TITLE OF THE INVENTION

"PROGNOSTIC SIGNATURE FOR CANCER RECURRENCE AND RESPONSIVENESS"

RELATED APPLICATIONS

[0001] This application claims priority to Australian Provisional Application No. 2024900337 entitled "Prognostic Signature" filed 13 February 2024, the contents of which are incorporated herein by reference in their entirety.

FIELD OF THE INVENTION

[0002] This invention relates generally to methods of determining the likelihood of a subject with cancer responding to an anti-cancer therapy. The invention also relates to methods of determining the likelihood of cancer recurrence in a subject. More particularly, the invention relates to novel response to therapy biomarkers, and cancer recurrence biomarkers, and the measurement of these biomarkers in prognostic methods.

BACKGROUND OF THE INVENTION

[0003] The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavor to which this specification relates.

[0004] Phosphatidylinositol is one of a number of phospholipids found in cell membranes which play an important role in intracellular signal transduction. Cell signaling via 3'-phosphorylated phosphoinositides has been implicated in a variety of cellular processes, e.g., malignant transformation, growth factor signaling, inflammation, and immunity (Rameh *et al.* (1999) *J. Biol. Chem.*, 274:8347-8350). The enzyme responsible for generating these phosphorylated signaling products, phosphoinositide 3-kinase (also referred to as PI3-kinase or PI3K), was originally identified as an activity associated with viral oncoproteins and growth factor receptor tyrosine kinases that phosphorylate phosphatidylinositol (PI) and its phosphorylated derivatives at the 3'-hydroxyl of the inositol ring (Panayotou *et al.* (1992) *Trends Cell Biol* 2: 358-60).

[0005] Phosphoinositide 3-kinases (PI3K) are lipid kinases that phosphorylate lipids at the 3-hydroxyl residue of an inositol ring (Whitman *et al.* (1988) *Nature*, 332:664). The 3-phosphorylated phospholipids (PIP3s) generated by PI3-kinases act as second messengers recruiting kinases with lipid binding domains (including pleckstrin homology (PH) regions), such as Akt and phosphoinositide-dependent kinase-1 (PDK1). Binding of Akt to membrane PIP3s causes the translocation of Akt to the plasma membrane, bringing Akt into contact with PDK1, which is responsible for activating Akt. The tumour-suppressor phosphatase, PTEN,

dephosphorylates PIP3 and therefore acts as a negative regulator of Akt activation. The PI3-kinases Akt and PDK1 are important in the regulation of many cellular processes including cell cycle regulation, proliferation, survival, apoptosis and motility and are significant components of the molecular mechanisms of diseases such as cancer, diabetes and immune inflammation (Vivanco *et al.* (2002) *Nature Rev. Cancer* 2:489; Phillips et al. (1998) Cancer 83:41).

[0006] The main PI3K isoform in cancer is the Class I PI3-kinase, p110 α (alpha) (as described in U.S. Pat. Nos. 5,824,492; 5,846,824; 6,274,327). Other isoforms are implicated in cardiovascular and immune-inflammatory disease (Workman P (2004) *Biochem Soc Trans* 32:393-396; Patel *et al.* (2004) *Proceedings of the American Association of Cancer Research* (Abstract LB-247) 95th Annual Meeting, March 27-31, Orlando, Florida, USA; Ahmadi K and Waterfield MD (2004) Encyclopedia of Biological Chemistry (Lennarz W J, Lane M D eds) Elsevier Academic Press). The PI3K/Akt/PTEN pathway is an attractive target for cancer drug development since such modulating or inhibitory agents would be expected to inhibit proliferation, reverse the repression of apoptosis and surmount resistance to cytotoxic agents in cancer cells (Folkes *et al.* (2008*) J. Med. Chem.* 51: 5522-5532; Yaguchi *et al.* (2006) *J. Nat. Cancer Inst.* 98(8):545-556).

SUMMARY OF THE INVENTION

[0007] The present invention is predicated in part on the discovery that a number of gene expression and protein biomarkers can be used to determine the likelihood of cancer recurrence in a subject, in addition to the likelihood that a subject with cancer would respond to an anti-cancer therapy. These methods have important clinical outcomes in prognostic methods for assessing cancer patients and determining treatment options.

[0008] In one aspect of the present invention, there is provided a method of monitoring the responsiveness of a cancer to an anti-cancer treatment in a subject, the method comprising the step of determining an expression level of one or a plurality of biomarkers in a biological sample obtained from the subject, wherein the biomarkers comprise one or more of *DCN*, *LCN2, LOX, HMOX1,* and *PDCD1LG2*; wherein an altered or modulated expression level of the one or plurality of markers indicates or correlates with relatively increased or decreased responsiveness of the cancer to the anti-cancer treatment. Alternatively, or in addition, the biomarkers could be measured by the presence or abundance of the expression product of *DCN*, *LCN2, LOX, HMOX1,* and/or *PDCD1LG2*.

[0009] In another aspect of the present invention, there is provided method of determining the likelihood of cancer recurrence in a subject, the method comprising the step of determining a level of a cancer recurrence biomarker in a biological sample obtained from the subject, wherein the biomarker comprises the level of interaction of p85β and H3K27Me3 in a cellular compartment, and an altered or modulated level of the biomarker indicates or correlates with a relatively increased likelihood or relatively decreased likelihood of cancer recurrence.

[0010] In some embodiments, a reduced level of interaction between p85β and H3K27Me3 as compared to a control or reference sample (*e.g.*, a predetermined threshold) indicates a relatively increased likelihood of cancer recurrence.

[0011] In some embodiments, the cellular compartment comprises the cell nucleus and/or the cell cytoplasm. In some preferred embodiments, the cellular compartment comprises the cell cytoplasm.

[0012] In some embodiments, the biological sample is a whole blood sample or a tumour sample.

[0013] In yet another aspect, the invention provides a method of determining the likelihood of a cancer responding to an anti-cancer treatment, the method comprising the step of determining presence or level of one or a plurality of response to therapy biomarkers in a biological sample obtained from the subject, wherein the response to therapy biomarkers are selected from IL6, EPCAM, ABCB5, and GBP2 and an altered or modulated level of the one or a plurality of biomarkers indicates or correlates with relatively increased or decreased likelihood of the cancer responding to the anti-cancer treatment.

[0014] In some embodiments, the biological sample comprises circulating tumour cells.

[0015] In still yet another aspect, there is provided a method of determining the likelihood of a cancer responding to an anti-cancer treatment, the method comprising the step of determining an expression level of one or a plurality of response to therapy biomarkers in a biological sample obtained from the subject, wherein the response to therapy biomarkers comprise one or both of IL6 and GBP2 and an altered or modulated expression level of the one or both markers indicates or correlates with a relatively increased or decreased likelihood of the cancer responding to the anti-cancer treatment.

[0016] In some embodiments, an increased level of the response to therapy biomarkers indicates or correlates with an increased likelihood of the subject responding the anti-cancer treatment. In some embodiments of this type, the increased level is as compared to a control or reference sample. In some of the same embodiments and some other embodiments, the increased level is as compared to a predetermined threshold.

[0017] In some embodiments, the anti-cancer treatment is an anti-PD1 immunotherapy, or a PARP inhibitor therapy.

[0018] Also provided is a kit for determining an indicator used in assessing a likelihood of a subject with cancer responding to cancer therapy, the kit comprising, consisting, or consisting essentially of (a) at least one reagent that allows quantification of a polynucleotide or polypeptide expression product of the biomarkers described above and/or elsewhere herein in a biological sample; and optionally (b) instructions for using the at least one reagent; wherein the cancer therapy comprises therapy with an immune checkpoint inhibitor. In one embodiment,

the kit further comprises at least one reagent that allows quantification of a polynucleotide or polypeptide expression product of *DCN*, *LCN2, LOX, HMOX1, PDCD1LG2,* IL6, EPCAM, ABCB5, and/or GBP2 in the biological sample.

[0019] Also provided is a kit for determining an indicator used in assessing a likelihood of a subject having cancer recurrence, the kit comprising, consisting, or consisting essentially of (a) at least one reagent that allows quantification of a polynucleotide or polypeptide expression product of the cancer recurrence biomarkers described above and/or elsewhere herein in a biological sample; and optionally (b) instructions for using the at least one reagent; wherein the cancer therapy comprises therapy with an immune checkpoint inhibitor. In one embodiment, the kit further comprises at least one reagent that allows quantification of a polynucleotide or polypeptide expression product of *PIK3R2* in the biological sample.

[0020] In a further aspect, provided is a solid support for determining an indicator used in assessing a likelihood of a subject with cancer (or who has had cancer) having a recurrence of the cancer, the solid support comprising, consisting, or consisting essentially of at least one first oligonucleotide primer or probe immobilized to the solid support, wherein the at least one first oligonucleotide primer or probe hybridizes to a PIK3R2 transcript or cDNA. In some embodiments, the support further comprises a PIK3R2 transcript or cDNA thereof hybridized to the at least one first oligonucleotide primer or probe. The cDNA may correspond to mRNA derived from a cell or cell population (e.g., is a tumour cell or tumour cell population) .

[0021] In another aspect of the invention, provided is a composition for determining an indicator used in assessing a likelihood of a cancer recurring in a subject, the composition comprising, consisting, or consisting essentially of tumour cells, a detection agent that binds to the polypeptide expression product of PIK3R2. In some examples, the detection agents are antibodies or antigen-binding fragments thereof. Suitably, the subject has undergone, or is undergoing, a cancer therapy (e.g., immunotherapy).

BRIEF DESCRIPTION OF THE FIGURES

[0022] The following figures form part of the present specification and are included to further demonstrate certain aspects of the present disclosure. The disclosure may be better understood by reference to one or more of these figures in combination with the detailed description of specific embodiments presented herein.

[0023] **Figure 1. PI3K-mTOR inhibitor treatment increases "anti-tumour" immune cell populations in 4T1 tumours.** Frozen tumour tissues were collected from Balb/c 4T1 mice treated +/- GDC-0084 (7.5 mg/kg) in combination with αPD1 (10 mg/kg). Tumours were subjected to spatial immune profiling and cellular phenotypes analyzed. (A) Differential location of all cellular phenotypes in αPD1 versus GDC-0084 combination with αPD1 tumour tissues. (B) Pie chart showing differences in immune cell populations in αPD1 versus GDC-0084 in combination with αPD1 tumour tissues, expressed as % of total cell populations. (C) Differential location of individual immune cell populations in αPD1 versus GDC-0084

combination with αPD1 tumour tissues. Red: B cells; blue: bone marrow blast cells; dark green: CD11b+ neutrophil; light green: CD11c+ dendritic cells; purple: CD24+ dendritic cells; light brown: CD38+ myeloid cells; orange: endothelial cells; yellow: erythrocytes; dark brown: neutrophils; dark pink: proliferating cells; pink: T cells; grey: vascular cells.

Figure 2. PI3K-mTOR inhibitor treatment increases B cell and T cell enriched cellular neighborhoods in 4T1 tumours. (A) Graphical representation of immune cell populations (% of total cells) expressed within ten identified cellular neighborhoods. Red: B cells; purple: Bone marrow blast cells; Blue: CD11b+ neutrophils; light green: CD11c+ dendritic cells; dark green: CD24+ dendritic cells; purple CD38+ myeloid cells; dark orange: endothelial cells; orange: erythrocytes; yellow: neutrophils; light brown: proliferating cells; dark brown: T cells; pink: unidentified; dark pink: vascular cells; grey: other. (B) Pie chart showing differences in cellular neighborhoods in αPD1 versus GDC-0084+αPD1 tumour tissues, expressed as % of total cellular neighborhoods. (C) Differential location of cellular neighborhoods in αPD1 versus GDC-0084+αPD1 tumour tissues. For (B) and (C) Red: CN 1 erythrocyte cell rich; blue: CN 2 other; green: CN 3 bone marrow blast rich; purple: CN 4 neutrophil rich; dark orange: CN 5 proliferating cell rich; orange: CN 6 B cell rich; yellow: CN 7 CD11c+ dendritic cell rich; brown: CN 8 unidentified; pink: CN 9 T cell rich; and grey: CN10 CD11b+ neutrophil rich. For representations without colour: clockwise from 0° - CN10 CD11b+ neutrophil rich, CN 9 T cell rich, ----, CN8 unidentified, CN 7 CD11c+ dendritic cell rich, CN 6 B cell rich, CN 5 proliferating cell rich, CN 4 neutrophil rich, CN 3 bone marrow blast rich, ---, CN 2 other, CN 1 erythrocyte cell rich

[0024] Figure 3. PI3K-mTOR inhibitor treatment increases the cross-talk between "anti-tumour" immune cells and bone marrow blasts. Spatial proximity analysis from αPD1 (filled columns; left-hand columns) versus GDC-0084 combination with αPD1 (open columns, right-hand columns) tumour tissues. (A) Mean number of immune cell populations within 50 µm of bone marrow blast cells, vascular cells and endothelial cells. (B) Mean distance (µm) of bone marrow blast cells, vascular cells and endothelial cells to nearest immune cell phenotypes.

[0025] Figure 4. Down-regulation of "pro-tumour" immune cell profile following PI3K-mTOR inhibitor treatment. (A) Nanostring nCounter cell abundance analysis of regulatory T cells (Treg), neutrophils, and mast cell immune populations from GDC-0084 with/without αPD1 treated tumours. (B) Toluidine staining for mast cell detection in αPD1 and GDC-0084 combination with αPD1 treated tumours.

[0026] Figure 5. Enhancement of "anti-tumour" immune cell profile following PI3K-mTOR inhibitor treatment. (A) Nanostring nCounter cell abundance analysis of dendritic cell, NK, B and T cell immune populations from GDC-0084 with/without αPD1 treated tumours. (B) Gene expression levels of cytotoxicity-related cytokines, including IFN-ɣ (IFNɣ) and Granzyme B (GzmB) in GDC-0084 with/without αPD1 treated tumours. (C) Nanostring nCounter

cell abundance analysis of exhausted T cell immune populations from GDC-0084 +/- αPD1 treated tumours.

[0027] Figure 6. Induction of "anti-tumour" immune-related pathways by PI3K-mTOR inhibitor treatment. Heat-map representing global significance scores from gene set analysis (GSA). Red depicts increased expression while blue indicates decreased expression of genes associated with indicated pathways.

[0028] Figure 7. PI3K-mTOR inhibitor treatment increases cytotoxicity, tumour suppression, immune checkpoint and "anti-tumour" related genes. Bubble plot depicting differential expression of PI3K regulated genes in GDC-0084 without/with αPD1 treated tumours.

[0029] Figure 8. Dual PI3K-mTOR inhibitors down-regulate IL-6 and increase viral mimicry gene expression in the MCF-7 resistance model. (A) Schematic of MCF-7 resistance model. (B) qPCR analysis of IL-6 and GBP2 mRNA expression following treatment with the PI3K inhibitors Idelalisib, LY294002, Dactolisib and GDC-0084 (expressed as % DMSO control). ***, $p<0.001$, ****, $p<0.0001$, versus DMSO control, Dunnett's post test.

[0030] Figure 9. PI3K-mTOR inhibition down-regulates IL-6 and increases viral mimicry gene expression in the MDA-MB-231 reversal maintenance model. (A) Schematic of the MDA-MB-231 reversal maintenance model. (B) qPCR analysis of IL-6 and GBP2 mRNA expression following treatment with the PI3K-mTOR dual inhibitor, GDC-0084 (expressed as % vehicle control). *, $p<0.05$, **, $p<0.01$, versus DMSO control, Dunnett's post test.

[0031] Figure 10. IL-6/EpCAM/ABCB5 CTC biomarker predicts treatment response in melanoma patients. Immunofluorescence image analysis from circulating tumour cells (CTCs) isolated from melanoma patients pre/post-treatment. Blood CTCs were stained with IL-6, EpCAM and ABCB5 antibodies and imaged and analysed using the ASI Digital pathology platform. Bar graph shows the % of IL-6/EpCAM/ABCB5+ cells (expressed as % of total CTCs) isolated from patient blood samples. *, $p<0.05$ versus pre-treatment group, Dunnett's post test.

[0032] Figure 11. p85β:H3K27Me3 interaction biomarker predicts recurrence in triple negative breast cancer patients. DUOLINK analysis was performed on PBMC isolated from TNBC patients, CH002 (non-recurrence), CH007 (non-recurrence), CH003 (recurrence) pre- and post-chemotherapy. (A) Representative super resolution images of DUOLINK staining profiles in PBMC isolated from TNBC patient blood samples. White spots indicate p85β:H3K27Me3 interactions within cell nuclei (blue). (B) QuPath image analysis of p85β:H3K27Me3 interactions within the whole cell, nucleus and cytoplasm pre- and post-chemotherapy.

[0033] Figure 12. Associations between PIK3R2 gene expression and survival probability amongst TNBC breast and ovarian cancer patient cohorts. (A) Distant metastasis-free survival (DMFS) in TNBC/basal-like and HER2+ breast cancer datasets with high and low expression levels for PIK3R2 – from analysis of the GSE158309 dataset (see, Heimes et al., 2020). Patients were grouped into PIK3R2-high or PIK3R2-low depending on the expression levels of PIK3R2: high is above 75th percentile (solid lines) and low is below 75th percentile (dotted lines). Kaplan-Meier survival curves for plotted using distant metastasis as the endpoint. Patients with high PIK3R2 (solid lines) showed poorer DMFS in both basal-like/TNBC and HER2+ subtypes. (B) Assessing PIK3R2 mRNA expression in stage IV metastatic triple-negative breast cancer patients post chemotherapy (carboplatin, Nab-paclitaxel) + Pembrolizumab durvalumab+Olaparib in a single-arm phase-2 trial {Wilkerson, 2024 #91}. Patients with higher mRNA expression for PIK3R2 showed poorer overall survival compared to patients with lower PIK3R2 mRNA expression. High = above 75th percentile in PIK3R2 expression; Low = lower than 75th percentile. (C) Assessing PIK3R2 mRNA expression levels in triple-negative breast cancer patients who had residual tumours after treatment with neoadjuvant chemotherapy (NACT) (see, Blaye et al., 2022). All patients here had residual tumours post NACT, however those who presented with a distant relapse had higher pre-treatment levels of PIK3R2 compared to those who did not present with a distant relapse.

