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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 September 2012 (27.09.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/129488 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01) *G01N 33/574* (2006.01)

(21) International Application Number:
PCT/US2012/030312

(22) International Filing Date:
23 March 2012 (23.03.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/466,615 23 March 2011 (23.03.2011) US
61/537,130 21 September 2011 (21.09.2011) US

(71) Applicant *(for all designated States except US)*: **VIRGINIA COMMONWEALTH UNIVERSITY** [US/US]; 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MANJILI, Masoud, H.** [IR/US]; Virginia Commonwealth University, 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US). **KMIECIAK, Maciej** [PL/US]; Virginia Commonwealth University, 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US). **TOOR, Amir, A.** [US/US]; Virginia Commonwealth University, 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US). **IDOWU, Michael, O.** [NG/US]; Virginia Commonwealth University, 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US). **BEAR, Harry, D.** [US/US]; Virginia Commonwealth University, 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US). **PAYNE, Kyle, K.** [US/US]; Virginia Commonwealth University, 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US). **MARINCOLA, Francesco, M.** [US/US]; Virginia Commonwealth University, 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US). **WANG, Ena** [US/US]; Virginia Commonwealth University, 800 East Leigh Street, Suite 113, Richmond, VA 23219 (US).

(74) Agents: **WHITHAM, Michael, E.** et al.; Whitman, Curtis, Christofferson & Cook, P.C., 11491 Sunset Hills Road, Suite 340, Reston, VA 20190 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME,

[Continued on next page]

(54) Title: GENE SIGNATURES ASSOCIATED WITH REJECTION OR RECURRENCE OF CANCER

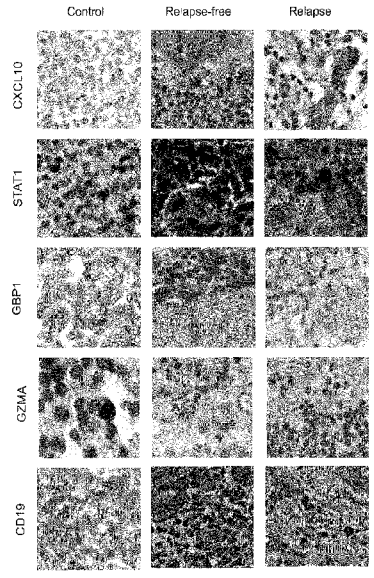


Figure 17A

(57) Abstract: Methods and tools for assessing the prognosis for patients following treatment of primary tumors are provided. The methods involve identifying immune- or cancer-related genetic markers whose differential expression patterns at tumor lesions are indicative of either tumor recurrence or recurrence-free survival. The methods and tools of the invention assist physicians by providing objective decision-making tools for planning patient treatment protocols.

WO 2012/129488 A2

MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

GENE SIGNATURES ASSOCIATED WITH REJECTION OR RECURRENCE OF CANCER

DESCRIPTION

BACKGROUND OF THE INVENTION

Field of the Invention

The invention generally relates to assessing the prognosis for cancer. In particular, the invention provides methods for identifying immune-related genetic markers whose expression patterns at tumor lesions are indicative of patient prognosis, e.g. patient outcome following initial treatment of the primary tumor.

Background of the Invention

Challenges in the immune therapy of cancers include a limited understanding of the requirements for tumor rejection and prevention of recurrences after successful therapy. Evaluation of T cell responses in human tumors, based predominantly on the metastatic melanoma model have clearly shown that the tumor bearing status primes systemic immune responses against tumor-associated antigens, which, however, are insufficient to induce tumor rejection (1, 2). Moreover, the experience gathered through the induction of tumor antigen-specific T cells by vaccines has shown that the presence of tumor antigen-specific T cells in the circulation (3, 4) or in the tumor microenvironment (5, 6) does not directly correlate with successful rejection or prevention of recurrence (7). Similarly, patients with pre-existing immune responses against HER-2/neu are not protected from the development of HER-2/neu expressing breast cancers (8). Although several and contrasting reasons have been proposed to explain this paradox, two lines of thoughts summarize these hypotheses; either tolerogenic and/or immune suppressive properties of tumors may hamper T cell function (9-11) or characteristics of the tumor microenvironment could induce tumor escape and evade the anti-tumor function of effector T cells (12, 13).

In spite of this paradoxical coexistence of tumor specific T cells and their target antigen-bearing cancer cells, sporadic observations in cancer patients suggest that T cells control tumor growth and mediate its rejection. Galon J et al. and others (14-16) observed that T cells modulate the growth of human colon cancer and T cell infiltration of primary

lesions may forecast a better prognosis. In addition, these authors observed that tumor-infiltrating T cells in cancers with good prognosis displayed transcriptional signatures typical of activated T cells such as the expression of interferon stimulated genes (ISGs), IFN-γ itself and cytotoxic molecules, in particular granzyme-B (15). Similar observations were reported by others in human ovarian carcinoma (17). Spontaneous regression of melanoma has also been reported to be mediated by complex and systemic immune responses rather than by T cell or antibody responses alone (18). These important observations derived from human tissues suggest a prognostic value for distinct signature of the immune function genes but cannot address the causality of the association between T cell infiltration and natural history of cancer. Recent reports based on adoptive transfer of tumor-specific T cells suggest a cause-effect relationship between the administration of T cells and tumor rejection (19). However, the complexity of the therapy associated with adoptive transfer of T cells which includes immune ablation and systemic administration of IL-2 prevents a clear interpretation of this causality.

The prior art has thus far failed to provide reliable methods to characterize successful and unsuccessful immune system responses to the presence of a tumor, and to accurately and consistently establish a prognosis prior to or during treatment of primary tumors. This is important because an accurate prognosis is extremely valuable in assessing treatment options for the patient.

SUMMARY OF THE INVENTION

The present invention introduces methods for the analysis of tumor tissue and tumor tissue microenvironments in order to assess the prognosis of patients with carcinomas, e.g. predict the success or failure of treatments, and/or of relapse after initial treatment. The invention is based on the identification of immune-related genetic markers whose expression patterns at tumor lesions can be used to predict whether or not the patient is mounting an effective immune response against the tumor that is likely to reject residual or recurring tumors, especially after intital or standard treatments such as surgery, radio- and chemotherapy, etc. In other words, the pattern of gene expression is indicative of the likelihood of relapse or recurrence of the cancer after treatment. In one embodiment, 299 genes have been identified which, when upregulated, are associated with a low risk of relapse, i.e. with a high probability of a relapse-free recovery (see Table 4). In another

embodiment, 50 genes have been identified which, when upregulated, are associated with a high risk of relapse, i.e. with a high probability of recurrence of the cancer. In yet another embodiment, 5 immune function genes have been identified and 8 cancer testis antigen (CTA) genes have been identified which, when upregulated, are associated with a low probability of relapse; conversely the absence of upregulation of the set of 5 and 8 genes is indicative of a high probability of relapse. The ability to identify these patterns or signatures of gene expression permit health practitioners to optimize treatment protocols for cancer patients. For example, a patient that is categorized as unlikely to experience a relapse need not be subjected to aggressive measures which can be traumatic and debilitating in and of themselves. On the other hand, patients identified as likely to experience a recurrence of the cancer can be aggressively treated in order to provide the best possible chance of long-term survival. As such, the methods of the invention are theranostic methods and constitute a personalized medicine approach to cancer treatment, relying on pharmacogenomics, molecular biology, microarray chip technology, etc. An embodiment of the invention therefore provides an objective decision-making tool for physicians regarding how to treat patients with, for example, ductal carcinoma *in situ* (DCIS), breast cancer, or other invasive carcinomas. An embodiment of the invention also provides kits containing ready-to-use microarray chips, two tier computer software data analysis and statistical methods for determining efficacious treatment options commensurate with the prognoses that are provided. An embodiment of the invention also provides a quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) kit containing 8 human CTA as well as 5 immune function genes. A panel of 5 housekeeping genes will be used for normalization of the data.

The invention provides an in vitro method for determining, in a cancer patient in need thereof, the likelihood of relapse. The method comprises the steps of i) obtaining a tumor tissue sample from said cancer patient; ii) quantifying a level of gene expression in said tumor tissue sample of at least one gene in a gene set, said gene set comprising at least one of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb; iii) comparing a quantification value for a level of gene expression of at least one of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb obtained in said quantifying step with a predetermined reference value for a level of gene expression of at least one of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5,

MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb in control tissue samples; and iv) providing a prognosis of a low likelihood of relapse for said patient when said quantification value is greater than said predetermined reference value; or v) providing a prognosis of a high likelihood of relapse for said patient when said quantification value is lower than said predetermined reference value. In one embodiment, the cancer is breast cancer. In another embodiment, the at least one gene includes the following genes: IGKC, IGLL5, STAT1, GBP1, and OCLN. In yet another embodiment, the at least one gene includes the following genes: MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb. In other embodiments, the at least one gene includes one or more housekeeping genes. In one embodiment of the invention, the control tissue samples include tissue samples from one or more of subjects without cancer, subject with stage I cancer, subjects with stage II cancer, subjects with stage III cancer, subjects with stage IV cancer, subject who have not relapsed after receiving conventional cancer treatment, and subjects who have relapsed after receiving conventional cancer treatment.

The invention also provides a theranostic method for developing a treatment protocol for a cancer patient. The method comprises the steps of i) obtaining a tumor tissue sample from said patient; ii) quantifying a level of gene expression of at least one of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb in said tumor tissue sample; iii) comparing a quantification value for a level of gene expression of at least one of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb obtained in said quantifying step with a predetermined reference value for a level of gene expression of at least one of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb in control tissue samples; and iv) when said quantification value is greater than said predetermined reference value, providing a prognosis of a low likelihood of relapse for said patient and recommending a tumor-burden reducing treatment for said patient; or v) when said quantification value is lower than said predetermined reference value, providing a prognosis of a high likelihood of relapse for said patient and recommending neoadjuvant therapy in conjunction with a tumor-burden reducing treatment for said patient. In one embodiment, the tumor-burden reducing treatment includes one or more treatments selected from the group consisting of surgical removal of tumor tissue, reduction in tumor volume by chemotherapy,

reduction in tumor volume by radiotherapy, and reduction in tumor volume by hormone therapy. In one embodiment, the neoadjuvant therapy includes administration of one more agents selected from the group consisting of 5-azacytidine, decitabine, histone deacetylation inhibitors.

The invention provides a system for determining a probability of relapse of a patient with a tumor. The system comprises: 1) means for obtaining measurements of expression of genes in tumors; 2) means for recognizing, using said measurements, patterns of gene expression, wherein said patterns of gene expression are correlated with said probability of relapse; and 3) means for assigning a probability of relapse to said patient with said tumor. wherein said genes comprise one or more of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb.

The invention further provides a microarray chip for analyzing the likelihood of relapse of a patient with a tumor, the microarray chip comprising primers specific for amplifying RNA corresponding to at least one gene selected from the group consisting of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb. In one embodiment, the at least one gene includes IGKC, IGLL5, STAT1, GBP1, and OCLN. In another embodiment, the at least one gene includes MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb. The microarray chip may include all 13 genes and may also include housekeeping genes.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A and B. T cells derived from wild-type FVB mice will induce apoptosis in MMC in vitro but fail to reject MMC in FVBN202 mice following AIT. A) Flow cytometry analysis of MMC after 24 hrs culture with splenocytes of FVB mice following three color staining. Gated neu positive cells were analyzed for the detection of annexin V+ and PI+ apoptotic cells. Data are representative of quadruplicate experiments. B) Donor T cells were enriched from the spleen of FVB mice using nylon wool column following the rejection of MMC. FVBN202 mice (n=4) were injected with CYP followed by inoculation with MMC (4 x 106 cells/mouse), and tail vein injection of donor T cells. Control groups were challenged with MMC in the presence or absence of CYP treatment. Tumor growth was monitored

twice weekly.

Figure 2A-C. Gene expression profiling and gene oncology pathway analyses in tumor regressing and tumor non-regressing groups. (A) Unsupervised cluster visualization of genes differentially expressed among regressing tumors (lanes 9-12) and non regressing tumors (lanes 1-8 = evasion model; lanes 13-18 = tolerogenic model). MMC tumors were harvested 10 days after challenge and hybridized to 36k oligo mouse arrays. 11256 genes with at least 3-fold ratio change and 80% presence call among all samples were projected using log2 intensity. (B) Supervised cluster analysis (Student t test, $p< 0.001$ and fold change >3) comparing regressing tumors (lanes 9-12) and non regressing tumors (lanes 1-8 = evasion model; lanes 13-18 = tolerogenic model). 2449 differentially expressed genes have been selected for further analysis. (C) Gene Ontology databank was queried to assign genes to functional categories and upregulated genes within the tumor regression group to functional categories. A = cytokine-receptor interaction; B, neuroactive ligand-receptor interaction; C = mitogen-activated protein kinase (MAPK) signaling pathway; D = regulation of actin cytoskeleton; E = cell adhesion molecules; F = natural killer mediated cytotoxicity; G = axon guidance; H = calcium signaling pathway; I = T cell receptor signaling pathway; J = insulin signaling pathway; K = Janus kinases, signal transducers and activators of transcription (JAK-STAT) signaling pathway; L = leukocyte transendothelial migration; M = Toll-like receptor signaling pathway.

Figure 3A-C. Gene expression profiling and gene oncology pathway analyses in tolerance and evasion models. (A) Supervised cluster analysis (Student t test, $p< 0.001$ and fold change >3) comparing evasion (Lanes 1-8) and tolerogenic group (Lanes 13-18). 1326 differentially expressed genes have been visualized including also tumor regression samples (Lanes 9-12). Gene ontology pathway analysis projecting either upregulated pathways in the evasion group (B, 854 genes, where A = focal adhesion; B = MAPK signaling; C = regulation of actin cytoskeleton; D = cell cycle; E = cytokine-receptor interaction; F = leukocyte transendothelial migration; G = axon guidance; H = small cell lung cancer; I = gap junction; J = p53 signaling pathway; K = calcium signaling pathway; L = cell communication; and M = glycan structures- biosynthesis); 1) or tolerogenic (immune suppression) tumor models (C, 475 genes, where A = cell communication; B = cell adhesion molecules, CAMs; C = insulin signaling pathway; D = cytokine-receptor interaction; E = extracellular matrix (ECM) receptor interaction; F = focal adhesion; G = tight junction; H =

glycan structures- biosynthesis 1; I = peroxisome proliferator-activated receptors (PPAR) signaling pathway; J = glutathione metabolism; K = glycan structures- biosynthesis 2; L = glycolysis/gluconeogenesis; and M = JAK-STAT signaling pathway).

Figure 4A-C. Representation in tabular form of : A,chemokines and their receptors and interferon stimulating genes differentially expressed in rejection model vs control; B, cytokines and signaling molecules (interleukins and receptors, cytotoxic and pro-apoptotic molecules, Toll-like receptors and lymphocyte signaling, FC-type receptors and immunoglobulins) differentially expressed in rejection model vs control; C, genes with immunological function (chemokines, interleukins and signaling genes, and ISGs) manually selected based on supervised comparison of evasion and tolerogenic (immune suppressed) tumor models and tumor rejection model.

Figure 5A and B are high level flow diagrams of processes of this invention implemented on a computer.

Figure 6. Schematic representation of the system of the invention.

Figure 7. Results of unsupervised clustering of gene expression of 9797 genes from tumor samples of human breast cancer patients. These genes exhibit at least a 3-fold ratio in change and 80% presence (average corrected) compared to control samples.

Figure 8. Student's T test comprising relapse vs relapse free ($P<0.001$) (80% corrected) of 349 genes from tumor samples of human breast cancer patients.

Figure 9A-N. Listing, as Table 4, of 299 genes, the upregulation of which is associated with decreased occurrence of human breast cancer relapse after treatment.

Figure 10A-C. Listing, as Table 5, of 50 genes, the upregulation of which is associated with increased occurrence of human breast cancer relapse after treatment.

Figure 11A and B. Significant canonical pathway analysis of immune system related pathways involved in breast cancer relapse or resistance to relapse. Solid bars = -log p value of the significance for genes upregulated in tumor lesions of patients who are relapse free vs relapsed patients, with cutoff of the significance $P<0.001$ (dotted line); ■'s connected by solid lines = ratio of number of genes in relapse free vs relapsed patients. Genes which inhibit effector immune responses are underlined.

Figures 12A-E. Immune system pathways identified as involved in cancer relapse. A, B cell development pathway; B, antigen presentation pathway; C, graft vs host disease signaling; D, interferon signaling; and E, primary immunodeficiency signaling. The degree

of gray shading of individual pathway components indicates the relative level of upregulation, with darker shading corresponding to a higher level of upregulation.

Figure 13A and B. Unsupervised gene clustering. A) Unsupervised cluster visualization of genes differentially expressed among relapse (n=8) and relapse free (n=9) patients. Tumors were hybridized to 36k oligo human array. Genes with at least 80% presence among all samples (9797) were projected using log2 intensity. Over-expression; under-expression; unchanged expression; and no detection of expression are indicated (intensity of both Cy3 and Cy5 below the cutoff value). Each row represents a single gene; each column represents a single sample. The dendrogram at the left of matrix indicates the degree of similarity among the genes examined by expression patterns. The dendrogram at the top of the matrix indicates the degree of similarity between samples. B) Multiple dimensional scaling based on the 36 k oligo array human platform comparing Relapse free (black circles) and relapse (gray circles)

Figure 14. Unsupervised gene cluster analysis. Five genes selected from the 299 genes by Complete Leave-One-Out Cross Validation (LOOCV) model as best predictors of diagnostic outcome. Black dots under the cluster indicate relapse free and underlined dots indicate relapse group.

Figure 15. Ingenuity pathway analysis. Forty six canonical pathways significant at the nominal 0.001 level of the unpaired Student's *t* test. The *p* value for each pathway is indicated by the bar and is expressed as -1 times the log of the *p* value. The line represents the ratio of the number of genes in a given pathway that meet the cutoff criteria divided by the total number of genes that make up that pathway.

Figure 16. qRT- PCR analysis of frozen tumor specimens of relapse-free vs. relapse patients. Two cohorts of patients were included in the validation group and their tumors were subjected to confirmatory qRT-PCR. Data are presented as average of mean of triplicate wells after normalization to GAPDH.

Figure 17A and B. IHC analysis of paraffin-embedded tumor specimens of relapse-free vs. relapse patients. A) Representative data (400X magnification) from 9 patients with relapse-free survival and 8 patients with relapse are presented. Human tonsil was stained as positive control. B) Cell counts are presented as percent positive cells of tumor infiltrating cells counted in five fields and averaged using a 400X magnification.

Figure 18A and B. qRT-PCR analysis of cancer testis antigen RNA extracted from tumor

lesions. A, patients who relapsed within 1-3 years or remained relapse-free for 4-5 years; and B, patients who relapsed within 5-6 years or remained relapse-free for 6-7 years (Figure 18B).

Figure 19. Nucleotide sequence encoding MAGE-A3 (SEQ ID NO: 1).

Figure 20. Nucleotide sequence encoding MAGE-A4 (SEQ ID NO: 2).

Figure 21. Nucleotide sequence encoding MAGE-A5 (SEQ ID NO: 3).

Figure 22. Nucleotide sequence encoding MAGE-A6 (SEQ ID NO: 4).

Figure 23. Nucleotide sequence encoding MAGE-C1 (SEQ ID NO: 5).

Figure 24. Nucleotide sequence encoding AKAP4 (SEQ ID NO: 6).

Figure 25. Nucleotide sequence encoding NY-ESO-1 (SEQ ID NO: 7).

Figure 26. Nucleotide sequence encoding SLLP1 (SEQ ID NO: 8).

Figure 27. Nucleotide sequence encoding SP17 (SEQ ID NO: 9).

Figure 28. Nucleotide sequence encoding SPANXb (SEQ ID NO: 10).

Figure 29. Nucleotide sequence encoding IGKC (SEQ ID NO: 11).

Figure 30. Nucleotide sequence encoding IGLL (SEQ ID NO: 12).

Figure 31. Nucleotide sequence encoding OCLN (SEQ ID NO: 13).

Figure 32. Nucleotide sequence encoding STAT1 (SEQ ID NO: 14).

Figure 33. Nucleotide sequence encoding GBP-1 (SEQ ID NO: 15).

DETAILED DESCRIPTION

The invention provides real-time identification of genetic "signatures" at tumor lesions that can be used to predict future tumor rejection and/or lack of recurrence, or failure in tumor rejection, and likely recurrence. The elucidation of differential patterns of immune system gene expression as described herein permit the classification of patients into either the category of patients who are likely to have a recurrence of the tumor, or the category of patients who are not likely to have a recurrence. 299 genes (listed in Table 4) have been identified as upregulated in breast cancer patients who do not experience recurrence after initial cancer treatments, and 50 genes (listed in Table 5) have been identified as upregulated in breast cancer patients who do experience recurrence after initial cancer treatments. In addition, further signatures of 5- and 8- genes (or of 13 when the two sets are combined) have been discovered, which, when upregulated, are associated with a positive ability of a patient to reject cancer cells (e.g. metastatic or residual cancer cells), but which, when not

upregulated, are associated with an inability to reject cancer cells and thus a high likelihood of relapse. An analysis which measures mRNA expression of these genes thus permits recognition of patterns of gene expression (also referred to as genetic or transcription "signatures") associated with increased or lowered risks of relapse, and can serve as a guide for health practitioners when prescribing treatment for the patient.

Thus, in one embodiment, it was found that gene expression patterns in individuals that could successfully reject tumor recurrence showed differential expression of about 349 genes. Of these, expression of 299 genes (listed in Table 4) was increased (upregulated) in patients who did not experience relapse whereas expression of 50 genes (listed in Table 5) was donwregulated in these patients, compared to normal control values. Conversely, expression of the 50 genes was upregulated in patients who did relapse and expression of the 299 genes was downregulated, compared to normal control gene expression levels.

In another embodiment, it was found that gene expression patterns in individuals that could successfully reject tumor recurrence showed differential expression (upregulation) of the 5 immune function genes IGKC (the locus of which is IGK@), IGLL5, STAT1, GBP1 and OCLN, whereas the absence of expression of the 5 genes was associated with a high probability of relapse. In another embodiment, expression (upregulation) of the 8 CTA genes MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb was found to be associated with a low probability of relapse, whereas the absence of expression was found to be associated with a high probability of relapse.

The patterns of gene expression for populations of patients who are likely to relapse vs those who are unlikely to do so are different or distinct from each other, and from normal control patterns, and the patterns of gene expression may be referred to herein as "differential". By "relapse" we mean that the patient, after completing initial (customary, conventional, etc.) cancer treatment (e.g. surgical tumor removal, radiation therapy, chemotherapy, etc.) and usually after being declared generally free of cancer, experiences a regrowth or reappearance of the tumor, either at the same location, or at a different location (i.e. metastatic spread of the tumor) usually within 1-5 years of completing cancer treatment. Recurrence may be due to the development of new tumor cells arising during or after treatment, or the persistence of residual tumor cells which escape the treatment that was provided.

The analysis of the invention may be carried out at any time during the life of a

cancer patient, e.g. soon after diagnosis of a tumor as malignant, and prior to initial treatment, since the results provide useful information to the clinicians who develop treatment protocols. A patient who is determined to be at high risk for recurrence would generally be treated more aggressively than would a patient who is identified as at low risk for recurrence. However, the test described herein can be carried out at any point, e.g. at any time during treatment or at any time after treatment or neoadjuvant therapy as long as tumor sampling is feasible. The analysis may be carried out multiple times for a patient, e.g. in order to check whether or not the gene expression status of the patient is constant, or to monitor the status of gene expression, etc.

The numerical cutoffs or guidelines for assigning a patient to a low-risk vs high risk group with respect to relapse is as follows: identification of a patient as being at high risk for recurrence is indicated when the patient's gene expression profile falls in a group of 50 genes upregulated ($p<0.001$) in the reference relapsed group (Table 5). In other words, in a high risk patient, canonical pathway analysis would not show upregulation of the five identified immune system pathways depicted in Figures 12 A-E, but would show downregulation of the 299 genes listed in Table 4. Conversely, identification of a patient as being at low risk for recurrence is indicated when the individual's gene expression profile falls in a group of 299 genes upregulated ($p<0.001$) in the reference relapse-free group. In other words, in a low risk patient, canonical pathway analysis would show upregulation of the five identified immune system pathways, as well as downregulation of the 50 genes listed in Table 5. The same holds true for the gene in the 5- and 8-gene signatures, whether used separately or in combination. Generally, the level of up- or downregulation is at a level which is at least 10%, 20%, 30%, 40%, 50%, 60%, 60%, 70%, 80%, 90%, or even 100% higher or lower, respectively, than that of the reference control, and may be even 2, 3, 4, 5, 6, 7, 8, 9, or 10-fold higher or lower, respectively, or even greater, e.g. 20, 30, 40, 50, 60, 70, 80, 90, 100, or more (e.g. 150, 200, 250, 500, 750, 1000, etc.) fold higher or lower, respectively, than the level of expression of a suitable control sample(s). These gene signatures/pathways were validated by detection of the selected genes by qRT-PCR and Immunohistochemistry (IHC). Herein, the tendency of a person to experience relapse may be referred to as the probability, likelihood, risk chance, etc. of a recurrence or relapse.

Many of the differentially expressed genes identified herein are associated with immune regulatory functions. For example, the identified genes are generally immune

system genes that are part of or are associated with an immune system pathway such as B cell development, antigen presentation, graft vs host disease signaling, interferon signaling and primary immunodeficiency signaling.

In some embodiments, all 349 genes are analyzed in an assay, although this need not always be the case. The invention also encompasses assays in which fewer than the 349 are analyzed, or in which more than the 349 are analyzed. However, generally gene expression of at least one gene from each of the two categories (low risk, Table 4, and high risk, Table 5) is determined and the levels are compared to each other, and to the level of expression in normal control non-tumor peripheral blood mononuclear cells (PBMC). In some embodiments, at least about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270 280, 290, 300, 310, 320, 330, or 340 total genes are included; with at least about 10, 15, 20, 25, 30, 35, 40, 45, or all 50 genes from the group of Table 5; and at least about 5,10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60,65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, or all 299 genes of the group of Table 4; or any desirable number of genes from each group.

The same is true for the 5- and 8-gene signatures, where from 1-5 of the genes in the 5-gene signature and/or from 1-8 of the genes in the 8-gene signature may be used together e.g. from about 1 up to about 13 of the genes (e.g. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 of the genes from either group) may be used in combination as a diagnostic to assess a patient's likelihood of relapse. The method to measure the expression of a 13 gene signature will typically be qRT-PCR using either frozen or paraffin-embedded tumor specimens.

Those of skill in the art will recognize that tumor samples (usually from solid tumors) generally contain both tumor cells and cells from the host, e.g. cells from the host immune system, blood vessels, etc. that have invaded the tumors. The "microenvironment" of a tumor, as used herein, includes such host-derived cells. As used herein, a "tumor sample" is understood to include cells from the microenvironment of the tumor and/or from the tumor itself.

Many methods for determining patterns or signatures of gene expression (e.g. differential gene expression) are known. Preferred methods include using "chip" technology, e.g. microarrays of a nucleic acid (usually DNA, although RNA is also sometimes used).

DNA microarrays consist of an arrayed series of hundreds or thousands of oligonucleotide probes which hybridize to target nucleic acids (e.g. cDNA, cRNA, etc.) in a sample under high-stringency conditions. Probe-target hybridization is then detected and quantified by, e.g. fluorescence-based detection of fluorophore-labeled targets to determine relative abundance of nucleic acid sequences in the sample. In the practice of the present invention, such chip technologies may be used, and several commercially available generic chips are known which would be suitable, examples of which include but are not limited to the Affymetrix 0133+2 whole genome chip, and other chips designed for array analysis of tumor lesions using immunohistochemistry (IHC), immunofluorescence (IF), etc. Microarray chips may also be developed specifically for use in the invention. Such chips are designed to probe only selected genes of interest, such as immune system genes, or useful subsets thereof e.g. any combination of the genes described herein. Suitable controls would be included on such a specialty chip. However, those of skill in the art will understand that gene products (e.g. proteins, polypeptides, peptides) corresponding to the genes identified herein may also be detected and/or quantified, either as a primary method of determining relapse risk, or as confirmation of genetic analysis.

In one embodiment of the invention, a chip is developed specifically for use in the methods of the invention. Such a chip would include probes capable of hybridizing to one or more genes or RNA expressed from genes as described herein, i.e. genes listed in Tables 4 and 5, or the genes of the 5- and/or 8-gene signatures. In addition, various useful subsets of the genes may be represented on a chip, and all such possible subsets are intended to be encompassed by the present invention, i.e. genes ranging in number from 1 to 349. In other words, those of skill in the art will recognize that, while generally an entire genetic profile or signature of a sample as described herein will be determined, this need not always be the case. For example, given the information provided herein, it is also possible to select certain genes or small groups of genes and compare their expression on a less comprehensive basis. All such variations and permutations of the technology disclosed herein are intended to be encompassed by this invention.

For the practice of the invention, the experimental or unknown samples for which a genetic signature is obtained are generally tumor samples such as biopsy samples, or samples of a tumor that has been surgically removed from a patient. Procedures for obtaining such samples are generally carried out by skilled medical personnel such as physicians, surgeons,

etc. Likewise, the treatment and handling of tumor samples in order to extract nucleic acids such as RNA for analysis from the samples may vary somewhat from circumstance to circumstance, but such methods are generally known, e.g. homogenizing samples in the presence of various agents such as nuclease inhibitors in order to promote or preserve the stability of the mRNA, apportioning samples, purifying fractions, adding reagents such as enzymes, labeling agents, etc., which is generally carried out manually or in a partially automated manner by those skilled in laboratory procedures. Any and all such suitable methods may be used to prepare the sample for analysis.

