(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 October 2006 (12.10.2006)

PCT

(10) International Publication Number
WO 2006/105642 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *C12Q 1/14* (2006.01)
C12Q 1/48 (2006.01)

(21) International Application Number:
PCT/CA2006/000497

(22) International Filing Date: 5 April 2006 (05.04.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/668,182 5 April 2005 (05.04.2005) US

(71) Applicant *(for all designated States except US)*: **BRITISH COLUMBIA CANCER AGENCY** [CA/CA]; 675 West 10th Avenue, Vancouver, British Columbia V5Z 1L3 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MACAULAY, Calum Eric** [CA/CA]; 338 East 37th Avenue, Vancouver, British Columbia V5W 1E7 (CA). **LAM, Wan L.** [CA/CA]; #4 - 3189 Ash Street, Vancouver, British Columbia V5Z 4L1 (CA). **GARNIS, Catherine Alice** [CA/US]; #2 - 32 Irving Street, Boston, Massachusetts 02114 (US). **LAM, Stephen** [CA/CA]; 5512 Wycliffe Road, Vancouver, British Columbia V6T 2E3 (CA).

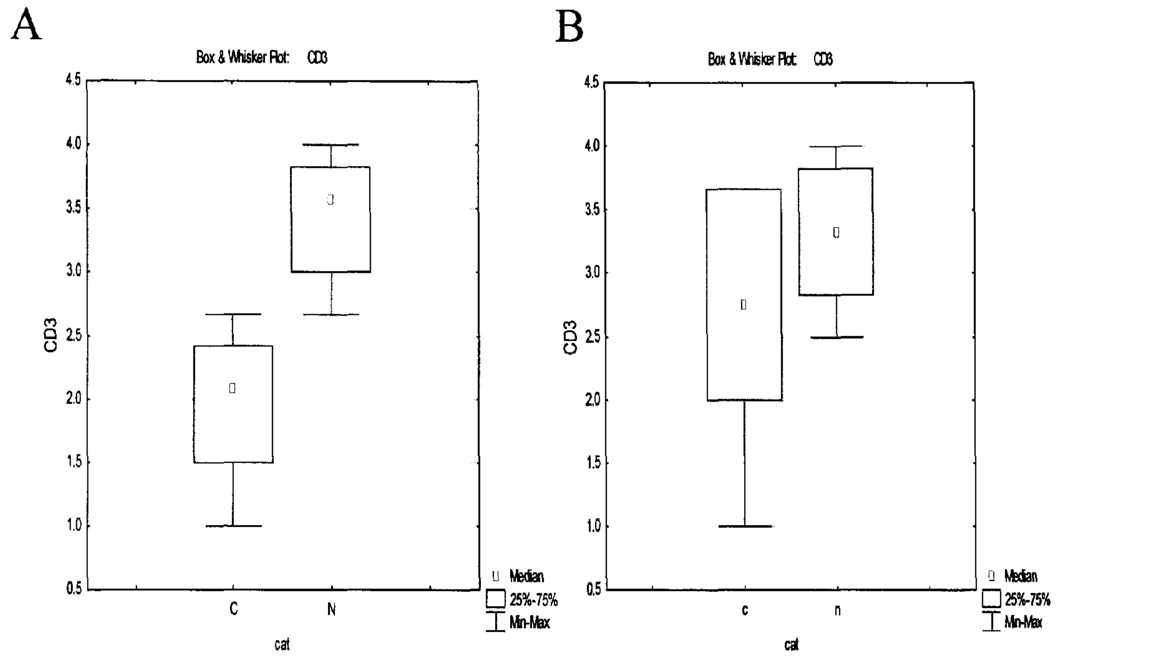
(74) Agent: **SMART & BIGGAR**; Box 11560 Vancouver Center, 2200 - 650 West Georgia Street, Vancouver, British Columbia V6B 4N8 (CA).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: BIOMARKERS FOR THE DETECTION OF LUNG CANCER AND USES THEREOF



(57) Abstract: The invention provides panels of biomarkers for the detection of lung cancer and malignancy associated changes in lung cancer. Said biomarkers find use in detecting malignancy-associated changes in a cell, for diagnosing or prognosing lung cancer in a subject by differential expression of such biomarkers, in methods of monitoring a response of a subject to lung cancer therapy, and for identifying compounds useful for lung cancer therapy.

WO 2006/105642 A1

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

BIOMARKERS FOR THE DETECTION OF LUNG CANCER AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. provisional application number 60/668,182, filed April 5, 2005, which is hereby incorporated by reference.

FIELD OF THE INVENTION

The invention is in the field of molecular diagnostics for cancer detection. More specifically, the invention provides biomarkers for the detection of lung cancer and uses thereof.

BACKGROUND OF THE INVENTION

Lung cancer is the most common cause of cancer death worldwide, with a low survival rate. Poor survival can be attributed in part to the lack of rapid cost-effective methods for early detection of the cancer[1]. Lung cancer is generally categorized into two classes, small cell lung carcinoma (SCLC), that accounts for about 20 to 30% of lung cancer patients, and non-small cell lung carcinoma (NSCLC), that accounts for approximately 70 to 80% of all lung cancers and can be further subdivided into three main histologic types, including squamous cell carcinoma, adenocarcinoma, and large cell carcinoma.

Lung cancer is often diagnosed by chest radiographs (X-rays), Computed Tomography (CT) imaging, spiral CT, Positron Emission Tomography (PET), scintigraphy, biopsy, biomarker analysis, or sputum cytology. As with any other diagnostic tests, lung cancer diagnostic tests are evaluated using the measures of sensitivity (the proportion of true positives that are correctly identified by the test) and specificity (the proportion of true negatives that are correctly identified by the test). Diagnostic tests often fail due to poor sensitivity and specificity. Chest X-rays can be capable of diagnosing NSCLC by detecting lesions or cavities formed by squamous cell carcinomas. In general, however, chest X-rays do not detect lung cancers until the cancer has metastasized and complete surgical resection is not possible. CT is used to track the spread of cancer cells, and may be more effective than a standard chest X-ray for the early detection of lung cancer. Spiral CT is a form of CT that may be more sensitive in diagnosing lung cancer at an early stage, however it has been reported to have low specificity and sensitivity with respect to detecting certain types of lung cancer. PET is a sensitive and non-invasive imaging technique that is capable of detecting lung cancers that have spread, for example into the mediastinum, as well as in the lungs. However, the costs associated with PET imaging make it relatively inaccessible for screening purposes. Scintigraphy is an imaging technique in which patients are administered radioactive agents that bind cancer cells. Biopsy involves obtaining lung tissue and cells for diagnosis, and may be performed by thoracoscopy, bronchoscopy (e.g., by bronchoalveolar lavage or BAL), or fine needle procedures. Biomarkers, such as pRb2/p130, p53, and ras have been implicated as diagnostic agents for lung cancer.

Sputum cytology is a non-invasive procedure for screening individuals at high risk for harboring or developing lung cancer. Although the specificity of sputum cytology is high, the technique exhibits low sensitivity due to the rarity of diagnostic (malignant) cells in sputum samples, compared to the amount of normal (non-malignant) cells [2]. The sensitivity of sputum cytology for detecting lung cancer is highly dependent on the number of sputum samples collected per patient and the quality of the sputum sample. Furthermore, the requirement for a skilled pathologist to analyse the sputum samples and identify changes in cellular morphology diagnostic of cancer results in increased costs for this technique and precludes it as a technique for general screening of high risk individuals.

Malignancy associated changes (MACs) are subtle sub-visual changes in histologically normal tissue and cells due to the presence of invasive or pre-invasive malignant cells in their vicinity[3,4]. Thus MACs may be present in cells adjacent to or distal from malignant cells. MACs resolve, i.e., are no longer present, upon successful removal of cancer cells, but persist in those patients with residual disease. MACs are thought to be induced in normal cells by soluble secreted factor(s) released from malignant cells [5]. MACs were first reported as qualitative observations (altered characteristics of chromatin structure) within the nuclei of cells from peripheral blood, buccal smears, sputum, bone marrow, uterus, pancreas, liver, and skin[6-10].

With the development of image cytometry devices, quantitative assessment of MACs in cell nuclei have been reported in a number of tissue types including breast, cervix, colon, thyroid and lung[11-17]. Morphological MACs have been used to detect early lung cancer using exfoliated (sputum) cells, using quantitative image cytometry and statistical analysis of cell populations[5,7,14,18]. However, quantitative morphological MACs techniques have to date resulted in less than optimal clinical diagnostic efficacy/performance characteristics.

SUMMARY OF THE INVENTION

The invention provides biomarkers for the detection of lung cancer and uses thereof.

In one aspect, the invention provides a method of detecting a malignancy associated change in a cell, the method comprising detecting differential expression of one or more nucleic acid molecules or expression products thereof in the cell, wherein said nucleic acid molecules or expression products thereof are selected from the group consisting of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), PTPRN2 (7q), e-cadherin and CD23 and wherein said differential expression is indicative of said malignancy associated change.

In some embodiments, the method comprises detecting differential expression of two or more of said nucleic acid molecules or expression products thereof in the cell or detecting differential expression of three or more of said nucleic acid molecules or expression products thereof in the cell. In some embodiments, the nucleic acid molecules or expression products thereof are selected from the group consisting of CD3, CD20cy, S100, and p63. In some embodiments, the cell is an epithelial cell or a lymphocyte. In some embodiments, the cell is a human cell. In some embodiments, the cell is derived from a subject at risk for lung cancer. In some embodiments, the cell is derived from a subject having lung cancer. In some embodiments, the expression product is a polypeptide. In some embodiments, the detecting is performed by an immunoassay. In some embodiments, the immunoassay is performed using an antibody that specifically binds said expression product. In some embodiments, the detecting is performed by a hybridisation assay. In some embodiments, the hybridization assay is performed using an probe that hybridizes to said nucleic acid molecule.

In alternative aspects, the invention provides a method of diagnosing or prognosing lung cancer in a subject comprising detecting the differential expression of one or more nucleic acid molecules or expression products thereof in a sample from the subject, wherein the nucleic acid molecules or expression products thereof are selected from the group consisting of one or more of MTGR2 [CBFA2T3], GALNS, KNS2, SOCS1, C1QTNF1, C1QTNF3, BAIAP2, RATOR [STX11], CD59, CAPN5, CD6, Kera, EMS1, PG11 [PG40, DCN, PROTEOGLYCAN II], LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), ALCAM, ADCY3, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, CASPR2, DTDST1, PCDH9, PROS1, RPL39, SNRPB, TSHR, UNC13, and PTPRN2 (7q).

In some embodiments, the method comprises detecting the differential expression of one or more nucleic acid molecules or expression products thereof in the sample, wherein the nucleic acid molecules are selected from the group consisting of one or more of PECAM1 [CD31], CD44, P100, FGFR 1.1, IL10, JAK2, SFTPA1, TP53, and UGCS [GCS]. In some embodiments, the method further comprises detecting the differential expression of one or more nucleic acid molecules or expression products thereof in the sample, wherein the nucleic acid molecules are selected from the group consisting of one or more of CD44, SFTPA1, CD3, CD20cy, S100, and p63. In some embodiments, the sample is sputum, cough, buccal mucosa or bronchoalveolar lavage. In some embodiments, the sample is substantially free of malignant cells.

In alternative aspects, the invention provides a method of diagnosing or prognosing lung cancer in a subject comprising detecting the differential expression of one or more nucleic acid molecules or expression products thereof in a sputum, cough or bronchoalveolar lavage sample from the subject, wherein the nucleic acid molecules or expression products thereof are selected from the group consisting of one or more of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), and PTPRN2 (7q). In some embodiments, the sample is substantially free of malignant cells.

In alternative aspects, the invention provides a panel of biomarkers for lung cancer or for malignancy associated changes associated with lung cancer comprising two or more nucleic acid molecules or expression products thereof selected from the group consisting of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), and PTPRN2 (7q)

In alternative aspects, the invention provides a method of monitoring the response of a subject undergoing therapy for lung cancer, the method comprising detecting differential expression of one or more nucleic acid molecules or expression products thereof in a cell derived from said subject, wherein said nucleic acid molecules or expression products thereof are selected from the group consisting of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), and PTPRN2 (7q), and wherein said differential expression is indicative of a malignancy associated change associated with said lung cancer. In some embodiments, the therapy is surgical resection of said lung cancer.

In alternative aspects, the invention provides a method of identifying a compound useful for lung cancer therapy, the method comprising detecting differential expression of one or more nucleic acid molecules or expression products thereof in a lung cancer cell in the presence or absence of said compound, wherein said nucleic acid molecules or expression products thereof are selected from the group consisting of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), and PTPRN2 (7q), and wherein said differential expression is indicative of the efficacy of said compound.

In alternative aspects, the invention provides use of the biomarkers described herein for preparation of a medicament for detecting a lung cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-H depict box and whisker plots for immunostaining intensity of bronchoalveolar lavage samples (in cytology microarray format) from subjects with (left box) or without (right box) lung cancer. **A,B**: CD3, **C,D**:CD20cy [L-26], **E,F**: S100 protein (both alpha and beta subunits) and **G,H**: p63 [4A4].

Figure 2 depicts the performance of e-cadherin and CD23 and various combinations thereof for the detection of lung cancer through immunostaining of BAL samples, which were presented and immunostained on cytology microarrays.

Figure 3 is a bar graph showing differential expression of ALCAM. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM). are shown on the y axis.

Figure 4 is a bar graph showing differential expression of ADCY3. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 5 is a bar graph showing differential expression of BAIAP2. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 6 is a bar graph showing differential expression of BCL2L1 first transcript form. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 7 is a bar graph showing differential expression of BCL2L1 second transcript form. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 8 is a bar graph showing differential expression of BDH first transcript form. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 9 is a bar graph showing differential expression of BDH second transcript form. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 10 is a bar graph showing differential expression of BOK. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 11 is a bar graph showing differential expression of C1QTNF1. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 12 is a bar graph showing differential expression of C1QTNF3. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 13 is a bar graph showing the ratio of C1QTNF1 expression / C1QTNF3 expression. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 14 is a bar graph showing differential expression of CAPN5. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 15 is a bar graph showing differential expression of CASPR2. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 16 is a bar graph showing differential expression of CD44. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 17 is a bar graph showing differential expression of CD59. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 18 is a bar graph showing differential expression of DTDST. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 19 is a bar graph showing differential expression of EMS1. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 20 is a bar graph showing differential expression of FGFR1. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 21 is a bar graph showing differential expression of GALNS. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 22 is a bar graph showing differential expression of IL10. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 23 is a bar graph showing differential expression of JAK2. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 24 is a bar graph showing differential expression of KNS2. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 25 is a bar graph showing differential expression of MTGR2. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 26 is a bar graph showing differential expression of UNC13. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 27 is a bar graph showing differential expression of p100. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 28 is a bar graph showing differential expression of PCDH9. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 29 is a bar graph showing differential expression of CD31. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 30 is a bar graph showing differential expression of PG11. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 31 is a bar graph showing differential expression of PROS1. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 32 is a bar graph showing differential expression of RPL39L. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 33 is a bar graph showing differential expression of SFTPA1. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 34 is a bar graph showing differential expression for SNRPB. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 35 is a bar graph showing differential expression of TP53. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 36 is a bar graph showing differential expression of TSHR. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

Figure 37 is a bar graph showing differential expression of UGCS. The SAGE libraries listed in Table 2 are shown on the x axis, while SAGE Tag counts, presented as tags per million (TPM), are shown on the y axis.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides, in part, biomarkers, e.g., nucleic acid molecules and expression products thereof, that are differentially expressed in histologically normal cells derived from subjects having a lung cancer and/or in malignant lung cancer cells, compared to normal cells derived from high risk subjects without cancer (e.g., heavy smokers) or normal cells derived from subjects who have undergone successful resection of their lung cancer. The biomarkers may be used individually or collectively. In some embodiments, the invention provides a rapid multifactorial assay for early detection of lung cancer.

In some embodiments, the biomarkers may be applied to histologically normal cells and tissues or body fluids readily accessible by non-invasive procedures and do not depend on the presence of malignant cells (often in very low abundance depending on the lung cancer and/or stage of the disease) in the sample being assayed. Assays including biomarkers according to the invention can be applied as a screening assay to detect presence of lung cancer in individuals at high risk to develop this disease. Since the assays can be applied to non-cancerous tissue/cells they are also useful in detecting the presence of residual disease after apparently curative therapy.

Cancers and Malignancy Associated Changes

By a "cancer" or "neoplasm" is meant any unwanted growth of cells serving no physiological function. In general, a cell of a neoplasm has been released from its normal cell division control, i.e., is a cell whose growth is not regulated by the ordinary biochemical and physical influences in the cellular environment. In most cases, a neoplastic cell proliferates to form a clone of cells which can be malignant. The term cancer includes cell growths that are technically benign but which carry the risk of becoming malignant. By "malignancy" is meant an abnormal growth of any cell type or tissue. The term malignancy includes cell growths that are pre-malignant. This term also includes any cancer, carcinoma, neoplasm, neoplasia, or tumor.

Most cancers fall within three broad histological classifications: carcinomas, which are the predominant cancers and are cancers of epithelial cells or cells covering the external or internal surfaces of organs, glands, or other body structures (e.g., skin, uterus, lung, breast, prostate, stomach, bowel), and which tend to mestastasize; sarcomas, which are derived from connective or supportive tissue (e.g., bone, cartilage, tendons, ligaments, fat, muscle); and hematologic tumors, which are derived from bone marrow and lymphatic tissue. Carcinomas may be adenocarcinomas (which generally develop in organs or glands capable of secretion, such as breast, lung, colon, prostate or bladder) or may be squamous cell carcinomas (which originate in the squamous epithelium and generally develop in most areas of the body).

Cancers may also be named based on the organ in which they originate i.e., the "primary site," for example, cancer of the breast, brain, lung, liver, skin, prostate, testicle, bladder, colon and rectum, cervix, uterus, etc. This naming persists even if the cancer metastasizes to another part of the body, that is different from the primary site. Cancers named based on primary site may be correlated with histological classifications. For example, lung cancer or bronchogenic carcinoma of the lung generally arises in epithelial cells in the lung, and is generally categorized into "small cell carcinoma" or "SCC" and "non-small cell lung carcinoma" or "NSCLC." NSCLC includes adenocarcinoma, squamous cell carcinoma, and large cell carcinoma.

"Malignancy associated changes" or "MACs" are subtle sub-visual changes in histologically normal tissue and cells due to the presence of invasive or pre-invasive malignant cells in their vicinity. Thus MACs may be present in cells adjacent to or distal from malignant cells. MACs may be identified morphologically, using image cytometry devices, such as those described in U.S. Patent No. 6,026,174, issued to Palcic et al. on February 15, 2000; U.S. Patent No. 6,493,460, issued to MacAulay et al. on December 10, 2002; U.S. Publication No. 2004/0042646, published March 4, 2004 by MacAulay et al., or using any other means known in the art. MACs associated with lung cancer may also be identified using biomarkers as described herein.

Biomarkers

The invention provides biomarkers, e.g., nucleic acid molecules and expression products thereof, that are differentially expressed in histologically normal cells derived from subjects having a lung cancer and/or in malignant lung cancer cells, compared to normal cells derived from high risk subjects without cancer (e.g., heavy smokers) or normal cells derived from subjects who have undergone successful resection of their lung cancer.

A "biomarker" is a molecular indicator of a specific biological property and as used herein is a nucleic acid molecule (e.g., a gene or gene fragment) or an expression product thereof (e.g., a polypeptide or peptide fragment or variant thereof) whose differential expression (presence, absence, over-expression or under-expression relative to a reference) within a cell or tissue indicates the presence or absence of a lung cancer. An "expression product" as used herein is a transcribed sense or antisense RNA molecule (e.g., an mRNA), or a translated polypeptide corresponding to or derived from a polynucleotide sequence. In some embodiments, an expression product can refer to an amplification product (amplicon) or cDNA corresponding to the RNA expression product transcribed from the polynucleotide sequence. A "panel" of biomarkers is a selection of two or more combinations of biomarkers.

In some embodiments, a biomarker is a nucleic acid molecule or an expression product thereof that is differentially expressed in a malignant lung cancer cell and/or in a normal cell derived from a subject having a lung cancer (i.e., a cell having a malignancy associated change), as compared to a reference or normal cell e.g., a cell derived from a high risk subject without lung cancer (e.g., heavy smokers) or a normal cell derived from a subject who has undergone successful resection of lung cancer. In some embodiments, the control or reference cell may be a cell derived from a healthy subject who is not considered to be at high risk for lung cancer e.g., a non-smoker. In some embodiments, a biomarker according to the invention is a nucleic acid molecule or an expression product thereof that is differentially expressed in a malignant lung cancer cell compared to a normal cell derived from a subject having a lung cancer (i.e., a cell having a malignancy associated change).

