

159159

Diagnostic Biomolecules



G00002

Field of the Invention

The present invention relates to biological molecules or biomolecules. In particular, to biomarkers. More in particular, the present invention relates to biomolecules used for diagnosis or prognostic evaluation of a disease or to monitor efficacy of a disease treatment.

Background of the Invention

Breast cancer is the second leading cause of cancer deaths in women today (after lung cancer) and is one of the most common cancers among women. The World Health Organization estimates that more than 1.2 million people worldwide each year will be diagnosed with breast cancer.

The American Cancer Society estimates that in 2005, approximately 210,000 women in the United States will be diagnosed with invasive breast cancer (Stages I-IV). The chance of developing invasive breast cancer during a woman's lifetime is approximately 1 in 7 (13.4%). Another 60,000 women will be diagnosed with *in situ* breast cancer, a very early form of the disease. Men can also suffer from breast cancer, though at a smaller probability. An estimated 1,700 cases will be diagnosed in men in 2005.

It is estimated that each year, 40,000 women and 460 men will die from breast cancer in the United States. According to the American Cancer Society, the chance that breast cancer will be responsible for a woman's death is about 1 in 33 (3%). The incidence rate of breast cancer (number of new breast cancers per 100,000 women) increased by approximately 4% during the 1980s but this has plateaued to 100.6 cases per 100,000 women in the 1990s. The death rates from breast cancer also declined significantly between 1992 and 1996, with the largest decreases among younger women. Medical experts attribute the decline in breast cancer deaths to earlier detection and more effective treatments.

Breast cancer is usually first detected as a palpable lump in the breast and/or by unusual presentation of the breast such as dimpling, puckering, bulging of the skin, a nipple that has changed position or an inverted nipple, redness, soreness, rash, swelling, or asymmetry of the breasts.

Detection by physical examination may be followed by mammography by x-ray or ultrasound. However, mammography is fraught with false positive results due to cysts, calcifications and benign lumps such as fibroadenomas. Consequently, further diagnoses based on cytology on tissue samples obtained from biopsies are usually necessary. Interpretation of cytological samples requires highly trained professionals.

As such, improved means to detect and treat breast cancer will be welcome.

Summary of the Invention

The present invention addresses the problems above and provides methods for detecting and/or treating a proliferative cell disorder such as cancer. In general, the disorder is cancer, particularly breast cancer.

In one aspect, the present invention provides a method of detecting and/or quantitating the presence of, predisposition to, and/or severity of, a proliferative cell disorder in a subject, the method comprising:

(a) providing at least one sample from a subject;

(b) determining the expression of a PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof; and

(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression

indicating the presence of, or predisposition to, and/or severity of a proliferative cell disorder in the subject.

There is also provided a method of monitoring the efficacy of a treatment for a proliferative cell disorder in a subject, the method comprising:

(a) providing at least two samples from a subject, each sample obtained at different time points;

(b) determining the expression of a PA2G4 gene, RNA, and/or mutation thereof, or PA2G4 protein, derivative, mutant and/or fragment thereof; and

(c) comparing the expression of the PA2G4 gene, RNA, and/or mutation thereof, and/or PA2G4 protein, derivative, mutant and/or fragment thereof, in the at least two samples.

There is also provided a method of prognosticating the outcome of a cell proliferative cell disorder in a subject, the method comprising:

(a) providing at least one sample from a subject;

(b) determining the expression of a PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof; and

(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression indicating the prognosis of a proliferative cell disorder in the subject.

The control may be, for example, is at least one subject not diagnosed with a cell proliferative disorder. Alternatively, the control may be at least one subject having a low level of expression of PA2G4 gene and/or protein. The control may also be an average value in expression obtained from a selected population.

The gene, RNA and/or mutation thereof may be human PA2G4 gene, RNA and/or mutation thereof, and the PA2G4 protein, derivative, mutant and/or fragment may be human PA2G4 protein, derivative, mutant and/or fragment thereof.

The detecting may be by labelling the PA2G4 protein, derivative, mutant and/or fragment thereof, with at least one sulfenyl halide agent followed by 2-dimensional gel electrophoresis and mass spectrometry. The sulfenyl halide agent may be 2-nitrobenzenesulfenyl chloride (NBS).

There is also provided a method of treating a proliferative cell disorder in a subject, the method comprising varying the expression of a PA2G4 gene, RNA, and/or mutation thereof.

The PA2G4 gene may have a sequence as given in SEQ ID NO:1. The varying may be obtained by administering nucleic acid hybridisable to all or part of the PA2G4 gene, RNA and/or mutation thereof. The hybridisable nucleic acid may be DNA or RNA. In particular, the hybridisable nucleic acid may be siRNA. The varying may be obtained by administering a compound hybridisable to all or part of the PA2G4 gene, RNA and/or mutation thereof.

There is also provided a method of treating a proliferative cell disorder in a subject, the method comprising varying the expression of a PA2G4 protein, derivative, mutant and/or fragment thereof.

The method may comprise varying the expression of the PA2G4 protein having the amino acid sequence of SEQ ID NO:2. The varying may be obtained by administering a polypeptide binding to all or part of the amino acid sequence of the PA2G4 protein, derivative, mutant and/or fragment thereof. The binding polypeptide may be an antibody. The varying may be obtained by administering a compound hybridisable to all or part of the amino acid sequence of the PA2G4 protein.

The varying of the expression may comprise reducing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof.

There is also provided a method of screening for at least one drug that causes a variation in expression of a at least one protein and/or fragment thereof, the method comprising:

(a) administering at least one drug to at least one subject;

(b) obtaining at least one protein from the subject;

(c) labelling at least one amino acid residue of the protein(s) with at least one agent; and

(d) analysing the protein(s) by mass spectrometry, thereby identifying the at least one protein and/or fragment thereof, and determining whether a variation of expression has occurred.

The analysing step may comprise separating the protein(s) based on a first characteristic, optionally separating the protein(s) based on a second characteristic, ionising the protein(s) and detecting and identifying the separated protein(s). The protein and/or fragment thereof may be a PA2G4 protein and/or fragment thereof. The PA2G4 protein has the sequence of SEQ ID NO:2. The agent may be a sulfenyl halide. In particular, the agent may be 2-nitrobenzenesulfenyl chloride (NBS). The amino acid residue may be tryptophan.

The separating steps may be based on at least one characteristic selected from the group consisting of mass, charge, functional group and isoelectric point. The separating steps may be steps for SDS-polyacrylamide gel electrophoresis (SDS-PAGE), two-dimensional gel electrophoresis (2DE), high performance liquid chromatography (HPLC) and/or free flow electrophoresis (FFE). The ionising may be selected from the group consisting of electrospray ionising (ESI), matrix assisted laser desorption/ionisation (MALDI), surface enhanced

laser desorption/ionisation (SELDI), nanospray ionisation, atmospheric pressure chemical ionisation (APCI), chemical ionisation (CI), electron impact (EI), fast atom bombardment (FAB), field desorption / field ionisation (FD/FI) and thermospray ionisation (TSP). The analysing step may be by tandem mass spectrometry. The method tandem mass spectrometry may use quadrupole time-of-flight geometry.

There is also provided a method of selecting at least one candidate for a clinical trial, experimentation and/or diagnostic test, comprising:

(a) providing at least one sample from at least one subject;

(b) determining the expression of the PA2G4 gene and/or mutation thereof, or PA2G4 protein, derivative, mutant and/or fragment thereof; and

(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression indicating the suitability of the subject as a candidate.

The control may be at least one subject not diagnosed with a cell proliferative disorder. The control may also be at least one subject having a low or reduced level of expression of PA2G4 gene and/or protein. The control may also be an average value in expression obtained from a selected population.

There is also provided a diagnostic and/or prognostic kit for the diagnosis and/or prognostic evaluation of a cell proliferative disorder in a subject, the diagnostic and/or prognostic kit comprising at least one molecule or compound hybridisable to a PA2G4 gene, RNA and/or mutation thereof; and/or binding to a PA2G4 protein, derivative, mutant and/or fragment thereof.

The molecule may be a nucleic acid sequence hybridisable to the PA2G4 gene, RNA and/or mutation thereof. The molecule may be a siRNA hybridisable to the PA2G4 gene transcript. The molecule may be a protein binding to the PA2G4 protein, derivative, mutant and/or fragment thereof. The molecule may be an

antibody binding to the PA2G4 protein, derivative, mutant and/or fragment thereof. The kit may further comprise packaging and information pertaining to the use of the compound for purposes such as diagnosis and/or prognostic evaluation of disease states or treatment conditions.

The proliferative cell disorder may be a cancer. In particular, the proliferative cell disorder may be breast cancer. The subject may be a mammal; in particular, a human being.

Brief Description of the Figures

Figure 1. (A) Modification of tryptophan residues with 2-nitrobenzensulfenyl chloride (NBS). * represents either ^{12}C (light) or ^{13}C (heavy). (B) Schematic representation of the NBS/2DE/MS protein identification/quantitation strategy. In this study, the selection of spots for excision was based on a prior analysis using conventional 2DE gels (see Results for details). However, it is also clearly feasible to excise all the observable spots on the NBS gel for subsequent analysis, if a more global proteomic survey is required.

Figure 2. (A) MALDI-TOF spectrum of tryptic peptides of chicken ovalbumin. Two pairs of NBS-labeled peptides (m/z 1734.7/1740.7 and m/z 2012.0/2018.0) are highlighted. (B) Mass spectra of the NBS-labeled peptides pair (m/z 1734.7/1740.7), showing the consistency between the amount of protein labeled and the resulting MALDI-TOF MS spectra.

Figures 3. (A) NBS-labeled peptide pair m/z 1135.5/1141.5 from protein Keratin-18 (K1C18_HUMAN) was acquired by MALDI-TOF MS using the CHCA/DHB matrices; (B) The same peptide pair was analysed using a combination of HNBA and CHCA/DHB matrices. To facilitate the comparison, the scale of the full range spectra (upper panels) was set to 100 mV, while the scale of the zoom-in spectra (lower panels) was set to 10 mV. The figures on

the right show that HNBA in combination with CHCA/DHB greatly enhanced the ionisation of NBS-labeled peptides.

Figures 4. 2DE map of NBS-labeled protein spots that were excised and identified by MS analysis. Fifty-eight protein spots were positively identified by peptide mass fingerprinting from 88 spots with differential fluorescence stain intensity as determined by 2DE analysis. Among them, 44 protein spots were detected with NBS-labeled peptide pairs, which were quantified by calculating $^{13}C/^{12}C$ peptide volume ratios.

Figures 5. Correlation of relative protein quantitation (ratios of TAM and vehicle treated MCF-7 cells) by NBS/2DE/MS and 2DE approaches. Among the 44 differentially expressed proteins, the correlation coefficient between the ratios estimated by the two methods was 0.51, indicating a statistically significant correlation between these two protein quantitation methods. Spots 14 and 21 are highlighted for further concordance study (see Figures 6 and 7).

Figure 6. Spot 14 (disulfide isomerase ER-60) has good concordance between the NBS/2DE/MS and 2DE quantitation. PDQuest 3D image analysis on this spot reveals a clear single image peak. In contrast, Spot 21 (CK19) has poor concordance between the NBS/2DE/MS and 2DE quantitation. 3D image analysis indicates that there are at least 2 image peaks detected in the protein spot, indicating co-migration of proteins.

Figure 7. Mass spectra of tryptic peptides of Spot-21 (CK19). The peptide sequencing mass spectra of the matched (m/z 1222.75) and unmatched peptide (m/z 1419.85) are shown with their y-ions. The matched peptide 1222.75 (TKFETEQALR) (SEQ ID NO:56) was found from CK19, while the unmatched peptide m/z 1419.85 (LEGLTDEINFLR) (SEQ ID NO:57) was from CK8.

Figure 8. Two isoforms of the heat shock 60 kDa protein (CH60_HUMAN) with different isoelectric points (Spots 7 and 8). (A) Based on the NBS labeling quantitation, this protein was up-regulated on Spot 8 but down-regulated on Spot 7. (B) The ratios of fluorescence labeling of 2DE spots were concordant with the NBS labeling ratios.

Figure 9.. Technical Reproducibility of the NBS/2DE/MS Platform. Quantitative fold changes of 40 protein spots from the same biological batch of MCF-7 cells, labelled twice, were compared. Considering 95% of the spots, the linear regression was $y = 0.8884x + 0.1452$ with $R2 = 0.71$. Considering 100% of the spots, the linear regression was $y = 0.7751x + 0.201$ with $R2 = 0.64$.

Figure 10. Validation of NBS quantification by immunoblotting. Two representative proteins GRP78 and CK19, which were found to change upon TAM treatment, were validated by immunoblotting with the protein specific antibodies. (A) GRP78 is up-regulated in TAM treated MCF-7 cells. (B) (Top) Immunoblot revealing the down-regulation of PA2G4 protein at 24 and 48 h after TAM treatment. Blot with ß-Actin antibody is shown for the equal loading of total cell lysates. (Below) RT-PCR amplification of PA2G4 and ß-actin transcripts from oligo-dT primed cDNA shows the relative down-regulation of PA2G4 mRNA upon TAM treatment.