The methods of the invention are practiced in order to predict the probability or likelihood or chance of relapse of a cancer patient after an initial treatment. By "initial treatment" we mean a standard or conventional treatment which removes or reduces the size of, i.e. which reduces the tumor burden of, a patient. Such tumor burden reducing treatment techniques include but are not limited to one or more of the following: surgery, chemotherapy, radiation therapy, hormone therapy, antibody therapy, T cell therapy, etc. Those of skill in the art will recognize that what is considered "conventional" or "frontline" or "standard" therapy may change with time as new techniques are developed and study results are interpreted. The methods of the invention are useful for predicting the likelihood that the patient will later have a recurrence or relapse of the cancer after the tumor burden-reducing treatment. Such relapses may be caused, for example, by residual tumor cells which were not removed or destroyed by the initial treatment, or by perpetuation of the conditions that allowed the tumor to develop in the patient so that new tumor cells arise after initial treatment(s). The methods of the invention make it possible to determine whether or not a patient has the ability (e.g. the intrinsic ability already in place) to destroy residual or newly developing cancer cells, once the majority of a primary tumor is no longer present. For example, if a patient does express the 5- and 8-gene patterns described herein, the prospect for doing so is high. If a patient does not express the 5- and 8-gene patterns described herein, then the prospect for doing so is low. In the latter case, a medical expert such as a physician will likely suggest or recommend that other more radical or far-reaching or even experimental therapies should also be undertaken. In other words, once the gene expression pattern of the patient is determined, it is possible to tailor the patient's treatment based on the results of the assessment, i.e. the treatment protocol of a patient may be adjusted to account for the tendency, or lack thereof, toward relapse (tumor recurrence, regrowth, or

redevelopment). For example, if the analysis suggests that the tumor is not likely to recur, non-aggressive treatment alternatives might include conservative surgery, lower frequency and duration of radiation and/or lower frequency of chemotherapy with a preference of low toxicity drugs. Alternatively, if the analysis suggests that the tumor is likely to recur, treatment with one or more of aggressive surgery, radiation and/or chemotherapy may be recommended. In addition, the use of neoadjuvant therapies prior to surgery and/or chemotherapy may be considered in order to convert a high risk patient into a low risk profile for relapse by means of the induction of the 5 genes and 8 genes. If neoadjuvant therapy is successful, treatment with surgery and less aggressive chemotherapy can be offered. Neoadjuvant therapies include but are not limited to 5-azacytidine, decitabine, histone deacetylation inhibitors, etc. Those of skill in the art will recognize that, in some circumstances, a "neoadjuvant therapy" may also be considered as a "standard" therapy (see the description of initial therapies provided above). However, in general, neoadjuvant therapy refers to the administration of therapeutic agents before a main treatment. Such therapies include, but are not limited to, immunotherapy, radiation therapy, chemotherapy, hormone therapy

In one embodiment, the neoadjuvant therapy is administration of decitabine. Decitabine is a demethylating pro-drug that has shown efficacy in patients with hematologic malignancies, particularly against myelodisplastic syndrom (MDS). Its efficacy has been attributed to the induction of tumor suppressor genes and CTA. Activation of decitabine by deoxycytidine kinase (DCK), which is selectively expressed in tumor cells and myeloid cells of some, but not all patients, leads to incorporation of decitabine into newly synthesized DNA strands during the S-phase of the cell-cycle. When a decitabine-containing DNA strand binds to the enzyme DNA methyltransferase (DNMT1), the decitabine in the strand forms a covalent complex with a serine residue at the DNMT1 active site, resulting in inactivation of the enzyme. This in turn results in hypomethylation of genes in the surrounding area. Accordingly, in one embodiment, the invention provides methods of converting a cancer patient with a high probability of relapse to a status of low relapse probablility by administering decitabine to the patient. Adminstration may be before or after administration or carrying out of other treatment modalities, and may involve one or multiple administrations of the drug. Technically, decitabine is a prodrug in that is must be activated

within the body via phosphorylation by the DCK enzyme. Thus, generally patients who are deemed eligible for decitabine are DCK positive, although this need not always be the case. Some forms of decitabine which do not require activation may exist or may be developed, and their use is contemplated herein. Alternatively, patients may be rendered DCK positive e.g. by the administration of gene therapy agents which cause expression of DCK. If a patient's tumor is negative for DCK, 5-azacytidine can be administered instead of decitabine in order to induce CTA expression in the tumor and in turn trigger an immune function gene signature in response to the CTA induction. Such expression may be systemic or may be targeted or localized to tumor cells. In some embodiments, neoadjuvant therapy may be combined with histone deacetylation inhibitors to achieve a sustained hypomethylation of genes. In some embodiments, the patient may be in an early stage of cancer; in other embodiments, the patient may have already relapsed, and the gene signature is determined e.g. for recurrent or metastatic tumors.

Immune responses to various types of cancers can be determined by the practice of the methods of the invention. Generally, such cancers will be of the types that form solid tumors, i.e. carcinomas. Examples of such cancers include but are not limited to breast cancer, ductal carcinoma in situ (DCIS), prostate cancer, stomach cancer, colon cancer, lung cancer, melanoma, and head and neck cancer, ovarian cancer, pancreatic cancer, etc. The tumors that are analyzed may be primary or secondary (metastasized) or recurring tumors, and the methods may be used to monitor the effects of treatments and patient progress.

Patients who may benefit from the analyses described herein are generally mammals, and may be humans, although that need not be the case. Veterinary applications of this technology are also contemplated.

Those of skill in the art will recognize that comparisons and analyses of gene expression patterns such as those described herein are generally automated to the extent possible, and are controlled by computer software analytical programs. The invention also provides computer implemented methods of determining, comparing and analyzing gene expression patterns, in order to assess the likelihood of tumor relapse in a patient. The analytical programs of the invention may be interfaced with, for example, programs that are part of an automated nucleic acid detection system so that data from the automated nucleic acid detection system is fed directly to the analytical programs of the invention. For example,

final identification of the genes that are expressed and measurement of the amount of gene expression products that are present in a sample is usually determined in an automated manner. Those of skill in the art will recognize that automated systems exist that are designed, once supplied with appropriate starting material (e.g. suitable nucleic acid sample to analyze, labeling reagents, etc.). Such programs are typically computer implemented and are able to output, for example, the identity of the genes associated with the nucleic acids in the sample and the quantity of the expressed genes, e.g. the degree of upregulation or downregulation. The interface between the analytical programs of such a system and those of the present invention may be direct or indirect, i.e. the programs of the invention may be merely linked to accept information from such a program, or one program with both capabilities may be developed. The analytical programs of the invention may contain (and in fact may be used to develop or update) a database of gene expression patterns from tumors (either from a specific individual, or from a plurality of individuals) and usually also have the capability to compare the results obtained with an experimental or unknown sample to the results of reference or control values in the database, or to any other value in the database. The computer programs are generally capable of statistically analyzing the data, including determining the significance of deviations from normal or control values, or between samples, or between sets of genes from a sample e.g. to determine differential expression of genes between the genes listed in Table 4 and those listed in Table 5, and/or to determine the expression or upregulation, or lack thereof, of the 5- and/or 8-gene signatures described herein.

The output of the programs of the invention may include, for example, absolute or relative levels of gene expression, e.g. identification of the expression of one or more genes of interest, identification of the absence of expression of one or more genes of interest, the levels of expression of one or more genes of interest (e.g. percentages, fold increases or decreases, etc.), and the like. In addition, the computer implemented analytical methods of the invention may interface directly or indirectly with cancer treatment protocol programs, i.e. two programs may be linked to merely accept data or output from another program, or one program encompassing both capabilities may be developed. In addition, all three programs (analysis of gene patterns, prognosis of patient response to tumor, and suggested treatment protocols) may be linked or integrated into one computer-implemented program. In this case, output from the program may be one or more suggested courses of treatment

(treatment protocols) for the patient from whom the tumor sample was obtained.

The invention also provides a system for characterizing an immune response of a patient with a tumor. A flow chart of the basic steps of one embodiment of the method is provided in Figure 5A and another is illustrated in Figure 5B. The system is illustrated schematically in Figure 6. The system includes measuring means 10 for obtaining measurements of differential expression of immune system genes in tumors. Those of skill in the art will recognize that such a system may include a microchip or "genechip" 11 for hybridizing nucleic acids from the sample of interest (in this case, a tumor sample), and that results from microchip hybridization experiments may be converted into a detectable signal, such as a fluorescent, luminescent, or other type of signal. The means for obtaining measurements may also comprise various detectors or other means of reading or measuring results 12 obtained with the microchip. Generally, the results will be expressed in the form of numeric values corresponding to levels of expression of the genes that were tested, e.g. lists of the amounts or relative amounts of detected transcription products associated with the genes of interest (e.g. immune system genes as described herein). Statistical significance data may also be provided.

The system of the invention also includes means for recognizing 20, using the measurements, patterns of differential gene expression. The means for recognizing 20 will generally be a computer processor or a network of computers comprising a computer implemented program (e.g. software with instructions, enclosed in a computer readable medium such as a diskette, hard disk, CD ROM, DVD, thumb drive, firmware, etc.) capable of receiving (inputting) the measurements from measuring means 10, and capable of statistically analyzing the measurements to identify or recognize one of three prototypic patterns, each of which correlates with one of the following three types or categories of immune responses: i) an immune response that is rejecting said tumor; ii) an immune response that is not rejecting said tumor due to the action of immune suppressing factors; and iii) an immune system response that is not rejecting said tumor due to changes in said tumor. The means for recognizing 20 can also output (i.e. comprises a means to output) the recognized pattern for display, for further processing and analysis, etc.

The system of the invention also includes means for assigning 30, capable of receiving the recognized pattern from recognizing means 20, and, based on the pattern, assigning a characteristic of interest (e.g. proclivity toward tumor recurrence, or lack thereof)

to the particular patient from whom the tumor sample was obtained. Assigning means 30 may also be a computer (the same or different from those described previously) comprising a computer implemented program (e.g. software, etc. as described previously) capable of receiving (inputting, i.e. containing a means to input) the pattern recognized by recognizing means 20. (In fact, recognizing means 20 and assigning means 30 may be integrated into a single computerized system.) This assignation can be outputted via an output means to a user, and used to establish a suitable treatment protocol. In fact, the system may optionally further include a means for developing and outputting a recommended treatment protocol 40 (which may or may not be integrated into a single computerized system with recognizing means 20 and assigning means 30). In all cases, output from each system means may be electronic (e.g. input or downloaded to another instrument, or to a computer screen or display). Alternatively, or in addition, hard copies of the output may be generated, e.g. by a printer that is linked to the system. The output may be in the form of, for example, a list, chart, diagram, photograph or photograph-like digitalized reproduction of results, and the like.

Those of skill in the art will recognize that instructions for causing a computer to carry out the computer-implemented analysis programs for one or more of measuring differential gene expression, recognizing a gene expression pattern as described herein and assigning or associating a pattern to/with a particular type of outcome, (and optionally for developing a treatment protocol) may be integrated into a single computer program or firmware.

The results obtained from the system of the invention are used or interpreted by a health care professional (e.g. a physician, or other skilled professional) to plan, recommend, adjust, modify or otherwise develop a treatment program that is tailored to the needs of the patient with the tumor. The treatment that is recommended is based on or takes into account the results of the analysis. Such a treatment will be more likely to provide benefit to the patient by working with or taking into account the patient's immune response or the status of the patient's immune system, rather than possibly aggravating the patient's immune response to the tumor, or rather than treating the tumor according to a protocol that does not take individual patient differences into account. For example, for patients who are likely to relapse, an aggressive treatment stance may be taken, including aggressive surgery and prolonged chemotherapy, radiation therapy, etc. In contrast, for patients who are unlikely to relapse, may receive more conservative, less aggressive treatments.

Like breast cancer patients, distinct signatures of immune function genes associated with recurrence or recurrence-free survival can be identified in mouse models of breast carcinoma, though with different patterns of gene signatures. In some embodiments, failure to reject tumor cells may be due to changes in the phenotype of the tumor itself, which allow it to evade the patient's immune system response, which is essentially normal and unimpaired. In this case, the patient's immune system initially mounts an appropriate response and, as the tumor changes, continues to attempt to deal with the tumor e.g. by switching from a predominantly Th-1 response (increased ISGs and granzyme) to a Th-2/humoral response. However, Th-2 and humoral responses are less successful in eradicating cancer cells, and rejection of tumor cells that arise after initial treatment does not occur, so a recurrence of another iteration of the tumor is likely. On the other hand, failure to reject may be due to direct suppression of the patient's immune system. In this scenario, immune suppressing factors (usually secreted by tumor cells) act on the immune system cells of the patient, causing it to shut down. These genes/pathways are shown in Figure 11 and include primary immunodeficiency signaling, calcium induced T cell apoptosis, CTLA4 signaling in CTL, production of NO and reactive oxygen species. Without being bound by theory, it is possible that upregulation of these genes even in relapse free patients may explain why these patients failed to reject their tumors in the first place, and that when the tumor is removed by surgery, these immune suppressors that are induced by tumor-derived factors will disappear and other immune effector genes will then be able to destroy residual tumors. In other words, in this scenario, primary rejection does not occur, and technically recurrence also does not occur since, from the beginning, the patient's immune effectors are suppressed, though are still present. At one time, it was thought that such a lack of rejection was due to "tolerance" of the patient's immune system for the tumor. However, it has now been discovered that tolerance to tumors never exists in cancer patients; a patient's immune system does not simply ignore or not respond to the presence of tumor cells. Instead, failure to reject tumors is due to the induction of immune regulatory mechanisms that suppress anti-tumor immune responses. These findings change the existing paradigm for understanding tumor formation in cancer patients.

In contrast, in some embodiments, different patterns of gene expression were observed in both categories of non-rejected tumors. In the case in which tumor evasion has occurred or is occurring, the chemokines such as those listed in Table 3 are differentially

expressed compared to individuals who reject the tumor or individuals whose immune system is suppressed. It is noted that some molecules such as Cxcl9 are increased in the rejection and recurrence models compared with the "tolerogenic" (suppression) model; this indicates that tumor recurrence has occurred under immune pressure such that the anti-tumor immune response rejected HER-2/neu positive tumors and at the same time induced loss of HER-2/neu and resulted in the recurrence of HER-2/neu negative tumor variants. In addition, when tumor evasion occurs, various Interleukins and Signaling genes such as those listed above the top-most solid line in Table 3 are differentially expressed. Further, when tumor evasion occurs, ISGs listed above the second (bottom) heavy line in Table 3 were also differentially expressed. Basically, when tumor evasion occurs, the host immune response contains some hallmarks of a Th-1 response, together will elements of a Th-2 and/or humoral response.

In the case of immune suppression, the pattern of gene expression is generally characterized by reduced Cxcl9 expression. Expression of some interleukins and signaling molecules such as Vpreb3 and Pias2 are also reduced, but those below the first (top-most) heavy line of Table 3 show increased expression compared to immune suppressed individuals, and, with some exceptions, in comparison to individuals who reject the tumor. Lastly, ISGs listed below the lower heavy line in Table 3 are generally upregulated in immune suppression individuals, compared to tumor evasion individuals, and with some exceptions, in comparison to individuals who rejected the tumor as well. Those of skill in the art will recognize that in some cases, the pattern of gene expression will be increased above a reference level whereas in other cases the pattern may show a decrease to below a reference level. Both of these deviations from the reference are valuable and can form a part of the overall genetic signature of the model. Similarly, some increases or decreases may overlap across models, e.g. may be increased in two out of the three and decreased in only one. Nevertheless, such markers are valuable because the genetic signature is based on the assessment of results obtained with many genes, and it is the overall patterns that are characteristic of a condition of interest.

These differentially expressed genes can be organized into three categories: 1) chemokines; 2) interferon-stimulated genes (ISGs); and 3) cytokines and signaling molecules. Table 1 (presented as Figure 4A) lists exemplary chemokines (and receptors) and exemplary ISGs and Table 2 (presented as Figure 4B) lists exemplary cytokines and

signaling molecules which may be differentially expressed in a tumor microenvironment of a patient that is mounting a robust, appropriate immune response to the tumor. Table 3 (presented as Figure 4C) lists other selected genes of interest with immunological function.

I. CHEMOKINES

Exemplary chemokines and chemokine related molecules such as receptors include but are not limited to:

CXC chemokines and receptors such As : chemokine (C-X-C motif) ligand 2 (Cxcl 2), chemokine (C-X-C motif) ligand 1 (Cxcl 1) and chemokine (C-X-C motif) ligand 11 (Cxcl 11).

CC chemokines and receptors such as: chemokine (C-C motif) ligand 1 (Ccl1); chemokine (C-C motif) ligand 4 (Ccl4); chemokine (C-C motif) ligand 5 (Ccl5); chemokine (C-C motif) ligand 6 (Ccl6); chemokine (C-C motif) ligand 8 (Ccl8); chemokine (C-C motif) ligand 9 (Ccl9); chemokine (C-C motif) ligand 11 (Ccl11); chemokine (C-C motif) ligand 22 (Ccl22); chemokine (C-C motif) receptor-like 2 (Ccrl 2); chemokine (C-C motif) receptor 10 (Ccr10); chemokine-like factor (Cklf); Duffy blood group, chemokine receptor (Darc); and chemokine-like factor, transcript variant 1 (Cklf).

Chemokines such as: chemokine(C-C motif) ligand 2 (Ccl2); chemokine (C-C motif) ligand 4 (Ccl4); chemokine (C-C motif) ligand 6 (Ccl6); chemokine (C-C motif) receptor 7 (Ccr7); chemokine (C-X-C motif) ligand 10 (Cxcl0); chemokine (C-X-C motif) ligand 9 (Cxc9); chemokine (Cmotif) ligand 1 (Xcl1); and chemokine (C-X3-C motif) ligand 1 (Cx3cl).

II. INTERFERON-STIMULATED GENES (ISGs)

Exemplary Interferon stimulated genes (ISGs) include but are not limited to: interferon alpha 2 (Ifna2); interferon gamma (Ifng); interferon activated gene 202B (Ifn202b); interferon, alpha-inducible protein 27 (Ifn27); interferon activated gene 204 (Ifn204); interferon induced transmembrane protein 1 (Ifntm1); interferon regulatory factor 6 (Inf6); interferon-induced protein with tetratricopeptide repeats 1 (Ifnt1); interferon regulatory factor 4 (Ifn4); myxovirus (influenza virus) resistance 1 (Mx1); signal transducer and activator of transcription 2 (Stat 2); signal transducer and activator of transcription 6 (Stat 6); and interferon regulatory factor 2 binding protein 1 (Ifn2bp1).

Exemplary ISG genes also include: interferon beta 1, fibroblast (Ifnb1); interferon regulatory factor 7 (Irf7); interferon δ-related developmental regulator 1 (Ifrd1); interferon

(alpha and beta) receptor 1 (Ifnar1); interferon gamma induced GTPase (Igtp); interferon regulatory factor 1 (Irf1); interferon regulatory factor 3 (Irf3); interferon regulatory factor 6 (Irf6); interferon gamma receptor1 (Irfgr1); and interferon alpah responsive gene (Ifrg15).

III. CYTOKINES AND SIGNALING MOLECULES

Exemplary cytokines and signaling molecules include but are not limited to:

Various interleukins and receptors such as: interleukin 1 alpha (Ila); interleukin 1 beta (Il1b); interleukin 1 family, member 9 (Il1f9); interleukin 5 (Il5); interleukin 7 (Il7); interleukin 17F (Il17f); interleukin 31 (Il31); interleukin 1 receptor accessary protien; transcript variant 2 (Il1rap); interleukin 2 receptor, gamma chain (Il2rg); interleukin 7 receptor (Il7r); interleukin 23 receptor (Il23r); and interleukin 17 receptor B (Il17rb).

Various cytotoxic and pro-apoptotic molecules such as: granzyme B (Gzmb); cytotoxic T lymphocyte-associated protein 2 alpha (Ctla2a); killer cell lectin-like receptor subfamily A, member 9 (Klra9); killer cell lectin-like receptor subfamily D, member 1 (Klrd1); Fas ligand (TNF superfamily, member 6) (Fas1); tumor necrosis factor (ligand) superfamily, member 11 (Tnfsf11); tumor necrosis factor receptor superfamily, member 1b (Tnfsf1b); and tumor necrosis factor receptor superfamily, member 4 (Tnfsf4).

Various Toll-like receptors and lymphocyte signaling molecules such as: toll-like receptor 4 (Tlr4); toll-like receptor 6 (Tlr6); interleukin 4 induced 1 (Il4i1); activated leukocyte cell adhesion molecule (Alcam); B-cell leukemia/lymphoma 2 related protein A1c (Bcl2a1c); IL-2-inducible T-cell kinase (Itk); early B-cell factor 4 (Ebf4); lymphocyte antigen 6 complex, locus A (Ly6a); lymphocyte antigen 6 complex, locus C (Ly6c); lymphocyte antigen 6 complex, locus F (Ly6f); lymphocyte protein tyrosine kinase (Lck); T-cell activation Rho GTPase-activating protein (Tagap); T-cell leukemia, homeobox 1 (Tls1); T-cell leukiemia/lymphoma 1B, 1 (Tcl1b1); NF-kappaB repressing factor (Nkrf); NFKB inhibitor interacting Ras-like protein 2 (Nkiras2); Nfkb light polypeptide gene enhancer in B-cells inhibitor, zeta (Nlfkbiz); and nuclear factor of activated T-cells 5, transcript variant b (Nfat5);

Various FC-type receptors such as: Leucocyte immunoglobulin-like receptor, subfamily B, member 4 (Ilrb4); macrophage galactose N-acetyl-galactosamine specific lectin 1 (Mgl1); macrophage galactose N-acetyl-galactosamine specific lectin 2 (Mgl2); and macrophage scavenger receptor 1 (Msr1).

Various immunoglobulin genes such as: immunoglobulin heavy chain 6 (Igh-6);

immunoglobulin heavy chain 6 (heavy chain of IgM) (Igh-6); immunoglobulin joining chain (Igj); immunoglobulin heavy chain 6 (Ign-6); immunoglobulin kappa chain variable 28 (Igk-V28); immunoglobulin lambda chain, variable 1 (Igl-V1); and immunoglobulin light chain variable region (Igkv4-90).

Various interleukin and signaling genes such as: interleukin 12b (Il12b); interleukin 13 (Il13); interleukin 17D (Il17d); interleukin 23 receptor (Il23r); interleukin 2 receptor, gamma chain (Il2rg); interleukin 4 (Il4); interleukin 4 induced 1 (Il4i1); interleukin 6 (Il6); interleukin 7 receptor (Il7r); interleukin 9 (Il9); toll-like recpetor 11 (Tlr11); B-ce.. Linker (Blnk); Bcl-2-related ovarian killer protein (Bok); pre-B lymphocyte gene 3 (Vpreb3); lymphocyte cytosolic protein 2 (Lcp2); lymphocyte antigen 6 complex, locus D (Ly6d); mfkb light chain gene enhancer 1, p105 (Nfkb1); protein inhibitor of activated STAT 2 (Pias2); protein inhibitor of activated STAT 3 (Pias3); signal transducer and activator of transcription 4 (Stat 4); interleukin 10 (Il10); interleukin1 receptor, type II (Il1r2); interleukin10 receptor, beta (Il10b); suppressor of cytokine signaling 1 (Socs1); suppressor of cytokine signaling 3 (Socs3); BCL-2-antagonist/killer 1 (Bak1); lymphocyte specific 1 (Lsp1); TRAF family member-associated Nf-kappa B activator (Tank); and toll-like receptor 6 (Tlr6).

The invention is further illustrated in the ensuing examples, which are provide to illustrate the invention but which should not be interpreted as limiting the invention in any way.

EXAMPLES

EXAMPLE 1. Signatures associated with rejection or recurrence in HER-2/neu positive mammary tumors

Summary: We have previously shown T cell-mediated rejection of the neu-overexpressing mammary carcinoma cells (MMC) in wild-type FVB mice. However, following rejection of primary tumors, a fraction of animals experience a recurrence of a neu antigen-negative variant (ANV) of MMC (tumor evasion model), after a long latency period. In the present study, we determined that T cells derived from wild-type FVB mice can specifically recognize MMC by secreting IFN-γ and can induce apoptosis of MMC *in vitro*. Neu-transgenic (FVBN202) mice develop spontaneous tumors that they cannot reject (tumor tolerance model). To dissect the mechanisms associated with rejection or tolerance of MCC tumors, we compared transcriptional patterns within the tumor microenvironment of MMC

undergoing rejection with those that resisted rejection either because of tumor evasion/antigen-loss recurrence (ANV tumors) or because of intrinsic tolerance mechanisms displayed by the transgenic mice. Gene profiling confirmed that immune rejection is primarily mediated through activation of interferon stimulated genes (ISGs) and T cell effector mechanisms. The tumor evasion model demonstrated combined activation of Th1 and Th2 with a deviation towards Th-2 and humoral immune responses that failed to achieve rejection likely because of lack of target antigen. Interestingly, the tumor tolerance model instead displayed immune suppression pathways through activation of regulatory mechanisms that included in particular the over-expression of IL-10, IL-10 receptor and suppressor of cytokine signaling (SOCS)-1 and SOCS-3. This data provides a road-map for the identification of novel biomarkers of immune responsiveness in clinical trials.

In order to carry out this work, we adopted an experimental model that could address the paradoxical relationship between adaptive immune responses against cancer antigens and rejection or persistence of antigen-bearing cancers with the intent of comparing functional signatures between the experimental model and previous human observation that could shed mechanistic information on this relationship and potentially provide novel predictive or prognostic biomarkers to be tested in the clinical settings. In this study, we compared transcriptional patterns of mammary tumors undergoing rejection to that of related tumors that evaded immune recognition through antigen loss (evasion model) or resided in tolerized transgenic mice (tolerogenic model). For this purpose, we used FVB mice that reject neu-overexpressing mammary carcinomas (MMC) because of the presence of a potent neu-specific T cell response. Although MMC are consistently rejected after a few weeks, occasionally MMC recur and in such instances they resist further immune pressure by invariably loosing HER-2/neu expression (tumor evasion model) (21, 22). Moreover, FVBN202 mice that constitutively express high levels of HER-2/neu fail to reject MMC because they cannot mount effective anti-tumor T cell responses (tolerogenic model). Thus, we compared the tumor microenvironment at salient moments of immune response/evasion/tolerance to gain, in this previously well-characterized model (21, 22), insights about the immune mechanisms leading to tumor rejection and their failure in conditions of tumor evasion or systemic tolerance. Interestingly, the tolerance model, which was expected to show tolerance, displayed immune suppression pathways through activation of regulatory mechanisms that included in particular the over-expression of IL-10, IL-10

receptor and suppressor of cytokine signaling (SOCS)-1 and SOCS-3.

MATERIALS AND METHODS

Mice

Wild-type FVB (Jackson Laboratories) and FVBN202 female mice (Charles River Laboratories) were used throughout these studies. FVBN202 is the rat neu transgenic mouse model in which 100% of females develop spontaneous mammary tumors by 6–10 mo of age, with many features similar to human breast cancer. These mice express an unactivated rat neu transgene under the regulation of the MMTV promoter (23). Because of the overexpression of rat neu protein, FVBN202 mice are expected to tolerate the neu antigen as self protein, and in cases where there might be a weak neu-specific immune response prior to the appearance of spontaneous mammary tumors are still well tolerated (24, 25). On the other hand, rat neu protein is seen as nonself antigen by the immune system of wild-type FVB mice, resulting in aggressive rejection of primary MMC (21, 26). The studies have been reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at Virginia Commonwealth University.

Tumor cell lines.

The MMC cell line was established from a spontaneous tumor harvested from FVBN202 mice as previously described (11, 15). Tumors were sliced into pieces and treated with 0.25% trypsin at 4 °C for 12-16 h. Cells were then incubated at 37 °C for 30 min, washed, and cultured in RPMI1640 supplemented with 10% Fetal Bovine Serum (FBS) (21, 22). The cells were analyzed for the expression of rat neu protein before use. Expression of rat neu protein was also analyzed prior to each experiment and antigen negative variants (ANV) were reported accordingly (see results).

In vivo tumor challenge.

Female FVB or FVBN202 mice were inoculated s.c. with MMC ($4\text{-}5\times10^6$ cells/mouse). Animals were inspected twice every week for the development of tumors. Masses were measured with calipers along the two perpendicular diameters. Tumor volume was calculated by: V(volume) = L(length) x W(width)$^2 \div 2$. Mice were sacrificed before a tumor mass exceeded 2000 mm^3.

IFN-γ ELISA

Secretion of MMC-specific IFN-γ by lymphocytes was detected by co-culture of lymphocytes (4×10^6 cells) with irradiated MMC or ANV (15,000 rads) at 10:1 E:T ratios in

complete medium (RPMI1640 supplemented with 10% FBS, 100 U/ml penicillin, 100 μg/ml streptomycin) for 24 hrs. Supernatants were then collected and subjected to IFN-γ ELISA assay using a Mouse IFN-γ ELISA Set (BD Pharmingen, San Diego, CA) according to the manufacturer's protocol. Results were reported as the mean values of duplicate ELISA wells.

Flow cytometry.

A three color staining flow cytometry analysis of the mammary tumor cells (10^6 cells/tube) was carried out using mouse anti-neu (Ab-4) Ab (Calbiochem, San Diego, CA), control Ig, FITC-conjugated anti-mouse Ig (Biolegend, San Diego, CA), PE-conjugated annexin V and propidium iodide (PI) (BD Pharmingen, San Diego, CA) at the concentrations recommended by the manufacturer. Cells were finally added with annexin V buffer and analyzed at 50,000 counts with the Beckman Coulter EPICS XL within 30 min.

Microarray performance and statistical analysis:

Total RNA from tumors was extracted after homogenization using Trizol reagent according to the manufacturer's instructions. The quality of secondarily amplified RNA was tested with the Agilent Bioanalyzer 2000 (Agilent Technologies, Palo Alto, CA) and amplified into anti-sense RNA (aRNA) as previously described (27, 28). Confidence about array quality was determined as previously described (29). Mouse reference RNA was prepared by homogenization of the following mouse tissues (lung, heart, muscle, kidneys and spleen) and RNA was pooled from 4 mice. Pooled reference and test aRNA were isolated and amplified in identical conditions during the same amplification/hybridization procedure to avoid possible inter-experimental biases. Both reference and test aRNA were directly labeled using ULS aRNA Fluorescent labeling Kit (Kreatech, Netherlands) with Cy3 for reference and Cy5 for test samples.