By "differential expression" or "differentially expressed" is meant a difference in the frequency or quantity, or both, of a biomarker in a cell or tissue or sample derived from a subject having a lung cancer compared to a reference cell or tissue or sample, e.g., in a malignant lung cancer cell and/or in a normal cell derived from a subject having a lung cancer (i.e., a cell having a malignancy associated change) compared to a reference or normal cell e.g., a cell derived from a high risk subject without cancer or with undetectable cancer (e.g., heavy smokers) or a normal cell derived from a subject who has undergone successful resection of lung cancer. In some embodiments, the control or reference cell may be a cell derived from a healthy subject who is not considered to be at high risk for lung cancer e.g., a non-smoker. In some embodiments, differential expression refers to a difference in the frequency or quantity, or both, of a biomarker in a malignant lung cancer cell compared to a normal cell derived from a subject having a lung cancer (i.e., a cell having a malignancy associated change). For example, differential expression of a biomarker can refer to an elevated level or at a decreased level of expression of the biomarker in samples of lung cancer patients compared to samples of reference subjects. Alternatively or additionally, differential expression of a biomarker can refer to detection at a higher frequency or at a lower frequency of the biomarker in samples of lung cancer patients compared to samples of reference subjects. A biomarker can be differentially present in terms of quantity, frequency or both. In some embodiments, differential expression of the biomarkers of the invention may be measured at different time points, e.g., before and after therapy. By "level of expression" is meant the level of mRNA, as well as pre-mRNA nascent transcript(s), transcript processing intermediates, mature mRNA(s), and degradation products, encoded by a gene in the cell, and/or the level of protein, protein fragments, and degradation products in a cell.

The difference in quantity or frequency or both of a biomarker may be measured by any suitable technique, such as a statistical technique. For example, a biomarker can be differentially expressed between a lung cancer sample and a reference sample, if the frequency of detecting the biomarker in a lung cancer sample is significantly higher or lower than in the reference sample, as measured by standard statistical analyses such as student's t-test, where $p<0.05$ is generally considered statistically significant. In some embodiments, a biomarker is differentially expressed if it is detected at least about 20, 30, 40, 50, 60, 70, 80, 90, 100 % or more or 2-, 5-, 10 or more fold more or less frequently in a lung cancer or MAC positive sample compared to a reference sample (MAC negative). Alternatively or additionally, a biomarker is differentially expressed if the amount of the biomarker in a lung cancer or MAC positive sample is statistically significantly different, e.g., by at least 20, 30, 40, 50, 60, 70, 80, 90, 100% or more or 2-, 5-, 10 or more fold when compared to the amount of the biomarker in a reference sample or if it is detectable in one sample and not detectable in the other. In some embodiments, differential expression may refer to an increase or decrease in expression of at least 20, 30, 40, 50, 60, 70, 80, 90, 100 % or more or 2-, 5-, 10 or more fold, in a test sample relative to a reference sample.

A "sample" can be any organ, tissue, cell, or cell extract isolated from a subject, such as a sample isolated from a mammal having a lung cancer or at risk for a lung cancer (e.g., based on family history or personal history, such a heavy smoking). For example, a sample can include, without limitation, cells or tissue (e.g., from a biopsy or autopsy) solid lung tumours, sputum, cough, bronchoalveolar lavage, bronchial brushings, buccal mucosa, peripheral blood, whole blood, red cell concentrates, platelet concentrates, leukocyte concentrates, blood cell proteins, blood plasma, platelet-rich plasma, a plasma concentrate, a precipitate from any fractionation of the plasma, a supernatant from any fractionation of the plasma, blood plasma protein fractions, purified or partially purified blood proteins or other components, serum, tissue or fine needle biopsy samples, and pleural fluid, etc. isolated from a mammal with a lung cancer, or any other specimen, or any extract thereof, obtained from a patient (human or animal), test subject, healthy volunteer, or experimental animal. A subject can be a human, rat, mouse, non-human primate, etc. A sample may also include sections of tissues such as frozen sections taken for histological purposes. In some embodiments, it may be desirable to separate lung cancer cells from non-lung cancer cells in a sample. In some embodiments, it may be desirable to not separate lung cancer cells from non-lung cancer cells in a sample. In some embodiments, a sample may be substantially free of malignants cells, for example, contain less than 1, 5, 10, 20, 30, 40, 50% malignant cells compared to non-malignant cells in the sample. In some embodiments, the non-malignant cells may be epithelial cells or lymphocytes. In some embodiments, the sample (e.g., a BAL sample) may be substantially free of epithelial cells, for example contain less than 1, 5, 10, 20, 30, 40, 50% epithelial cells. In some embodiments, the sample (e.g., a BAL sample) may contain substantial amounts of non-epithelial cells, e.g., at least 20, 30, 40, 50, 60, 70, 80, 90, 100% alveolar macrophages and lymphocytes.

A sample may also include, without limitation, products produced in cell culture by normal or transformed cells (e.g., via recombinant DNA or monoclonal antibody technology). A "sample" may also be a cell or cell line created under experimental conditions, that is not directly isolated from a subject. A sample can also be cell-free, artificially derived or synthesised. A sample may be from a cell or tissue known to be cancerous, suspected of being cancerous, or believed not be cancerous (e.g., normal or control).

A "control" or "reference" includes a sample obtained for use in determining base-line expression or activity. Accordingly, a control sample may be obtained by a number of means including from non-cancerous cells or tissue e.g., from cells surrounding a tumor or cancerous cells of a subject; from subjects not having a cancer; from subjects not suspected of being at risk for a cancer; or from cells or cell lines derived from such subjects. A control also includes a previously established standard. Accordingly, any test or assay conducted according to the invention may be compared with the established standard and it may not be necessary to obtain a control sample for comparison each time.

Biomarkers for malignancy associated changes and/or for lung cancer, according to the invention, include CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), PTPRN2 (7q), e-cadherin or CD23. These biomarkers may be used individually or collectively. For example, one or more, e.g., 2, 3, 4, 5, 6, 7, etc. of the biomarkers, up to all of the biomarkers, may be used together in any combination in an assay. In some embodiments, one or more of the biomarkers may be specifically excluded from an assay. In some embodiments, when a panel comprises only two biomarkers, various combinations of two of CD3, CD20cy, S100, p63, CD44, or SFTPA1 may be specifically excluded. In some embodiments, one or more of the biomarkers may be used in combination with an existing biomarker for lung cancer, such as pRb2/p130, p53, and/or ras.

Biomarkers according to the invention include substantially identical homologues and variants of the nucleic acid molecules and expression products thereof described herein, for example, a molecule that includes nucleotide sequences encoding polypeptides functionally equivalent to the biomarkers of the invention, e.g, sequences having one or more nucleotide substitutions, additions, or deletions, such as allelic variants or splice variants or species variants or molecules differing from the nucleic acid molecules and polypeptides referred to in the Tables herein due to the degeneracy of the genetic code.

Species variants are nucleic acid sequences that vary from one species to another, although the resulting polypeptides generally will have significant amino acid identity and functional similarity relative to each other. A polymorphic variant (e.g., a single nucleotide polymorphism or SNP) is a variation in the nucleic acid sequence of a particular gene between individuals of a given species.

A "substantially identical" sequence is an amino acid or nucleotide sequence that differs from a reference sequence only by one or more conservative substitutions, as discussed herein, or by one or more non-conservative substitutions, deletions, or insertions located at positions of the sequence that do not destroy the biological function of the amino acid or nucleic acid molecule. Such a sequence can be any integer from 10% to 99%, or more generally at least 10%, 20%, 30%, 40%, 50, 55% or 60%, or at least 65%, 75%, 80%, 85%, 90%, or 95%, or as much as 96%, 97%, 98%, or 99% identical when optimally aligned at the amino acid or nucleotide level to the sequence used for comparison using, for example, the Align Program (Myers and Miller, CABIOS, 1989, 4:11-17) or FASTA. For polypeptides, the length of comparison sequences may be at least 2, 5, 10, or 15 amino acids, or at least 20, 25, or 30 amino acids. In alternate embodiments, the length of comparison sequences may be at least 35, 40, or 50 amino acids, or over 60, 80, or 100 amino acids. For nucleic acid molecules, the length of comparison sequences may be at least 5, 10, 15, 20, or 25 nucleotides, or at least 30, 40, or 50 nucleotides. In alternate embodiments, the length of comparison sequences may be at least 60, 70, 80, or 90 nucleotides, or over 100, 200, or 500 nucleotides. Sequence identity can be readily measured using publicly available sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, or BLAST software available from the National Library of Medicine, or as described herein). Examples of useful software include the programs Pile-up and PrettyBox. Such software matches similar sequences by assigning degrees of homology to various substitutions, deletions, substitutions, and other modifications.

Alternatively, or additionally, two nucleic acid sequences may be "substantially identical" if they hybridize under high stringency conditions. In some embodiments, high stringency conditions are, for example, conditions that allow hybridization comparable with the hybridization that occurs using a DNA probe of at least 500 nucleotides in length, in a buffer containing 0.5 M $NaHPO_4$, pH 7.2, 7% SDS, 1 mM EDTA, and 1% BSA (fraction V), at a temperature of 65°C, or a buffer containing 48% formamide, 4.8x SSC, 0.2 M Tris-Cl, pH 7.6, 1x Denhardt's solution, 10% dextran sulfate, and 0.1% SDS, at a temperature of 42°C. (These are typical conditions for high stringency northern or Southern hybridizations.) Hybridizations may be carried out over a period of about 20 to 30 minutes, or about 2 to 6 hours, or about 10 to 15 hours, or over 24 hours or more. High stringency hybridization is also relied upon for the success of numerous techniques routinely performed by molecular biologists, such as high stringency PCR, DNA sequencing, single strand conformational polymorphism analysis, and in situ hybridization. In contrast to northern and Southern hybridizations, these techniques are usually performed with relatively short probes (e.g., usually about 16 nucleotides or longer for PCR or sequencing and about 40 nucleotides or longer for in situ hybridization). The high stringency conditions used in these techniques are well known to those skilled in the art of molecular biology, and examples of them can be found, for example, in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y., 1998, which is hereby incorporated by reference.

Preparation of Reagents Using Biomarkers

The biomarkers described herein may be used to prepare oligonucleotide probes and antibodies that hybridize to or specifically bind the biomarkers listed in the Tables herein, and homologues and variants thereof.

Antibodies

An "antibody" includes molecules having antigen binding regions, such as whole antibodies of any isotype (IgG, IgA, IgM, IgE, etc.) and fragments thereof. Antibody fragments include Fab', Fab, F(ab')2, single domain antibodies, Fv, scFv, etc. Antibodies may be prepared using standard techniques of preparation as, for example, described in Harlow and Lane (Harlow and Lane Antibodies; A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1988), or known to those skilled in the art. For example, a coding sequence for a polypeptide biomarker of the invention may be purified to the degree necessary for immunization of rabbits. To attempt to minimize the potential problems of low affinity or specificity of antisera, two or three polypeptide constructs may be generated for each protein, and each construct may be injected into at least two rabbits. Antisera may be raised by injections in a series, preferably including at least three booster injections. Primary immunizations may be carried out with Freund's complete adjuvant and subsequent immunizations with Freund's incomplete adjuvant. Antibody titres may be monitored by Western blot and immunoprecipitation analyses using the purified protein. Immune sera may be affinity purified using CNBr-Sepharose-coupled protein. Antiserum specificity may be determined using a panel of unrelated proteins. Antibody fragments may be prepared recombinantly or by proteolytic cleavage.

Peptides corresponding to relatively unique immunogenic regions of a polypeptide biomarker of the invention may be generated and coupled to keyhole limpet hemocyanin (KLH) through an introduced C-terminal lysine. Antiserum to each of these peptides may be affinity purified on peptides conjugated to BSA, and specificity tested in ELISA and Western blots using peptide conjugates and by Western blot and immunoprecipitation.

Monoclonal antibodies which specifically bind any one of the polypeptide biomarkers of the invention are prepared according to standard hybridoma technology (see, e.g., Kohler et al., Nature 256:495, 1975; Kohler et al., Eur. J. Immunol. 6:511, 1976; Kohler et al., Eur. J. Immunol. 6:292, 1976; Hammerling et al., In Monoclonal Antibodies and T Cell Hybridomas, Elsevier, N.Y., 1981). Alternatively monoclonal antibodies may be prepared using the polypeptides of the invention and a phage display library (Vaughan et al., Nature Biotech 14:309-314, 1996). Once produced., monoclonal antibodies may also be tested for specific recognition by Western blot or immunoprecipitation.

In some embodiments, antibodies may be produced using polypeptide fragments that appear likely to be immunogenic, by criteria such as high frequency of charged residues. Antibodies can be tailored to minimise adverse host immune response by, for example, using chimeric antibodies contain an antigen binding domain from one species and the Fc portion from another species, or by using antibodies made from hybridomas of the appropriate species.

Antibodies which specifically bind the biomarkers of the invention are considered to be useful. An antibody "specifically binds" an antigen when it recognises and binds the antigen, for example, a biomarker as described herein, but does not substantially recognise and bind other molecules in a sample. Such an antibody has, for example, an affinity for the antigen which is at least 2, 5, 10, 100, 1000 or 10000 times greater than the affinity of the antibody for another reference molecule in a sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular biomarker. For example, a polyclonal antibody raised to a biomarker from a specific species such as rat, mouse, or human may be selected for only those polyclonal antibodies that are specifically immunoreactive with the biomarker and not with other proteins, except for polymorphic variants and alleles of the biomarker. In some embodiments, a polyclonal antibody raised to a biomarker from a specific species such as rat, mouse, or human may be selected for only those polyclonal antibodies that are specifically immunoreactive with the biomarker from that species and not with other proteins, including polymorphic variants and alleles of the biomarker.

Antibodies that specifically bind any of the biomarkers described herein may be employed in an immunoassay by contacting a sample with the antibody and detecting the presence of a complex of the antibody bound to the biomarker in the sample. The antibodies used in an immunoassay may be produced as described herein or known in the art, or may be commercially available from suppliers, such as Dako Canada, Inc., Mississauga, ON. The antibody may be fixed to a solid substrate (e.g., nylon, glass, ceramic, plastic, etc.) before being contacted with the sample, to facilitate subsequent assay procedures. The antibody-biomarker complex may be visualized or detected using a variety of standard procedures, such as detection of radioactivity, fluorescence, luminescence, chemiluminescence, absorbance, or by microscopy, imaging, etc. Immunoassays include immunohistochemistry, enzyme-linked immunosorbent assay (ELISA), western blotting, and other methods known to those of skill in the art. Immunoassays can be used to determine presence or absence of a biomarker in a sample as well as the amount of a biomarker in a sample. The amount of an antibody-biomarker complex can be determined by comparison to a reference or standard, such as a polypeptide known to be present in the sample. The amount of an antibody-biomarker complex can also be determined by comparison to a reference or standard, such as the amount of the biomarker in a reference or control sample. Accordingly, the amount of a biomarker in a sample need not be quantified in absolute terms, but may be measured in relative terms with respect to a reference or control.

Probes and Primers

A "probe" or "primer" is a single-stranded DNA or RNA molecule of defined sequence that can base pair to a second DNA or RNA molecule that contains a complementary sequence (the target). The stability of the resulting hybrid molecule depends upon the extent of the base pairing that occurs, and is affected by parameters such as the degree of complementarity between the probe and target molecule, and the degree of stringency of the hybridization conditions. The degree of hybridization stringency is affected by parameters such as the temperature, salt concentration, and concentration of organic molecules, such as formamide, and is determined by methods that are known to those skilled in the art. Probes or primers specific for the nucleic acid biomarkers described herein, or portions thereof, may vary in length by any integer from at least 8 nucleotides to over 500 nucleotides, including any value in between, depending on the purpose for which, and conditions under which, the probe or primer is used. For example, a probe or primer may be 8, 10, 15, 20, or 25 nucleotides in length, or may be at least 30, 40, 50, or 60 nucleotides in length, or may be over 100, 200, 500, or 1000 nucleotides in length. Probes or primers specific for the nucleic acid biomarkers described herein may have greater than 20-30% sequence identity, or at least 55-75% sequence identity, or at least 75-85% sequence identity, or at least 85-99% sequence identity, or 100% sequence identity to the nucleic acid biomarkers described herein.

Probes or primers may be derived from genomic DNA or cDNA, for example, by amplification, or from cloned DNA segments, and may contain either genomic DNA or cDNA sequences representing all or a portion of a single gene from a single individual. A probe may have a unique sequence (e.g., 100% identity to a nucleic acid biomarker) and/or have a known sequence. Probes or primers may be chemically synthesized. A probe or primer may hybridize to a nucleic acid biomarker under high stringency conditions as described herein.

Probes or primers can be detectably-labeled, either radioactively or non-radioactively, by methods that are known to those skilled in the art. Probes or primers can be used for MACs or lung cancer detection methods involving nucleic acid hybridization, such as nucleic acid sequencing, nucleic acid amplification by the polymerase chain reaction (e.g., RT-PCR), single stranded conformational polymorphism (SSCP) analysis, restriction fragment polymorphism (RFLP) analysis, Southern hybridization, northern hybridization, in situ hybridization, electrophoretic mobility shift assay (EMSA), fluorescent in situ hybridization (FISH), and other methods that are known to those skilled in the art.

By "detectably labelled" is meant any means for marking and identifying the presence of a molecule, e.g., an oligonucleotide probe or primer, a gene or fragment thereof, or a cDNA molecule. Methods for detectably-labelling a molecule are well known in the art and include, without limitation, radioactive labelling (e.g., with an isotope such as ^{32}P or ^{35}S) and nonradioactive labelling such as, enzymatic labelling (for example, using horseradish peroxidase or alkaline phosphatase), chemiluminescent labeling, fluorescent labeling (for example, using fluorescein), bioluminescent labeling, or antibody detection of a ligand attached to the probe. Also included in this definition is a molecule that is detectably labeled by an indirect means, for example, a molecule that is bound with a first moiety (such as biotin) that is, in turn, bound to a second moiety that may be observed or assayed (such as fluorescein-labeled streptavidin). Labels also include digoxigenin, luciferases, and aequorin.

Arrays and Kits

Antibodies, probes, primers and other reagents prepared using the biomarkers of the invention may be used to prepare arrays for use in detecting of MACs and/or lung cancer. By "array" or "matrix" is meant refer to a pattern or arrangement of addressable locations or "addresses," each representing an independent site, on a surface. Arrays generally require a solid support (for example, nylon, glass, ceramic, plastic, etc.) to which the nucleic acid molecules, polypeptides, antibodies, tissue etc. are attached in a specified dimensional arrangement, such that the pattern of hybridization to a probe is easily determinable.

Generally, a probe (e.g., an antibody, nucleic acid probe or primer, polypeptide, etc.) is immobilized on an array surface and contacted with a sample containing a target binding partner (i.e., in the case of an antibody, a polypeptide that specifically binds the antibody, or in the case of a probe, a nucleic acid molecule that hybridizes to the probe) under conditions suitable for binding. If desired, unbound material in the sample may be removed. The bound target is detected and the binding results are analysed using appropriate statistical or other methods. The probe or the target may be detectably labeled for ease of detection and subsequent analysis. Multiple probes corresponding to the biomarkers described herein may be used. The multiple probes may correspond to one or more of the biomarkers described herein. In addition to probes capable of binding the biomarkers described herein, the arrays may control and reference nucleic acid molecules, polypeptides, or antibodies, to allow for normalization of results from one experiment to another and the comparison of multiple experiments on a quantitative level. Accordingly, the invention provides biological assays using nucleic acid, polypeptide, antibody, or cytology arrays.

The invention also provides kits for detecting MACs and/or lung cancer. The kits may include one or more reagents corresponding to the biomarkers described herein, e.g, antibodies that specifically bind the biomarkers, nucleic acid probes or primers that hybridize to the biomarkers, etc.. In some embodiments, the kits may include a plurality of reagents, e.g., on an array, corresponding to the biomarkers described herein. The kits may include detection reagents, e.g., reagents that are detectably labelled. The kits may include written instructions for use of the kit in detection of MACs and/or lung cancer, and may include other reagents and information such as control or reference standards, wash solutions, analysis software, etc.

Diagnostic and Other Methods

Lung cancers or MACs may be diagnosed by detecting the differential expression of one or more of the biomarkers described herein, by immunoassay, such as immunohistochemistry, ELISA, western blotting, or any other method known to those of skill in the diagnostic arts. The detecting may be carried out in vitro or in vivo.

