632	6.042	23.3	2.612	1.576	4.89	8.09	3.165	3.165
STMN1	7.417	0.281	0	0.452	0.423	0.499	0.175	4.668	4.668
EP4BP1	1.21	0.334	0	0.757	0.564	0.614	1.193	1.879	1.879
SEPTIN8	1.756	44.411	0.349	43.903	50.016	60.511	24.876	17.933	17.933
LAMB3	1.997	0.549	1.708	0.512	0.554	0.275	1.147	0.949	0.949
ADRF	1.268	1.847	4.299	2.041	1.963	2.083	1.73	1.645	1.645
MTX	6.645	5.248	0.255	16.396	15.835	8.031	5.87	22.385	22.385
PTTG1	1.892	0.074	0	0.554	0.354	0.201	0.326	3.044	3.044
COL17A1	1.585	0.033	0	0.195	0.347	0.04	0.364	0.541	0.541
FABP7	0	0.007	0	0	0	0.003	0	0.006	0.006
MTS	0.059	1.345	0	1.558	0.98	1.377	0.275	1.753	1.753
RAC2	1.635	0.323	0.51	0.422	0.434	0.295	0.425	0.422	0.422
VIM	7.344	0.829	0.51	1.015	0.747	0.559	1.037	0.971	0.971
PDZK1P1	0.506	3.396	12.215	1.427	0.797	2.546	5.365	1.633	1.633
PI3	0.15	4.927	4.08	1.201	0.254	3.717	26.644	3.485	3.485
HES1	1.111	0.9	1.375	0.669	5.085	1.282	0.866	0.804	0.804
SOC1	1.724	0.139	0.448	0.161	0.353	0.129	0.193	0.233	0.233
MME4	0.913	0.491	0	0.398	0.485	0.633	0.226	0.584	0.584
IFIT3	1.324	1.131	1.325	0.756	0.758	1.501	1.468	1.068	1.068
LRX1B	0.652	5.085	0.766	4.913	3.455	6.361	5.67	4.253	4.253
RHAG1	1.837	2.748	0.882	0.367	4.166	0.312	0.31	1.309	1.309
ITF	0.39	0.576	0.255	0.696	0.353	1.29	0.315	1.084	1.084
ALDH3A1	2.534	6.453	2.741	7.353	5.391	10.418	4.816	5.703	5.703
CLU	2.603	6.659	0	4.952	4.376	7.605	2.769	1.931	1.931
CCML1	1.343	0.964	6.753	0.741	2.194	0.964	0.603	0.703	0.703
ELC42	0.362	0.145	0	1.277	0.847	0.743	1.894	1.298	1.298
CRABP2	0.882	0.971	0	0.509	0.425	0.911	5.799	2.393	2.393
MTHFD2	1.242	0.448	0.233	0.259	0.517	0.34	0.556	0.605	0.605

Gene	Neoplastic AG1	Neoplastic AG2	Neoplastic Basal1	Neoplastic Basal2	Neoplastic Basal3	Neoplastic Basal4	Neoplastic Basal5	Neoplastic KIPK1	Neoplastic KIPK2
TNFAIP3	6.462	2.793	1.89	2.493	0.715	2.904	2.301	1.021	1.012
IGKV3-20	2.888	139.076	0.37	0.536	0.159	0.431	0.568	0.743	0.1
TNFRM45A	0.497	0.136	0.242	0.158	0.027	0.303	0.181	0.245	0.296
MUC16	0.464	0.321	0.096	0.106	0.029	0.131	0.122	0.085	0.054
ATF3	4.088	1.972	2.993	1.494	0.648	1.065	2.23	0.588	8.587
KRT7	0.338	0.355	0.371	0.658	0.323	0.489	0.344	0.218	0.444
CXCL17	0.097	0.329	0.056	0.071	0.02	0.268	0.05	0.137	0.097
EGR1	4.92	2.325	5.015	2.415	1.514	1.099	3.397	2.053	16.007
KRT14	1.449	0.46	1.607	1.495	0.224	0.252	4.63	8.445	7.806
CXCL2	0.271	0.507	0.507	0.627	0.688	0.587	0.479	0.348	0.561
CXCL13	1.801	3.134	1.752	5.446	1.104	3.523	3.436	1.827	2.449
MYC	1.253	0.877	1.214	1.546	5.825	0.86	0.923	0.767	1.979
MUC20	0.278	0.071	0.089	0.069	0.681	0.111	0.056	0.131	0.07
BCAM	2.047	0.957	2.73	1.402	0.537	1.092	1.738	7.185	5.87
IGKV4-1	2.638	25.927	0.221	0.417	0.115	0.474	0.51	0.545	0.208
MYD88	0.298	0.115	0.374	0.461	0.31	1.488	0.307	0.17	0.194
SYNPO2	5.789	2.502	2.31	2.068	0.871	1.837	1.76	1.753	2.065
IGFBP7	0.742	0.714	1.224	1.178	0.223	1.116	1.017	4.364	1.804
MUC20-CT1	1.647	0.584	0.471	0.429	0.57	0.442	0.386	0.441	0.403
IGKV3-11	6.12	40.653	0.16	0.943	0.105	0.282	0.543	0.343	0.113
ZNF322	0.507	0.411	0.442	0.34	0.23	0.244	0.413	0.453	0.363
KYNU	0.507	0.222	0.162	0.297	1.511	0.477	0.216	0.105	0.103
GERM	1.361	0.742	1.037	0.89	0.396	0.215	0.975	0.475	0.898
IGLV2-23	1.108	21.321	0.163	0.243	0.066	0.255	0.328	0.501	0.059
S100P	0.071	0.104	0.044	0.02	0.009	0.132	0.063	0.028	0.056
CFB	0.295	0.371	0.29	0.29	0.146	0.194	0.258	0.346	0.313
MUC4	0.246	0.163	0.051	0.07	0.052	0.155	0.064	0.097	0.128
MUC7	1.077	0.829	0.767	0.934	0.902	1.003	0.59	0.771	0.523
FAM18D	0.023	0.049	0.021	0.023	0.014	0.034	0.012	0	0.005
IKK α	3.794	55.327	0.36	0.787	0.112	2.55	0.375	0.414	0.206

Gene	Neoplastic: MPB3	Neoplastic: PB1	Neoplastic: PB2	Neoplastic: SPB1	Neoplastic: SPB2	Neoplastic: SPB3	Neoplastic: SPB4	Neoplastic: SPB5	Non-neoplastic: AB1	Non-neoplastic: AB2
TNFAIP3	0.324	1.489	2.187	0.7	1.149	0.475	1.942	1.11	6.477	1.328
IGKV3-JO	0.204	0.46	2.968	0.682	5.53	0.442	0.447	0.247	20.431	125.371
TNFRM45A	1.604	0.174	0.153	0.65	0.844	0.758	0.903	4.76	0.306	0.236
MLUC16	0.014	0.048	0.055	0.637	0.705	4.757	0.316	0.139	8.871	0.508
ATF3	0.11	1.13	1.696	0.603	12.57	1.311	2.541	4.578	2.764	1.602
KRT17	1.924	0.474	0.551	0.585	0.361	2.538	0.71	0.194	1.301	0.354
CXCL17	0.015	0.051	0.039	0.587	0.402	4.128	0.406	1.157	1.364	1.406
ESR1	1.041	2.481	2.638	0.549	13.274	1.772	3.053	2.496	5.123	2.472
KRT14	5.377	0.752	1.382	0.543	0.659	0.008	4.363	0.538	1.967	2.72
CKS2	0.24	4.446	1.102	0.515	0.696	0.316	0.441	0.509	0.178	0.366
CXCL13	0.164	1.893	2.108	0.513	0.502	0	2.295	0.108	0.805	0.686
MYC	1.665	1.179	1.333	0.508	1.671	0.173	0.869	1.468	0.665	0.678
MLUC20	0.015	0.065	0.055	0.504	0.403	3.653	0.215	0.579	1.166	0.186
BCAM	0.941	1.557	1.58	0.497	0.904	0.488	0.705	0.32	1.536	1.439
IGKV4-1	0.064	0.247	0.872	0.484	0.111	0.353	0.29	0.237	3.489	6.801
MAP7	7.631	0.985	0.393	0.484	0.151	1.427	2.346	0.381	0.385	0.196
SYNP02	1.053	1.596	1.788	0.46	0.932	0.3	0.637	0.551	1.464	0.325
IGFBP7	2.083	0.914	0.921	0.459	0.561	0.126	0.789	0.406	1.398	2.221
MLUC20-OT1	0.24	0.434	0.484	0.442	0.462	0.764	0.345	1.052	1.441	0.248
IGKV3-11	0.26	0.266	0.535	0.441	0.084	0.077	0.382	0.283	2.907	41.103
ZNF322	0.345	0.181	0.372	0.437	0.456	0.244	0.158	0.591	0.286	0.187
KYNU	0.166	0.242	0.209	0.424	0.429	0.786	0.441	0.892	0.493	0.162
GEM	6.95	0.565	0.765	0.421	1.107	0.056	1.134	0.15	0.973	0.209
IGLV2-23	0.082	0.09	1.082	0.418	0.196	0.15	0.228	0.398	2.154	34.371
S100P	0.007	0.018	0.021	0.412	0.274	3.869	0.492	0.316	1.054	0.451
CFB	0.261	0.146	0.177	0.403	0.549	1.564	0.873	0.235	1.602	0.674
MLK4	0	0.063	0.06	0.39	0.342	1.924	0.144	0.187	5.155	0.68
MCM7	0.661	2.35	2.533	0.387	0.442	0.358	0.273	1.075	0.564	0.479
FAM130	0.004	0.015	0.014	0.379	0.156	2.034	0.133	0.201	1.294	0.561
IGKC	0.151	0.727	1.476	0.371	0.417	0.159	0.33	0.355	4.887	49.592