Whole genome mouse 36 k oligo arrays were printed in the Infectious Disease and Immunogenetics Section of Transfusion Medicine (IDIS), Clinical Center, National Institute of Health, Bethesda using oligos purchased from Operon (Huntsville, AL). The Operon Array-Ready Oligo Set (AROS™) V 4.0 contains 35,852 longmer probes representing 25,000 genes and about 38,000 gene transcripts and also includes 380 controls. The design is based on the Ensembl Mouse Database release 26.33b.1, Mouse Genome Sequencing Project, NCBI RefSeq, Riken full-length cDNA clone sequence, and other GenBank sequence. The microarray is composed of 48 blocks and one spot is printed per probe per slide. Hybridization was carried out in a water bath at 42 °C for 18-24 hours and the arrays were

then washed and scanned on a Gene Pix 4000 scanner at variable PMT to obtain optimized signal intensities with minimum (< 1% spots) intensity saturation.

Resulting data files were uploaded to the mAdb databank (http://nciarray.nci.nih.gov) and further analyzed using BRBArrayTools developed by the Biometric Research Branch, National Cancer Institute (30) (web site located at linus.nci.nih.gov/BRB-ArrayTools.html) and Cluster and Treeview software (31). The global gene-expression profiling consisted of 18 experimental samples. Subsequent filtering (80% gene presence across all experiments and at least 3-fold ratio change) selected 11,256 genes for further analysis. Gene ratios were average-corrected across experimental samples and displayed according to uncentered correlation algorithm (32).

Statistical analysis

Rate of tumor growth was compared statistically by un-paired Student's t test. Unsupervised analysis was performed for class confirmation using the BRBArrayTools and Stanford Cluster program (32). Class comparison was performed using parametric unpaired Student's t test or three-way ANOVA to identify differentially-expressed genes among tumor-bearing, tumor-rejection and relapse groups using different significance cut-off levels as demanded by the statistical power of each comparison. Statistical significance and adjustments for multiple test comparisons were based on univariate and multivariate permutation test as previously described (33, 34).

RESULTS

T cell-mediated rejection of MMC and relapse of its neu antigen negative variant, ANV, in wild-type FVB mouse

Wild-type FVB mice are capable of rejecting MMC within 3 weeks because of specific recognition of rat neu protein by their T cells as opposed to their transgenic counterparts, FVBN202, that tolerate rat neu protein and fail to reject MMC (21, 26). In order to determine whether aggressive rejection of primary MMC by T cells may lead to relapse-free survival in wild-type FVB mice, we performed follow-up studies. Animals (n=15) were challenged with MMC by subcutaneous (s.c.) inoculation at the right groin. Animals were then monitored for tumor growth twice weekly. All mice rejected MMC within 3 weeks after the challenge. However, a fraction of these animals (8 out of 15 mice) developed recurrent tumors at the site of inoculation. These relapsed tumors had lost neu expression under immune pressure (21, 26). Relapsed-free groups were maintained as

breeding colonies and did not show any relapse during their life span. Splenocytes of FVB mice secreted IFN-γ in the presence of MMC only (2200 pg/ml) while no appreciable IFN-γ was detected when lymphocytes were stimulated with ANV (110 pg/ml). No IFN-γ was secreted by splenocytes or tumor cells alone (data not shown).

T cells derived from wild-type FVB mice will induce apoptosis in MMC

In order to determine whether neu-specific recognition of MMC by T cells may induce apoptosis in MMC, *in vitro* studies were performed. Splenocytes of naïve FVB mice were stimulated with irradiated MMC for 24 hrs followed by 3 day expansion in the presence of IL-2 (20 U/ml). Lymphocytes were then co-cultured with MMC (E:T ratio of 2.5:1 and 10:1) for 48 hrs in the presence of IL-2 (20 U/ml). Control wells were seeded with MMC or splenocytes alone in the presence of IL-2. Cells (floaters and adherents) were collected and subjected to a three color flow cytometry analysis using mouse anti-rat neu Ab (Ab-4), PE-conjugated anti mouse Ig, control Ig, annexin V, and PI. Gated neu positive cells were analyzed for the detection of annexin V+ and PI+ apoptotic cells. As shown in Figure 1A, 80% of MMC were annexin V- and PI- in the absence of lymphocytes while only 49% of MMC were annexin V- and PI- in the presence of lymphocytes at 10:1 E:T ratio. At a lower E:T ratio (2.5:1) there was a slight dropping in the number of viable MMC (from 80% to 74%), but marked increase in the number of early apoptotic cells (annexin V+/PI-) from 1% to 10%. At a higher E:T ratio (10:1) early (Annexin V+/PI-) or late (annexin V+/PI+) apoptotic cells and necrotic cells (annexin V-/PI+) were markedly increased.

Adoptive immunotherapy (AIT) of FVBN202 mice using T cells derived from wild-type FVB donors failed to reject MMC

In order to determine whether T cells of FVB mice with neu-specific and anti-tumor activity may protect FVBN202 mice against MMC challenge, AIT was performed. Using nylon wool column, T cells were enriched from the spleen of FVB donor mice following the rejection of MMC. FVBN202 recipient mice were injected i.p. with cyclophosphamide (CYP; 100 μg/g) in order to deplete endogenous T cells. After 24 hrs animals were challenged with MMC tumors (4 x 10^6 cells/mouse). Four-five hrs after tumor challenge, donor T cells were transferred into FVBN202 mice (6 x 10^7 cells/mouse) by tail vein injections. Control FVBN202 mice were challenged with MMC in the presence or absence of CYP treatment. Animals were then monitored for tumor growth. As shown in Figure 1B, CYP treatment of animals resulted in retardation of tumor growth in FVBN202 mice as

expected. Student's t test analysis on days 14, 21, and 28 post-challenge showed significant differences between these two groups (P= 0.005, 0.007, 0.01, respectively). Adoptive transfer of neu-specific effector T cells from MMC-sensitized FVB mice into CYP-treated FVBN202 groups did not significantly inhibit tumor growth compared to CYP-treated control groups ($P > 0.05$). Adoptive transfer of neu-specific effector T cells from untreated FVB mice into CYP-treated FVBN202 groups showed similar trend of tumor growth (data not shown). These experiments suggest that T cell responses associated with MMC rejection in wild-type FVB mice (21) may represent an epiphenomenon with no true cause-effect relationship or that FVBN202 mice retain tolerogenic properties in spite of CYP treatment that can hamper the function of potentially effective anti-cancer T cell responses. We favor the second hypothesis based on our previous depletion experiments that demonstrated the requirement of endogenous effector T cells of FVB mice for rejection of MMC tumors (21).

Genetic signatures defining rejection or tolerance of MMC tumors

To ascertain whether the presence of neu-specific effector T cells may trigger a cascade of events which may determine success or failure in tumor rejection, wild-type FVB and FVBN202 mice were inoculated with MMC. Historically, all FVB mice reject MMC, however a fraction develop a latent tumor relapse. In contrast, FVBN202 mice fail to reject transplanted MMC. Ten days after the tumor challenge, transplanted MMC tumors were excised and RNAs were extracted from both FVB and FVBN202 carrier mice based on the presumption that the biology of the former would be representative of active tumor rejection and that of the latter representative of tumor tolerance. Thus, the timing of tumor harvest was chosen to capture transcriptional signatures associated with the active phase of the tumor rejection process in wild-type FVB mice in comparison with the corresponding tolerance of spontaneous mammary tumors in the FVBN202 mice. We speculated that this comparison would allow distinguishing whether tolerance was due to inhibition of T cell function within the tumor microenvironment of spontaneous mammary tumors or to a complete absence of such responses. To enhance the robustness of the comparison, a similar analysis was performed extracting total RNA from spontaneous tumor in FVBN202 mice. In addition, RNA was extracted from MMC tumors in wild-type FVB mice that experienced tumor recurrence following the initial rejection of MMC. This second analysis allowed the comparison of mechanisms of tumor evasion in the absence of known tolerogenic effects. Micro array analyses were then performed on the amplified RNA (aRNA) extracted from

these tumors using 36k oligo mouse arrays. Hence, genes considered as differentially expressed in the study groups could either represent MMC tumor cells or host cells infiltrating the tumor site. Probes with missing values greater than 80% or a change less than 3 fold were excluded from further analysis. Unsupervised clustering demonstrated outstanding differences among the three experimental groups (Figure 2A). Genes of spontaneous mammary tumors (samples 13, 15, 16, 17) clustered closely to those of transplanted MMC (14 and 19) excised from FVBN202 mice suggesting that the biology of MMC tumors remains comparable between these two experimental models of tolerance. Global transcriptional patterns associated with tumor relapse (samples 1–8) were instead clearly different from those of spontaneous mammary tumors or MMC transplanted in tolerant FVBN202 mice suggesting that a completely different biological process was at the basis of tumor evasion through loss of target antigen expression. Finally, MMC tumors undergoing rejection (samples 9–12) were clearly separated from the either kind of non-regressing tumors.

Biomarkers of rejection

Our first class comparison searched for differences between the four tumor samples undergoing rejection and the rest of the MMC tumors whether belonging to the tolerogenic or the evasion process. This approach followed the exclusion principle whereby factors determining the occurrence of a phenomenon should be discernible from unrelated ones independent of the causes preventing its occurrence. An unpaired Student t test with a cut-off set at $p < 0.001$ identified 2,449 genes differentially expressed between regressing and non regressing tumors (permutation p value = 0) of which 1,003 genes were upregulated in regressing tumors clearly distinguishing the two categories (Figure 2B). Of those, a large number were associated with immune regulatory functions. Gene Ontology databank was queried to assign genes to functional categories and upregulated pathways were ranked according to the number of genes identified by the study belonging to each category (Figure 2C). The top categories of genes that were upregulated in primary rejected MMC tumors were Cytokine-Cytokine interaction, MAPK signaling, Cell Adhesion related transcripts and axon guidance, T cell receptor, JAK-STAT and Toll-like receptor signaling pathways. NK cell mediated cytotoxicity and calcium signaling pathways were also enriched in upregulated genes. In contrast, very little evidence of immune activation could be observed in either category of non-regressing tumors suggesting that lack of immune rejection is due to absent

or severely hampered immune responses in the tumor microenvironment independent of the mechanisms leading to this resistance.

To better describe the immununological pathways associated with tumor regression we organized genes with immune function into three categories including chemokines, IFN-α2, IFN-γ and interferon-stimulated genes (ISGs) (Table 1, presented in Figure 4A), and cytokines and signaling molecules (Table 2, presented in Figure 4B). From this analysis, it became clear that T cell infiltration into tumors was associated with activation of various pathways leading to the expression of IFN-α, IFN-γ and several ISGs including interferon regulatory factor (IRF)-4, IRF-6 and STAT-2. In addition, several cytotoxic molecules were overexpressed including calgranulin-a, calgranulin-b and granzyme-B; all of them representing classical markers of effector T cell activation in humans (10) and in mice (35). Thus, tumor rejection in this model clearly recapitulates patterns observed in various human studies in which expression of ISGs is associated with the activation of cytotoxic mechanisms among which granzyme-B appears to play a central role.

Is there a difference between signatures of immune evasion and immune tolerance?

As shown in Table 3 (presented in Figure 4C), the high expression of IL-10 and the IL-10 receptor-β chain concordant with IRF-1 in the tolerogenic model strongly suggests the presence of regulatory mechanism within the microenvironment of MCC-bearing FVBN202 mice. Preferential expression of SOCS-1 and SOCS-3 in the microenvironment of MMC tumors of FVBN202 mice also strongly suggest a marked activation of regulatory functions present in the tolerized host (Table 3, presented in Figure 4C).

To further investigate whether similar mechanisms were involved in failure of tumor rejection in the tolerance model and the evasion model, we characterized potential differences between the two models of immune resistance; we compared statistical differences between the tolerogenic and the evasion model comparing the two non-regressing groups by unpaired Student t test using as a significance threshold a p-value < 0.001. This analysis was performed on pre-selected genes that had been filtered for an at least 80% presence of data in the whole data set and a minimal fold increase of 3 in at least one experiment (Figure 3). This analysis identified 1,369 genes differentially expressed by the two groups (multivariate permutation test p-value = 0) of which 462 were upregulated in the tolerogenic model and 907 were upregulated in the tumor evasion model (Figure 3A). Several of these genes where specifically expressed by either group although the expression

of a few of them was shared by the regressing MCC tumors. Annotations and functional analysis based on Genontology data base (GEO) demonstrated that the predominant functional classes of genes transcriptionally active in one of the other type of non-responding MMC tumors were not associated with classical activation of T cell effector functions but rather were associated with more general metabolic processes (Figures 3B and 3C). However, detailed analysis of transcripts associated with immunological function (Table 3, presented in Figure 4C) defined dramatic differences between the two mechanisms of immune resistance.

DISCUSSION

FVB mice reject primary MMC by T cell-mediated neu-specific immune responses. However, a fraction of animals develop tumor relapse after a long latency. On the other hand their transgenic counterparts, FVBN202, fail to mount effective neu-specific immune responses and develop tumors (21). Although FVBN202 mice appear to elicit weak immune responses against the neu protein within a certain window of time (24), the neu expressing MMC tumors are still well tolerated and animals develop spontaneous mammary tumors. Despite the observation that T cells derived from FVB mice were capable of recognizing MMC and inducing apoptosis in these tumors in vitro, adoptive transfer of such effector T cells into FVBN202 mice failed to protect these animals against challenge with MMC.

It has been suggested that T cells play a significant role in determining the natural history of colon (14-16) and ovarian (17) cancer in humans. Transcriptional signatures have been identified that suggest not only T cell localization but also activation through the expression of IFN-γ, ISGs and cytotoxic effector molecules such as granzyme-B (10). We have recently shown that rejection of basal cell cancer induced by the activation of Toll-like receptor agonists also is mediated, at least in part, by localization and activation of CD8 expressing T cells with increased expression of cytotoxic molecules (36). Yet, a comprehensive experimental overview of the biological process associated with tumor rejection in its active phase has not been reported. Thus, our first class comparison searched for differences between the four tumor samples undergoing rejection and the rest of the MMC tumors whether belonging to the tolerogenic or the evasion process. Unlike non-regressing tumors (tolerance and evasion models), regressing tumors (rejection model) showed upregulation of immune activation genes, suggesting that failure in tumor rejection is due to immune evasion or severely hampered immune responses in the tumor

microenvironment. A particularly interesting observation was the relative low expression of ISGs, with the exception of IRF-2bp1. While the transcriptional patterns differentiating regressing from non-regressing tumors were striking and in many ways representative of previous observations in humans by our and others groups (37), differences among MMC tumors non-regressing in FVBN202 mice and those relapsing after regression in FVB mice were subtle. We have previously proposed that lack of regression of human tumors is primarily associated with indolent immune responses rather than dramatic changes in the tumor microenvironment enacted to counterbalance a powerful effector immune response (13, 33, 37). The MMC tolerance model allowed investigating this hypothesis at least in this restricted case. Spontaneous mammary tumors or transplanted MMC tumors in FVBN202 mice displayed immune suppressive properties that were identified by transcriptional profiling through the activation of genes associated with regulatory function. This would occur only in case an indolent adaptive immune response occurred in these transgenic mice and was hampered at the tumor site by a mechanism of peripheral suppression. If however, central tolerance was the reason for the lack of rejection, minimal changes should be observed in tolerogenic model similar to those detectable in the tumor evasion model where MMC tumors lost expression of HER-2/neu and become irrelevant targets for HER-2/neu-specific T cell responses. The presence of regulatory mechanism within the microenvironment of MCC-bearing FVBN202 mice was associated with increased IL-10 as well as increased expression of SOCS-1 and SOCS-3. It has recently been shown that myeloid-derived suppressor cells (MDCS) induce macrophages to secret IL-10 and suppress anti-tumor immune responses (38). Importantly, it was shown that high levels of MDCS in neu transgenic mice would suppress anti-tumor immune responses against tumors (39). Interleukin-10 is increasingly recognized to be strongly associated with regulatory T cell (40) and M2 type tumor-associated macrophage function (41) and its expression is mediated in the context of chronic inflammatory stimuli by the over-expression of IRF-1. SOCS-1 inhibits type I IFN response, CD40 expression in macrophages, and TLR signaling (42-44). Expression of SOCS-3 in DCs converts them into tolerogenic DCs and support Th-2 differentiation (45). Importantly, tumors that express SOCS-3 show IFN-γ resistance (46). Unlike tolerance model, recurrence model revealed expression of Igtp, suggesting the involvement of IFN-γ in this model (Table 3). This observation is consistent with our previous findings on the role of IFN-γ in neu loss and tumor recurrence (21).

MMC tumors evading immune recognition had undergone a process of complex immune editing that resulted not only in the loss of the HER-2/neu target antigen but also in the upregulation of various Th2 type cytokines such as IL-4 and, IL-13 (47) and the corresponding transcription factor IRF-7 over-expression predominantly associated to a deviation from cellular Th-1 to Th-2 and humoral type immune responses (48). In addition, the microenvironment of recurrent tumors was characterized by the coordinate expression of STAT-4, IL12b, IL-23r and IL-17; this cascade has been associated with the development of Th17 type immune responses that play a dominant role in autoimmune inflammation (49, 50) and T-cell dependent cancer rejection (51, 52). Since both humoral and cellular immune responses are potentially involved in the rejection of HER-2/neu expressing tumors (53), this data suggests that a cognitive and active immune response is still attempting to eradicate MMC tumors which may still express subliminal levels of the target antigen. However, the overall balance between host and cancer cells favors, in the end, tumor cell growth because the expression of HER-2/neu, the primary target of both cellular and humoral responses, is critically reduced.

Altogether, these observations suggest that neu antigen loss and subsequent immunological evasion from cellular Th-1 to Th-2 and humoral type immune response is a major mechanism in evasion model while peripheral suppression such as sustained IL-10, SOCS1 and SOCS3 expression is a major player in tolerance model. This conclusion provides a satisfactory explanation for the lack of rejection of MMC tumors in FVBN202 mice receiving adoptively transferred HER-2/neu-specific T cells. In this case, effective T cell responses exclude central tolerance or peripheral ignorance as the only mechanism potentially hampering their effector function at the tumor site suggesting that other regulatory mechanisms such as peripheral suppression could be responsible for inactivation of donor effector T cells. High levels of MDSC in neu transgenic mice support this possibility, and the role and mechanisms of MDSC in suppression of adoptively transferred neu-specific T cells remain to be determined in FVBN202 mice.

EXAMPLE 2. Immune function gene signatures in human breast cancer

In order to validate signatures of immune function genes that had been detected in the animal (mouse) model of breast carcinoma, gene expression (RNA) in tumor lesions of human breast cancer patients who either remained relapse-free or developed relapse within 1-5 years after the initial treatment was analyzed. As shown in Figure 7, unsupervised

analysis of gene arrays from breast cancer patients compared to reference (control) human peripheral blood mononuclear cells (PBMCs) showed that relapse-free and relapse patients exhibit different, distinct clusterings of expressed genes. 9,797 genes were identified which exhibited at least a 3-fold change in expression between relapsed and relapse-free individuals.

Supervised analysis at p value < 0.001 revealed upregulation of 50 genes in patients with relapse as well as upregulation of 299 genes in patients with relapse-free survival (Figure 8). These genes are listed in tabular form in Figures Tables 4 and 5, respectively. Table 4 is presented in Figures 9A-L and Table 5 is presented in Figures 10A and B. Further canonical pathway analysis of immune function genes revealed that relapse-free survival is associated with upregulation of a network of genes related to B cell development, antigen presentation, graft vs host disease signaling, interferon signaling and primary immunodeficiency signaling (Figures 12A-E). In addition, representative genes involved in these pathways were further validated by detecting protein products using IHC where commercial antibodies were available (data not shown).

Novel findings resulting from this work include the following: 1) one single gene/cell of the immune response cannot predict the outcome with respect to relapse; rather, a network (pattern, signature, etc.) of immune cell activation is required for prognosis; and 2) upregulation of genes involved in immune suppression along with upregulation of genes that are involved in host protection can explain why patients with an intact immune system can still develop cancer. Conventional cancer therapy would change this balance in favor of immune effector signature because genes involved in immune suppression (such as NO) are upregulated along with immune effector genes (listed in Figure 11) and are usually induced because of the presence of the tumors.

In summary, 50 genes associated with tumor relapse and 299 genes associated with relapse-free survival have been identified. Expression patterns of these disease-outcome standard genes can be used as references to analyze gene expression in tumor samples obtained from cancer patients, and to determine or predict the likely prognosis for the patient, thereby influencing the choice of treatments for the patient.

EXAMPLE 3. A signature of immune function genes associated with recurrence-free survival in breast cancer patients

ABSTRACT

The clinical significance of tumor-infiltrating immune cells has been reported in a variety of human carcinomas including breast cancer. However, molecular signature of tumor-infiltrating immune cells and their prognostic value in breast cancer patients remain elusive. We hypothesize that a distinct network of immune function genes at the tumor site can predict a low-risk vs. high-risk of distant relapse in breast cancer patients regardless of the status of ER, PR or HER-2/neu in their tumors. We conducted retrospective studies in a diverse cohort of breast cancer patients with a 1-5 year tumor relapse vs. those with up to 7 years relapse-free survival. The RNA were extracted from the frozen tumor specimens at the time of diagnosis and subjected to microarray analysis and real-time RT-PCR. Paraffin-embedded tissues were also subjected to immunohistochemistry staining. We determined that a network of immune function genes involved in B cell development, interferon signaling associated with allograft rejection and autoimmune reaction, antigen presentation pathway, and cross talk between adaptive and innate immune responses were exclusively upregulated in patients with relapse-free survival. Among the 299 genes, five genes which included B cell response genes were found to predict with >85% accuracy relapse-free survival. Real-time RT-PCR confirmed the 5-gene prognostic signature that was distinct from an FDA-cleared 70-gene signature of MammaPrint panel and from the Oncotype DX recurrence score assay panel. These data suggest that neoadjuvant immunotherapy in patients with high risk of relapse may reduce tumor recurrence by inducing the immune function genes.

INTRODUCTION

The findings descried in the previous Examples led to a retrospective study in breast cancer patients for whom we had outcome data available. Based on preclinical findings we hypothesized that distinct networks (signatures) of immune function genes expressed by tumor-infiltrating and/or tumor-associated cells at the time of diagnosis could predict breast cancer recurrence or relapse-free survival following conventional therapies, and could offer immunotherapeutic strategies to overcome tumor relapse.

MATERIALS AND METHODS

Clinical specimens Tissue specimens were collected from female breast cancer patients prior

to any treatment and maintained in the VCU Massey Cancer Center Tissue & Data Acquisition and Analysis Core (TDAAC) over the past 7 years. According to the follow-up history thus far, we have corresponding annotated patient outcome data available which can stratify patients into 1-5 years relapse (n= 8) and up to 7 years relapse-free survival (n=9). Of these, frozen tissues were used for RNA extraction and paraffin-embedded tissues were used for immunohistochemistry (IHC) staining. Two pathologists identified specimens that contained 10-70% tumor-infiltrating cells for analysis of immune-function genes. Variation in the percent infiltrating cells among specimens was due to different methods of tumor collection used. These studies have been reviewed and approved by the Institutional Review Board (HM10920) at Virginia Commonwealth University.

RNA amplification, probe preparation and microarray hybridization For expression studies based on oligo array techniques, total RNA from tumors was amplified into antisense RNA (aRNA) as previously described [1, 2]. Reference control in human arrays was obtained by pooling peripheral blood mononuclear cells (PBMC) from 4 normal donors. Both, human reference and test total RNA were amplified into antisense RNA in large amounts using identical conditions [1, 2]. Confidence about array quality was confirmed as previously described [3]. For 36k human array performances, both reference and test aRNA were directly labelled using ULS aRNA Fluorescent Labeling kit (Kreatech) with Cy3 for reference and Cy5 for test samples. Whole-genome human 36K oligo arrays, representing 25,100 unique genes of the Operon Human Genome Array–Ready OligoSet version 4.0, were printed in house, using oligos purchased from Operon. The design is based on the Ensembl Human Databasebuild NCBI-35c, with a full coverage on the NCBI human Refseq dataset (04/04/2005). Hybridization was carried out in a water bath at +42°C for 20 hs and the arrays were then washed and scanned on an Agilent Microarray Scanner. Resulting data files were analyzed using BRB-Array-Tools developed by the Biometric Research Branch, National Cancer Institute, National Institutes of Health and visualized using Cluster and TreeView software. The global gene expression profiling consisted of 17 experimental samples. Subsequent filtering (80% gene presence across all experiments) selected 9797 genes for further analysis. Gene ratios were average corrected across experimental samples and displayed according to uncentered algorithm.

Statistical analysis Unsupervised analysis was performed for class confirmation using the

BRB-Array-Tools and Stanford Cluster Program. Class comparison was performed using parametric unpaired Student's t test or three-way ANOVA to identify genes differentially expressed among relapse and relapse –free groups using different significance cutoff levels as demanded by statistical power of each comparison. Statistical significance and univariate and multivariate permutation test as previously described [4]. Functional gene network analysis was performed using the Ingenuity Pathway Analysis system (IPA) which transforms large data sets into a group of relevant networks containing direct and indirect relationships between genes based on known interactions in the literature.

Complete leave-one-out cross validation (LOOCV) model Models to predict which breast cancer patients are likely to relapse were developed using BRB-Array-Tools [5]. Complete leave-one-out cross validation (LOOCV) based prediction accuracy estimates were 100% for the compound covariate predictor (CCP) and the diagonal linear discriminant analysis (DLDA) classifier. CCP is a weighted linear combination of log-ratios for genes that are univariately significant at the specified level. The univariate t-statistics for comparing the classes are used as the weights. DLDA is a version of linear discriminant analysis that ignores correlations among the genes in order to avoid over-fitting the data. Based on 1000 random permutations, the compound covariate predictor and the diagonal linear discriminant analysis classifier both had p-value of 0.001.

Immunohistochemistry (IHC) Immunohistochemistry of paraffin-embedded tumor specimens was performed using Dako automated immunostainer (Dako, Carpinteria, CA). We used anti-human antibodies towards CXCL10 (Santa Cruz Biotechnology, 1:300), signal transducer and activator of transcription 1 (STAT1) (BD Biosciences; 1:100), guanylate binding protein 1 (GBP1) (Abnova, 1:75), granzyme A (GZMA) (SeroTec, 1:50), and CD19 (Abcam, 1:1000) which represent T and B cell responses as well as antigen presentation pathways. The antigen retrieval was achieved using a rice steamer. In order to circumvent the endogenous biotin activity, we used Dako Envision Dual Link System-HRP (Dako, Capinteria CA) in a two-step IHC technique, based on HRP labeled polymer which is conjugated with secondary antibodies. The labeled polymer does not contain avidin or biotin, thereby avoiding the non specific endogenous avidin-biotin activity in the sections.

*Real-time PCR*The RNAs were extracted using Trizol asusing known methods. The cDNA was prepared from 1 μg of total RNA using the Superscript II Kit (Invitrogen) with a dT18

oligonucleotide primer at 42°C for 2 hs. The SensiMix SYBR & Fluorescein Kit (BIOLINE, Taunton, MA) was used according to manufacturer's instructions and real-time PCR was performed using the Bio-Rad's real-time PCR detection system. Suitable primers are known in the art or readily ascertainable by one of skill in the art. Data were normalized to GAPDH housekeeping gene.

This study was conducted under Institutional Review Board (IRB) protocol# HM10920 at Virginia Commonwealth University. All patients gave informed consent under the IRB protocol #247: Tissue Acquisition System for Cancer Research at Virginia Commonwealth University.

RESULTS

Patients at high risk or low risk of tumor recurrence show distinct clustering of genes in the tumor microenvironment at the time of diagnosis We have previously shown that microarray analysis of immune function genes in the tumor microenvironment of a mouse model of breast carcinoma had prognostic value for predicting tumor rejection, tumor progression and recurrence [6]. In the present study we sought to determine whether taking a similar approach with a focus of the specimens with tumor infiltrating cells may predict disease outcome in breast cancer patients. We performed total RNA extraction from frozen tumor specimens with at least 10% infiltrating cells derived from patients with no evidence of recurrence up to 7 years follow-up (n=9) and those with recurrence in the first 5 years of follow-up (n=8). All patients included in this discovery group had differences in age, ethnicity, tumor stage, and status of ER, PR or HER-2/neu in their tumors (Table 6). Microarray analyses were performed on the amplified RNA using 36K oligo human arrays. Genes with missing values >80% were excluded from further analysis trimming the final working set to 9797 genes. Unsupervised clustering showed strong differences between the two groups of patients (Figure 13A). Multiple dimensional scaling based on the complete data set demonstrated that the relapse-free group (black circles) segregated completely in Euclidian space from those who had suffered a relapse (gray circles) (Figure 13B).

Table 6. Patient characteristics and their tumor specimens with > 10% infiltrating cells

Patients	Infiltrated cells (%)	Recurrence	Follow up (years)	Stage	ER	PR	HER-2
VBR-V12c	20	No	4	NA	-	-	-
VBR-V10	20	No	3	IIA	-	-	-
VBR-13	20	No	7	IA	-	-	+
VBR-52	20	No	4	IIIA	-	-	borderline
VBR-04b	25	No	4	IIIA	-	-	-
VBR-70	25	No	6	IIB	-	-	+
VBR-12	30	No	7	IA	+	+	+
VBR-51	20	No	4	IIIC	+	+	-
VBR-23	25	No	6	IIB	-	-	-
VBR-73	50	Yes	3	IA	+	+	+
VBR-31	15	Yes	3	IIIA	NA	NA	NA
VBR-80	10	Yes	1	IIA	-	-	-
VBR-8	20	Yes	5	IIB	-	-	-
VBR-77	10	Yes	3	IIA	+	-	+
VBR-28	50	Yes	3	IIA	-	-	-
VBR-76	10	Yes	3	IA	-	-	-
VBR-7	20	Yes	1	IIA	+	+	+

Differential expression of immune function genes at the tumor microenvironment is

associated with breast cancer outcome In order to determine overall differences between relapse-free and relapsed patients, direct comparison between the two clinical outcomes was performed using Student's t test with 10000 random permutations test. The differentially expressed genes were selected based on permutation p value <0.005 and parametric p value <0.001. The comparison identified 349 genes differentially expressed between the two groups with the zero probabilities of getting at least 349 genes significant by chance (at the 0.001 level) if there are no real differences between the relapse and relapse-free group (Global test p value=0). Among the 349 genes, 299 were up-regulated in relapse-free patients compared to relapsed patients (Table 7, presented in Figure 14). These genes included a dominant cluster of co-modulated genes involved in T cell response (*CXCL10, CXCL9, GZMA, GZMB, HLA-B, HLA-C, CCR7, AIM2, APOE, GBP1, IL7R*), B cell activation (*CD19, C1QA, CD8A*) and antigen presentation (*APOE, CD74, CR1, IL23A, STAT1*). In addition, genes such as cytotoxic T lymphocyte antigen 4 (CTLA-4) and IL-23 Rα that are involved in negative regulation of effector immune responses were also up-regulated in relapse-free patients (Table 7). Conversely, 50 genes that were down-regulated in relapse-free patients were not associated with immune function except for a few genes involved in viral defense mechanisms (integrin B5 --ITGB5) (Table 8).