"Diagnosis" as used herein means identifying a lung cancer or a malignancy associated change associated with the presence of a lung cancer. Lung cancer diagnostic methods are evaluated using the measures of "sensitivity" and "specificity". "Sensitivity" refers to the proportion of true positives that are correctly identified by a diagnostic test (i.e., the proportion of subjects with lung cancer that test positive), while "specificity" refers to the proportion of true negatives that are correctly identified by the test (i.e., the proportion of subjects without lung cancer that test negative). Sensitivity therefore measures a test's ability to make a correct diagnosis of the disease in the subject being tested, while specificity measures a test's ability to correctly identify subject who do not have the disease being tested for. In some embodiments, the sensitivity of diagnostic tests conducted using the biomarkers described herein may be at least 50%, or at least 55%, 60%, 65% or 70%, or at least 75%, 80%, 85%, 95%, or 99%. In some embodiments, the sensitivity of diagnostic tests conducted using the biomarkers described herein may range from any value between 50% to 100%. In some embodiments, the specificity of diagnostic tests conducted using the biomarkers described herein may be at least 50%, or at least 55%, 60%, 65% or 70%, or at least 75%, 80%, 85%, 95%, or 99%. In some embodiments, the specificity of diagnostic tests conducted using the biomarkers described herein may range from any value between 50% to 100%. In some embodiments, diagnosis of a lung cancer includes making a prognosis.

While individual biomarkers are useful diagnostics, multiple biomarkers can be measured, as the use of multiple biomarkers can increase the predictive value of a diagnostic test, by increasing the sensitivity and/or specificity of the diagnostic test. Variation in differential expression across multiple biomarkers in different samples can diagnose or predict the presence or absence of lung cancer, the response to a particular therapy for lung cancer, or better assess the risk for developing a lung cancer. For example, a combination of two or more, or three or more, or all, of CD3, CD20cy, S100, and p63 can be used to detect the presence of a lung cancer or MACs in a sample, such as sputum, cough, BAL, or buccal mucosa. Suitable statistical methods and algorithms, e.g., logistical regression algorithm, may be used to analyse and use multiple biomarkers for diagnostic, prognostic, or other purposes. The biomarkers (or specific combination of the biomarkers) can be detected and measured multiple times, for example, before, during and after a therapy for lung cancer.

Detection of the biomarkers described herein may be performed as an initial screen for the detection of early stages of lung cancer and/or may be used in conjunction with conventional methods of lung cancer diagnosis, such as sputum cytology, chest X-ray, CT scans, spiral CT, PET, scintigraphy, biopsy, traditional morphological MACs analysis, etc. Detection of the biomarkers described herein may also be performed in conjunction with previously recognized biomarkers for lung cancer, such as pRb2/p130, p53, and/or ras. Detection of the biomarkers described herein may be performed as part of a routine examination, for example, of heavy smokers over a certain age (e.g., over 60), or may be performed to determine baseline levels of the biomarkers in subjects at risk for lung cancer (e.g., heavy smokers).

Detection of the biomarkers described herein may enable a medical practitioner to determine the appropriate course of action for a subject (e.g, further testing, surgery, no action, etc.) based on the diagnosis. Detection of the biomarkers described herein may also help determine the presence or absence of lung cancer, early diagnosis of lung cancer, prognosis for lung cancer, efficacy of a therapy for lung cancer, monitoring a lung cancer therapy in a subject, or detecting relapse of lung cancer in a subject who has undergone therapy for lung cancer and is in remission.

In alternative aspects, the biomarkers and reagents prepared using the biomarkers may be used to identify lung cancer therapeutics. The kits and arrays can be used to measure biomarkers according to the invention, to diagnose a lung cancer or MACs. The kits can also be used to monitor a subject's response to a lung cancer therapy, enabling the medical practitioner to modify the treatment based upon the results of the test. The kits can also be used to identify and validate lung cancer therapeutics, such as small molecules, peptides, etc.

Various alternative embodiments and examples of the invention are described herein. These embodiments and examples are illustrative and should not be construed as limiting the scope of the invention.

EXAMPLE

A comparison of the genomic changes and gene expression patterns in lung cancer was undertaken as follows.

Genomic changes were identified using a whole genome high resolution copy number alteration array based platform (SMRT array) based upon the tiling of 27K to 32K BAC derived DNA fragments which cover 99% of the sequenced genome (Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, Ling V, MacAulay C, Lam WL. (2004). A tiling resolution DNA microarray with complete coverage of the human genome. Nat Genet. Mar;36(3):299-303; Krzywinski M, Bosdet I, Smailus D, Chiu R, Mathewson C, Wye N, Barber S, Brown-John M, Chan S, Chand S, Cloutier A, Girn N, Lee D, Masson A, Mayo M, Olson T, Pandoh P, Prabhu AL, Schoenmakers E, Tsai M, Albertson D, Lam W, Choy CO, Osoegawa K, Zhao S, de Jong PJ, Schein J, Jones S, Marra MA. A set of BAC clones spanning the human genome. Nucleic Acids Res. 2004 Jul 9;32(12):3651-60.). The copy number profiles generated from multiple SMRT array experiments were aligned to enable the recognition of frequent alteration sites in invasive and CIS lung cancer genomes (10 CIS, 6 lung SCC and 7 lung Adenoma carcinoma samples) and used to detect the amplification or deletion of gene copy numbers. Common areas of copy number changes (at least 3) across these multiple profiles were identified. This approach resulted in the identification of 173 genes.

Gene expression patterns were measured using a SAGE (serial analysis of gene expression) technique as follows.

SAGE tags for the 173 genes from the SMRT array were identified using the SAGE Anatomic Viewer (SAV, March 2005 version; the National Cancer Institute, U.S. National Institutes of Health). The SAGE tags were then matched against a proprietary database of Lung SAGE libraries to determine their expression across 5 carcinoma-in-situ (CIS) samples, 6 invasive small cell carcinoma (SCC) samples, and 1 metaplasia sample (collectively "lung cancer"), 17 normal bronchial epithelial brushings from lung cancer patients ("MAC+"), and 6 normal bronchial epithelial brushings from patients 6 months post successful resection and 4 normal bronchial epithelial brushings from high-risk normal subjects without lung cancer (heavy smokers) (collectively "MAC-") (Table 1).

Table 1. Proprietary SAGE libraries used and their groupings

Lung Cancer	5 CIS	CISKL4 , CISRWL4, CISJAL4, CISRL, CISWL
	6 Invasive	InvSSL, InvKL, InvPL, InvW2, InvApoolL, InvBpoolL
	1 Metaplasia	MetSAL4
MAC+	17 Mac Positive	MJA, MR, MwlA, MWlB, MSS, MDD1738, MTF00, MLP00, MADV00, MSK1682, MSK1724, MCS1194, MSK1524, MAVE1337, MCS1454, MRW1672, MRW1706
MAC-	6 Post Resection	NTF01, NLP01, NADV01, NCS1419, NSK1735, NAVE1623
	4 High Risk Normals	ACA1494JLB, ACA1776JLT, ACA1582WTB, ACA1762WTT

SAGE libraries were constructed according to the MicroSAGE protocol, using *Nla* III as the anchoring enzyme and *Bsm* FI as the tagging enzyme (Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995) Science **270**(5235), 484-487). Reagents, primers and restriction enzymes were purchased from Dynal Biotech (Brown Deer, Wisconsin), Integrated DNA Technologies (Toronto, Ontario), Fisher Scientific (Nepean, Ontario), and New England Biolabs (Pickering, Ontario). The I-SAGE kit and Platinum Taq polymerase were purchased from Invitrogen Life Technologies (Burlington, Ontario). Tag counts were scaled to 106 tags/library and presented as tags per million (TPM). Tag-to-gene mapping was according to the SAGE Genie database (Boon, K., Osorio, E. C., Greenhut, S. F., Schaefer, C. F., Shoemaker, J., Polyak, K., Morin, P. J., Buetow, K. H., Strausberg, R. L., De Souza, S. J., and Riggins, G. J. (2002) Proc Natl Acad Sci U S A **99**(17), 11287-11292), with reference to SAGEmap.

During construction of the SAGE libraries, an interim analysis of a subset of the SAGE data was performed. 20 commercially available antibodies to target genes were initially evaluated by immunohistochemistry across 12 MAC (standard morphological MAC) positive and 12 MAC negative bronchial alveolar lavage (BAL) samples, as well as 36 lung cancer cell lines. To test these target antibodies across cytological samples from multiple high risk normal subjects and multiple cancer patients a technique for spotting multiple (up to 120) cytological samples on to individual slides and making multiple similar copies (each slide has all the same distribution samples on it) was used as described in MacAulay, C & Korbelik, J., Microvolume liquid dispenser suitable for microarrays and methods related thereto. United States Patent Application 20030003025, filed June 19, 2002. This cytology microarray technology enables the rapid evaluation of antibody expression (immunostaining) across multiple samples in a single experiment.

The following four markers exhibited the best immunohistochemistry performance, based on greatest differential immunostaining intensity (Figures 1A-H, Table 2).

Table 2

Gene Name	Antibody (Exemplary or Used in Assay)	Accession No.	UniProt/Swiss -Prot No.:
CD3	Polyclonal Rabbit anti-human, A0452 (Dako Canada, Inc., Missisauga, ON)	NM_000733.2	P07766
CD20cy	Monoclonal Mouse anti-human, clone L26 (Dako Canada, Inc., Missisauga, ON)	NM_152866	P11836
S100 protein (both alpha and beta subunits)	Polyclonal Rabbit anti-S100, Z0311 Ig Fraction (Dako Canada, Inc., Missisauga, ON)	NM_002960.1	P33764
p63	Monoclonal Mouse anti-human, clone 4A4 (Dako Canada, Inc., Missisauga, ON)	NM_003722	Q9H3D4

These four markers/antibodies were reevaluated on cytology microarrays containing 30 BAL samples from lung cancer patients and 30 samples from subjects without lung cancer (Figures 1A-H). The BAL samples were collected from well studied and classified (white light and fluorescence bronchoscopy, high resolution spiral CT, biopsy confirmed diagnosis) subjects and patients. The results were analysed using STATISTICA software, version 6.0 (Tulsa, OK).

SMRT array genomic profiles (DNA copy number changes) for these 4 markers in CIS, adenocarcinoma and squamous cell carcinoma indicate that, for the S100A genes located on 1q21, amplification is seen in 3 of 10 CIS samples, 0 out of 6 SCC samples, and 3 out 7 of the adenomacarcinoma sample); for the TP73L (p63) gene located on 3q27, amplification is seen in 7 of the 10 CIS samples, 1 out of 6 for the SCC samples, and 1 out of 7 of the adenomacarcinoma samples; for the CD3 gene located on 11q23, deletions are seen in 2 of the 10 CIS samples, amplification in 3 out of 6 SCC samples, and no changes in the 7 adenomacarcinoma samples; for the CD20cy gene located on 11q12.1, no changes were seen in the 10 CIS samples, no changes in the 6 SCC samples, and no changes in the 7 adenomacarcinoma samples.

In addition, the performance of E-cadherin (Accession Nos. NM_004360 and P12830) and CD23 (NM_002002.3 and P06734) and various combinations thereof were analysed (Figure 2). The sensitivity and specificity of these biomarkers for detection of lung cancer is shown in Table 5.

Table 5: Performance of various combinations of cd23 and e-cadherin immunohistochemical markers for the detection of lung cancer through immunostaining of BAL samples.

Immunostain	Sensitivity (%)	Specificity (%)
CD23	70	82
e-cadherin T1	92	70
e-cadherin T2	65	83
CD23 + e-cadherin T1	77	81
e-cadherin red staining	77	69
CD23 + e-cadherin T2	61	89

T1= Threshold 1: more than 50 % cells stain positive
T2=Threshold 2: more than 5% cells stain positive

Gene Annotation and Protein Data:

In addition to the interim analysis, the expression profiles for each gene (TAG) across the SAGE libraries were examined to find genes in which the expression was altered between lung cancer samples, MAC+, and MAC- samples, as indicated in Figures 3-37 and summarized in Table 3. Attention was focussed on genetic and expression changes associated with signaling pathways. The results from the application of this approach identified a number of gene expression changes in signaling pathways as seen by SAGE, out of the 173 frequent genetic changes (genes) of the SMRT array (Table 3), where a (+) refers to over-expression and a (-) refers to under-expression.

Table 3. Summary of Different Expression Results of Figs. 3-37

Name	OMIM No.	Accession No.	UniProt/Swiss-Prot No.:	Lung Cancer	MAC+	MAC-
ALCAM	601662	NM_001627	Q13740	-	+	+
ADCY3	600291	NM_004036	O60266	+	+	-
BAIAP2	605475	NM_017450.1	Q9UQB8	-	+	+
BCL2L 1.1	600039	NP_001182	Q07817	-	+	+
BCL2L 1.2	600039	NP_001182	Q07817	+	-	-
BDH 1	603063	NM_004051.4	Q02338	-	+	+
BDH 2				+	-	-
BOK		NM_032515	Q9UMX3	+	+	-
C1QTNF1	Hs.201398	NM_030968	Q9BXJ1	+	-	-
C1QTNF3	Hs.201394	NM_181435	Q9BXJ4	-	+	+
CAPN5	602537	NM_004055.3	O15484	-	+	+
CASPR2	604569	NM_014141	Q9UHC6	+	+	-
CD44	107269	NM_000610	P16070	+	+	-
CD59	107271	NM_203330	P13987	-	+	+
DTDST1	606718	NM_000112	P50443	+	+	-
EMS1	164765	NM_005231	Q14247	+	-	+
FGFR 1.1	136350	NM_000604	P11362	+	-	+
GALNS	253000	NM_000512	P34059	-	-	+
IL10	124092	NM_000572.2	P22301	+	+	-
JAK2		NM_004972.2	O60674	-	-	+
KNS2	600025	NM_005552.3	Q07866	-	-	+
MTGR2, CBFA2T3	603870	NM_005187	O75081	-	+	+
P100	605917	NM_012112	Q9ULW0	+	-	-
PCDH9	603581	NM_203487	Q9HC56	-	+	-
PECAM1, CD31	173445	NM_000442.2	P16284	+	+	-
PG11, PG40, DCN, PROTEOGLYCAN II	125255	NM_001920	P07585	+	+	-
PROS1	176880	NM_000313.1	P07225	-	-	+

RPL39	607547	NM_052969.1	Q96EH5	+	+	-
SFTPA1		NM_005411.3	Q8IWL2	-	+	-
SNRPB	182282	NM_198216	P14678	+	-	+
TP53		NM_000546.2	P04637	+	-	-
TSHR	603372	NM_000369	P16473	-	+	-
UNC13	605836	NM_006377.2	O14795	-	+	-
UGCS, GCS		NM_001498.2	P48506	-	+	-

Additional genes exhibiting differential expression across lung cancer samples, MAC+, and MAC- samples are indicated in Table 4.

Table 4: Differentially Expressed Genes

Gene Name	**OMIM No.**	**Accession No.**	**UniProt/Swiss-Prot No.:**
SOCS1	603597	NM_003745.1	O15524
RATOR, STX11,	604206	NM_003764	O75558
CD6	186720	NM_006725.2	P30203
Kera	603288	NM_007035	O60938
LUM, LDC,	600616	NM_002345.3	P51884
BCAP31	300398	NM_005745.6	P51572
EFNB1	300035	NM_004429	P98172
FLNA	300017	NM_001456	P21333
SH2D1A	300490	NM_002351.1	O60880
ADAMTS1	608990	NM_006988	Q9UHI8
MYO7A	276903	NM_000260	Q13402
WNK1, PRKWNK1	605232	NM_018979.1	Q9H4A3
RPL39L-associated with 60S subunit	607547	NM_052969.1	Q96EH5
MAD1L1 (7p)		NM_003550	Q9Y6D9
CARD11(7p)		NM_032415.2	Q9BXL7
PTPRN2 (7q)		NM_002847.2	Q92932

REFERENCES

1. Fry, W.A. et al. (1996) The National Cancer Data Base report on lung cancer. *Cancer* 77 (9), 1947-1955.
2. Palcic B, Garner DM, Beveridge J, Sun XR, Doudkine A, MacAulay C, Lam S, Payne PW. (2002) Increase of sensitivity of sputum cytology using high-resolution image cytometry: field study results. Cytometry. Jun 15;50(3):168-76.

3. Nieburgs, H.E., Zak, F.G., Allen, D.C. et al. Systemic cellular changes in material from human and animal tissues in presence of tumors. In: Transactions of the 7th Annual Meeting of the Inter-Society Cytology Council, 1959, p.137.
4. Palcic, B., MacAulay, C. Malignancy associated changes: can they be emlployed clinically ? In: Wied, G.L., Bartels, P.H., Rosenthal, D.L., Schenk, U. (eds): Compendium on the Computerized Cytology and Histology Laboratory. Chicago: Tutorials of Cytology, 1994, p. 157.
5. Sun, X.R. et al. (2003) In vitro model for studying malignancy associated changes. *Anal Cell Pathol* 25 (2), 95-102
6. Nieburgs, H.E. (1959) Tubeless gastric analysis and cytologic procedures. *Surg Clin North Am* 39, 1213-1222
7. van Oppen Toth, B. et al. (1975) Malignancy Associated Changes in Sputum: A Correlated Study of 315 Patients. *Acta Cytologica* 19 (6), 573-576
8. Nieburgs, H.E. and Goldberg, A.F. (1968) Changes in polymorphonuclear leukocytes as a manifestation of malignant neoplasia. *Cancer* 22 (1), 35-42
9. Finch, R.R. (1971) A classification of nuclear aberrations in relation to malignancy associated changes 9MAC). *Acta Cytol* 15, 553-558
10. De Potter, C. et al. (1987) Feulgen DNA content and mitotic activity in proliferative breast disease: A comparison with ductal carcinoma in situ. *Acta Cytol* 11, 1307-1309
11. Klawe, H. and Rowinski, J. (1974) Malignancy associated changes (MAC) in cells of buccal smears detected by means of objective image analysis. *Acta Cytol* 18 (1), 30-33
12. Susnick, B. et al. (1995) Malignancy-Associated Changes in the Breast. *Analytical and Quantitative Cytology and Histology* 17 (1), 62-68
13. Bibbo, M. et al. (1989) Karyometric marker features in tissue adjacent to invasive cervical carcinomas. *Analyt Quant Cytol Histol* 11, 281-285
14. MacAulay, C. et al. (1995) Malignancy-associated changes in bronchial epithelial cells in biopsy specimens. *Anal Quant Cytol Histol* 17 (1), 55-61
15. Kasper, H.U. et al. (1997) Diagnostic and prognostic relevance of malignancy-associated changes in cervical smears. *Anal Quant Cytol Histol* 19 (6), 482-488

16. Wied, G.L. et al. (1980) Cytomorphometric markers for uterine cancer in intermediate cells. *Anal Quant Cytol* 2 (4), 257-263
17. Montag, A. et al. (1991) Karyometric features in nuclei near colonic adenocarcinoma:statistical analysis. *Analyt Quant Cytol Histol* 13, 159-167
18. Palcic, B. et al. (1991) Detection and localization of early lung cancer by imaging techniques. *Chest* 99 (3), 742-743
19. MacAulay, C & Korbelik, J. Microvolume liquid dispenser suitable for microarrays and methods related thereto. United States Patent Application 20030003025. Filed June 19, 2002.
20. Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, Ling V, MacAulay C, Lam WL. (2004). A tiling resolution DNA microarray with complete coverage of the human genome. Nat Genet. Mar;36(3):299-303.

OTHER EMBODIMENTS

Although various embodiments of the invention are disclosed herein, many adaptations and modifications may be made within the scope of the invention in accordance with the common general knowledge of those skilled in this art. Such modifications include the substitution of known equivalents for any aspect of the invention in order to achieve the same result in substantially the same way. Accession numbers, as used herein, refer to Accession numbers from multiple databases, including GenBank, the European Molecular Biology Laboratory (EMBL), the DNA Database of Japan (DDBJ), or the Genome Sequence Data Base (GSDB), for nucleotide sequences, and including the Protein Information Resource (PIR), SWISSPROT, Protein Research Foundation (PRF), and Protein Data Bank (PDB) (sequences from solved structures), as well as from translations from annotated coding regions from nucleotide sequences in GenBank, EMBL, DDBJ, or RefSeq, for polypeptide sequences. Numeric ranges are inclusive of the numbers defining the range. In the specification, the word “comprising” is used as an open-ended term, substantially equivalent to the phrase “including, but not limited to”, and the word “comprises” has a corresponding meaning. Citation of references herein shall not be construed as an admission that such references are prior art to the present invention. All publications are incorporated herein by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein and as though fully set forth herein. The invention includes all embodiments and variations substantially as hereinbefore described and with reference to the examples and drawings.