[illegible]

Gene	Non-neoplastic: G00612	Non-neoplastic: G00613	Non-neoplastic: G00614	Non-neoplastic: G00615	Non-neoplastic: G00616	Non-neoplastic: G00617	Non-neoplastic: G00618	Non-neoplastic: G00619	Non-neoplastic: G00620	Non-neoplastic: G00621	Non-neoplastic: G00622	Non-neoplastic: G00623	Non-neoplastic: G00624
TP53	0.217	0.221	0.561	1.075	0.051	0.748	0.275	0.535	0.272	0.234			
BRCA1	1.394	0.974	3.171	2.123	1.066	1.309	0.82	0.432	0.546	0.19			
TP53	0.145	0.097	0.055	0.175	0.183	0.032	0.171	0.172	0.298	0.12			
BRCA1	0.152	0.089	0.053	0.084	0.104	0.001	0.104	0.06	0.206	0.086			
ATM	4.527	1.083	2.23	1.086	1.459	3.437	0.651	8.784	0.368	0.312			
BRCA1	21.504	5.716	26.311	14.527	2.704	18.589	0.206	0.109	1.137	0.773			
BRCA1	18.458	12.458	16.007	10.375	3.23	0.123	0.714	0.577	1.489	2.531			
BRCA1	2.425	0.697	1.411	1.585	1.59	2.041	2.291	21.293	1.719	0.391			
BRCA1	0.509	0.034	0.173	0.022	12.32	0	20.069	25.309	23.544	18.336			
BRCA1	0.055	0.01	0.107	0.063	0.075	0.215	0.223	0.559	0.293	4.67			
BRCA1	0.017	0.032	0.106	0.145	0.14	0.021	0.379	0.036	0.046	1.638			
BRCA1	0.257	0.020	0.297	0.106	0.005	0.159	0.092	3.434	0.388	0.913			
BRCA1	0.174	0.023	1.371	0.444	0.261	13.266	0.167	0.154	0.185	0.706			
BRCA1	0.091	0.054	0.342	0.24	11.364	0.303	14.157	11.299	7.316	2.43			
BRCA1	1.123	0.391	10.983	2.094	1.289	0.854	0.451	0.854	0.824	0.063			
BRCA1	3.343	0.085	25.033	0.343	2.35	0.254	0.28	0.653	0.331	0.337			
BRCA1	0.103	0.01	0.018	0.102	0.342	0	0.53	0.534	0.122	0.337			
BRCA1	0.129	0.072	0.319	0.275	9.464	0.326	13.214	11.118	15.979	2.302			
BRCA1	0.357	0.189	0.407	0.78	0.503	1.379	0.355	0.397	0.318	0.419			
BRCA1	0.389	0.163	0.337	0.558	0.884	0.323	2.097	0.21	0.403	0.011			
BRCA1	0.203	0.166	0.306	0.26	0.271	0.365	0.352	0.276	0.322	0.209			
BRCA1	0.021	0.024	0.265	0.251	0.361	0.236	0.224	0.183	0.183	0.09			
BRCA1	0.01	0.010	0.078	0.114	0.134	0.025	0.17	0.525	0.204	0.073			
BRCA1	1.503	0.465	5.549	7.887	2.292	0.429	0.353	0.259	0.567	0.038			
BRCA1	0.105	0.052	2.82	2.042	0.13	0	0.028	0.206	0.611	0.03			
BRCA1	1.038	0.162	12.403	3.192	1.169	1.116	1.301	0.242	0.023	0.597			
BRCA1	0.073	0.036	1.379	0.673	1.314	0.209	0.094	0.611	1.091	2.057			
BRCA1	0.123	0.036	0.135	0.146	0.213	0.29	0.3	0.26	1.213	1.057			
BRCA1	4.435	4.071	3.714	6.822	0.398	0	0.012	0.015	0.119	0.387			
BRCA1	0.543	0.276	1.863	1.273	1.466	0.362	1.083	0.256	0.866	0.278			

Gene	Non-neoplastic: P62	Non-neoplastic: SP8	Globet transition2	Non-neoplastic: SP8.1	Non-neoplastic: SP6.1	Non-neoplastic: SP6.3	Non-neoplastic: SP8.6	Non-neoplastic: SP8.8
TM6SF3	4.482	0.526	1.485	0.153	0.443	0.2	0.262	0.185
IGKJ2-20	2.348	1.244	0.116	1.411	1.079	0.709	0.708	0.637
TMEM45A	0.118	1.73	0	0.552	0.728	1.013	0.527	0.42
MLC16	0.28	11.54	5.634	2.369	3.458	5.502	1.436	1.002
ATF5	1.673	0.939	10.471	0.279	8.627	1.345	0.944	0.887
PRF7	0.234	12.355	8.788	2.111	1.396	7.155	2.212	1.344
CXCL17	0.568	15.908	9.903	5.414	5.098	11.106	3.6	3.504
ESR1	2.822	0.93	0	1.019	11.38	1.695	1.145	1.977
KRT14	25.598	0.016	1.119	2.237	7.652	0.378	110.916	27.85
CXCL13	1.593	0.368	0.126	0.284	0.833	0.269	0.481	1.334
RPLC	1.145	0.191	0	0.516	2.504	0.412	0.502	0.921
MLC20	0.253	3.591	2.147	1.187	1.314	2.433	0.911	0.641
PCAM	3.379	0.188	0.11	0.915	1.35	0.533	0.533	1.047
IGKX4.1	1.616	6.021	0.627	0.841	0.332	0.432	0.551	9.41
KIMPT	0.314	0.821	0	0.53	0.531	0.958	0.361	0.162
SYMPD2	1.09	0.05	0	0.045	0.111	0.077	0.074	0.031
IGFBP7	2.776	0.286	0	1.752	1.985	1.26	1.225	2.018
MLC20-C71	0.386	0.639	0	0.521	0.484	0.524	0.481	0.288
IGKX3.11	1.719	0.598	0	0.589	0.218	0.852	0.291	0.287
ZNF312	0.389	0.279	0	0.295	0.272	0.34	0.247	0.176
KYNU	0.897	0.649	0.777	0.331	0.323	0.47	0.549	0.388
GRM	0.173	0.176	0	0.049	0.26	0.191	0.158	0.046
RGLY2-23	0.417	3.252	0.51	0.389	0.27	0.231	0.505	0.185
S100P	0.048	17.045	46.144	1.215	1.948	6.558	7.944	1.702
CFB	0.805	6.922	3.19	0.773	0.712	2.009	1.02	0.551
MLC3	0.525	6.873	14.41	2.951	2.918	4.329	5.755	2.278
MCM7	2.433	0.162	3.255	0.171	0.268	0.179	0.133	1.472
FABP3D	0.033	6.289	2.851	2.56	2.462	6.816	1.653	0.388
IKKC	0.443	2.038	0.86	0.561	4.285	0.315	0.797	0.334

Gene	Neoplastic AB1	Neoplastic AB2	Neoplastic Basal1	Neoplastic Basal2	Neoplastic Basal3	Neoplastic Basal4	Neoplastic Basal5	Neoplastic MPB1	Neoplastic MPB2
PRSS13	0.18	0.271	0.12	0.197	0.021	0.448	0.157	1.08	0.689
CH3L2	0.055	0.16	0.182	0.397	0.026	0.028	0.323	0.101	0.047
AC019349.1	0.028	0.064	0.154	0.079	0.032	0.032	0.117	0.085	0.168
SRGN	0.901	2.274	0.262	1.25	0.088	5.284	2.073	0.402	0.175
CtDC190	0.284	0.04	0.138	0.099	0.013	0.072	0.064	0.077	0.121
ADAM28	0.556	0.401	0.243	0.218	0.356	0.168	0.32	0.295	0.167
SERPINF1	0.946	0.697	1.153	0.942	0.116	3.297	0.328	1.907	1.796
C18orf33	0.071	0.109	0.091	0.067	0.008	0.156	0.079	0.33	0.22
IGLV3-19	1.527	20.314	0.292	0.326	0.045	0.176	0.22	0.605	0.021
POU2F3.1	2.178	0.949	0.479	0.441	0.268	0.537	0.415	0.393	0.367
BUCAL2	0.304	0.045	0.108	0.094	0.049	0.177	0.089	0.099	0.068
PTPRCAP	0.316	1.101	0.068	0.132	0.073	0.138	2.4	0.057	0.061
IGLV3-21	0.843	30.091	0.091	0.204	0.044	0.185	0.2	0.518	0.017
RASSF6	0.523	0.478	0.296	0.284	0.713	0.465	0.274	0.505	0.389
FOSB	3.886	1.731	1.137	0.447	0.273	0.368	1.112	0.319	3.768
IFITM2	0.361	1.356	0.125	0.246	0.089	0.587	1.939	0.481	0.324
SFRP1	0.272	0.21	0.464	0.766	0.165	0.716	0.251	0.21	0.228
CD3E	0.218	0.996	0.053	0.091	0.055	0.092	1.778	0.055	0.061
KBT4	0.016	0.046	0.003	0.004	0.068	0.001	0.012	0.019	0
IL1R2	0.822	0.403	1.092	0.997	0.517	0.725	1.276	4.766	5.47
FYB1	0.265	0.931	0.15	0.203	0.114	3.107	1.044	0.153	0.112
CD52	0.244	0.738	0.065	0.099	0.078	0.128	1.932	0.085	0.039
CORD14	0.345	0.973	0.096	0.171	0.14	0.884	1.758	0.15	0.097
HMEB2	0.351	0.737	0.233	0.356	0.158	1.093	0.366	0.198	0.172
AC022509.1	0.588	0.218	0.335	0.356	0.108	0.295	0.23	0.204	0.213
PTPRC	0.342	1.592	0.077	0.122	0.083	0.141	1.357	0.066	0.055
TYMS	0.245	0.653	0.304	0.442	0.469	0.472	0.301	0.483	0.183
IGKV1-5	1.918	80.496	0.666	1.064	0.131	0.438	0.323	0.311	0.285
CXCL14	3.272	0.432	5.24	3.588	15.106	0.261	4.85	13.42	27.962
PHL2	0.131	0.1	0.24	0.126	0.067	0.073	0.201	0.706	0.719