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

[0034] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, preferred methods and materials are described. For the purposes of the present invention, the following terms are defined below.

[0035] The articles “a” and “an” are used herein to refer to one or to more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

[0036] The term “about” as used herein refers to the usual error range for the respective value readily known to the skilled person in this technical field. Reference to “about” a value or parameter herein includes (and describes) embodiments that are directed to that value or parameter *per se*.

[0037] The “amount” or “level” of a biomarker is a detectable level in a sample. These can be measured by methods known to one skilled in the art and also disclosed herein. The expression level or amount of biomarker assessed can be used to determine the response to treatment.

[0038] As used herein, "and/or" refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations when interpreted in the alternative (or).

[0039] Throughout this specification, unless the context requires otherwise, the words "comprise", "comprises" and "comprising" will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements. Thus, use of the term "comprising" and the like indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present. By "consisting of" is meant including, and limited to, whatever follows the phrase "consisting of". Thus, the phrase "consisting of" indicates that the listed elements are required or mandatory, and that no other elements may be present. By "consisting essentially of" is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase "consisting essentially of" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present depending upon whether or not they affect the activity or action of the listed elements.

[0040] The term "control subject", as used in the context of the present invention, may refer to a subject known to have a particular state (e.g., responsive to a cancer therapy, or has a cancer that does not recur) (positive control), or to a subject known not have the particular state (e.g., not responsive to a cancer therapy or has a cancer that recurs) (negative control). It is understood that control subjects include data obtained and used as a standard, i.e. it can be used over and over again for multiple different subjects. In other words, for example, when comparing a subject sample to a control sample, the data from the control sample could have been obtained in a different set of experiments, for example, it could be an average obtained from a number of healthy subjects and not actually obtained at the time the data for the subject was obtained.

[0041] The terms "correlated" and "associated" are used interchangeably herein to refer to the association between two measurements (or measured entities). The disclosure provides genetic and/or epigenetic variations, the level(s) of which are associated with disease diagnosis and/or prognosis and/or response to treatment.

[0042] The terms "decrease", "reduced", "reduction", "inhibit", "suppress", "attenuate" and the like are all used herein to mean a decrease by a statistically significant amount. In some embodiments, these terms typically mean a decrease by at least 10% as compared to a reference level (e.g., the absence of a given treatment or agent) and can include, for example, a decrease by at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about

75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, at least about 99%, or more. As used herein “reduction”, “suppression”, and “inhibition” does not necessitate a complete inhibition or reduction as compared to a reference level. “Complete inhibition” and the like is a 100% inhibition as compared to a reference level. A decrease can be preferably down to a level accepted as within the range of normal (e.g., for an individual without a given disorder).

[0043] The term ‘immobilized" means that a molecular species of interest is fixed to a solid support, suitably by covalent linkage. This covalent linkage can be achieved by different means depending on the molecular nature of the molecular species. Moreover, the molecular species may be also fixed on the solid support by electrostatic forces, hydrophobic or hydrophilic interactions or Van-der-Waals forces. The above described physicochemical interactions typically occur in interactions between molecules. In particular embodiments, all that is required is that the molecules (e.g., nucleic acids or polypeptides) remain immobilized or attached to a support under conditions in which it is intended to use the support, for example in applications requiring nucleic acid amplification and/or sequencing or in in antibody-binding assays. For example, oligonucleotides or primers are immobilized such that a 3' end is available for enzymatic extension and/or at least a portion of the sequence is capable of hybridizing to a complementary sequence. In some embodiments, immobilization can occur via hybridization to a surface attached primer, in which case the immobilized primer or oligonucleotide may be in the 3'-5' orientation. In other embodiments, immobilization can occur by means other than base-pairing hybridization, such as the covalent attachment.

[0044] The terms “increased”, “increase”, “enhance”, or “activate” are all used herein to mean an increase by a statistically significant amount. In some embodiments, the terms “increased”, “increase”, “enhance”, or “activate” can mean an increase of at least 10% as compared to a reference level (e.g., the absence of a given treatment or agent) and can include, for example, of at least about 10% as compared to a reference level, for example an increase of at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, at least about 99%, or up to and including a 100% increase or any increase between 10-100% as compared to a reference level or at least about a 2-fold, or at least about a 3-fold, or at least about a 3-fold, or at least about a 4-fold, or at least about a 5-fold, or at least about a 10-fold increase, or any increase between 2-fold and 10-fold or greater as compared to a reference level. In the context of a marker or symptom, an “increase” is a statistically significant increase in such level.

[0045] As used herein, the term “label” and grammatical equivalents thereof, refer to any atom or molecule that can be used to provide a detectable and/or quantifiable signal. In particular, the label can be attached, directly or indirectly, to a nucleic acid or protein. Suitable labels that can be attached include, but are not limited to, radioisotopes, fluorophores,

quenchers, chromophores, mass labels, electron dense particles, magnetic particles, spin labels, molecules that emit chemiluminescence, electrochemically active molecules, enzymes, cofactors, and enzyme substrates. A label can include an atom or molecule capable of producing a visually detectable signal when reacted with an enzyme. In some embodiments, the label is a "direct" label which is capable of spontaneously producing a detectible signal without the addition of ancillary reagents and is detected by visual means without the aid of instruments. For example, colloidal gold particles can be used as the label. Many labels are well known to those skilled in the art. In specific embodiments, the label is other than a naturally-occurring nucleoside. The term "label" also refers to an agent that has been artificially added, linked or attached via chemical manipulation to a molecule.

[0046] As used herein, "level" with reference to a biomarker refers to the amount or concentration of the biomarker in a sample. The amount or concentration may be absolute or may be relative, and can be determined using any method known in the art.

[0047] As used herein, the term "likelihood" is used as a measure of whether subjects with a particular biomarker profile actually have a particular state (e.g., a cancer that is likely to recur) based on a given mathematical model. An increased likelihood for example may be relative or absolute and may be expressed qualitatively or quantitatively. For instance, an increased risk may be expressed as simply determining the subject's level of a given biomarker and placing the test subject in an "increased risk" category, based upon previous population studies. Alternatively, a numerical expression of the test subject's increased risk may be determined based upon biomarker level analysis.

[0048] "Measuring" or "measurement" means assessing the presence, absence, quantity or amount (which can be an effective amount) of a given substance within a sample, including the derivation of qualitative or quantitative concentration levels of such substances, or otherwise evaluating the values or categorization of a subject's clinical parameters. Alternatively, the term "assaying," "detecting" or "detection" may be used to refer to all measuring or measurement as described in this specification.

[0049] By "obtained" is meant to come into possession. Samples so obtained include, for example, nucleic acid extracts or polypeptide extracts isolated or derived from a particular source. For instance, the extract may be isolated directly from a biological fluid or tissue of a subject.

[0050] As used herein, the terms "overexpress," "overexpression," "overexpressing" or "overexpressed" interchangeably refer to a gene (e.g., *PI3KCA* gene) that is transcribed or translated at a detectably greater level, usually in a cancer cell, in comparison to a normal cell. Overexpression, therefore, refers to both overexpression of protein and RNA (due to increased transcription, post transcriptional processing, translation, posttranslational processing, altered stability and altered protein degradation), as well as local overexpression due to altered protein traffic patterns and augmented functional activity, for example, as in an

increased enzyme hydrolysis of substrate. Overexpression can also be by 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more in comparison to a normal cell or comparison cell (e.g., a breast cell).

[0051] The term "PI3K inhibitor" and grammatical variants thereof are used herein to refer to a molecule that decreases or inhibits at least one function or biological activity of PI3K. For example, PI3K inhibitors may inhibit or reduce the enzymatic activity of PI3K and/or may inhibit or reduce the expression of PI3K.

[0052] As used herein, the term "positive response" means that the result of a treatment regimen includes some clinically significant benefit, such as the prevention, or reduction of severity, of symptoms, or a slowing of the progression of the condition. For example, a reduction in tumour size or tumour burden, or a slowing in the rate of tumour growth or spread (i.e., metastasizing), can indicate a positive response. By contrast, the term "negative response" or "non-response" means that a treatment regimen provides no or minimal clinically significant benefit, such as the prevention, or reduction of severity, of symptoms, or increases the rate of progression of the condition. In some instances, the Response Evaluation Criteria in Solid Tumors (RECIST) is used to assess positive or negative response to therapy (Eisenhauer et at. (2009) *Eur J Cancer.* 45: 228-47; and http://www. irrecist.com/recist/). A positive response may include a "partial response" or "complete response", such as defined by RECIST 1.1, while a negative response may be equivalent to "stable disease".

[0053] As used herein, the term "predetermined threshold" refers to a value, above or below which, indicates a particular state, such as a subject that is likely (or unlikely) to respond to a cancer therapy, or a cancer that is likely (or unlikely) to recur. For example, for the purposes of the present invention, a predetermined threshold may represent the level of a biomarker in a sample from an appropriate control subject, such as a subject known to respond to a cancer therapy, or in pooled samples from multiple control subjects or medians or averages of multiple control subjects. Thus, a level above or below the threshold indicates whether a subject is likely to respond to a cancer therapy, as taught herein. In other examples, a predetermined threshold may represent a value larger or smaller than the level determined for a control subject so as to incorporate a further degree of confidence that a level above or below the predetermined threshold is indicative of the presence of the particular state. For example, the predetermined threshold may represent the average or median level of a biomarker in a group of control subjects, plus or minus 1, 2, 3 or more standard deviations. Those skilled in the art can readily determine an appropriate predetermined threshold based on analysis of biological samples from appropriate control subjects.

[0054] The term "solid support" as used herein refers to a solid inert surface or body to which a molecular species, such as a nucleic acid and polypeptides can be immobilized. Non-limiting examples of solid supports include glass surfaces, plastic surfaces, latex, dextran, polystyrene surfaces, polypropylene surfaces, polyacrylamide gels, gold

surfaces, and silicon wafers. In some embodiments, the solid supports are in the form of membranes, chips or particles. For example, the solid support may be a glass surface (e.g., a planar surface of a flow cell channel). In some embodiments, the solid support may comprise an inert substrate or matrix which has been "functionalized", such as by applying a layer or coating of an intermediate material comprising reactive groups which permit covalent attachment to molecules such as polynucleotides. By way of non-limiting example, such supports can include polyacrylamide hydrogels supported on an inert substrate such as glass. The molecules (e.g., polynucleotides) can be directly covalently attached to the intermediate material (e.g., a hydrogel) but the intermediate material can itself be non-covalently attached to the substrate or matrix (e.g., a glass substrate). The support can include a plurality of particles or beads each having a different attached molecular species.

[0055] As used herein, a "subject" means a human or animal. Usually the animal is a vertebrate such as a primate, rodent, domestic animal or game animal. Primates include chimpanzees, cynomolgus monkeys, spider monkeys, and macaques (e.g., Rhesus). Rodents include mice, rats, woodchucks, ferrets, rabbits, and hamsters. Domestic and game animals include cows, horses, pigs, deer, bison, buffalo, feline species (e.g., domestic cat), canine species (e.g., dog, fox, wolf), avian species (e.g., chicken, emu, ostrich), and fish (e.g., trout, catfish, and salmon). In some embodiments the subject is a mammal (e.g., a primate (e.g., a human)). The terms "individual", "patient" and "subject" are used interchangeably herein.

[0056] As used herein, the terms "treat", "treatment", "treating" and the like, refer to therapeutic treatments, wherein the object is to reverse, alleviate, ameliorate, inhibit, slow down or stop the progression or severity of a condition associated with a disease or disorder (e.g., cancer or tumour). The term "treating" includes reducing or alleviating at least one adverse effect or symptom of a condition, disease or disorder associated with cancer or tumour. Treatment is generally "effective" if one or more symptoms or clinical markers are reduced. Alternatively, treatment is "effective" if the progression of a disease is reduced or halted. That is, "treatment" includes not just the improvement of symptoms or markers, but also a cessation of, or at least slowing of, progress or worsening of symptoms compared to what would be expected in the absence of treatment. Beneficial or desired clinical results include, but are not limited to, alleviation of one or more symptom(s), diminishment of extent of disease, stabilized (i.e., not worsening) state of disease, delay or slowing of disease progression, amelioration, or palliation of the disease state, remission (whether partial or total), and/or decreased mortality, whether detectable or undetectable. The term "treatment" of a disease also includes providing relief from the symptoms or side-effects of the disease (including palliative treatment). A treatment need not cure a disorder (i.e., complete reversal or absence of disease) to be considered effective.

[0057] As used herein, the term "tumour" refers to any neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues. The terms "cancer" and "cancerous" refer to or describe the physiological condition in mammals that is typically characterized in part by unregulated cell growth. As used herein, the

term "cancer" refers to non-metastatic and metastatic cancers, including early stage and late stage cancers. The term "precancerous" refers to a condition or a growth that typically precedes or develops into a cancer. The term "non-metastatic" refers to a cancer that is benign or that remains at the primary site and has not penetrated into the lymphatic or blood vessel system or to tissues other than the primary site. Generally, a non-metastatic cancer is any cancer that is a stage 0, I or II cancer. By "early stage cancer" is meant a cancer that is not invasive or metastatic or is classified as a stage 0, I or II cancer. The term "late stage cancer" generally refers to a stage III or IV cancer, but can also refer to a stage II cancer or a sub stage of a stage II cancer. One skilled in the art will appreciate that the classification of a stage II cancer as either an early stage cancer or a late stage cancer depends on the particular type of cancer. Illustrative examples of cancer include, but are not limited to, breast cancer, prostate cancer, ovarian cancer, cervical cancer, pancreatic cancer, colorectal cancer, lung cancer, hepatocellular cancer, gastric cancer, liver cancer, bladder cancer, cancer of the urinary tract, thyroid cancer, renal cancer, carcinoma, melanoma, brain cancer, non-small cell lung cancer, squamous cell cancer of the head and neck, endometrial cancer, multiple myeloma, mesothelioma, rectal cancer and esophageal cancer. In an exemplary embodiment, the cancer is breast cancer.

[0058] It will be appreciated that the terms used herein and associated definitions are used for the purpose of explanation only and are not intended to be limiting.

[0059] Each embodiment described herein is to be applied *mutatis mutandis* to each and every embodiment unless specifically stated otherwise.

2. Methods of determining cancer recurrence

[0060] The present invention discloses that the association between phosphatidylinositol 3-kinase (PI3K) 85 kDa regulatory subunit beta (p85β) and histone 3, tri-methylated at lysine 27 (H3K27Me3) can be measured as a biomarker to indicate the likelihood of cancer recurrence in as subject. The biomarker described herein as being useful for indicating the likelihood of cancer recurrence (also referred to herein as a "cancer recurrence biomarker") is the association (i.e., the proximity) between p85β and H3K27Me3. In this regard, the present invention provides a method of determining the likelihood of cancer recurrence in a subject, the method comprising the step of determining a level (i.e., proximity) of at least one biomarker in a biological sample obtained from the subject, wherein the biomarker comprises p85β:H3K27Me3, and an altered or modulated level of the biomarker indicates or correlates with relatively increased or likelihood of cancer recurrence. In some preferred embodiments of the invention, the level of the biomarker is measured by determining the cellular compartment of PI3K p85β and histone 3, as described in more detail below and/or elsewhere herein.

[0061] The wild-type human p85β amino acid sequence has a UniProt accession no. O00459, and provided below:

MAGPEGFQYRALYPFRRERPEDLELLPGDVLVVSRAALQALGVAEGGERCPQSVGW
MPGLNERTRQRGDFPGTYVEFLGPVALARPGPRPRGPRPLPARPRDGAPEPGLTLP
DLPEQFSPPDVAPPLLVKLVEAIERTGLDSESHYRPELPAPRTDWSLSDVDQWDTAAL
ADGIKSFLLALPAPLVTPEASAEARRALREAAGPVGPALEPPTLPLHRALTLRFLLQHL
GRVASRAPALGPAVRALGATFGPLLLRAPPPPSSPPPGGAPDGSEPSPDFPALLVEK
LLQEHLEEQEVAPPALPPKPPKAKPASTVLANGGSPPSLQDAEWYWGDISREEVNEK
LRDTPDGTFLVRDASSKIQGEYTLTLRKGGNNKLIKVFHRDGHYGFSEPLTFCSVVDLI
NHYRHESLAQYNAKLDTRLLYPVSKYQQDQIVKEDSVEAVGAQLKVYHQQYQDKSRE
YDQLYEEYTRTSQELQMKRTAIEAFNETIKIFEEQGQTQEKCSKEYLERFRREGNEKE
MQRILLNSERLKSRIAEIHESRTKLEQQLRAQASDNREIDKRMNSLKPDLMQLRKIRDQ
YLVWLTQKGARQKKINEWLGIKNETEDQYALMEDEDDLPHHEERTWYVGKINRTQAE
EMLSGKRDGTFLIRESSQRGCYACSVVVDGDTKHCVIYRTATGFGFAEPYNLYGSLK
ELVLHYQHASLVQHNDALTVTLAHPVRAPGPGPPPAAR [SEQ ID NO: 1].