WHAT IS CLAIMED IS:

1. A method of detecting a malignancy associated change in a cell, the method comprising detecting differential expression of one or more nucleic acid molecules or expression products thereof in the cell, wherein said nucleic acid molecules or expression products thereof are selected from the group consisting of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), PTPRN2 (7q), e-cadherin and CD23 and wherein said differential expression is indicative of said malignancy associated change.

2. The method of claim 1 comprising detecting differential expression of two or more of said nucleic acid molecules or expression products thereof in the cell.

3. The method of claim 2 comprising detecting differential expression of three or more of said nucleic acid molecules or expression products thereof in the cell.

4. The method of any one of claims 1 to 3 wherein said nucleic acid molecules or expression products thereof are selected from the group consisting of CD3, CD20cy, S100, and p63.

5. The method of any one of claims 1 to 4 wherein said cell is an epithelial cell or a lymphocyte.

6. The method of any one of claims 1 to 5 wherein said cell is a human cell.

7. The method of any one of claims 1 to 6 wherein said cell is derived from a subject at risk for lung cancer.

8. The method of any one of claims 1 to 6 wherein said cell is derived from a subject having lung cancer.

9. The method of any one of claims 1 to 8 wherein said expression product is a polypeptide.

10. The method of any one of claims 1 to 9 wherein said detecting is performed by an immunoassay.

11. The method of claim 10 wherein said immunoassay is performed using an antibody that specifically binds said expression product.

12. The method of any one of claims 1 to 9 wherein said detecting is performed by a hybridisation assay.

13. The method of claim 12 wherein said hybridization assay is performed using an probe that hybridizes to said nucleic acid molecule.

14. A method of diagnosing or prognosing lung cancer in a subject comprising detecting the differential expression of one or more nucleic acid molecules or expression products thereof in a sample from the subject, wherein the nucleic acid molecules or expression products thereof are selected from the group consisting of one or more of MTGR2 [CBFA2T3], GALNS, KNS2, SOCS1, C1QTNF1, C1QTNF3, BAIAP2, RATOR [STX11], CD59, CAPN5, CD6, Kera, EMS1, PG11 [PG40, DCN, PROTEOGLYCAN II], LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), ALCAM, ADCY3, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK,

CASPR2, DTDST1, PCDH9, PROS1, RPL39, SNRPB, TSHR, UNC13, and PTPRN2 (7q).

15. The method of claim 14 further comprising detecting the differential expression of one or more nucleic acid molecules or expression products thereof in the sample, wherein the nucleic acid molecules are selected from the group consisting of one or more of PECAM1 [CD31], CD44, P100, FGFR 1.1, IL10, JAK2, SFTPA1, TP53, and UGCS [GCS].

16. The method of claim 15 further comprising detecting the differential expression of one or more nucleic acid molecules or expression products thereof in the sample, wherein the nucleic acid molecules are selected from the group consisting of one or more of CD44, SFTPA1, CD3, CD20cy, S100, and p63.

17. The method of any one of claims 14 to 16, wherein said sample is sputum, cough, buccal mucosa or bronchoalveolar lavage.

18. The method of any one of claims 14 to 17, wherein said sample is substantially free of malignant cells.

19. A method of diagnosing or prognosing lung cancer in a subject comprising detecting the differential expression of one or more nucleic acid molecules or expression products thereof in a sputum, cough or bronchoalveolar lavage sample from the subject, wherein the nucleic acid molecules or expression products thereof are selected from the group consisting of one or more of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA,

SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), and PTPRN2 (7q).

20. The method of claim 19 wherein the sample is substantially free of malignant cells.

21. A panel of biomarkers for lung cancer or for malignancy associated changes associated with lung cancer comprising two or more nucleic acid molecules or expression products thereof selected from the group consisting of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), and PTPRN2 (7q), with the proviso that when the panel comprises only two biomarkers, CD3, CD20cy [L-26], S100, p63 [4A4], CD44, or SFTPA1 may not both be present.

22. A method of monitoring the response of a subject undergoing therapy for lung cancer, the method comprising detecting differential expression of one or more nucleic acid molecules or expression products thereof in a cell derived from said subject, wherein said nucleic acid molecules or expression products thereof are selected from the group consisting of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p),

CARD11(7p), and PTPRN2 (7q), and wherein said differential expression is indicative of a malignancy associated change associated with said lung cancer.

23. The method of claim 22 wherein said therapy is surgical resection of said lung cancer.

24. A method of identifying a compound useful for lung cancer therapy, the method comprising detecting differential expression of one or more nucleic acid molecules or expression products thereof in a lung cancer cell in the presence or absence of said compound, wherein said nucleic acid molecules or expression products thereof are selected from the group consisting of CD3, CD20cy, S100, p63, ALCAM, ADCY3, BAIAP2, BCL2L 1.1, BCL2L 1.2, BDH 1, BDH 2, BOK, C1QTNF1, C1QTNF3, CAPN5, CASPR2, CD44, CD59, DTDST1, EMS1, FGFR 1.1, GALNS, IL10, JAK2, KNS2, MTGR2 [CBFA2T3], P100, PCDH9, PECAM1 [CD31], PG11 [PG40, DCN, PROTEOGLYCAN II], PROS1, RPL39, SFTPA1, SNRPB, TP53, TSHR, UNC13, UGCS [GCS], SOCS1, RATOR [STX11], CD6, Kera, LUM [LDC], BCAP31, EFNB1, FLNA, SH2D1A, ADAMTS1, MYO7A, WNK1 [PRKWNK1], RPL39L-associated with 60S subunit, MAD1L1 (7p), CARD11(7p), and PTPRN2 (7q), and wherein said differential expression is indicative of the efficacy of said compound.

25. Use of the panel of claim 21 for preparation of a medicament for detecting a lung cancer.

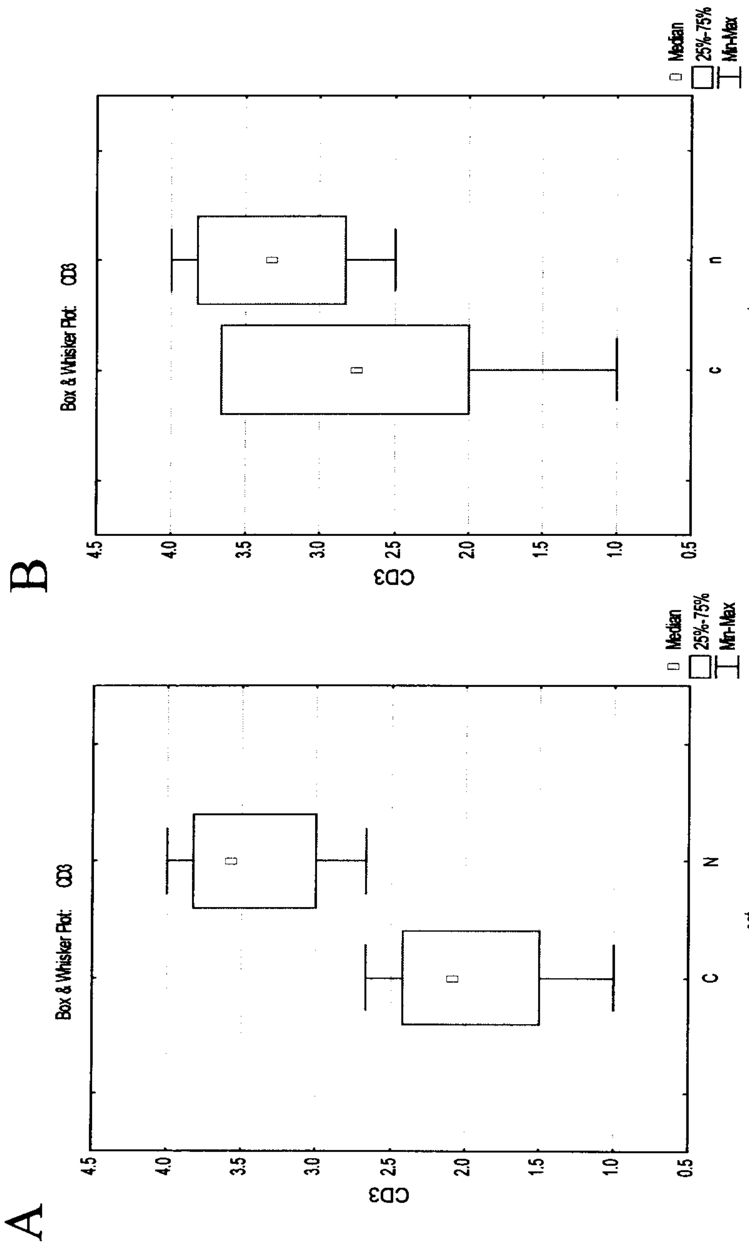

A
Box & Whisker Plot: CD3
CD3
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
C
N
cat
Median
25%-75%
Min-Max
B
Box & Whisker Plot: CD3
CD3
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
c
n
cat
Median
25%-75%
Min-Max

FIG. 1

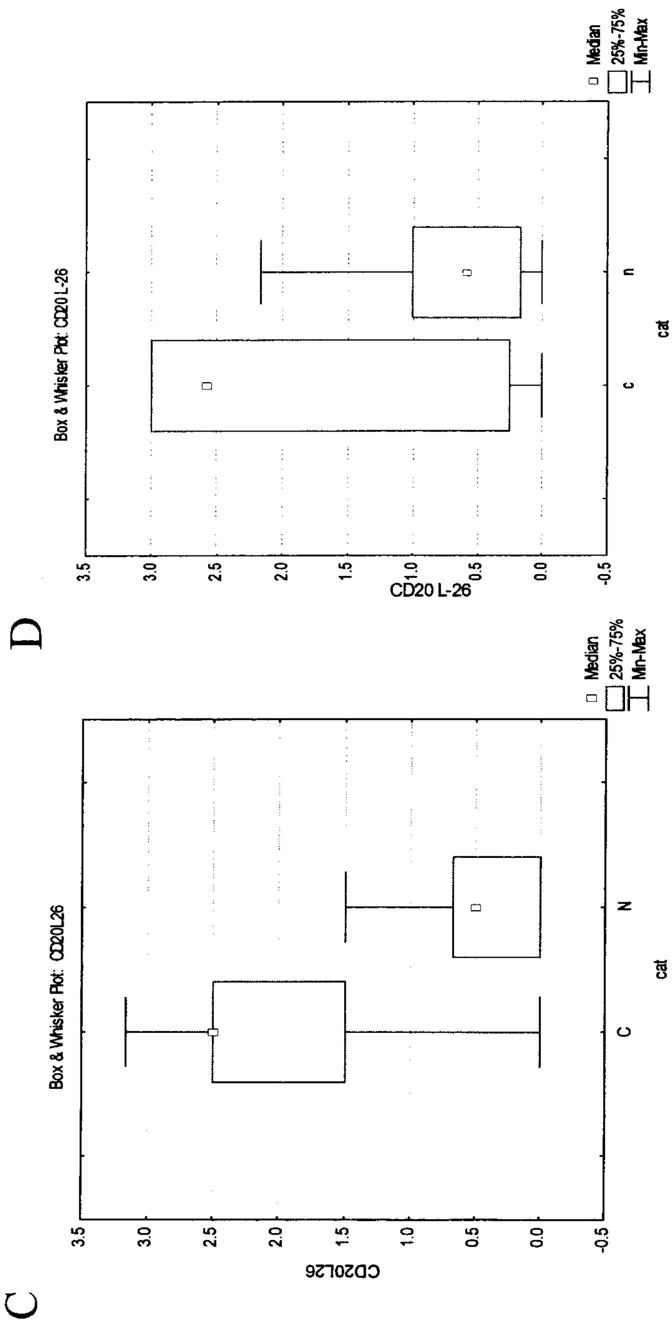

C
Box & Whisker Plot: CD20L26
CD20L26
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
C
N
cat
Median
25%-75%
Min-Max
D
Box & Whisker Plot: CD20 L-26
CD20 L-26
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
c
n
cat
Median
25%-75%
Min-Max

FIG. 1

F

E

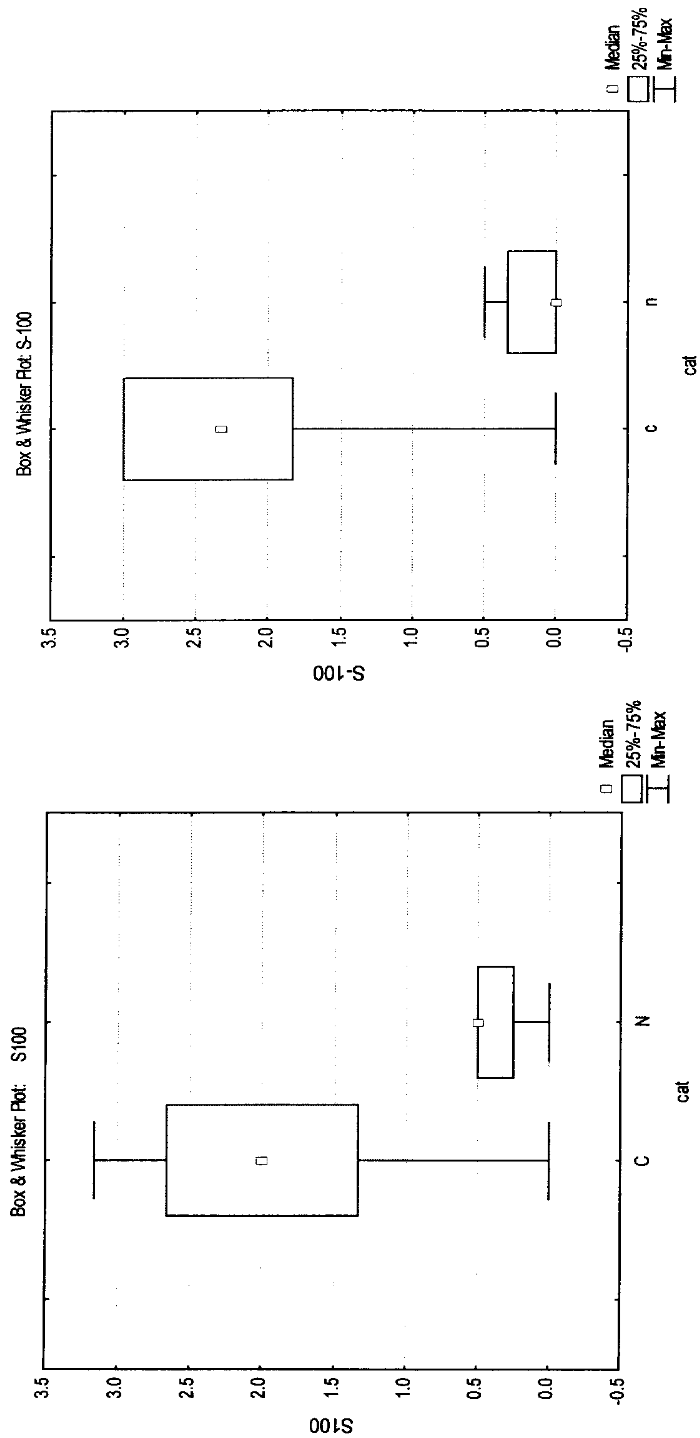

Box & Whisker Plot: S-100
S-100
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
c
n
cat
Median
25%-75%
Min-Max
Box & Whisker Plot: S100
S100
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
C
N
cat
Median
25%-75%
Min-Max

FIG. 1

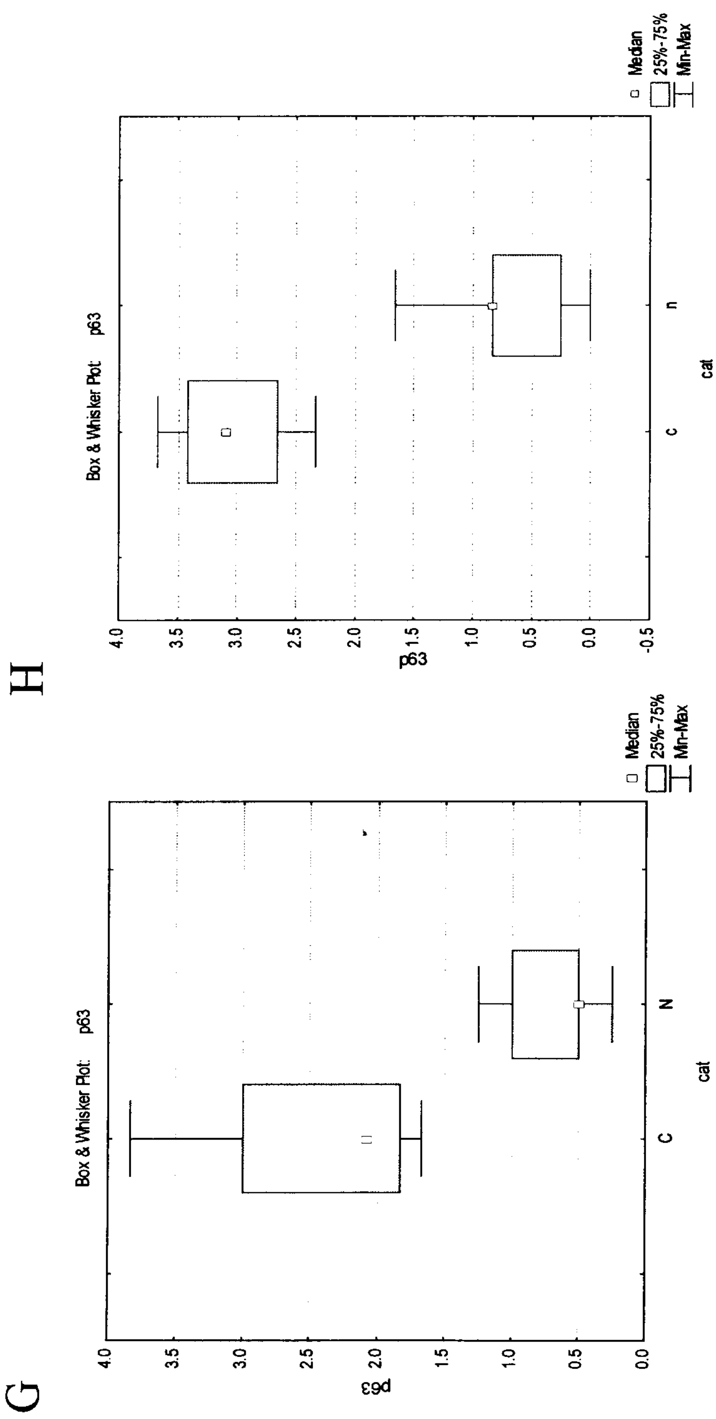

G
H
Box & Whisker Plot: p63
p63
cat
C
N
c
n
Median
25%-75%
Min-Max
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5