Gene	Neoplastic: MPB.1	Neoplastic: PB1	Neoplastic: PB2	Neoplastic: SPB1	Neoplastic: SPB2	Neoplastic: SPB3	Neoplastic: SPB4	Neoplastic: SPB5	Non-neoplastic: AB1	Non-neoplastic: AB2
PRSS23	0.06	0.079	0.133	0.36	0.175	0.919	0.26	0.063	2.662	1.311
CH3L2	0.371	0.052	0.055	0.358	0.981	0.732	11.418	0.219	0.194	0.12
AC019349.1	0	0.068	0.105	0.328	0.838	0.059	0.301	0.084	0	0.043
SRGN	0.178	0.838	1.019	0.328	0.425	0.143	0.505	0.259	0.92	3.44
CDC190	0.004	0.072	0.083	0.314	0.149	0.684	0.097	0.478	0.178	0.002
ADAM28	0.088	0.125	0.161	0.306	0.221	0.936	0.455	0.356	1.628	0.36
SERPINF1	0.417	1.002	0.912	0.303	0.539	0.47	0.542	0.089	0.779	1.089
C15orf93	0.019	0.04	0.039	0.289	0.383	1.314	0.243	0.19	0.988	0.604
IGLV3-19	0.03	0.085	0.206	0.283	0.066	0.128	0.131	0.162	0.698	2.66.497
PDIRJ3.1	0.295	0.483	0.602	0.281	0.296	0.254	0.348	0.498	1.143	0.423
BICD1	0.046	0.071	0.083	0.279	0.19	0.854	0.21	0.27	1.186	0.083
PTPRCAP	0.081	0.148	0.151	0.273	0.126	0.078	0.111	0.084	0.642	1.662
IGLV3-21	0.105	0.468	1.004	0.268	0.084	0.309	0.089	0.214	1.2	142.799
RASSF6	0.141	0.192	0.2	0.253	0.306	0.626	0.128	0.34	0.972	0.828
FOSB	0.392	0.531	0.809	0.238	4.106	0.512	1.186	2.445	4.529	1.436
HITM2	0.155	0.223	0.296	0.223	0.175	0.686	0.197	0.081	1.735	1.879
SFRP1	5.505	0.635	0.539	0.211	0.143	0.135	0.046	0.295	0.014	0.013
CDSE	0.091	0.113	0.113	0.209	0.107	0.044	0.111	0.062	0.342	1.242
KRT4	0	0.007	0.002	0.206	0.089	1.778	0.062	0.076	0.259	0.06
IL1K2	0.144	0.593	0.842	0.202	0.235	0.017	1.797	0.039	0.232	0.371
F6B1	0.055	0.219	0.191	0.201	0.172	0.158	0.178	0.075	0.95	0.907
CD52	0.039	0.113	0.117	0.195	0.147	0.07	0.105	0.049	0.574	1.303
CORO1A	0.074	0.255	0.243	0.191	0.205	0.126	0.129	0.041	0.892	2.229
HMGB2	0.285	7.116	2.332	0.187	0.209	0.163	0.119	0.552	0.403	0.889
AC02509.1	4.347	0.265	0.327	0.179	0.223	0.069	0.357	0.332	0.195	0.153
PTPRC	0.053	0.111	0.164	0.139	0.139	0.077	0.234	0.069	2.228	1.886
TYMS	0.71	3.48	2.52	0.177	0.062	0.013	0.073	0.81	0.368	0.459
IGRV1-5	0.061	0.136	2.896	0.173	0.249	0.051	0.181	0.143	18.574	320.189
CXCL14	0.354	2.046	2.593	0.172	0.642	0.305	2.946	0.088	1.534	0.189
PHL2	0.801	0.175	0.138	0.169	0.117	0.114	0.721	0.04	0.678	0.951

Item	Non-metals: Base1	Non-metals: Base2	Non-metals: Base3	Non-metals: Base5	Non-metals: Chloride1	Non-metals: Chloride2	Non-metals: Chloride3	Non-metals: Chloride4	Non-metals: Chloride5
PH2527	0.631	0.49	0.51	0.404	0.591	0.525	0.263	1.157	1.045
CH942	0.027	0.683	0.071	0.156	0.012	0.053	0.028	0.017	0.111
AC013346.1	0.2	0.136	0.073	0.115	0	0	0.006	0	0.063
SH35N	0.2	0.595	0.979	2.178	0.184	0.487	0.145	0.999	1.925
CDK130	0.005	0.004	0	0.001	2.002	3.363	2.394	0.092	0.092
ADAM28	0.899	0.379	0.03	0.441	0.077	0.364	0.577	0.185	0.812
SD09873	0.016	0.536	0.496	1.046	0.65	0.532	0.902	0.278	0.596
CH30413	3.403	0.639	0.861	0.393	2.444	0.482	3.399	2.223	0.882
GLV3-13	0.094	15.25	0.085	0.487	0.077	0.525	0.163	1.106	1.116
PO14213.1	0.339	0.766	0.227	0.369	0.254	1.36	0.421	0.135	0.54
GL13.1	0.131	0.053	0.152	0.145	1.204	2.053	0.943	1.036	0.684
PP99240	0.183	0.383	0.295	0.69	0.795	0.897	0.323	1.205	1.916
GLV3-21	0.122	0.222	0.063	0.16	0.172	1.312	0.509	2.078	7.475
HS5946	0.779	0.259	2.769	0.305	0.543	0.359	0.395	0.278	0.535
FG9	1.766	0.878	0.807	0.583	0.343	1.884	1.041	0.636	7.485
PT1602	3.394	1.236	0.322	0.405	1.734	0.976	2.373	1.329	1.836
PH01	0.049	0.151	0.069	0.018	0.503	0	0.001	0	0.159
CD94	0.15	0.266	0.153	0.134	0.331	0.391	0.082	0.913	1.34
K973	0.447	0.029	0	0.019	0.243	0.086	0.773	0.173	0.887
GL92	1.784	0.971	0.039	0.105	0.04	0.034	0.06	0	0.125
PH01	0.186	0.588	0.143	7.952	0.128	0.363	0.287	0.523	1.07
CD32	0.139	0.417	0.043	0.338	0.212	0.46	0.206	0.973	1.312
CD921A	0.432	1.359	1.359	1.03	0.235	0.886	0.233	1.156	1.808
HAN282	0.203	0.432	0.365	0.83	0.262	0.537	0.367	0.703	0.955
AC022506.1	0.021	0.265	0	0	0	0	0	0	0.372
PT18C	0.164	0.246	0.134	0.89	0.26	0.706	0.264	0.663	2.338
TY026	0.299	0.489	0.233	0.237	0.506	0.616	0.329	0.18	0.243
GL001.5	0.441	0.307	0.805	0.628	0.516	2.875	0.425	7.373	2.11
CK114	7.986	25.63	60.743	0.978	0.019	0.026	0.648	0.114	2.021
PH2	2.319	0.412	0.454	0.046	0.03	0.058	0.434	0.058	0.486

[illegible]

Gene	Non-neoplastic: PB2	Non-neoplastic: SP8	Gabbet transition1	Non-neoplastic: SP6	Gabbet transition2	Non-neoplastic: SPB3	Non-neoplastic: SPB2	Non-neoplastic: SPB3	Non-neoplastic: SPB4	Non-neoplastic: SPB5
PHS23	2.129	11.921	0	9.168	6.309	11.216	1.476	2.112		
CH3L2	0.058	0.037	0	0.04	0.027	0.135	0.512	0.009		
AL2119349.1	0.128	0.018	1.321	0.235	0.085	0.126	2.239	0.699		
SRGN	1.607	0.23	0.51	0.184	0.306	0.23	0.179	0.202		
CCDC190	0.04	0.326	0	0.105	0.038	0.197	0.095	0.152		
MDM178	0.255	1.634	1.541	1.326	1.388	1.545	0.755	0.523		
SEPM41	1.431	1.257	0	2.323	2.578	2.07	0.645	1.752		
C19orf33	1.191	5.297	4.405	6.665	6.236	6.701	6.411	3.931		
GLIS-19	0.363	0.31	0.116	0.05	0.075	0.133	0.223	0.024		
POLR2B.1	0.762	0.248	0	0.156	0.26	0.162	0.24	0.135		
BRD4.2	0.017	1.315	1.212	0.47	0.7	0.826	0.928	0.238		
PTPRCAP	0.62	0.388	0	0.344	0.244	0.322	0.212	0.283		
SLY5-21	0.246	0.362	1.321	0.208	0.074	0.15	0.486	0.074		
PA356	0.554	0.266	0.116	0.556	0.323	0.411	0.129	0.565		
FO58	1.62	1.146	1.146	0.459	7.143	1.362	0.992	1.154		
PTTM2	1.498	2.114	0.255	1.94	1.399	2.327	1.268	1.746		
SPBP1	0.07	0.017	0	0.02	0.012	0.033	0.002	0.063		
CDSE	0.373	0.242	0	0.206	0.212	0.216	0.102	0.159		
KRT3	0.32	2.53	28.575	1.557	3.474	3.458	10.692	4.688		
ELB2	1.295	0.019	0	0.127	0.136	0.075	0.386	0.331		
RYB1	0.861	0.215	0.255	0.296	0.265	0.237	0.34	0.357		
ITK2	0.485	0.323	1.021	0.243	0.134	0.202	0.329	0.274		
CCND1A	1.158	0.342	1.276	0.329	0.34	0.241	0.199	0.213		
HMG2	3.59	0.246	0.116	0.389	0.482	0.311	0.195	1.988		
AC022509.1	0.112	0.004	0	0.006	0.008	0.004	0.005	0.002		
PTPRC	2.328	0.282	0.51	0.196	0.277	0.166	0.185	0.266		
TNKS	2.722	0.037	0	0.26	0.145	0.071	0.043	2.233		
SPX1.5	1.189	2.067	0.116	0.39	0.357	0.187	0.129	0.144		
TXN14	9.265	0.96	0.58	0.401	0.166	1.29	1.378	1.04		
PH2	1.084	1.022	2.671	1.29	1.648	1.292	0.884	1.343		