[0062] Histone 3 is a DNA packing protein. The tri-methylation is known to be associated with the downregulation of nearby genes via the formation of heterochromatic regions. The wild-type histone H3 protein has an UniProt accession no. P68431, and an amino acid sequence as set out, below:

MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKS
TELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAIHAKRVT
IMPKDIQLARRIRGERA [SEQ ID NO: 2].

[0063] Biological samples that can be used with the present invention include cancer cells. Cancer cells for the practice of the present invention can be obtained from any suitable cancer-cell containing patient sample, illustrative examples of which include tumour biopsies, circulating tumour cells (CTCs), primary cell cultures, or cell lines derived from tumours or exhibiting tumour-like properties, as well as preserved tumour samples, such as formalin-fixed, paraffin-embedded tumour samples, or frozen tumour samples. In some embodiments, the sample is obtained prior to treatment with a therapy. In other embodiments, the sample is obtained after treatment with a therapy. In some embodiments, the sample comprises a tumour tissue sample, which can be formalin-fixed and paraffin-embedded, archival, fresh or frozen. In some embodiments, the whole blood comprises immune cells, circulating tumor cells, and any combinations thereof.

[0064] Presence and/or levels/amount of a biomarker (e.g., the association between p85β and H3K27Me3) can be determined qualitatively and/or quantitatively based on any suitable criterion known in the art. In certain embodiments, presence and/or expression levels/amount of a biomarker in a first sample is increased or elevated as compared to presence/absence and/or expression levels/amount in a second sample (e.g., before treatment with a therapy). In certain embodiments, presence/absence and/or levels/amount of a biomarker in a first sample is decreased or reduced as compared to presence and/or levels/amount in a

second sample. In certain embodiments, the second sample is a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue. Additional disclosures for determining presence/absence and/or levels/amount of an association biomarker are described herein.

[0065] In some embodiments of any of the methods, an elevated level refers to an overall increase of about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater, in the level of biomarker, as detected by standard art known methods such as those described herein, as compared to a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue. In certain embodiments, an elevated level refers to the increase in level/amount of a biomarker in the sample wherein the increase is at least about any of 1.5×, 1.75×, 2×, 3×, 4×, 5×, 6×, 7×, 8×, 9×, 10×, 25×, 50×, 75×, or 100× the level/amount of the respective biomarker in a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue. In some embodiments, an elevated level refers to an overall increase of greater than about 1.5-fold, about 1.75-fold, about 2-fold, about 2.25-fold, about 2.5-fold, about 2.75-fold, about 3-fold, or about 3.25-fold as compared to a reference sample, reference cell, reference tissue, control sample, control cell, control tissue, or internal control.

[0066] In some embodiments of any of the methods, a reduced level refers to an overall reduction of about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater, in the level of biomarker, detected by standard art known methods such as those described herein, as compared to a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue. In certain embodiments, reduced level refers to a decrease in level/amount of a biomarker in the sample wherein the decrease is at least about any of 0.9×, 0.8×, 0.7×, 0.6×, 0.5×, 0.4×, 0.3×, 0.2×, 0.1×, 0.05×, or 0.01× the level/amount of the respective biomarker in a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue.

[0067] Provided herein are predictive/prognostic methods and kits that are based on the determination that p85β co-localizes in the cytoplasm and the nucleus with H3K27Me3 and that this co-localization contributes at least in part to the recurrence of the cancer. The predictive methods suitably comprise: (i) obtaining a sample from a subject, wherein the sample comprises a cancer cell (e.g., a CTC); (ii) contacting the sample with a first antigen-binding molecule that binds to p85β in the sample and a second antigen-binding molecule that binds to H3K27Me3 in the sample; and (iii) detecting localization of the first and second antigen-binding molecule(s) in the nucleus of the cancer cell, wherein localization of the first and second antigen binding molecules in the nucleus of the cancer cell is indicative that the cancer cell has reduced likelihood of recurrence.

[0068] Assessment of the localization of p85β and H3K27Me3 in the cytoplasm and/or the nucleus of a cancer cell may be performed using any suitable localization

identification technique, e.g., by immunohistochemistry (IHC), typically using an p85β antibody that has a different detectable moiety or label than a H3K27Me3 partner antibody. A number of localization assays are known in the art, for example, chemiluminescence assays, FRET, sequential CHIP assays, co-immunoprecipitaion, ELISAs, etc. In some embodiments, spatial proximity assays (also referred to as "proximity assays") are employed, which can be used to assess the formation of a complex between p85β and H3K27Me3. Proximity assays rely on the principle of "proximity probing", wherein an analyte, typically an antigen, is detected by the coincident binding of multiple (i.e., two or more, generally two, three or four) binding agents or probes, which when brought into proximity by binding to the analyte (hence "proximity probes") allow a signal to be generated.

[0069] In some embodiments, at least one of the proximity probes comprises a nucleic acid domain (or moiety) linked to the analyte-binding domain (or moiety) of the probe, and generation of the signal involves an interaction between the nucleic acid moieties and/or a further functional moiety which is carried by the other probe(s). Thus, signal generation is dependent on an interaction between the probes (more particularly by the nucleic acid or other functional moieties/domains carried by them) and hence only occurs when both the necessary two (or more) probes have bound to the analyte, thereby lending improved specificity to the detection system.

[0070] The concept of proximity probing has been developed in recent years and many assays based on this principle are now well known in the art.

[0071] Proximity assays are typically used to assess whether two particular proteins or portions thereof are in close proximity, e.g., proteins that are bound to each other, fusion proteins, and/or proteins that are positioned in close proximity. One such assay, known as proximity ligation assay (PLA), and which is used in some embodiments of the present invention, features two antibodies (raised in different species) bound to the targets of interest (see, *Nature Methods*, 3, 995-1000 (2006)). PLA probes, which are species-specific secondary antibodies with a unique oligonucleotide strand attached, are then bound to the appropriate primary antibodies. In the case of the targets being in close proximity, the oligonucleotide strands of the PLA probes can interact with additional ssDNA and DNA ligase such they can be circulated and amplified via rolling circle amplification (RCA). When highly processive DNA polymerases such as Phi 29 DNA polymerase is used, the circular DNA template can be replicated hundreds to thousands of times longer and as a result producing ssDNA molecules from hundreds of nanometers to microns in length (see, Angewandte Chemie International Edition, 2008, 47, 6330-6337). After the amplification, the replicated DNA can be detected via detection systems. Thus, a visible signal is indicative that the targets of interest are in close proximity. These assays feature the use of several DNA-antibody conjugates as well as enzymes such as DNA ligase and DNA polymerase.

[0072] In other embodiments, a dual binders (DB) assay is employed, which utilizes a bi-specific detection agent consisting of two Fab fragments with fast off-rate kinetics joined by a flexible linker (Van dieck et al., 2014 *Chemistry & Biology* 21(3) :357-368). In principle, because the dual binders comprise Fab fragments with fast off-rate kinetics, the dual binders are washed off if only one of the Fab fragments is bound to its epitope (simultaneous cooperative binding of both Fab fragments of the dual binder prevents dissociation of the dual binder and leads to positive staining/visibility).

[0073] According to another approach disclosed in International Patent Publication No. WO2014/139980, which is encompassed in the practice of the present invention, proximity assays and tools are described, which employ a biotin ligase substrate and an enzyme to perform a proximity assay. The method provides detection of target molecules and proximity while maintaining the cellular context of the sample. The use of biotin ligase such as an enzyme from *E. coli* and peptide substrate such as amino-acid substrate for that enzyme provides for a sensitive and specific detection of protein-protein interactions in FFPE samples. Because biotin ligase can efficiently biotinylate appropriate peptide substrate in the presence of biotin and the reaction can only occur when the enzyme makes physical contact with the peptide substrate, biotin ligase and the substrate can be separately conjugated to two antibodies that recognize targets of interest respectively.

[0074] Cancer recurrence biomarkers of the present disclosure include *PIK3R2*, and more particularly the expression products of on *PIK3R2*. The on *PIK3R2* gene encodes the P85β subunit of phosphatidylinositol 3-kinase (PI3K), which (as described above) regulates the activity of the PI3K enzyme.

[0075] Expression from the human on *PIK3R2* gene results in the transcript set forth in SEQ ID NO: 3 (represented as mDNA; as set forth in NCBI Reference Sequence: NM_005027.4) and the P85β subunit protein set forth in SEQ ID NO:1.

[0076] As demonstrated herein, *PIK3R2* can be used as a cancer recurrence biomarker to predict whether a cancer in a subject is likely to recur. Expression products of *PIK3R2*, including polynucleotide (e.g., mRNA) and polypeptide expression products of PIK3R2, correlate with the recurrence of cancer therapy and can be assessed so as to predict whether a cancer is likely to recur in the subject. Thus, in specific embodiments the PIK3R2 biomarker may be used either by itself or in combination with one or more other cancer therapy biomarkers for the determination of the indicator for assessing the likelihood of a subject responding to cancer therapy.

[0077] A method for determining an indicator used in assessing a likelihood of a cancer recurring in a subject, the method comprising, consisting or consisting essentially of:

(1) determining a biomarker value for at least one cancer recurrence biomarker in a sample from the subject, wherein the, or one of the, cancer recurrence biomarker(s) is an expression product of PIK3R2; and

(2) determining the indicator using the biomarker value(s), wherein the indicator is at least partially indicative of the likelihood of recurrence of the cancer.

[0078] In some embodiments, the expression product of PIK3R2 is a polynucleotide and the biomarker value for PIK3R2, is indicative of the abundance or concentration of the polynucleotide in the sample.

[0079] In some embodiments of this type, wherein the polynucleotide expression product comprises a nucleotide sequence having at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence set forth in SEQ ID NO: 3, or a complement thereof.

[0080] In some embodiments, wherein the expression product of PIK3R2 is a polypeptide and the biomarker value for PIK3R2 is indicative of the abundance or concentration of the polypeptide in the sample.

[0081] In some embodiments, the polypeptide expression product comprises an amino acid sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence set forth in SEQ ID NO: 1.

[0082] In some embodiments, the abundance or concentration of the polynucleotide or polypeptide expression product of PIK3R2 is about the same as the abundance or concentration that correlates with a cancer that known to recur, and the indicator is thereby determined to be at least partially indicative of cancer that is likely to recur. Alternatively, the abundance or concentration of the polynucleotide or polypeptide expression product of PIK3R2 is decreased relative the abundance or concentration that correlates with a cancer that is known to recur, and the indicator is thereby determined to be at least partially indicative of a cancer that is not likely to recur. Alternatively, the abundance or concentration of the polynucleotide or polypeptide expression product of PIK3R2 is increased relative to the abundance or concentration that correlates with a cancer that is known to recur, and the indicator is thereby determined to be at least partially indicative of a cancer that is known to recur.

[0083] In some embodiments, the abundance or concentration of the polynucleotide or polypeptide expression product of PIK3R2 is about the same as the abundance or concentration that correlates with a positive response to cancer therapy, and the indicator is thereby determined to be at least partially indicative of a positive response to therapy.

[0084] In some embodiments, the sample is a clinical sample. In some embodiments, the sample is obtained from a primary or metastatic tumour. Tissue biopsy is often used to obtain a representative piece of tumour tissue. Alternatively, tumour cells can be obtained indirectly in the form of tissues or fluids that are known or thought to contain the

tumour cells of interest. For instance, samples of lung cancer lesions may be obtained by resection, bronchoscopy, fine needle aspiration, bronchial brushings, or from sputum, pleural fluid or blood. Genes or gene products can be detected from cancer or tumour tissue or from other body samples such as urine, sputum, serum or plasma. The same techniques discussed above for detection of target genes or gene products in cancerous samples can be applied to other body samples. Cancer cells may be sloughed off from cancer lesions and appear in such body samples. By screening such body samples, a simple early diagnosis can be achieved for these cancers. In addition, the progress of therapy can be monitored more easily by testing such body samples for target genes or gene products.

[0085] In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is a single sample or combined multiple samples from the same subject or individual that are obtained at one or more different time points than when the test sample is obtained. For example, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is obtained at an earlier time point from the same subject or individual than when the test sample is obtained. Such reference sample, reference cell, reference tissue, control sample, control cell, or control tissue may be useful if the reference sample is obtained during initial diagnosis of cancer and the test sample is later obtained when the cancer becomes metastatic.

[0086] In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is a combination of multiple samples from one or more healthy individuals who are not the subject or individual. In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is a combination of multiple samples from one or more individuals with a disease or disorder (e.g., cancer) who are not the subject or individual. In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is pooled RNA samples from normal tissues or pooled plasma or serum samples from one or more individuals who are not the subject or individual. In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is pooled RNA samples from tumour tissues or pooled plasma or serum samples from one or more individuals with a disease or disorder (e.g., cancer) who are not the subject or individual.

[0087] In some embodiments, the sample is a tissue sample from the individual. In some embodiments, the tissue sample is a tumour tissue sample (e.g., biopsy tissue). In some embodiments, the tissue sample is lung tissue. In some embodiments, the tissue sample is renal tissue. In some embodiments, the tissue sample is skin tissue. In some embodiments, the tissue sample is pancreatic tissue. In some embodiments, the tissue sample is gastric tissue. In some embodiments, the tissue sample is bladder tissue. In some embodiments, the tissue sample is esophageal tissue. In some embodiments, the tissue sample is mesothelial tissue. In some embodiments, the tissue sample is breast tissue. In some embodiments, the tissue sample is thyroid tissue. In some embodiments, the tissue sample is colorectal tissue. In

some embodiments, the tissue sample is head and neck tissue. In some embodiments, the tissue sample is osteosarcoma tissue. In some embodiments, the tissue sample is prostate tissue. In some embodiments, the tissue sample is ovarian tissue, HCC (liver), blood cells, lymph nodes, and/or bone/bone marrow tissue. In some embodiments, the tissue sample is colon tissue. In some embodiments, the tissue sample is endometrial tissue. In some embodiments, the tissue sample is brain tissue (e.g., glioblastoma, neuroblastoma, and so forth).

[0088] In some embodiments, a tumour tissue sample (the term "tumour sample" is used interchangeably herein) may encompass part or all of the tumour area occupied by tumour cells. In some embodiments, a tumour or tumour sample may further encompass tumour area occupied by tumour associated intratumoral cells and/or tumour associated stroma (e.g., contiguous peri-tumoral desmoplastic stroma). Tumour associated intratumoral cells and/or tumour associated stroma may include areas of immune infiltrates immediately adjacent to and/or contiguous with the main tumour mass.

[0089] In some embodiments, tumour cell staining is expressed as the percentage of all tumour cells showing staining (e.g., membranous, cytoplasmic or nuclear staining) of any intensity.

[0090] Infiltrating immune cell staining may be expressed as the percentage of the total tumour area occupied by immune cells that show staining of any intensity. The total tumour area encompasses the malignant cells as well as tumour-associated stroma, including areas of immune infiltrates immediately adjacent to and contiguous with the main tumour mass. In addition, infiltrating immune cell staining may be expressed as the percent of all tumour infiltrating immune cells.

[0091] In some embodiments, the tumour is a malignant cancerous tumour (i.e., cancer). In some embodiments, the tumour and/or cancer is a solid tumour or a non-solid or soft tissue tumour. Examples of soft tissue tumours include leukemia (e.g., chronic myelogenous leukemia, acute myelogenous leukemia, adult acute lymphoblastic leukemia, acute myelogenous leukemia, mature B-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia, prolymphocytic leukemia, or hairy cell leukemia) or lymphoma (e.g., non-Hodgkin's lymphoma, cutaneous T-cell lymphoma, or Hodgkin's disease). A solid tumour includes any cancer of body tissues other than blood, bone marrow, or the lymphatic system. Solid tumours can be further divided into those of epithelial cell origin and those of non-epithelial cell origin. Examples of epithelial cell solid tumours include tumours of the gastrointestinal tract, colon, colorectal (e.g., basaloid colorectal carcinoma), breast, prostate, lung, kidney, liver, pancreas, ovary (e.g. , endometrioid ovarian carcinoma), head and neck, oral cavity, stomach, duodenum, small intestine, large intestine, anus, gall bladder, labium, nasopharynx, skin, uterus, male genital organ, urinary organs (e.g., urothelium carcinoma, dysplastic urothelium carcinoma, transitional cell carcinoma), bladder, and skin. Solid tumours of non-epithelial origin include

sarcomas, brain tumours, and bone tumours. In some embodiments, the cancer is non-small cell lung cancer (NSCLC). In some embodiments, the cancer is a second-line or third-line locally advanced or metastatic non-small cell lung cancer. In some embodiments, the cancer is adenocarcinoma. In some embodiments, the cancer is a squamous cell carcinoma. In some embodiments, the cancer is non-small cell lung cancer (NSCLC), glioblastoma, neuroblastoma, melanoma, breast carcinoma (e.g., triple-negative breast cancer), gastric cancer, colorectal cancer (CRC), or hepatocellular carcinoma. In some embodiments, the cancer is a primary tumour. In some embodiments, the cancer is a metastatic tumour at a second site derived from any of the above types of cancer.

3. Methods of determining treatment responsiveness

[0092] The biomarkers described herein as being useful for indicating the likelihood of a cancer responding to an anti-cancer therapy (also referred to herein as "response to therapy biomarkers") are IL-6, EpCAM, ABCB5, and GBP2. In this regard, the present invention provides a method of determining the likelihood of a cancer responding to an anti-cancer therapy, the method comprising the step of determining an expression level of one or a plurality of biomarkers in a biological sample obtained from the subject, wherein the biomarkers comprise one or more of IL-6, EpCAM, ABCB5, and GBP2, and an altered or modulated expression level of the one or the one or a plurality of markers indicates or correlates with relatively increased or decreased likelihood of the cancer responding to the anti-cancer treatment.