Figure 11. 2DE images of MCF-7 cells treated with TAM (A) or vehicle (B) for 48 hr, and the NBS-labeled combined TAM or vehicle treated cells (C). All the images show very similar protein patterns. The results of qualitative and quantitative image analyses are presented in Figure 12.

Figure 12. PDQuest image analysis results comparing the 2DE images of TAM treated or vehicle treated (untreated) MCF-7 cells. Triplicate gels were run for each treatment. (A) Match rates for the two sets of gels was ca. 85-95% among the ca. 340 visible gel spots; (B) Numbers of differentially displayed protein spots at qualitative and quantitative levels.

Detailed Description

Bibliographic references mentioned in the present specification are for convenience listed in the form of a list of references and added at the end of the examples. The whole content of such bibliographic references is herein incorporated by reference.

Definitions

Biological material – any material from a biological source. The material may comprise biomolecules, cells, tissue and/or fluid. The material may be obtained from cell culture, an animal, mammalian or human subject.

Biomolecule – A biological molecule such as protein acid, nucleic acid (DNA and/or RNA), lipid, carbohydrate, and their derivatives.

Biomarker - Biomarkers are biochemical or molecular molecules (proteins, polypeptides, carbohydrates, lipids, etc, and their derivatives) associated with the presence and severity of specific disease states or different treatment conditions. Under the present invention, a protein or protein derivative may be used as a biomarker. Biomarkers are detectable and measurable by a variety of methods. To identify a biomarker, it is usually necessary to detect a difference or change in the expression or abundance of a biological molecule and identifying the particular molecule.

Expression or abundance of a biomarker – the expression of a biomarker may be determined from the presence or abundance of its gene, gene transcript, and gene product. The terms "gene" and "gene transcript" include RNA sequences complementary to the gene and cDNA sequences obtained by reverse transcription of the gene transcript. The terms also include the wild type gene, variations and mutations of the gene and gene transcript wherein the variation or mutation share substantial identity with the gene or gene transcript. Similarly, "gene product" includes wild type gene product, variations, fragments or derivatives thereof. "Substantial identity" means that the variants of the gene or gene product retain sufficient identity for them to be detectable by methods

and probes used for the wild type gene or gene product and retain the same functions as the wild type gene or gene products. The determination may be qualitative such as whether the biomarker is expressed, or the determination may be quantitative, or the determination may be semi-quantitative by any method known in the art such as by microarray technology or polymerase chain reaction. Departure (increase or decrease) from normal levels in the undiseased state can indicate a disease state or predisposition to a disease state. For example, the overexpression or increased abundance of a protein may be indicative of a disease state, the severity of the disease state and hence the prognosis for the subject in which the biomarker was determined, and/or it might be an indicator of susceptibility to a disease state. The difference in expression and abundance may be determined between different biomarkers or between the same biomarker under different conditions or time points. The expression of the biomarker(s) may then be compared and correlated to a reference value or to other values obtained at different time points or between different biomarkers to determine the correlation with the presence, severity of a disease state, to determine the efficacy of a treatment, or to determine a prognosis of the disease outcome for the subject. Reference values may be determined from a statistically significant number of subjects suffering from or not suffering from. The expression of the biomarker(s) may also be used in conjunction with other suitable diagnostic or prognostic markers, biomarkers or indices to obtain a higher level of confidence. When a biomarker is said to be "over expressed" when compared to controls, it is meant that the expression of that biomarker is at an abundance or level that is statistically significantly more than that naturally expressed by at least one wild type or non-mutant control subject not diagnosed with the disease state or condition. Similarly, when a biomarker is "under expressed", the expression of that biomarker is statistically significantly less than that naturally expressed by at least one wild type or non-mutant control subject not diagnosed with the disease state or condition. Under this definition, a subject genetically deficient for that biomarker cannot be said to be under expressing that biomarker. Similarly, when transfected with the gene for that biomarker so that that

biomarker is expressed, the deficient subject cannot be said to be over expressing the biomarker as he was originally deficient in that biomarker.

Control – A reference subject, experiment or value by which values obtained in tests can be compared against. Control values or ranges s usually represent the "normal" state so that a statistically difference or deviation of the control values or ranges represents an abnormal or disease state. A person skilled in the art will know how to select and/or obtain control subjects, experiments or values for use as references.

Selecting a subject or candidate – to determine the suitability of a subject or candidate for a clinical trial, experimentation, diagnostic and/or other tests by measuring at least one characteristic of the subject or candidate. The value representing the characteristic is then compared to a reference or control value or range of values, and the result is used to determine if the subject or candidate is suitable for the trial, experimentation, diagnostic and/or other test. A person skilled in the art will know how to select subjects or candidates based on their amenability to a particular treatment, or their susceptibility to a particular challenge or disease.

Cell proliferation disorder – a condition in which there is abnormal growth of cells.

Cancer – A malignant and uncontrolled growth of cells in one part of the body that can spread to other parts of the body.

Nucleic acid – "Nucleic acids" are linear polymers of nucleotides, linked by 3', 5' phosphodiester linkages. In DNA, deoxyribonucleic acid, the sugar group is deoxyribose and the bases of the nucleotides adenine, guanine, thymine and cytosine. RNA or ribonucleic acid, has ribose as the sugar and uracil replaces thymine. An "isolated nucleic acid" is a nucleic acid the structure of which is not identical to that of any naturally occurring nucleic acid or to that of any fragment of a naturally occurring genomic nucleic acid. The term therefore covers, for example, (a) a DNA which has the sequence of part of a naturally occurring genomic DNA molecule but is not flanked by both of the coding sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (b) a nucleic acid incorporated into a vector or into the

genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. The invention also features an isolated nucleic acid at least 50% (including any percentage between 50% and 100%, e.g., 85%, 95%, or 100%) identical to SEQ ID NO: 1, which can be used, e.g., for detection of the human PA2G4 gene in a subject, or screening of therapeutic compounds for treating a cell proliferation-associated disorder.

Gene product – Under the present invention, the gene product refers to a gene product other than a protein. The gene product may be a length of RNA with a biological functional other than a gene transcript coded for by a particular gene.

Gene transcript – the messenger RNA sequence coded for by a gene.

Protein - A biological molecule composed of one or more chains of amino acids in a specific order, coded for by a particular gene. As such, any mutation in the gene may be reflected in a mutant protein being expressed. Under the present invention, a mutant protein that has substantially the same biological functions and activities of the non-mutant protein is considered to be the same as the non-mutant protein. Proteins may have derivatives such as isoforms. A protein isoform is a version of a protein with some small differences, usually a splice variant or the product of some post-translational modification. Under the present invention, a protein also encompasses fragments that are sufficiently large enough for the protein to be detected, identified and/or quantified by the method(s) used. A protein complex is a mixture of different proteins.

Proteome – The set of proteins expressed by a cell or organ at a particular time and under specific conditions. Accordingly, proteomic analysis is the study of the full set of proteins encoded by a genome.

Percentage Identity - The "percent identity" of two nucleic acid sequences is determined using the algorithm of Karlin and Altschul (1990), modified as in Karlin and Altschul (1993). Such an algorithm is incorporated into the XBLAST programs of Altschul et al. (1990). BLAST nucleic acid searches are performed

with the XBLAST program. Where gaps exist between two sequences, Gapped BLAST is utilised as described in Altschul et al. (1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST) are used. See the World Wide Web address http://ncbi.nih.gov. The percentage identity of other sequences such as protein sequences may similarly be determined.

Antibody - An immunoglobulin protein produced by B-lymphocytes of the immune system that binds to a specific antigen molecule. The term includes monoclonal antibodies, polyclonal antibodies as well as fragments thereof, such as Fab, F(ab')2, and Fv fragments.

Drug – any compound used for the treatment (amelioration of, reduction of, or cure for) any disease or pathological state or condition.

Screening – to pick out a compound that has an effect in any disease or pathological condition. Screening may involve detecting, identifying and/or quantifying any biomolecule such as a protein, the expression of which is affected by the compound.

Treating - the administration of a composition or drug to a subject with a cell proliferation-associated disorder; with the purpose to cure, alleviate, relieve, remedy, prevent, or ameliorate the disorder, the symptom of the disorder, the disease state secondary to the disorder, or the predisposition toward the disorder. An "effective amount" is an amount of the composition that is capable of producing a medically desirable result, e.g., as described above, in a treated subject. This method can be performed alone or in conjunction with other drugs or therapy. Treating includes reduction of the biomarker to a level found in the non-disease state.

Hybridizing, hybridisable – a biomolecule such as a sequence of nucleic acids may hybridize to another sequence of nucleic acids if there are sufficient stretches of complementary nucleotides between them. For peptides and proteins, one protein such as an antibody or fragment thereof may hybridize to another protein if there are sufficient complementary structures between these two biomolecules.

Diagnose, diagnosis – To diagnose a disease is to is to determine the nature or the identity of a disease. A diagnosis may be accompanied by a determination as to the severity of the disease.

Prognosticate, prognosis, - To prognosticate is to predict the outcome or prognosis of a disease, such as to give a chance of survival based on observations and results of clinical tests.

Predisposition – The likelihood of being diagnosed with, or susceptibility to a particular disease.

Blood derivative – Blood is made up of a cellular and a liquid component. A blood derivative is any of one or more of the blood components less than whole blood.

The present invention addresses the problems above and provides methods for detecting and/or treating a proliferative cell disorder such as cancer. In general, the disorder is cancer, particularly breast cancer.

In one aspect, the present invention provides a method of detecting and/or quantitating the presence of, predisposition to, and/or severity of, a proliferative cell disorder in a subject, the method comprising:

(a) providing at least one sample from a subject;

(b) determining the expression of a PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof;

(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression indicating the presence of, or predisposition to, and/or severity of a proliferative cell disorder in the subject.

There is also provided a method of monitoring the efficacy of a treatment for a proliferative cell disorder in a subject, the method comprising:

(a) providing at least two samples from a subject, each sample obtained at different time points;

(b) determining the expression of a PA2G4 gene, RNA, and/or mutation thereof, or PA2G4 protein, derivative, mutant and/or fragment thereof; and

(c) comparing the expression of the PA2G4 gene, RNA, and/or mutation thereof, and/or PA2G4 protein, derivative, mutant and/or fragment thereof, in the at least two samples.

There is also provided a method of prognosticating the outcome of a cell proliferative cell disorder in a subject, the method comprising:

(a) providing at least one sample from a subject;

(b) determining the expression of a PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof; and

(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression indicating the prognosis of a proliferative cell disorder in the subject.

The control may be, for example, is at least one subject not diagnosed with a cell proliferative disorder. The control may also be a subject having a low level of expression of PA2G4 gene and/or protein. The control may also be an average value in expression obtained from a selected population.

The gene, RNA and/or mutation thereof may be human PA2G4 gene, RNA and/or mutation thereof, and the PA2G4 protein, derivative, mutant and/or fragment may be human PA2G4 protein, derivative, mutant and/or fragment thereof.

The detecting may be by labelling the PA2G4 protein, derivative, mutant and/or fragment thereof, with at least one sulfenyl halide agent followed by 2-

dimensional gel electrophoresis and mass spectrometry. The sulfenyl halide agent may be 2-nitrobenzenesulfenyl chloride (NBS).

There is also provided a method of treating a proliferative cell disorder in a subject, the method comprising varying the expression of a PA2G4 gene, RNA, and/or mutation thereof.

The PA2G4 gene may have a sequence as given in SEQ ID NO:1. The varying may be obtained by administering nucleic acid hybridisable to all or part of the PA2G4 gene, RNA and/or mutation thereof. The hybridisable nucleic acid may be DNA or RNA. In particular, the hybridisable nucleic acid may be siRNA. The varying may be obtained by administering a compound hybridisable to all or part of the PA2G4 gene, RNA and/or mutation thereof.

There is also provided a method of treating a proliferative cell disorder in a subject, the method comprising varying the expression of a PA2G4 protein, derivative, mutant and/or fragment thereof.

The method may comprise varying the expression of the PA2G4 protein having the amino acid sequence of SEQ ID NO:2. The varying may be obtained by administering a polypeptide binding to all or part of the amino acid sequence of the PA2G4 protein, derivative, mutant and/or fragment thereof. The binding polypeptide may be an antibody. The varying may be obtained by administering a compound hybridisable to all or part of the amino acid sequence of the PA2G4 protein.

The varying of the expression may comprise reducing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof.

There is also provided a method of screening for at least one drug that causes a variation in expression of a at least one protein and/or fragment thereof, the

method comprising: administering at least one drug to at least one subject; obtaining at least one protein from the subject; labelling at least one amino acid residue of the protein(s) with at least one agent; and analysing the protein(s) by mass spectrometry, thereby identifying the at least one protein and/or fragment thereof, and determining whether a variation of expression has occurred.

The analysing step may comprise separating the protein(s) based on a first characteristic, optionally separating the protein(s) based on a second characteristic, ionising the protein(s) and detecting and identifying the separated protein(s). The protein and/or fragment thereof may be a PA2G4 protein and/or fragment thereof. The PA2G4 protein has the sequence of SEQ ID NO:2. The agent may be a sulfenyl halide. In particular, the agent may be 2-nitrobenzenesulfenyl chloride (NBS). The amino acid residue may be tryptophan.