Gene	Neoplastic A21	Neoplastic A22	Neoplastic A23	Neoplastic Basal1	Neoplastic Basal2	Neoplastic Basal3	Neoplastic Basal4	Neoplastic Basal5	Neoplastic A281	Neoplastic A282
IER3	0.739	0.917	0.301	0.702	0.15	0.973	0.791	0.267	0.721	0.721
NREP	0.744	0.656	0.8	0.589	3.133	0.27	0.47	0.398	0.883	0.883
SCNN1B	0.209	0.17	0.213	0.073	0.019	0.03	0.162	0.117	0.361	0.361
IGKV2-30	1.201	20.04	0.04	0.206	0.037	0.213	0.034	1.041	0.036	0.036
MMP10	0.444	0.222	0.48	0.153	0.011	0.031	0.594	2.862	1.482	1.482
IGHA1	1.755	23.603	0.159	0.573	0.07	0.233	0.345	0.112	0.073	0.073
ZG16B	0.032	0.068	0.058	0.046	0.05	0.081	0.05	0.154	0.135	0.135
DEFB1	0.039	0.042	0.041	0.027	0.001	0.001	0.073	0.11	0.04	0.04
CAPS	0.131	0.071	0.187	0.108	0.047	0.2	0.067	0.143	0.18	0.18
GOLM1	0.386	0.119	0.311	0.356	0.305	0.698	0.278	0.22	0.12	0.12
OLFML4	0.379	0.029	0.022	0.038	0.103	0	0.015	0.029	0.027	0.027
RARBES1	0.063	0.152	0.013	0.039	0.014	0.082	0.015	0.03	0.029	0.029
RNASE1	0.119	0.066	0.03	0.045	0.015	2.46	0.082	0	0.013	0.013
CD3D	0.153	0.621	0.038	0.069	0.044	0.043	1.21	0.026	0.022	0.022
SPIB	0.989	0.378	0.599	2.261	0.814	13.844	0.439	0.382	0.211	0.211
IGHV3-23	1.088	30.097	0.056	0.291	0.051	0.068	0.128	0.103	0.068	0.068
IGHV4-34	1.242	7.766	0.106	0.167	0.02	0.082	0.06	0.111	0.025	0.025
HEX6	0.049	0.031	0.085	0.115	0.73	0.222	0.09	0.033	0.038	0.038
CLCA4	0.102	0.015	0.036	0.033	0.002	0.006	0.028	0.026	0.04	0.04
CEACF141	4.531	0.646	1.697	1.156	0.587	0.837	0.446	0.349	0.637	0.637
CLDN3	0.689	0.131	0.415	1.136	0.342	7.744	0.475	0.09	0.039	0.039
TMC5	0.06	0.042	0.014	0.017	0.005	0.041	0.025	0.015	0.02	0.02
IGHV5-51	0.582	16.3	0.032	0.106	0.032	0.034	0.185	0.091	0.154	0.154
MTLE	0.171	0.187	0.163	0.481	0.511	0.324	0.221	0.178	0.307	0.307
RHCG	0.059	0	0.003	0.01	0.002	0.073	0.015	0	0.037	0.037
IGHV1-18	0.337	10.712	0.048	0.18	0.03	0.076	0.13	0.724	0.019	0.019
FOXSP	2.335	3.168	3.423	21.386	0.352	3.765	2.9	1.944	0.034	0.034
LXN	0.064	0.033	0.07	0.057	0.086	0.044	0.085	0.08	0.062	0.062
TK1	0.127	0.508	0.117	0.173	0.129	0.502	0.078	0.866	0.04	0.04
ANKRD36C	0.521	0.297	0.08	0.105	0.142	0.075	0.099	0.035	0.025	0.025

Gene	Neoclastic MPB.3	Neoclastic PB1	Neoclastic PB2	Neoclastic PB3	Neoclastic SPB.1	Neoclastic SPB.2	Neoclastic SPB.3	Neoclastic SPB.4	Neoclastic SPB.5	Non-neoclastic AB1	Non-neoclastic AB2
HER3	0.124	0.275	0.496	0.161	0.513	0.759	0.406	0.17	0.12	1.389	1.091
NREP	0.263	0.511	0.481	0.161	0.395	0.148	0.17	0.17	0.12	0.216	0.086
SCNN1B	0.018	0.059	0.086	0.161	0.358	0.31	0.089	0.165	0.165	0.157	0.146
IGKV2-30	0.014	0.033	0.846	0.154	0.059	0.068	0.09	0.078	0.078	4.207	35.327
MMP10	0.071	0.025	0.086	0.154	0.047	0.018	2.843	0.016	0.016	2.11	2.134
IGHA1	0.063	0.176	0.603	0.153	0.319	0.067	0.204	0.376	0.376	4.795	24.54
ZG18B	0.028	0.036	0.034	0.151	0.188	0.716	0.122	0.072	0.072	0.305	0.237
DEFB1	0.004	0.011	0.011	0.143	0.066	0.995	0.076	0.065	0.065	0.181	0.067
CAP5	0.18	0.081	0.086	0.142	0.164	0.2	0.071	0.081	0.081	0.381	0.118
GOLM1	0.193	0.433	0.388	0.141	0.219	0.641	0.197	0.19	0.19	0.395	0.167
OLFML4	0.005	0.049	0.058	0.137	0.328	0.153	0.71	1.995	1.995	0.017	0.075
RARRS1	0.232	0.025	0.029	0.137	0.005	1.234	0.043	0.107	0.107	1.111	0.294
RNAS1	0.003	0.062	0.055	0.122	0.015	0.592	0.033	0.618	0.618	0.151	0.25
CD3D	0.023	0.084	0.081	0.12	0.026	0.013	0.069	0.026	0.026	0.448	0.835
SPB	0.063	1.357	1.267	0.117	0.37	0.11	0.198	0.068	0.068	0.75	0.303
IGHV3-23	0.08	0.072	1.614	0.113	0.027	0	0.839	0.027	0.027	1.217	17.759
IGHV4-34	0	0.059	0.35	0.113	0.077	0	0.045	0.03	0.03	0.495	179.55
HE36	0.026	0.186	0.136	0.109	0.043	0.068	0.04	0.038	0.038	0.062	0.001
C1CA4	0.091	0.012	0.019	0.102	0.101	0.157	0.109	0.239	0.239	0.341	0.61
Ccorf141	0.153	0.658	1.21	0.095	0.161	0.058	0.41	0.051	0.051	0.594	0.117
CLDN3	0.042	0.892	0.994	0.094	0.462	0.195	0.136	0.044	0.044	0.434	0.086
TNKS	0.029	0.006	0.013	0.081	0.032	0.921	0.038	0.175	0.175	1.479	0.11
IGHV5-51	0.142	0.041	0.286	0.083	0.031	0.016	0.091	0.051	0.051	20.289	98.312
MT1E	0.162	0.254	0.303	0.077	0.194	0.584	0.074	0.067	0.067	0.465	0.356
RHCG	0	0.001	0.005	0.073	0.002	0	0.342	0.03	0.03	0.005	0
IGHV1-18	0.019	0.179	0.534	0.07	0.046	0.274	0.089	0.104	0.104	14.22	91.672
FXR3P	205.955	14.367	11.858	0.069	0.022	0	5.791	0.232	0.232	3.162	0.852
LNN	0.041	0.054	0.057	0.067	0.093	0.151	0.068	0.129	0.129	0.087	0.107
TK1	0.093	2.541	1.549	0.067	0.045	0.021	0.037	0.611	0.611	0.068	0.44
ANKRD36C	0.182	0.104	0.14	0.066	0.094	0.419	0.134	0.438	0.438	3.484	0.157

[illegible]