4. Sample preparation

[0093] Generally, a sample from a subject with cancer is processed prior to cancer recurrence biomarker and/or cancer therapy biomarker detection or quantification. For example, nucleic acid and/or proteins may be extracted, isolated, and/or purified from a sample prior to analysis. Various DNA, mRNA, and/or protein extraction techniques are well known to those skilled in the art. Processing may include centrifugation, ultracentrifugation, ethanol precipitation, filtration, fractionation, resuspension, dilution, concentration, etc. In some embodiments, methods and systems provide analysis (e.g. , quantification of RNA or protein biomarkers) from raw sample (e.g., biological fluid such as blood, serum, etc.) without or with limited processing . In some examples, whole cells or tissue sections are isolated and analysed for cancer recurrence and/or cancer therapy biomarker expression, such as using immunohistochemistry (IHC) or flow cytometry.

[0094] Methods may comprise steps of homogenizing a sample in a suitable buffer, removal of contaminants and/or assay inhibitors, adding a cancer recurrence and/or cancer therapy biomarker capture reagent (e.g., a magnetic bead to which is linked an oligonucleotide complementary to a target cancer recurrence and/or cancer therapy biomarker), incubated under conditions that promote the association (e.g., by hybridization) of the target biomarker with the capture reagent to produce a target biomarker: capture reagent complex,

incubating the target biomarker: capture complex under target biomarker-release conditions. In some embodiments, multiple cancer recurrence and/or cancer therapy biomarkers are isolated in each round of isolation by adding multiple cancer recurrence and/or cancer therapy biomarker capture reagents (e.g., specific to the desired biomarkers) to the solution. For example, multiple cancer therapy biomarker capture reagents, each comprising an oligonucleotide specific for a different cancer recurrence and/or cancer therapy biomarker can be added to the sample for isolation of multiple cancer recurrence and/or cancer therapy biomarkers. It is contemplated that the methods encompass multiple experimental designs that vary both in the number of capture steps and in the number of target cancer recurrence and/or cancer therapy biomarkers captured in each capture step.

[0095] In some embodiments, capture reagents are molecules, moieties, substances, or compositions that preferentially (e.g., specifically and selectively) interact with a particular biomarker sought to be isolated, purified, detected, and/or quantified. Any capture reagent having desired binding affinity and/or specificity to the particular cancer recurrence and/or cancer therapy biomarker can be used in the present technology.

[0096] For example, the capture reagent can be a macromolecule such as a peptide, a protein (e.g., an antibody or other ligand that binds to a cancer therapy biomarker), an oligonucleotide, a nucleic acid (e.g. , nucleic acids capable of hybridizing with the cancer therapy biomarkers), oligosaccharides, carbohydrates, lipids, or small molecules, or a complex thereof. As illustrative and non-limiting examples, an avidin target capture reagent may be used to isolate and purify targets comprising a biotin moiety, an antibody may be used to isolate and purify targets comprising the appropriate antigen or epitope, and an oligonucleotide may be used to isolate and purify a complementary polynucleotide.

[0097] Any nucleic acids, including single-stranded and double-stranded nucleic acids, that are capable of binding, or specifically binding, to a target cancer recurrence and/or cancer therapy biomarker can be used as the capture reagent. Examples of such nucleic acids include DNA, RNA, aptamers, peptide nucleic acids, and other modifications to the sugar, phosphate, or nucleoside base. Thus, there are many strategies for capturing a target and accordingly many types of capture reagents are known to those in the art.

[0098] In addition, cancer recurrence and/or therapy biomarker capture reagents may comprise a functionality to localize, concentrate, aggregate, etc. the capture reagent and thus provide a way to isolate and purify the target cancer therapy biomarker when captured (e.g., bound, hybridized, etc.) to the capture reagent (e.g., when a target: capture reagent complex is formed). For example, in some embodiments the portion of the capture reagent that interacts with the cancer recurrence and/or cancer therapy biomarker (e.g., an oligonucleotide) is linked to a solid support (e.g., a bead, surface, resin, column, and the like) that allows manipulation by the user on a macroscopic scale. Often, the solid support allows the use of a mechanical means to isolate and purify the target: capture reagent complex from a

heterogeneous solution. For example, when linked to a bead, separation is achieved by removing the bead from the heterogeneous solution, e.g., by physical movement. In embodiments in which the bead is magnetic or paramagnetic, a magnetic field is used to achieve physical separation of the capture reagent (and thus the target cancer therapy biomarker) from the heterogeneous solution.

[0099] The cancer recurrence or cancer therapy biomarkers may be quantified or detected using any suitable technique. In specific embodiments, the cancer therapy biomarkers are quantified using reagents that determine the level, abundance or amount of individual cancer recurrence and/or cancer therapy biomarkers, either as isolated biomarker or as expressed in or on a cell. Non-limiting reagents of this type include reagents for use in nucleic acid- and protein-based assays.

[0100] In particular examples, when assessing the BRAF/NRAS mutational status, the genomic DNA from a sample is isolated and subjected to sequencing {e.g. next-generation sequencing, pyrosequencing, Sanger sequencing etc., as is well-known to those skilled in the art).

[0101] Cancer cells for the practice of the present invention can be obtained from any suitable cancer-cell containing patient samples, illustrative examples of which include tumour biopsies, circulating tumour cells (CTCs), primary cell cultures, or cell lines derived from tumours or exhibiting tumour-like properties, as well as preserved tumour samples, such as formalin-fixed, paraffin-embedded tumour samples, or frozen tumour samples. In some embodiments, the sample is obtained prior to treatment with a therapy. In other embodiments, the sample is obtained after treatment with a therapy. In some embodiments, the sample comprises a tumour tissue sample, which can be formalin fixed and paraffin embedded, archival, fresh or frozen. In some embodiments, the sample is -. In some embodiments, the whole blood comprises immune cells, circulating tumor cells, and any combinations thereof.

[0102] Presence and/or levels/amount of a biomarker (e.g., any one or more of the association between p85β and H3K27Me3; PI3K2R; IL-6, EpCAM, ABCB5, and GBP2) can be determined qualitatively and/or quantitatively based on any suitable criterion known in the art, including but not limited to proteins and protein fragments. In certain embodiments, presence and/or expression levels/amount of a biomarker in a first sample is increased or elevated as compared to presence/absence and/or expression levels/amount in a second sample (e.g., before treatment with a therapy). In certain embodiments, presence/absence and/or levels/amount of a biomarker in a first sample is decreased or reduced as compared to presence and/or levels/amount in a second sample. In certain embodiments, the second sample is a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue. Additional disclosures for determining presence/absence and/or levels/amount of a gene are described herein.

[0103] In some embodiments of any of the methods, an elevated level refers to an overall increase of about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater, in the level of biomarker (e.g., protein or nucleic acid), detected by standard art known methods such as those described herein, as compared to a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue. In certain embodiments, an elevated level refers to the increase in level/amount of a biomarker in the sample wherein the increase is at least about any of 1.5×, 1.75×, 2×, 3×, 4×, 5×, 6×, 7×, 8×, 9×, 10×, 25×, 50×, 75×, or 100× the level/amount of the respective biomarker in a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue. In some embodiments, an elevated level refers to an overall increase of greater than about 1.5-fold, about 1.75-fold, about 2-fold, about 2.25-fold, about 2.5-fold, about 2.75-fold, about 3.0-fold, or about 3.25-fold as compared to a reference sample, reference cell, reference tissue, control sample, control cell, control tissue, or internal control (e.g., housekeeping gene).

[0104] In some embodiments of any of the methods, a reduced level refers to an overall reduction of about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater, in the level of biomarker (e.g., protein or nucleic acid (e.g., gene or mRNA)), detected by standard art known methods such as those described herein, as compared to a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue. In certain embodiments, reduced level refers to a decrease in level/amount of a biomarker in the sample wherein the decrease is at least about any of 0.9×, 0.8×, 0.7×, 0.6×, 0.5×, 0.4×, 0.3×, 0.2×, 0.1×, 0.05×, or 0.01× the level/amount of the respective biomarker in a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue.

[0105] Presence and/or level/amount of various biomarkers in a sample can be analyzed by a number of methodologies, many of which are known in the art and understood by the skilled artisan, including, but not limited to, immunohistochemistry ("IHC"), Western blot analysis, immunoprecipitation, molecular binding assays, ELISA, ELIFA, fluorescence activated cell sorting ("FACS"), MassARRAY, proteomics, quantitative blood based assays (as for example serum ELISA), biochemical enzymatic activity assays, *in situ* hybridization, Southern analysis, Northern analysis, whole genome sequencing, polymerase chain reaction ("PCR") including quantitative real time PCR ("qRT-PCR") and other amplification type detection methods, such as, for example, branched DNA, SISBA, TMA and the like), RNA-Seq, FISH, microarray analysis, gene expression profiling, and/or serial analysis of gene expression ("SAGE"), as well as any one of the wide variety of assays that can be performed by protein, gene, and/or tissue array analysis. Typical protocols for evaluating the status of genes and gene products are found, for example in Ausubel et al., eds., 1995, Current Protocols In Molecular Biology, Units 2 (Northern Blotting), 4 (Southern Blotting), 15 (Immunoblotting) and 18 (PCR Analysis). Multiplexed immunoassays such as those available from Rules Based Medicine or Meso Scale Discovery ("MSD") may also be used.

[0106] In some embodiments, presence and/or level/amount of a biomarker, particularly for non-PTM biomarkers such as the biomarkers disclosed herein, is determined using a method comprising: (a) performing gene expression profiling, PCR (such as RT-PCR or qRT-PCR), RNA-seq, microarray analysis, SAGE, MassARRAY technique, or FISH on a sample (such as a subject cancer sample); and (b) determining presence and/or expression level/amount of a biomarker in the sample. In some embodiments, the microarray method comprises the use of a microarray chip having one or more nucleic acid molecules that can hybridize under stringent conditions to a nucleic acid molecule encoding a gene mentioned above or having one or more polypeptides (such as peptides or antibodies) that can bind to one or more of the proteins encoded by the genes mentioned above. In one embodiment, the PCR method is qRT-PCR. In one embodiment, the PCR method is multiplex-PCR. In some embodiments, gene expression is measured by microarray. In some embodiments, gene expression is measured by qRT-PCR. In some embodiments, expression is measured by multiplex-PCR.

[0107] Methods for the evaluation of mRNAs in cells are well known and include, for example, hybridization assays using complementary DNA probes (such as *in situ* hybridization using labeled riboprobes specific for the one or more genes, Northern blot and related techniques) and various nucleic acid amplification assays (such as RT-PCR using complementary primers specific for one or more of the genes, and other amplification type detection methods, such as, for example, branched DNA, SISBA, TMA and the like).

[0108] Samples from mammals can be conveniently assayed for mRNAs using Northern blot, dot blot or PCR analysis. In addition, such methods can include one or more steps that allow one to determine the levels of target mRNA in a biological sample (e.g., by simultaneously examining the levels a comparative control mRNA sequence of a "housekeeping" gene such as an actin family member). Optionally, the sequence of the amplified target cDNA can be determined.

[0109] Optional methods include protocols which examine or detect mRNAs, such as target mRNAs, in a tissue or cell sample by microarray technologies. Using nucleic acid microarrays, test and control mRNA samples from test and control tissue samples are reverse transcribed and labeled to generate cDNA probes. The probes are then hybridized to an array of nucleic acids immobilized on a solid support. The array is configured such that the sequence and position of each member of the array is known. For example, a selection of genes whose expression correlates with increased or reduced clinical benefit of a therapy may be arrayed on a solid support. Hybridization of a labeled probe with a particular array member indicates that the sample from which the probe was derived expresses that gene.

[0110] In preferred embodiments, presence and/or level/amount is measured by observing protein levels. In certain embodiments, the method comprises contacting the biological sample with an antibody to at least one cancer recurrence biomarker (e.g., p85β

and/or H3K27Me3) or response to therapy biomarker (e.g., IL-6, EpCAM, ABCB5, and/or GBP2) under conditions permissive for binding of the biomarker(s), and detecting whether a complex is formed between the antibody or antibodies and the biomarker(s). Such method may be an *in vitro* or *in vivo* method. In some embodiments, one or more anti-biomarker antibodies are used to select subjects eligible for a therapy e.g., a cytotoxic therapy or an immunotherapy. In some embodiments of this type the therapy is an anti-PD1 immunotherapy, or a PARP inhibitor therapy (e.g., Olaparib).

[0111] In certain embodiments, the presence and/or expression level/amount of biomarker proteins in a sample is examined using IHC and staining protocols. IHC staining of tissue sections has been shown to be a reliable method of determining or detecting presence of proteins in a sample. In some embodiments, the level of a response to therapy biomarker (e.g., IL-6, EpCAM, ABCB5, and/or GBP2) in a sample from an individual is an elevated level and, in further embodiments, is determined using IHC. In one embodiment, the level of biomarker is determined using a method comprising: (a) performing IHC analysis of a sample (such as a subject cancer sample) with an antibody; and (b) determining the level of a biomarker in the sample. In some embodiments, IHC staining intensity is determined relative to a reference. In some embodiments, the reference is a reference value. In some embodiments, the reference is a reference sample (e.g., control cell line staining sample or tissue sample from non-cancerous patient).

[0112] In some embodiments, expression of at least one response to therapy biomarker (e.g., IL-6, EpCAM, ABCB5, and/or GBP2) is evaluated on a tumour or tumour sample. As used herein, a tumour or tumour sample may encompass part or all of the tumour area occupied by tumour cells. In some embodiments, a tumour or tumour sample may further encompass tumour area occupied by tumour associated intratumoural cells and/or tumour associated stroma (e.g., contiguous peri-tumoral desmoplastic stroma). Tumour associated intratumoural cells and/or tumour associated stroma may include areas of immune infiltrates (e.g., tumour infiltrating immune cells as described herein) immediately adjacent to and/or contiguous with the main tumour mass. In some embodiments, response to therapy biomarker expression is evaluated on tumour cells.

[0113] In alternative methods, the sample may be contacted with an antibody specific for said biomarker under conditions sufficient for an antibody-biomarker complex to form, and then detecting said complex. The presence of the biomarker may be detected in a number of ways, such as by Western blotting and ELISA procedures for assaying a wide variety of tissues and samples, including plasma or serum. A wide range of immunoassay techniques using such an assay format are available, see, e.g., U.S. Patent Nos. 4,016,043, 4,424,279 and 4,018,653. These include both single-site and two-site or "sandwich" assays of the non-competitive types, as well as in the traditional competitive binding assays. These assays also include direct binding of a labeled antibody to a target biomarker.

[0114] In certain embodiments, the samples are normalized for both differences in the amount of the biomarker assayed and variability in the quality of the samples used, and variability between assay runs. Such normalization may be accomplished by detecting and incorporating the expression of certain normalizing biomarkers, including expression products of well-known housekeeping genes. Alternatively, normalization can be based on the mean or median signal of all of the assayed genes or a large subset thereof (global normalization approach). On a gene-by-gene basis, measured normalized amount of a subject tumour mRNA or protein is compared to the amount found in a reference set. Normalized expression levels for each mRNA or protein per tested tumor per subject can be expressed as a percentage of the expression level measured in the reference set. The presence and/or expression level/amount measured in a particular subject sample to be analyzed will fall at some percentile within this range, which can be determined by methods well known in the art.

[0115] In some embodiments, the sample is a clinical sample. In some embodiments, the sample is obtained from a primary or metastatic tumour. Tissue biopsy is often used to obtain a representative piece of tumour tissue. Alternatively, tumour cells can be obtained indirectly in the form of tissues or fluids that are known or thought to contain the tumour cells of interest. For instance, samples of lung cancer lesions may be obtained by resection, bronchoscopy, fine needle aspiration, bronchial brushings, or from sputum, pleural fluid or blood. Genes or gene products can be detected from cancer or tumour tissue or from other body samples such as urine, sputum, serum or plasma. The same techniques discussed above for detection of target genes or gene products in cancerous samples can be applied to other body samples. Cancer cells may be sloughed off from cancer lesions and appear in such body samples. By screening such body samples, a simple early diagnosis can be achieved for these cancers. In addition, the progress of therapy can be monitored more easily by testing such body samples for target genes or gene products.

[0116] In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is a single sample or combined multiple samples from the same subject or individual that are obtained at one or more different time points than when the test sample is obtained. For example, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is obtained at an earlier time point from the same subject or individual than when the test sample is obtained. Such reference sample, reference cell, reference tissue, control sample, control cell, or control tissue may be useful if the reference sample is obtained during initial diagnosis of cancer and the test sample is later obtained when the cancer becomes metastatic.

[0117] In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is a combination of multiple samples from one or more healthy individuals who are not the subject or individual. In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is a combination of multiple samples from one or more individuals with a disease or disorder

(e.g., cancer) who are not the subject or individual. In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is pooled RNA samples from normal tissues or pooled plasma or serum samples from one or more individuals who are not the subject or individual. In certain embodiments, a reference sample, reference cell, reference tissue, control sample, control cell, or control tissue is pooled RNA samples from tumour tissues or pooled plasma or serum samples from one or more individuals with a disease or disorder (e.g., cancer) who are not the subject or individual.