The separating steps may be based on at least one characteristic selected from the group consisting of mass, charge, functional group and isoelectric point. The separating steps may be steps for SDS-polyacrylamide gel electrophoresis (SDS-PAGE), two-dimensional gel electrophoresis (2DE), high performance liquid chromatography (HPLC) and/or free flow electrophoresis (FFE). The ionising may be selected from the group consisting of electrospray ionisation (ESI), matrix assisted laser desorption/ionisation (MALDI), surface enhanced laser desorption/ionisation (SELDI), nanospray ionisation, atmospheric pressure chemical ionisation (APCI), chemical ionisation (CI), electron impact (EI), fast atom bombardment (FAB), field desorption / field ionisation (FD/FI) and thermospray ionisation (TSP). The analysing step may be by tandem mass spectrometry. The method tandem mass spectrometry may use quadrupole time-of-flight geometry.

There is also provided a method of selecting at least one candidate for clinical trial(s), experimentation and/or diagnostic test, comprising:

(a) providing at least one sample from at least one subject;

(b) determining the expression of the PA2G4 gene and/or mutation thereof, or PA2G4 protein, derivative, mutant and/or fragment thereof; and

(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression indicating the suitability of the subject as a candidate.

There is also provided a diagnostic and/or prognostic kit for the diagnosis and/or prognostic evaluation of a cell proliferative disorder in a subject, the diagnostic and/or prognostic kit comprising at least one molecule or compound hybridisable to a PA2G4 gene, RNA and/or mutation thereof; and/or binding to a PA2G4 protein, derivative, mutant and/or fragment thereof.

The molecule may be a nucleic acid sequence hybridisable to the PA2G4 gene, RNA and/or mutation thereof. The molecule may be a siRNA hybridisable to the PA2G4 gene transcript. The molecule may be a protein binding to the PA2G4 protein, derivative, mutant and/or fragment thereof. The molecule may be an antibody binding to the PA2G4 protein, derivative, mutant and/or fragment thereof.

The proliferative cell disorder may be a cancer. In particular, the proliferative cell disorder may be breast cancer. The subject may be a mammal; in particular, a human being.

The present application describes the novel use of a gene, gene transcript, RNA gene product and/or mutation thereof, and/or protein, derivative, mutant and/or fragment thereof, for the diagnosis of breast cancer, for determining a prognosis and for the monitoring of the efficacy of a treatment for breast cancer. Through the use of mass spectrometry (MS), the applicant made a surprising discovery that the proliferation associated gene 2G4 (PA2G4) gene is suitable as a biomarker for these purposes.

Mass spectrometry (MS) may be used to analyse molecules. In an MS system, molecules of interest are ionised, separated in an analyzer based on their m/z ratios and then detected by a detector. The results may be displayed in the form of a m/z spectrum. The MS system usually includes a data analysis sub-system for the analysis of the data.

Ionisation of the molecules may be achieved by many methods, including: Atmospheric Pressure Chemical Ionisation (APCI), Thermospray Ionisation (TSP), Chemical Ionisation (CI), Electron Impact (EI), Electrospray Ionisation (ESI), Fast Atom Bombardment (FAB), Field Desorption / Field Ionisation (FD/FI), Surface Enhanced Laser Desorption Ionisation (SELDI) and Matrix Assisted Laser Desorption Ionisation (MALDI).

The analyser may utilize different principles or geometries for the analysis of the ions, including quadrupoles, magnetic sectors, and both Fourier transform, quadrupole ion traps and time-of-flight (TOF) geometries. Two or more analysers may be arranged in tandem. Commonly used tandem geometries include quadrupole-quadrupole, magnetic sector-quadrupole, and quadrupole-time-of-flight geometries.

In MS, a person skilled in the art will be able to select different ionisation, analysis and detection methods to best study a biomolecule of interest such as proteins.

Applicants used 2-nitrobenzenesulfenyl chloride (NBS) labeling technology, (Kuyama et al. 2003) in combination with 2DE and MS for quantitative proteome analysis (NBS/2DE/MS). NBS labeling involves the binding of light and heavy NBS moieties to tryptophan residues, resulting in a 6 Da mass difference between light and heavy NBS labeled tryptophan peptides.

The NBS/2DE/MS strategy to study proteomic alterations in MCF-7 breast cancer cells induced by tamoxifen (TAM) treatment. TAM is the first line

hormonal treatment for estrogen receptor positive (ER+) breast cancer, and although previous reports have described the mRNA gene expression patterns in breast cancers and cell lines associated with TAM treatment (Saji et al, 2005; Zhang et al, 1999). Using NBS/2DE/MS, inventors selected 88 protein spots that were quantitatively altered between TAM-treated and control MCF-7 cells, of which 44 could be successfully identified and quantified by NBS labeling. Three proteins, CK19, GRP78, and PA2G4, were also validated as bona-fide TAM-responsive proteins in MCF-7 cells by Western blotting, and further shown that in primary breast cancers, PA2G4 expression behaves as a novel prognostic factor in two independent patient cohorts.

Among the proteins studied, Proliferation Associated 2G4 (PA2G4), has very good potential for use as a biomarker for breast cancer. The gene sequence is given as SEQ ID NO: 1 and the gene product (protein) sequence is given in SEQ ID NO: 2.

PA2G4 (Ebp-1) belongs to a large family of cell cycle regulatory proteins or replication proteins that maintain the cell cycle activities of proliferating cells. It appears to be a proliferation-dependent gene that probably encodes a nuclear DNA-binding protein involved in the control of cell replication. PA2G4/Ebp-1 has been shown to participate in the differentiation of human ErbB receptor-positive breast and prostate cancer cells, and is capable of inducing anti-proliferative effects by virtue of its interactions with the retinoblastoma (Rb) tumor suppressor protein (Rao et al., 2002). This protein is found in both mammals and humans.

The mitogenic effects of E2 have been attributed in large part to its ability to increase the expression of key cell cycle regulatory genes. There was a report of comprehensive gene expression analysis of the effects of E2; using a classical E2-responsive human model, and the identification of some E2 targeted genes. It was found that PA2G4 was over-expressed in MCF-7 even

after 3 hours of E2 treatment. In the study, they used various ER+ and ER- breast cancer cell lines (Pompeia et al, 2004).

Two other proteins, CK19 and GRP78, were validated as bona-fide TAM responsive proteins in MCF-7 cells.

The identification of drug-responsive biomarkers in complex protein mixtures is an important goal of quantitative proteomics. In general, we developed a novel approach for identifying such drug-induced protein alterations which combine labelling of proteins with an agent followed by separation of the proteins and analysis by mass spectrometry. In particular, as an example, we used 2-nitrobenzenesulfenyl chloride (NBS) tryptophan labeling with two-dimensional gel electrophoresis (2DE)/mass spectrometry (MS). Lysates from drug-treated and control samples are labeled with light or heavy NBS residues, then separated on a common 2DE gel. Under the method of the present invention, the proteins of interest are separated and isolated by the 2DE step before analysis by MS. Protein alterations identified by MS through the differential intensity of paired NBS peptide peaks.

Using NBS/2DE/MS, we profiled the proteomic alterations induced by tamoxifen (TAM) in the estrogen receptor (ER) positive MCF-7 breast cancer cell line. Of 88 protein spots that significantly changed upon TAM treatment, 44 spots representing 23 distinct protein species were successfully identified with NBS paired peptides. Of these 23 TAM-altered proteins, 16 (70%) have not been previously associated with TAM or ER activity. We found the NBS labeling procedure to be both technically and biologically reproducible, and the NBS/2DE/MS alterations exhibited good concordance with conventional 2DE differential protein quantitation, with discrepancies due largely to the co-migration of distinct proteins in the regular 2DE gels. To validate the NBS/2DE/MS results, we used immunoblotting to confirm GRP78, CK19, and PA2G4 as bona-fide TAM regulated proteins, and in two independent breast

cancer patient cohorts we demonstrate that PA2G4 expression can serve as a novel prognostic factor for disease free survival.

As PA2G4 has been identified as a reliable biomarker for breast cancer, its gene, gene transcript, RNA gene product, gene derivative, gene mutation, protein, protein derivative, protein mutant, and protein fragment may be used accordingly for this purpose. Any method known in the art may be used to qualitatively detect or quantitatively measure the presence of PA2G4.

Gene Expression

A nucleic acid of the invention may be expressed in vitro by DNA transfer into a suitable host cell by methods known in the art. For example, the nucleic acid may be inserted into a recombinant expression vector. A variety of host-expression vector systems may be utilised to express a nucleic acid of the invention. These include; but are not limited to, micro-organisms such as bacteria transformed with recombinant bacteriophage DNA, plasmid DNA, or cosmid DNA expression vectors; yeast transformed with recombinant yeast expression vectors; and human cell lines infected with recombinant virus or plasmid expression vectors. Isolation and purification of recombinant polypeptides, or fragments thereof, can be carried out by conventional means including preparative chromatography and immunological separations involving monoclonal or polyclonal antibodies.

The invention also features methods for determining prognosis for, diagnosing, treating, and identifying therapeutic compounds for treating a cell proliferation-associated disorder.

A prognostic method of the present invention involves comparing expression of the biomarker in samples from patients with reference values obtained from one or more patients with known disease outcomes, or from known benign or malignant tissues or cells. Thereafter, it will be possible to determine the prognoses of these patients.

A diagnostic method of the invention involves comparing one or more of the PA2G4, GRP78 and CK19 genomic DNA levels in a sample prepared from a subject (i.e., an animal or a human) with that in a sample prepared from a normal subject, i.e., a subject who does not suffer from or at risk for developing a cell proliferation-associated disorder.

For both prognostic and diagnostic methods, a higher PA2G4 genomic DNA level indicates that the subject is suffering from, or at risk for, a cell proliferation-associated disorder. This methods can be used on their own or in conjunction with other procedures to diagnose a cell proliferation-associated disorder in appropriate subjects and to provide a prognosis as to the outcome of the disease.

Amplification of a gene locus can be detected by a variety of methods known in the art. For example, the copy number of a gene locus can be determined and compared by PCR amplification of genomic DNA prepared from a test sample and a control sample. Amplification of a gene locus can also be identified by Southern blot analysis. Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location and an amount of the DNA sequence present in the chromosome.

The invention also provides a method for identifying compounds (e.g., proteins, peptides, peptidomimetics, peptoids, antibodies, or small molecules) that decrease the PA2G4 genomic DNA level, modulate the expression, decrease the level of PA2G4 or PA2G4 mRNA or protein, or decrease the level of PA2G4 interaction with other molecules recognizing it. The resulting identified compounds can then be used, e.g., for preventing and treating a cell proliferation-associated disorder.

The candidate compounds of the invention may be obtained using any of the numerous approaches in combinatorial library methods known in the art. Such libraries include: peptide libraries, peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone that is resistant to enzymatic degradation); spatially addressable parallel solid phase or solution phase libraries; synthetic libraries obtained by deconvolution or affinity chromatography selection; and the "one-bead one-compound" libraries. See, e.g., Lam (1997).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example, in: Gallop et al. (1994).

Libraries of compounds may be presented in solution (e.g., Houghten, 1992), or on beads (Lam, 1991), chips (Fodor, 1993), bacteria (U.S. Pat. No. 5,223,409), spores (U.S. Pat. No. 5,223,409), plasmids (Cull et al, 1992), or phages (U.S. Pat. No. 5,223,409).

To identify compounds that decrease the PA2G4 genomic DNA level, modulate expression of a marker nucleotide sequence under the control of a promoter associated with PA2G4, decrease the level of PA2G4 mRNA or protein, decrease the level of PA2G4, or decrease the level of PA2G4 in a system; the system is contacted with a candidate compound and the PA2G4 genomic DNA level, the expression of the marker nucleotide sequence, the level of PA2G4 mRNA or protein, the level of PA2G4 interaction, is evaluated relative to that in the absence of the candidate compound. In a cell system, the cell can contain amplified PA2G4 genomic DNA (e.g., a cancer cell), naturally express PA2G4, be modified to express a recombinant nucleic acid, for example, having a PA2G4 gene fused to a heterologous promoter. If the level of PA2G4 genomic DNA is changed due to the candidate compound; that is, either lower in the presence of the candidate compound than in its absence, or if the expression of the marker nucleotide sequence in the presence of the candidate compound differs from that in the absence of the candidate compound, the candidate

compound is identified as being useful for preventing and treating a cell proliferation-associated disorder.

The PA2G4, GRP78 and CK19 genomic DNA levels can be determined by methods described above and any other methods well known in the art.

Gene Product Expression

The expression of the marker nucleotide sequence, PA2G4, GRP78 and CK19 can be determined either at the mRNA level or at the protein level. Methods of measuring mRNA levels in a tissue or a body fluid sample are known in the art. In order to measure mRNA levels, cells can be lysed and the levels of mRNA in the lysates or in RNA purified or semi-purified from the lysates can be determined by any of a variety of methods including; without limitation, hybridization assays using detectably labeled DNA or RNA probes and quantitative or semi-quantitative RT-PCR methodologies using appropriate oligonucleotide primers. Alternatively, quantitative or semi-quantitative in situ hybridization assays can be carried out using, for example, tissue sections or non-lysed cell suspensions, and detectably (e.g., fluorescently or enzyme) labeled DNA or RNA probes. Additional methods for quantifying mRNA include RNA protection assay (RPA) and serial analysis of gene expression (SAGE).