Gene	Non-neoplastic PB2	Non-neoplastic SPP6-Isaklet transition1	Non-neoplastic SPP6-Coblet transition1	Non-neoplastic SPP61	Non-neoplastic SPP62	Non-neoplastic SPP63	Non-neoplastic SPP64	Non-neoplastic SPP65
BR3	1.577	1.521	5.923	1.022	5.043	1.502	1.413	1.934
NRCP	0.34	0.02	0	0.019	0.004	0.076	0.011	0.002
SCNN1B	0.031	0.381	0.78	0.243	0.347	0.331	0.149	0.178
NRX2-30	0.371	0.234	0.56	0.313	0.151	0.085	0.253	0.027
MRPT0	1.087	0.307	0	0.09	0.582	0.236	1.602	0.174
NRX41	0.912	1.006	2.391	0.422	2.696	0.155	0.312	0.307
NRX41	0.153	3.065	8.783	0.545	0.427	1.107	0.357	0.611
2636B	0.267	1.172	0	2.628	1.383	1.163	2.315	2.21
TRF03	0.041	0.647	0	0.592	0.748	0.885	0.319	0.129
CAPS	0.196	1.518	0.355	0.251	0.212	0.61	0.316	0.175
GCM1	0.001	0.021	0	0.003	0	0	0.114	0
OLFM4	0.305	8.547	0	1.394	1.053	5.04	0.863	0.423
BAKRE51	0	1.041	0	0.023	0.078	0.416	0.221	0.021
RP45E1	0.41	0.161	0	0.161	0.128	0.383	0.048	0.059
CR30	0.683	0.035	0	0.026	0.013	0.011	0.032	0
SPP8	0.546	3.577	0	0.088	0.103	0.152	0.054	0.021
IGHV3-23	0.177	0.063	0	0.101	0.098	0.027	0.078	0.012
IGHV4-34	0.063	0.084	0	0.024	0.037	0.041	0.041	0.042
HE56	0.038	0.201	1.238	1.355	1.233	0.542	2.073	0.557
C1C44	0.192	0.012	0	0.017	0.012	0.014	0.041	0.035
C6orf141	0.132	1.584	0.766	0.082	0.543	0.455	0.341	0.054
CLDN5	0.03	2.477	0.698	0.345	0.379	0.858	0.657	0.195
TMCS	0.185	0.155	0	0.14	0.168	0.082	0.025	0.043
IGHV5-51	2.931	3.426	0	3.823	4.216	3.543	0.767	3.891
M11C	0.014	0.003	32.803	0.058	0.028	0.117	6.217	0.257
RHC3	0.384	0.308	0.233	0.189	0.178	0.212	0.184	0.046
IGHV2-18	0.293	0.05	0.255	0.008	0.002	0.012	0.032	0.055
FDC3P	0.065	0.307	0	0.372	0.309	0.426	0.082	0.184
LXN	1.753	0.032	0	0.054	0.205	0.083	0.139	2.084
TK1	0.907	1.913	3.583	0.293	0.405	0.568	1.119	0.251
ANKRD36C								

Gene	Neoplastic AG1	Neoplastic AG2	Neoplastic Basal1	Neoplastic Basal2	Neoplastic Basal3	Neoplastic Basal4	Neoplastic Basal5	Neoplastic AGP1	Neoplastic AGP2
MZB1	0.249	3.546	0.016	0.042	0.008	0.019	0.091	0.04	0.004
TMEM176A	0.141	0.305	0.154	0.154	0.162	0.52	0.131	0.182	0.063
VIMU1	0.044	0.039	0.005	0.005	0.002	0.002	0.008	0.009	0.014
ERN2	0.059	0.035	0.016	0.012	0.012	0.012	0.011	0.002	0.015
UNC02384	0.349	0.239	0.253	0.454	1.178	0.538	0.146	0.112	0.054
CEMPF	0.403	0.48	0.126	0.154	0.143	0.274	0.159	0.089	0.036
BPIFA1	0.283	0.361	0.005	0.024	0.018	0.04	0.002	0.002	0
BCL2A1	0.843	2.85	0.316	3.216	0.886	6.676	0.778	0.161	0.725
TRF3	0.06	0.311	0.087	0.068	0.255	0.047	0.105	0.046	0.014
GOLGA8A	0.305	0.163	0.081	0.058	0.036	0.092	0.074	0.081	0.091
PLEK	0.431	0.713	0.302	0.538	0.081	4.27	0.301	0.08	0.175
STAT4	0.029	0.103	0.01	0.043	0.296	0.045	0.001	0	0.009
KIT	0.546	0.321	0.084	0.147	1.503	0.143	0.048	0.056	0.111
SPRR3	0	0	0	0	0	0.003	0	0	0
PCLAF	0.137	0.335	0.103	0.133	0.162	0.318	0.065	0.061	0.085
CEACAM5	0.004	0	0	0	0	0	0	0	0
FAM193A	0.048	0.06	0.054	0.042	0.018	0.011	0.042	0.057	0.057
THCP2L1	0.03	0.016	0.013	0.013	0.023	0.021	0.018	0.021	0.017
C15orf48	0.186	0.477	0.05	0.09	0.252	0.425	0.19	0.015	0.013
MIA	0.106	0.032	0.074	0.36	0.18	0.881	0.125	0.059	0.054
BRCC3	0.224	0.131	0.091	0.101	0.069	0.047	0.06	0.063	0.101
CINB1	0.132	0.531	0.092	0.121	0.115	0.141	0.073	0.054	0.063
CDK1	0.109	0.218	0.066	0.076	0.098	0.084	0.078	0.084	0.051
DNAH3	0.046	0.009	0.029	0.018	0.002	0.026	0.031	0.039	0.028
IL33	0.012	0.022	0.042	0.009	0.016	0.001	0.062	0.238	0.101
AZGP1	0.036	0.001	0.021	0.072	1.243	0.048	0.007	0.014	0.019
MAL	0.036	0.062	0.023	0.058	0.011	0.058	0.156	0.016	0.047
TPPP3	0.013	0.028	0.023	0.024	0.017	0.131	0.01	0.028	0.04
CYP2F1	0.01	0	0.003	0.003	0.003	0	0.007	0.005	0.006
HCK	0.209	0.245	0.112	0.262	0.425	1.754	0.137	0.05	0.054

Gene	Neoplastic MPB3	Neoplastic PB1	Neoplastic PB2	Neoplastic SPB1	Neoplastic SPB2	Neoplastic SPB3	Neoplastic SPB4	Neoplastic SPB5	Non-neoplastic AB1	Non-neoplastic AB2
MTB1	0.019	0.026	0.111	0.063	0.037	0.005	0.039	0.064	0.676	6.805
TMEM176A	0.069	0.076	0.088	0.062	0.046	0.528	0.154	0.316	0.039	0.568
VMO1	0.021	0.006	0.004	0.062	0.005	0.303	0.003	0.038	0.972	0.776
ERH2	0	0.006	0.009	0.058	0.028	0.261	0.011	0.027	1.222	0.054
LRX12394	0.039	0.42	0.437	0.058	0.032	0.018	0.223	0.155	0.088	0.088
CENPF	0.12	4.257	1.534	0.055	0.038	0.074	0.024	0.398	0.353	0.265
BPRF1	0	0.011	0.017	0.054	0	0	0.008	0	1.642	0.028
BCL2A1	0.89	1.667	1.925	0.048	0.04	0.022	0.35	0.026	0.477	0.578
TFH3	0.005	0.054	0.073	0.048	0.048	1.124	0.044	0.216	1.015	0.269
GOLGA8A	0.051	0.058	0.068	0.047	0.087	0.052	0.053	0.048	0.629	0.276
PLEK	0.454	0.469	0.474	0.047	0.109	0.016	0.078	0.048	0.385	0.692
STAT1	0.007	0.026	0.011	0.047	0.004	1.916	0.063	0.002	0.11	0.09
KIT	0.019	0.119	0.128	0.044	0.116	0.379	0.076	0.021	0.416	0.03
SPRR3	0	0.002	0	0.044	0	0.128	0.505	0.003	0.107	0
PCLAF	0.024	1.94	1.303	0.042	0.033	0.025	0.029	0.367	0.086	0.464
CEACAM5	0	0.002	0	0.039	0.073	1.672	0.169	0.032	0.382	0.015
FAM183A	0.039	0.032	0.034	0.039	0.03	0	0.023	0.022	0.041	0.015
TFP211	0.038	0.022	0.023	0.039	0.039	0.14	0.007	0.012	0.693	0.013
C15orf48	0.015	0.052	0.102	0.038	0.019	0.437	0.812	0.295	0.338	0.56
MIA	4.774	0.25	0.196	0.038	0.059	0.344	0.09	0.013	0.038	0.11
DRC3	0.189	0.053	0.079	0.034	0.081	0	0.025	0.017	0.129	0.004
C10B1	0.068	3.16	0.788	0.033	0.026	0.14	0.071	0.18	0.075	0.329
CDK1	0.022	2.317	0.734	0.033	0.033	0.047	0.097	0.24	0.095	0.399
DNAH3	0.009	0.013	0.015	0.031	0.038	0.044	0.015	0.003	0.036	0.084
IL33	0.009	0.009	0.006	0.031	0.067	0	0.078	0.035	0.402	0.376
AZGP1	0.014	0.092	0.078	0.028	0.004	0.043	0.084	0.035	0.317	0.02
NAL	0.004	0.022	0.019	0.027	0.046	0.013	0.026	0.009	0.094	0.203
TPP3	0.003	0.016	0.018	0.027	0.052	0.049	0.013	0.006	0.087	0
CYP2F1	0	0.001	0.004	0.026	0.026	0.993	0.001	0.01	0.497	0.021
HCK	0.014	0.152	0.199	0.026	0.045	0.015	0.238	0.18	0.266	0.199

[illegible]

[illegible]