[0118] In some embodiments, the sample is a tissue sample from the individual. In some embodiments, the tissue sample is a tumour tissue sample (e.g., biopsy tissue). In some embodiments, the tissue sample is lung tissue. In some embodiments, the tissue sample is renal tissue. In some embodiments, the tissue sample is skin tissue. In some embodiments, the tissue sample is pancreatic tissue. In some embodiments, the tissue sample is gastric tissue. In some embodiments, the tissue sample is bladder tissue. In some embodiments, the tissue sample is esophageal tissue. In some embodiments, the tissue sample is mesothelial tissue. In some embodiments, the tissue sample is breast tissue. In some embodiments, the tissue sample is thyroid tissue. In some embodiments, the tissue sample is colorectal tissue. In some embodiments, the tissue sample is head and neck tissue. In some embodiments, the tissue sample is osteosarcoma tissue. In some embodiments, the tissue sample is prostate tissue. In some embodiments, the tissue sample is ovarian tissue, HCC (liver), blood cells, lymph nodes, and/or bone/bone marrow tissue. In some embodiments, the tissue sample is colon tissue. In some embodiments, the tissue sample is endometrial tissue. In some embodiments, the tissue sample is brain tissue (e.g., glioblastoma, neuroblastoma, and so forth).

[0119] In some embodiments, a tumour tissue sample (the term "tumour sample" is used interchangeably herein) may encompass part or all of the tumour area occupied by tumour cells. In some embodiments, a tumour or tumour sample may further encompass tumour area occupied by tumour associated intratumoral cells and/or tumour associated stroma (e.g., contiguous peri-tumoral desmoplastic stroma). Tumour associated intratumoral cells and/or tumour associated stroma may include areas of immune infiltrates immediately adjacent to and/or contiguous with the main tumour mass.

[0120] In some embodiments, tumour cell staining is expressed as the percentage of all tumour cells showing staining (e.g., membranous, cytoplasmic or nuclear staining) of any intensity.

[0121] Infiltrating immune cell staining may be expressed as the percentage of the total tumour area occupied by immune cells that show staining of any intensity. The total tumour area encompasses the malignant cells as well as tumour-associated stroma, including areas of immune infiltrates immediately adjacent to and contiguous with the main tumour mass. In addition, infiltrating immune cell staining may be expressed as the percentage of all tumour infiltrating immune cells.

[0122] In some embodiments, the tumour is a malignant cancerous tumour (i.e., cancer). In some embodiments, the tumour and/or cancer is a solid tumour or a non-solid or soft tissue tumour. Examples of soft tissue tumours include leukemia (e.g., chronic myelogenous leukemia, acute myelogenous leukemia, adult acute lymphoblastic leukemia, acute myelogenous leukemia, mature B-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia, prolymphocytic leukemia, or hairy cell leukemia) or lymphoma (e.g., non-Hodgkin's lymphoma, cutaneous T-cell lymphoma, or Hodgkin's disease). A solid tumour includes any cancer of body tissues other than blood, bone marrow, or the lymphatic system. Solid tumours can be further divided into those of epithelial cell origin and those of non-epithelial cell origin. Examples of epithelial cell solid tumours include tumours of the gastrointestinal tract, colon, colorectal (e.g., basaloid colorectal carcinoma), breast, prostate, lung, kidney, liver, pancreas, ovary (e.g. , endometrioid ovarian carcinoma), head and neck, oral cavity, stomach, duodenum, small intestine, large intestine, anus, gall bladder, labium, nasopharynx, skin, uterus, male genital organ, urinary organs (e.g., urothelium carcinoma, dysplastic urothelium carcinoma, transitional cell carcinoma), bladder, and skin. Solid tumours of non-epithelial origin include sarcomas, brain tumours, and bone tumours. In some embodiments, the cancer is non-small cell lung cancer (NSCLC). In some embodiments, the cancer is a second-line or third-line locally advanced or metastatic non-small cell lung cancer. In some embodiments, the cancer is adenocarcinoma. In some embodiments, the cancer is a squamous cell carcinoma. In some embodiments, the cancer is non-small cell lung cancer (NSCLC), glioblastoma, neuroblastoma, melanoma, breast carcinoma (e.g., triple-negative breast cancer), gastric cancer, colorectal cancer (CRC), or hepatocellular carcinoma. In some embodiments, the cancer is a primary tumour. In some embodiments, the cancer is a metastatic tumour at a second site derived from any of the above types of cancer.

[0123] In some embodiments, the at least one response to therapy biomarker is detected in the sample using a method selected from the group consisting of FACS, Western blot, ELISA, immunoprecipitation, immunohistochemistry, immunofluorescence, radioimmunoassay, dot blotting, immunodetection methods, HPLC, surface plasmon resonance, optical spectroscopy, mass spectrometry, HPLC, qPCR, RT-qPCR, multiplex qPCR or RT-qPCR, RNA-seq, microarray analysis, SAGE, MassARRAY technique, and FISH, and combinations thereof. In some embodiments, the least one response to therapy biomarker is/ detected using FACS analysis. In some embodiments, the at least one response to therapy biomarker is detected in blood samples. In some embodiments, the at least one response to therapy biomarker is detected in circulating tumour cells in blood samples. Any suitable method to isolate/enrich such population of cells may be used including, but not limited to, cell sorting. In some embodiments, expression of IL-6 is elevated in samples from individuals that respond to treatment with a therapy, suitably chemotherapy, and/or an immunotherapy (e.g., one that comprises an anti-immune checkpoint molecule antibody, illustrative examples of which include an anti-PD-1 antibody; or a PARP inhibitor therapy). In some embodiments, expression of

EpCAM is elevated in samples from individuals that respond to treatment with a therapy, suitably chemotherapy, and/or an immunotherapy (e.g., one that comprises an anti-immune checkpoint molecule antibody, illustrative examples of which include an anti-PD-1 antibody; or a PARP inhibitor therapy). In some embodiments, expression of and ABCB5 is elevated in samples from individuals that respond to treatment with a therapy, suitably chemotherapy, and/or an immunotherapy (e.g., one that comprises an anti-immune checkpoint molecule antibody, illustrative examples of which include an anti-PD-1 antibody; or a PARP inhibitor therapy).

[0124] Also provided herein are methods for monitoring the pharmacodynamic activity (PD activity) of a therapy (e.g., PI3K inhibitor therapy and/or an immunotherapy), by determining the level or amount of at least one treatment monitoring biomarker (e.g., *DCN*, *LCN2*, *LOX*, *HMOX1*, and/or *PDCD1LG2*) in a sample comprising cancer cells obtained from a subject, wherein the subject has been treated with the therapy, and determining the treatment as demonstrating pharmacodynamic activity based on the expression level of the at least one treatment monitoring biomarker in the sample obtained from the subject, as compared with a reference, where: (1) an unchanged level of at least one treatment monitoring biomarker in the cancer cell relative to a suitable control (e.g., a cancer cell of the patient, which expresses the treatment monitoring biomarker, before exposure to the therapy) indicates no or weak pharmacodynamic activity to the therapy, (2) an elevated level of the treatment monitoring biomarker in the cancer cell relative to a suitable control (e.g., a cancer cell of the patient before exposure to the therapy) indicates significant or strong pharmacodynamic activity to the therapy, (3) a decreased level of the treatment monitoring biomarker in the cancer cell relative to a suitable control (e.g., a cancer cell of the patient before exposure to the therapy) indicates no or weak pharmacodynamic activity to the therapy. Expression level of the biomarker(s) and/or cellular composition may be measured by one or more methods as described herein.

[0125] In some embodiments, the expression level of one or more biomarker genes, proteins and/or cellular composition may be compared to a reference which may include a sample from a subject not receiving a therapy (e.g., a PI3K inhibitor therapy). In some embodiments, a reference may include a sample from the same subject before receiving a therapy (e.g., a cytotoxic therapy or an immunotherapy). In some embodiments, a reference may include a reference value from one or more samples of other subjects receiving a therapy (e.g., a cytotoxic therapy or an immunotherapy). For example, a population of patients may be treated, and a mean, average, or median value for expression level of the at least one response to therapy biomarker, cancer recurrence biomarker, and/or treatment monitoring biomarker, may be generated from the population as a whole. A set of samples obtained from cancers having a shared characteristic (e.g., the same cancer type and/or stage, or exposure to a common therapy) may be studied from a population, such as with a clinical outcome study. This set may be used to derive a reference, e.g., a reference number, to which a subject's sample may be

compared. Any of the references described herein may be used as a reference for monitoring PD activity.

[0126] Certain aspects of the present disclosure relate to measurement of the expression level of one or more biomarkers (e.g., gene expression products including mRNAs and proteins) in a sample. In some embodiments, a sample may include cancer cells. In some embodiments, the sample may be a peripheral blood sample (e.g., from a patient having a tumour). In some embodiments, the sample is a tumour sample. A tumour sample may include cancer cells, lymphocytes, leukocytes, stroma, blood vessels, connective tissue, basal lamina, and any other cell type in association with the tumor. In some embodiments, the sample is a tumour tissue sample containing tumor-infiltrating leukocytes. In some embodiments, the sample may be processed to separate or isolate one or more cell types (e.g., leukocytes). In some embodiments, the sample may be used without separating or isolating cell types.

[0127] A tumour sample may be obtained from a subject by any method known in the art, including without limitation a biopsy, endoscopy, or surgical procedure. In some embodiments, a tumour sample may be prepared by methods such as freezing, fixation (e.g., by using formalin or a similar fixative), and/or embedding in paraffin wax. In some embodiments, a tumour sample may be sectioned. In some embodiments, a fresh tumour sample (i.e., one that has not been prepared by the methods described above) may be used. In some embodiments, a tumor sample may be prepared by incubation in a solution to preserve mRNA and/or protein integrity.

[0128] In some embodiments, the sample may be a peripheral blood sample. A peripheral blood sample may include white blood cells, PBMCs, and the like. Any technique known in the art for isolating leukocytes from a peripheral blood sample may be used. For example, a blood sample may be drawn, red blood cells may be lysed, and a white blood cell pellet may be isolated and used for the sample. In another example, density gradient separation may be used to separate leukocytes (e.g., PBMCs) from red blood cells. In some embodiments, a fresh peripheral blood sample (i.e., one that has not been prepared by the methods described above) may be used. In some embodiments, a peripheral blood sample may be prepared by incubation in a solution to preserve mRNA and/or protein integrity.

[0129] In some embodiments, responsiveness to therapy may refer to any one or more of: extending survival (including overall survival and progression free survival); resulting in an objective response (including a complete response or a partial response); or improving signs or symptoms of cancer. In some embodiments, responsiveness may refer to improvement of one or more factors according to the published set of RECIST guidelines for determining the status of a tumour in a cancer patient (i.e., responding, stabilizing, or progressing). For a more detailed discussion of these guidelines, see, Eisenhauer et al. (2009, *Eur. J. Cancer* 45: 228-47), Topalian et al. (2012, *N. Engl. J. Med.* 366:2443-54), Wolchok et al. (2009, *Clin. Can. Res.* 15: 7412-20) and Therasse et al. (2000, *J. Natl. Cancer Inst.* 92: 205-16). A responsive subject

may refer to a subject whose cancer(s) show improvement, e.g., according to one or more factors based on RECIST criteria. A non-responsive subject may refer to a subject whose cancer(s) do not show improvement, e.g., according to one or more factors based on RECIST criteria.

[0130] Conventional response criteria may not be adequate to characterize the anti-tumour activity of therapeutic agents of the invention, which can produce delayed responses that may be preceded by initial apparent radiological progression, including the appearance of new lesions. Therefore, modified response criteria have been developed that account for the possible appearance of new lesions and allow radiological progression to be confirmed at a subsequent assessment. Accordingly, in some embodiments, responsiveness may refer to improvement of one of more factors according to immune-related response criteria (irRC) (see, e.g., Wolchok et al. (2009, supra)). In some embodiments, new lesions are added into the defined tumour burden and followed, e.g., for radiological progression at a subsequent assessment. In some embodiments, presence of non-target lesions is included in assessment of complete response and not included in assessment of radiological progression. In some embodiments, radiological progression may be determined only on the basis of measurable disease and/or may be confirmed by a consecutive assessment >4 weeks from the date first documented.

[0131] In some embodiments, responsiveness may include immune activation. In some embodiments, responsiveness may include treatment efficacy. In some embodiments, responsiveness may include immune activation and treatment efficacy.

5. Biomarker panels

[0132] The biomarkers of the present invention can be used in predictive and/or prognostic tests to assess, determine, and/or qualify (used interchangeably herein) response to therapy signature status in a patient and therefore, direct treatment of the patient. The phrase "response to therapy signature status" includes a high response to therapy signature (RT high) and a low response to therapy signature (RT low). Based on this status, further procedures may be indicated, including additional tests or therapeutic procedures or regimens.

[0133] These and other biomarkers are disclosed herein, and it is understood that when combinations, subsets, interactions, groups, etc., of these biomarkers 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. Thus, if a panel of biomarkers A, B, and C are disclosed as well as a class of biomarkers D, E, and F and an example of a combination panel A-D is disclosed, then even if each is not individually recited each is individually and collectively contemplated meaning combinations, A-E, A-F, B-D, B-E, B-F, C-D, C-E, and C-F are considered disclosed. Likewise, any subset or combination of these is also disclosed. Thus, for example, the sub-group of A-E, B-F, and C-E would be considered disclosed. This concept applies to all aspects of this

application including, but not limited to, steps in methods of using the disclosed biomarkers. 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 embodiments of the disclosed methods.

[0134] The response to therapy signature panel suitably includes one or more of IL-6, EpCAM, ABCB5, and/or GBP2. Non-limiting examples of these signatures include IL-6, and at least one other biomarker selected from the following biomarker combinations: (a) EpCAM; (b) EpCAM and ABCB5; (c) EpCAM and GBP2; (d) EpCAM, ABCB5, and GBP2; (e) ABCB5; (f) ABCB5, and GBP2; (g) EpCAM, ABCB5, and GBP2; and (h) GBP2.

[0135] The power of an assay to correctly predict response to therapy is commonly measured as the sensitivity of the assay, the specificity of the assay or the area under a receiver operated characteristic ("ROC") curve. Sensitivity is the percentage of true positives that are predicted by a test to be positive, while specificity is the percentage of true negatives that are predicted by a test to be negative. An ROC curve provides the sensitivity of a test as a function of 1-specificity. The greater the area under the ROC curve, the more powerful the predictive value of the test. Other useful measures of the utility of a test are positive predictive value and negative predictive value. Positive predictive value is the percentage of people who test positive that are actually positive. Negative predictive value is the percentage of people who test negative that are actually negative.

[0136] In particular embodiments, the biomarker signatures of the present invention may show a statistical difference in different response to therapy statuses of at least $p<0.05$, $p<10^{-2}$, $p<10^{-3}$, $p<10^{-4}$ or $p<10^{-5}$. Predictive or prognostic tests that use these biomarkers may show an ROC of at least 0.6, at least about 0.7, at least about 0.8, or at least about 0.9.

[0137] In certain embodiments, the biomarkers are measured in a patient sample using the methods described herein and a response to therapy signature status is calculated. In particular embodiments, the measurement(s) may then be compared with a relevant predictive or prognostic amount(s), cut-off(s), or multivariate model scores that distinguish a high therapy response signature (RT high) status from a low therapy response signature (RT low) status. The predictive or prognostic amount(s) represents a measured amount of a biomarker(s) above which or below which a patient is classified as having a particular therapy response signature status. As is well understood in the art, by adjusting the particular predictive or prognostic cut-off(s) used in an assay, one can increase sensitivity or specificity of the assay depending on the preference of the skilled person. In particular embodiments, the particular predictive or prognostic cut-off can be determined, for example, by measuring the level or amount of biomarkers in a statistically significant number of samples from patients with different response to therapy signature statuses, and drawing the cut-off to suit the desired levels of specificity and sensitivity.

[0138] Furthermore, in certain embodiments, the values measured for biomarkers of a biomarker panel are mathematically combined and the combined value is correlated to the underlying predictive or prognostic question of high or low response to therapy signature.

[0139] Biomarker values may be combined by any appropriate mathematical method known in the art. Well-known mathematical methods for correlating a biomarker combination to a disease status employ methods like discriminant analysis (DA) (e.g., linear-, quadratic-, regularized-DA), discriminant functional analysis (DFA), kernel methods (e.g., SVM), multidimensional scaling (MDS), nonparametric methods (e.g., k-nearest-neighbor classifiers), PLS (partial least squares), tree-based methods (e.g., logic regression, CART, random forest methods, boosting/bagging methods), generalized linear models (e.g., logistic regression), principal components based methods (e.g., SIMCA), generalized additive models, fuzzy logic based methods, neural networks and genetic algorithms based methods. The skilled artisan will have no problem in selecting an appropriate method to evaluate a biomarker combination of the present invention. In some embodiments, the method used in a correlating a biomarker combination of the present invention is selected from DA (e.g., linear-, quadratic-, regularized discriminant analysis), DFA, kernel methods (e.g., SVM), MDS, nonparametric methods (e.g., k-nearest-neighbor classifiers), PLS (partial least squares), tree-based methods (e.g., logic regression, CART, random forest methods, boosting methods), or Generalized linear models (e.g., logistic regression), and principal components analysis. Details relating to these statistical methods are found in the following references: Ruczinski et al., *J. Comp. Graph. Stat.* 475-511 (2003); Friedman, J. H., *J. Am. Stat. Ass.* 165-75 (1989); Hastie, Tibshirani, Friedman, The Elements of Statistical Learning, Springer Series in Statistics (2001); Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J. Classification and regression trees, California : Wadsworth (1984); Breiman, L., 45 *Machine Learning* 5-32 (2001); Pepe, M. S., The Statistical Evaluation of Medical Tests for Classification and Prediction, Oxford Statistical Science Series, 28 (2003); and Duda, R. O., Hart, P. E., Stork, D. G., Pattern Classification, Wiley Interscience, 2nd Edition (2001).