Table 7. List of up-regulated genes in tumor lesions of patients with relapse-free survival

Gene Name	Fold-change	Gene Name	Fold-change	Gene Name	Fold-change
IGKC	15.50	PVRIG	2.83	SMARCC2	2.15
IGL @	12.31	IFIT3	2.82	IGKV1D-8	2.14
IGKC	11.21	DCC	2.81	Unknown	2.13
LOC440871	9.92	HLA-DRA	2.81	PPP1R9B	2.13
Unknown	9.21	HLA-G	2.81	APOE	2.13
IGKV3-20	8.92	Unknown	2.80	NR1H3	2.12
LOC440871	8.81	CD19	2.79	LILRB2	2.12
Unknown	8.70	KDM2A	2.77	HLA-G	2.11
LOC652493	8.67	TBC1D10C	2.76	DOK2	2.10
CXCL9	8.52	HLA-DRB5	2.76	LILRB4	2.10
IGL	8.43	IRF8	2.75	TNFRSF9	2.09
LIPT1	8.10	Unknown	2.73	SLAMF8	2.09
MS4A1	7.87	GMFG	2.71	AOAH	2.09
LOC96610	7.78	HLA-DQA2	2.71	PAG1	2.08
LOC100290481	7.49	GZMB	2.71	LAIR1	2.07
Unknown	7.48	HLA-DRB1	2.70	Unknown	2.07
LOC100132941	7.32	PLEK	2.68	SAMD3.	2.06
IGKC	7.06	IFI30	2.67	FLJ31306	2.05
IGLC1	7.01	IL4I1	2.67	C1QB	2.05
Unknown	6.86	PARP14	2.67	Unknown	2.05

KIAA1632	6.56	OASL	2.67	MCL1	2.03
CCL23	6.56	IL2RB	2.66	CD53	2.03
Unknown	6.04	IGKC	2.64	CCR7	2.02
IGLC1	5.98	PLAC8	2.62	CD8A	2.01
POU2AF1	5.95	MGC29506	2.62	GPR89B	2.00
STAT1	5.48	HLA-E	2.60	Unknown	2.00
Unknown	5.23	CCL21	2.60	APOL3	2.00
GBP4	5.19	SLC25A11	2.59	PPA1	2.00
GBP1	5.10	CYFIP2	2.58	GIMAP2	2.00
IL7R	5.09	IFI16	2.57	RPS27	1.99
Unknown	5.08	Unknown	2.57	CCDC108	1.99
LOC642838	5.08	GMFG	2.57	Unknown	1.99
LYZ	4.97	PACRG	2.56	STAT2	1.98
CXCL10	4.89	HOXC6	2.55	HTATIP2	1.93
UBD	4.78	IRF1	2.55	PPP1R3F	1.93
GBP5	4.76	HLA-L	2.54	SAMSN1	1.93
IGKC	4.61	NKG7	2.54	NFKBIL2	1.92
Unknown	4.57	Unknown	2.54	Unknown	1.92
PTPRC	4.57	CASP1	2.53	C5orf20	1.92
LOC100293559	4.53	FKBP11	2.52	CTLA4	1.91
IGKV3-20	4.50	FAM173A	2.49	ZAP70	1.91
IGLC1	4.45	VNN2	2.49	EFR3A	1.91
LOC100289290	4.44	ZDHHC2	2.48	DNAH	1.90
LOC100287372	4.39	HLA-H	2.48	IL15RA	1.90
OCLN	4.36	TARP	2.47	OS9	1.90
IL23A	4.10	PSMB8	2.46	PSME1	1.90
GZMA	4.07	TAP1	2.45	SAMHD1	1.90
GBP3	4.06	PIM2	2.44	Unknown	1.90
CXCL11	4.02	AIF1	2.43	C1QA	1.88
BMS1P4	4.00	Unknown	2.42	TAGAP	1.87
GPR171	3.99	HLA-F	2.42	ICAM2	1.87
Unknown	3.90	PLTP	2.41	RPL26	1.87
TDR @	3.88	FYB	2.40	LOC100128203	1.85
PSMB9	3.86	HLA-DMA	2.40	CNTNAP5	1.84
LOC642424	3.80	CNN2	2.40	BTN3A1	1.84
LOC100130100	3.80	EPSTI1	2.40	SH2D2A	1.83
IL7R	3.74	Unknown	2.40	PSMB9	1.82
IGF2	3.73	TRAF3IP3	2.39	IGHA1	1.82
MAFG	3.72	LOC440839	2.39	CRTAM	1.82
CD48	3.65	WIPF1	2.39	Unknown	1.81
CCR2	3.64	CYBB	2.38	ZNF90	1.81
ISG20	3.63	ARHGAP25	2.38	GCET2	1.80
GIMAP5	3.61	IGKC	2.38	ZBP1	1.80
CD74	3.59	HLA-L	2.38	NR3C1	1.79
LCK-	3.58	TFEC	2.37	LILRB1	1.78
LTB	3.56	RPS27	2.37	HMGN4	1.78
GIMAP6	3.50	NCF1	2.36	C1QC	1.77
IRF4	3.47	POU2F2	2.35	S1PR4	1.77
GNLY	3.40	EMB	2.35	LOC439949	1.76
IFI44L	3.39	ICAM3	2.34	PUM2	1.76
MMP10	3.38	GIMAP4	2.33	LEPROTL1	1.76
CD247	3.37	HLA-G	2.33	ACAP1	1.75
LGALS2	3.33	Unknown	2.32	RASSF2	1.75
Unknown	3.31	SLAMF1	2.31	RPL38	1.75

ZNF341	3.27	EVI2B	2.30	C2orf60	1.74
GIMAP7	3.27	SP110	2.30	NR1H3	1.71
B2M	3.23	DNAJC5G	2.30	CYP2S1	1.71
FGF19	3.21	NYNRIN	2.29	GCH1	1.71
STAT1	3.18	MBP	2.28	FNBP1	1.70
ITGAL	3.18	SIT1	2.28	TRIM25	1.70
CARD16	3.12	ROBO4	2.28	CD97	1.70
CD2	3.12	SP140	2.28	TBCC	1.69
ABCC6	3.10	APOBEC3G	2.27	LAP3	1.69
ITM2A	3.07	PRKCB	2.27	Unknown	1.69
IGJ	3.07	BTG1	2.27	Unknown	1.68
IGLV6	3.05	TMEM149	2.26	AQP12A	1.64
HLA-B	3.05	MIR155HG	2.26	IL2RA	1.63
B2M	3.03	AIM2	2.26	PFN1	1.63
FLJ40330	3.00	WARS	2.26	NBPF15	1.62
ASB16	2.98	IGKC	2.23	XCL2	1.61
P2RX5	2.97	LILRB2	2.22	CR1	1.61
CTSC	2.97	PSME2	2.22	LAX1	1.61
FLJ40330	2.95	HLA-C	2.20	FGD2	1.58
CD79A	2.95	ARHGAP9	2.20	C19orf35	1.58
CST7	2.95	TNFSF13B	2.20	NCRNA00183	1.51
HLA-DPA1	2.94	PPP1R16B	2.19	CD28	1.49
LCP1	2.92	Unknown	2.19	IL18RAP	1.48
ERICH1	2.91	RPS27	2.17	EDEM1	1.48
HLA-DRA	2.90	IFIT5	2.16	FPR3	1.45
ARL8A	2.90	IL21R	2.16		

Table 8. List of down-regulated genes in tumor lesions of patients with relapse-free survival

Gene name	Fold-change	Gene Name	Fold-Change
C1QTNF3	0.21	SLC35E1	0.53
BHMT2	0.36	FAM108C1	0.54
JDP2	0.38	CERCAM	0.55
GJB2	0.39	RIC8A	0.55
PYGL	0.41	PTDSS2	0.56
PER3	0.42	LOC650226	0.57
COMP	0.42	MAN1B1	0.59
LOC100132800	0.43	TK2	0.59
ULK1	0.43	VPS37C	0.60
C14orf45	0.45	LOC116437	0.60
TCEAL3	0.45	NP_001013662.1	0.60
BBS2	0.46	PHF13	0.60
TMC4	0.47	CTSF	0.60
TNFAIP8L1	0.49	SLC16A5	0.61
COPS7A	0.49	GPR107	0.61
IER2	0.50	IFT20	0.62
ITGB5	0.50	SLC12A4	0.62
C1QTNF6	0.51	CREB5	0.62
VKORC1	0.52	PHF2	0.63

ADAM20	0.52	ZFP14	0.63
C22orf39	0.53	MYOM1	0.64
SIRPB1	0.53	SNUPN	0.64
VPS24	0.53	ICMT	0.64
PYROXD2	0.53	ANXA4	0.67
GFPT1	0.53	GATSL3	0.67

A heat map of the gene signature, identified by Student's *t* test $p < 0.001$ showed a perfect segregation of the two groups of patients (not shown). Interestingly, none of these genes was found in the 70-gene MammaPrint signature or 16-gene Oncotype DX panel. Moreover, among the selected 9797 genes derived from 80% filtering, we could find 16 (out of 70) and 13 (out of 16) genes belong respectively to MammaPrint and Oncotype DX panel. An unsupervised cluster based on the 16 MammaPrint genes and on 13 Oncotype DX genes did not show a clear segregation between relapse and relapse free groups (not shown). On the contrary, Complete Leave-One-Out Cross Validation (LOOCV) based prediction model applied to the 349 genes derived from the Student's *t* test $p < 0.001$ identified the genes immunoglobulin kappa chain (IGKC or IGK@), glutamate binding protein 1 (GBP1), signal transducer and activator of transcription 1 (STAT1), immunoglobulin lambda chain 5 (IGLL5), and occludin (OCLN) as best predictor of relapse or relapse-free survival (Figure 14). Among these 5 genes, the first four were selected in all LOOCV models, while OCLN was the only missed one in all LOOCV models, reflecting very stable feature sets. Unsupervised clustering based on these five genes showed a perfect segregation between relapsed and relapse free patients.

Detection of distinct immune function pathways at the time of diagnosis can predict relapse-free survival in breast cancer patients The pathway analysis was performed on the 349 genes in the supervised comparisons using the gene set expression comparison kit implemented in BRB-Array-Tools. The human pathway lists determined by "Ingenuity System Database" was selected. Samples with no recurrence showed significant up-regulation of genes involved in antigen presentation pathway, allograft rejection, graft versus host disease (GVHD), B cell development, dendritic cell maturation and interferon signaling (Figure 15). Interestingly, genes involved in T cell apoptosis, immunodeficiency signaling, CTLA-4 signaling and production of NO and reactive oxygen species were also up-regulated in the tumor specimens of patients with relapse-free survival.. An independent

cohort of patients were also included in the validation group and confirmatory real-time PCR was performed on the selected genes representing IFN-stimulated genes (ISGs) and T cell response (CXCL10, GZMA, GZMB, IL-7 Rα, IRF1) as well as the 5-gene signatures identified as best predictor of relapse or relapse-free survival. We used RNAs extracted from tumor lesions of 12 patients who served as validation group (7 with relapse-free survival and 5 with relapse). Patients with relapse-free survival had significantly higher expression of the immune function genes (>85% with an exception of patient#6) compared to patients with relapse (Figure 16).

In order to determine cellular sources of genes representing immune function pathways, IHC analysis of paraffin-embedded tumor specimens was performed according to the availability of commercial Abs and also intensity of the genes that would allow detection of their protein products. The IHC further confirmed higher expression of CXCL10, STAT1, GBP1, GZMA, and CD19 in the relapse-free group compared to those from the relapse group: human tonsils are shown as positive controls (Figure 17A). CXCL10 was expressed both in infiltrating cells and tumor cells of relapse-free patients while it was weakly expressed in tumor cells of patients with relapse. STAT1 showed strong staining in infiltrating cells and tumor cells of relapse-free group while it was expressed to a lesser extent mainly in infiltrating cells of the relapse group. GBP1 was expressed primarily in the infiltrating cells and also in tumor cells of the relapse-free group while it was weakly expressed only in tumor cells of the relapse group. GZMA was barely detectable even in human tonsils, yet it was detected only in tumor-infiltrating cells of the relapse-free group. The CD19 positive infiltrating cells were also present at a higher frequency in the tumor lesions of patients with relapse-free survival compared to only scattered presence in those with tumor relapse. Expression level of the proteins was quantified by showing the proportion of tumor infiltrating cells that stained positive. As shown in Figure 17B, there were significant differences between the two groups in the expression of CXCL10 ($p= 0.013$), STAT1 ($p= 0.033$), GBP1 ($p= 0.005$), GZMA ($p= 0.010$) and CD19 ($p= 0.046$).

DISCUSSION

We showed that microarray analysis of breast tumor specimens with tumor-infiltrating cells can detect a network of 349 genes which included 299 genes encompassing immune function genes that had a prognostic value in a diverse cohort of

breast cancer patients. Importantly, among these genes a 5-gene signature IGKC, GBP1, STAT1, IGLL5, and OCLN] was identified as best predictor of relapse-free survival with >85% accuracy. Although limitation with the availability of clinical specimens and patients outcome data did not allow including a larger cohort of specimens in the validation group, such a significant predictive value in a diverse cohort of patients is promising and has led to an ongoing prospective studies.

The network of immune function genes that were exclusively up-regulated in the tumor lesions of breast cancer patients with relapse-free survival included those involved in B cell development, interferon signaling associated with allograft rejection and autoimmune reaction, antigen presentation pathway, and cross talk between adaptive and innate immune responses. On the other hand, these genes were down-regulated in tumor specimens of patients with subsequent relapse, compared to those in the standard PBMC. Interestingly, genes involved in primary immunodeficiency signaling, T cell apoptosis, CTLA4 signaling and production of NO and reactive oxygen species were also up-regulated in the tumor specimens of relapse-free patients. Such paradoxical findings as to simultaneous up-regulation of immune effector genes and immune suppressor genes may suggest that tumor-derived factors were responsible for the expression of immune suppressor genes thereby facilitating cancer progression even in the presence of the increased immune effector genes. However, removal of breast tumors by conventional therapy must have eliminated the source of immune suppressive factors and resulted in down-regulation of the suppressor genes; subsequently, the immune effector genes may have protected the patients from their residual micrometastases and relapse. Previous finding have shown that primary tumors secrete soluble factors including GM-CSF, VEGF, and MCP-1 that result in increases in myeloid-derived suppressor cells (MDSCs) such that successful immunotherapy was not possible unless MDSC were depleted or reduced by chemotherapy or shrinking of the primary tumors. Real-time PCR analysis of selected genes in a validation cohort of patients showed significant differences between two groups in the expression of immune function genes. Importantly, variability within a group in the expression of a single immune function gene suggests that a signature rather than a single gene can predict relapse-free survival.

This novel signature associated with favorable outcome included 299 genes encompassing the immune function genes that were distinct from the 70-gene MammaPrint

signature and from 16-gene signature of the Oncotype DX panel. Moreover, an unsupervised clustering based on MammaPrint and Oncotype DX genes did not show a clear segregation between relapsed and relapse-free groups. Oncotype DX was originally validated in patients with ER+ and node negative tumors, though it is now being expanded to patients with node positive breast cancer. Therefore, it was not surprising that Oncotype DX could not segregate the patients in this study, because of majority of these patients were ER negative and/or node positive. Interestingly, a perfect segregation was shown by an unsupervised cluster analysis based on 5 genes (IGKC, GBP1, STAT1, IGLL5, and OCLN) among the 299 genes which were identified by the LOOCV prediction model as best predictor of relapse or relapse-free survival. Real-time PCR also confirmed higher expression of these 5 genes in a greater than 85% of patients with relapse-free survival.

The IHC analysis of tumor specimens further confirmed upregulation of the immune function genes representing immune mechanism pathways mainly in tumor infiltrating cells of the relapse-free group. For instance, CXCL10 and GBP1 are IFN-stimulated genes (ISGs) that showed strong staining in tumor infiltrating cells of relapse-free patients compared to the relapsed group. CXCL10 binds CXCR3 on DCs, macrophages and T cells. Increased expression of CXCL10 in tumor lesions of relapse-free patients may suggest CXCL10-induced DC maturation and antigen cross presentation that results in Th1-type immune responses [28]. GBP1 is a key mediator of angiostatic effects of the immune responses, inflammation in particular, and its expression in the tumors and tumor infiltrating immune cells is associated with favorable prognosis [29], as was the case in our study. As expected, upregulation of these ISGs in relapse-free patients compared to relapsed groups was associated with higher expression of STAT1 and IRF1 genes as well as an increased expression of nuclear STAT1 in their tumor infiltrating cells. However, nuclear expression of STAT1 in tumor cells was comparable between the two groups. This may explain progression of primary breast cancer in the two groups. It was shown that increased nuclear STAT1 resulted in the induction of apoptosis in the tumors [31] as well as the metastatic ability of the tumors that escaped from apoptosis [31]. GZMA and GZMB which are involved in T cell responses and the graft-versus-host disease (GVHD) pathway, as well as CD19, which is involved in the B cell response were also uniformly increased in tumor lesions of relapse-free patients. Although there was a significant difference between the two groups in the expression of each molecule, a panel of differentially expressed molecules, i.e.

signature of immune function genes provides more valid prognostic marker compare to one molecule alone.

In summary, we performed comparative analysis between two cohorts of patients and also validated our signature in the two cohorts to determine whether some patients with relapse may also show expression of the immune function genes in their tumors. Our data suggest that the 5-gene signature of immune response (IGKC, GBP1, STAT1, IGLL5, and OCLN) can not only be used as prognostic biomarker for breast cancer patients it would also offer therapeutic strategies for preventing breast cancer recurrence. Absence of the 5-gene signature in the tumor lesions at the time of diagnosis suggest that patients may be at a high risk of relapse.

EXAMPLE 5. Cancer Testis Antigens as prognostic biomarkers for breast cancer patients

The previous Examples descried the finding that an immune function gene signature at the tumor lesions of patients with early stage breast cancer could predict relapse-free survival following conventional therapies. We hypothesized that expression of cancer testis antigens (CTA) may be responsible for converting weakly immunogenic breast tumors into highly immunogenic tumors, and result in relapse-free survival. To test this hypothesis, we performed qRT-PCR analysis of RNA extracted from tumor lesions of 2 cohorts of patients with breast cancer. The first group relapsed within 1-3 years or remained relapse-free for 4-5 years (Figure 18A). The second group relapsed within 15-6 years or remained relapse-free for 6-7 years (Figure 18B). We detected an increased expression of a number of CTA in tumor lesions of patients with relapse-free but not in those with tumor relapse. Namely, the genes MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NYESO-1, SP17 and SPANXb were upregulated and expressed in relapse-free patients.

We also showed that treatment of human breast tumor cell lines with a demethylating agent, decitabine, induced expression of CTA in the tumors. Together, these data suggest that lack of CTA expression in tumor lesions of breast cancer patients at the time of diagnosis predict high risk of tumor relapse, and that using Decitabine in a neoadjuvant setting may convert patients with high risk into those with low risk of tumor relapse (see Example 7 below).

EXAMPLE 6. An immunological biomarker as a predictor of therapeutic efficacy in patients with advanced breast cancer.

Data presented in the preceding Examples suggest that a gene signature which includes immune function genes and CTA, may serve as a predictor or surrogate of therapeutic efficacy in breast cancer. This Example describes the development of a test system that can predict therapeutic efficacy of conventional therapies and immunotherapy in patients with locally advanced or metastatic breast cancer. In order to develop the test, retrospective studies are conducted in patients with locally advanced tumor or metastatic breast cancer for outcome data id available. Detection of s gene signature in the tumor tissue can be used as predictor and surrogate of the efficacy of conventional therapies whereas lack of the signature would predict poor outcome defined by tumor recurrence following conventional therapies.

Briefly, a total of 408 tumor specimens are collected from patients with advanced breast cancer in the past 10 years as well as associated outcome data. Among these patients, ~50% have tumor relapse. Using an algorithm, a cut off score is developed for the expression of the signature that can determine whether a patient with advanced breast cancer (stage III-IV) will or will not respond to conventional therapies or neoadjuvant immunotherapy. A 13-gene signature gene panel will be used: 5 immune function genes (IGKC, IGLLS, STAT1, GBP1 and OCLN) and 8 CTA (MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb) Preliminary data described in the Examples above showed that this 13-gene signature is an independent predictor of outcome regardless of age, sex, race, tumor size, nodal status, the status of ER, PR, HER-2/neu and neoadjuvant or adjuvant chemotherapy.

MATERIALS AND METHODS

Clinical specimens. Frozen tumor specimens as well as FFPE tissues have been collected over the past 10 years with corresponding annotated patient outcome data available. Retrospective studies are conducted in 408 breast cancer patients with locally advanced (stage III) or metastatic tumors (stage IV) (229 patients who did not respond to chemotherapy and 179 patients who showed prolonged survival after chemotherapy).

Extraction of RNA from FFPE tumor specimens. Recovery and extraction of RNA from FFPE tissues provides a number of challenges because of RNA degradation and its cross linking to other molecules as a result of the addition of hydroxymethyl groups and dimerization through methylene bridges. To overcome such problems, protocols for

Agencourt FormPure and the MagMAX 96 for microarrays are combined in a semi-automated fashion on the MagMAX Express instrument. The protease digestion conditions of the kit are combined with the use of magnetic beads designed to release a maximal amount of RNA of all sizes. Automated magnetic particle processors, such as the Applied Biosystems MagMAX Express, are used to perform total RNA isolation, both from fresh-frozen and from FFPE specimens, with high quality, as assessed by the RNA Integrity Number (RIN) from Bioanalyzer runs. RIN values are smaller in FFPE specimens compared to fresh-frozen samples. Nevertheless, partially degraded RNA is still a valid template for qRT-PCR, since small amplicons are generated. RNA isolated from the FFPE specimens is subjected to TaqMan Gene Expression Assays corresponding to the gene signature of interest.

Expression pattern of a 13-gene panel (5 immune function + 8 CTA) in tumor specimens of patients with advanced breast cancer. Five immune function genes representing T cell pathway (STAT1, GBP1, OCLN), B cell pathway (IGKC, IGLL5) as well as a panel of 8 human CTA (MAGE-a3, MAGE-a4, MAGE-a-5, MAGE-a6, Mage C1, NY-ESO1, AKAP4, and SPNAKXb) are included in qRT-PCR TaqMan analysis. Expression of these genes is normalized to five housekeeping reference genes: *ACTB* (*b*-actin), *GAPDH* (glyceraldehyde 3-phosphate dehydrogenase), *GUS* (beta-glucuronidase), *RPLPO* (large ribosomal protein), and *TFRC* (transferrin receptor protein 1). Test data (not shown) has shown that reference-normalized expressions range from 0 to 10, with a 1-unit increase reflecting a doubling of RNA.

Algorithm and statistical analysis. The 408 specimens are divided into two groups, with 40% randomized into a test group and 60% randomized into a validation group. Samples from patients who did not respond to chemotherapy are randomized separately from samples from patients who had showed a prolonged cancer free survival to ensure equal proportions in the two sets (test and validation groups). For the test group data (92 chemotherapy non-responders; 72 chemotherapy responders), a logistic discriminant function is used to create the classification algorithm, since logistic regression is optimal for categorical outcomes like remission status (30, 31). The resulting model yields estimates of the probability of chemotherapy response ($\hat{\pi}_i$) as

$$\hat{\pi}_i = \frac{1}{1-\exp(u_i)}, u_i = \beta_0 + \sum_{j=1}^{5} \beta_j IF_{ji} + \sum_{k=1}^{8} \beta_k CTA_{ki} + \theta x_i$$

where both immune function genes (IF_{ji}) and human CTA genes (CTA_{ki}) are included as binary indicators (1 or 0) of whether the gene is up-(or down-)wardly expressed in the i^{th} subject by 2-fold as compared to 5 housekeeping genes, x_i is a vector consisting of any covariates (age, gender, tumor stage, and ER/PR/HER2 status) related to chemotherapy response status, and the β and θ are the estimated regression parameters. Subjects are classified as "likely to respond to chemotherapy" if the probability ($\hat{\pi}_i$) is greater than a value P (defined below). A ROC curve is estimated by varying the cut-off probability for classification (P) from 0.05 to 0.95, with the optimal value of P chosen to maximize the distance from the ROC curve to the chance line. The validation group (137 chemotherapy non-responders; 107 chemotherapy responders) are used to validate the classification algorithm by entering their 5 IF and 8 CTA gene expressions into the generated algorithm to determine the likelihood of response to chemotherapy; if that probability is above the threshold P, then that patient is deemed to have a poor outcome, otherwise the patient is not deemed to have a poor outcome.

The algorithm is successful in that >75%, or 80%, or 85%, or >90% of the subjects from the validation group are correctly classified based on their remission status.

EXAMPLE 7. An immunological biomarker as a predictor of the efficacy of neoadjuvant immunotherapy

Induction of tumor antigen-specific immune responses in patients with breast cancer is not always associated with an objective response and relying on the immune response genes or genes that are involved in tumor invasion and metastasis alone cannot accurately predict an effective response to therapy. However, focusing on a gene signature which is involved in tumor-immune interactions should provide a more reliable predictor of response to therapy. The 13-gene signature described herein is indicative of a favorable response to conventional therapy. Logically then, induction of expression of the genes involved in the signature should protect patients from relapse.

Decitabine is a demethylating pro-drug that has shown efficacy in patients with hematologic malignancies, particularly against myelodisplastic syndrom (MDS). Its efficacy

has been attributed to the induction of tumor suppressor genes and CTA. Activation of Decitabine by deoxycytidine kinase deoxycytidine kinase (DCK), which is selectively expressed in tumor cell and myeloid cells of some, but not all patients, leads to its incorporation into newly synthesized DNA strands during the S-phase of the cell-cycle. When a decitabine-containing DNA strand binds to the enzyme DNA methyltransferase (DNMT1), the decitabine in the strand forms a covalent complex with a serine residue at the DNMT1 active site, resulting in its inactivation. This in turn results in hypomethylation of genes in the surrounding area.

When decitabine is administered to a cancer patient, if the patient is DCK positive and thus expresses DCK, then hypomethylation of genes involved in cancer should occur, e.g. the tumor suppressor genes and the CTA antigens of the 13-gene signature, rendering the tumor cells susceptible to immune-mediated apoptosis. Further, the selective expression of DCK in tumor cells and myeloid cells should prevent T and B cells from the demethylating effects of Decitabine. Notably, DCK is generally overexpressed in poor outcome breast cancer patients. Taken together, this indicates that poor outcome patients whose tumors lack expression of the 13-gene signature are likely to respond to Decitabine for the expression of CTA and in turn the induction of CTA-reactive immune responses. When applied in a neoadjuvant setting, such a therapy triggers the induction of the 13-gene signature. When the tumor burden is reduced by standard therapy, induction of the 13-gene signature ultimately results in the elimination of minimal residual disease and prevents metastatis and/or relapse.

MATERIALS AND METHODS

Patients' specimens. Two needle biopsy specimens are obtained from locally advanced and/or metastatic breast cancer patients; one sample at the time of diagnosis and another sample at the initiation of standard adjuvant therapy (or 10 days after neoadjuvant Decitabine or Decitabine + IL-2). RNA is extracted from fine-needle aspiration specimens. Patients whose tumors express DCK, determined by qRT-PCR, are eligible to receive neoadjuvant Decitabine. A 100 ml blood is drawn at the time of diagnosis and 10 days after Decitabine therapy, or on the day of starting standard adjuvant therapy.

Treatment schedule. The following hematological parameters are required for the patients: Hb $\geq$ 9 g/dl untransfused; platelet count $\geq 100 \times 10^9$/liter untransfused; and an absolute neutrophil count of $>1.5 \times 10^9$/liter without growth factors. Decitabine is given as a daily s.c.

injections of a safe dose of 0.25 mg/kg for 5 days (Eisai Inc.). If Decitabine-induced CTA expression is not sufficient to induce the immune function genes in a patient, the patient also receives low-dose IL-2 plus Decitabine in order to augment the immune response gene signature. Injections of low-dose IL-2 (1 μg/kg Proleukin, Prometheus Laboratories) are given s.c. in the abdominal region for four consecutive days, starting 24 h after Decitabine injections.

qRT-PCR. Since we only need 10 ng of total RNA per sample, small samples from metastatic tumors do not offer any setback. Five immune function genes representing T cell pathway (STAT1, GBP1, OCLN), B cell pathway (IGKC, IGLL5) as well as a panel of 8 human CTA (Mage A3-6, Mage C1, NY-ESO1, AKAP4, SPNAKXb) and, optionally, SLIP1 and SP17, are included in qRT-PCR TaqMan analysis. Expression of these genes is normalized to five housekeeping reference genes (*ACTB*, *GAPDH*,*GUS*, *RPLPO*, and *TFRC*). Reference-normalized expressions typically range from 0 to 10, with a 1-unit increase reflecting a doubling of RNA. DNA is also extracted to determine hypomethylation of CTA promoter using bisulfite genomic sequencing.

CTA-reactive immune responses.

IFN-γ ELISA and flow cytometry analysis is performed to determine T cell responses to human recombinant NY-ESO-1 or MAGE-A4. Monocyte-derived DC and lymphocytes are prepared. Based on experience with patients with multiple myeloma, an optimal dose of 10 ug/ml antigen in a T cell: DC ratio of 4:1 is sufficient to determine T cell responses in a 24 hr culture in vitro. In order to determine cellular source of the antigen-specific IFN-γ production, a three color flow cytometry analysis is performed using FITC-CD3, APC-CD8, APC-CD4, and PE-IFN-γ antibodies (Biolegend).