While the method of NBS/2DE/MS has been described, other methods of measuring protein levels in a tissue or body fluid sample are also known in the art. Many such methods employ antibodies (e.g., monoclonal or polyclonal antibodies) that bind specifically to the target protein. In such assays, the antibody itself or a secondary antibody that binds to it can be detectably labeled. Alternatively, after the antibody is conjugated with biotin, binding with detectably labeled avidin (a polypeptide) can be used to detect its presence. Combinations of these approaches (including "multi-layer sandwich" assays) familiar to those in the art can be used to enhance the sensitivity of the methodologies. Some of these protein-measuring assays (e.g., ELISA or Western blot) can be applied to lysates of cells, and others can be applied to

histological sections or non-lysed cell suspensions, e.g., immunohistological methods or fluorescence flow cytometry.

Methods of measuring the amount of label will be depend on the nature of the label and are well known in the art. Appropriate labels include, without limitation, radionuclides (e.g., ^{125}I, ^{131}I, ^{35}S, ^{3}H, or ^{32}P), enzymes (e.g., alkaline phosphatase, horseradish peroxidase, luciferase, or beta-galactosidase), fluorescent moieties or proteins (e.g., fluorescein, rhodamine, phycoerythrin, GFP, or BFP), or luminescent moieties (e.g., Qdot.TM. nanoparticles supplied by the Quantum Dot Corporation, Palo Alto, California). Other applicable assays include quantitative immunoprecipitation or complement fixation assays.

The level of PA2G4 can be determined by any method known in the art, e.g., by *in vitro* binding assay or by using a yeast hybrid system. The binding domains of PA2G4 can be identified, e.g., by mutagenesis, and used in the screening assays.

Treatment

This invention further provides a method for preventing and treating a cell proliferation-associated disorder. Subjects to be treated can be identified; for example, by determining the PA2G4 genomic DNA, DNA transcript or DNA product level in a sample prepared from a subject by methods described above. If the PA2G4 genomic DNA level is higher in the sample from the subject than that in a sample (control) from a normal subject, the subject is a candidate for treatment with an effective amount of compound that decreases the PA2G4 genomic DNA, DNA transcript or DNA product level.

In an *in vivo* approach, a therapeutic composition (e.g., a composition containing a compound identified as described above) is administered to the subject. Generally, the compound will be suspended in a pharmaceutically-acceptable carrier (e.g., physiological saline) and administered orally or by intravenous infusion, or injected or implanted subcutaneously, intramuscularly,

intrathecally, intraperitoneally, intrarectally, intravaginally intranasally, intragastrically, intratracheally, or intrapulmonarily. For prevention and treatment of cancer, the compound can be delivered directly to the cancer tissue.

The dosage required depends on the choice of the route of administration; the nature of the formulation; the nature of the subject's illness; the subject's size, weight, surface area, age, and sex; other drugs being administered; and the judgment of the attending physician. Suitable dosages are in the range of 0.01-100.0 mg/kg. Wide variations in the needed dosage are to be expected in view of the variety of compounds available and the different efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization as is well understood in the art. Encapsulation of the compound in a suitable delivery vehicle (e.g., polymeric microparticles or implantable devices) may increase the efficiency of delivery, particularly for oral delivery.

Another way to achieve uptake of the nucleic acid is using liposomes, prepared by standard methods. The vectors can be incorporated alone into these delivery vehicles or co-incorporated with tissue-specific antibodies. Alternatively, one can prepare a molecular conjugate composed of a plasmid or other vector attached to poly-L-lysine by electrostatic or covalent forces. Poly-L-lysine binds to a ligand that can bind to a receptor on target cells (Cristiano et al, 1995). Alternatively, tissue specific targeting can be achieved by the use of tissue-specific transcriptional regulatory elements (TRE) which are known in the art. Delivery of "naked DNA" (i.e., without a delivery vehicle) to an intramuscular, intradermal, or subcutaneous site is another means to achieve in vivo expression.

In the relevant polynucleotides (e.g., expression vectors), the nucleic acid sequence encoding an anti-sense PA2G4 RNA is operatively linked to a

promoter or enhancer-promoter combination. Enhancers provide expression specificity in terms of time, location, and level. Unlike a promoter, an enhancer is able to function when located at variable distances from the transcription initiation site, provided a promoter is present. An enhancer can also be located downstream of the transcription initiation site.

The inhibition of gene expression may be achieved through use of small inhibitory RNA sequences (siRNA). Expression of the PA2G4 gene may also be inhibited using RNA interference (RNAi). This is a technique for post-transcriptional gene silencing ("PTGS"), in which target gene activity is specifically abolished with cognate double-stranded RNA ("dsRNA"). In many embodiments, dsRNA of about 18-26 nucleotides, in particular, 20-24 nucleotides, more particularly 21-23 nucleotides. For example a 21-nucleotide sequence homologous to the target gene, is introduced into the cell and a sequence specific reduction in gene activity is observed. RNA interference provides a mechanism of gene silencing at the mRNA level. It offers an efficient and broadly applicable approach for gene knock-out as well as for therapeutic purposes.

Sequences capable of inhibiting gene expression by RNA interference may be of any desired length. For instance, the sequence can have at least 15, 20, 25, or more consecutive nucleotides. The sequence may be dsRNA or any other type of polynucleotide, provided if the sequence may form a functional silencing complex to degrade the target mRNA transcript.

In one embodiment, the sequence consists of or comprises a short interfering RNA (siRNA). The siRNA may be, for example, dsRNA having 19-25 nucleotides. siRNAs may be produced endogenously by degradation of longer dsRNA molecules by an RNase III-related nuclease called Dicer. siRNAs may also be introduced into a cell exogenously or by transcription of an expression construct. Once formed, the siRNAs and protein components assemble into endoribonuclease-containing complexes known as RNA-induced silencing

complexes (RISCs). An ATP-generated unwinding of the siRNA activates the RISCs, which in turn targets the complementary mRNA transcript by Watson-Crick base pairing, thereby cleaving and destroying the mRNA. Cleavage of the mRNA takes place near the middle of the region bound by the siRNA strand. This sequence-specific mRNA degradation results in gene silencing.

Polynucleotides may be administered in a pharmaceutically acceptable carrier. Pharmaceutically acceptable carriers are biologically compatible vehicles that are suitable for administration to an animal or a human, e.g., physiological saline or liposomes. A preferred dosage for administration of polynucleotide is from approximately 10^6 to 10^{12} copies of the polynucleotide molecule. This dose may be repeatedly administered, as needed. Routes of administration may be any of those listed above.

Antibodies (monoclonal or polyclonal) to PA2G4 may be used to reduce the level of PA2G4 protein, or to decrease the level of PA2G4 in a subject. The term "antibody" includes intact molecules as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv which are capable of binding to an epitopic determinant present in the PA2G4 protein. Methods of making monoclonal and polyclonal antibodies and fragments thereof are known in the art. See, for example, Harlow and Lane, (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, New York.

Other compounds that may be used to inhibit PA2G4 expression include a peptide or polypeptide containing an amino acid sequence at the C-terminus of the PA2G4 protein.

In addition, the invention provides a method of developing a procedure for treating a cell proliferation-associated disorder by providing a plurality of subjects suffering from a cell proliferation-associated disorder (e.g., subjects having amplified PA2G4 genomic DNA or PA2G4 protein expression); administering to each subject an effective amount of an PA2G4 inhibitor

followed by an effective amount of a nucleotide or nucleoside analogue (e.g., gemcitabine), each at a unique time point; and selecting an optimal time point at which the cell proliferation-associated disorder is inhibited to the greatest extent. Once the optimal time point has been identified, the procedure may be used to treat a cell proliferation-associated disorder in appropriate subjects.

The specific examples below are to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. Without further elaboration, it is believed that one skilled in the art may, based on the description herein, utilize the present invention fully. All publications recited herein are hereby incorporated by reference in their entirety.

Having now generally described the invention, the same will be more readily understood through reference to the following examples that are provided by way of illustration, and are not intended to be limiting of the present invention.

Examples

Materials and Methods

Example 1 – Chemicals and Materials

MCF-7 breast cancer cells were obtained from the American Type Culture Collection (Manassas, VA, USA). Tamoxifen (TAM), chicken ovalbumin, streptomycin, L-glutamine, penicillin, 4-hydroxycinnamic acid (CHCA), 3-hydroxy-4-nitrobenzoic acid (HNBA), 2,5-dihydroxybenzoic acid (DHB), protease inhibitor cocktail, phenyl methyl sulfonyl fluoride (PMSF), sodium deoxychlate and sodium orthovanadate were from Sigma (St Louis, MO, USA). The sulfenyl halide agent and 2-nitrobenzenesulfenyl chloride (NBS, heavy and light forms) were from Shimadzu Biotech (Kyoto, Japan). Tris (2-carboxyethyl) phosphine (TCEP) hydrochloride was from Pierce (Rockford, IL, USA). Deep Purple fluorescent dye, dithiotritol (DTT) and horseradish peroxidase conjugated secondary antibodies were from GE HealthCare (Uppsala,

Sweden). Ethylene diamine tetra acetic acid disodium salt dihydrate (EDTA) and polyvinylidene difluoride membrane (PVDF) were from Bio-Rad (Hercules, CA, USA). C18 ZipTips were from Millipore (Bedford, MA, USA). Sequencing grade modified trypsin was from Promega (Madison, WI, USA). Dulbecco modified Eagle medium 5 (DMEM) was from Gibco (Grand Island, NY, USA). Dextran charcoal-stripped fetal bovine serum (FBS) was from HyClone Laboratories (Pittsburg, PA, USA). Anti-vinculin and anti-PA2G4 antibodies were from Upstate Biotechnology (VA, USA). Anti-cytokeratin 19 and anti-GRP78 antibodies were from Santa Cruz Biotechnology (Santa Cruz Inc, CA, USA). Trizol and Reverse transcriptase enzymes were from Invitrogen (CA, USA). Oligonucleotide primers used in cDNA synthesis and RT-PCR were synthesized from Sigma Proligo (Sigma, Singapore).

Example 2 - Cell Culture and TAM Treatment of MCF-7 Cells

MCF-7 breast cancer cells were cultured in DMEM supplemented with 10% FBS, 100 U/mL penicillin, 100 U/mL streptomycin, and 2 mM L-glutamine. Before TAM treatment, cells were washed in PBS and maintained in phenol red free DMEM with 5 % Dextran charcoal-stripped FBS for 24 hrs. Subsequently, cells were treated with 10 µM TAM at 40-50% confluence and harvested at 24 and 48 hrs. As control, sister cultures were treated with an equivalent volume of the vehicle (0.1% ethanol).

NBS/2DE/MS

Example 3 - Protein Extraction and NBS Labeling

TAM-treated and control MCF-7 cells were lysed in 2DE lysis buffer (7M urea, 2M thiourea, 4% CHAPS, 1mM PMSF, 50 µg/mL Dnase I, 50 µg/mL Rnase I and protease inhibitor cocktail) and the extracted proteins buffer exchanged into 8M urea and 5mM EDTA using a 2D-sample clean up kit (Bio-Rad). Proteins were quantitated by the Bradford method (Bradford, 1976) and 100 µg of each protein lysate was labeled in the dark with a 20-fold molar excess of 12C (light) or 13C (heavy) 2-nitrobenzenesulfenyl chloride (NBS) in 25 µL of acetic acid with shaking. In NBS labeling, a 12C or 13C NBS moiety is

selectively introduced onto tryptophan residues, which imparts a 6 Da mass difference between light and heavy NBS labeled tryptophan peptides. Using mass spectroscopy (MS), the ratio intensity of the paired peptide peaks can then be determined, providing a relative quantitation of the respective light and heavy populations (Figure 1A) (Kuyama et al. 2003).

After labeling, the two protein mixtures were combined and the unlabeled NBS moieties removed by a Sephadex LH-20 column (GE HealthCare). Notably, we have previously shown that the use of a Sephadex LH-20 column to remove excess NBS label does not significantly interfere with protein recovery (see Figure 2 in Matsuo et al, 2006), and that protein recovery is robust regardless of whether the samples are denatured with either urea or GdnHCl. The sample was dried with a vacuum concentrator and resuspended in 50 µL of 50mM Tris buffer (pH 8.5) containing 0.1% SDS and reduced with 4mM TCEP for 30 min at 37 °C, followed by alkylation with 10mM iodoacetetamide for 45 min at room temperature in the dark. The lysate was then buffer exchanged into 2DE lysis buffer (7M urea, 2M thiourea and 4% CHAPS).

Example 4 - Two-dimensional Gel Electrophoresis.

Prior to the first dimensional separation, each IPG strip (18 cm, pH 4-7, GE HealthCare) was rehydrated with 340 µL of rehydration buffer (7M urea, 2M thiourea, 4% CHAPS, 0.5% IPG buffer pH 4-7 and 20mM DTT) for 8 hrs. The first dimensional separation was performed on the IPGphor IEF system (GE HealthCare). Protein lysates of 100 µg in 80 µL of lysis buffer, 20mM DTT and 0.5% IPG buffer, pH 4-7, were applied through cup loading anodically. The strips were then focused at 20°C: 200V for 3 hr; 500V for 1 hr; from 500 to 8000V for 7 hr; and 8000V until 67,000 Vh was reached. After focusing, the IPG strips were equilibrated in two steps: (1) 15 min in 50mM Tris-HCl, pH 8.8, 6M urea, 30% glycerol, 2% SDS, 1% DTT and a trace of bromophenol blue; (2) 15 min in a similar solution containing 2.5% iodoacetamide instead of DTT.