6. Generation of classification algorithms for qualifying response to therapy signature status

[0140] In some embodiments, data that are generated using samples such as "known samples" can then be used to "train" a classification model. A "known sample" is a sample that has been pre-classified. The data that are used to form the classification model can be referred to as a "training data set". The training data set that is used to form the classification model may comprise raw data or pre-processed data. Once trained, the classification model can recognize patterns in data generated using unknown samples. The classification model can then be used to classify the unknown samples into classes. This can be useful, for example, in predicting whether or not a particular biological sample is associated with a certain biological condition.

[0141] Classification models can be formed using any suitable statistical classification or learning method that attempts to segregate bodies of data into classes based on objective parameters present in the data. Classification methods may be either supervised or unsupervised. Examples of supervised and unsupervised classification processes are described in Jain, "Statistical Pattern Recognition: A Review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, January 2000, the teachings of which are incorporated by reference.

[0142] In supervised classification, training data containing examples of known categories are presented to a learning mechanism, which learns one or more sets of relationships that define each of the known classes. New data may then be applied to the learning mechanism, which then classifies the new data using the learned relationships. Examples of supervised classification processes include linear regression processes (e.g., multiple linear regression (MLR), partial least squares (PLS) regression and principal components regression (PCR)), binary decision trees (e.g., recursive partitioning processes such as CART), artificial neural networks such as back propagation networks, discriminant analyses (e.g., Bayesian classifier or Fischer analysis), logistic classifiers, and support vector classifiers (support vector machines).

[0143] Another supervised classification method is a recursive partitioning process. Recursive partitioning processes use recursive partitioning trees to classify data derived from unknown samples. Further details about recursive partitioning processes are provided in U.S. Patent Publication No. 2002/0138208 to Paulse et al., "Method for analyzing mass spectra."

[0144] In other embodiments, the classification models that are created can be formed using unsupervised learning methods. Unsupervised classification attempts to learn classifications based on similarities in the training data set, without pre-classifying the spectra from which the training data set was derived. Unsupervised learning methods include cluster analyses. A cluster analysis attempts to divide the data into "clusters" or groups that ideally should have members that are very similar to each other, and very dissimilar to members of other clusters. Similarity is then measured using some distance metric, which measures the distance between data items, and clusters together data items that are closer to each other. Clustering techniques include the MacQueen's K-means algorithm and the Kohonen's Self-Organizing Map algorithm.

[0145] Learning algorithms asserted for use in classifying biological information are described, for example, in PCT International Patent Publication No. WO 01/31580, U.S. Patent Publication Nos. 2002/0193950, 2003/0004402, and 2003/0055615

[0146] The classification models can be formed on and used on any suitable digital computer. Suitable digital computers include micro, mini, or large computers using any standard or specialized operating system, such as a Unix, Windows® or Linux™ based operating system. In embodiments utilizing a mass spectrometer, the digital computer that is used may be

physically separate from the mass spectrometer that is used to create the spectra of interest, or it may be coupled to the mass spectrometer.

[0147] The training data set and the classification models according to embodiments of the invention can be embodied by computer code that is executed or used by a digital computer. The computer code can be stored on any suitable computer readable media including optical or magnetic disks, sticks, tapes, etc., and can be written in any suitable computer programming language including R, C, C++, visual basic, etc.

[0148] The learning algorithms described above are useful both for developing classification algorithms for the biomarkers already discovered, and for finding new biomarkers. The classification algorithms, in turn, form the base for diagnostic/prognostic tests by providing diagnostic/prognostic values (e.g., cut-off points) for biomarkers used singly or in combination.

[0149] In some embodiments any of the classification methods disclosed herein may be performed at least in part by one or more computers and/or may be stored in a database on a non-transitory computer medium. In some embodiments any of the classification methods disclosed herein may be embodied or stored at least in part on a computer-readable medium having computer-executable instructions thereon. In some embodiments a computer-readable medium comprises any non-transitory and/or tangible computer-readable medium.

7. Deriving biomarker values

[0150] Biomarker values can be measured biomarker values, which are values of biomarkers directly measured for the subject, or alternatively could be "derived" biomarker values, which are values that have been derived from one or more measured biomarker values, for example by applying a function to the one or more measured biomarker values. As used herein, biomarkers to which a function has been applied are referred to as "derived biomarkers".

[0151] The biomarker values may be determined in any one of a number of ways that are well known in the art. For example, a comprehensive description of biomarker value determination can be found in Inti. Pat. Pub. No. WO 2015/117204, which is incorporated herein by reference in its entirety. In one example, the process of determining biomarker values can include measuring the biomarker values, for example by performing tests on the subject or on sample(s) obtained from the subject.

[0152] More typically, however, the step of determining the biomarker values includes having an electronic processing device receive or otherwise obtain biomarker values that have been previously measured or derived. This could include for example, retrieving the biomarker values from a data store such as a remote database, obtaining biomarker values that have been manually input, using an input device, or the like. Suitably, the indicator may be determined using a combination of a plurality of biomarker values, the indicator being at least partially indicative of cancer recurrence or responsiveness to cancer therapy. Assuming the

method is performed using an electronic processing device, an indication of the indicator is optionally displayed or otherwise provided to the user.

[0153] In some embodiments, biomarker values are combined, for example by adding, multiplying, subtracting, or dividing biomarker values to determine an indicator value. This step is performed so that multiple biomarker values can be combined into a single indicator value, providing a more useful and straightforward mechanism for allowing the indicator to be interpreted and hence used in determining the likelihood of a subject having recurrence of cancer and/or responding to cancer therapy.

[0154] It will be understood that in this context, the biomarkers used within the above-described method can define a biomarker profile for cancer recurrence or cancer therapy responsiveness, which includes a minimal number of biomarkers (e.g., at least one biomarker), whilst maintaining sufficient performance to allow the biomarker profile to be used in making a clinically relevant determination. Minimizing the number of biomarkers used minimizes the costs associated with performing diagnostic or prognostic tests and in the case of polypeptide biomarkers, allows the test to be performed utilizing relatively straightforward techniques such as quantitative RT-PCR and/or immunofluorescence, and allowing the test to be performed rapidly in a clinical environment. In this regard, the indication provided by the methods described herein could be a graphical or alphanumeric representation of an indicator value. Alternatively, however, the indication could be the result of a comparison of the indicator value to predefined thresholds or ranges, or alternatively could be an indication of the likelihood of cancer recurrence or responsiveness of a subject to cancer therapy.

[0155] Furthermore, producing a single indicator value allows the results of the test to be easily interpreted by a clinician or other medical practitioner, so that test can be used for reliable diagnosis in a clinical environment.

[0156] Solely by way of an illustration, the indicator-determining methods suitably include determining at least one biomarker value, wherein the biomarker value is a value measured or derived for at least one cancer therapy biomarker of the subject and is at least partially indicative of a concentration or abundance of the cancer recurrence or cancer therapy biomarker in a sample taken from the subject. By way of an example, in some embodiments the cancer recurrence biomarker comprises an expression product of MAP1LC3B.

[0157] The derived biomarker value is then used to determine the indicator for use in determining the likelihood of a subject responding to cancer therapy, either by using the derived biomarker value as an indicator value, or by performing additional processing, such as comparing the derived biomarker value to a reference or the like, as generally known in the art and as described in more detail below, or to another biomarker value. In some embodiments, the indicator is indicative of a level, concentration or abundance of an expression product of PIK3R2. In other embodiments, the indicator is indicative of a level or abundance of an expression product of PIK3R2.

[0158] The derived biomarker values could be combined using a combining function such as an additive model; a linear model; a support vector machine; a neural network model; a random forest model; a regression model; a genetic algorithm; an annealing algorithm; a weighted sum; a nearest neighbor model; and a probabilistic model. In some embodiments, the indicator is compared to an indicator reference, with a likelihood of responsiveness to cancer being determined in accordance with results of the comparison. The indicator reference may be derived from indicators determined for a number of individuals in a reference population. The reference population typically includes individuals having different characteristics, such as a plurality of individuals of different sexes; and/or ethnicities, with different groups being defined based on different characteristics, with the subject's indicator being compared to indicator references derived from individuals with similar characteristics. The reference population can include a plurality of individuals known to have a recurrence of cancer or those that are responsive to cancer therapy (including completely responsive and/or partially responsive), and in particular therapy using an immune checkpoint inhibitor; or a plurality of individuals known to be non-responsive to cancer therapy, and in particular therapy using an immune checkpoint inhibitor.

[0159] In specific embodiments, the indicator-determining methods of the present invention are performed using at least one electronic processing device, such as a suitably programmed computer system or the like. In this case, the electronic processing device typically obtains at least one measured biomarker value, either by receiving this from a measuring or other quantifying device, or by retrieving these from a database or the like. The processing device then determines the indicator by any suitable means.

[0160] The processing device can then generate a representation of the indicator, for example by generating a sign or alphanumeric indication of the indicator, a graphical indication of a comparison of the indicator to one or more indicator references or an alphanumeric indication of the likely responsiveness of the subject to have a recurrence or cancer and/or respond to a cancer therapy.

[0161] The indicator-determining methods of the present invention typically include obtaining a sample from a subject who has been diagnosed with cancer, and quantifying or otherwise assessing at least one of the biomarkers within the sample to determine biomarker values. This can be achieved using any suitable technique, and will depend on the nature of the biomarker, as described above. Suitably, an individual measured or biomarker value corresponds to the level, abundance or concentration of a cancer therapy biomarker or to a function that is applied to that level or amount. For example, if the indicator in some embodiments of the indicator-determining method of the present invention, which uses a plurality of cancer recurrence and/or cancer therapy biomarkers, is based on a ratio of concentrations of two polynucleotides or two polypeptides, this process would typically include quantifying the polynucleotides or polypeptides by any means known in the art, including quantitative RT-PCR or immunofluorescence, or by a functional assay.

[0162] In some embodiments, the likelihood of a subject responding to cancer is established by determining one or more cancer therapy biomarker values, wherein an individual cancer therapy biomarker value is indicative of a value measured or derived for a cancer therapy biomarker in a subject or in a sample obtained from the subject. These biomarkers are referred to herein as "sample cancer therapy biomarkers." In accordance with the present invention, a sample cancer therapy biomarker will correspond to a reference cancer therapy biomarker (also referred to herein as a "corresponding cancer therapy biomarker"). By "corresponding cancer therapy biomarker" is meant a cancer therapy biomarker that is structurally and/or functionally similar to a reference cancer therapy biomarker. Representative corresponding cancer therapy biomarkers include expression products of allelic variants (same locus), homologues (different locus), and orthologues (different organism) of reference cancer therapy biomarker genes. Nucleic acid variants of reference cancer therapy biomarker genes and encoded cancer therapy biomarker polypeptides can contain nucleotide substitutions, deletions, inversions and/or insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product). For nucleotide sequences, conservative variants include those sequences that, because of the degeneracy of the genetic code, encode the amino acid sequence of a reference cancer therapy polypeptide. The same approach applies for the cancer recurrence biomarkers.

[0163] Corresponding cancer biomarkers include amino acid sequences that display substantial sequence similarity or identity to the amino acid sequence of a reference cancer biomarker polypeptide. In general, an amino acid sequence that corresponds to a reference amino acid sequence will display at least about 80, 81, 82, 83, 84, 85, 86, 97, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% or even up to 100% sequence similarity or identity to a reference amino acid sequence. Corresponding biomarkers also include nucleic acid sequences that display substantial sequence similarity or identity to the nucleic acid sequence of a reference biomarker polynucleotide. In general, a nucleic acid sequence that corresponds to a reference nucleic acid sequence will display at least about 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 97, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% or even up to 100% sequence similarity or identity to a reference nucleic acid sequence.

[0164] In some embodiments, calculations of sequence similarity or sequence identity between sequences are performed as follows:

[0165] To determine the percentage identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In some embodiments, the length of a reference sequence aligned for comparison purposes is at least 30%, usually at least 40%, more usually at least 50%, 60%, and even more usually at least 70%, 80%, 90%, 100% of the length of the reference sequence. The amino acid

residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide at the corresponding position in the second sequence, then the molecules are identical at that position. For amino acid sequence comparison, when a position in the first sequence is occupied by the same or similar amino acid residue (i.e. , conservative substitution) at the corresponding position in the second sequence, then the molecules are similar at that position.

[0166] The percentage identity between the two sequences is a function of the number of identical amino acid residues shared by the sequences at individual positions, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. By contrast, the percentage similarity between the two sequences is a function of the number of identical and similar amino acid residues shared by the sequences at individual positions, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. For purposes herein, the sequence of a cDNA of a mRNA transcript, and the sequence of the mRNA itself, are deemed to have 100% sequence identity, although it is understood that one molecule is a DNA molecule and thus comprises “T” while the other is a RNA molecule and this comprises “U”.

[0167] The comparison of sequences and determination of percentage identity or percentage similarity between sequences can be accomplished using a mathematical algorithm. In certain embodiments, the percentage identity or similarity between amino acid sequences is determined using the Needleman and Wiinsch (1970, *J. Mol. Biol.* 48 : 444-453) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg .com), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In specific embodiments, the percent identity between nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.gcg.com), using a NWSgapdna.CM P matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. An non-limiting set of parameters (and the one that should be used unless otherwise specified) includes a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

[0168]

[0169] In some embodiments, the percentage identity or similarity between amino acid or nucleotide sequences can be determined using the algorithm of E. Meyers and W. Miller (1989, Cabios, 4: 11-17) which has been incorporated into the ALIGN program (version 2.0), using a PAM 120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

[0170] The nucleic acid and protein sequences described herein can be used as a “query sequence” to perform a search against public databases to, for example, identify other

family members or related sequences. Such searches can be performed using the NBLAST and X BLAST programs (version 2.0) of Altschul, et al., (1990, *J. Mol. Biol.*, 215 : 403- 10). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to 53010 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997, *Nucleic Acids Res.* 25: 3389-3402). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used .

[0171] Corresponding cancer therapy biomarker and cancer recurrence biomarker polynucleotides also include nucleic acid sequences that hybridize to reference cancer therapy biomarker polynucleotides, or to their complements, under stringency conditions described below. As used herein, the term "hybridizes under low stringency, medium stringency, high stringency, or very high stringency conditions" describes conditions for hybridization and washing. "Hybridization" is used herein to denote the pairing of complementary nucleotide sequences to produce a DNA-DNA hybrid or a DNA-RNA hybrid. Complementary base sequences are those sequences that are related by the base-pairing rules. In DNA, A pairs with T and C pairs with G. In RNA, U pairs with A and C pairs with G. In this regard, the terms "match" and "mismatch" as used herein refer to the hybridization potential of paired nucleotides in complementary nucleic acid strands. Matched nucleotides hybridize efficiently, such as the classical A-T and G-C base pair mentioned above. Mismatches are other combinations of nucleotides that do not hybridize efficiently.

[0172] Guidance for performing hybridization reactions can be found in Ausubel et at., (1998, supra), Sections 6.3.1-6.3.6. Aqueous and non-aqueous methods are described in that reference and either can be used. Reference herein to low stringency conditions include and encompass from at least about 1% v/v to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridization at 42°C, and at least about 1 M to at least about 2 M salt for washing at 42°C. Low stringency conditions also may include 1% Bovine Serum Albumin (BSA), 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization at 65°C, and (i) 2 x SSC, 0.1% SDS; or (ii) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 5% SDS for washing at room temperature. One embodiment of low stringency conditions includes hybridization in 6 M sodium chloride/sodium citrate (SSC) at about 45°C, followed by two washes in 0.2 x SSC, 0.1% SDS at least at 50°C (the temperature of the washes can be increased to 55°C for low stringency conditions). Medium stringency conditions include and encompass from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5 M to at least about 0.9 M salt for hybridization at 42°C, and at least about 0.1 M to at least about 0.2 M salt for washing at 55°C. Medium stringency conditions also may include 1% Bovine Serum Albumin (BSA), 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization

at 65°C, and (i) 2 x SSC, 0.1% SDS; or (ii) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 5% SDS for washing at 60-65°C. One embodiment of medium stringency conditions includes hybridizing in 6 x SSC at about 45°C, followed by one or more washes in 0.2 x SSC, 0.1% SDS at 60°C. High stringency conditions include and encompass from at least about 31% v/v to at least about 50% v/v formamide and from about 0.01 M to about 0.15 M salt for hybridization at 42°C, and about 0.01 M to about 0.02 M salt for washing at 55°C. High stringency conditions also may include 1% BSA, 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization at 65°C, and (i) 0.2 x SSC, 0.1% SDS; or (ii) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 1% SDS for washing at a temperature in excess of 65°C. One embodiment of high stringency conditions includes hybridizing in 6 x SSC at about 45°C, followed by one or more washes in 0.2 x SSC, 0.1% SDS at 65°C.

[0173] In certain embodiments, a corresponding cancer therapy biomarker or cancer recurrence biomarker polynucleotide is one that hybridizes to a disclosed nucleotide sequence under very high stringency conditions. One embodiment of very high stringency conditions includes hybridizing 0.5 M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2 x SSC, 1% SDS at 65°C.

[0174] Other stringency conditions are well known in the art and a skilled addressee will recognize that various factors can be manipulated to optimize the specificity of the hybridization. Optimization of the stringency of the final washes can serve to ensure a high degree of hybridization. For detailed examples, see Ausubel et al supra at pages 2.10.1 to 2.10. 16 and Sambrook et al. (1989, supra) at sections 1.101 to 1.104.