Gene	Non-neoplastic SPB	Goblet transition1	Non-neoplastic SPB	Goblet transition2	Non-neoplastic SPB1	Non-neoplastic SPB2	Non-neoplastic SPB3	Non-neoplastic SPB4	Non-neoplastic SPB5
MB31	0.128	0.169	0	0	0.018	0.029	0.022	0.059	0.202
TMEM176A	0.155	0.385	0	0	0.047	0.006	0.091	0.134	0.005
ANKK1	0.158	11.489	2.56	0	2.257	1.681	5.559	0.875	1.439
ERK2	0.036	1.87	0	0	0.567	0.734	0.952	0.509	0.187
UNC02394	0.556	0.021	0	0	0.003	0	0	0.026	0.135
CENPF	3.898	0.013	0	0	0.083	0.046	0.048	0.024	1.908
SPRY41	0.259	92.289	0	0	0.548	0.138	12.124	0.01	0.06
SC12A1	2.121	0.088	0	0	0.06	0.056	0.034	0.144	0.021
TF31	0.351	56.566	0	0	0.262	0.518	2.939	0.474	0.314
SC3.6A8A	0.244	0.455	0.11	0	0.169	0.244	0.295	0.099	0.094
PLEK	0.365	0.046	0	0	0.032	0.026	0.028	0.053	0.002
PAH	0.425	1.434	9.057	0	0.108	0.082	0.65	0.238	1.52
KIT	0.365	0.075	0	0	0.041	0.06	0.067	0	0.021
SPR3	0.002	1.329	303.143	0	0.159	0.029	0.59	19.346	1.181
PCAF	1.753	0.044	0	0	0.071	0.067	0.053	0.216	1.983
CEACAM5	0.002	5.894	29.167	0	0.442	0.205	1.241	1.997	0.549
FAM183A	0.024	0.084	0	0	0.084	0.048	0.041	0.019	0.016
PCP2L1	0.125	0.512	0	0	0.408	0.44	0.511	0.049	0.159
C15orf46	0.239	6.693	90.756	0	0.559	0.432	2.61	5.872	0.663
MA	0.154	0.354	0.582	0	0.106	0.047	0.252	0.582	0.025
PRC3	0.048	0.02	0	0	0.015	0.019	0.012	0.012	0.014
COR1	1.303	0.068	0	0	0.045	0.061	0.053	0.102	0.933
FOX1	1.125	0.058	0	0	0.033	0.053	0.039	0.13	0.635
UNAH3	0.001	0.003	0	0	0.018	0.017	0.016	0	0.008
IL33	0.695	0.349	0	0	1.67	1.231	0.646	0.53	0.818
AZGP1	0.027	0.522	0.372	0	0.031	0.006	0.154	0.071	0.124
MA1	0.014	0.06	64.018	0	0.019	0.021	0.016	0.383	0.012
PPP43	0.037	0.189	0	0	0.155	0.166	0.1	0.27	0.127
CYP2F1	0.008	5.164	0	0	0.398	1.203	2.274	0.06	0.056
CK	0.222	0.023	0	0	0.01	0.014	0.02	0.054	0.002

Gene	Neoplastic A61	Neoplastic A62	Neoplastic B2a1	Neoplastic B4a2	Neoplastic B2a3	Neoplastic B2a4	Neoplastic B2a5	Neoplastic A61	Neoplastic A62
CYR61	0.822	0.213	0.794	0.31	0.329	0.13	0.318	0.918	3.01
TOP2A	0.206	0.158	0.043	0.036	0.021	0.144	0.062	0.037	0.024
MUC5AC	0.095	0.325	0.002	0.006	0.013	0.001	0.012	0.001	0
DNAH11	0.345	0.045	0.102	0.075	0.033	0.049	0.09	0.064	0.09
LYZ	0.611	2.805	0.249	0.997	0.85	32.67	0.17	0.05	0.021
FGFR1	0.234	0.086	0.174	0.255	0.145	0.126	0.116	0.118	0.109
FSCA	0.052	0.067	0.007	0.008	0.004	0.009	0.002	0.001	0.01
SNTN	0.004	0	0.006	0.004	0.002	0.004	0.009	0.001	0.013
MUC5B	0.035	0.045	0.004	0.013	0.024	0.015	0	0.001	0
NUSAP1	0.221	0.274	0.062	0.078	0.054	0.124	0.051	0.015	0.026
RSPH1	0.044	0.025	0.03	0.049	0.052	0.08	0.026	0.016	0.013
DNAH5	0.036	0.038	0.008	0.011	0.018	0.004	0.007	0.006	0
BPIFB1	0.044	0.317	0.006	0.051	0.284	0.093	0.001	0	0
P1FO	0.018	0.004	0.015	0.017	0.01	0.007	0.006	0.012	0.032
PGC8	0.047	0.079	0.001	0.007	0.019	0.129	0	0	0
RP1	0.116	0.109	0.038	0.031	0.015	0.019	0.014	0.016	0.045
CFTR	0.001	0	0.001	0.001	0.005	0.023	0	0	0
CRYM	0.003	0.036	0.001	0.005	0.025	0.001	0	0	0
MK67	0.216	0.152	0.043	0.059	0.012	0.167	0.11	0.005	0.024
BIRC5	0.074	0.216	0.034	0.05	0.035	0.161	0.023	0.002	0.012
CFAP126	0.008	0.001	0.003	0.007	0.008	0.035	0.005	0	0
BARRES2	0.01	0.066	0.003	0.003	0.001	0.002	0.002	0	0.002
TNFRSF11B	0.438	0.161	0.04	0.281	0.205	8.762	0.072	0.004	0.007
UBE2C	0.082	0.336	0.044	0.089	0.043	0.153	0.026	0.016	0.021
PRB3	0.01	0	0.005	0.024	0.156	0.017	0	0.005	0.003
SCGB1A1	0.272	0.027	0.004	0.034	0.02	0.053	0.009	0.027	0
CDHR3	0.053	0.051	0.03	0.04	0.015	0.025	0.034	0.049	0.038
PRSS27	0.016	0.002	0.005	0.004	0.001	0.001	0.006	0	0.01
CLCNK8	0	0	0.001	0.001	0	0.002	0	0.005	0
DNAH12	0.009	0.008	0.005	0.005	0.004	0.001	0.001	0.002	0.006

Gene	Neoplastic t0P83	Neoplastic PB1	Neoplastic PB2	Neoplastic SPB1	Neoplastic SPB2	Neoplastic SPB3	Neoplastic SPB4	Neoplastic SPB5	Non-neoplastic AB1	Non-neoplastic AB2
CYR61	0.222	0.436	0.508	0.025	0.364	0.01	0.283	0.049	0.654	0.486
TOP2A	0.043	3.604	0.827	0.025	0.014	0	0.016	0.175	0.128	0.264
MUC5AC	0	0.004	0.003	0.021	0	0.018	0.001	0.003	4.935	0.024
DNAH11	0.078	0.043	0.081	0.02	0.047	0.024	0.039	0.01	0.123	0.002
LYZ	0.033	1.191	0.761	0.02	0.076	2.089	0.142	0.15	0.922	5.332
FCGR1	4.173	0.306	0.267	0.017	0.007	0.077	0.012	0.026	0.07	0.053
PSCA	0.014	0.007	0.004	0.017	0.008	0.73	0.018	0.011	0.865	0.009
SMTN	0.007	0.002	0.003	0.016	0.006	0	0.003	0.004	0.005	0.014
MUC5B	0.087	0.006	0.012	0.013	0.003	0.839	0.004	0.005	1.79	0.359
NUSAP1	0.028	2.385	0.768	0.012	0.014	0.016	0.016	0.126	0.212	0.54
KSP41	0.045	0.017	0.038	0.011	0.025	0.013	0.01	0.038	0.052	0.001
DNAH5	0.003	0.01	0.012	0.01	0.007	0.046	0.017	0.013	0.039	0.02
BPIFB1	0	0.032	0.022	0.009	0	0.844	0.094	0.007	2.335	0.392
P1FO	0.01	0.036	0.006	0.009	0.023	0	0.02	0.002	0.076	0
PIGR	0	0.01	0.003	0.009	0.02	2.181	0.011	0.014	1.544	0.138
RP1	0	0.023	0.03	0.009	0.029	0	0.011	0.004	0.04	0
CFTR	0.004	0.002	0.001	0.008	0	0.111	0.003	0	0.173	0
CRYM	0	0.012	0.007	0.008	0	0.152	0.003	0.001	0.19	0
NR67	0.036	2.703	0.978	0.008	0.023	0	0.016	0.243	0.531	0.324
BIRC5	0.025	2.194	0.733	0.006	0.032	0	0.006	0.189	0.095	0.059
CFAP126	0.011	0.004	0.002	0.006	0.01	0	0.001	0.001	0.006	0.099
KARRES2	0	0.002	0.007	0.006	0.007	0	0.004	0.003	0.021	0.396
TNFRSF11B	0.069	0.575	0.368	0.006	0	0	0.116	0.004	0.009	0.029
UBE2C	0.036	5.892	1.175	0.006	0.009	0	0.026	0.192	0.238	0.407
PRB3	0	0.02	0.01	0.005	0	0	0.025	0.001	0	0
SCGB1A1	0.005	0.036	0.017	0.005	0.001	1.458	0.012	4.58	3.699	0.184
CDHR3	0.176	0.042	0.036	0.004	0.029	0.018	0.025	0.007	0.02	0.059
PRSS17	0.003	0.004	0.003	0.004	0.008	0.013	0.02	0.006	0.006	0.003
CLCA4	0	0.001	0.001	0.003	0	0	0	0	0	0
DNAH12	0	0.004	0.006	0.003	0.003	0	0	0.004	0.005	0