FIG. 1

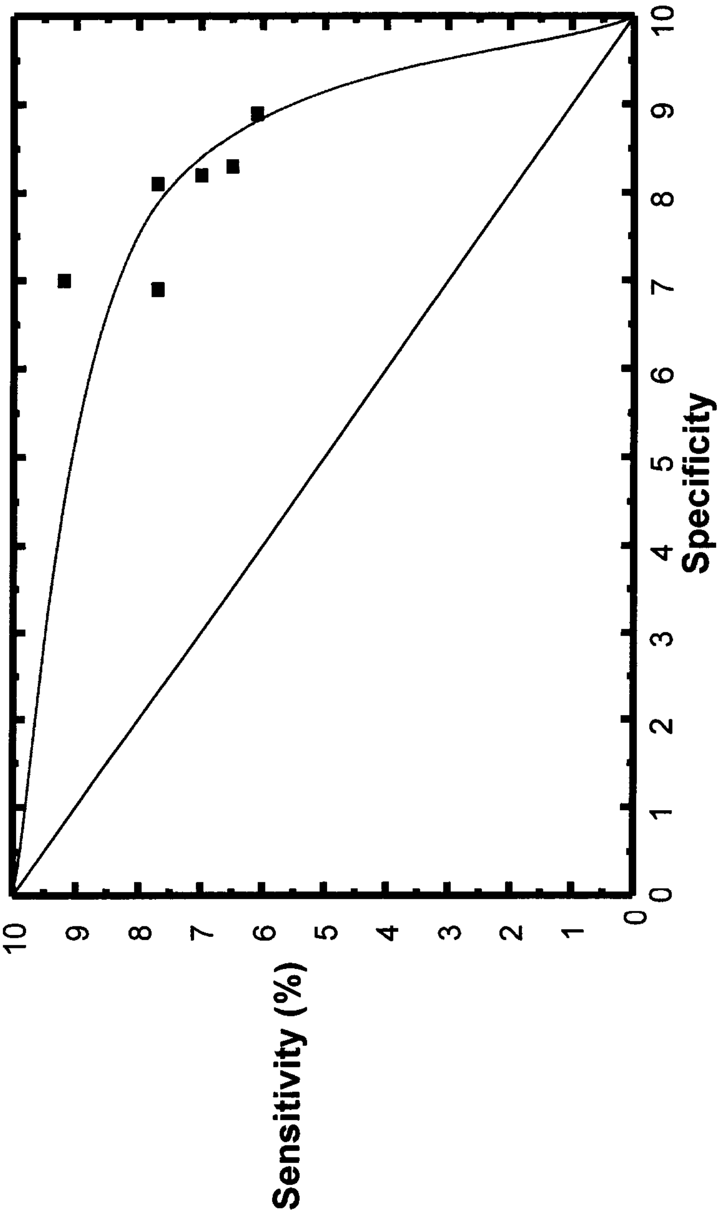


FIG. 2

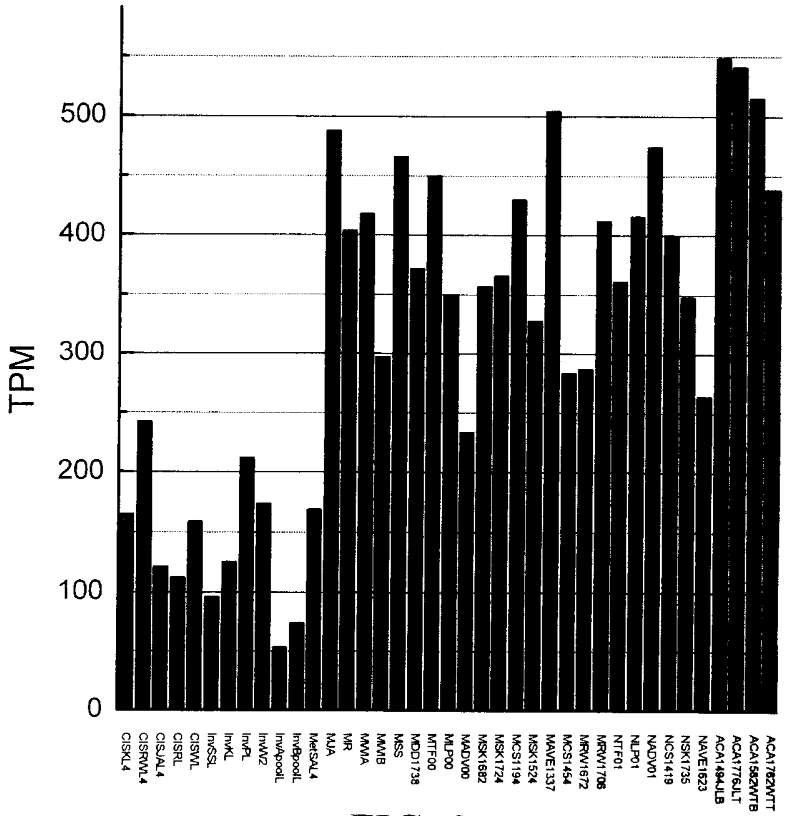

TPM
0
100
200
300
400
500
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1782WTT

FIG. 3

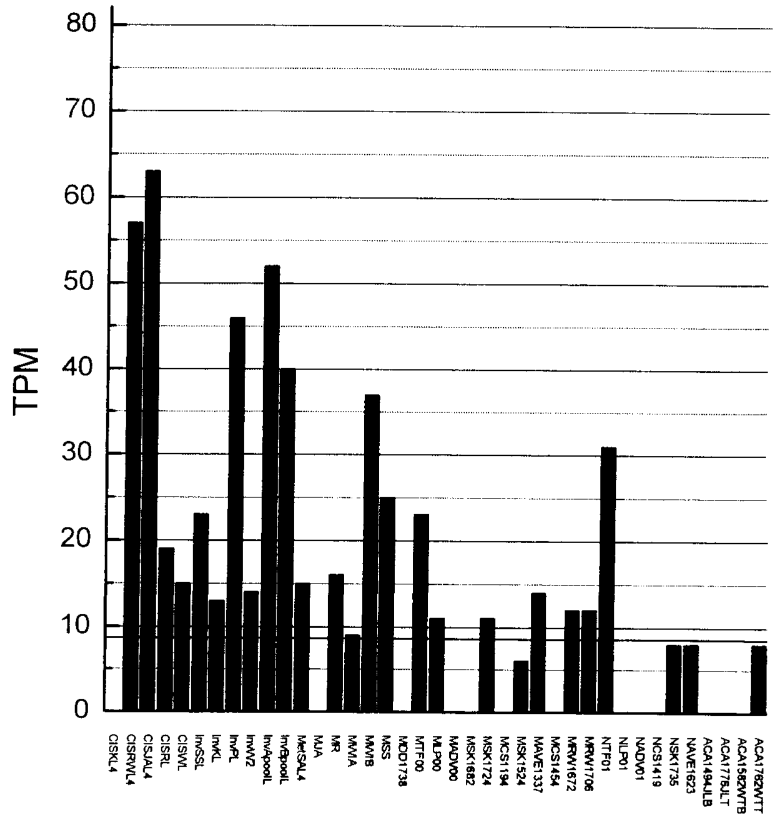

TPM
0
10
20
30
40
50
60
70
80
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1782WTT

FIG. 4

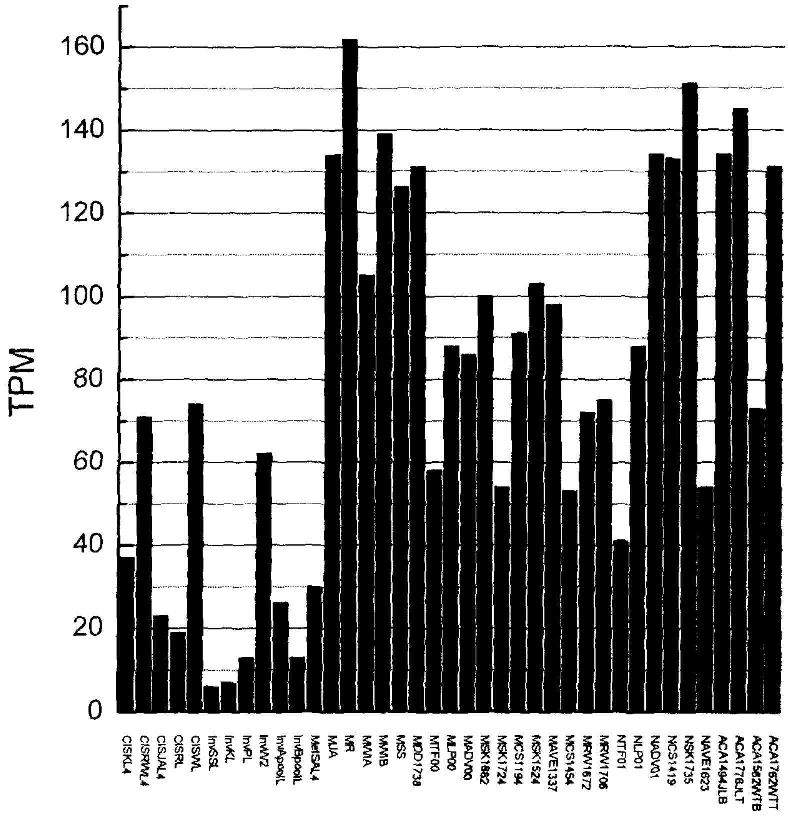

TPM
0
20
40
60
80
100
120
140
160

FIG. 5

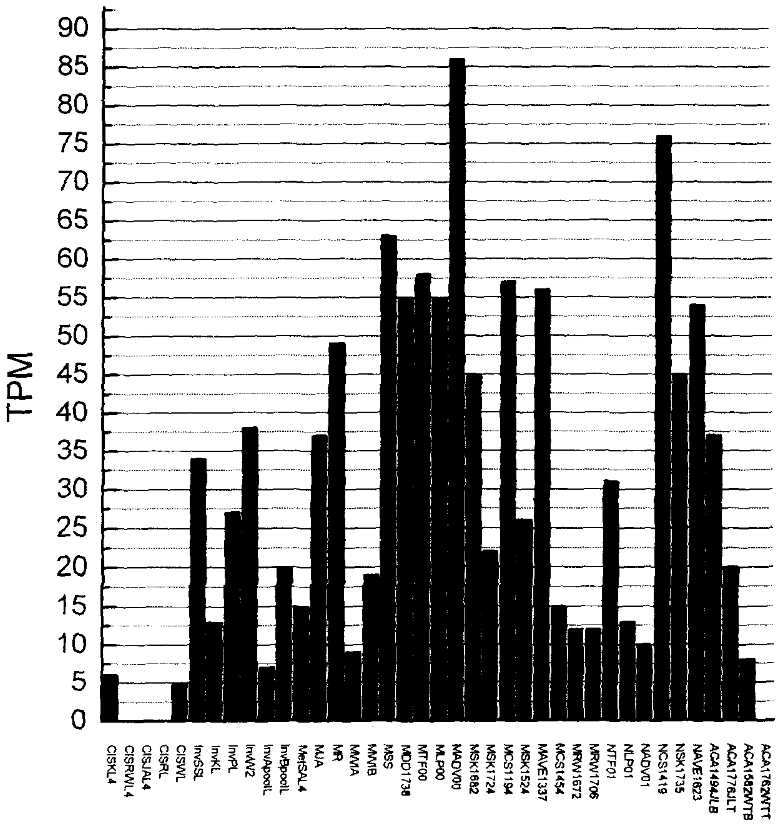

TPM
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90

FIG. 6

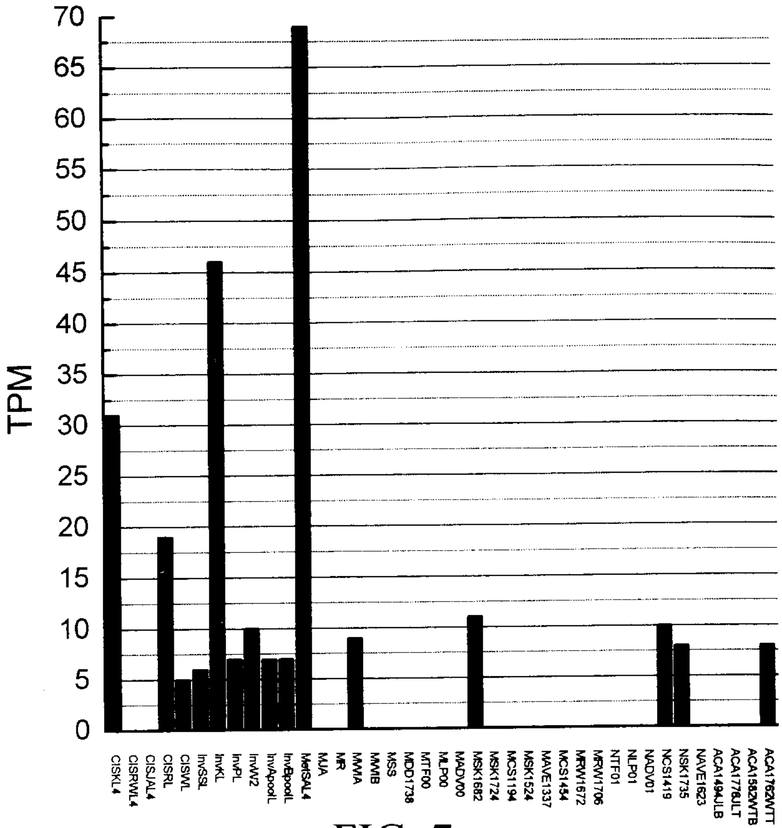

TPM
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRV1672
MRV1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 7

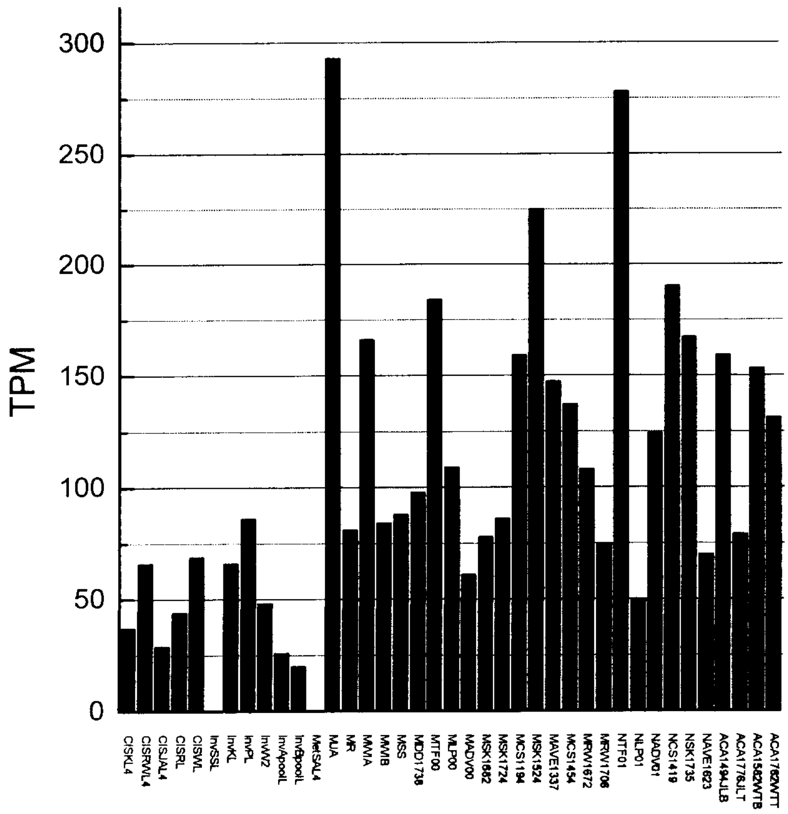

TPM
300
250
200
150
100
50
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRV1672
MRV1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 8

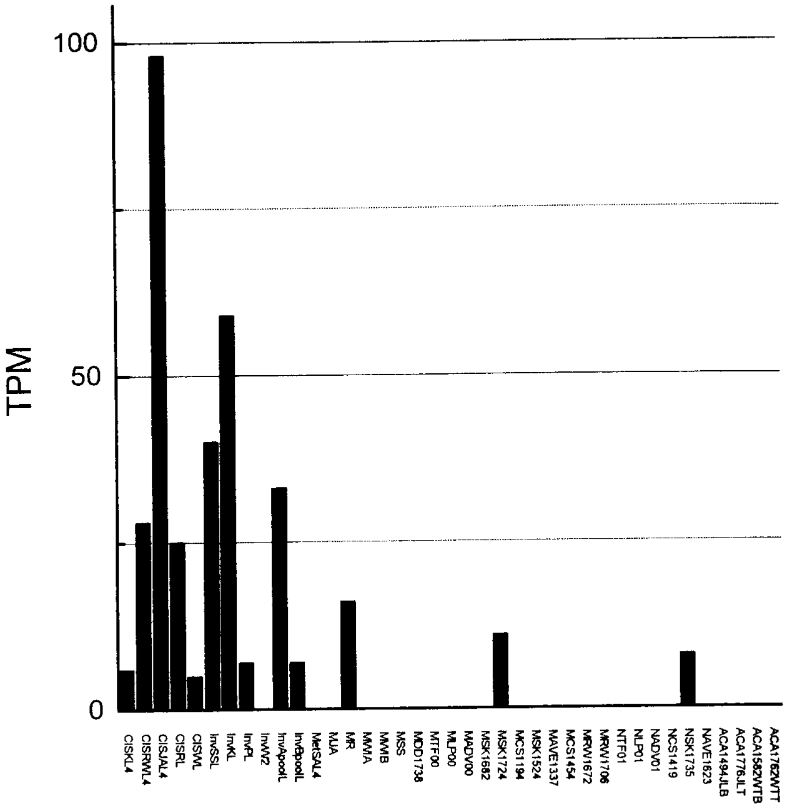

TPM
100
50
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 9

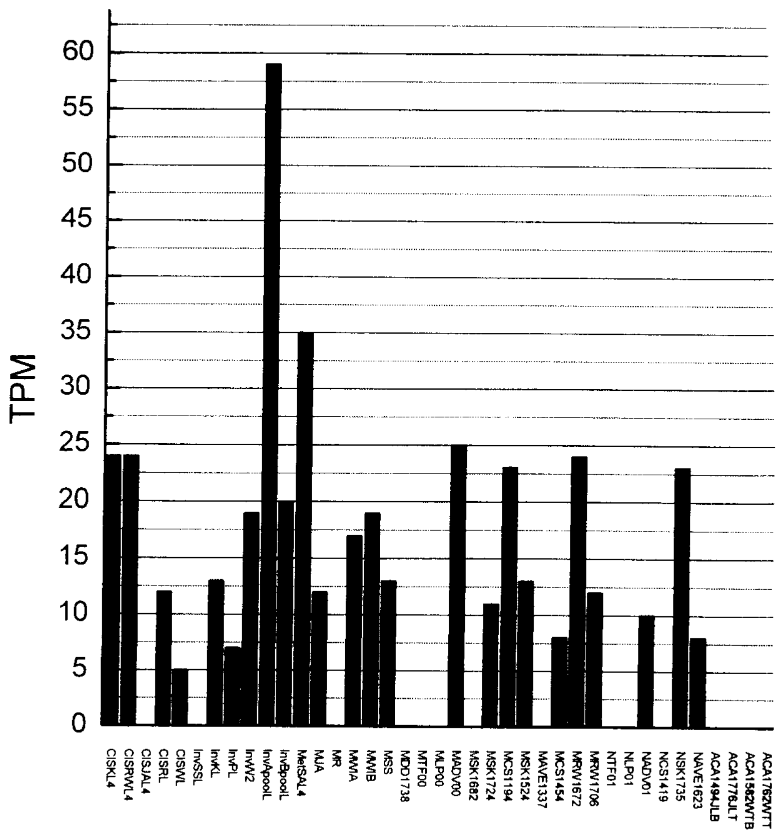

TPM
60
55
50
45
40
35
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 10

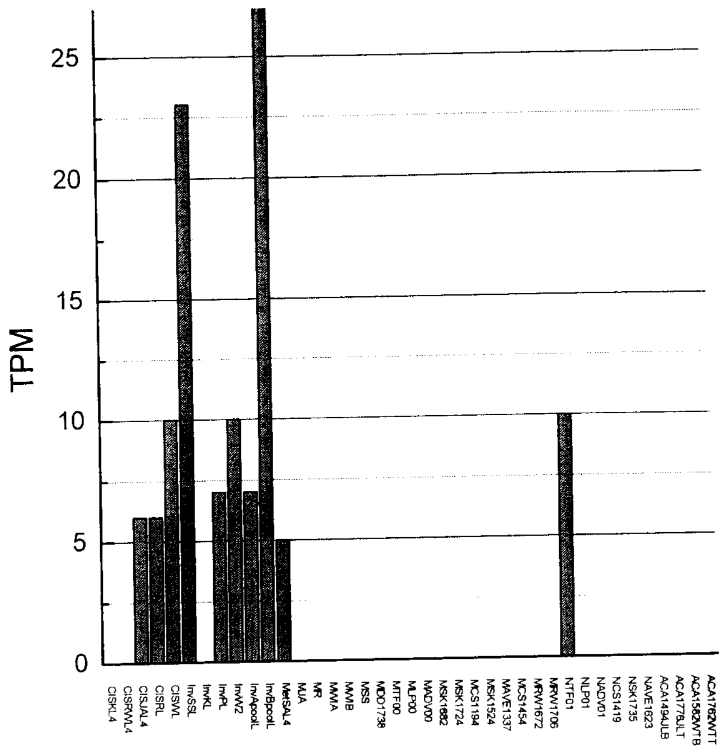

TPM
0
5
10
15
20
25
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1562WTB
ACA1762WTT