Once equilibrated, the strips were transferred onto 10% isocratic polyacrylamide gels (18cm x 20cm x 0.75mm). The IPG gels were then sealed with 0.5% w/v agarose in running buffer (25mM Tris-HCl, pH8.3, 192mM glycine and 0.1% SDS). The second dimensional separation was then performed using the Protean II XL system (Bio-Rad). The SDS-PAGE was carried out at 17°C with 8mA/gel for 1 hr followed by 24mA/gel constant current, until the dye front reached the bottom of the gel. The gels were then labeled with the Deep Purple fluorescence dye according to the manufacturer's protocol (GE HealthCare). The gels were digitized using a FX Molecular Imager (Bio-Rad), and image analysis performed using PDQuest 7.3 image analysis software (Bio-Rad).

Example 5 - Protein Identification by Mass Spectrometry

Fluorescence labeled protein gel spots were excised and in-gel trypsin digested using the Xcise robotic system (Shimadzu Biotech, Kyoto, Japan). The automatic operation involved washing the gel pieces with water twice, shrinking with 100% acetonitrile (ACN), and drying prior to digestion. Thirty microliters of trypsin (3.33ng/µL in 50mM ammonium bicarbonate, pH 8.5) was added to each gel plug and the digestion was performed overnight at 30 °C with shaking. The samples were then cleaned and concentrated with C18 ZipTips.

Finally, the peptide mixtures were eluted onto the MALDI sample plate with 1 µL of 50% ACN/0.5% TFA and mixed with 1 µL matrix containing 10mg/mL a-cyano-4-hydroxycinnamic acid (CHCA), 2,5-dihydroxybenzoic acid (DHB), and 3-hydroxy-4-nitrobenzoic acid (HNBA) in 50% ACN/0.5% TFA prior to MS analysis. This is a novel matrix formulation based on Matsuo et al. (2006) instead of the regular CHCA/DHB matrix (Laugesen and Roepstorff, 2003) MALDI-TOF MS analyses were performed using the AXIMA-CFR plus mass spectrometer (Shimadzu Corporation, Kyoto, Japan and Kratos Analytical, Manchester, UK), under the following settings : nitrogen laser (337 nm); reflectron mode; detection of positive ions. The acceleration potential was set to 35 kV using a gridless-type electrode. MALDI-TOF MS spectra were acquired in

manual mode, from m/z 800 to 3000 and internally calibrated with two trypsin autolysis peaks (m/z 842.51 and 2211.10). Peak lists from peptide mass mapping spectra were automatically extracted and submitted to an in-house Mascot server (Matrix Science, London, UK) to search against UniProt and NCBInr databases.

To assist searching for NBS-labeled peptides, two modification values of 12CNBS (+153 Da) and 13CNBS (+159 Da) were added to the Mascot server as variable modifications. The NBS-labeled paired peaks were manually selected from the TOF MS spectrum and the relative quantitation was achieved by calculating the volume ratio of each NBS labeled peak pairs using the Kompact software (Shimadzu/Kratos). MS/MS peptide sequencing was performed using the AXIMA-QIT MALDI quadrupole ion trap time-of-flight mass spectrometer (Shimadzu Corporation, Kyoto, Japan and Kratos Analytical, Manchester, UK) equipped with a 337 nm nitrogen laser. One microliter of tryptic-digested sample was deposited on to the MALDI plate and mixed with 1 µL of DHB (10 mg/mL in 50% ACN/0.5% TFA). The TOF spectrum was externally calibrated using fullerite pre-deposited onto the MALDI sample stage. All spectra were acquired with standard instrument settings for optimum transmission at medium and high masses. Data acquisition and processing were performed using the Kompact software (Shimadzu/Kratos).

Example 6 - Statistical analysis

To investigate the biological and experimental variations associated with NBS/2DE/MS, and to compare protein quantities between NBS/2DE/MS and conventional 2DE analysis, we applied a series of Pearson's correlation tests to assess the similarity of the two datasets under investigation. To determine the significance of any observed correlation value, a bootstrap permutation test was performed by scrambling the index of the two datasets being compared, and re-calculating the correlation coefficient for the randomized datasets. For the correlation tests, to determine the significance of any observed correlation value, a bootstrap permutation test was performed by scrambling the order of

the proteins in the two datasets being compared, and re-calculating the correlation coefficient for the randomized datasets. This procedure was repeated 10,000 times, the frequency at which the randomized coefficients exceeded the observed value was computed, and this was reflected as an empirical p-value.

Example 7 - Western Blotting Analysis

Whole cell lysates were prepared by lysing TAM-treated and control cells in RIPA buffer (0.15 M NaCl, 50 mM Tris-HCl (pH 7.4), 1 mM EDTA, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS) containing 1 mM PMSF and 1 mM sodium orthovanadate. The lysates were quantitated with the Bradford reagent using BSA as a standard. Equal amounts of the cell lysate (50 µg) were run in 12% SDS-polyacrylamide gel and electroblotted onto a PVDF membrane. The blots were probed with anti-vinculin, anti-cytokeratin 19 and anti-GRP78 antibodies respectively. After incubating the membrane with horseradish peroxidase-conjugated secondary antibody, the reactivity was visualized on X-ray films (Kodak) using a chemiluminescent detection kit (Bio-Rad).

Example 8 – Semi-quantitative PCR analysis

For cDNA synthesis and RT-PCR, total RNA was extracted from TAM treated MCF-7 cells using Trizol reagent. RNA was quantitated and equal quantities of RNA samples were electrophoresed in a 2% agarose gel to confirm the normalization. Equal quantities of vehicle treated and TAM treated RNA samples were reverse transcribed by Superscript II Reverse transcriptase using oligo-dT primers in the conditions suggested by the manufacturer. RT-PCR was performed from the cDNA using PA2G4 and ß-actin specific primers. Sequences of the oligos used in the study were, PA2G4-RTF: CACGTGCCTTCTTCAGTGAG (SEQ ID NO:58); PA2G4-RTR: ACTCTGGAGGAGGGCCTTTA ((SEQ ID NO:59); ß-actin-RTF: CGGGAAATCGTGCGTGACATT (SEQ ID NO: 60); ß-actin-RTR: TGATCTCCTTCTGCATCCTGT (SEQ ID NO:61).

Example 9 – Quantitative NBS/2DE/MS analyses of TAM-treated MCF-7 cells

We then applied the NBS/2DE/MS approach to complex biological samples, by profiling MCF-7 breast cancer cells treated with either TAM or vehicle. First, to provide a basis of comparison, the MCF-7 samples were analysed using a conventional 2DE platform. In this conventional 2DE analysis, triplicate gels for TAM treated and vehicle treated MCF-7 cells were generated, visualized with Deep Purple fluorescence labelling, and PDQuest software used to identify differences in fluorescence label intensities. The resulting 2DE gels showed highly similar protein patterns for samples treated with either TAM or vehicle, with a 85 to 95% match rate across the approximately 340 protein spots in each gel. Qualitatively, one protein spot was unique in the TAM treated cells, while seven were unique in the vehicle treated cells. At a quantitative level, 27 and 47 protein spots exhibited a >2-fold intensity increase and decrease respectively upon TAM treatment. (Figure 11 and Table 3). Second, the same protein lysates from the TAM-treated and control cells were also subjected to NBS/2DE/MS analysis.

In Table 1 below, the 44 identified NBS protein spots representing 23 different protein species. The 13C/12C ratios were calculated based on the NBS-paired peak intensities. Three proteins were identified with more than one pair of NBS peptides. For them, the 13C/12C ratios were the average of all the paired peptide ratios of the same protein. To test the technical reproducibility, protein lysates from the exact same batch of TAM and vehicle treated MCF-7 cells were re-labelled with NBS, separated on 2DE gels, and re-quantified based on the 13C/12C ratios. Guided by the conventional 2DE gel image analysis, 88 spots were 26 excised from the NBS-labelled 2DE gels, out of which 58 protein spots were positively identified by peptide mass fingerprinting. Spots 1-44 were proteins detected with NBS-labelled peptide peaks. Among them 44 protein spots were detected with NBS-labelled peptide pairs, which were quantified by calculating 13C/12C peptide volume ratio.

TABLE 1

S/N	Acc No	Protein Name	NBS peptide	Sequence	13C/12C (1)	13C/ 12C (2)	Mean ± std
1	P05783	K1C18_HUMAN	1135/1141	DWSHYFK (SEQ ID NO:3)	0.15	0.13	0.13 ± 0.01
2	P05783	K1C18_HUMAN	1135/1141	DWSHYFK (SEQ ID NO:4)	0.24	0.56	0.41 ± 0.22
3	P05783	K1C18_HUMAN	1135/1141	DWSHYFK (SEQ ID NO:5)	0.28	0.38	0.32 ± 0.06
4	P05783	K1C18_HUMAN	1135/1141	DWSHYFK (SEQ ID NO:6)	0.54	0.78	0.65 ± 0.17
5	P10809	CH60_HUMAN	1497/1503	TVIIEQSWGSPK (SEQ ID NO:7)	0.45	0.47	0.46 ± 0.01
6	P10809	CH60_HUMAN	1497/1503	TVIIEQSWGSPK (SEQ ID NO:8)	0.95	1.09	1.02 ± 0.09
7	P10809	CH60_HUMAN	1497/1503	TVIIEQSWGSPK (SEQ ID NO:9)	0.36	0.35	0.35 ± 0.01
8	P10809	CH60_HUMAN	1497/1503	TVIIEQSWGSPK (SEQ ID NO:10)	1.09	1.85	1.46 ± 0.53
9	P04792	HSPB1_HUMAN	1094/1100	HTADRWR (SEQ ID NO:11)	0.09	0.08	0.08 ± 0.01
10	P04792	HSPB1_HUMAN	1094/1100	HTADRWR (SEQ ID NO:12)	0.13	0.50	0.32 ± 0.25
11	P04792	HSPB1_HUMAN	1113/1119	DWYPHSR (SEQ ID NO:13)	0.56	0.70	0.63 ± 0.09
12	P30101	PDIA3_HUMAN	935/941	GSNYWR (SEQ ID NO:14)	0.98	1.07	1.03 ± 0.06

13	P30101	PDIA3_HUMAN	935/941	GSNYWR (SEQ ID NO:15)	0.57	0.76	0.66 ± 0.13
14	P30101	PDIA3_HUMAN	935/941	GSNYWR (SEQ ID NO:16)	1.15	1.27	1.21 ± 0.07
15	P11021	GRP78_HUMAN	1583/1589	TWNDPSVQQDIK (SEQ ID NO:17)	1.81	1.16	1.46 ± 0.45
16	P11021	GRP78_HUMAN	1583/1589	TWNDPSVQQDIK (SEQ ID NO:18)	1.21	1.63	1.41 ± 0.29
17	P11021	GRP78_HUMAN	1583/1589	TWNDPSVQQDIK (SEQ ID NO:19)	1.35	0.92	1.13 ± 0.30
18	P11021	GRP78_HUMAN	1583/1589	TWNDPSVQQDIK (SEQ ID NO:20)	1.31	0.84	1.07 ± 0.33
19	P08727	K1C19_HUMAN	1161/1167 1235/1241	IRDWYQK (SEQ ID NO:21) / DAEAWFTSR (SEQ ID NO:22)	0.35	0.48	0.41 ± 0.18
20	P08727	K1C19_HUMAN	1161/1167 1235/1242	IRDWYQK (SEQ ID NO:23) / DAEAWFTSR (SEQ ID NO:24)	0.30	0.71	0.51 ± 0.30
21	P08727	K1C19_HUMAN	1161/1167 1235/1242	IRDWYQK (SEQ ID NO:25) / DAEAWFTSR (SEQ ID NO:26)	0.19	0.16	0.17 ± 0.11
22	P08727	K1C19_HUMAN	1161/1167 1235/1242	IRDWYQK (SEQ ID NO:27) / DAEAWFTSR (SEQ ID NO:28)	0.29	0.19	0.24 ± 0.15