Statistical analysis. The primary ouatcome of treatment is induced expression of a panel of CTA (8 of 10 CTA) and the 5-gene signature of immune function at the tumor site. A covariance analysis compares tumor specimens before and one week after neoadjuvant therapy between both those patients receiving Decitabine (or plus IL-2). The baseline tumor specimen levels and therapy (as well as an interaction term) are included as model predictors, with post-therapy tumor specimens as the response. To adjust for multiple comparisons across the 8 CTA and 5 immune function genes, we can adjust the overall

false-positive rate (two-sided $\alpha = 0.05$) using the step-down approach (48), where the gene with the largest tumor specimen increase uses significance level ($\alpha/13$), while the gene with the smallest tumor specimen increase uses (α). IL-2 may optionally be administered with Decitabine as described herein. In addition, the presence of CTA-specific T cell or antibody responses is determined in the blood of patients before and one week after neoadjuvant therapy using an analysis of covariance. Disease free survival is demonstrated with Kaplan-Meier curves for patients who receive either neoadjuvant immunotherapy or standard therapy, which are compared using a log-rank test.

If patient experiences relapse despite neoadjuvant therapy, the expression of CTA at the time of relapse can be used to determine whether hypomethylation of the CTA promoter is reversible. In such cases, neoadjuvant therapy is combined with histone deacetylation inhibitors to achieve a sustained hypomethylation of CTA.

RESULTS

A 13-gene signature immunological biomarker is used as predictor and surrogate of the efficacy of neoadjuvant immunotherapy. The 13-gene signature serves as a predictor or a surrogate of the efficacy of neoadjuvant therapy in patients with advanced breast carcinoma who had failed to respond to initial adjuvant chemotherapy. Patients whose tumors express deoxycitidine kinase (DCK) were eligible for the treatment, because Decitabine is a pro-drug that requires DCK for activation. Decitabine is administered in a neoadjuvant setting to induce CTA expression, in situ, and in turn trigger the induction of CTA-reactive immune responses (in situ immunization) prior to standard therapies. If Decitabine-induced CTA expression fails to induce all 5 immune function genes in a patient, that patient receives low dose IL-2 plus Decitabine in order to augment the immune response gene signature. Needle biopsy specimens are analyzed before and after neoadjuvant therapy (at the time of standard adjuvant therapy) in order to determine whether this treatment induced expression of CTA and the immune function genes. Blood samples also are analyzed before and after neoadjuvant treatment in order to determine the induction of CTA-reactive T cell and antibody responses. Patients with triple negative tumors are treated because these patients would otherwise relapse faster.

These studies demonstrate an immunological biomarker that can predict who will benefit from neoadjuvant immunotherapy.

EXAMPLE 8. Clinical applications

During a routine mammogram or during routine self-examination, a "lump" is discovered in the breast tissue of a human patient. Diagnostic biopsy samples are taken from the lump and from the surrounding tissue and it is determined that the patient has breast cancer. In addition to routine biopsy analysis, the samples are assessed using the methods of the invention to determine whether the patient is likely to relapse after treatment or is likely to be relapse free after treatment, as follows:

Scenario 1: Analysis of the gene expression patterns in the microenvironment of the tumor reveal a pattern of gene expression that is biased toward upregulation of the 299 genes listed in Table 4. As a result of this finding, the patient's health care team concludes that the prognosis for the patient (e.g. after removal of the tumor) is favorable, and that recurrence (relapse) is unlikely. Recommended treatment may include, for example, conservative surgical removal of the tumor, vaccination and/or drugs that boost the immune system such as revlimid. But the patient may be spared the inconvenience and discomfort of more aggressive therapy such as mastectomy, chemotherapy and radiation therapy.

Scenario 2: Analysis of the gene expression patterns in the microenvironment of the tumor reveal a pattern of gene expression that is biased toward upregulation of the 50 genes listed in Table 5. As a result of this finding, the patient's health care team concludes that the prognosis for the patient (e.g. after removal of the tumor) is not favorable, and that recurrence is likely. Recommended treatment includes: extensive surgery to remove surrounding tissues, using an aggressive chemotherapeutic regimen (e.g. drugs such as fludarabine, cyclophosphamide, IFN-γ, fludarabine, cyclophosphamide, amd gemcitabine) and aggressive radiation therapies.

Scenario 3: Analysis of the gene expression patterns in the tumor of microenvironment of the tumor reveal a pattern of gene expression that is biased towards expression and/or upregulation of the 5 genes in the 5-gene signature and the 8 genes in the 8-gene signature. As a result of this finding, the patient's health care team concludes that the prognosis for the patient after initial conventional tumor removal is favorable, and that recurrence is not likely. Recommended treatment may include, for example, conservative surgical removal of the tumor, and/or chemo- or hormone therapy to shrink the tumor, or even vaccinogens and/or drugs that boost the immune system such as revlimid. But the patient may be spared the inconvenience and discomfort of more aggressive therapy such as mastectomy.

Scenario 4: Analysis of the gene expression patterns in the tumor or microenvironment of the tumor reveal the absence of expression of 5 genes of the 5-gene signature and the 8 genes in the 8-gene signature. As a result of this finding, the patient's health care team concludes that the prognosis for the patient after initial conventional tumor removal or reduction in volume is not favorable, and that recurrence is likely. As a result, aggressive treatment measures are taken, e.g. extensive surgery to remove surrounding tissues, use of an aggressive chemotherapeutic regimen (e.g. drugs such as fludarabine, cyclophosphamide, IFN-γ, fludarabine, cyclophosphamide, amd gemcitabine), aggressive radiation therapies, etc. In addition, if the patient is DCK positive, then she receives neoadjuvant treatment with decitabine to convert her gene expression profile to expression of the genes of the signature(s).

The complete contents of all references, patent applications and issued patents cited herein are hereby incorporated by reference.

While the invention has been described in terms of its preferred embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims. Accordingly, the present invention should not be limited to the embodiments as described above, but should further include all modifications and equivalents thereof within the spirit and scope of the description provided herein.

REFERENCES

1. Marincola FM, Rivoltini L, Salgaller ML, Player M, Rosenberg SA. Differential anti-MART-1/MelanA CTL activity in peripheral blood of HLA-A2 melanoma patients in comparison to healthy donors: evidence for in vivo priming by tumor cells. J Immunother 1996;19:266-77.

2. D'Souza S, Rimoldi D, Lienard D, Lejeune F, Cerottini JC, Romero P. Circulating Melan-A/Mart-1 specific cytolytic T lymphocyte precursors in HLA-A2+ melanoma patients have a memory phenotype. Int J Cancer 1998;78:699-706.

3. Lee K-H, Wang E, Nielsen M-B, et al. Increased vaccine-specific T cell frequency after peptide-based vaccination correlates with increased susceptibility to in vitro stimulation but does not lead to tumor regression. J Immunol 1999;163:6292-300.

4. Marincola FM. A balanced review of the status of T cell-based therapy against cancer. J Transl Med 2005;3:16.

5. Panelli MC, Riker A, Kammula US, et al. Expansion of Tumor-T cell pairs from Fine Needle Aspirates of Melanoma Metastases. J Immunol 2000;164:495-504.

6. Zippelius A, Batard P, Rubio-Godoy V, et al. Effector function of human tumor-specific CD8 T cells in melanoma lesions: a state of local functional tolerance. Cancer Res 2004;64:2865-73.

7. Rosenberg SA, Sherry RM, Morton KE, et al. Tumor progression can occur despite the induction of very high levels of self/tumor antigen-specific CD8+ T cells in patients with melanoma. J Immunol 2005;175:6169-76.

8. Disis ML, Calenoff E, McLaughlin G, et al. Existent T-cell and antibody immunity to HER-2/neu protein in patients with breast cancer. Cancer Res 1994;54:16-20.

9. Lee PP, Yee C, Savage PA, et al. Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients. Nat Med 1999;5:677-85.

10. Monsurro' V, Wang E, Yamano Y, et al. Quiescent phenotype of tumor-specific CD8+ T cells following immunization. Blood 2004;104:1970-8.

11. Nagorsen D, Voigt S, Berg E, Stein H, Thiel E, Loddenkemper C. Tumor-infiltrating macrophages and dendritic cells in human colorectal cancer: relation to local regulatory T cells, systemic T-cell response against tumor-associated antigens and survival. J Transl Med 2007;5:62.

12. Marincola FM, Jaffe EM, Hicklin DJ, Ferrone S. Escape of human solid tumors from T cell recognition: molecular mechanisms and functional significance. Adv Immunol 2000;74:181-273.

13. Marincola FM, Wang E, Herlyn M, Seliger B, Ferrone S. Tumors as elusive targets of T cell-based active immunotherapy. Trends Immunol 2003;24:335-42.

14. Pages F, Berger A, Camus M, et al. Effector memory T cells, early metastasis, and survival in colorectal cancer. N Engl J Med 2005;353:2654-66.

15. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science 2006;313: 1960-4.

16. Galon J, Fridman WH, Pages F. The adaptive immunologic microenvironment in colorectal cancer: a novel perspective. Cancer Res 2007;67:1883-6.

17. Zhang L, Conejo-Garcia JR, Katsaros D, et al. Intratumoral T cells, recurrence, and

survival in epithelial ovarian cancer. N Engl J Med 2003;348:203-13.

18. Kalialis LV, Drzewiecki KT, Mohammadi M, Mehlsen AB, Klyver H. Spontaneous regression of metastases from malignant melanoma: a case report. Melanoma Res. 2008 Aug;18(4):279-83.

19. Dudley ME, Wunderlich JR, Yang JC, et al. Adoptive cell transfer therapy following non-myeloablative but lymphodepleting chemotherapy for the treatment of patients with refractory metastatic melanoma. J Clin Oncol 2005;23:2346-57.

20. Muranski P, Boni A, Wrzesinski C, Citrin DE, Rosenberg SA, Childs R, Restifo NP. Increased intensity lymphodepletion and adoptive immunotherapy--how far can we go? Nat Clin Pract Oncol. 2006 Dec;3(12):668-81

21. Kmieciak M, Knutson KL, Dumur CI, Manjili MH. HER-2/neu antigen loss and relapse of mammary carcinoma are actively induced by T cell-mediated anti-tumor immune responses. Eur J Immunol 2007;37:675-85.

22. Manjili MH, Arnouk H, Knutson KL, et al. Emergence of immune escape variant of mammary tumors that has distinct proteomic profile and a reduced ability to induce "danger signals". Breast Cancer Res Treat. 2006;96:233-41.

23. Guy CT, Webster MA, Schaller M, Parsons TJ, Cardiff RD, Muller WJ. Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. Proc Natl Acad Sci U S A 1992;89:10578-82.

24. Takeuchi N, Hiraoka S, Zhou XY, et al. Anti-HER-2/neu immune responses are induced before the development of clinical tumors but declined following tumorigenesis in HER-2/neu transgenic mice. Cancer Res 2004;64:7588-95.

25. Kmieciak M, Morales JK, Morales J, Grimes M, Manjili MH. Danger signal and nonself entity of tumor antigen are both required for eliciting effective immune responses against HER-2/neu positive mammary carcinoma: implications for vaccine design. Cancer Immunol Immunother 2008 (In press).

26. Knutson KL, Almand B, Dang Y, Disis ML. Neu antigen-negative variants can be generated after neu-specific antibody therapy in neu transgenic mice. Cancer Res 2004;64:1146-51.

27. Wang E, Miller L, Ohnmacht GA, Liu E, Marincola FM. High fidelity mRNA amplification for gene profiling using cDNA microarrays. Nature Biotech 2000;17:457-9.

28. Wang E. RNA amplification for successful gene profiling analysis. J Transl Med

2005;3:28.

29. Jin P, Zhao Y, Ngalame Y, et al. Selection and validation of endogenous reference genes using a high throughput approach. BMC Genomics 2004;5:55.

30. Rubinfeld B, Robbins P, el Gamil M, Albert I, Porfiri E, Polakis P. Stabilization of beta-catenin by genetic defects in melanoma cell lines. Science 1997;275:1790-2.

31. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863-8.

32. Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines. Nature Genetics 2000;24:227-35.

33. Wang E, Miller LD, Ohnmacht GA, et al. Prospective molecular profiling of subcutaneous melanoma metastases suggests classifiers of immune responsiveness. Cancer Res 2002;62:3581-6.

34. Basil CF, Zhao Y, Zavaglia K, et al. Common cancer biomarkers. Cancer Res 2006;66:2953-61.

35. Kaech SM, Hemby S, Kersh E, Ahmed R. Molecular and functional profiling of memory CD8 T cell differentiation. Cell 2002;111:837-51.

36. Panelli MC, Stashower M, Slade HB, et al. Sequential gene profiling of basal cell carcinomas treated with Imiquimod in a placebo-controlled study defines the requirements for tissue rejection. Genome Biol 2006;8:R8.

37. Mantovani A, Romero P, Palucka AK, Marincola FM. Tumor immunity: effector response to tumor and the influence of the microenvironment. Lancet 2007;in press.

38. Sinha P, Clements VK, Bunt SK, Albelda SM, Ostrand-Rosenberg S. Cross-talk between myeloid-derived suppressor cells and macrophages subverts tumor immunity toward a type 2 response. J Immunol 2007;179:977-83.

39. Melani C, Chiodoni C, Forni G, Colombo MP. Myeloid cell expansion elicited by the progression of spontaneous mammary carcinomas in c-erbB-2 transgenic BALB/c mice suppresses immune reactivity. Blood 2003;102:2138-45.

40. Wu K, Bi Y, Sun K, Wang C. IL-10-producing type 1 regulatory T cells and allergy. Cell Mol Immunol 2007;4:269-75.

41. Biswas SK, Gangi L, Paul S, et al. A distinct and unique transcriptional programme expressed by tumor-associated macrophages: defective NF-{kappa}B and enhanced IRF-3/STAT1 activation. Blood 2006;107:2112-22.

42. Fenner JE, Starr R, Cornish AL, et al. Suppressor of cytokine signaling 1 regulates the immune response to infection by a unique inhibition of type I interferon activity. Nat Immunol 2006;7:33-9.

43. Mansell A, Smith R, Doyle SL, et al. Suppressor of cytokine signaling 1 negatively regulates Toll-like receptor signaling by mediating Mal degradation. Nat Immunol 2006;7:148-55.

44. Qin H, Wilson CA, Lee SJ, Benveniste EN. IFN-beta-induced SOCS-1 negatively regulates CD40 gene expression in macrophages and microglia. FASEB J 2006;20:985-7.

45. Li Y, Chu N, Rostami A, Zhang GX. Dendritic cells transduced with SOCS-3 exhibit a tolerogenic/DC2 phenotype that directs type 2 Th cell differentiation in vitro and in vivo. J Immunol 2006;177:1679-88.

46. Fojtova M, Boudny V, Kovarik A, et al. Development of IFN-gamma resistance is associated with attenuation of SOCS genes induction and constitutive expression of SOCS 3 in melanoma cells. Br J Cancer 2007;97:231-7.

47. Kroemer G, Hirsch F, Gonzalez-Garcia A, Martinez C. Differential involvement of Th1 and Th2 cytokines in autoimmune diseases. Autoimmunity 1996;24:25-33.

48. Sasaki S, Amara RR, Yeow WS, Pitha PM, Robinson HL. Regulation of DNA-raised immune responses by cotransfected interferon regulatory factors. J Virol 2002;76:6652-9.

49. Hunter CA. New IL-12-family members: IL-23 and IL-27, cytokines with divergent functions. Nat Rev Immunol 2005;5:521-31.

50. Afzali B, Lombardi G, Lechler RI, Lord GM. The role of T helper 17 (Th17) and regulatory T cells (Treg) in human organ transplantation and autoimmune disease. Clin Exp Immunol 2007;148:32-46.

51. Hao JS, Shan BE. Immune enhancement and anti-tumour activity of IL-23. Cancer Immunol Immunother 2006;55:1426-31.

52. Shan BE, Hao JS, Li QX, Tagawa M. Antitumor activity and immune enhancement of murine interleukin-23 expressed in murine colon carcinoma cells. Cell Mol Immunol 2006;3:47-52.

53. Fulton A, Miller F, Weise A, Wei WZ. Prospects of controlling breast cancer metastasis by immune intervention. Breast Dis 2006;26:115-27.

54. Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM (2000) High fidelity mRNA amplification. Nat Biotechnol 18: 457-459.

55. Wang E (2005) RNA amplification for successful gene profiling analysis. J Transl Med 3:28.

56. Jin P, Zhao Y, Ngalame Y, Panelli MC, Nagorsen D, Monsurró V, Smith K, Hu N, Su H, Taylor PR, Marincola FM, Wang E (2004) Selection and validation of endogenous endogenous reference genes using a high throughput approach. BMC Genomics 5: 55.

57. Wang E, Miller LD, Ohnmacht GA, Mocellin S, Perez-Diez A, Petersen D, Zhao Y, Simon R, Powell JI, Asaki E, Alexander HR, Duray PH, Herlyn M, Restifo NP, Liu ET, Rosenberg SA, Marincola FM (2002) Prospective molecular profiling of subcutaneous melanoma metastases suggests classifiers of immune responsiveness. Cancer Res 62: 3581–3586.

58. Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y (2007) Analysis of Gene Expression Data Using BRB-Array Tools. Cancer Inform 3: 11-17.

59. Worschech A, Kmieciak M, Knutson KL, Bear HD, Szalay AA, Wang E, Marincola FM, Manjili MH (2008) Signatures associated with rejection or recurrence in HER-2/neu-positive mammary tumors. Cancer Res 68: 2436-2446.

60. Boyman O, Hefti HP, Conrad C, Nickoloff BJ, Suter M, Nestle FO (2004) Spontaneous Development of Psoriasis in a New Animal Model Shows an Essential Role for Resident T Cells and Tumor Necrosis Factor. J Exp Med 199: 731–736.

61. Dominguez F, Martínez S, Quiñonero A, Loro F, Horcajadas JA, Pellicer A, Simón C (2008) CXCL10 and IL-6 induce chemotaxis in human trophoblast cell lines. Mol Hum Reprod 14: 423-430.

62. Sun W, Xu W, Snyder M, He W, Ho H, Ivashkiv LB, Zhang JJ (2005) The Conserved Leu-724 Residue Is Required for Both Serine Phosphorylation and Co-activator Recruitment for Stat1-mediated Transcription Activation in Response to Interferon-γ . J Biol Chem 280: 41844-41851.

63. Lundmark F, Duvefelt K, Iacobaeus E, Kockum I, Wallström E, Khademi M, Oturai A, Ryder LP, Saarela J, Harbo HF, Celius EG, Salter H, Olsson T, Hillert J (2007) Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. Nat Genet 39: 1108-1113.

CLAIMS

We claim:

1. An in vitro method for determining, in a cancer patient in need thereof, the likelihood of relapse, comprising the steps of

i) obtaining a tumor tissue sample from said cancer patient;

ii) quantifying a level of gene expression in said tumor tissue sample of at least one gene selected from the group consisting of MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb;

iii) comparing a quantification value for said level of gene expression of at least one of MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb obtained in said quantifying step with a predetermined reference value for a level of gene expression of at least one of MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb in control tissue samples; and

iv) providing a prognosis of a low likelihood of relapse for said patient when said quantification value is greater than said predetermined reference value; or

v) providing a prognosis of a high likelihood of relapse for said patient when said quantification value is lower than said predetermined reference value.

2. The method of claim 1, wherein said cancer is breast cancer.

3. The method of claim 1, further comprising the steps of

quantifying a level of gene expression in said tumor tissue sample for each gene of a genes set which includes each of IGKC, IGLL5, STAT1, GBP1, and OCLN;

comparing quantification values for said level of gene expression of IGKC, IGLL5, STAT1, GBP1, and OCLN obtained in said quantifying step with predetermined levels of gene expression for IGKC, IGLL5, STAT1, GBP1, and OCLN in control tissue samples, and

wherein said steps of providing a prognosis of a low likelihood of relapse or providing a prognosis of a high likelihood or relapse considers quantification levels of gene expression for said gene set being respectively greater or less than said predetermined levels.

4. The method of claim 1 wherein said at least one gene includes at least a plurality of the following: MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1

and SPANXb.

5. The method of claim 1 wherein said at least one gene includes each of the following: MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb.

6. The method of claim 1, wherein said at least one gene includes one or more housekeeping genes.

7. The method of claim 1, wherein said control tissue samples include tissue samples from one or more of subjects without cancer, subject with stage I cancer, subjects with stage II cancer, subjects with stage III cancer, subjects with stage IV cancer, subject who have not relapsed after receiving conventional cancer treatment, and subjects who have relapsed after receiving conventional cancer treatment.

8. A theranostic method for developing a treatment protocol for a cancer patient, comprising the steps of

i) obtaining a tumor tissue sample from said patient;

ii) quantifying a level of gene expression of at least one MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb in said tumor tissue sample;

iii) comparing a quantification value for a level of gene expression of at least one of MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb obtained in said quantifying step with a predetermined reference value for a level of gene expression of at least one of MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb in control tissue samples; and

iv) when said quantification value is greater than said predetermined reference value, providing a prognosis of a low likelihood of relapse for said patient and recommending a tumor-burden reducing treatment for said patient; or

v) when said quantification value is lower than said predetermined reference value, providing a prognosis of a high likelihood of relapse for said patient and recommending neoadjuvant therapy in conjunction with a tumor-burden reducing treatment for said patient.

9. The theranostic method of claim 8, wherein said tumor-burden reducing treatment includes one or more treatments selected from the group consisting of surgical removal of tumor tissue, reduction in tumor volume by chemotherapy, reduction in tumor volume by radiotherapy, and reduction in tumor volume by hormone therapy.

10. The theranostic method of claim 8, wherein said neoadjuvant therapy includes administration of one more agents selected from the group consisting of 5-azacytidine, decitabine, histone deacetylation inhibitors.

11. A system for determining a probability of relapse of a patient with a tumor, said system comprising

means for obtaining measurements of expression of genes in tumors;

means for recognizing, using said measurements, patterns of gene expression, wherein said patterns of gene expression are correlated with said probability of relapse; and

means for assigning a probability of relapse to said patient with said tumor.

wherein said genes comprise at least a plurality of ~~one or more of~~ IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb.

12. A microarray chip for analyzing the likelihood of relapse of a patient with a tumor, comprising

primers specific for amplifying RNA corresponding to at least a plurality of genes selected from the group consisting of IGKC, IGLL5, STAT1, GBP1, OCLN, MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb.

13. The microarray chip of claim 12, wherein said plurality of genes includes each of IGKC, IGLL5, STAT1, GBP1, and OCLN.

14. The microarray chip of claim 12, wherein said plurality of genes includes at least one gene selected from MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb.

15. A theranostic method for developing a treatment protocol for a cancer patient, comprising the steps of

i) obtaining a tumor tissue sample from said patient;

ii) quantifying a level of gene expression of each of IGKC, IGLL5, STAT1, GBP1, and OCLN in said tumor tissue sample;

iii) comparing a quantification value for a level of gene expression of each of IGKC, IGLL5, STAT1, GBP1, and OCLN obtained in said quantifying step with a predetermined reference value for a level of gene expression of each of IGKC, IGLL5, STAT1, GBP1, and OCLN in control tissue samples; and

iv) when said quantification value is greater than said predetermined reference value, providing a prognosis of a low likelihood of relapse for said patient and recommending a tumor-burden reducing treatment for said patient; or

v) when said quantification value is lower than said predetermined reference value, providing a prognosis of a high likelihood of relapse for said patient and recommending neoadjuvant therapy in conjunction with a tumor-burden reducing treatment for said patient.

16. The theranostic method of claim 15, wherein said tumor-burden reducing treatment includes one or more treatments selected from the group consisting of surgical removal of tumor tissue, reduction in tumor volume by chemotherapy, reduction in tumor volume by radiotherapy, and reduction in tumor volume by hormone therapy.

17. The theranostic method of claim 15, wherein said neoadjuvant therapy includes administration of one more agents selected from the group consisting of 5-azacytidine, decitabine, histone deacetylation inhibitors.

18. An in vitro method for determining, in a cancer patient in need thereof, the likelihood of relapse, comprising the steps of

i) obtaining a tumor tissue sample from said cancer patient;

ii) quantifying a level of gene expression in said tumor tissue sample each of IGKC, IGLL5, STAT1, GBP1, and OCLN;

iii) comparing a quantification value for said level of gene expression of each of

IGKC, IGLL5, STAT1, GBP1, and OCLN obtained in said quantifying step with a predetermined reference value for a level of gene expression of each of IGKC, IGLL5, STAT1, GBP1, and OCLN in control tissue samples; and

iv) providing a prognosis of a low likelihood of relapse for said patient when said quantification value is greater than said predetermined reference value; or

v) providing a prognosis of a high likelihood of relapse for said patient when said quantification value is lower than said predetermined reference value.

19. The method of claim 18, wherein said cancer is breast cancer.

20. The method of claim 18, wherein said quantifying step includes quantifying one or more housekeeping genes.

21. The method of claim 18, wherein said control tissue samples include tissue samples from one or more of subjects without cancer, subject with stage I cancer, subjects with stage II cancer, subjects with stage III cancer, subjects with stage IV cancer, subject who have not relapsed after receiving conventional cancer treatment, and subjects who have relapsed after receiving conventional cancer treatment.

22. The method of claim 18, further comprising the steps of

quantifying a level of gene expression in said tumor tissue sample at least one gene of the following: MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb;

comparing quantification values for said level of gene expression of at least one gene of the following: MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXbobtained in said quantifying step with predetermined levels of gene expression for at least one gene of the following: MAGE-a3, MAGE-a4, MAGE-a5, MAGE-a6, AKAP4, MAGE-C1, NY-ESO-1 and SPANXb in control tissue samples, and

wherein said steps of providing a prognosis of a low likelihood of relapse or providing a prognosis of a high likelihood or relapse considers quantification levels of gene expression for said gene set being respectively greater or less than said predetermined levels.