FIG. 11

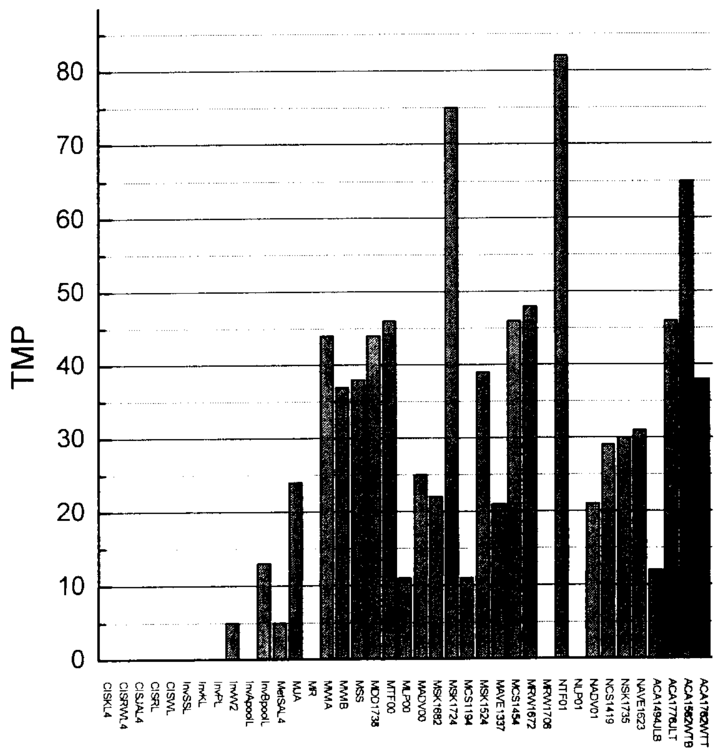

TMP
0
10
20
30
40
50
60
70
80
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1562WTB
ACA1762WTT

FIG. 12

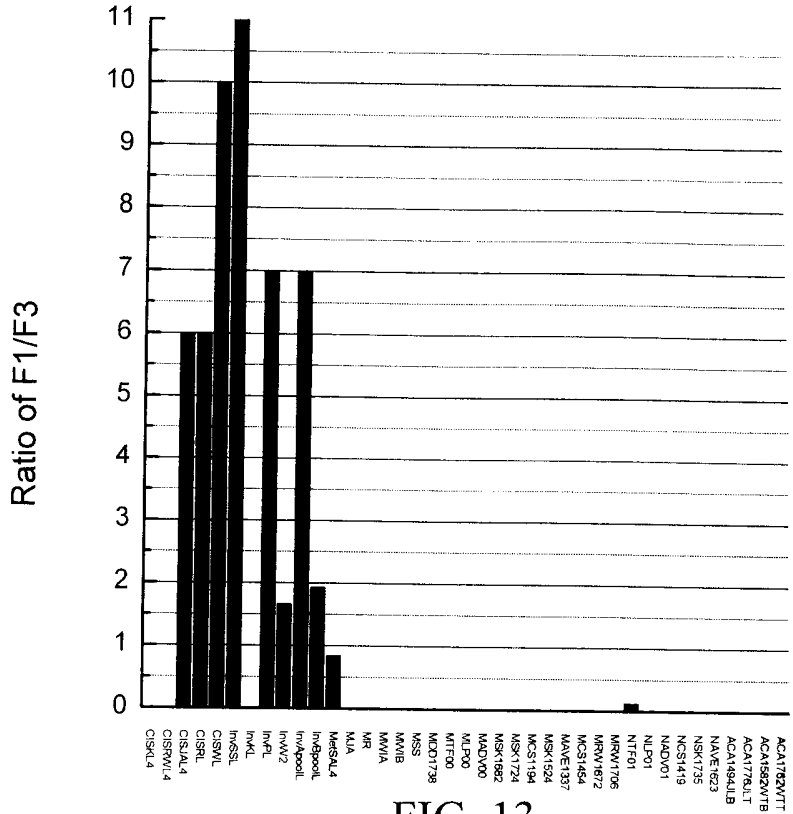

Ratio of F1/F3
11
10
9
8
7
6
5
4
3
2
1
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRV1672
MRV1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 13

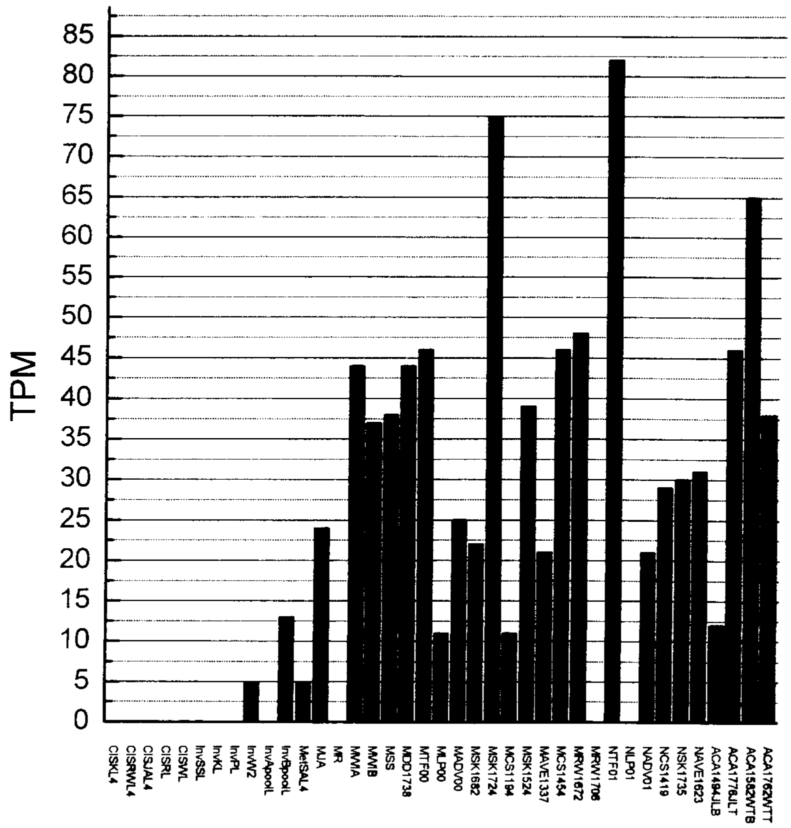

TPM
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRV1672
MRV1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 14

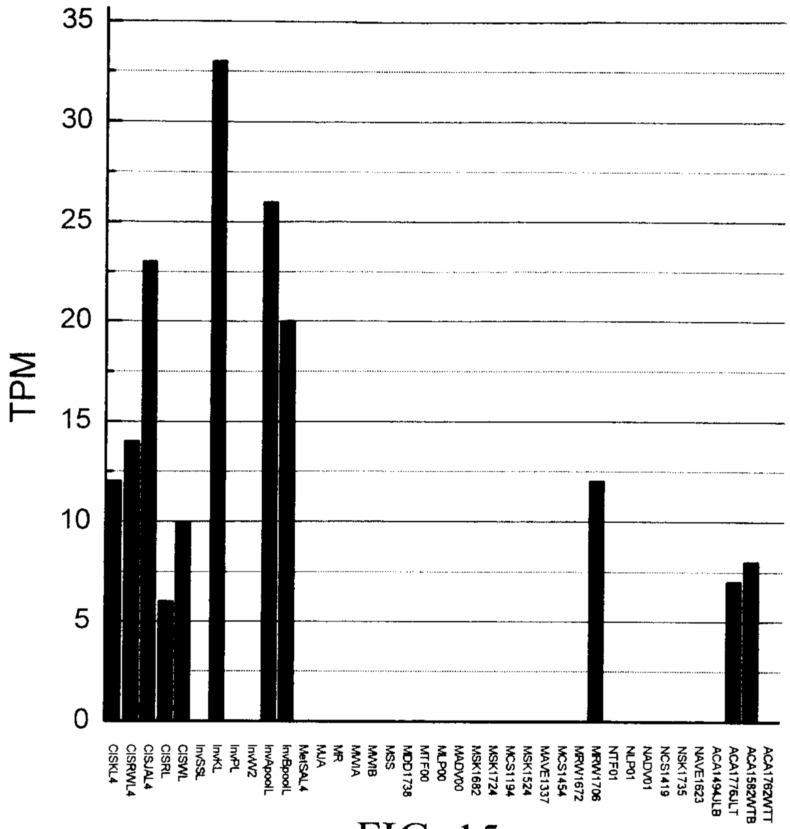


FIG. 15

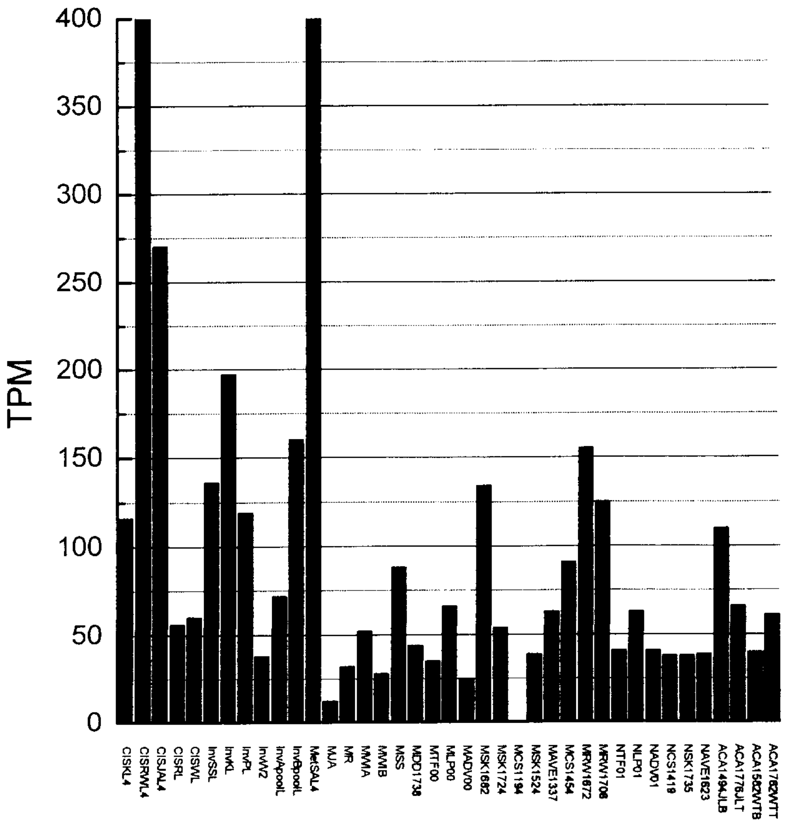


FIG. 16

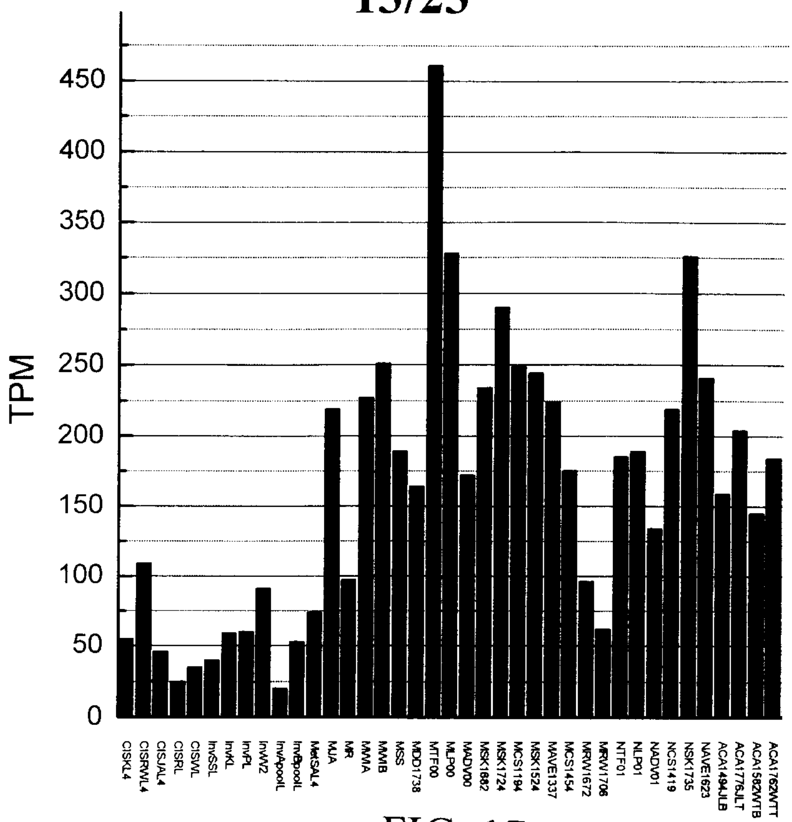

TPM
0
50
100
150
200
250
300
350
400
450
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1882
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1708
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 17

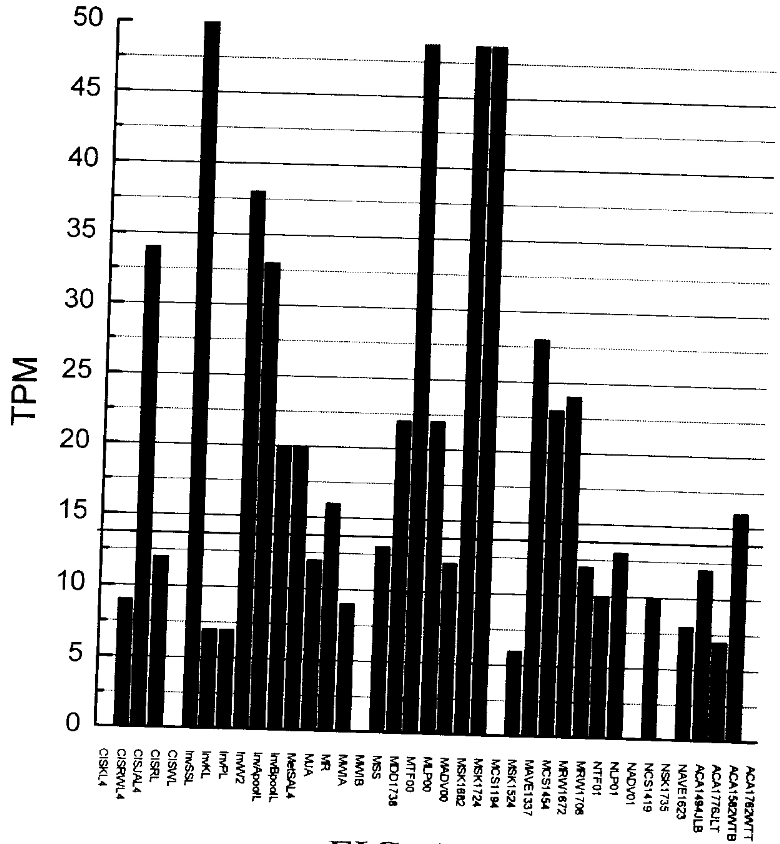

TPM
0
5
10
15
20
25
30
35
40
45
50
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1882
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1708
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 18

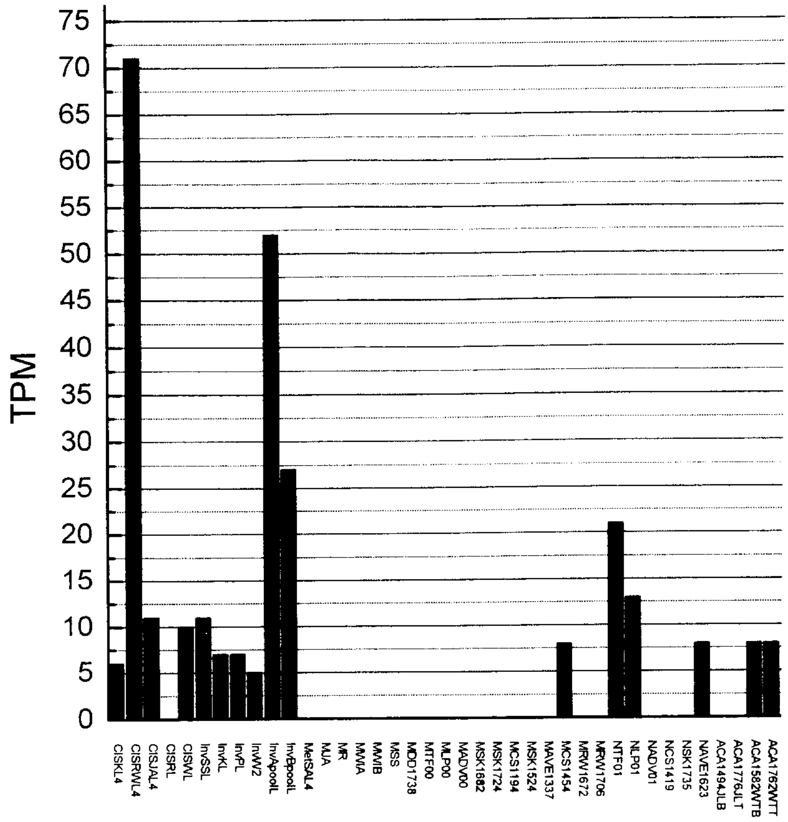

TPM
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MOD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 19

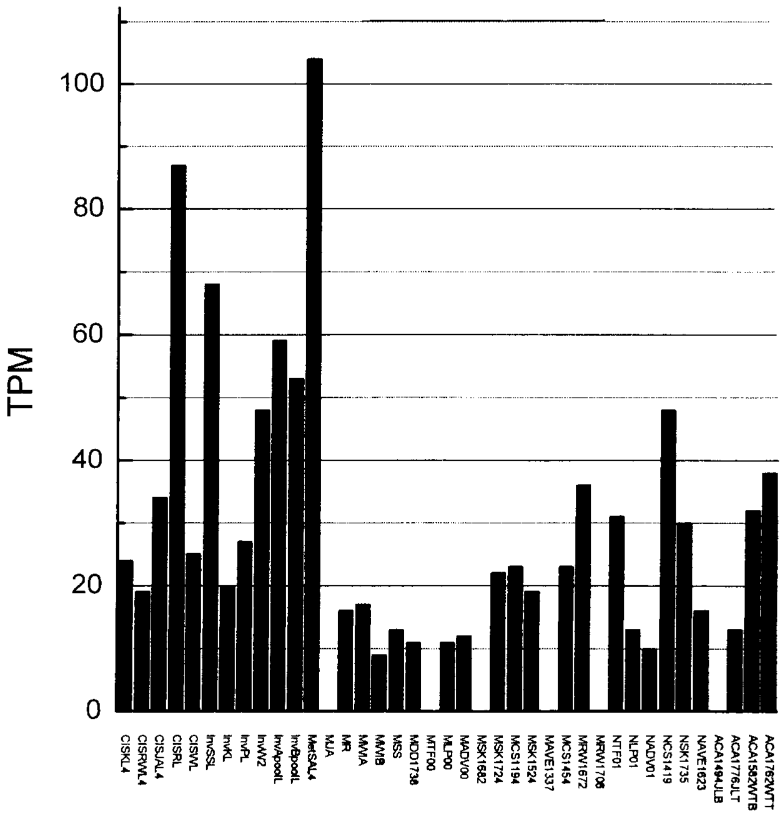
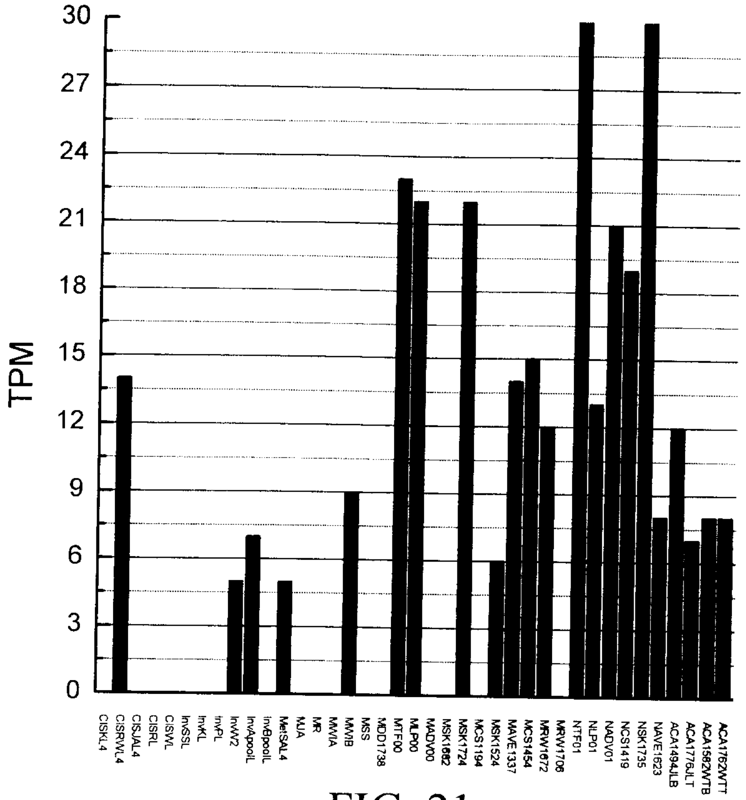
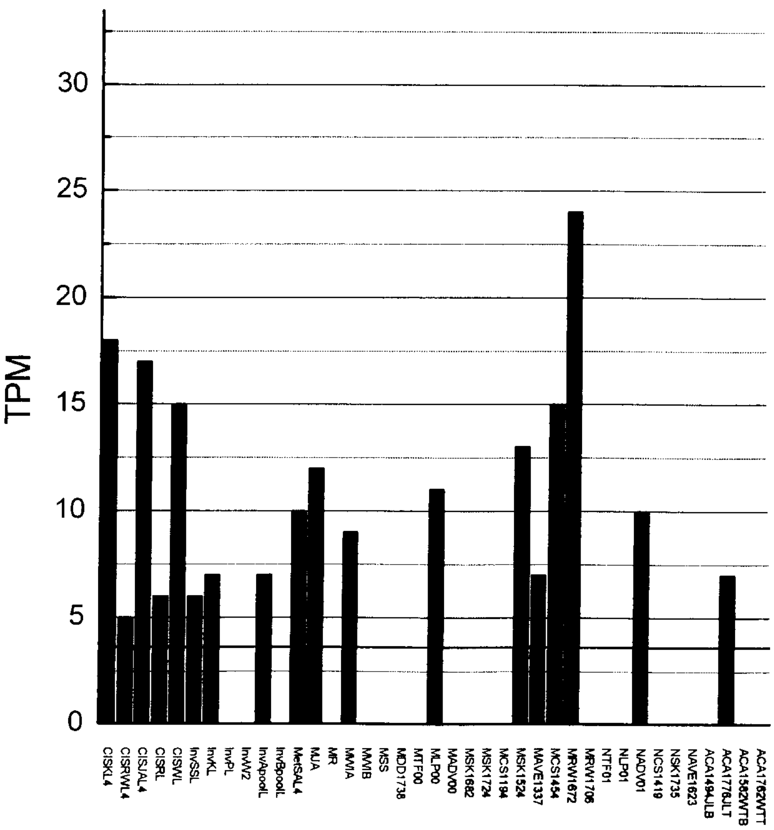