23	P08727	K1C19_HUMAN	1161/1167 1235/1242	IRDWYQK (SEQ ID NO:29) / DAEAWFTSR (SEQ ID NO:30)	0.68	0.86	0.77 ± 0.21
24	P30040	ERP29_HUMAN	1090/1096 1218/1224	WAEQYLK (SEQ ID NO:31)/ KWAEQYLK (SEQ ID NO:32)	0.87	1.14	1.01 ± 0.18
25	P30040	ERP29_HUMAN	1223/1229	VGAIQRWLK (SEQ ID NO:33)	NA	0.71	0.71
26	P07237	PDIA1_HUMAN	1355/1361	EADDIVNWLK (SEQ ID NO:34)	0.21	0.29	0.25 ± 0.05
27	P07237	PDIA1_HUMAN	1355/1361	EADDIVNWLK (SEQ ID NO:35)	0.43	0.43	0.62 ± 0.25
28	P08238	HS90B_HUMAN	2409/2415	HNDDEQYAWESSAGGSFTVR (SEQ ID NO:36)	0.98	0.37	0.67 ± 0.42
29	P08238	HS90B_HUMAN	2409/2415	HNDDEQYAWESSAGGSFTVR (SEQ ID NO:37)	0.32	0.47	0.39 ± 0.10
30	P05787	K2C8_HUMAN	1183/1189	WSLLQQQK (SEQ ID NO:38)	0.73	0.93	0.83 ± 0.14
31	P05787	K2C8_HUMAN	1183/1189	WSLLQQQK (SEQ ID NO:39)	0.92	0.87	0.89 ± 0.03
32	P52907	CAZA1_HUMAN	946/952	NFWNGR (SEQ ID NO:40)	0.94	0.88	0.91 ± 0.04
33	P15311	EZRI_HUMAN	1000/1006 1328/1334	GFPTWLK (SEQ ID NO:41) / IQVWHAEHR	0.27	0.38	0.32 ± 0.13

				(SEQ ID NO:42)			
34	P48643	TCPE_HUMAN	1891/1897	WVGGPEIELIAIATGGR (SEQ ID NO:43)	0.61	0.94	0.77 ± 0.23
35	Q9BQE3	TBA6_HUMAN	2482/2488	AFVHWYVGEGMEEGEFSEAR (SEQ ID NO:44)	0.37	0.27	0.32 ± 0.07
36	P07437	TBB2_HUMAN	1850/1856	NSSYFVEWIPNNVK (SEQ ID NO:45)	0.32	0.24	0.27 ± 0.05
37	Q9UQ80	PA2G4_HUMAN	2080/2086	LVKPGNQNTQVTEAWNK (SEQ ID NO:46)	0.20	0.19	0.19 ± 0.01
38	P55072	TERA_HUMAN	1482/1488	WALSQSNPSALR (SEQ ID NO:47)	0.90	1.12	1.01 ± 0.15
39	P25786	PSA1_HUMAN	1932/1938	NQYDNDVTVWSPQGR (SEQ ID NO:48)	0.38	0.53	0.45 ± 0.10
40	Q02790	FKBP4_HUMAN	1789/1795	VFVHYTGWLLDGTK (SEQ ID NO:49)	NA	0.21	0.21
41	P55809	SCOT_HUMAN	1518/1524	YGDLANWMIPGK (SEQ ID NO:50)	NA	2.61	2.61
42	P04181	OAT_HUMAN	2253/2259	WLAVDYENVRPDIVLLGK (SEQ ID NO:51)	1.08	0.87	0.97 ± 0.15

43	P78371	TCPB_HUMAN	1444/1450	IHPQTIIAGWR (SEQ ID NO:52)	0.38	0.45	0.41 ± 0.04
44	P30041	PRDX6_HUMAN	2049/2055	LIALSIDSVEDHLAWSK (SEQ ID NO:53)	0.29	0.52	0.41 ± 0.15

In Table 1, some protein spots possessed more than one NBS peptide pair, indicating the presence of multiple tryptophan containing peptides in these proteins. For these proteins, the final NBS ratio was calculated by averaging the ratios of all NBS paired peptides associated with the same protein spot. In total, the 44 identified NBS labelled protein spots represented 23 different protein species (Table 1), where 10 species, including K1C18_HUMAN (CK18), were represented by multiple isoforms/spots on the 2DE gel. Using NBS ratio analysis, 20 protein spots showed a >2-fold intensity decrease upon TAM treatment, 1 spot showed a >2-fold increase, and the remaining 23 spots changed less than 2-fold between TAM and vehicle.

The 44 identified NBS protein spots representing 23 different protein species. The 13C/12C ratios were calculated based on the NBS-paired peak intensities (Column 13C/12C (1)). Three proteins were identified with more than one pair of NBS peptides. For these proteins, the 13C/12C ratios were the average of all the paired peptide ratios of the same protein. To test the technical reproducibility, protein lysates from the exact same batch of TAM and vehicle treated MCF-7 cells were re-labelled with NBS, separated on 2DE gels, and re-quantified based on the 13C/12C ratios (Column 13C/12C (2)). The mean and SD values are based on combining values from the 13C/12C (1) and 13C/12C (2) columns.

Table 2 (below): Guided by the conventional 2DE gel image analysis, 88 spots were excised from the NBS-labeled 2DE gels. A total of 58 protein spots were positively identified by peptide mass fingerprinting (Figure 4 displays the NBS/2DE reference map with the identified protein spots), of which 44 contained NBS peptide pairs (Table 2). Some protein spots possessed more than one NBS peptide pair, indicating the presence of multiple tryptophan containing peptides in these proteins. For these proteins, the final NBS ratio was calculated by averaging the ratios of all NBS paired peptides associated with the same protein spot. Using NBS ratio analysis, 21/44 protein spots

showed a >2-fold intensity alteration upon TAM treatment, and the remaining 23/44 spots changed less than 2-fold between TAM and vehicle. In comparison, for the conventional 2DE analysis, 17/44 of these spots were associated with a >2 fold difference, while the remaining 27/44 exhibited a <2 fold difference. The 44 identified NBS labeled protein spots represented 23 different protein species (Table 2), where 10 species, including K1C18_HUMAN (CK18), were represented by multiple isoforms/spots on the 2DE gel.

TABLE 2

Spot No.	Acc. No.	Protein Name	Protein Description	Theoretical *Mw*	Theoretical *pI*	Peptide Coverage%	Peptide Matched	Score	Expectation
1	P05783	K1C18_HUMAN	Cytokeratin-18	47897	5.34	31	16	65	4.14E-03
2	P05783	K1C18_HUMAN	Cytokeratin-18	47897	5.34	32	17	95	4.44E-06
3	P05783	K1C18_HUMAN	Cytokeratin-18	47897	5.34	39	16	95	4.64E-06
4	P05783	K1C18_HUMAN	Cytokeratin-18	47897	5.34	48	21	140	1.31E-10
5	P10809	CH60_HUMAN	60 kDa heat shock protein	61016	5.70	32	15	96	3.36E-06
6	P10809	CH60_HUMAN	60 kDa heat shock protein	61016	5.70	24	14	68	2.03E-03
7	P10809	CH60_HUMAN	60 kDa heat shock protein	61016	5.70	37	15	78	2.07E-04
8	P10809	CH60_HUMAN	60 kDa heat shock protein	61016	5.70	16	8	64	4.75E-03
9	P04792	HSPB1_HUMAN	Heat-shock protein beta-1 (HspB1)	22768	5.98	44	12	79	1.65E-04
10	P04792	HSPB1_HUMAN	Heat-shock protein beta-1 (HspB1)	22768	5.98	44	10	55	4.33E-02
11	P04792	HSPB1_HUMAN	Heat-shock protein beta-1 (HspB1)	22768	5.98	32	9	54	5.46E-02

TABLE 2 (CONTINUED)

12	P30101	PDIA3_HUMAN	Protein disulfide-isomerase A3 precursor	56747	5.98	30	14	89	1.54E-05
13	P30101	PDIA3_HUMAN	Protein disulfide-isomerase A3 precursor	56747	5.98	25	11	73	6.12E-04
14	P30101	PDIA3_HUMAN	Protein disulfide-isomerase A3 precursor	56747	5.98	31	14	63	6.12E-03
15	P11021	GRP78_HUMAN	78 kDa glucose-regulated protein precursor	72288	5.07	33	18	102	8.26E-07
16	P11021	GRP78_HUMAN	78 kDa glucose-regulated protein precursor	72288	5.07	36	20	120	1.31E-08
17	P11021	GRP78_HUMAN	78 kDa glucose-regulated protein precursor	72288	5.07	46	26	189	1.65E-15
18	P11021	GRP78_HUMAN	78 kDa glucose-regulated protein precursor	72288	5.07	48	31	231	1.04E-19
19	P08727	K1C19_HUMAN	Cytokeratin-19	44079	5.04	50	24	146	3.29E-11
20	P08727	K1C19_HUMAN	Cytokeratin-19	44079	5.04	47	22	139	1.65E-10
21	P08727	K1C19_HUMAN	Cytokeratin-19	44079	5.04	59	25	185	4.14E-15
22	P08727	K1C19_HUMAN	Cytokeratin-19	44079	5.04	61	28	184	5.21E-15

TABLE 2 (CONTINUED)

23	P08727	K1C19_HUMAN	Cytokeratin-19	44079	5.04	69	33	271	1.04E-23
24	P30040	ERP29_HUMAN	Endoplasmic reticulum protein ERp29 precursor	28975	6.77	29	12	63	6.71E-03
25	P30040	ERP29_HUMAN	Endoplasmic reticulum protein ERp29 precursor	28975	6.77	31	11	56	3.00E-02
26	P07237	PDIA1_HUMAN	Protein disulfide-isomerase precursor	57081	4.76	33	19	98	2.03E-06
27	P07237	PDIA1_HUMAN	Protein disulfide-isomerase precursor	57081	4.76	30	15	69	1.57E-03
28	P08238	HS90B_HUMAN	Heat shock protein HSP 90-beta	83081	4.97	20	13	51	9.70E-02
29	P08238	HS90B_HUMAN	Heat shock protein HSP 90-beta	83081	4.97	25	17	62	8.45E-03
30	P05787	K2C8_HUMAN	Cytokeratin-8	53510	5.52	52	23	136	3.29E-10
31	P05787	K2C8_HUMAN	Cytokeratin-8	53510	5.52	30	16	113	6.56E-08
32	P52907	CAZA1_HUMAN	F-actin capping protein alpha-1 subunit	32902	5.45	40	12	62	8.26E-03
33	P15311	EZRI_HUMAN	Ezrin (p81)	69225	5.95	31	26	63	6.71E-03

TABLE 2 (CONTINUED)

34	P48643	TCPE_HUMAN	T-complex protein 1	59633	5.45	21	12	56	3.36E-02
35	Q9BQE3	TBA6_HUMAN	Tubulin alpha-6 chain	49863	4.96	42	15	76	3.61E-04
36	P07437	TBB2_HUMAN	Tubulin beta-2 chain	49639	4.78	36	20	118	2.07E-08
37	Q9UQ80	PA2G4_HUMAN	Proliferation-associated protein 2G4	43759	6.13	27	12	56	3.07E-02
38	P55072	TERA_HUMAN	Transitional endoplasmic reticulum ATPase	89135	5.14	33	19	101	1.04E-06
39	P25786	PSA1_HUMAN	Proteasome subunit alpha type 1	29537	6.15	22	7	50	1.44E-01
40	Q02790	FKBP4_HUMAN	FK506-binding protein 4	51641	5.35	30	11	60	1.22E-02
41	P55809	SCOT_HUMAN	Succinyl-CoA:3-ketoacid-coenzyme A transferase 1	56122	7.14	26	7	49	1.77E-01
42	P04181	OAT_HUMAN	Ornithine aminotransferase	48504	6.57	30	13	73	6.12E-04
43	P78371	TCPB_HUMAN	T-complex protein 1	57321	6.02	29	11	90	1.28E-05
44	P30041	PRDX6_HUMAN	Peroxiredoxin 6	24888	6.02	54	14	76	3.00E-04

TABLE 2 (CONTINUED)

45	P08559	ODPA_HUMAN	Pyruvate dehydrogenase E1 component alpha subunit	43268	8.35	20	11	67	2.80E-03
46	Q9BX68	HINT2_HUMAN	Histidine triad nucleotide-binding protein 2	17151	9.20	41	4	62	8.85E-03
47	P11142	HSP7C_HUMAN	Heat shock cognate 71 kDa protein	70854	5.37	28	15	62	9.06E-03
48	P06576	ATPB_HUMAN	ATP synthase beta chain	56525	5.26	47	18	120	1.31E-08
49	P08865	RSSA_HUMAN	40S ribosomal protein SA (p40)	32702	4.79	20	5	44	5.09E-01
50	P32119	PRDX2_HUMAN	Peroxiredoxin 2	21747	5.67	27	8	62	7.89E-03
51	Q9NX01	TXN4B_HUMAN	Thioredoxin-like protein 4B	17004	5.63	24	3	33	6.56E+00
52	P30048	PRDX3_HUMAN	Thioredoxin-dependent peroxide reductase	27675	7.67	30	8	56	3.07E-02
53	P05388	RLA0_HUMAN	60S acidic ribosomal protein	34252	5.71	29	5	45	4.24E-01

TABLE 2 (CONTINUED)

54	P31040	DHSA_HUMAN	Succinate dehydrogenase [ubiquinone] flavoprotein subunit	72645	7.06	12	5	41	9.93E-01
55	P54284	CACB3_HUMAN	Voltage-dependent L-type calcium channel beta-3 subunit	54498	5.93	17	9	50	1.47E-01
56	P06576	ATPB_HUMAN	ATP synthase beta chain	56525	5.26	35	12	161	1.04E-12
57	P35232	PHB_HUMAN	Prohibitin	29786	5.57	46	10	73	6.26E-04
58	P07741	APT_HUMAN	Adenine phosphoribosyltransferase	19464	5.79	35	5	69	1.65E-03

Example 10 – Validation of Differentially Expressed Proteins by Immunoblotting

Finally, to definitively validate the quantitative information obtained from NBS/2DE/MS, we conducted immunoblotting experiments for two representative proteins. Total cell lysates from independent batches of MCF-7 cells treated with TAM or vehicle for 24 hr and 48 hrs were separated on 1D gels and immunoblotted with protein specific antibodies (Figure 10). The 78 kDa glucose-regulated protein (GRP78), a multi-drug resistant protein, had four isoforms displayed in the NBS/2DE map (Spots 15-18 in Figure 4). Based on the NBS (13C /12C) ratios and 2DE quantitation results, the expression of all four isoforms were expected to increase upon 48 hr TAM treatment (Table 1 and Figure 5). Immunoblotting with anti-GRP78 confirmed the up-regulation of this protein even at 24 hr after TAM treatment and this increased expression was stable up to 48 hrs (Figure 10). Similarly, there were five CK19 isoforms identified by 2DE analysis (Spot 19 to 23 in Figure 4). Based on the conventional 2DE image analysis, all but Spot 19 exhibited a significant increase (> 2-fold) upon 48 hr of TAM treatment. However, based on the NBS/2DE/MS quantitation, all the CK19 isoforms showed a clear down-regulation upon TAM treatment (Spot 19 to 23 in Figure 5). Immunoblotting with anti-CK19 antibodies unambiguously validated the down-regulation of the protein upon TAM treatment at 48 hr (Figure 10). This data indicates that the NBS/2DE-based protein quantitative approach is likely to provide more accurate quantitative information regarding individual proteins compared to the 2DE only quantitative technology.