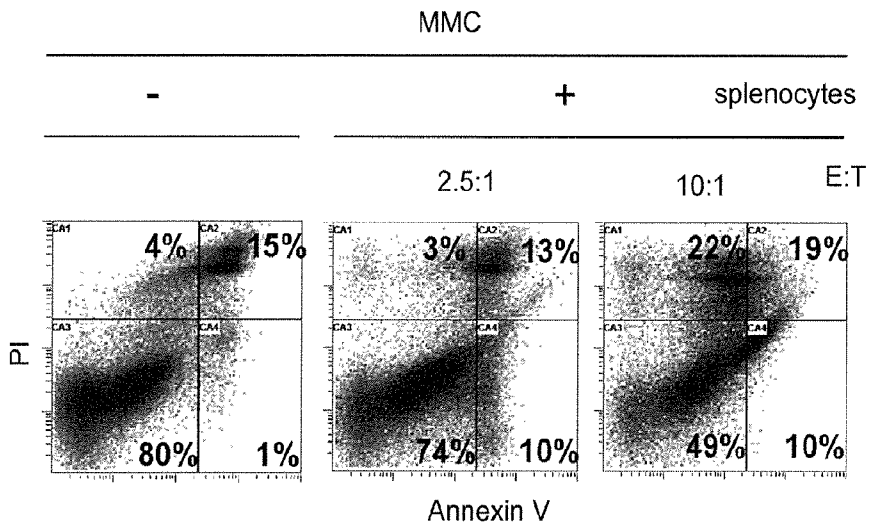


Figure 1A

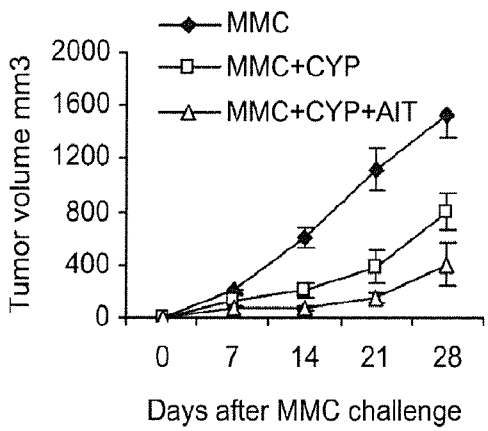


Figure 1B

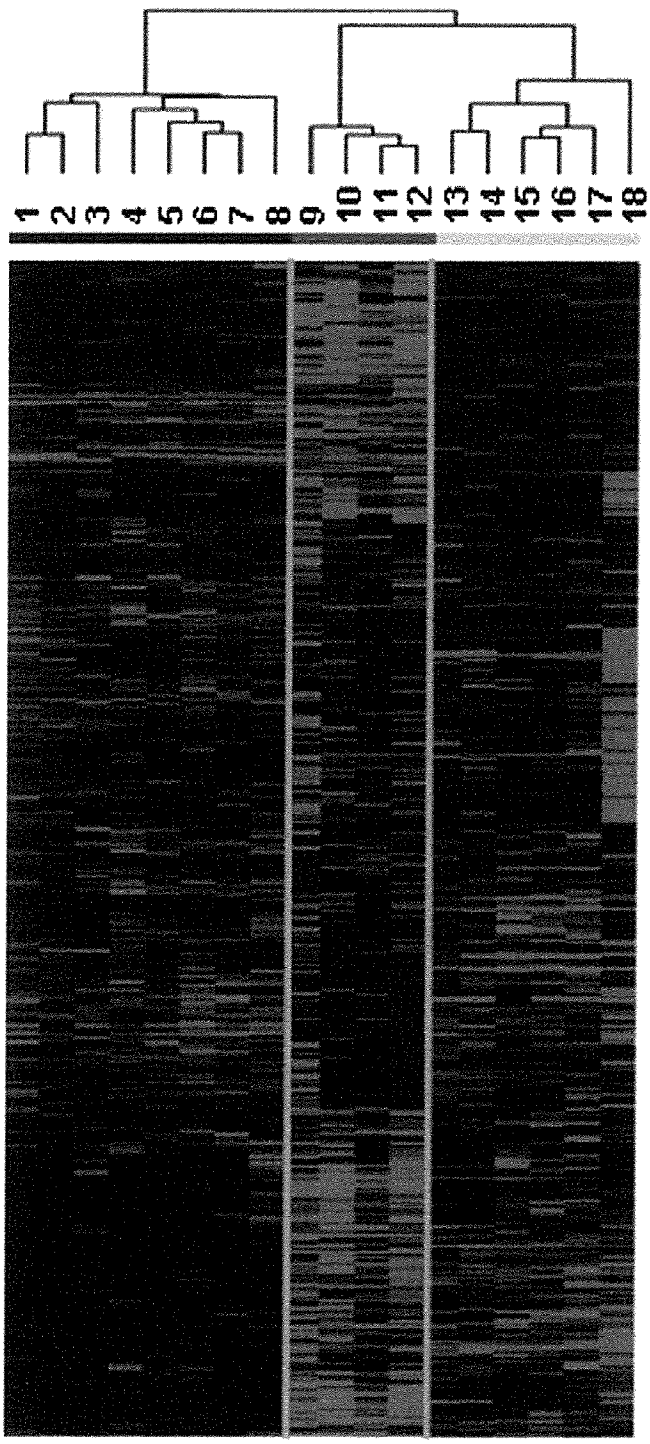


Figure 2A

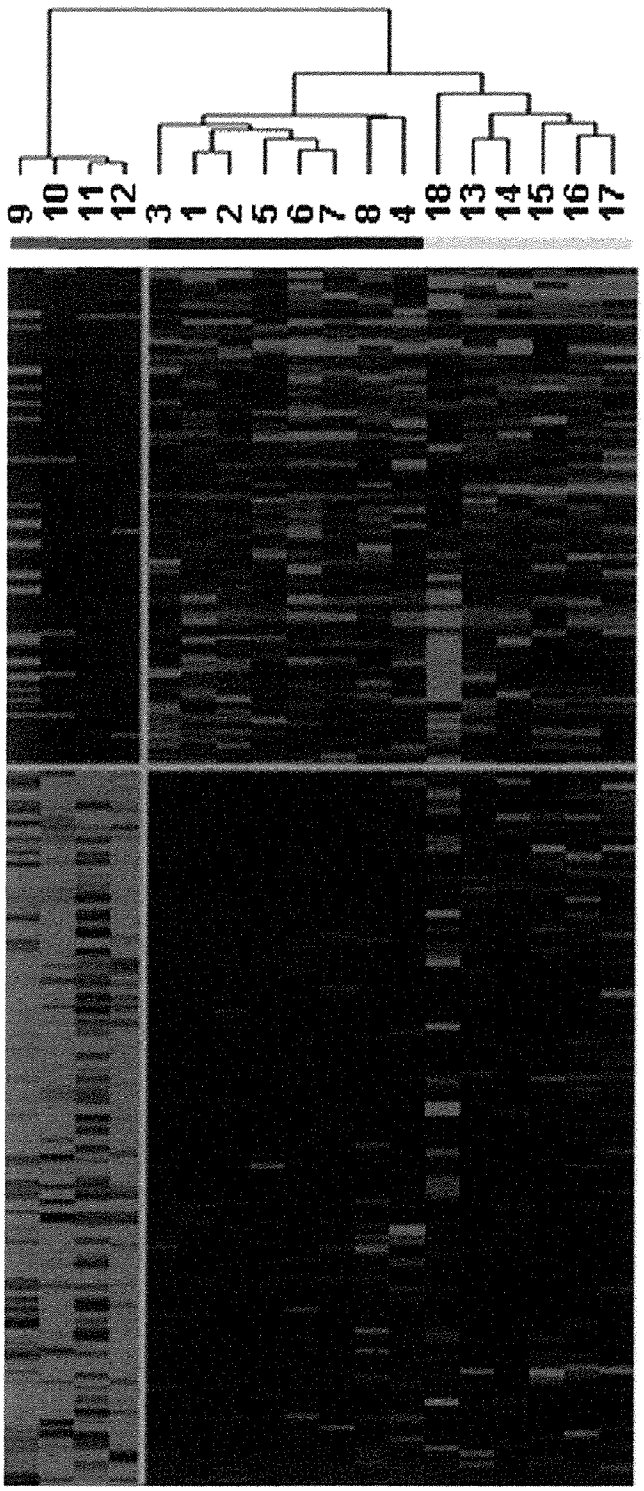


Figure 2B

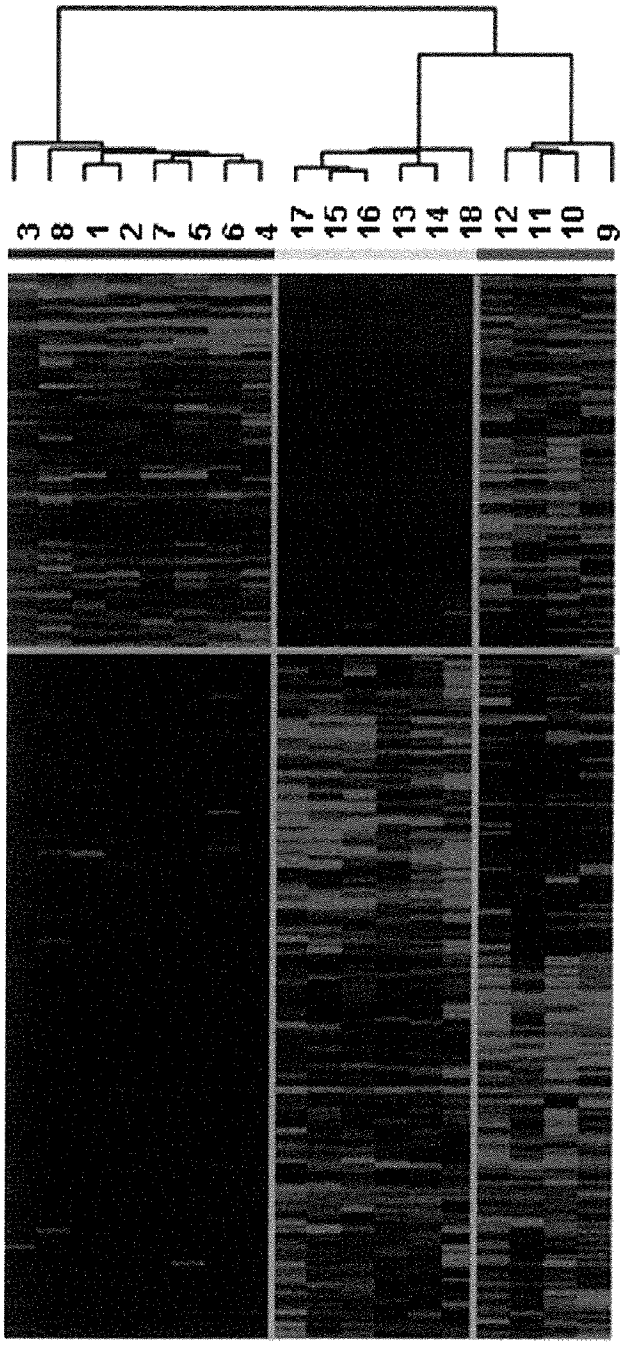


Figure 3A

Table 1

	CXC Chemokines and receptors	Mouse	Human		rejection	controls
Cxcl2	chemokine (C-X-C motif) ligand 2	GRO; MIP-2;KC?	GROξ;MGSA-ξ	CXCR2	14.54	0.44
Cxcl1	chemokine (C-X-C motif) ligand 1	GRO; MIP-2;KC?	GROa;MGSA-ξ	CXCR2>1	5.40	0.70
Cxcl11	chemokine (C-X-C motif) ligand 11	I-TAC	I-TAC	CXCR3	3.93	0.68
	CC Chemokines and receptors					
Ccl1	chemokine (C-C motif) ligand 1	TCA-2;P500	I-309	CCR8	2.15	0.80
Ccl4	chemokine (C-C motif) ligand 4	MIP-1ξ	MIP-1ξ	CCR5	5.06	0.63
Ccl5	chemokine (C-C motif) ligand 5	RANTES	RANTES	CCR1;3,5	5.42	0.62
Ccl5	chemokine (C-C motif) ligand 5	RANTES	RANTES	CCR1;3,5	3.96	0.67
Ccl6	chemokine (C-C motif) ligand 6	C10;MRP-1	Unknown	Unknown	3.79	0.68
Ccl8	chemokine (C-C motif) ligand 8	MCP-2	MCP-2	CCR3;5	5.94	0.60
Ccl9	chemokine (C-C motif) ligand 9	MRP-2;CCF18;MIP-1ξ	Unknown	CCR1	6.72	0.58
Ccl11	small chemokine (C-C motif) ligand 11	Eotaxin	Eotaxin	CCR3	4.34	0.64
Ccl22	chemokine (C-C motif) ligand 22	ABCD-1	MDC/STCP-1	CCR4	4.08	0.74
Ccrl2	chemokine (C-C motif) receptor-like 2			CCL2,7,12,13,16	6.58	0.62
Ccr10	chemokine (C-C motif) receptor 10			CCL27,28	2.22	0.83
Cklf	chemokine-like factor				2.86	0.74
Darc	Duffy blood group, chemokine receptor				2.83	0.86
Cklf	chemokine-like factor, transcript variant 1				2.82	0.74

Figure 4A-1

Table 1 - Continued

	Interferon Stimulated Genes (ISGs)	Mouse	Human	rejection	controls
Ifna2	interferon alpha 2			2.91	0.80
Ifng	interferon gamma			2.89	0.72
Ifi202b	interferon activated gene 202B			6.05	0.60
Ifi27	interferon, alpha-inducible protein 27			4.68	0.64
Ifi204	interferon activated gene 204			4.14	0.67
Ifitm1	interferon induced transmembrane protein 1			3.73	0.75
Irf6	Interferon regulatory factor 6			3.59	0.76
Ifit1	interferon-induced protein with tetratricopeptide repeats 1			3.33	0.77
Irf4	interferon regulatory factor 4			2.97	0.73
Mx1	myxovirus (influenza virus) resistance 1			2.63	0.76
Stat2	signal transducer and activator of transcription 2			2.35	0.78
Stat6	signal transducer and activator of transcription 6			2.14	0.80
Irf2bp1	nterferon regulatory factor 2 binding protein 1			0.29	1.42

Figure 4A-2

Table 2

Genes up-regulated in the rejection model

Symbol	Description	reject	contr.
	Interleukins and receptors		
Il1a	interleukin 1 alpha	2.71	0.81
Il1b	interleukin 1 beta	8.91	0.54
Il1f9	interleukin 1 family, member 9	1.97	0.82
Il5	interleukin 5	1.64	0.87
Il7	interleukin 7	3.11	0.84
Il17f	interleukin 17F	3.86	0.68
Il31	interleukin 31	2.16	0.80
		1.00	1.00
Il1rap	interleukin 1 receptor accessory protein, transcript variant 2	3.56	0.70
Il2rg	interleukin 2 receptor, gamma chain	1.75	0.85
Il7r	interleukin 7 receptor	2.46	0.88
Il23r	interleukin 23 receptor	7.38	0.63
Il17rb	interleukin 17 receptor B	4.46	0.65
		1.00	1.00
	Cytotoxic and pro-apoptotic molecules	1.00	1.00
Gzmb	granzyme B	1.90	0.83
Gzmb	granzyme B	1.57	0.88
Ctla2a	cytotoxic T lymphocyte-associated protein 2 alpha	3.60	0.69
Klra9	killer cell lectin-like receptor subfamily A, member 9	1.95	0.87
Klrd1	killer cell lectin-like receptor, subfamily D, member 1	7.36	0.57
Fasl	Fas ligand (TNF superfamily, member 6)	2.78	0.75
Tnfsf11	tumor necrosis factor (ligand) superfamily, member 11	2.56	0.80
Tnfrsf1b	tumor necrosis factor receptor superfamily, member 1b	2.21	0.83
Tnfrsf4	tumor necrosis factor receptor superfamily, member 4	2.77	0.73

Genes down-regulated in the rejection model

Symbol	Description	reject	contr.
Tnfaip1	tumor necrosis factor, alpha-induced protein 1	0.51	1.21
Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	0.14	1.52
Ngfrap1	nerve growth factor receptor (TNFRSF16) associated protein 1	0.38	1.15

Figure 4B-1

Table 2 - Continued

Genes up-regulated in the rejection model				Genes down-regulated in the rejection model			
Symbol	Description	reject	contr.	Symbol	Description	reject	contr.
	Toll-like receptors and lymphocyte signaling						
Tlr4	toll-like receptor 4	2.14	0.80				
Tlr6	toll-like receptor 6	2.98	0.86				
Il4i1	interleukin 4 induced 1	3.34	0.77				
Alcam	activated leukocyte cell adhesion molecule	2.12	0.81				
Bcl2a1c	B-cell leukemia/lymphoma 2 related protein A1c	3.41	0.70				
Itk	IL2-inducible T-cell kinase	2.63	0.76	Ilf3	interleukin enhancer binding factor 3, transcript variant 2	0.44	1.27
Ebf4	early B-cell factor 4	6.32	0.75				
Ly6a	lymphocyte antigen 6 complex, locus A	2.83	0.74				
Ly6c	lymphocyte antigen 6 complex, locus C	3.18	0.72	Ly6e	Lymphocyte antigen 6 complex, locus E	0.22	1.53
Ly6f	Lymphocyte antigen 6 complex, locus F	3.62	0.67	Ly6e	lymphocyte antigen 6 complex, locus E	0.47	1.18
Ly6f	lymphocyte antigen 6 complex, locus F	2.41	0.78	Btla	B and T lymphocyte associated, transcript variant 2	0.38	1.32
Lck	lymphocyte protein tyrosine kinase	2.26	0.79	Bcap31	B-cell receptor-associated protein 31	0.34	1.36
Tagap	T-cell activation Rho GTPase-activating protein	2.36	0.78	Tiam2	T-cell lymphoma invasion and metastasis 2	0.47	1.24
Tlx1	T-cell leukemia, homeobox 1	7.36	0.65	Ikbkap	inhibitor of kappa light polypeptide enhancer in B-cells	0.44	1.27
Tcl1b1	T-cell leukemia/lymphoma 1B, 1	2.37	0.78				
Nkrf	NF-kappaB repressing factor	3.90	0.75				
Nkiras2	NFKB inhibitor interacting Ras-like protein 2	2.48	0.82	Irak1bp1	interleukin-1 receptor-associated kinase 1 binding protein 1	0.33	1.26
Nfkbiz	Nfkb light polypeptide gene enhancer in B-cells inhibitor, zeta	3.83	0.68	Irak1	interleukin-1 receptor-associated kinase 1	0.31	1.40
Nfat5	nuclear factor of activated T-cells 5, transcript variant b	3.34	0.71				

Figure 4B-2

Table 2 - Continued

Genes up-regulated in the rejection model				Genes down-regulated in the rejection model			
Symbol	Description	reject	contr.	Symbol	Description	reject	contr.
	FC-type receptors						
Lilrb4	Leukocyte immunoglobulin-like receptor, subfamily B, member 4	3.60	0.74				
Mgl1	macrophage galactose N-acetyl-galactosamine specific lectin 1	6.23	0.59				
Mgl2	macrophage galactose N-acetyl-galactosamine specific lectin 2	2.25	0.79				
Msr1	macrophage scavenger receptor 1	2.78	0.80				
	Immunoglobulins						
Igh-6	Immunoglobulin heavy chain 6	7.43	0.56				
Igh-6	Immunoglobulin heavy chain 6 (heavy chain of IgM)	2.39	0.78				
Igj	immunoglobulin joining chain	2.86	0.74				
Igk-V28	Immunoglobulin kappa chain variable 28	2.82	0.79				
Igl-V1	Immunoglobulin lambda chain, variable 1	3.69	0.76				
Igl-V1	Immunoglobulin lambda chain, variable 1	2.03	0.82				
Igkv4-90	Immunoglobulin light chain variable region	1.82	0.84				

Figure 4B-3

Table 3

Symbol	Description	evasion	tolerogenic	rejection
	Chemokines			
Ccl2	chemokine (C-C motif) ligand 2	2.304	0.614	0.392
Ccl4	chemokine (C-C motif) ligand 4	1.899	0.478	0.838
Ccl6	chemokine (C-C motif) ligand 6	2.525	0.536	0.400
Ccr7	chemokine (C-C motif) receptor 7	1.441	0.696	0.829
Cxcl10	chemokine (C-X-C motif) ligand 10	2.341	0.781	0.264
Cxcl9	chemokine (C-X-C motif) ligand 9	1.440	0.354	2.288
Xcl1	chemokine (C motif) ligand 1	4.913	0.279	0.282
Cx3cl1	chemokine (C-X3-C motif) ligand 1	5.120	0.393	0.155
	Interleukins and signaling			
Il12b	interleukin 12b	1.768	0.866	0.397
Il13	interleukin 13	1.475	0.779	0.669
Il17d	interleukin 17D	1.690	0.674	0.633
Il23r	interleukin 23 receptor	2.048	0.457	0.772
Il2rg	interleukin 2 receptor, gamma chain	1.798	0.742	0.484
Il4	interleukin 4	2.342	0.497	0.521
Il4i1	interleukin 4 induced 1	2.290	0.701	0.325
Il6	interleukin 6	1.105	0.504	2.291
Il7r	interleukin 7 receptor	1.218	0.439	3.063
Il9	interleukin 9	1.711	0.801	0.477
Tlr11	toll-like receptor 11	2.656	0.435	0.540
Blnk	B-cell linker	1.457	0.749	0.726
Bok	Bcl-2-related ovarian killer protein	2.071	0.432	0.822
Vpreb3	pre-B lymphocyte gene 3	2.712	0.148	2.398
Lcp2	lymphocyte cytosolic protein 2	1.640	0.588	0.824
Ly6d	lymphocyte antigen 6 complex, locus D	4.093	0.223	0.567
Nfkb1	nfkb light chain gene enhancer 1, p105	1.737	0.785	0.477
Pias2	Protein inhibitor of activated STAT 2	1.336	0.502	1.458
Pias3	protein inhibitor of activated STAT 3	1.763	0.828	0.427

Figure 4C-1

Table 3 - Continued

Symbol	Description	evasion	tolerogenic	rejection
	Chemokines			
Stat4	signal transducer and activator of transcription 4	1.579	0.685	0.630
Il10	interleukin 10	0.788	2.528	0.400
Il1r2	interleukin 1 receptor, type II	0.804	2.504	0.391
Il10rb	interleukin 10 receptor, beta	0.422	2.839	1.174
Socs1	suppressor of cytokine signaling 1	0.566	4.683	0.308
Socs3	suppressor of cytokine signaling 3	0.621	1.751	1.120
Bak1	BCL2-antagonist/killer 1	0.663	2.698	0.514
Lsp1	lymphocyte specific 1	0.609	1.468	1.514
Tank	TRAF family member-associated Nf-kappa B activator	0.718	2.204	0.351
Tlr6	toll-like receptor 6	0.480	1.141	3.559
	ISGs			
Ifnb1	interferon beta 1, fibroblast	1.388	0.644	1.003
Irf7	interferon regulatory factor 7	1.500	0.677	0.406
Ifrd1	interferon-related developmental regulator 1	2.248	0.337	1.011
Ifnar1	interferon (alpha and beta) receptor 1	1.884	0.855	0.356
Igtp	interferon gamma induced GTPase	1.863	0.671	0.524
Irf1	interferon regulatory factor 1	0.549	2.715	0.671
Irf3	interferon regulatory factor 3	0.845	1.759	0.479
Irf6	interferon regulatory factor 6	0.601	1.816	1.282
Ifngr1	interferon gamma receptor 1	0.582	4.515	0.308
Ifrg15	interferon alpha responsive gene	0.591	1.913	1.168

Figure 4C-2

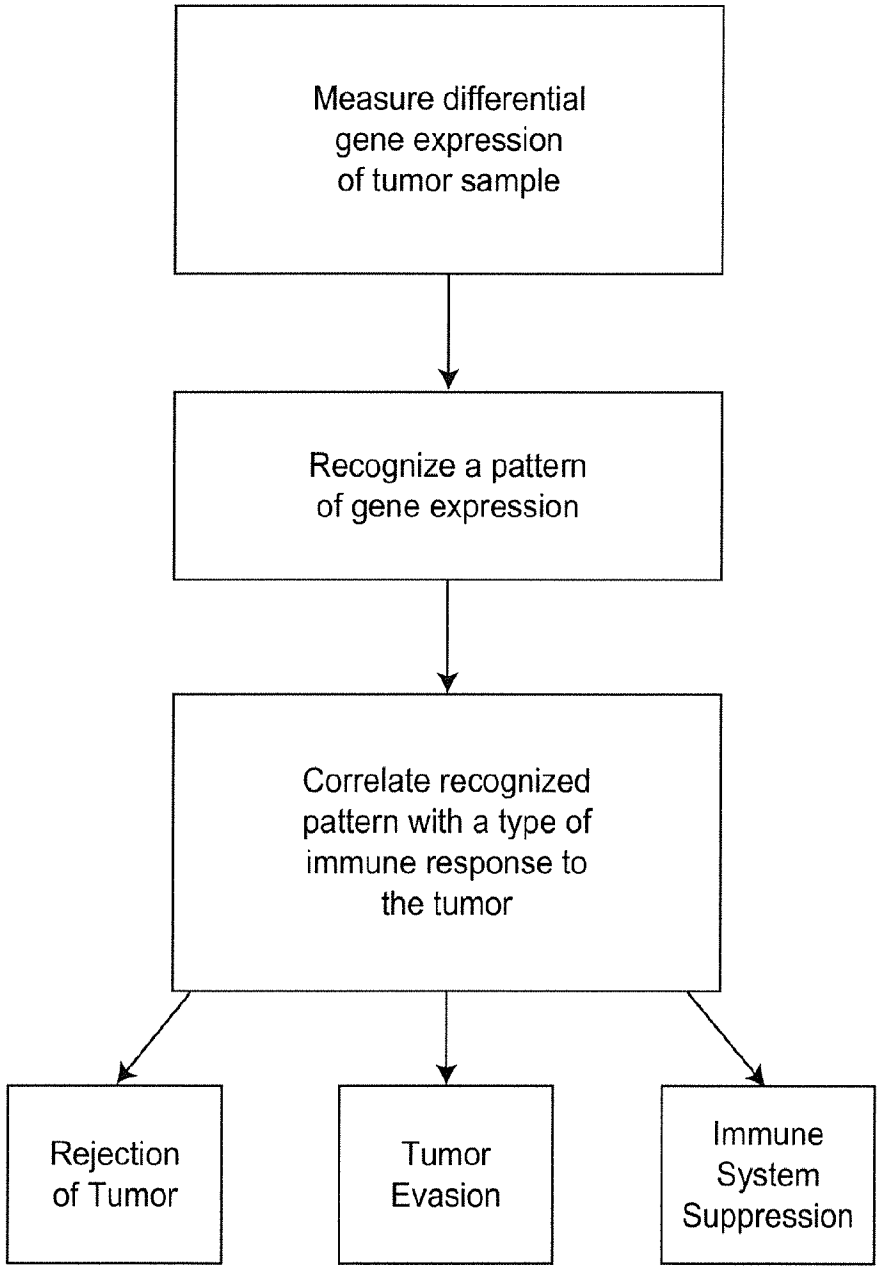

Measure differential gene expression of tumor sample
Recognize a pattern of gene expression
Correlate recognized pattern with a type of immune response to the tumor
Rejection of Tumor
Tumor Evasion
Immune System Suppression

Figure 5A

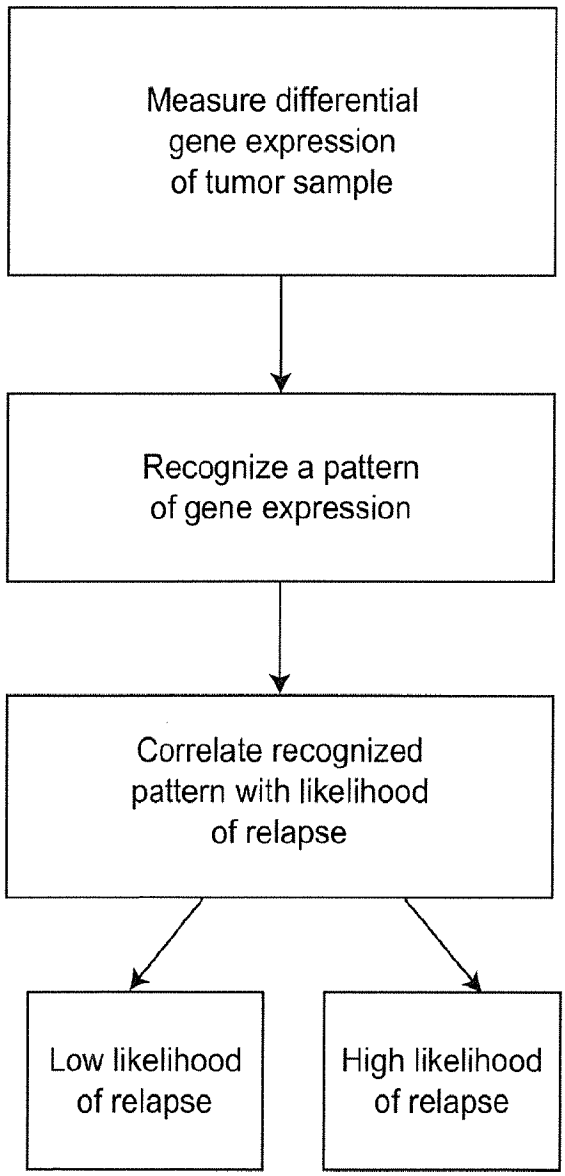

Measure differential gene expression of tumor sample
Recognize a pattern of gene expression
Correlate recognized pattern with likelihood of relapse
Low likelihood of relapse
High likelihood of relapse

Figure 5B

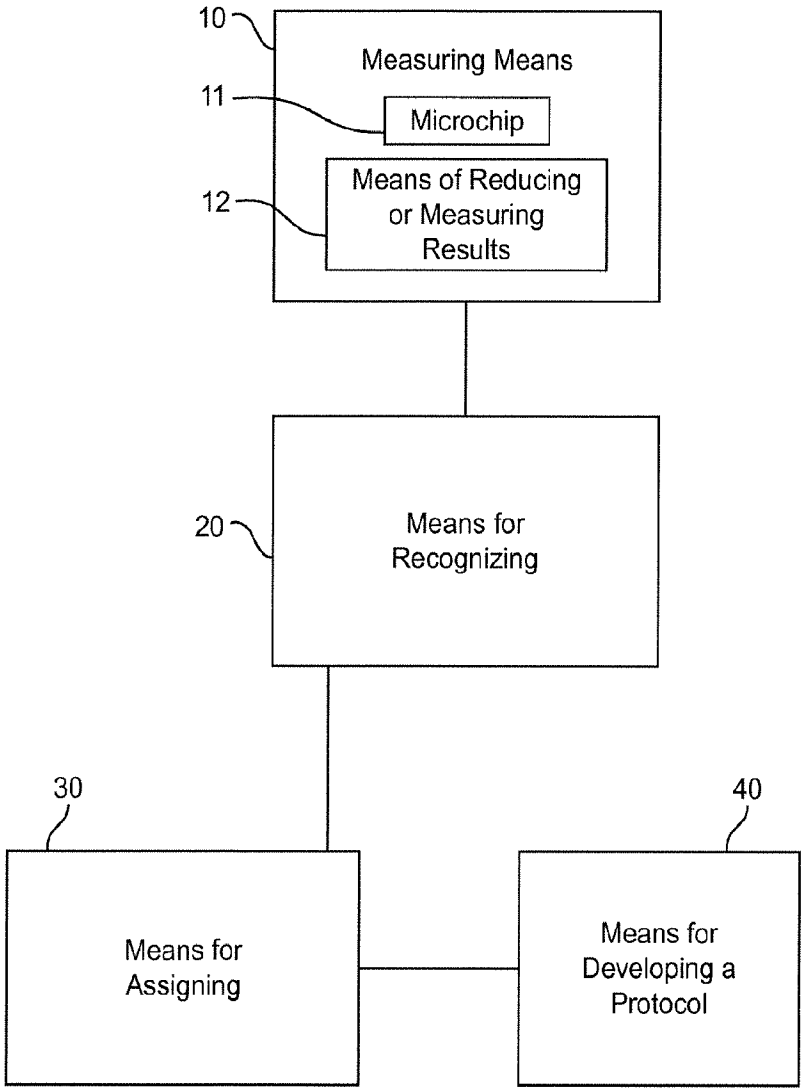

10
Measuring Means
11
Microchip
12
Means of Reducing or Measuring Results
20
Means for Recognizing
30
Means for Assigning
40
Means for Developing a Protocol

Figure 6

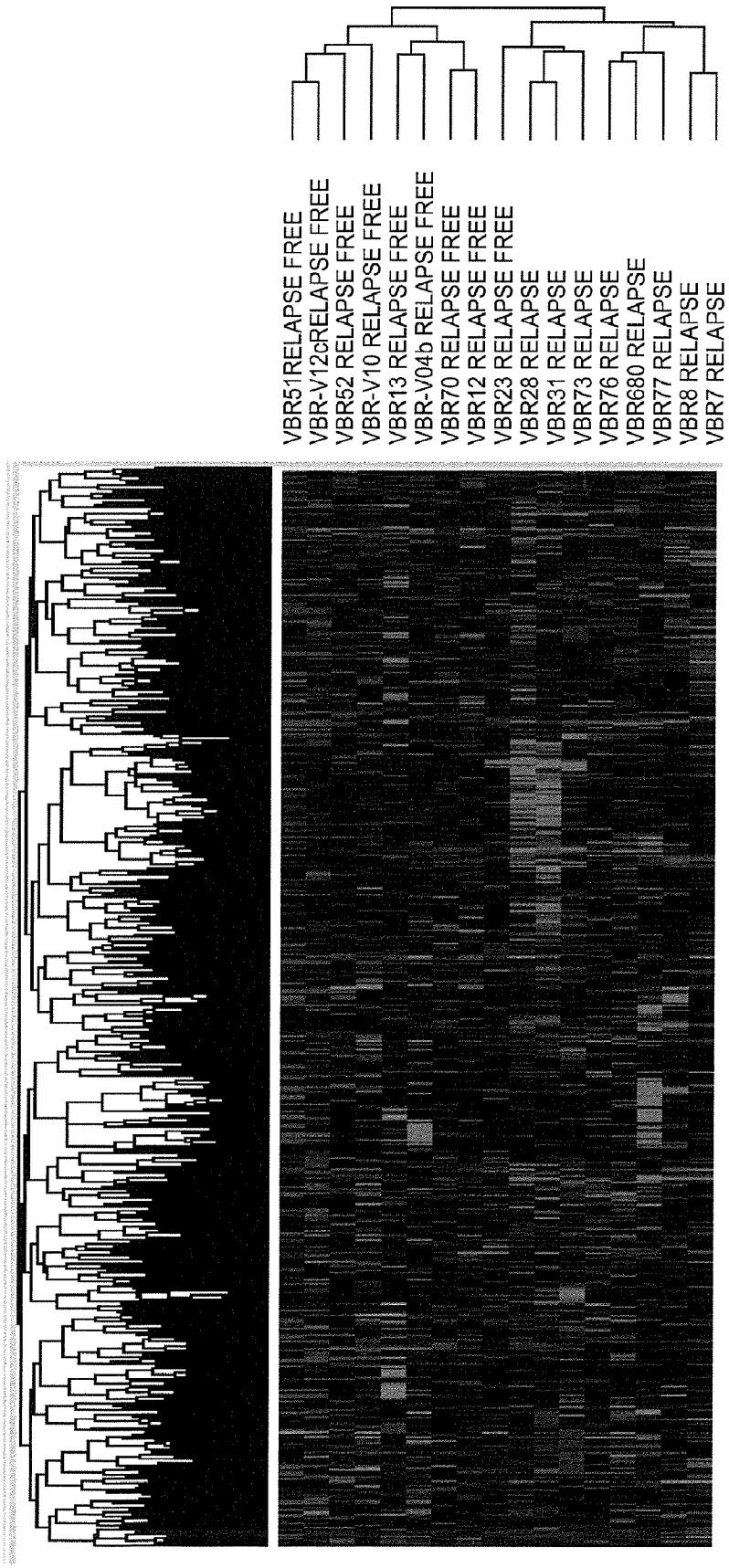

VBR51RELAPSE FREE
VBR-V12cRELAPSE FREE
VBR52 RELAPSE FREE
VBR-V10 RELAPSE FREE
VBR13 RELAPSE FREE
VBR-V04b RELAPSE FREE
VBR70 RELAPSE FREE
VBR12 RELAPSE FREE
VBR23 RELAPSE FREE
VBR28 RELAPSE
VBR31 RELAPSE
VBR73 RELAPSE
VBR76 RELAPSE
VBR680 RELAPSE
VBR77 RELAPSE
VBR8 RELAPSE
VBR7 RELAPSE