[illegible]

[illegible]

Gene	Non-neoplastic SPB	Coeliac transition1	Non-neoplastic SPB-Coeliac transition2	Non-neoplastic SPB1	Non-neoplastic SPB2	Non-neoplastic SPB3	Non-neoplastic SPB4	Non-neoplastic SPB5
CYR61	0.83	0.419	0	0.097	1.599	0.475	0.298	0.556
ICP2A	2.618	0.004	0	0.013	0.084	0.018	0.023	1.052
MUC5AC	0.086	36.515	0	0.02	0.167	0.166	0.087	0.004
DNAH11	0.029	0.024	0	0.005	0.009	0.036	0.015	0.005
LYZ	1.535	3.179	2.552	0.184	0.378	0.476	0.43	2.335
FCBPL1	0.056	0.069	0	0.051	0.098	0.09	0.015	0.032
PSCA	0.007	9.554	29.962	0.116	0.178	1.114	0.493	0.98
SNRN	0.001	0.021	0	0.053	0.003	0.003	0.003	0
MUC5B	0.014	11.532	0.255	0.057	0.122	0.191	0.075	0.065
MUCAP1	2.46	0.012	0	0.014	0.01	0.014	0.002	0.617
KCRA1	0.006	0.076	0	0.019	0.016	0.07	0.022	0.007
DNAH5	0.004	0.081	0	0.019	0.007	0.022	0.005	0.004
BRP3	0.454	74.768	2.808	0.273	0.151	2.172	0.186	1.182
HFO	0.004	0.016	0	0.015	0.006	0.009	0	0
HGR	0.11	20.27	1.225	0.6	0.794	4.62	1.049	0.276
BP1	0.005	0.013	0	0.009	0.024	0.005	0.045	0.003
CFTR	0	0.089	0	0.032	0.068	0.116	0.024	0.013
CRPM	0	2.142	0.135	0.106	0.227	1.11	0.131	0.031
MIR57	3.262	0.003	0.56	0.023	0.058	0.025	0.03	1.365
BRP5	0.94	0.008	0	0.013	0.061	0.018	0.029	0.68
CFAP126	0	0.016	0	0.003	0.007	0.011	0	0
BARP52	0.187	0.197	0	0.083	0.047	0.153	0.011	0.29
TNFRSF11B	0.236	0.008	0	0.002	0	0	0.004	0.086
URFX	1.309	0.027	0	0.008	0.011	0.039	0.002	0.923
PRR3	0.216	0.997	2.042	0.036	0.015	0.046	0.105	0.77
SCGB1A3	0.303	167.041	0.952	2.674	7.047	26.195	1.219	1.296
CDHR3	0.019	0.006	0	0.005	0.001	0.009	0.004	0
PRSS27	0.001	0.108	20.243	0.06	0.018	0.057	0.951	0.11
CLCNK3	0	0.006	0	0.003	0.011	0.021	0	0.006
DNAH12	0.009	0.012	0	0.007	0.004	0.003	0.034	0.004

[illegible]

[illegible]

Gene	Non-neoplastic Sublet2	Non-neoplastic Sublet3	Non-neoplastic Sublet4	Non-neoplastic Sublet5	Non-neoplastic Sublet6	Non-neoplastic Sublet7	Non-neoplastic Sublet8	Non-neoplastic Sublet9	Non-neoplastic Sublet10	Non-neoplastic Sublet11	Non-neoplastic Sublet12	Non-neoplastic Sublet13	Non-neoplastic Sublet14	Non-neoplastic Sublet15	Non-neoplastic Sublet16	Non-neoplastic Sublet17	Non-neoplastic Sublet18	Non-neoplastic Sublet19	Non-neoplastic Sublet20
SLP1	0.003	0.002	0.025	0.024	0	0	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
SLP2	0.012	0.006	0.024	0.006	0.006	0	0.017	0.047	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046
SLP3	0.036	0.003	0	0	0	0	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
SLP4	0.053	0.007	0.028	0.048	0.006	0.006	0	0.004	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
SLP5	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP6	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP7	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP8	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP9	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP10	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP11	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP12	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP13	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP14	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP15	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP16	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP17	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP18	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP19	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
SLP20	0.003	0.002	0.006	0.006	0.006	0.006	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012

Gene	Non-neoplastic: FB.2	Non-neoplastic: SFB	Goblet transition1	Non-neoplastic: SFB.1	Non-neoplastic: SFB.2	Non-neoplastic: SFB.3	Non-neoplastic: SFB.4	Non-neoplastic: SFB.5
STAP1	0.006	0.012	0	0.015	0.01	0.013	0.008	0.004
C9orf24	0.007	0.058	0	0.025	0.029	0.021	0.011	0.017
EMPF2	0.116	0.006	0	0	0.002	0	0.011	0
G0S2	0.111	0.037	0	0.023	0.002	0.012	0.11	0.009
HYDM	0	0.005	0	0.017	0.004	0.017	0	0
A4C13	0	0.009	0	0.002	0	0.004	0	0
BPFB2	0.061	1.128	0	0.004	0	0.004	0.027	0.117
CLICF188	0.006	0.033	0	0.037	0.003	0.103	0.008	0.001
CLICF194	0	0.03	0	0.006	0.019	0.014	0	0
C3orf185	0.005	0.099	0	0.012	0.005	0.023	0	0
CDC20B	0.006	0.061	0	0.013	0.005	0.011	0	0
C1orf157	0	0.015	0	0.008	0.017	0.011	0	0
TM6ST1	0.006	0.053	0	0.021	0.008	0.14	0.021	0.035
OM4A6.1	0	0.015	0	0	0.002	0.005	0	0.002
FOXN1	0.004	0.001	0	0.007	0	0.001	0	0
FOX11	0.003	0.066	0	0.02	0.048	0.004	0.005	0.013
HEPACAM2	0.001	0.125	0	0.008	0.002	0.000	0	0
YPO2	0.016	31.15	236.594	0.135	0.379	2.671	2.243	0.461
R5P	0.128	0.053	0	0.007	0.013	0.007	0.116	0
MT3	0.597	0.066	0	0.032	0.087	0.319	0.004	0.052
ODAM	0.026	0.005	0	0	0	0.001	0	0
PR	0.02	0.383	0.51	0.024	0.006	0.113	0.039	0.177
PR64	0.072	0.303	1.276	0.019	0.008	0.014	0.024	0.309
PRH2	0.223	0.487	2.297	0.062	0.029	0.009	0.134	0.925
PR64	1.131	1.437	4.594	0.13	0.109	0.239	0.469	2.687
SCOR3A1	0.026	46.641	0.255	0.234	0.91	3.208	0.044	0.205
TF1	0	2.674	0	0	0	0.016	0	0
TMEM190	0	0.017	0	0.003	0.005	0.011	0	0
TMPC3118	0	0	22.195	0	0	0	0.072	0.008

Claims

- [Claim 1] A method for detecting the presence of NPC relapse (“NPCR”) in a subject, comprising: detecting in a biological sample obtained from the subject, the presence of one or more markers selected from the group consisting of:
- (a) a genetic mutation,
 - and/or
 - (b) a cytometric marker, wherein the cytometric marker comprises LY6D+ cells, and/or optionally, a biomarker expression,
- wherein the genetic marker is:
- i) a single-nucleotide variation (SNV) in (a) HMGN2P3 (high mobility group nucleosomal binding domain 2 pseudogene 3) (chr16: 26, 032, 755); (b) ARL5A (ADP Ribosylation Factor Like GTPase 5A) (chr2: 151, 828, 247), DHX57 (DExH-box helicase 57) (chr2: 38, 868, 300); (c) IL32 (interleukin 32) (chr16: 3, 065, 801) and /or (d) ATPAF1 (ATP Synthase Mitochondrial F1 Complex Assembly Factor 1) (chr1: 46, 668, 177); or
 - ii) an InDel in: (a) DNAJC11 (DnaJ Heat Shock Protein Family (Hsp40) Member C11) (chr1: 6, 667, 742); (b) EIF2AK1 (Eukaryotic Translation Initiation Factor 2 Alpha Kinase 1) (chr7: 6, 046, 108); (c) FAM234A (Family With Sequence Similarity 234 Member A) (chr16: 254, 566); (d) PARPBP (PARP1 Binding Protein) (chr12: 102, 123, 938) and/or (e) HDAC2 (Histone deacetylase 2) (chr6: 113, 970, 875).
- [Claim 2] The method of claim 1, wherein the SNV is selected from the group consisting of (a) [T>C] in IL32, (b) [T>A] in DHX57; (c) [G>A] HMGN2P3 (d); ([A>G] in ARL5A (e) [A>C] (in ATPAF1, and/or (f) [A>T] in LMO4.
- [Claim 3] The method of claim 1, wherein the InDel is selection from the group consisting of (a) [A>AT] InDel in EIF2AK1; (b) TTC>T. TTCTC] InDel in DNAJC11; (c) [G>GT] InDel in FAM234; (d) [G>A. GA] InDel in PARPBP; (e) [CT>C, CTT, CTTT, CTTTT, CTTTTT] InDel in HDAC2 and (f) CCAG>C] InDel in TACSRD2.
- [Claim 4] The method of claim 1, wherein the cytometric marker is a cytometric mutation or a cytometric specific expression.
- [Claim 5] The method of claim 3, wherein a cytometric (cell type) mutation is selected from the group consisting of:

- (i) HDAC2 (chr6: 113, 970, 875) [CT>C, CTT, CTTT, CTTTT, CTTTTT] InDel in neoplastic SPB1 cells present in the sample;
- (ii) an [Ato C] missense SNV ATPAF1 (chr1: 46668177) in LY6D+ neoplastic SPB1 cells present in the sample;
- (iii) a [CCAGC] in frame deletion in TACSRD2 (chr 1: 5857099) present in LY6D+ neoplastic SPB1 cells present in the sample; and/or
- (iv) an (Ato T) 5' UTR single-nucleotide in LMO4 (chr 1: 87328973) in LY6D+ neoplastic SPB1 cells present in the sample.