Comparative 2DE Analysis lines and use of PA2G4 as a diagnostic and prognostic indicator in breast cancer

Example 11 – Proteomics of PA2G4 in breast cancer cells. a) 2DE analysis of 2 breast cancer cell lines (MCF-7, HCC-38) and a CRL cell line indicated that PA2G4 was over-expressed in the ER+ cell lines but not the ER- cell lines.

TABLE 3

P2AG4_Human	**CRL (ER-)**	**HCC-38 (ER-)**	**MCF-7 (ER+)**
Stain Intensity r1	4949	6488	13437
r2	8012	8285	10918
r3	4067	5232	9609
mean+/- SE	5676+/-1195	6668+/-886	11321+/-1123
TTest compared with HCC		0.54	0.03

Comparing to the CRL cell line, the PA2G4 expression showed no significant different in HCC-38 but significantly different in MCF-7 (Table 3). This is consistent with the finding of Charpentier et al. (2000)

Example 12 - PA2G4 as a prognostic marker for breast cancer

b) The over-expression of PA2G4 in MCF-7 cells can be suppressed by tamoxifen treatment. It was found a 5x decrease using NBS analysis and a 3x decrease using 2DE analysis.

To demonstrate the relevance of our study to breast cancer and clinical outcome, we focused on PA2G4, originally identified as an ERBB3 receptor interacting protein. We performed 24 immunoblotting experiments to confirm the down-regulation of PA2G4 protein in MCF7 cells by TAM treatment. As shown in Figure 10b, exposure of MCF7 cells to TAM induced a modest but noticeable down-regulation of PA2G4 protein after 24h, which persisted until 48hr. In an independent experiment, we have also found that the expression of

PA2G4 protein is >2 fold higher in MCF-7 (ER+) than in HCC-38 (ER-) and CCD-1059sk (ER-) breast cell lines.

We then performed a series of semi-quantitative RT-PCR experiments and found that PA2G4 mRNA was clearly down-regulated by TAM exposure at both the 24h and 48h time points (Figure 10b). Given this transcriptional regulation, we then performed data-mining on two independent microarray data sets of ER+ primary tumors. The first set (GIS; Miller et al, 2005) consisted of data from 201 ER+ breast cancer patients, where 67 patients received only adjuvant hormonal therapy after surgery. The second set (Oxford; Sotiriou et al, 2006) consisted of data from 65 ER+ patients, where 33 patients received TAM only and the remaining patients did not receive treatment. By performing univariate statistical analysis, we found that PA2G4 transcript expression demonstrated a highly significant correlation with disease-free survival in both data sets (p-values are 1.75x10-5 (GIS) and 0.001 (Oxford) respectively), with patients with high PA2G4 expressing tumors exhibiting poorer clinical outcome (Table 4A).

TABLE 4

A)	**Univariate Analysis**				**B)**	**Multivariate Analysis**			
GIS (all ER+, n = 201)									
	P-value	Regression Coefficient	95% Confidence Interval for regression coefficient			P-value	Regression Coefficient	95% Confidence Interval for regression coefficient	
			Lower	Upper				Lower	Upper
PA2G4	**2E-05**	1.001	1.001	1.002	*PA2G4*	**0.001**	1.001	1.000	1.002
					size	0.015	1.033	1.006	1.061
					p53	0.916	0.960	0.452	2.040
					node	**0.004**			
					grade	0.469			
					age	0.531	1.007	0.986	1.028

GIS (Tamoxifen-treated, n = 67)

	P-value	Regression Coefficient	95% Confidence Interval for regression coefficient			P-value	Regression Coeffi-cient	95% Confidence Interval for regression coefficient	
			Lower	Upper				Lower	Upper
PA2G4	**0.002**	1.001	1.001	1.002	*PA2G4*	**0.005**	1.001	1.000	1.002
					size	0.796	1.006	0.961	1.053
					p53	0.762	1.202	0.366	3.940
					node	0.294			
					grade	0.320			
					age	**0.039**	1.043	1.002	1.086

Oxford (all ER+, n = 65)

	P-value	Regression Coefficient	95% Confidence Interval for regression coefficient			P-value	Regress-ion Coefficient	95% Confidence Interval for regression coefficient

			Lower	Upper					Lower	Upper
PA2G4	**0.001**	1.002	1.001	1.004		*PA2G4*	**0.043**	1.002	1.000	1.004
						grade	0.897	1.048	0.519	2.113
						age	0.974	0.999	0.933	1.069
						size	0.302	1.263	0.810	1.969

Oxford (Tamoxifen-treated, n = 33)

	P-value	Regression Coefficient	95% Confidence Interval for regression coefficient				Regression Coefficient	95% Confidence Interval for regression coefficient	
			Lower	Upper					Lower	Upper
PA2G4	**0.008**	1.003	1.001	1.005	*PA2G4*	0.185	1.002	0.999	1.005
					grade	0.709	1.225	0.421	3.563
					age	0.260	0.938	0.838	1.049
					node	0.790	1.418	0.108	18.537
					size	0.447	1.260	0.695	2.282

Table 4. Univariate and multivariate analysis of PA2G4 as a survival associated gene in ER+ breast cancer. The significant association ($p<0.05$) is highlighted in bold.

The significant association between PA2G4 expression and survival was also observed in both data sets when the univariate analysis was restricted to only those patients receiving adjuvant hormonal therapy (p values are 0.002 (GIS) and 0.008 (Oxford) respectively; Table 4a). We then tested if PA2G4 expression might behave as an independent prognostic factor in comparison to other clinical variables, such as tumor grade, patient age, and tumor size. In this multivariate analysis, PA2G4 expression still retained its prognostic significance in both patient cohorts (p values are 0.001 (GIS) and 0.04 (Oxford) respectively) and also in the GIS cohort of patients receiving adjuvant hormonal therapy (p = 0.005) (Table 4b). Thus, in two independent microarray data sets of ER+ breast cancers, high PA2G4 expression in primary tumors is associated with decreased disease-free survival compared to patients with low PA2G4 expressing tumors. These results show that PA2G4 is a novel survival-associated breast cancer gene and is suitable for use as a prognostic biomarker.

Example 13 – Antibodies against PA2G4

Suitable antibodies against PA2G4 protein are available from commercial sources and may be used to detect and/or quantify PA2G4 in any immunological binding/hybridization reaction known in the art. For example, a PA2G4 polyclonal antibody is available from the Abnova Corporation (Taiwan; Product Number H00005036-A01). This is a mouse polyclonal antibody raised against a partial recombinant PA2G4 immunogen. Alternatively, another example is an anti-Ebp1 polyclonal antibody raised from rabbit available from Upstate (USA; Product Number 07-397).

Example 14 – Suitable recombinant expression vectors for PA2G4

A person skilled in the art will be able to select a suitable vector and insert the PA2G4 gene sequence into it for subsequent expression of the gene. For PA2G4, vectors with the PA2G4 inserted are commercially available. For

example, Origene (USA) has the PA2G4 sequence in its pCMV6-XL5 vector as Product Number SC103017.

Example 12 – Use of PA2G4 in diagnosis

The present invention discloses the differential expression of the PA2G4 gene and protein in ER+ breast cancer cells and disease patients, and the suppression effect of tamoxifen (TAM) in its expression. Based on this finding, the invention provides a method for diagnosing or prognosticating ER+ breast cancer disease in a subject, and for determining whether a subject is at increased risk of developing breast cancer disease. Furthermore, this invention provides therapeutic and prophylactic methods for treating or preventing breast caner disease using the PA2G4 gene and its corresponding gene products.

For the diagnostic application, the anti-human PA2G4 antibody can be used to analyse (detect and measure) this protein in human samples such as serum, nipple aspirate fluid, extracellular fluids and primary tumour tissues. A level of expression of PA2G4 higher than those from reference values obtained from controls, such as patients not diagnosed with or suffering from breast cancer, indicates a possibility of breast cancer.

Example 15 – Use of PA2G4 to monitor therapy

Based on our observation of the relationship between the PA2G4 gene/protein expression and inhibitory effect of TAM on the said molecules expression, as well as the survival data study in two breast cancer patient cohorts, we belief that PA2G4 gene/protein has the potential use for monitoring therapy of breast cancer. For further study, we will use the siRNA for gene silencing to observe the physiological change of ER+ cells (e.g. MCF-7) after the transfection of siRNA. Then we can evaluate the role of PA2G4 gene in the ER+ cell survival and determine whether we can suppress the ER+ cancer cells by specifically inhibiting the PA2G4 gene. The assay can be a dose-dependant time-course response study and has particular application to serum and other body fluidic

samples. The result will be able to show the feasibility of using PA2G4 to monitor therapy in breast cancer patients.

Example 16 – Therapeutic use of PA2G4

Molecules for the treatment of a proliferative cell disorder such as breast cancer can be prepared using PA2G4 gene sequence (SEQ ID NO:1) or PA2G4 protein sequence (SEQ ID NO:2). A person skilled in the art will be able to produce nucleic acids complementary to the PA2G4 gene or mRNA such as siRNA, by any method known in the art, to inhibit or decrease the translation of the mRNA. Similarly, antibodies such as those cited in Example 10 based on the PA2G4 protein, can be raised by any method known in the art, to produce antibodies that bind to, and hence, inactivate the actions of the PA2G4 protein. Such therapeutic molecules can be mixed with suitable excipients, vehicles or carriers and administered into the diseased tissues or cells to obtained the desired therapeutic effect. In particular, the desired therapeutic effect is to lower or decrease the expression of the PA2G4 gene and/or protein.

Kit

Example 17 – Kit

The present invention also provides a kit comprising at least one biomarker complementary or hybridisable to PA2G4 gene, gene transcript, gene product, variant or fragment thereof. Such complementary or hybridisable biomarkers include nucleic acid probes recognizing and binding to the PA2G4 gene or gene products or fragments or variations thereof, and antibodies recognizing and binding to the PA2G4 gene product or fragments thereof. The gene can be wild type or mutants thereof, and the proteins can be wild type or mutants thereof as long as sufficient homology exists to allow recognition of the gene or protein by the complementary biomarker. The complementary or hybridisable biomolecules can be labeled and hence be used for prognostic evaluation

(predicting the susceptibility to), or for the diagnosis of, the cell proliferative disease, or to indicate the severity of the disease state, or to monitor the efficacy of a course of treatment.

Example 18 – Other biomarkers for breast cancer

In this study, inventors found another TAM-regulated protein, CK18 neoepitope M30, which has been shown to be a useful serum marker of apoptotic epithelium in septic patients (de Graauw et al, 2005). During apoptosis, DNA fragmentation, nuclear condensation and CK18 cleavage are known events, and a neoepitope in CK18 called M30 neoantigen, becomes available at an early caspase cleavage event. The observation of TAM mediated down-regulation of CK18 show that measurement of CK18 downregulation is useful in TAM treated patients as a measure of drug responsiveness. Notably, even though TAM is the primary hormonal therapy for breast cancer, a major adverse side effect of TAM therapy is undesirable endometrial proliferation. As such, the down-regulation of CK18 indicates a predisposition to breast cancer and CK18 is useful as a biomarker for breast cancer.

Interestingly, a previous report on TAM mediated changes in Ishikawa endometrial adenocarcinoma cell lines using 2DE showed a significant increase in CK18 expression (Shah et al. 2004), which is opposite to our observation in breast cancer cells. The current observation on selective estrogen receptor modulator (SERM) mediated down-regulation of CK18 in the MCF-7 cell line and the exactly opposite observation in endometrial cancer cell line allows CK18 to represent an important tissue-specific target of SERMs. Finally, GRP78 is another protein identified by NBS quantitation as being TAM-regulated.