Figure 7

Figure 8

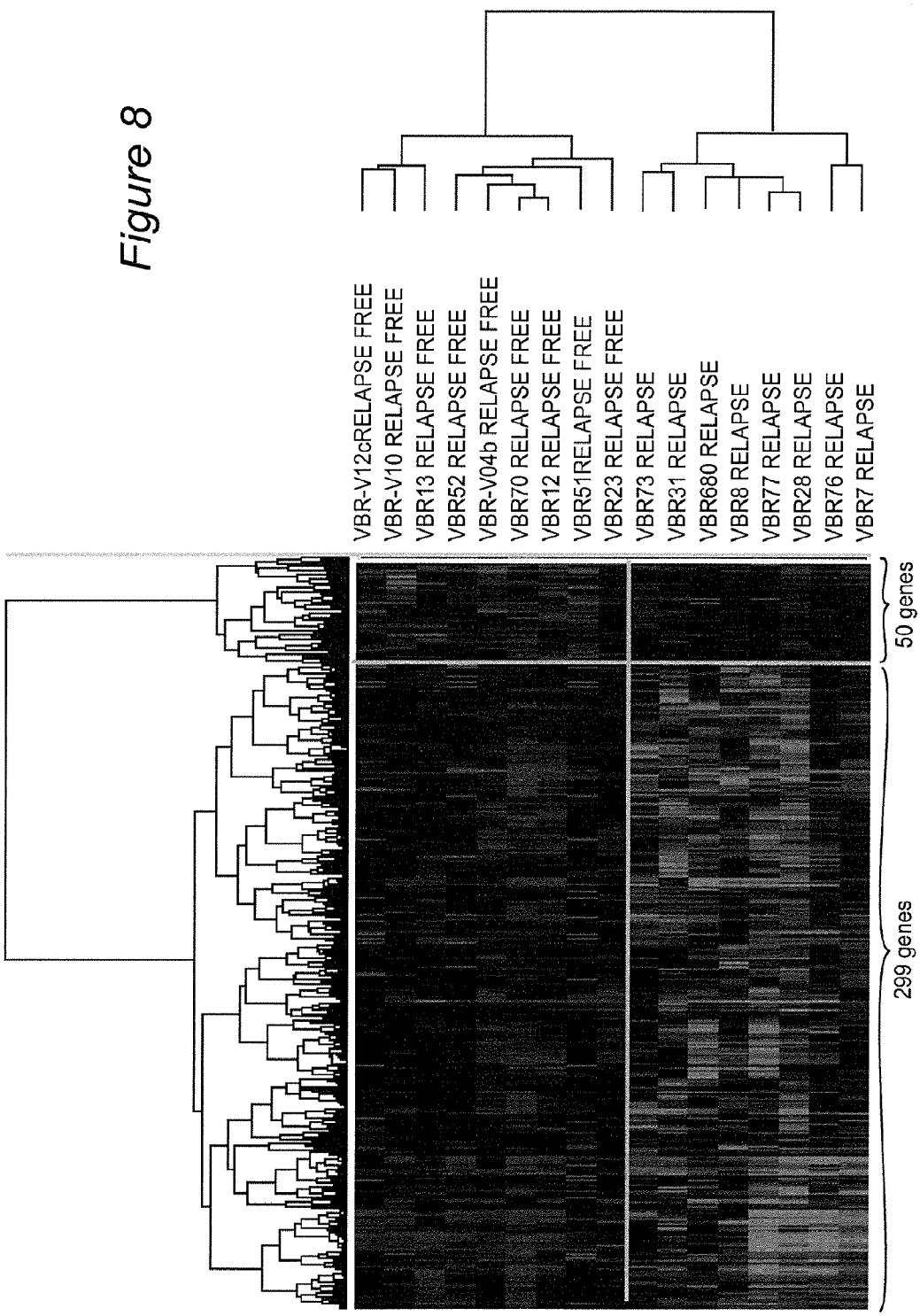

VBR-V12cRELAPSE FREE
VBR-V10 RELAPSE FREE
VBR13 RELAPSE FREE
VBR52 RELAPSE FREE
VBR-V04b RELAPSE FREE
VBR70 RELAPSE FREE
VBR12 RELAPSE FREE
VBR51RELAPSE FREE
VBR23 RELAPSE FREE
VBR73 RELAPSE
VBR31 RELAPSE
VBR680 RELAPSE
VBR8 RELAPSE
VBR77 RELAPSE
VBR28 RELAPSE
VBR76 RELAPSE
VBR7 RELAPSE
50 genes
299 genes

Table 4

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
12.6047915	0.8130229	15.5036128	IGKC--Immunoglobulin kappa constant
7.0918936	0.5762246	12.3075153	IGL@--Immunoglobulin lambda locus
18.9015207	1.6860658	11.2104292	IGKC--Ig kappa chain C region. [Source:Uniprot/SWISSPROT;Acc:P01834]
19.343305	1.9502818	9.9182102	LOC440871--PREDICTED: similar to hCG2043206 (LOC440871), mRNA.
13.1368088	1.4265401	9.2088606	--Unknown
13.5919842	1.5245249	8.9155542	IGKV3-20--Immunoglobulin kappa variable 3-20
25.2647688	2.866613	8.8134565	LOC440871--PREDICTED: similar to hCG2043206 (LOC440871), mRNA.
7.414271	0.8519614	8.7025907	--Unknown
16.8056924	1.9389905	8.667238	LOC652493--PREDICTED: similar to pre-B lymphocyte gene 1 (LOC652493), mRNA.
28.6339462	3.3627215	8.5151109	CXCL9--chemokine (C-X-C motif) ligand 9 (CXCL9), mRNA.
11.2005112	1.3280892	8.4335535	--Clone ds1-1 immunoglobulin lambda chain VJ region, (IGL)
27.355858	3.3787658	8.0964054	LIPT1--lipoyltransferase 1 (LIPT1), nuclear gene encoding mitochondrial protein, transcript variant 5, mRNA.
0.9218951	0.1171789	7.8674181	MS4A1--membrane-spanning 4-domains, subfamily A, member 1 (MS4A1), transcript variant 3, mRNA.
8.1274472	1.0445817	7.7805756	LOC96610--BMS1 homolog, ribosome assembly protein (yeast) pseudogene (LOC96610), non-coding RNA.
4.5321044	0.6051429	7.4893128	LOC100290481--PREDICTED: similar to immunoglobulin lambda locus (LOC100290481), mRNA.
1.5310405	0.2047893	7.4761727	--Clone 167-7 anti-HIV-1 gp41 immunoglobulin heavy chain
9.4480336	1.2902836	7.3224473	LOC100132941--PREDICTED: similar to Ig heavy chain (LOC100132941), mRNA.
32.7051356	4.6356912	7.055072	IGKC--Immunoglobulin kappa constant
11.5961501	1.6552401	7.0057209	IGLC1--Ig lambda chain C regions. [Source:Uniprot/SWISSPROT;Acc:P01842]
10.1273856	1.4759433	6.8616357	--Isolate RSV22LB immunoglobulin light chain variable region
4.6967966	0.7154825	6.5645167	KIAA1632--KIAA1632
3.6785377	0.5604657	6.563359	CCL23--chemokine (C-C motif) ligand 23 (CCL23), transcript variant CKbeta8, mRNA.
8.1568032	1.3495946	6.0438913	--Unknown
			IGLC1--Ig lambda chain C regions.

Figure 9A

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
11.6708878	1.9516815	5.9799141	[Source:Uniprot/SWISSPROT;Acc:P01842]
3.712414	0.6241482	5.9479687	POU2AF1--POU class 2 associating factor 1 (POU2AF1), mRNA.
2.1368458	0.3897602	5.4824622	STAT1--signal transducer and activator of transcription 1, 91kDa (STAT1), transcript variant beta, mRNA.
23.681266	4.5287701	5.2290723	--EmptyWell
1.4696916	0.2829853	5.193526	GBP4--guanylate binding protein 4 (GBP4), mRNA.
5.4857212	1.0748638	5.1036432	GBP1--guanylate binding protein 1, interferon-inducible, 67kDa (GBP1), mRNA.
0.2064003	0.0405223	5.0934983	IL7R--interleukin 7 receptor (IL7R), mRNA.
14.9694217	2.9449662	5.0830539	--Clone 996-K1-9 immunoglobulin kappa light chain variable region mRNA, partial sequence
6.0436156	1.1890092	5.0829006	LOC642838--PREDICTED: similar to hCG1742442 (LOC642838), mRNA.
0.2016847	0.0405944	4.9682874	LYZ--lysozyme (renal amyloidosis) (LYZ), mRNA.
22.9637998	4.6932157	4.8929777	CXCL10--chemokine (C-X-C motif) ligand 10 (CXCL10), mRNA.
19.3383998	4.0434917	4.7825991	UBD--ubiquitin D (UBD), mRNA.
0.5611412	0.1178536	4.7613422	GBP5--guanylate binding protein 5 (GBP5), transcript variant 2, mRNA.
6.6191849	1.4369997	4.6062536	IGKC--Immunoglobulin kappa constant
3.9328938	0.8607381	4.5692109	--Unknown
0.3423117	0.0749785	4.5654653	PTPRC--protein tyrosine phosphatase, receptor type, C (PTPRC), transcript variant 2, mRNA.
6.2008817	1.3703002	4.5251996	LOC100293559--PREDICTED: similar to hCG1812074 (LOC100293559), miscRNA.
4.6588373	1.0351145	4.5007939	IGKV3-20--Immunoglobulin kappa variable 3-20
8.4176078	1.8899696	4.4538325	IGLC1--Ig lambda chain C regions. [Source:Uniprot/SWISSPROT;Acc:P01842]
6.4692631	1.4576197	4.438238	LOC100289290--PREDICTED: similar to hCG2042717 (LOC100289290), mRNA.
6.3853994	1.4529483	4.3947877	LOC100287372--PREDICTED: similar to hCG2029236 (LOC100287372), mRNA.
5.1789822	1.1891382	4.35524	OCLN--occludin (OCLN), mRNA.
0.2920193	0.07121	4.1008167	IL23A--Interleukin 23, alpha subunit p19
0.6931297	0.170253	4.0711754	GZMA--granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3) (GZMA), mRNA.

Figure 9B

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
5.1465171	1.2661796	4.0646028	GBP3--guanylate binding protein 3 (GBP3), mRNA.
1.4911578	0.3711934	4.0171994	CXCL11--chemokine (C-X-C motif) ligand 11 (CXCL11), mRNA.
4.208898	1.0520741	4.0005717	BMS1P4--BMS1 pseudogene 4
0.257644	0.064621	3.987	GPR171--G protein-coupled receptor 171 (GPR171), mRNA.
4.8504688	1.2450823	3.8957012	--Transcribed locus, moderately similar to XP_001085196.1 PREDICTED: guanylate binding protein 1, interferon-inducible, 67kD isoform 2 [Macaca mulatta]
0.2538812	0.0654211	3.8807208	TRD@--T cell receptor delta locus
0.9103647	0.235969	3.8579847	PSMB9--proteasome (prosome, macropain) subunit, beta type, 9 (large multifunctional peptidase 2) (PSMB9), transcript variant 2, mRNA.
5.0381404	1.3243602	3.804207	LOC642424--PREDICTED: similar to hCG1742442 (LOC642424), mRNA.
6.3057394	1.6581539	3.8028674	LOC100130100--PREDICTED: similar to hCG26659 (LOC100130100), mRNA.
0.2032809	0.0543498	3.7402305	IL7R--Interleukin 7 receptor
6.7281597	1.8021737	3.733358	IGF2--Insulin-like growth factor 2 (somatomedin A)
2.0384404	0.5475036	3.7231546	MAFG--V-maf musculoaponeurotic fibrosarcoma oncogene homolog G (avian)
0.1599341	0.043832	3.6488014	CD48--CD48 molecule (CD48), mRNA.
0.7054506	0.1936223	3.6434362	CCR2--chemokine (C-C motif) receptor 2 (CCR2), transcript variant A, mRNA.
0.705033	0.1942112	3.630239	ISG20--interferon stimulated exonuclease gene 20kDa (ISG20), mRNA.
1.0428554	0.2890232	3.6082061	GIMAP5--GTPase, IMAP family member 5 (GIMAP5), mRNA.
3.5563149	0.9896536	3.5934946	CD74--CD74 molecule, major histocompatibility complex, class II invariant chain (CD74), transcript variant 1, mRNA.
0.213742	0.0597011	3.5802004	LCK--lymphocyte-specific protein tyrosine kinase (LCK), transcript variant 1, mRNA.
0.3968556	0.1114939	3.5594384	LTB--lymphotoxin beta (TNF superfamily, member 3) (LTB), transcript variant 2, mRNA.
0.4188335	0.1195464	3.5035225	GIMAP6--GTPase, IMAP family member 6 (GIMAP6), transcript variant 1, mRNA.
2.3993517	0.6920893	3.4668238	IRF4--interferon regulatory factor 4 (IRF4), mRNA.
0.258028	0.0759855	3.3957544	GNLY--granulysin (GNLY), transcript variant NKG5, mRNA.
			IFI44L--interferon-induced protein 44-like (IFI44L),

Figure 9C

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
5.7702966	1.7018262	3.3906498	mRNA.
7.3486883	2.1770817	3.3754766	MMP10--matrix metallopeptidase 10 (stromelysin 2) (MMP10), mRNA.
0.2015088	0.0597884	3.3703681	CD247--CD247 molecule (CD247), transcript variant 2, mRNA.
0.2046517	0.0614415	3.3308407	LGALS2--lectin, galactoside-binding, soluble, 2 (LGALS2), mRNA.
2.2947826	0.6930262	3.3112492	--Isolate RSV108LB immunoglobulin light chain variable region
2.8098096	0.8580681	3.2745765	ZNF341--zinc finger protein 341 (ZNF341), mRNA.
0.3488308	0.106728	3.2684107	GIMAP7--GTPase, IMAP family member 7 (GIMAP7), mRNA.
1.6031656	0.4968678	3.2265438	B2M--beta-2-microglobulin (B2M), mRNA.
2.8773615	0.8967135	3.2087856	FGF19--fibroblast growth factor 19 (FGF19), mRNA. STAT1--signal transducer and activator of transcription 1, 91kDa (STAT1), transcript variant
4.5263174	1.4249092	3.1765654	alpha, mRNA.
0.2137026	0.0673027	3.1752462	ITGAL--integrin, alpha L (antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide) (ITGAL), transcript variant 2, mRNA.
0.1786038	0.0572166	3.1215395	CARD16--caspase recruitment domain family, member 16 (CARD16), transcript variant 1, mRNA.
0.4995361	0.160212	3.1179705	CD2--CD2 molecule (CD2), mRNA.
4.067416	1.3110366	3.1024428	ABCC6--ATP-binding cassette, sub-family C (CFTR/MRP), member 6 (ABCC6), transcript variant 1, mRNA.
0.4492907	0.146221	3.0726817	ITM2A--integral membrane protein 2A (ITM2A), transcript variant 2, mRNA.
1.5883002	0.5174895	3.0692414	IGJ--immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides (IGJ), mRNA.
2.8927904	0.948385	3.0502279	IGLV6-57--Immunoglobulin lambda variable 6-57
3.5607414	1.1687084	3.0467322	HLA-B--major histocompatibility complex, class I, B (HLA-B), mRNA.
1.2179612	0.4014484	3.0339169	B2M--beta-2-microglobulin (B2M), mRNA.
1.7266048	0.5764712	2.9951274	FLJ40330--hypothetical LOC645784 (FLJ40330), non-coding RNA.
2.466832	0.82768	2.9804176	ASB16--ankyrin repeat and SOCS box-containing 16 (ASB16), mRNA.

Figure 9D

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
0.3086322	0.1038164	2.9728674	P2RX5--purinergic receptor P2X, ligand-gated ion channel, 5 (P2RX5), transcript variant 2, mRNA.
2.272018	0.7642568	2.9728464	CTSC--cathepsin C (CTSC), transcript variant 1, mRNA.
1.2538735	0.4243354	2.9549115	FLJ40330--hypothetical LOC645784 (FLJ40330), non-coding RNA.
0.1670732	0.0566102	2.9512891	CD79A--CD79a molecule, immunoglobulin-associated alpha (CD79A), transcript variant 2, mRNA.
0.0918304	0.0311214	2.9507138	CST7--cystatin F (leukocystatin) (CST7), mRNA.
0.6711725	0.2283264	2.9395309	HLA-DPA1--major histocompatibility complex, class II, DP alpha 1 (HLA-DPA1), mRNA.
0.3483993	0.1194719	2.9161614	LCP1--lymphocyte cytosolic protein 1 (L-plastin) (LCP1), mRNA.
0.1574962	0.0540588	2.9134232	ERICH1--glutamate-rich 1 (ERICH1), mRNA.
2.1972253	0.7571821	2.9018453	HLA-DRA--Major histocompatibility complex, class II, DR alpha
1.3741981	0.4744537	2.8963798	ARL8A--ADP-ribosylation factor-like 8A (ARL8A), mRNA.
0.14027	0.0496129	2.8272906	PVRIG--poliovirus receptor related immunoglobulin domain containing (PVRIG), mRNA.
9.8385794	3.4880133	2.8206829	IFIT3--interferon-induced protein with tetratricopeptide repeats 3 (IFIT3), transcript variant 2, mRNA.
6.1516697	2.1857113	2.8144932	DCC--deleted in colorectal carcinoma (DCC), mRNA.
1.1953319	0.4257006	2.807917	HLA-DRA--major histocompatibility complex, class II, DR alpha (HLA-DRA), mRNA.
2.49586	0.888883	2.8078612	HLA-G--Major histocompatibility complex, class I, G
2.9941311	1.0706539	2.7965444	--CDNA FLJ12284 fis, clone MAMMA1001757
0.737882	0.2649327	2.785168	CD19--CD19 molecule (CD19), mRNA.
2.1337654	0.7691954	2.7740225	KDM2A--lysine (K)-specific demethylase 2A (KDM2A), transcript variant 1, mRNA.
0.1700928	0.0616218	2.7602703	TBC1D10C--TBC1 domain family, member 10C (TBC1D10C), mRNA.
0.904787	0.3281995	2.7568205	HLA-DRB5--major histocompatibility complex, class II, DR beta 5 (HLA-DRB5), mRNA.
1.1261336	0.4095097	2.7499558	IRF8--interferon regulatory factor 8 (IRF8), mRNA.
1.3956392	0.5113438	2.7293559	--Unknown
0.153668	0.0567167	2.7093932	GMFG--glia maturation factor, gamma (GMFG), mRNA.
			HLA-DQA2--major histocompatibility complex, class II,

Figure 9E

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
2.6517167	0.9788878	2.7089078	DQ alpha 2 (HLA-DQA2), mRNA.
0.4096803	0.1512458	2.7087059	GZMB--granzyme B (granzyme 2, cytotoxic T-lymphocyte-associated serine esterase 1) (GZMB), mRNA.
4.1164024	1.5261451	2.6972549	HLA-DRB1--major histocompatibility complex, class II, DR beta 1 (HLA-DRB1), mRNA.
0.3377807	0.1261538	2.6775312	PLEK--pleckstrin (PLEK), mRNA.
1.2129867	0.4537244	2.6734001	IFI30--interferon, gamma-inducible protein 30 (IFI30), mRNA.
7.5663595	2.8346123	2.669275	IL4I1--interleukin 4 induced 1 (IL4I1), transcript variant 1, mRNA.
3.1267034	1.1714032	2.6691948	PARP14--poly (ADP-ribose) polymerase family, member 14 (PARP14), mRNA.
1.4470043	0.5428081	2.6657751	OASL--2'-5'-oligoadenylate synthetase-like (OASL), transcript variant 2, mRNA.
0.3030279	0.1138043	2.6627112	IL2RB--interleukin 2 receptor, beta (IL2RB), mRNA.
2.8488036	1.0800318	2.6377035	IGKC--Immunoglobulin kappa constant
0.1850349	0.0704933	2.6248559	PLAC8--placenta-specific 8 (PLAC8), transcript variant 2, mRNA.
2.3159816	0.8823648	2.624744	MGC29506--Plasma cell-induced ER protein 1
1.6382152	0.6295552	2.6021786	HLA-E--major histocompatibility complex, class I, E (HLA-E), mRNA.
2.3314803	0.8984287	2.5950644	CCL21--chemokine (C-C motif) ligand 21 (CCL21), mRNA.
2.7208561	1.0490429	2.5936557	SLC25A11--solute carrier family 25 (mitochondrial carrier; oxoglutarate carrier), member 11 (SLC25A11), nuclear gene encoding mitochondrial protein, transcript variant 3, mRNA.
0.3622064	0.1404241	2.579375	CYFIP2--cytoplasmic FMR1 interacting protein 2 (CYFIP2), transcript variant 1, mRNA.
1.3044432	0.5068151	2.5738047	IFI16--interferon, gamma-inducible protein 16 (IFI16), mRNA.
3.3199436	1.2912869	2.5710348	--Transcribed locus
0.1389923	0.0540728	2.5704658	GMFG--Glia maturation factor, gamma
1.0372059	0.4053788	2.5586091	PACRG--PARK2 co-regulated (PACRG), transcript variant 1,mRNA.
2.1520049	0.8430503	2.5526411	HOXC6--homeobox C6 (HOXC6), transcript variant 1, mRNA.
1.3317794	0.522333	2.549675	IRF1--interferon regulatory factor 1 (IRF1), mRNA.
			HLA-L--major histocompatibility complex, class I, L,

Figure 9F

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
2.5772928	1.0138469	2.5420927	pseudogene (HLA-L), non-coding RNA.
0.1827319	0.0719113	2.5410731	NKG7--natural killer cell group 7 sequence (NKG7), mRNA.
2.217368	0.8740928	2.536765	--CDNA FLJ12169 fis, clone MAMMA1000643
0.3209957	0.1270821	2.5258925	CASP1--caspase 1, apoptosis-related cysteine peptidase (interleukin 1, beta, convertase) (CASP1), transcript variant epsilon, mRNA.
2.4479293	0.970219	2.5230689	FKBP11--FK506 binding protein 11, 19 kDa (FKBP11), transcript variant 1, mRNA.
1.4483553	0.5814015	2.4911448	FAM173A--family with sequence similarity 173, member A (FAM173A), mRNA.
0.3906636	0.1570838	2.4869748	VNN2--vanin 2 (VNN2), transcript variant 2, mRNA.
0.6887819	0.2772033	2.484754	ZDHHC2--zinc finger, DHHC-type containing 2 (ZDHHC2), mRNA.
2.0303475	0.8199518	2.4761791	HLA-H--major histocompatibility complex, class I, H (pseudogene) (HLA-H), non-coding RNA.
0.0636543	0.0258202	2.4652903	TARP--TCR gamma alternate reading frame protein (TARP), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
2.201645	0.8948794	2.46027	PSMB8--proteasome (prosome, macropain) subunit, beta type, 8 (large multifunctional peptidase 7) (PSMB8), transcript variant 2, mRNA.
3.0226661	1.2338226	2.4498384	TAP1--transporter 1, ATP-binding cassette, sub-family B (MDR/TAP) (TAP1), mRNA.
1.821444	0.7469356	2.4385557	PIM2--pim-2 oncogene (PIM2), mRNA.
0.1694583	0.0698192	2.4271008	AIF1--allograft inflammatory factor 1 (AIF1), transcript variant 1, mRNA.
1.9550915	0.8085205	2.41811	--Transcribed locus
2.2467871	0.9303111	2.4150923	HLA-F--major histocompatibility complex, class I, F (HLA-F), transcript variant 3, mRNA.
8.6633695	3.6001342	2.4064018	PLTP--phospholipid transfer protein (PLTP), transcript variant 2, mRNA.
0.4290108	0.1785868	2.4022545	FYB--FYN binding protein (FYB-120/130) (FYB), transcript variant 1, mRNA.
1.868924	0.7782029	2.4015895	HLA-DMA--major histocompatibility complex, class II, DM alpha (HLA-DMA), mRNA.
4.5591	1.8989027	2.4009129	CNN2--calponin 2 (CNN2), transcript variant 2, mRNA.
3.3112292	1.3809449	2.3977996	EPSTI1--epithelial stromal interaction 1 (breast) (EPSTI1), transcript variant 2, mRNA.