TPM
100
80
60
40
20
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MOD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 20



TPM
30
27
24
21
18
15
12
9
6
3
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 21



TPM
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 22

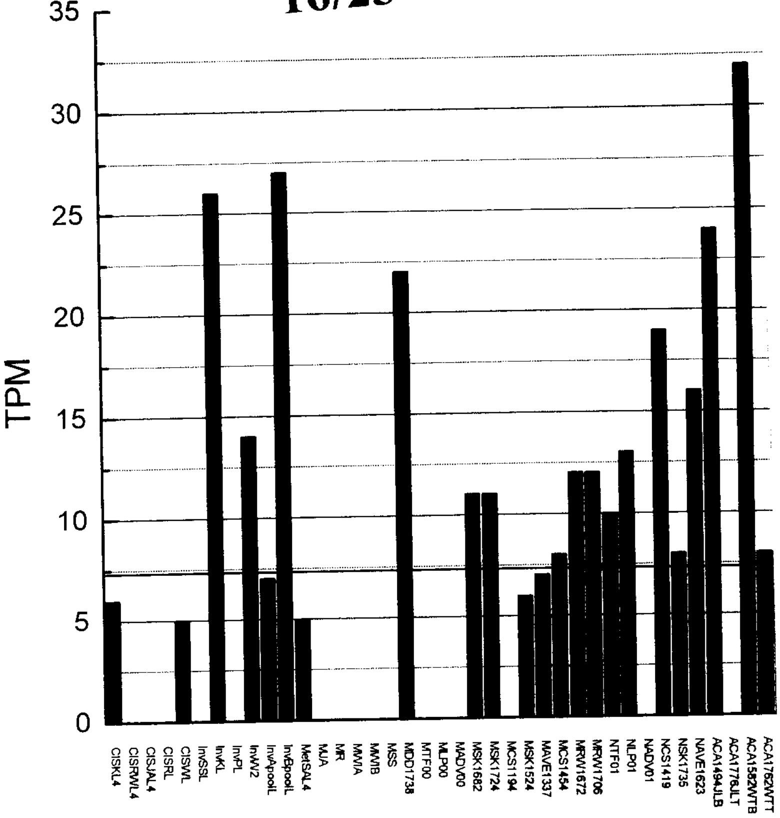

TPM
0
5
10
15
20
25
30
35
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRV1672
MRV1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 23

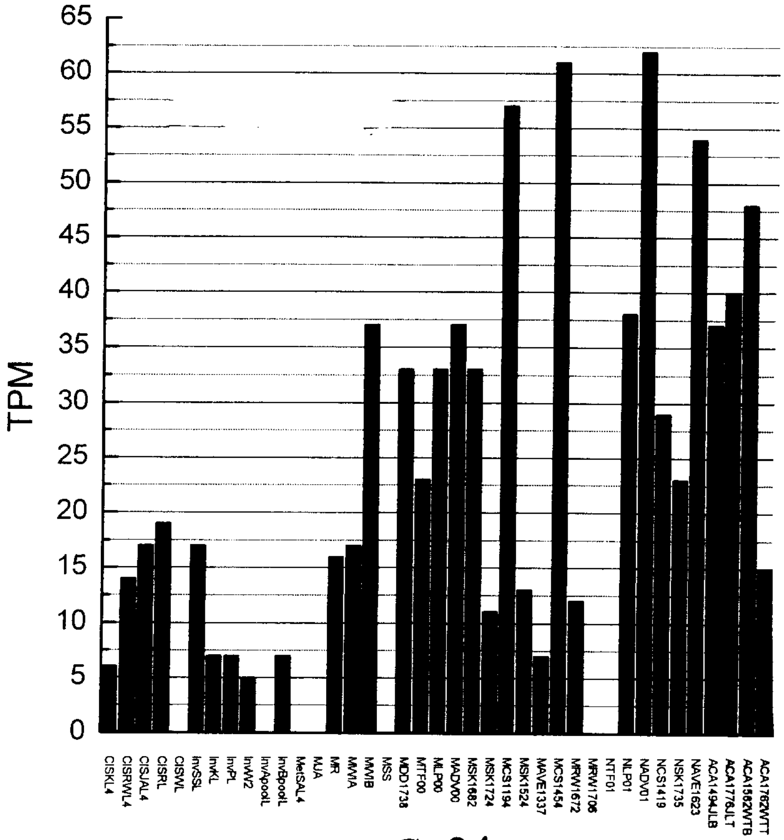

TPM
0
5
10
15
20
25
30
35
40
45
50
55
60
65
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRV1672
MRV1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 24

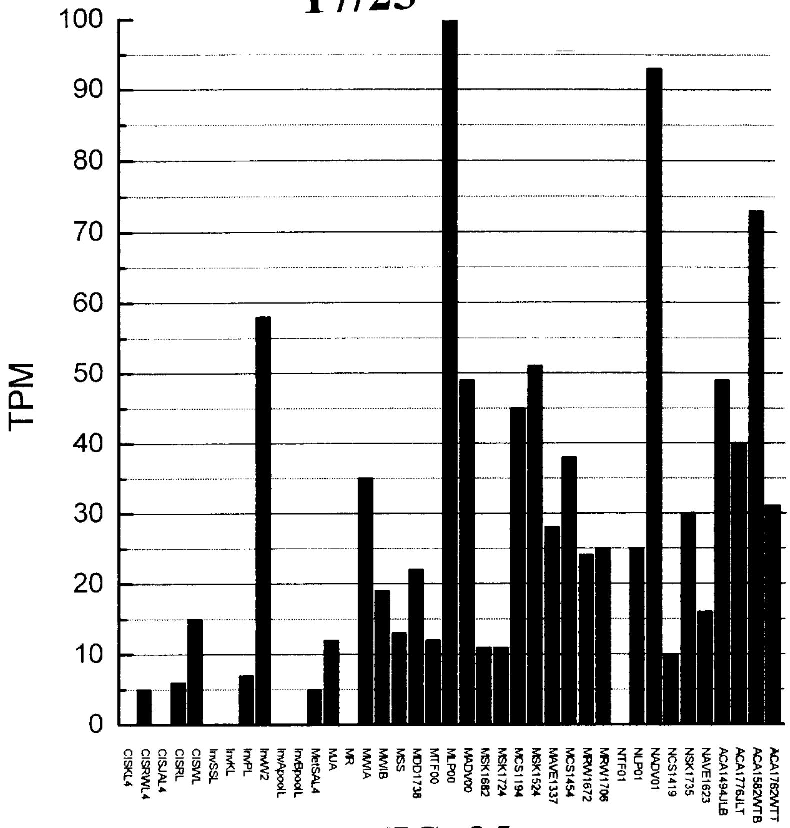

TPM
0
10
20
30
40
50
60
70
80
90
100
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 25

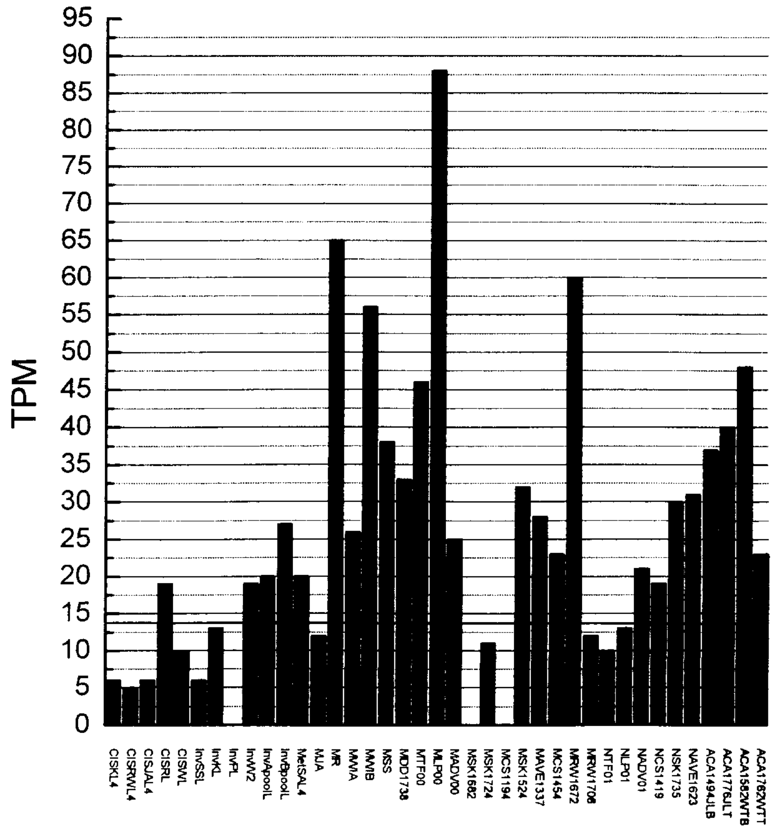

TPM
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 26

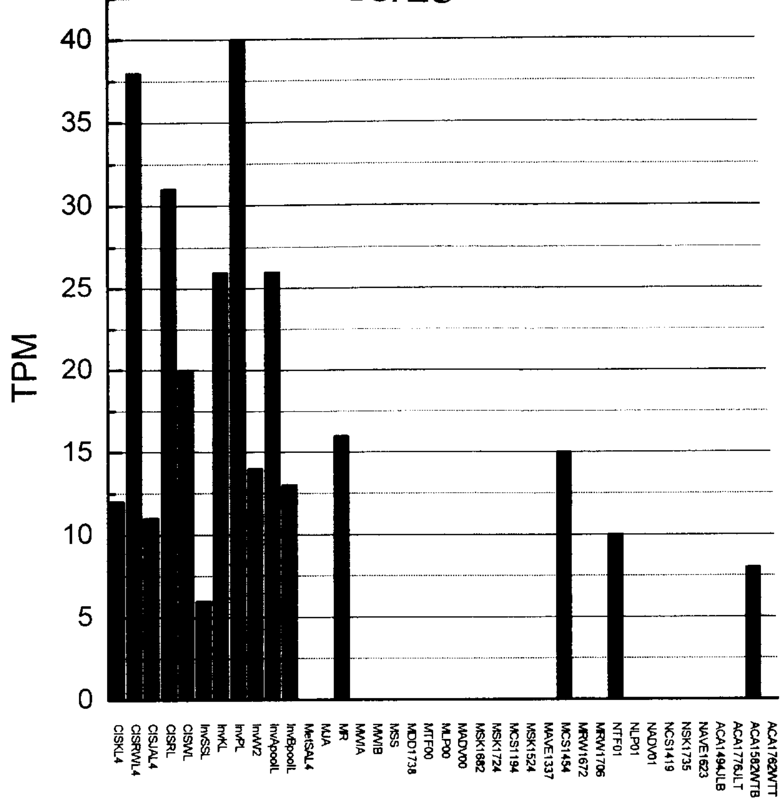

TPM
0
5
10
15
20
25
30
35
40
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1782WTT

FIG. 27

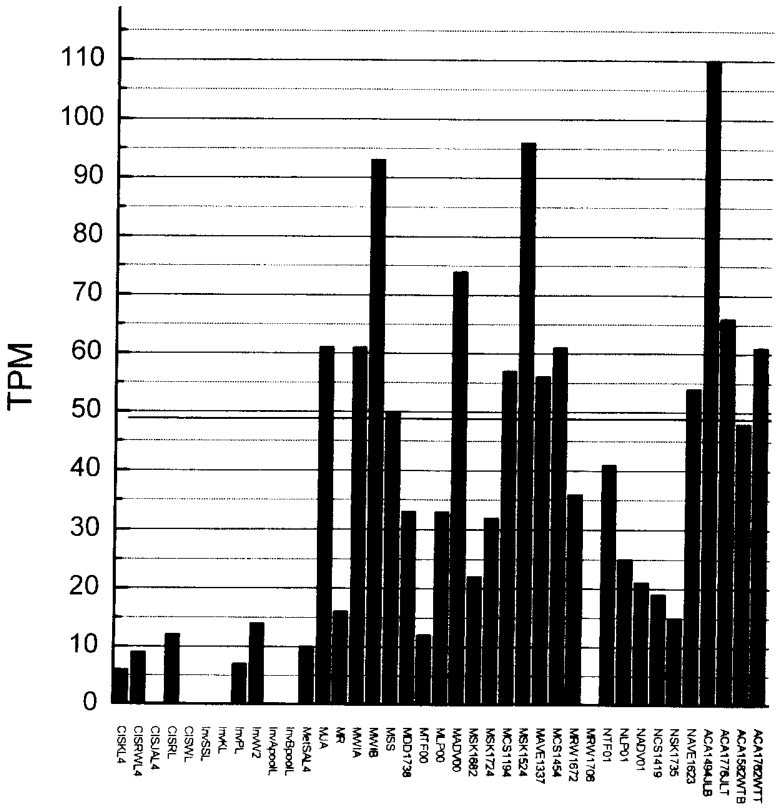
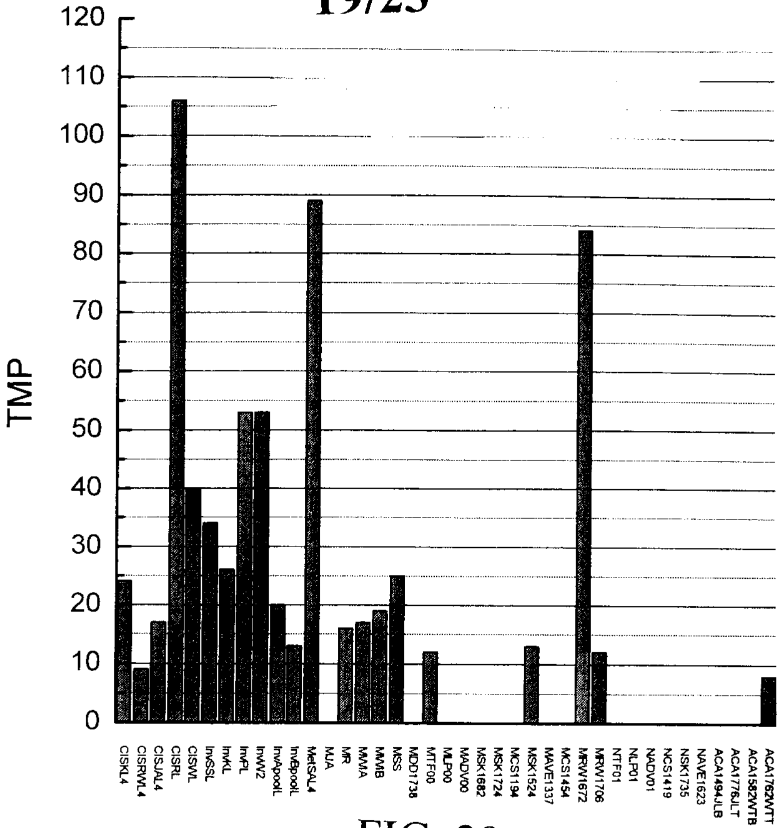
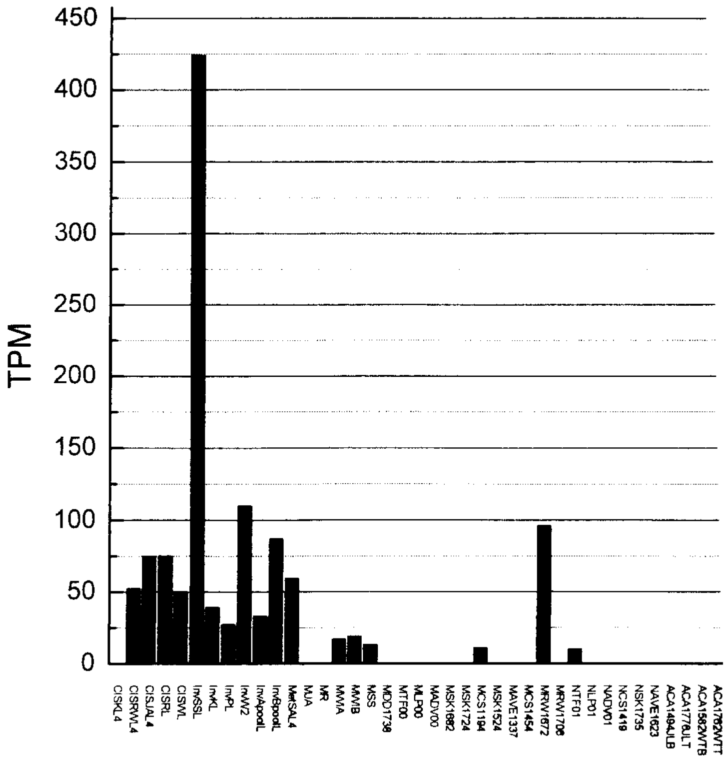

TPM
0
10
20
30
40
50
60
70
80
90
100
110
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1782WTT

FIG. 28



120
110
100
90
80
70
60
50
40
30
20
10
0
TMP
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApodL
InvBpodL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 29



450
400
350
300
250
200
150
100
50
0
TPM
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApodL
InvBpodL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 30