GRP78 is a known drug resistance effector and is also a stress response chaperon. Lung cancer and neuroblastoma patients with a positive GRP78 expression tended to show a better prognosis and improved survival (Uramoto et al., 2005). The observed up-regulation of GRP78 in the TAM treated MCF-7

cells represent the immediate stress response to the drug treatment, and also could couple cellular stress to the cell death program (Rao et al., 2002; Ding et al., 2004). This is supported by accompanied up-regulation of another ER stress gene, HERP, at the mRNA level upon TAM treatment.

Two independently validated proteins, GRP78 and CK19, are also found to be bona-fide up and down-regulated proteins during TAM treatment. The results indicate that NBS/2DE/MS is a reliable approach for cellular protein quantitation than conventional 2DE approaches.

It will be apparent to a person skilled in the art that the present invention may also be used in veterinary medicine for animals. While specific examples to practice the invention have been provided, it will be appreciated that various modifications and improvements may be made by a person skilled in the art without departing from the spirit and scope of the present invention.

References

Altschul et al., J Mol Biol 215, 403-410 (1990)

Altschul et al., Nucleic Acids Res 25, 3389-3402 (1997)

Bradford, M. M. (1976) Anal. Biochem. 72, 248

Carell et al. Angew Chem Int Ed Engl 33, 2061(1994)

Carrell et al. Angew Chem Int Ed Engl 33, 2059 (1994)

Charpentier AH, Bednarek AK, Daniel RL, Hawkins KA, Laflin KJ, Gaddis S, MacLeod MC, Aldaz CM (2000) Effects of estrogen on global gene expression: identification of novel targets of estrogen action. Cancer Res 60: 5977-5983

Cristiano et al. J Mol Med 73, 479 (1995)

Cull et al. PNAS USA 89, 1865-1869 (1992)

Cwirla et al. PNAS USA 87, 6378-6382 (1990)

de Graauw, M.; Tijdens, I.; Cramer, R.; Corless, S.; Timms, J. F.; van de, W. B. J.Biol.Chem., 280, 29885-29898 (2005)

Ding, S. J.; Li, Y.; Tan, Y. X.; Jiang, M. R.; Tian, B.; Liu, Y. K.; Shao, X. X.; Ye, S. L.; Wu, J. R.; Zeng, R.; Wang, H. Y.; Tang, Z. Y.; Xia, Q. C. Mol Cell Proteomics, 3, 73-81 (2004)

Felici J Mol Biol 222, 301-310 (1991)

Fodor Nature 364, 555-556) (1993)

Gallop et al. J Med Chem 37,1233 (1994)

Houghten Biotechniques 13, 412-421 (1992)

Hsu, W. M.; Hsieh, F. J.; Jeng, Y. M.; Kuo, M. L.; Tsao, P. N.; Lee, H.; Lin, M. T.; Lai, H. S.; Chen, C. N.; Lai, D. M.; Chen, W. J. Int.J.Cancer, 113, 920-927 (2005)

Karlin and Altschul Proc Natl Acad Sci USA 87, 2264-2268) (1990)

Karlin and Altschul, Proc Natl Acad Sci USA 90, 5873-5877) (1993)

Kuyama, H.; Watanabe, M.; Toda, C.; Ando, E.; Tanaka, K.; Nishimura, O. Rapid Commun.Mass Spectrom. 17, 1642-1650 (2003)

Lam Nature 354, 82-84, (1991)

Lam Anticancer Drug Des 12, 145, (1997)

Laugesen, S.; Roepstorff, P. J.Am.Soc.Mass Spectrom. 14, 992-1002, (2003)

Matsuo, E.; Toda, C.; Watanabe, M.; Ojima, N.; Izumi, S.; Tanaka, K.; Tsunasawa, S.;

Nishimura, O. Proteomics, 6, 2042-2049 (2006)

Miller, L. D.; Smeds, J.; George, J.; Vega, V. B.; Vergara, L.; Ploner, A.; Pawitan, Y.; Hall, P.; Klaar, S.; Liu, E. T.; Bergh, J. Proc.Natl.Acad.Sci.U.S.A 2005, 102, 13550-13555.

Sotiriou, C.; Wirapati, P.; Loi, S.; Harris, A.; Fox, S.; Smeds, J.; Nordgren, H.; Farmer, P.; Praz, V.; Haibe-Kains, B.; Desmedt, C.; Larsimont, D.; Cardoso, F.; Peterse, H.; Nuyten, D.; Buyse, M.; Van de Vijver, M. J.; Bergh, J.; Piccart, M.; Delorenzi, M. J Natl.Cancer Inst. 2006, 98, 262-272.

Pompeia C, Hodge DR, Plass C, Wu YZ, Marquez VE, Kelley JA, Farrar WL. Cancer Res 64: 3465-3473 (2004)

Rao, R. V.; Castro-Obregon, S.; Frankowski, H.; Schuler, M.; Stoka, V.; del Rio, G.; Bredesen, D. E.; Ellerby, H. M. J.Biol.Chem. 277, 21836-21842, (2002)

Saji, S.; Kawakami, M.; Hayashi, S.; Yoshida, N.; Hirose, M.; Horiguchi, S.; Itoh, A.; Funata, N.; Schreiber, S. L.; Yoshida, M.; Toi, M. Oncogene, 24, 4531-4539 (2005)

Scott and Smith Science 249, 386-390, (1990)

Shah, Y. M.; Basrur, V.; Rowan, B. G. Mol Cell Endocrinol. 219, 127-139, (2004)

U.S. Pat. No. 5,223,409

Uramoto, H.; Sugio, K.; Oyama, T.; Nakata, S.; Ono, K.; Yoshimastu, T.; Morita, M.; Yasumoto, K. Lung Cancer 49, 55-62 (2005)

Zhang, G. J.; Kimijima, I.; Onda, M.; Kanno, M.; Sato, H.; Watanabe, T.; Tsuchiya, A.; Abe, R.; Takenoshita, S. Clin Cancer Res 5, 2971-2977 (1999)

Zuckernann et al. J Med Chem 37, 2678-85 (1994)

Claims

1. A method of detecting and/or quantitating the presence of, predisposition to, and/or severity of, a proliferative cell disorder in a subject, the method comprising:
(a) providing at least one sample from a subject;
(b) determining the expression of a PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof;
(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression indicating the presence of, or predisposition to, and/or severity of a proliferative cell disorder in the subject.

2. A method of monitoring the efficacy of a treatment for a proliferative cell disorder in a subject, the method comprising:
(a) providing at least two samples from a subject, each sample obtained at different time points;
(b) determining the expression of a PA2G4 gene, RNA, and/or mutation thereof, or PA2G4 protein, derivative, mutant and/or fragment thereof; and
(c) comparing the expression of the PA2G4 gene, RNA, and/or mutation thereof, and/or PA2G4 protein, derivative, mutant and/or fragment thereof, in the at least two samples.

3. A method of prognosticating the outcome of a cell proliferative cell disorder in a subject, the method comprising:
(a) providing at least one sample from a subject;
(b) determining the expression of a PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof;

(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression indicating the prognosis of a proliferative cell disorder in the subject.

4. The method according to any one of the claims 1 to 3, wherein the control is at least one subject not diagnosed with a cell proliferative disorder.

5. The method according to any one of claims 1 to 3, wherein the gene, RNA and/or mutation thereof is human PA2G4 gene, RNA and/or mutation thereof, and the PA2G4 protein, derivative, mutant and/or fragment is human PA2G4 protein, derivative, mutant and/or fragment thereof.

6. The method according to any of claims 1 to 5, wherein the detecting is by labelling the PA2G4 protein, derivative, mutant and/or fragment thereof, with at least one sulfenyl halide agent followed by 2-dimensional gel electrophoresis and mass spectrometry.

7. The method according to claim 6, wherein the sulfenyl halide agent is 2-nitrobenzenesulfenyl chloride (NBS).

8. A method of treating a proliferative cell disorder in a subject, the method comprising varying the expression of a PA2G4 gene, RNA, and/or mutation thereof

9. The method according to any of claims 1 to 8, wherein the PA2G4 gene has a sequence as given in SEQ ID NO:1.

10. The method according to claim 8or 9, wherein the varying is obtained by administering nucleic acid hybridisable to all or part of the PA2G4 gene, RNA and/or mutation thereof.

11. The method according to claim 10, wherein the hybridisable nucleic acid is DNA or RNA.

12. The method according to claim 11, wherein the hybridisable nucleic acid is siRNA.

13. The method according to claim 8 or 9, wherein the varying is obtained by administering a compound hybridisable to all or part of the PA2G4 gene, RNA and/or mutation thereof.

14. A method of treating a proliferative cell disorder in a subject, the method comprising varying the expression of a PA2G4 protein, derivative, mutant and/or fragment thereof.

15. The method according to claim 14, wherein the method comprises varying the expression of the PA2G4 protein having the amino acid sequence of SEQ ID NO:2.

16. The method of claim 14 or 15, the varying is obtained by administering a polypeptide binding to all or part of the amino acid sequence of the PA2G4 protein, derivative, mutant and/or fragment thereof.

17. The method according to claim 16, wherein the binding polypeptide is an antibody.

18. The method according to claim 14, wherein the varying is obtained by administering a compound hybridisable to all or part of the amino acid sequence of the PA2G4 protein.

19. The method according to any one of claims 8 to 18, wherein varying the expression comprises reducing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof.

20. A method of screening for at least one drug that causes a variation in expression of a at least one protein and/or fragment thereof, the method comprising:
(a) administering at least one drug to at least one subject;
(b) obtaining at least one protein from the subject;
(c) labelling at least one amino acid residue of the protein(s) with at least one agent; and
(d) analysing the protein(s) by mass spectrometry, thereby identifying the at least one protein and/or fragment thereof, and determining whether a variation of expression has occurred.

21. The method according to claim 20, wherein step (d) analysing comprises separating the protein(s) based on a first characteristic, optionally separating the protein(s) based on a second characteristic, ionising the protein(s), detecting and identifying the separated protein(s).

22. The method according to any one of claims 20 to 21, wherein the protein and/or fragment thereof is a PA2G4 protein and/or fragment thereof.

23. The method according to claim 22, wherein the PA2G4 protein has the sequence of SEQ ID NO:2.

24. The method according to any one of claims 20 to 23, wherein the agent is a sulfenyl halide.

25. The method according to any one of claims 20 to 24, wherein the agent is 2-nitrobenzenesulfenyl chloride (NBS).

26. The method according to any one of claims 20 to 25, wherein the amino acid residue is tryptophan.

27. The method according to any one of claims 21 to 26, wherein the separating steps are based on at least one characteristic selected from the group consisting of mass, charge, functional group and isoelectric point.

28. The method according to any one of claims 21 to 27, wherein the separating steps are steps for SDS-polyacrylamide gel electrophoresis (SDS-PAGE), two-dimensional gel electrophoresis (2DE), high performance liquid chromatography (HPLC) and/or free flow electrophoresis (FFE).

29. The method according to any one of claims 21 to 28, wherein the ionising is selected from the group consisting of electrospray ionisation (ESI), matrix assisted laser desorption/ionisation (MALDI), surface enhanced laser desorption/ionisation (SELDI), nanospray ionisation, atmospheric pressure chemical ionisation (APCI), chemical ionisation (CI), electron impact (EI), fast atom bombardment (FAB), field desorption / field ionisation (FD/FI) and thermospray ionisation (TSP).

30. The method according to any one of claims 20 to 29, wherein step (d) analysing is by tandem mass spectrometry.

31. The method according to claim 30, wherein tandem mass spectrometry uses quadrupole time-of-flight geometry.

32. A method of selecting at least one candidate for clinical trial(s), experimentation and/or diagnostic test, comprising:

(a) providing at least one sample from at least one subject;

(b) determining the expression of the PA2G4 gene and/or mutation thereof, or PA2G4 protein, derivative, mutant and/or fragment thereof;and

(c) comparing the expression of the PA2G4 gene, RNA and/or mutation thereof; and/or PA2G4 protein, derivative, mutant and/or fragment thereof, with that of at least one control, a difference in expression indicating the suitability of the subject as a candidate.

33. The method according to claim 32, wherein the control is at least one subject not diagnosed with a cell proliferative disorder.

34. wherein the control is a subject not diagnosed with a cell proliferative disorder.

35. The method according to any one of claims 1 to 34, wherein the proliferative cell disorder is a cancer.

36. The method according to any one of claims 1 to 34, wherein the proliferative cell disorder is breast cancer.

37. The method according to any one of claims 1 to 36, wherein the subject is a mammal.

38. The method according to any one of claims 1 to 36 wherein the subject is a human being.

39. A diagnostic and/or prognostic kit for the diagnosis and/or prognostic evaluation of a cell proliferative disorder in a subject, the kit comprising at least one molecule or compound hybridisable to a PA2G4 gene, RNA and/or mutation thereof; and/or binding to a PA2G4 protein, derivative, mutant and/or fragment thereof.

40. The diagnostic and/or prognostic kit according to claim 39, wherein the molecule is a siRNA hybridisable to a PA2G4 gene transcript.

41. The diagnostic and/or prognostic kit according to claim 39, wherein the molecule is a protein binding to the PA2G4 protein, derivative, mutant and/or fragment thereof.

42. The diagnostic and/or prognostic kit according to either of claims 39 or 41, wherein the molecule is an antibody binding to the PA2G4 protein, derivative, mutant and/or fragment thereof.