[Claim 6]

The method of claim 3, comprising detecting the presence in the sample, neoplastic LY6D+ cells optionally wherein the neoplastic LY6D+ cells express markers selected from the group consisting of:

- (a) KRT16 (Keratin 16); CEBPD (CCAAT enhancer binding protein delta); CDKN1A (cyclin-dependent kinase inhibitor 1A); PGM2 (Phosphoglucomutase 2);
- (b) MEG3 (Maternally Expressed 3); and CTNNBIP1 (Catenin Beta Interacting Protein 1);
- (c) IGF2BP3 (Insulin-like growth factor 2 mRNA-binding protein 3); FAF1 (Fas Associated Factor 1; DUSP11 (Dual-specificity phosphatase 11); and CLDND1 (claudin domain containing 1);
- (d) BAG4 (BAG Cochaperone 4); SIPA1L2 (Signal Induced Proliferation Associated 1 Like 2); AP3M2 (Adaptor Related Protein Complex 3 Subunit Mu 2); and SERPINB12 (Serpin Family B Member 12); and
- (e) CALML3 (Calmodulin Like 3); CLCA4 (Chloride Channel Accessory 4); GPX2 ((glutathione peroxidase 2), and LSP1 (Lymphocyte Specific Protein 1); the presence or increased abundance (compared to epithelial cells) of which is indicative of relapse.

[Claim 7]

The method of claim 3, detecting the presence in the sample, non-neoplastic LY6D+ cells, wherein the non-neoplastic LY6D+ cells express markers selected from the group consisting of:

- (a) CLCA4 (Chloride Channel Accessory 4); SYT8 (Synaptotagmin 8); FGFR3 (Fibroblast growth factor receptor 3); and/or
- (b) SUS4 (Sushi Domain Containing 4); TNNT3 (Troponin T3, Fast Skeletal Type); and NSG1 (Neuronal Vesicle Trafficking Associated 1), wherein the absence of the non- neoplastic LY6D+ cells or low levels of the non-neoplastic LY6D+ cells indicate relapse.

[Claim 8]

The method of any one of claims 1-7, comprising detecting the presence mutations in IL32 and DHX57.

- [Claim 9] The method of any one of claims 1-8, wherein the method detects a relapse with a specificity of at least 80%.
- [Claim 10] The method of any one of claims 1-7, wherein the method detects a relapse with a 100% sensitivity.
- [Claim 11] A method for detecting the presence of indicators of NPC ("NPCD"), in a subject, comprising: detecting in a biological sample obtained from the subject:
- (a) a SNVs is present in one or more genes from the group consisting of (i) EEF2KMT ([G>A] (chr16: 5, 091, 852); (ii) SOCS1 ([A>C] (chr16: 11, 255, 404; (iii) TESMIN ([A>T] (chr11: 68, 750, 666); and (iv) IGFBP7 ([T>C] (chr4: 57, 110, 068), optionally in combination with EBV.
- (b) an InDel selected from the group consisting of (i) EMP2 ([C>CA] (chr16: 10, 543, 633); (ii) IL32 ([AAGGTGACT...CGCCAGC>A, AAGC] (chr16: 3, 065, 653); and(iii) CSTA ([G>GA] I (chr3: 122, 337, 565); and/or
- (c) expression of a biomarker selected from the group consisting of: (i) CKAP4 (cytoskeleton associated Protein 4), (ii) SYNGR2 (Synaptogyrin-2), (iii) CFL1 (Cofilin 1) and (iv) RPMS1; and/or (v) CCL20, (vi) IL32, (viii) LCN2 wherein $CCL20_{high} + IL32_{high} + LCN2_{low}$ indicates the presence of NPC, wherein $CCL20_{high}$ refers to high levels of CCL20, $IL32_{high}$ refers to high levels of IL32 and $LCN2_{low}$ refers to low levels of LCN2.
- [Claim 12] The method of claim 11, wherein the presence of one or more mutations or expression of the one or more biomarkers indicates NPC in the subject with at least about 80% sensitivity, preferably, at least 85, 90 or up to 95% sensitivity.
- [Claim 13] The method of any one of claims 1-11 comprising detecting mRNA expression in the sample comprising subjecting the sample to DNA or RNA sequencing, optionally, wherein the DNA or RNA sequencing single-cell RNA sequencing and RT-PCR.
- [Claim 14] The method of any one of claims 1-11, comprising detecting RNA and protein /peptide, or fragment thereof, the method comprising contacting a biological sample with a binding partner binds to the one or more biomarkers and detecting binding between the antibody and the one or more biomarkers.
- [Claim 15] The method of any one of claims 1-12, wherein the mutation is detected by subjecting the sample to a DNA or RNA sequencing process.

- [Claim 16] The method of claim 14, wherein the binding is determined by immunohistochemistry (IHC), Enzyme-linked Immunosorbent Assay (ELISA) and in a lateral flow assay.
- [Claim 17] The method of claim 16, wherein the lateral flow assay comprises:
(a) reacting the biological sample with
(i) a binding agent that selectively binds the biomarker to form a capture complex of the binding agent and analyte, and
(ii) a capture agent that selectively binds to the capture complex but not free analyte to form a sandwich complex of the binding agent, capture agent, and analyte, and
(b) measuring sandwich complex formation;
wherein the amount of sandwich complex formation is directly related to the amount of the analyte in the sample.
- [Claim 18] A lateral flow device comprising a solid substrate having an application point 26, a conjugate zone 16, a capture zone 18, and an absorbent zone 20, the conjugation zone comprises (i) a binding partner to a biomarker indicative of NPC relapse, wherein the biomarker is selected from the group consisting of : PSMA4 (proteasome 20S subunit alpha 4), CALML3 (calmodulin like 3), SLC2A1 (solute carrier family 2 member 1), SNX3 (sorting nexin 3), LY6D (lymphocyte antigen 6 family member D), YBX1 (Y box binding Protein 1), and RPMS1 (an open reading frame of BamHI-Arightward transcripts of Epstein-Barr virus); or (ii) a binding partner to a biomarker indicative of NPC, wherein the biomarker is selected from the group consisting of: presence of NPC, selected from the group consisting of CKAP4; SYNGR2; CFL1 and RPMS1.
- [Claim 19] The device of claim 18, therein binding partner to the biomarker is an antibody.
- [Claim 20] A nasopharyngeal swab device comprising a “head” portion for insertion into the nasal cavity, a handle a series of preferably cylindrical “poles”, of varying diameters connecting head to the handle, wherein the head is preferably a hollow cylinder and comprises hollow cylindrical shapes (“hairs”) disposed on its surface, for collecting cells and secretions.
- [Claim 21] The swab of claim 20, wherein the head portion measures about 3mm in outer diameter and about 12-15 mm in length (without taking into account the hairs) or if the hairs are taken to account, from about 5mm in outer diameter/width and 14-17 mm in length.

- [Claim 22] The swab of claim 20 or 21, wherein each hair is a hollow cylinder with an outer radius (R), an inner radius (r), a height (which is the distance between its first plane and second plane), a thickness (R-r), a first plane disposed on the outer surface of the head, and a second plane disposed exterior to the swab.
- [Claim 23] The swab of claim 22, wherein the hollow cylindrical hairs are about 0.3 mm thick, about 2 mm in outer diameter and about and about 1.3 mm high.
- [Claim 24] The swab of any one of claims 20-23, wherein the hollow cylindrical hairs are located throughout the outer surface of the head of the device, about 0.75 mm apart.
- [Claim 25] The swab of any one of claims 20-24 comprising dome-shaped cross wires (about 0.1 mm thick) disposed on the cylindrical hairs, configured for use to gently scratch cells off, press against tissue and squeeze secretion out.
- [Claim 26] The swab of claim 25, wherein dome-shaped cross wires project about 0.2 mm from the second plane of the hairs, thus, the length from the top of the dome to the base of the hollow cylindrical hair (i. e., the portion resting on the outer surface of the head) is about 1.5 mm.
- [Claim 27] The swab of any one of claims 20-26, wherein, connecting the head and handle of the device are a first pole (connected to the head) of about 1 mm in diameter, a second/middle pole of about 2mm in diameter and which is hollow and is configured for extension or retraction of the device pushing the second pole over the first pole (retraction) or away from the second pole (extension) and a third pole, which is about 1 mm in diameter.
- [Claim 28] The swab of claim 27, wherein the first pole is about 20-50 mm in length, the second/middle pole is about 20-50 mm in length and/or the third pole is about 10-30 mm in length.
- [Claim 29] The device of any one of claims 20-28, wherein the handle is a hollow bulb, configured for squeezing, wherein, upon squeezing, materials can be forced out of the swab and upon releasing, materials can be sucked into the swab.
- [Claim 30] A sample collection tube comprising a cap, and a hollow interior with brushes/bristles disposed therein, wherein the brushes/bristles are arranged in a spiral or helical pattern along the length of the tube's-interior, creating a textured surface.