8. Compositions and Solid Supports

[0175] Also provided are compositions and solid supports for determining an indicator used in assessing a likelihood of a subject with cancer responding to cancer therapy or cancer recurrence in a subject. The compositions and solid support may be relevant to applications where the cancer therapy biomarker or cancer recurrence biomarker is polypeptide or a polynucleotide.

[0176] By way of an illustrative embodiment, the composition contains PIK3R2 transcript or cDNA thereof, and at least one oligonucleotide primer or probe that hybridizes to the PIK3R2 transcript or cDNA. As would be appreciated, the level of transcript or cDNA in the composition is reflective of the level of transcript in the subject from which the sample is taken and obtained, and can thus be used in the method of the present invention to assess the likelihood of a subject responding to cancer therapy or cancer recurrence.

[0177] Alternatively, the compositions may comprise a polypeptide expression product of p85β and a detection agent that binds to the polypeptide expression product of PIK3R2. In particular embodiments, the composition contains tumour cells that comprise the polypeptide expression product of PIK3R2. Typically, the detection agents are antibodies or

antigen-binding fragments thereof that are specific for the polypeptide expression product of PIK3R2.

[0178] Solid supports of the present invention include those to which at least one oligonucleotide primer or probe that hybridizes to a, for example, PIK3R2 transcript or cDNA thereof, are immobilized. In some embodiments, a PIK3R2 transcript or cDNA thereof is/are hybridized to their respective oligonucleotides or probes.

9. Kits

[0179] The present invention also extends to kits for determining expression of biomarkers, including the response to therapy biomarkers, the cancer recurrence biomarkers, and/or the treatment monitoring biomarkers, disclosed herein, which include reagents that allow detection and/or quantification of the biomarkers. Such reagents include, for example, compounds or materials, or sets of compounds or materials, which allow quantification of the biomarkers. In specific embodiments, the compounds, materials or sets of compounds or materials permit determining the expression level of a gene (e.g., a treatment monitoring biomarker gene), including without limitation the extraction of RNA material, the determination of the level of a corresponding RNA, etc., primers for the synthesis of a corresponding cDNA, primers for amplification of DNA, and/or probes capable of specifically hybridizing with the RNAs (or the corresponding cDNAs) encoded by the genes,

[0180] The kits may also optionally include appropriate reagents for detection of labels, positive and negative controls, washing solutions, blotting membranes, microtiter plates, dilution buffers and the like. For example, a nucleic acid-based detection kit may include (i) a treatment monitoring biomarker polynucleotide (which may be used as a positive control), (ii) a primer or probe that specifically hybridizes to a treatment monitoring biomarker polynucleotide. Also included may be enzymes suitable for amplifying nucleic acids including various polymerases (reverse transcriptase, Taq, Sequenase™, DNA ligase etc. depending on the nucleic acid amplification technique employed), deoxynucleotides and buffers to provide the necessary reaction mixture for amplification. Such kits also generally will comprise, in suitable means, distinct containers for each individual reagent and enzyme as well as for each primer or probe.

[0181] Alternatively, a protein-based detection kit may include (i) at least one response to therapy biomarker polypeptide, which is suitably selected from IL-6, EpCAM, ABCB5, and/or GBP2, or a fragment of IL-6, EpCAM, ABCB5, and/or GBP2, (ii) one or more antigen binding molecules that bind specifically to IL-6, EpCAM, ABCB5, and/or GBP2. The antigen-binding molecules are suitably detectably labeled. The kit can also feature various devices (e.g., one or more) and reagents (e.g., one or more) for performing one of the assays described herein; and/or printed instructional material for using the kit to quantify the expression of a biomarker gene. The reagents described herein, which may be optionally associated with detectable labels, can be presented in the format of a microfluidics card, a chip or chamber, a

microarray or a kit adapted for use with the assays described in the examples or below, e.g., RT-PCR or qPCR techniques described herein.

[0182] Materials suitable for packing the components of the diagnostic kits may include crystal, plastic (polyethylene, polypropylene, polycarbonate and the like), bottles, vials, paper, envelopes and the like. Additionally, the kits of the invention can contain instructional material for the simultaneous, sequential or separate use of the different components contained in the kit. The instructional material can be in the form of printed material or in the form of an electronic support capable of storing instructions such that they can be read by a subject, such as electronic storage media (magnetic disks, tapes and the like), optical media (CD-ROM, DVD) and the like. Alternatively or in addition, the media can contain internet addresses that provide the instructional material.

10. Patient classification and treatment management

[0183] The present invention extends to methods of selecting or identifying individuals who are appropriate candidates for treatment with a therapy (e.g., a cytotoxic therapy, an immunotherapy, etc.) for treatment of cancer. Such individuals include patients that are predicted to be responsive to the therapy and thus have an increased likelihood of benefiting from administration of the therapy relative to other patients having different characteristic(s) (e.g., non-responsiveness to the therapy). In certain embodiments an appropriate candidate is one who is reasonably likely to benefit from treatment or at least sufficiently likely to benefit as to justify administering the treatment in view of its risks and side effects. The invention also encompasses methods of selecting or identifying individuals who are not appropriate candidates for treatment with a therapy (e.g., a cytotoxic therapy, an immunotherapy, etc.) for treatment of cancer. Such individuals include patients that are predicted to be non-responsive or weakly responsive to the therapy and thus have a decreased likelihood of benefiting from administration of the therapy relative to other patients having different characteristic(s) (e.g., responsiveness to the therapy), or a low or substantially no likelihood of benefiting from such treatment, such that it may be desirable to use a different or additional treatment. In some embodiments, whether a subject is an appropriate candidate for therapy with a therapy is determined based on an assay of at least one response to therapy biomarker.

[0184] In some aspects described herein are methods of determining, for example based on an assay of at least one response to therapy biomarker, the likelihood that a subject in need of treatment for cancer will respond to treatment with a therapy (e.g., a cytotoxic therapy, an immunotherapy, etc.) and/or of identifying and/or selecting a subject to receive such treatment. In specific embodiments, the therapy is a PI3K inhibitor (e.g., GDC-0084 inhibitor, paxalisib), suitably in combination with an anti-immune checkpoint inhibitor. The phrase “treatment with an immune checkpoint inhibitor”, also referred to as “immune checkpoint inhibitor treatment”, “therapy with an immune checkpoint inhibitor”, or “immune checkpoint

inhibitor therapy", encompasses embodiments pertaining to treatment with a single immune checkpoint inhibitor and embodiments pertaining to treatment with two or more immune checkpoint inhibitors in combination. In some embodiments immune checkpoint inhibitor treatment comprises inhibiting two or more different immune checkpoint pathways using a single agent or using two or more separate agents.

[0185] In order that the invention may be readily understood and put into practical effect, particular preferred embodiments will now be described by way of the following non-limiting experimental examples.

EXPERIMENTAL

PI3K-mTOR induces an anti-tumour immune profile.

[0186] To investigate the differential expression of immune cell populations following treatment with the PI3K-mTOR inhibitor, GDC-0084 ("paxilasib"), the present inventors employed spatial immune profiling using the Phenocycler Spatial Tissue Exploration Program (STEP) platform (Akoya Biosciences). In particular, analysis was carried out using the 4T1 triple negative breast cancer (TNBC) model. FFPE tumour tissues from mice receiving either a αPD1 treatment or a combination of GDC-0084 and αPD1 treatment were subjected to an 18-marker STEP core mouse panel (Figure 1A). Unsupervised clustering phenotyping analysis revealed an increase in anti-tumour immune cell phenotypes including B cells (28% vs 2%), T cells (15% vs 10%), CD24+ and CD11c+ dendritic cells (6% vs 2%, 12% vs 10%) and CD38+ myeloid cells (4% vs 1%) following GDC-0084 and αPD1 combination treatment (Figure 1B).

[0187] Notably, the GDC-0084 and αPD1 combination treatment resulted in reduced pro-tumour phenotypes, such as a reduced proportion of neutrophils and bone marrow blasts. Differential distribution of these cell populations were also observed between treatments, which may be a key indicator of disease progression and treatment response (Figure 1C).

[0188] In order to better understand the crosstalk occurring between immune cell populations, the present inventors conducted cellular neighbourhood and spatial proximity analysis. The cellular neighbourhood (CN) investigation revealed ten unique cellular neighbourhoods, including CN1 enriched in erythrocytes; CN6, CN7, CN9, and CN10 with abundance of B cells, CD11c+ dendritic cells, T cells, and CD11b+ neutrophils (Figure 2A). GDC-0084 and αPD1 combination treatment induced cellular neighbourhoods enriched in anti-tumour immune cell phenotypes such as T cells (CN9; 12% vs 5%), B cells (CN6; 22% vs 0.2%) and CD11c+ dendritic cells (CN7; 10% vs 6%) (Figure 2B). Furthermore, cellular neighbourhoods enriched in pro-immune cell populations such as neutrophils (CN4; 4% vs 8%) and bone marrow blasts (CN3; 5% vs 11%) were all reduced following GDC-0084+αPD1 combination treatment (Figure 2B). Differential distribution of these cellular neighbourhoods were also observed between mice receiving the αPD1 treatment and those administered with the GDC-0084 and αPD1 combination treatment (Figure 2C).

[0189] Spatial proximity analysis performed on tumours treated with GDC-0084 and αPD1 combination therapy revealed a significant increase in the number of anti-tumour immune cells (B cells, T cells, and dendritic cells) surrounding bone marrow blasts, vascular and endothelial cells (Figure 3A). Additionally, the distance between these cell populations was reduced (Figure 3B). These data suggest an increased crosstalk between anti-tumour immune cells within tumours exposed to the GDC-0084 and αPD1 combination treatment.

PI3K-mTOR inhibition diminishes mast cell and Treg infiltration into the tumour

[0190] Next, the inventors conducted NanoString nCounter analysis to investigate differential cellular composition and gene expression in RNA prepared from GDC-0084 treated tumours. Through rigorous cell abundance analysis, it was observed that both GDC-0084 and combination therapies exert inhibitory effects on pro-tumor immune cell populations, notably mast cells, neutrophils and Treg cells, in comparison to control and anti-PD-1 treated mice (Figure 4A). Importantly, a reduction of mast cell populations was observed using Toluidine staining on αPD1 therapy and GDC-0084 and αPD1 combination therapy tumour tissues (Figure 4B).

[0191] NanoString nCounter analysis of relative cell abundance also revealed heightened total tumour infiltrating lymphocyte (TIL) counts in both GDC-0084 monotherapy and combination therapy. This elevation was characterized by enrichment in subsets of immune cells with potent anti-tumour properties, including dendritic cells (DCs), natural killer (NK) cells, T cells and B cells (Figure 5A). Importantly, elevated TIL levels have more favourable therapeutic outcomes in the context of metastatic breast cancer. GDC-0084 treatment also enhanced cytotoxic cell numbers and associated IFN-γ and Granzyme B mRNA expression (Figure 5B), confirming re-invigoration of exhausted T cell populations. Moreover, the presence of exhausted T cells was more pronounced in both the control and anti-PD-1 treated groups, indicative of their potential involvement in immune suppression (Figure 5C).

Impact of PI3K-mTOR inhibition on tumour-associated genes and signalling pathways.

[0192] To elucidate the impact of PI3K-mTOR inhibition on gene expression programs, the inventors meticulously analyzed more than 40 signalling pathways, spanning tumour biology, immune response, evasion, and tumor microenvironment remodelling (Figure 6). GDC-0084 monotherapy and combination treatment with αPD1 significantly induced pathways related to anti-tumour immune function, and in particular the interferon response, interleukin signalling, T cell co-stimulation, cytotoxicity, and antigen presentation (Figure 6). The combination treatment led to a notable reduction in extracellular matrix remodeling and metastasis pathways (Figure 6).

Generation of a gene signature in mice.

[0193] Further examination of the PI3K-mTOR regulated Nanostring nCounter analysis led to the discovery of a tumour microenvironment signature (Table 1). This panel of 16 genes, which were all up-regulated following GDC-0084 treatment, includes the tumour

suppressor genes *CDH1*, *DCN*, *LCN2*, *LOX*, and *SOCS1*; the anti-tumour genes *CXCR6*, *NKG7*, and *NOS2*; the immune checkpoint-associated genes *CD274* (PD-L1) and *PDCD1LG2* (PD-L2) and the cytotoxicity-related genes *GZMB*, *IFNG,* and *HMOX1* (Figure 7; and Table 1). Additionally, the pro-tumour genes *FAM83A*, *PDGFRB*, and *FOXP3* were all down-regulated in response to PI3K-mTOR inhibition (Figure 7; Table 1).

TABLE 1

Tumour Microenvironment (TME) Signature	**Cancer treatment responsiveness signature**	
GZMB	*EIF4EBP1*	*KIT*
CDH1	*VEGFA*	*FLT4*
CXCR6	*MTHFD2*	*TPSAB1*
NKG7	*NFAM1*	*CPA3*
DCN	*LCN2*	*FOS*
LCN2	*CD274*	*CDC25C*
CD274	*CEACAM1*	*CARD11*
LOX	*CDKN1A*	*PCLAF*
NOS2	*HMOX1*	*DNA2*
HMOX1	*SERPINE1*	*CCNE1*
SOCS1	*IL15RA*	*SQLE*
IFNG	*PDCD1LG2*	*CDC20*
PDCD1LG2	*CCR5*	*PLK1*
FAM83A	*CD53*	*MKI67*
PDGFRB	*MMP9*	*KIF2C*
FOXP3	*ACKR3*	*JUP*
	LOX	*FANCA*
	DCN	*BIRC5*
	FSTL1	*LTBR*

Validation of the cancer responsiveness gene signature.

[0194] To validate the relevance of the PI3K-mTOR-regulated gene signature in human breast cancer patients, the inventors conducted whole transcriptome GeoMx analysis on longitudinal human TNBC patient samples collected pre-chemotherapy, post-chemotherapy, and at the time of recurrence. The novel technique of overlaying human data with the mouse Nanostring nCounter analysis was used to generate the cancer treatment responsiveness signature (Table 1). This panel of 38 differently expressed genes included the tumour suppressor genes *DCN*, *LCN2*, and *LOX*; the cytotoxicity-related gene *HMOX1*, the immune checkpoint-associated genes *PDCD1LG2* and *CD274*, and the cytokine/chemokine receptors *CCR5* and *IL15RA* (Table 1).

CTC biomarker predicts patient response to therapy.

[0195] These spatial immune profiling and Nanostring nCounter data confirm the differential regulation of immune populations following PI3K-mTOR inhibition. Hence, the inventors next sought to examine the expression of critical mediators of the immune response. This included a key driver of inflammation, *IL6*, and the viral mimicry gene, *GBP2*. Pre-treatment of MCF-7 cells (breast cancer cell line) with PI3K-mTOR inhibitors (GDC-0084 and Dactolisib)

prior to PMA/TGFβ-induced epithelial-to-mesenchymal transition induction, resulted in a significant reduction in *IL6* mRNA expression and an up-regulation of *GBP2* (Figure 8). Similar results were also confirmed using the highly aggressive breast cancer cell line, MDA-MB-231 (Figure 9).

[0196] Based on this information, the inventors established a blood biomarker incorporating the immune cell mediator, IL-6. Circulating tumour cells (CTCs) were prepared from melanoma patient blood samples collected pre- and post-treatment. Immunofluorescence staining was performed using antibodies directed against IL-6 and the CTC markers, EpCAM and ABCB5. As anticipated, elevated levels of IL-6/EpCAM/ABCB5+ cells were observed pre-treatment when inflammation is prominent (Figure 10). However, following treatment, the proportion of CTCs expressing *IL6* dramatically reduced confirming the utility of IL-6/EpCAM/ABCB5+ CTCs as a biomarker for patient response to treatment.

PBMC biomarker predicts cancer recurrence

[0197] PI3Ks are composed of p110 catalytic subunits and p85 regulatory subunits. Recent studies have demonstrated that p85β promotes tumorigenesis by inducing nuclear translocation (see, Hao et al., 2022). Nuclear p85β recruits deubiquitinase USP7 to stabilize the histone methyltransferases, EZH1 and EZH2, and enhances histone H3 lysine 27 trimethylation (H3K27Me3). To investigate the direct interaction of p85β and H3K27Me3, the inventors employed the DUOLINK® proximity ligation assay. PBMCs were prepared from TNBC patient blood samples (pre- and post-chemotherapy) and subjected to the DUOLINK® proximity ligation assay (Figure 11A).

[0198] Elevated p85β:H3K27Me3 interactions were observed post-chemotherapy both within the nucleus and the cytoplasm in patients in which the cancer did not recur (see, CH002, CH007, Figure 11B). Importantly, patients which experienced cancer recurrence maintained low levels of p85β:H3K27Me3 interactions post-chemotherapy, notably in the cytoplasm. These data show the clear clinical potential for the use of this novel biomarker as a key indicator of cancer recurrence in a subject.

Associations between PIK3R2 gene expression and patient survival.

[0199] The inventors analysed whether expression of *PIK3R2*, which encodes p85β, was associated with outcomes in TNBC/basal-like breast cancer (GSE158309 dataset, as described in Heimes et al., 2020). Patients with basal-like/TNBC expressing high levels of *PIK3R2* (solid lines) showed poorer distant metastasis-free survival (DMFS) than patients with low expression (Figure 12A).