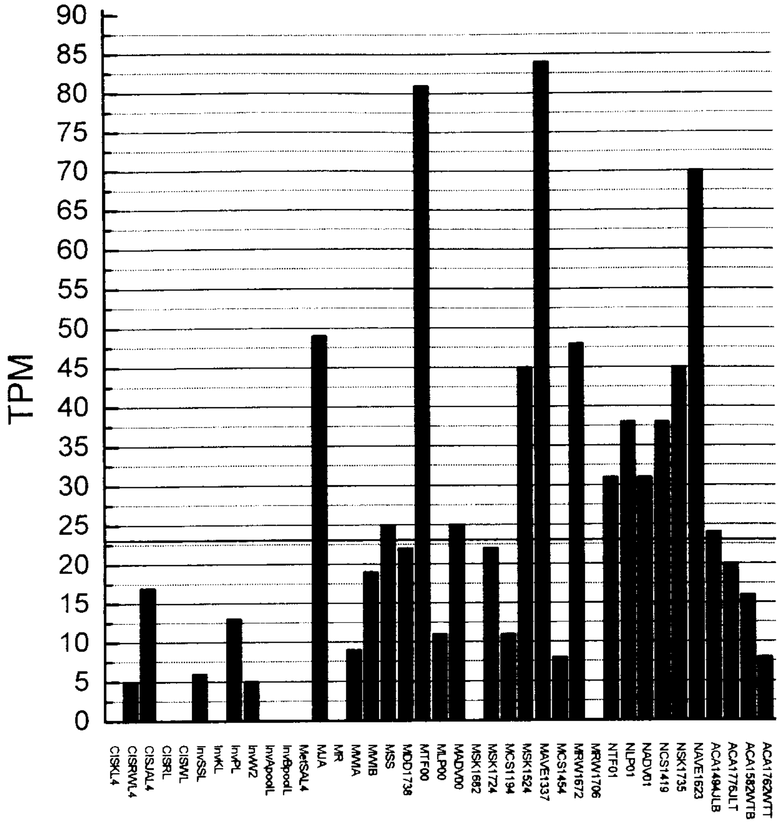

TPM
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 31

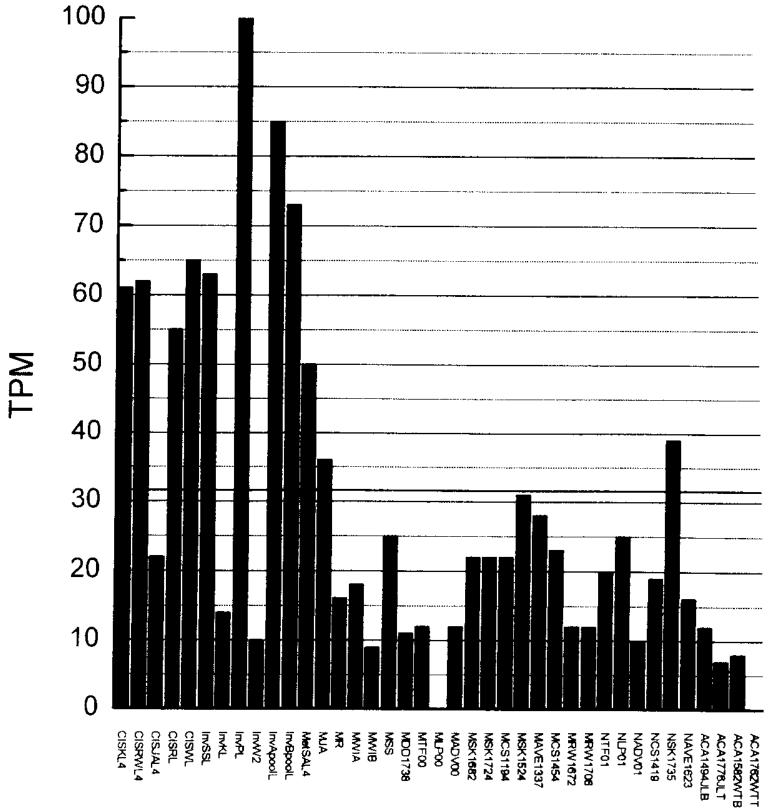

TPM
100
90
80
70
60
50
40
30
20
10
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 32

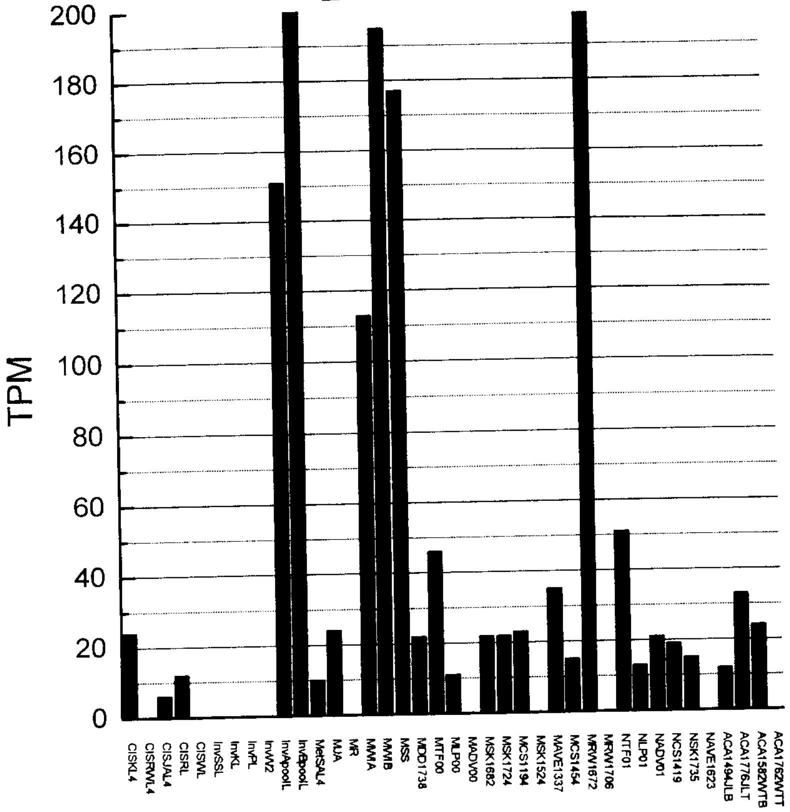

TPM
200
180
160
140
120
100
80
60
40
20
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 33

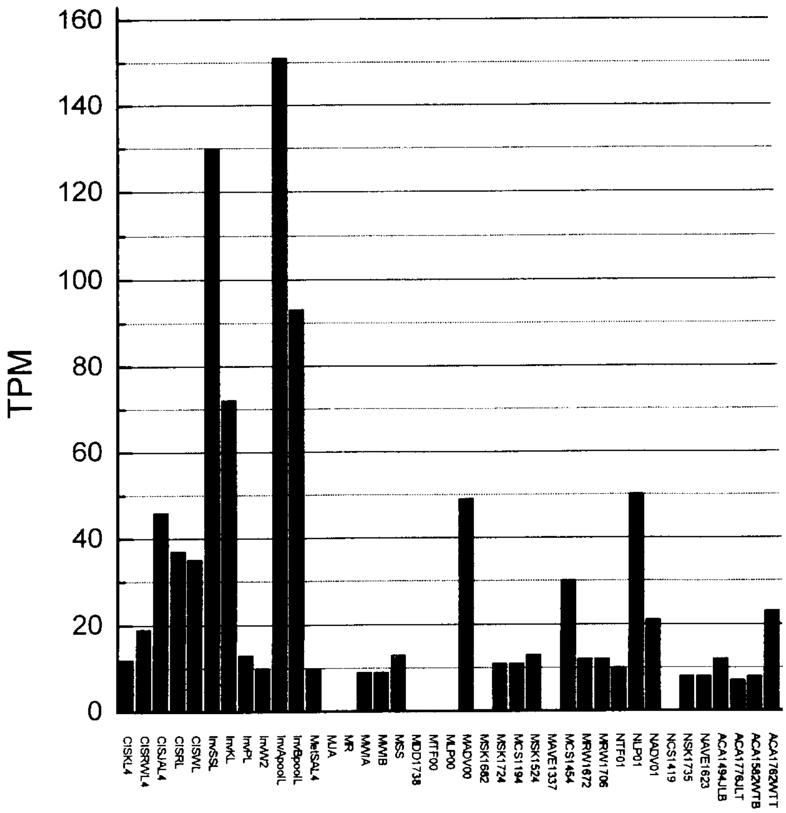

TPM
160
140
120
100
80
60
40
20
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWA
MWB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 34

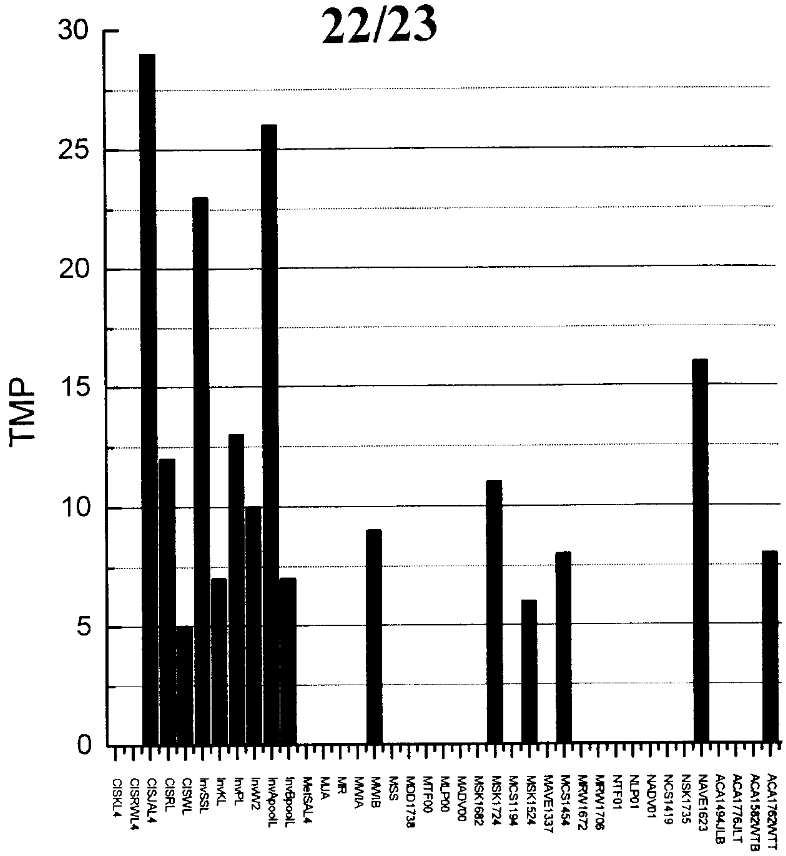

TMP
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 35

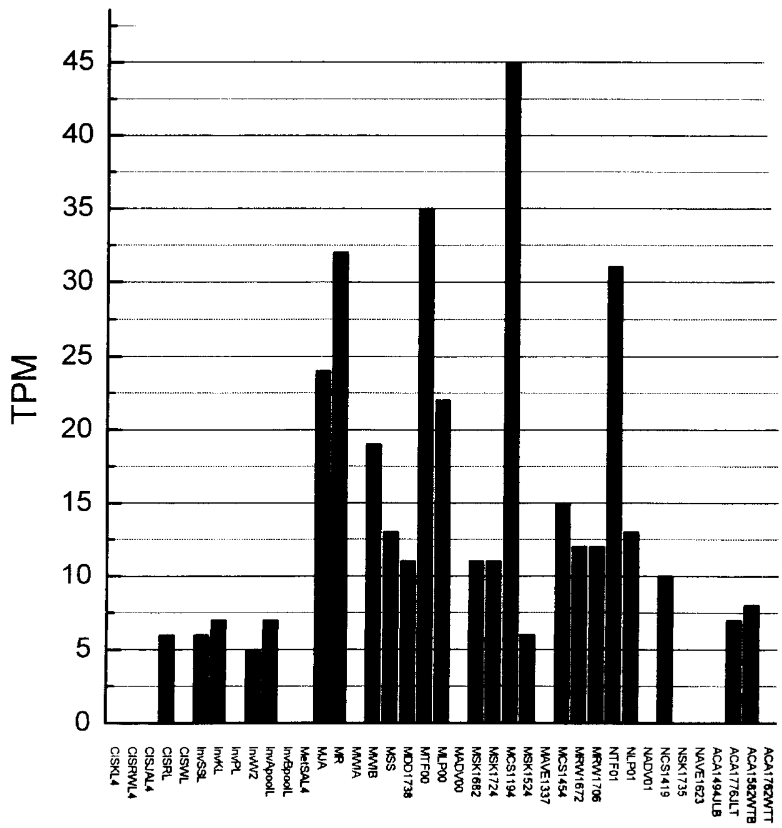

TPM
45
40
35
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MWIA
MWIB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 36

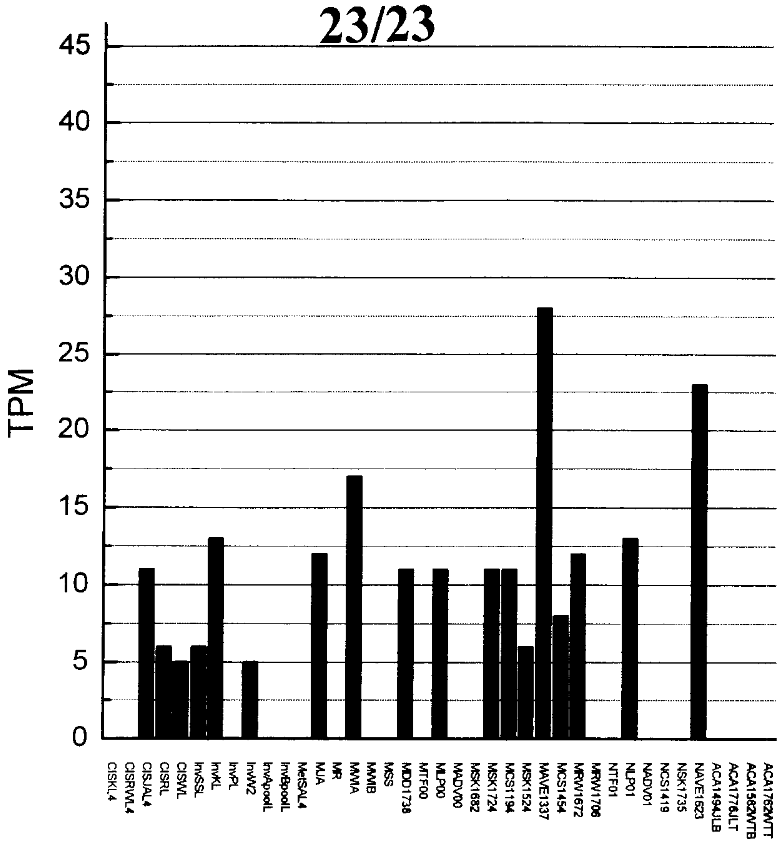

TPM
45
40
35
30
25
20
15
10
5
0
CISKL4
CISRWL4
CISJAL4
CISRL
CISWL
InvSSL
InvKL
InvPL
InvW2
InvApoolL
InvBpoolL
MetSAL4
MJA
MR
MMA
MMB
MSS
MDD1738
MTF00
MLP00
MADV00
MSK1682
MSK1724
MCS1194
MSK1524
MAVE1337
MCS1454
MRW1672
MRW1706
NTF01
NLP01
NADV01
NCS1419
NSK1735
NAVE1623
ACA1494JLB
ACA1776JLT
ACA1582WTB
ACA1762WTT

FIG. 37

INTERNATIONAL SEARCH REPORT	International application No. PCT/CA2006/000497

A. CLASSIFICATION OF SUBJECT MATTER
IPC: ***C12Q 1/68*** (2006.01) , ***C12Q 1/48*** (2006.01) , ***C12Q 1/14*** (2006.01)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC: ***C12Q 1/68*** (2006.01) , ***C12Q 1/48*** (2006.01) , ***C12Q 1/14*** (2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)
WEST; Delphion; Canadian Patent Database; Scopus; Caplus; Keywords: lung cancer; differential expression ; biomarker(s); expression profile; array(s); diagnos(is, ing); prognos(is, ing)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO2004/031413 A2 (ONCOTHERAPY SCIENCE, INC. (JP); JAPAN as represented by THE PRESIDENT OF THE UNIVERSITY OF TOKYO (JP)) 15 April 2004 - claims; pages 52-56; Tables 1 and 2 (pages 21, 40 and 41)	1-25
X A	WO2005/000098 A2 (THE TRUSTEES OF BOSTON UNIVERSITY (US)) 6 January 2005 - claims 16-52	1, 5-14, 17-20, 25 2-4, 15, 16, 21-24
X A	US2003/0175704 A1 (LASEK, A.K.W. et al (US)) 18 September 2003 - paragraphs 0008, 0009, 0088-0092; claims 8-10; Table 1 (Seq ID 48)	1, 5, 6, 8, 12, 13, 22, 23, 25 2-4, 7, 9-11, 14-21, 24

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents :
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 August 2006 (11-08-2006)	21 August 2006 (21-08-2006)

Name and mailing address of the ISA/CA	Authorized officer
Canadian Intellectual Property Office Place du Portage I, C114 - 1st Floor, Box PCT 50 Victoria Street Gatineau, Quebec K1A 0C9 Facsimile No.: 001(819)953-2476	Michael W. De Vouge (819) 997-2952

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2006/000497

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	WO2004/111603 A2 (GENOMIC HEALTH, INC. (US)) 23 December 2004 - claims 1-7, 27, 32-45	1-3, 5-8, 12, 13, 17, 19, 21-23, 25 4, 9-11, 14-16, 18, 20, 24
X A	US2004/0063120 A1 (THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)) 1 April 2004 - paragraphs 0018, 0149-0170, 0186 and 0187	1-13, 21-25 14-20
X A	WO00/61612 A2 (CORIXA CORPORATION, INC. (US)) 19 October 2000 - claims 40-57; page 11, line 4; page 16, lines 6-7;	1-6, 8-13, 17, 19, 22, 23, 25 7, 14-16, 18, 20, 21, 24

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CA2006/000497

Patent Document Cited in Search Report	Publication Date	Patent Family Member(s)	Publication Date
WO2004031413 A2	15-04-2004	AU2003272088 A1	23-04-2004
		CA2500151 A1	15-04-2004
		CN1705753 A	07-12-2005
		EP1551998 A2	13-07-2005
		JP2006500949T T	12-01-2006
		US2006024692 A1	02-02-2006
		WO2005089735 A2	29-09-2005
		WO2005090991 A1	29-09-2005
WO2005000098 A2	06-01-2005	CA2528572 A1	06-01-2005
		EP1633892 A2	15-03-2006
		US2006154278 A1	13-07-2006
		WO2004111197 A2	23-12-2004
US2003175704 A1	18-09-2003	NONE	
WO2004111603 A2	23-12-2004	AU2004248120 A1	23-12-2004
		CA2527285 A1	23-12-2004
		EP1631689 A2	08-03-2006
		US2005064455 A1	24-03-2005
US2004063120 A1	01-04-2004	AU2003253860 A1	23-01-2004
		WO2004005891 A2	15-01-2004
WO0061612 A2	19-10-2000	AU2894002 A	24-06-2002
		AU3094999 A	11-10-1999
		AU4069200 A	14-11-2000
		AU7314901 A	08-01-2002
		AU2003230849 A1	27-10-2003
		BR0009505 A	11-06-2002
		BR9908823 A	21-11-2000
		CA2323093 A1	23-09-1999
		CA2369578 A1	19-10-2000
		CA2537555 A1	17-02-2005
		CN1299414 A	13-06-2001
		CN1370182 A	18-09-2002
		CZ20013527 A3	16-10-2002
		EA5140 B1	30-12-2004
		EP0997970 A1	03-05-2000
		EP1064372 A2	03-01-2001
		EP1169347 A2	09-01-2002
		EP1319069 A2	18-06-2003
		EP1351967 A2	15-10-2003
		EP1496743 A2	19-01-2005
		EP1646869 A2	19-04-2006
		EP1675213 A1	28-06-2006
		HU0200760 A2	29-06-2002

...continued on **Extra Sheet**

... continued from **Patent Family Annex**

WO0061612 A2	19-10-2000	IL138330D D0	31-10-2001
		IL145731D D0	25-07-2002
		JP3466408B2 B2	10-11-2003
		JP10215116 A	11-08-1998
		JP2002533056T T	08-10-2002
		JP2002543769T T	24-12-2002
		JP2004513615T T	13-05-2004
		JP2004523222T T	05-08-2004
		JP2005522209T T	28-07-2005
		MXPA01009962 A	20-08-2002
		NO20004631 A	15-11-2000
		NZ506699 A	19-12-2003
		NZ514818 A	30-04-2004
		PL354348 A1	12-01-2004
		US6210883 B1	03-04-2001
		US6259411 B1	10-07-2001
		US6271804 B1	07-08-2001
		US6312695 B1	06-11-2001
		US6426072 B1	30-07-2002
		US6482597 B1	19-11-2002
		US6518256 B1	11-02-2003
		US6531315 B1	11-03-2003
		US6660838 B1	09-12-2003
		US6696247 B2	24-02-2004
		US6706262 B1	16-03-2004
		US6737514 B1	18-05-2004
		US6821518 B1	23-11-2004
		US6960570 B2	01-11-2005
		US7049063 B2	23-05-2006
		US2002115139 A1	22-08-2002
		US2002147143 A1	10-10-2002
		US2003138438 A1	24-07-2003
		US2003236209 A1	25-12-2003
		US2004235072 A1	25-11-2004
		US2005142620 A1	30-06-2005
		US2005261166 A1	24-11-2005
		WO0200174 A2	03-01-2002
		WO0247534 A2	20-06-2002
		WO9833232 A1	30-07-1998
		WO9947674 A2	23-09-1999
		WO03086175 A2	23-10-2003
		WO2005013899 A2	17-02-2005
		WO2006031363 A2	23-03-2006
		ZA200004866 A	13-12-2001