[0200] Next, the inventors assessed *PIK3R2* mRNA expression in stage IV metastatic TNBC post chemotherapy (carboplatin, nab-paclitaxel) + pembrolizumab durvalumab + olaparib in a single-arm phase-2 trial (Wilkerson et al., 2024) (Figure 12B). Patients with higher mRNA expression of *PIK3R2* showed poorer overall survival compared with patients with lower *PIK3R2* mRNA expression. In addition, higher pre-treatment levels of *PIK3R2* were

positively associated with distant relapse in TNBC patients who had residual tumors after treatment with neoadjuvant chemotherapy (NACT) (Blye et al., 2022) (Figure 12C). On the basis of these data, the strength of *PIK3R2* as a useful prognostic and predictive biomarker is clearly evidenced.

Materials & Methods

[0201] All materials and reagents used in the synthesis and testing of the compositions described are commercially available, for example, from Sigma-Aldrich Co., Novabiochem, Abeam, and American Type Culture Collection (ATCC) unless otherwise stated.

Animal studies

[0202] Female BALB/c mice, 6-8 weeks of age, were procured from the Animal Resources Centre (ARC). Following acquisition, the mice were allowed a one-week acclimatization period within the containment suites at QIMR Berghofer Medical Research Institute (QIMRB). All experimental procedures were performed in accordance with the guidelines and regulations endorsed by the QIMRB Ethics Committee. 1 x 10^5 4T1 cells, suspended in phosphate-buffered saline (PBS) were administered into the mammary fat pad of the BALB/c mice. Treatments were started when tumours reached approximately 50-100 mm^3. In the 4T1 model, mice were subjected to anti-PD-1 treatment (10 mg/kg) administered intraperitoneally every 5 days, and daily oral administration of GDC-0084 (7.5 mg/kg). Tumours were harvested prior to reaching ethical limits and fixed in 3.7% paraformaldehyde solution (FFPE) or frozen in OCT for downstream analysis.

NanoString nCounter assay

[0203] Total RNA was extracted from FFPE tumor tissue using the RNeasy FFPE kit according to the manufacturer's protocols. The concentration of RNA was measured using the Qubit RNA HS assay kit.

[0204] The RNA samples were subjected to hybridize with a multiplexed mouse tumour signaling 360 panel codeset. The hybridized samples were subsequently loaded onto the nCounter prep station chip, and data acquisition was accomplished using the nCounter Digital Analyzer, following manufacturer-prescribed procedures. Data analysis was undertaken via the utilization of the nSolver Analysis Software version 4.0, supplied by NanoString. Counts were subjected to normalization in accordance with the expression levels of housekeeping genes and positive/negative-control probes.

STEP Spatial immune profiling and analysis

Staining

[0205] OCT frozen mouse tumour tissues were stained with an 18-panel antibody cocktail (CD19, CD21/35, CD3, TCRb, CD4, CD90.2, CD11b, CD45, Ly6G, IgM, CD24, CD11c, CD38, CD31, Ter119, CD49f, CD71, Ki67) and subjected to the Spatial Tissue Exploration Program (STEP) platform (Akoya Biosciences).

Unsupervised clustering phenotyping and cellular neighborhood analysis

[0206] Data Analysis was performed using Akoya's internal software, Multiplexed Image Analysis (MIA), which combines multiple existing analysis tools with a Graphical User Interface. Cell segmentation was first performed using StarDist[1] method, where the deep learning model was retrained by Akoya team on DeepCell TissueNet database[2] which contains more than one million cells annotated for segmentation purposes from different tissues and different platforms. Cytoplasm segmentation was then estimated from nuclear expansion by morphological dilation. Marker mean fluorescent intensity (MFI) was calculated for each segmented cell and z-score normalization was applied across all cells for each marker such that each marker has a mean equals to zero and standard deviation equals to one. Only QC-passed lineage markers were used for downstream analysis. In this study, all markers are passed QC after manually reviewing based on appropriate cellular localization and pattern of staining. Cells without any marker expression or abnormally high signal sum were removed from the downstream analysis. After cell segmentation and filtering, there were total 528,659 cells across two samples.

[0207] The normalized MFI matrices for all markers (columns) and all cells (rows) from each QPTiff image were the input for unsupervised clustering. Leiden[3, 4] algorithm and GPU-accelerated[5, 6] method were used for clustering. For Leiden clusters, resolution equals to 1-6 were tested and the optimal resolution = 2 was chosen for manual annotations in order to assign cell phenotypes based on marker expression pattern in hierarchical clustering heatmap and their location on the images. Clusters with similar expression profiles were combined. After phenotyping, the percentage of each phenotype *i* within each sample *j* was calculated and plotted as a pie chart.

$$percentage_{i,j} = \frac{\#\ of\ cells\ for\ phenotype\ i\ in\ sample\ j}{total\ \#\ of\ cells\ in\ sample\ j}.$$

[0208] For every annotated cell in a dataset, we use a k-nearest-neighbors approach to enumerate the cell types of its ten nearest spatial neighbors in Euclidean space, this cell's local neighborhood was called a "window", then use a k-means-clustering approach to group these windows into ten cell neighborhoods (k) based on its mean composition of cell types. Each cell is then allocated to the same cell neighborhood as its surrounding window.[7] After cellular neighborhoods analysis, the percentage of each CN i within each sample j was calculated as a pie chart.

$$percentage_{i,j} = \frac{\#\ of\ cells\ for\ cellular\ neighborhood\ i\ in\ sample\ j}{total\ \#\ of\ cells\ in\ sample\ j}.$$

[0209] The distances between each cell to its nearest specific cell phenotypes (excluding others and unidentified) was calculated using CytoMAP[8]. The number of specific cell phenotypes (excluding other and unidentified) within 50 µm of every cell was also calculated. After calculating, the mean distances between pair-wise cell types and mean number of

"LocalDensityOf" cells within 50 µm radius of a single specific cell type within each sample were calculated as bar plots.

$$Mean\ Number/distance_{i,k,j} = \frac{sum\ \#\ of\ LocalDensityOf/Dist\ to\ cell\ type\ i}{total\ \#\ of\ cell\ type\ k\ in\ sample\ j}.$$

Cell culture

[0210] All breast cancer cell lines used were sourced from ATCC. MDA-MB-231 cell lines were maintained and cultured in DMEM (Invitrogen) supplemented with 10% FBS, 2 mM L-glutamine, and 1% PSN. MCF-7 cells were stimulated with 1.29 ng/mL phorbol 12-myristate 13-acetate (PMA) (Sigma-Aldrich) or 5 ng/mL recombinant TGF-β1 (R&D Systems) for 24 hours. For inhibitor studies, 2 x 10^5 cells were seeded in 2 mL complete media in 6 well plates and incubated overnight at 37°C/5% CO_2. Cells were treated with PI3K-mTOR inhibitors including GDC--0084 (10 µM), Dactolisib (10 µM), Idealisib (90 µM), LY294002 (50 µM) for 24 hrs and then removed prior to stimulation with PMA/TGF-β1 (MCF-7 only).

RNA extraction and cDNA synthesis

[0211] Cells were lysed in 500 µl of TRIzol Reagent (Invitrogen) before RNA isolation using the Direct-zol RNA Miniprep Kit (Zymo, R2052) and treated with RNase-free DNase I (Qiagen, 79254) following the manufacturer's protocol. The Nanodrop (Thermofisher) was then used to determine the RNA concentration and purity (A260/A280). RNA (1 µg) was then reverse-transcribed to cDNA using the Superscript VILO IV Master Mix (Thermo Fisher, 11756050) following the supplier's protocol. cDNA was diluted 1:20 with RNase-free water for RT-qPCR.

RT-qPCR for IL-6 and GBP2

[0212] Gene transcription was analysed using Taqman Gene Expression Master Mix (Life Technologies, 4369016) on the ABI Viia 7 real-time PCR machine (Applied Biosystems). Cts were converted to arbitrary copy numbers and normalized to the geomean of the house-keeping genes ACTB and PPIA. The following human Taqman probes were used: Human: IL-6, GBP2. No template and no reverse-transcriptase controls were included to exclude genomic DNA contamination.

CTC enrichment

[0213] Whole blood from melanoma cancer patients was stored in EDTA tubes for circulating tumour cell identification. The RosetteSep™ Human CD45 Depletion Cocktail was used to enrich tumour cells (CTCs) from whole blood by depleting CD45+ cells. Unwanted cells were targeted for depletion with Tetrameric Antibody Complexes recognizing CD45, CD66b and glycophorin A on red blood cells (RBCs). Unwanted cells were then removed via centrifugation over a buoyant density medium LYMPHOPREP™ (Catalogue #07801). The purified epithelial tumour cells were then extracted as a highly enriched population from the interface between the plasma and the buoyant density medium and were harvested in 20% FBS in PBS.

IL-6/EpCAM/ABCB5 biomarker imaging

[0214] CTCs were permeabilised by incubating with 0.5% Triton X-100 for 15 min, blocked with 1% BSA in PBS and were probed with either IL-6, EpCAM or ABCB5 and visualized with a donkey anti-rabbit AF 488, anti-mouse AF 568, donkey anti-goat 647. Cover slips were mounted on glass microscope slides with ProLong Glass Antifade reagent (Life Technologies). Protein targets were localised by digital pathology laser scanning microscopy. Single 0.5 µm sections were obtained using the ASI Digital pathology (ASI) platform. Digital images were analysed using automated ASI software as described previously (Applied Spectral Imaging, Carlsbad, CA).

Healthy donor PBMC isolation from stage III TNBC patients

[0215] Whole blood was stored in EDTA tubes for circulating tumour cell identification. Unwanted cells were targeted for using glycophorin A on red blood cells (RBCs) and were then removed via centrifugation over a buoyant density medium LYMPHOPREP™ (Catalogue #07801). The purified PBMCs were then extracted as a highly enriched population from the interface between the plasma and the buoyant density medium and were harvested in 20% FBS in PBS.

DUOLINK proximity assay

[0216] PBMCs were fixed in 3.7% paraformaldehyde and cytospinning was used to mount onto glass microscope slides. The DUOLINK proximity ligation assay (MilliporeSigma) was performed using antibodies directed against P85β and H3K27Me3 and digital images were analyzed using ImageJ software (ImageJ, NIH, Bethesda, MD, USA).

Population-based cancer survival analysis

[0217] Human cancer datasets meeting the following criteria were downloaded from Array Express (AE), Gene Expression Omnibus (GEO), and Sequence Read Archive (SRA): (a) at least 20 expression profiles from cancer patient tissues (not cell lines or preclinical models); (b) annotated with clinical outcomes of at least one of (i) categorical outcomes after treatment, e.g., progressive disease, stable disease, pathological complete response, response, no response, and distant metastasis; and/or (ii) continuous outcome data including survival and event follow-up, e.g., overall, disease-specific, distant metastasis-free, or progression-free survival.

[0218] For cohort analysis, within each cohort, the mRNA expression levels of *PIK3R2* were compared by categorizing the samples into high- and low-risk subgroups, e.g., high risk = distant metastasis and low risk = no metastasis. Samples were also categorized into high and low subgroups based on *PIK3R2* expression, where high = ≥75th percentile of *PIK3R2* expression and low = <75th percentile. Where mutation or copy-number data were available, samples were categorized into altered and unaltered for the *PIK3R2* gene: altered = harbors a mutation or copy number change in *PIK3R2*; unaltered = otherwise. Box and scatter plots comparing *PIK3R2* expression between the above subgroups were plotted using the ggplot

package in R version>4.0 (21). Statistical significance between subgroups was compared with the Wilcoxon test. Kaplan-Meier survival curves between high- and low-risk groups were plotted using the ggsurve package and the statistical difference between the curves were computed using log-rank test.

REFERENCES

Hao, Y., He, B. ,Wu, L., Li, Y., Wang, C., Wang, T., Sun, L., Zhang, Y., Zhan, Y., Zhao, Y., Markowitz, S., Veigl, M., Conlon, R. A., Wang, Z., Nuclear translocation of p85β promotes tumorigenesis of *PIK3CA* helical domain mutant cancer. *Nature Comm.*, 2002, 13:1974.

Hassan B, Akcakanat A, Holder AM, Meric-Bernstam F. Targeting the PI3-kinase/Akt/mTOR signaling pathway. Surg Oncol Clin N Am. 2013 Oct;22(4):641-64.

Jiang, N., Dai, Q., Su, X. et al. Role of PI3K/AKT pathway in cancer: the framework of malignant behavior. Mol Biol Rep 47, 4587–4629 (2020).

Chen X, Cao Y, Sedhom W, Lu L, Liu Y, Wang H, Oka M, Bornstein S, Said S, Song J, Lu SL. Distinct roles of PI3KCA in the enrichment and maintenance of cancer stem cells in head and neck squamous cell carcinoma. Mol Oncol. 2020 Jan;14(1):139-158.

Xia P, Xu XY. PI3K/Akt/mTOR signaling pathway in cancer stem cells: from basic research to clinical application. Am J Cancer Res. 2015 Apr 15;5(5):1602-9.

Yang L, Huang F, Mei J, Wang X, Zhang Q, Wang H, Xi M, You Z. Posttranscriptional Control of PD-L1 Expression by 17β-Estradiol via PI3K/Akt Signaling Pathway in ERα-Positive Cancer Cell Lines. Int J Gynecol Cancer. 2017 Feb;27(2):196-205.

Fusco N, Malapelle U, Fassan M, Marchiò C, Buglioni S, Zupo S, Criscitiello C, Vigneri P, Dei Tos AP, Maiorano E, Viale G. PI3KCA Mutations as a Molecular Target for Hormone Receptor-Positive, HER2-Negative Metastatic Breast Cancer. Front Oncol. 2021 Mar 25;11:644737.

Mosele F, Stefanovska B, Lusque A, et al. Outcome and molecular landscape of patients with PI3KCA-mutated metastatic breast cancer. Annals of Oncology; Published online 24 January 2020.

Lanczky A, Gyorffy B: Web-Based Survival Analysis Tool Tailored for Medical Research (KMplot): Development and Implementation, J Med Internet Res, 2021 Jul 26;23(7):e27633.

Heimes AS, Hartner F, Almstedt K, Krajnak S, Lebrecht A, Battista MJ, *et al.* Prognostic Significance of Interferon-gamma and Its Signaling Pathway in Early Breast Cancer Depends on the Molecular Subtypes. *Int J Mol Sci* 2020; **21**(19).

Wilkerson AD, Parthasarathy PB, Stabellini N, Mitchell C, Pavicic PG, Jr., Fu P, *et al.* Phase II Clinical Trial of Pembrolizumab and Chemotherapy Reveals Distinct Transcriptomic Profiles by Radiologic Response in Metastatic Triple-Negative Breast Cancer. *Clin Cancer Res* 2024; **30**(1): 82-93.

Blaye C, Darbo E, Debled M, Brouste V, Velasco V, Pinard C, *et al.* An immunological signature to predict outcome in patients with triple-negative breast cancer with residual disease after neoadjuvant chemotherapy. *ESMO Open* 2022; **7**(4): 100502.

WHAT IS CLAIMED IS:

1. A method of monitoring the responsiveness of a cancer to an anti-cancer treatment in a subject, the method comprising the step of determining an expression level of one or a plurality of biomarkers in a biological sample obtained from the subject, wherein the biomarkers comprise one or more of *DCN*, *LCN2, LOX, HMOX1, PDCD1LG2,* and *PIK3R2*; wherein an altered or modulated expression level of the one or plurality of markers indicates or correlates with relatively increased or decreased responsiveness of the cancer to the anti-cancer treatment.

2. A method of determining the likelihood of cancer recurrence in a subject, the method comprising the step of determining an expression level of a *PIK3R2* gene in the subject, and an increased expression level of the *PIK3R2* gene indicates or correlates with a relatively increased likelihood of cancer recurrence.

3. A method of determining the likelihood of cancer recurrence in a subject, the method comprising the step of determining a level of a cancer recurrence biomarker in a biological sample obtained from the subject, wherein the biomarker comprises the level of interaction of p85β and H3K27Me3 in a cellular compartment, and an altered or modulated level of the biomarker indicates or correlates with a relatively increased likelihood or relatively decreased likelihood of cancer recurrence.

4. The method of claim 3, wherein a reduced level of interaction between p85β and H3K27Me3 as compared to a control or reference sample indicates a relatively increased likelihood of cancer recurrence.

5. The method of claim 3 or claim 4, wherein the cellular compartment comprises the cell nucleus and the cell cytoplasm, and preferably the cell cytoplasm.

6. The method of any one of claims 3 to 5, wherein the biological sample is a whole blood sample, or a tumour sample.

7. A method of determining the likelihood of a cancer responding to an anti-cancer treatment, the method comprising the step of determining presence or level of one or a plurality of response to therapy biomarkers in a biological sample obtained from the subject, wherein the response to therapy biomarkers are selected from IL-6, EpCAM, ABCB5, and GBP2, and an altered or modulated level of the one or a plurality of biomarkers indicates or correlates with relatively increased or decreased likelihood of the cancer responding to the anti-cancer treatment.

8. The method of claim 7, wherein the biological sample comprises circulating tumour cells.

9. A method of determining the likelihood of a cancer responding to an anti-cancer treatment, the method comprising the step of determining an expression level of one or a plurality of biomarkers in a biological sample obtained from the subject, wherein the biomarkers

comprise one or both of IL6 and GBP2 and an altered or modulated expression level of the one or both markers indicates or correlates with a relatively increased or decreased likelihood of the cancer responding to the anti-cancer treatment.

10. The method of any one of claims 7 to 9, wherein an increased level of the response to therapy biomarkers indicates or correlates with an increased likelihood of the subject responding the anti-cancer treatment.

11. The method of any one of claims 7 to 9, wherein the anti-cancer treatment is an anti-PD1 immunotherapy, or a PARP inhibitor therapy.