Figure 9G

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
0.3060981	0.1276936	2.3971305	--Unknown
0.2822402	0.1178964	2.3939685	TRAF3IP3--TRAF3 interacting protein 3 (TRAF3IP3), mRNA.
1.6088375	0.6725689	2.3920783	LOC440839--tigger transposable element derived 1 pseudogene (LOC440839), non-coding RNA.
0.9280237	0.3881602	2.3908263	WIPF1--WAS/WASL interacting protein family, member 1 (WIPF1), transcript variant 1, mRNA.
0.437314	0.183371	2.3848599	CYBB--cytochrome b-245, beta polypeptide (CYBB), mRNA.
0.4334499	0.1822223	2.3786878	ARHGAP25--Rho GTPase activating protein 25 (ARHGAP25), transcript variant 4, mRNA.
3.298774	1.3870047	2.3783439	IGKC--Immunoglobulin kappa constant
1.1828791	0.4973795	2.3782226	HLA-L--major histocompatibility complex, class I, L, pseudogene (HLA-L), non-coding RNA.
1.0254113	0.4320061	2.3736034	TFEC--transcription factor EC (TFEC), transcript variant 2, mRNA.
0.1929315	0.081404	2.3700491	RPS27--Ribosomal protein S27
0.2981614	0.1262	2.3626094	NCF1--neutrophil cytosolic factor 1 (NCF1), mRNA.
0.1837255	0.0780536	2.3538371	POU2F2--POU class 2 homeobox 2 (POU2F2), mRNA.
0.4363389	0.185689	2.3498377	EMB--embigin homolog (mouse) (EMB), mRNA.
0.503905	0.2154222	2.3391511	ICAM3--intercellular adhesion molecule 3 (ICAM3), mRNA.
0.5100916	0.2186014	2.3334323	GIMAP4--GTPase, IMAP family member 4 (GIMAP4), mRNA.
0.9336839	0.4003369	2.3322453	HLA-G--major histocompatibility complex, class I, G (HLA-G), mRNA.
0.7029693	0.3024095	2.3245605	--CDNA FLJ23860 fis, clone LNG08308
1.4791611	0.6409984	2.3075892	SLAMF1--signaling lymphocytic activation molecule family member 1 (SLAMF1), mRNA.
0.2327509	0.1010205	2.3039965	EVI2B--ecotropic viral integration site 2B (EVI2B), mRNA.
0.5741927	0.249965	2.2970922	SP110--SP110 nuclear body protein (SP110), transcript variant a, mRNA.
1.3331717	0.5808084	2.2953727	DNAJC5G--DnaJ (Hsp40) homolog, subfamily C, member 5 gamma (DNAJC5G), mRNA.
3.3945798	1.4798477	2.293871	NYNRIN--NYN domain and retroviral integrase containing (NYNRIN), mRNA.
			MBP--myelin basic protein (MBP), transcript variant 8,

Figure 9H

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
1.0216569	0.4473436	2.2838302	mRNA.
1.4195403	0.6230992	2.2781932	SIT1--signaling threshold regulating transmembrane adaptor 1 (SIT1), mRNA.
3.5629555	1.5642298	2.2777699	ROBO4--Roundabout homolog 4, magic roundabout (Drosophila)
0.7714668	0.3389666	2.2759377	SP140--SP140 nuclear body protein (SP140), transcript variant 1, mRNA.
1.1811289	0.5209687	2.2671782	APOBEC3G--apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G (APOBEC3G), mRNA.
0.3834102	0.1691616	2.2665325	PRKCB--protein kinase C, beta (PRKCB), transcript variant 2, mRNA.
0.3134774	0.1383962	2.2650725	BTG1--B-cell translocation gene 1, anti-proliferative (BTG1), mRNA.
1.0872069	0.4800808	2.2646333	TMEM149--transmembrane protein 149 (TMEM149), mRNA.
2.4767249	1.0945237	2.2628335	MIR155HG--MIR155 host gene (non-protein coding) (MIR155HG), non-coding RNA.
1.1917184	0.5275769	2.2588523	AIM2--absent in melanoma 2 (AIM2), mRNA.
1.1904504	0.5273925	2.2572382	WARS--tryptophanyl-tRNA synthetase (WARS), transcript variant 3, mRNA.
2.7917405	1.2543619	2.225626	IGKC--Immunoglobulin kappa constant
0.1015575	0.0457346	2.2205834	LILRB2--leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2 (LILRB2), transcript variant 2, mRNA.
1.1585957	0.5229379	2.2155514	PSME2--proteasome (prosome, macropain) activator subunit 2 (PA28 beta) (PSME2), mRNA.
2.0224333	0.9178708	2.2033964	HLA-C--major histocompatibility complex, class I, C (HLA-C), mRNA.
0.1979883	0.0898573	2.2033627	ARHGAP9--Rho GTPase activating protein 9 (ARHGAP9), transcript variant 3, mRNA.
0.716873	0.3260291	2.1988009	TNFSF13B--tumor necrosis factor (ligand) superfamily, member 13b (TNFSF13B), transcript variant 1, mRNA.
0.6811418	0.3104281	2.1942014	PPP1R16B--protein phosphatase 1, regulatory (inhibitor) subunit 16B (PPP1R16B), transcript variant 2, mRNA.
0.1787107	0.0817326	2.1865287	--Unknown
0.1527546	0.0703309	2.171941	RPS27--ribosomal protein S27 (RPS27), mRNA.
2.4804908	1.1475198	2.1616104	IFIT5--interferon-induced protein with tetratricopeptide repeats 5 (IFIT5), mRNA.
			IL21R--interleukin 21 receptor (IL21R), transcript

Figure 9I

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
0.8498774	0.3934556	2.1600337	variant 1, mRNA.
2.1981814	1.021726	2.1514391	SMARCC2--SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c member 2 (SMARCC2), transcript variant 1, mRNA.
2.4987441	1.1701532	2.1353991	IGKV1D-8--Immunoglobulin kappa variable 1D-8
0.2338724	0.1095757	2.1343456	--40S RIBOSOMAL PROTEIN S27 HOMOLOG
1.3513318	0.6336913	2.1324764	PPP1R9B--protein phosphatase 1, regulatory (inhibitor) subunit 9B (PPP1R9B), mRNA.
16.1618288	7.6023984	2.1258855	APOE--apolipoprotein E (APOE), mRNA.
3.539066	1.6678796	2.1218954	NR1H3--nuclear receptor subfamily 1, group H, member 3 (NR1H3), transcript variant 3, mRNA.
0.1394283	0.0658896	2.1160876	LILRB2--leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2 (LILRB2), transcript variant 2, mRNA.
1.4207834	0.6723416	2.1131869	HLA-G--major histocompatibility complex, class I, G (HLA-G), mRNA.
0.0845117	0.0402999	2.0970709	DOK2--docking protein 2, 56kDa (DOK2), mRNA.
1.1951396	0.5704474	2.0950917	LILRB4--leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 4 (LILRB4), transcript variant 2, mRNA.
2.7364096	1.3070393	2.093594	TNFRSF9--tumor necrosis factor receptor superfamily, member 9 (TNFRSF9), mRNA.
4.1970712	2.0062134	2.0920363	SLAMF8--SLAM family member 8 (SLAMF8), mRNA.
0.2225488	0.1067236	2.0852819	AOAH--acyloxyacyl hydrolase (neutrophil) (AOAH), mRNA.
0.2120642	0.1017799	2.0835557	PAG1--phosphoprotein associated with glycosphingolipid microdomains 1 (PAG1), mRNA.
0.3625851	0.1748283	2.07395	LAIR1--leukocyte-associated immunoglobulin-like receptor 1 (LAIR1), transcript variant b, mRNA.
0.7127879	0.3437045	2.0738395	--CDNA FLJ11890 fis, clone HEMBA1007256
0.3305634	0.1603867	2.0610396	SAMD3--sterile alpha motif domain containing 3 (SAMD3), transcript variant 1, mRNA.
1.4153082	0.68921	2.0535225	FLJ31306--hypothetical LOC379025 (FLJ31306), transcript variant 2, non-coding RNA.
			C1QB--complement component 1, q subcomponent, B

Figure 9J

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
14.8821925	7.2686459	2.0474505	chain (C1QB), mRNA.
2.5933583	1.2668513	2.0470898	--MRNA; cDNA DKFZp564G103 (from clone DKFZp564G103)
1.7601831	0.8678035	2.0283199	MCL1--myeloid cell leukemia sequence 1 (BCL2 -related) (MCL1), transcript variant 1, mRNA.
1.1717005	0.5786024	2.025053	CD53--CD53 molecule (CD53), transcript variant 2, mRNA.
0.5798639	0.2865651	2.0234975	CCR7--chemokine (C-C motif) receptor 7 (CCR7), mRNA.
1.1949684	0.5947747	2.0091111	CD8A--CD8a molecule (CD8A), transcript variant 2, mRNA.
2.9635191	1.4796427	2.0028613	GPR89B--G protein-coupled receptor 89B (GPR89B), mRNA.
0.6153218	0.3076254	2.0002304	--Unknown
1.4882668	0.7456641	1.9958942	APOL3--apolipoprotein L, 3 (APOL3), transcript variant alpha/d, mRNA.
2.9688619	1.4875397	1.9958204	PPA1--pyrophosphatase (inorganic) 1 (PPA1), mRNA.
0.4527862	0.2268728	1.9957711	GIMAP2--GTPase, IMAP family member 2 (GIMAP2), mRNA.
0.1739881	0.0873653	1.991501	RPS27--ribosomal protein S27 (RPS27), mRNA.
1.6329699	0.8206203	1.9899214	CCDC108--coiled-coil domain containing 108 (CCDC108), transcript variant 1, mRNA.
0.2594353	0.1305324	1.9875158	--Unknown
1.2364835	0.6240647	1.9813385	STAT2--signal transducer and activator of transcription 2, 113kDa (STAT2), mRNA.
1.6313059	0.8437595	1.9333779	HTATIP2--HIV-1 Tat interactive protein 2, 30kDa (HTATIP2), transcript variant 2, mRNA.
3.2292849	1.6762122	1.926537	PPP1R3F--Protein phosphatase 1, regulatory (inhibitor) subunit 3F
0.4513523	0.2343138	1.9262723	SAMSN1--SAM domain, SH3 domain and nuclear localization signals 1 (SAMSN1), mRNA.
1.2554541	0.6539053	1.9199327	NFKBIL2--nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 2 (NFKBIL2), mRNA.
1.413855	0.7372466	1.9177504	--Unknown
1.2212251	0.6376385	1.9152312	C5orf20--chromosome 5 open reading frame 20 (C5orf20), mRNA.
0.8980528	0.4693521	1.9133882	CTLA4--cytotoxic T-lymphocyte-associated protein 4 (CTLA4), transcript variant 2, mRNA.
			ZAP70--zeta-chain (TCR) associated protein kinase

Figure 9K

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
0.3546162	0.1858474	1.908104	70kDa (ZAP70), transcript variant 2, mRNA.
1.5679927	0.8229066	1.9054322	EFR3A--EFR3 homolog A (S. cerevisiae) (EFR3A), mRNA.
1.7276753	0.9070891	1.9046369	DNAH6--dynein, axonemal, heavy chain 6 (DNAH6), mRNA.
1.7710022	0.9315393	1.9011567	IL15RA--interleukin 15 receptor, alpha (IL15RA), transcript variant 1, mRNA.
0.9078905	0.4776226	1.9008532	OS9--osteosarcoma amplified 9, endoplasmic reticulum lectin (OS9), transcript variant 3, mRNA.
1.8756064	0.9889619	1.8965405	PSME1--proteasome (prosome, macropain) activator subunit 1 (PA28 alpha) (PSME1), transcript variant 1, mRNA.
0.4449865	0.234665	1.8962627	SAMHD1--SAM domain and HD domain 1 (SAMHD1), mRNA.
1.0644845	0.5616393	1.895317	--CDNA FLJ34038 fis, clone FCBBF2005645
6.3377947	3.365398	1.8832229	C1QA--complement component 1, q subcomponent, A chain (C1QA), mRNA.
0.4142254	0.2209566	1.874691	TAGAP--T-cell activation RhoGTPase activating protein (TAGAP), transcript variant 1, mRNA.
0.3041359	0.1625613	1.8708999	ICAM2--intercellular adhesion molecule 2 (ICAM2), transcript variant 4, mRNA.
0.2129259	0.1138246	1.8706487	RPL26--ribosomal protein L26 (RPL26), mRNA.
1.1115032	0.6022825	1.845485	LOC100128203--PREDICTED: similar to hCG2040272 (LOC100128203), mRNA.
3.2371634	1.7578951	1.8414998	CNTNAP5--contactin associated protein-like 5 (CNTNAP5), mRNA.
0.9944294	0.5407174	1.8390924	BTN3A1--butyrophilin, subfamily 3, member A1 (BTN3A1), transcript variant 3, mRNA.
0.3750566	0.2054621	1.8254298	SH2D2A--SH2 domain protein 2A (SH2D2A), transcript variant 4, mRNA.
0.6217485	0.340878	1.8239621	PSMB9--Proteasome (prosome, macropain) subunit, beta type, 9 (large multifunctional peptidase 2)
1.4913772	0.8177582	1.8237385	IGHA1--Immunoglobulin heavy constant alpha 1
0.5253144	0.2889291	1.8181431	CRTAM--cytotoxic and regulatory T cell molecule (CRTAM), mRNA.
2.3055398	1.2707731	1.8142812	--Isolate P228K immunoglobulin kappa light chain variable region (IGKV1D-43) mRNA, IGKV1D-43*01 allele
1.0149149	0.5607632	1.8098814	ZNF90--Zinc finger protein 90
0.5158057	0.2862114	1.8021846	GCET2--germinal center expressed transcript 2 (GCET2), transcript variant 1, mRNA.

Figure 9L

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
0.6039561	0.3361639	1.796612	ZBP1--Z-DNA binding protein 1
0.4480985	0.2507615	1.7869506	NR3C1--nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor) (NR3C1), transcript variant 4, mRNA.
0.5790948	0.3244772	1.7847012	LILRB1--leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 1 (LILRB1), transcript variant 2, mRNA.
1.4520862	0.8167358	1.7779143	HMGN4--high mobility group nucleosomal binding domain 4 (HMGN4), mRNA.
22.9944842	12.9848932	1.7708643	C1QC--complement component 1, q subcomponent, C chain (C1QC), transcript variant 2, mRNA.
0.5473897	0.3098581	1.766582	S1PR4--sphingosine-1-phosphate receptor 4 (S1PR4), mRNA.
0.8602084	0.4876222	1.764088	LOC439949--PREDICTED: hypothetical protein LOC439949 (LOC439949), mRNA.
1.823593	1.0372648	1.7580784	PUM2--pumilio homolog 2 (Drosophila) (PUM2), mRNA.
0.1471318	0.0837999	1.755752	LEPROTL1--leptin receptor overlapping transcript-like 1 (LEPROTL1), transcript variant 1, mRNA.
0.4008678	0.2288622	1.7515684	ACAP1--ArfGAP with coiled-coil, ankyrin repeat and PH domains 1 (ACAP1), mRNA.
0.2628029	0.1500755	1.7511382	RASSF2--Ras association (RalGDS/AF-6) domain family member 2 (RASSF2), transcript variant 2, mRNA.
1.3806817	0.7896441	1.748486	RPL38--ribosomal protein L38 (RPL38), transcript variant 1, mRNA.
1.3352264	0.7655143	1.7442213	C2orf60--chromosome 2 open reading frame 60 (C2orf60), transcript variant 1, mRNA.
2.6759212	1.5616511	1.7135205	NR1H3--nuclear receptor subfamily 1, group H, member 3 (NR1H3), transcript variant 2, mRNA.
1.3890647	0.8125466	1.7095201	CYP2S1--cytochrome P450, family 2, subfamily S, polypeptide 1 (CYP2S1), mRNA.
1.2961979	0.7583284	1.7092831	GCH1--GTP cyclohydrolase 1 (GCH1), transcript variant 1, mRNA.
0.4860473	0.2859263	1.6999045	FNBP1--formin binding protein 1 (FNBP1), mRNA.
0.8342972	0.4909102	1.6994904	TRIM25--tripartite motif-containing 25 (TRIM25), mRNA.
0.1879446	0.1106941	1.6978739	CD97--CD97 molecule (CD97), transcript variant 2, mRNA.

Figure 9M

Table 4 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
0.9439545	0.5582367	1.6909575	TBCC--tubulin folding cofactor C (TBCC), mRNA.
2.0137433	1.1914838	1.6901138	LAP3--leucine aminopeptidase 3 (LAP3), mRNA.
1.5880954	0.9421267	1.6856495	--Unknown
0.3888413	0.2321098	1.6752477	--Human DNA sequence from clone RP3-336K20 on chromosome 6 Contains parts of 2 genes for novel protein
0.8819085	0.5380899	1.6389613	AQP12A--aquaporin 12A (AQP12A), mRNA.
1.4447688	0.8851784	1.6321781	IL2RA--interleukin 2 receptor, alpha (IL2RA), mRNA.
1.3717441	0.8415628	1.629996	PFN1--profilin 1 (PFN1), mRNA.
1.9659229	1.2130635	1.6206266	NBPF15--neuroblastoma breakpoint family, member 15 (NBPF15), transcript variant 1, mRNA.
0.7399832	0.4588757	1.6126005	XCL2--chemokine (C motif) ligand 2 (XCL2), mRNA.
0.7600897	0.472468	1.6087643	CR1--complement component (3b/4b) receptor 1 (Knops blood group) (CR1), transcript variant F, mRNA.
1.2379365	0.7710457	1.6055293	LAX1--lymphocyte transmembrane adaptor 1 (LAX1), transcript variant 2, mRNA.
0.1847703	0.1167514	1.5825962	FGD2--FYVE, RhoGEF and PH domain containing 2 (FGD2), mRNA.
1.79192	1.1353552	1.5782902	C19orf35--chromosome 19 open reading frame 35 (C19orf35), mRNA.
1.7019732	1.1270768	1.5100774	NCRNA00183--Non-protein coding RNA 183
0.5586752	0.3740648	1.4935252	CD28--CD28 molecule (CD28), mRNA.
0.3007145	0.2029006	1.4820783	IL18RAP--interleukin 18 receptor accessory protein (IL18RAP), mRNA.
0.8818769	0.5950284	1.4820753	EDEM1--ER degradation enhancer, mannosidase alpha-like 1 (EDEM1), mRNA.
1.9792207	1.3692456	1.4454826	FPR3--formyl peptide receptor 3 (FPR3), mRNA.

Figure 9N

Table 5

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
0.9303583	1.3842925	0.6720821	GATSL3--GATS protein-like 3 (GATSL3), mRNA.
1.3297317	1.9895554	0.6683562	ANXA4--annexin A4 (ANXA4), mRNA.
1.9190275	2.9841036	0.6430834	ICMT--isoprenylcysteine carboxyl methyltransferase (ICMT), mRNA.
2.073402	3.2352174	0.6408849	SNUPN--snurportin 1 (SNUPN), transcript variant 3, mRNA.
0.7979328	1.2552768	0.6356628	MYOM1--myomesin 1, 185kDa (MYOM1), transcript variant 2, mRNA.
1.5990019	2.5275325	0.6326336	ZFP14--zinc finger protein 14 homolog (mouse) (ZFP14), mRNA.
1.0074162	1.6006794	0.6293678	PHF2--PHD finger protein 2 (PHF2), mRNA.
0.2468642	0.3979965	0.6202673	CREB5--cAMP responsive element binding protein 5 (CREB5), transcript variant 4, mRNA.
1.0493231	1.6931362	0.6197512	SLC12A4--solute carrier family 12 (potassium/chloride transporters), member 4 (SLC12A4), transcript variant 5, mRNA.
0.9703596	1.5672031	0.6191664	IFT20--intraflagellar transport 20 homolog (Chlamydomonas) (IFT20), mRNA.
1.4646061	2.3849381	0.6141065	GPR107--G protein-coupled receptor 107 (GPR107), transcript variant 3, mRNA.
0.8396772	1.376139	0.6101689	SLC16A5--solute carrier family 16, member 5 (monocarboxylic acid transporter 6) (SLC16A5), mRNA.
4.1086023	6.7959067	0.6045701	CTSF--cathepsin F (CTSF), mRNA.
1.3596031	2.251731	0.6038035	PHF13--PHD finger protein 13 (PHF13), mRNA.
0.7069344	1.1736149	0.6023564	NP_001013662.1--Unknown
1.2520582	2.0796506	0.6020522	LOC116437--hypothetical LOC116437 (LOC116437), non-coding RNA.
1.8478321	3.0992014	0.5962285	VPS37C--vacuolar protein sorting 37 homolog C (S. cerevisiae) (VPS37C), mRNA.
1.5212941	2.5691825	0.5921316	TK2--thymidine kinase 2, mitochondrial (TK2), transcript variant 2, mRNA.
1.2247562	2.0746593	0.5903409	MAN1B1--mannosidase, alpha, class 1B, member 1 (MAN1B1), mRNA.
0.9830808	1.7379351	0.5656603	LOC650226--ankyrin repeat domain 26 pseudogene (LOC650226), non-coding RNA.
0.7038196	1.2542699	0.5611389	PTDSS2--phosphatidylserine synthase 2 (PTDSS2), mRNA.
1.3520219	2.468806	0.547642	RIC8A--resistance to inhibitors of cholinesterase 8 homolog A (C. elegans) (RIC8A), mRNA.
3.2528513	5.9419578	0.5474376	CERCAM--cerebral endothelial cell adhesion molecule (CERCAM), mRNA.
			FAM108C1--family with sequence similarity 108,

Figure 10A

Table 5 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
3.0197554	5.579572	0.5412163	member C1 (FAM108C1), mRNA.
3.0593703	5.677145	0.5388924	SLC35E1--solute carrier family 35, member E1 (SLC35E1), mRNA.
3.781783	7.0735401	0.5346379	GFPT1--glutamine-fructose-6-phosphate transaminase 1 (GFPT1), mRNA.
1.3063471	2.4561498	0.5318678	PYROXD2--pyridine nucleotide-disulphide oxidoreductase domain 2 (PYROXD2), mRNA.
2.6912041	5.0678499	0.5310347	VPS24--vacuolar protein sorting 24 homolog (S. cerevisiae) (VPS24), transcript variant 2, mRNA.
1.0099103	1.904454	0.5302886	SIRPB1--signal-regulatory protein beta 1 (SIRPB1), transcript variant 3, mRNA.
0.6936946	1.3128172	0.5284015	C22orf39--chromosome 22 open reading frame 39 (C22orf39), transcript variant 2, mRNA.
0.4174286	0.7974344	0.5234645	ADAM20--ADAM metallopeptidase domain 20 (ADAM20), mRNA.
2.8418385	5.4982447	0.5168629	VKORC1--vitamin K epoxide reductase complex, subunit 1 (VKORC1), transcript variant 2, mRNA.
2.0532447	4.0472441	0.5073192	C1QTNF6--C1q and tumor necrosis factor related protein 6 (C1QTNF6), transcript variant 2, mRNA.
7.4588517	14.7768327	0.5047666	ITGB5--integrin, beta 5 (ITGB5), mRNA.
0.993252	1.989044	0.4993615	IER2--immediate early response 2 (IER2), mRNA.
2.704991	5.5093111	0.4909853	COPS7A--COP9 constitutive photomorphogenic homolog subunit 7A (Arabidopsis) (COPS7A), transcript variant 4, mRNA.
0.4047136	0.8304736	0.4873287	TNFAIP8L1--tumor necrosis factor, alpha-induced protein 8-like 1 (TNFAIP8L1), transcript variant 2, mRNA.
1.793315	3.77661	0.4748478	TMC4--transmembrane channel-like 4 (TMC4), transcript variant 2, mRNA.
0.6875513	1.5101135	0.4552978	BBS2--Bardet-Biedl syndrome 2 (BBS2), mRNA.
0.669445	1.4760492	0.4535384	TCEAL3--transcription elongation factor A (SII)-like 3 (TCEAL3), transcript variant 2, mRNA.
1.8322637	4.0718716	0.4499807	C14orf45--chromosome 14 open reading frame 45 (C14orf45), mRNA.
0.8849078	2.0458265	0.4325429	ULK1--unc-51-like kinase 1 (C. elegans) (ULK1), mRNA.
2.234576	5.1711924	0.4321201	LOC100132800--Hypothetical protein LOC100132800
2.8600432	6.7893429	0.4212548	COMP--cartilage oligomeric matrix protein (COMP), mRNA.

Figure 10B

Table 5 (Cont.)

Geom mean of ratios in class 1 (relapse free)	Geom mean of ratios in class 2 (Relapse)	Fold-change	Name
2.091045	5.0322269	0.4155307	PER3--period homolog 3 (Drosophila) (PER3), mRNA.
0.4393936	1.0698667	0.4106994	PYGL--phosphorylase, glycogen, liver (PYGL), transcript variant 2, mRNA.
2.4901603	6.4396556	0.3866915	GJB2--gap junction protein, beta 2, 26kDa (GJB2), mRNA.
0.3765888	0.983926	0.382741	JDP2--Jun dimerization protein 2 (JDP2), transcript variant 3, mRNA.
2.1586802	6.046314	0.3570242	BHMT2--betaine-homocysteine methyltransferase 2 (BHMT2), mRNA.
1.6105298	7.5546461	0.213184	C1QTNF3--C1q and tumor necrosis factor related protein 3 (C1QTNF3), transcript variant 1, mRNA.

Figure 10C

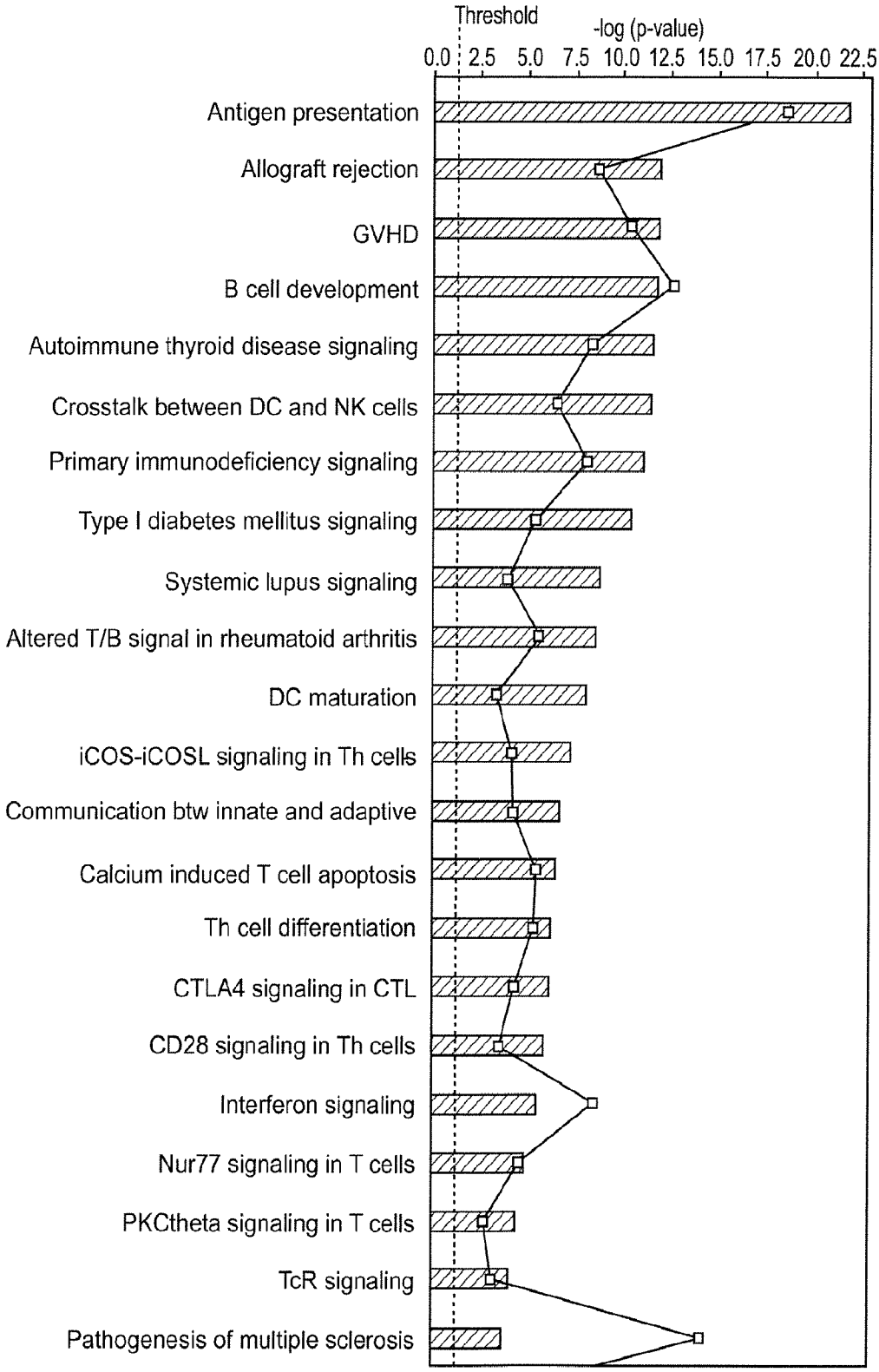

Threshold
-log (p-value)
0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5 20.0 22.5
Antigen presentation
Allograft rejection
GVHD
B cell development
Autoimmune thyroid disease signaling
Crosstalk between DC and NK cells
Primary immunodeficiency signaling
Type I diabetes mellitus signaling
Systemic lupus signaling
Altered T/B signal in rheumatoid arthritis
DC maturation
iCOS-iCOSL signaling in Th cells
Communication btw innate and adaptive
Calcium induced T cell apoptosis
Th cell differentiation
CTLA4 signaling in CTL
CD28 signaling in Th cells
Interferon signaling
Nur77 signaling in T cells
PKCtheta signaling in T cells
TcR signaling
Pathogenesis of multiple sclerosis

Figure 11A

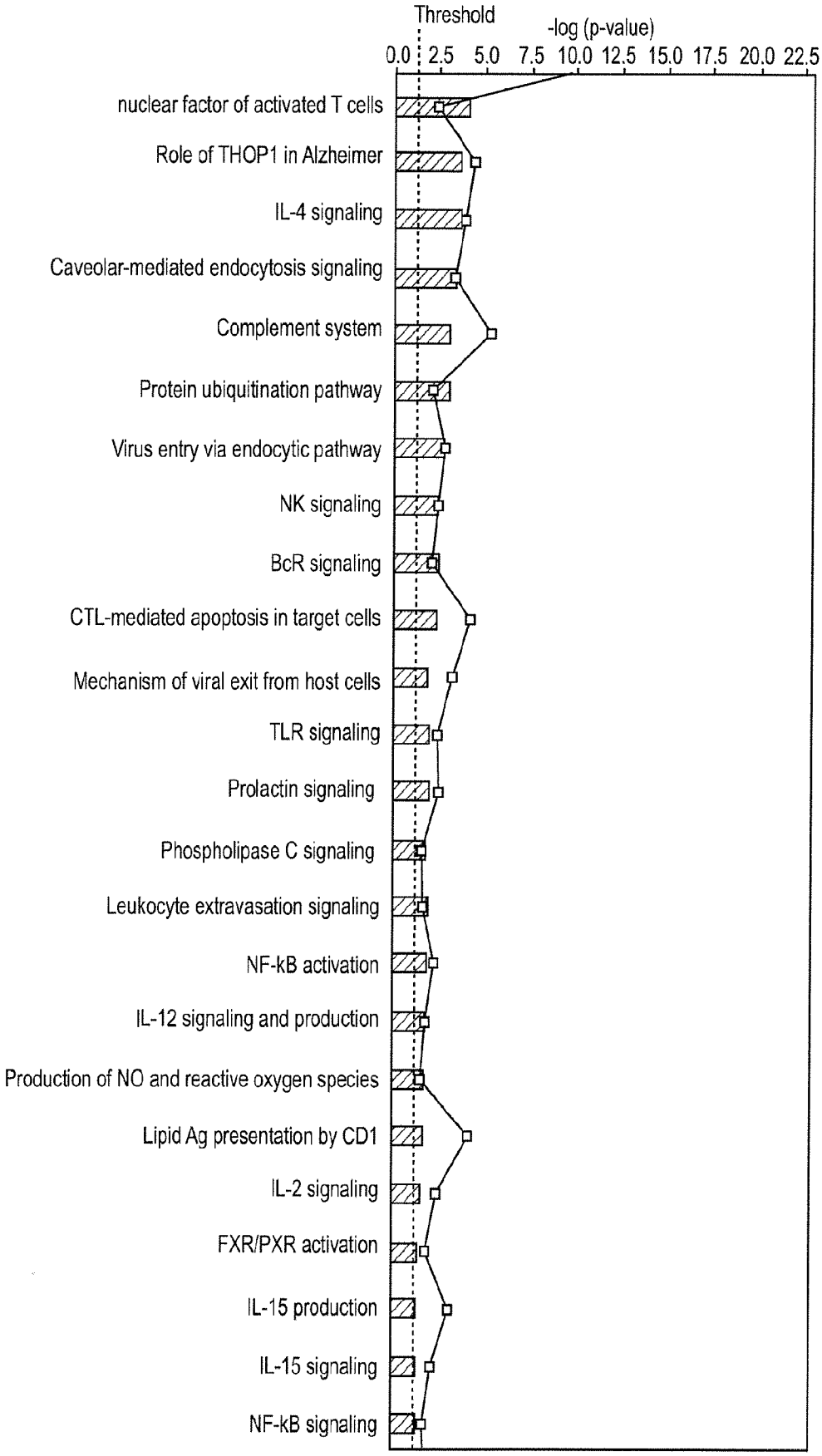

Threshold
-log (p-value)
0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5 20.0 22.5
nuclear factor of activated T cells
Role of THOP1 in Alzheimer
IL-4 signaling
Caveolar-mediated endocytosis signaling
Complement system
Protein ubiquitination pathway
Virus entry via endocytic pathway
NK signaling
BcR signaling
CTL-mediated apoptosis in target cells
Mechanism of viral exit from host cells
TLR signaling
Prolactin signaling
Phospholipase C signaling
Leukocyte extravasation signaling
NF-kB activation
IL-12 signaling and production
Production of NO and reactive oxygen species
Lipid Ag presentation by CD1
IL-2 signaling
FXR/PXR activation
IL-15 production
IL-15 signaling
NF-kB signaling

Figure 11B

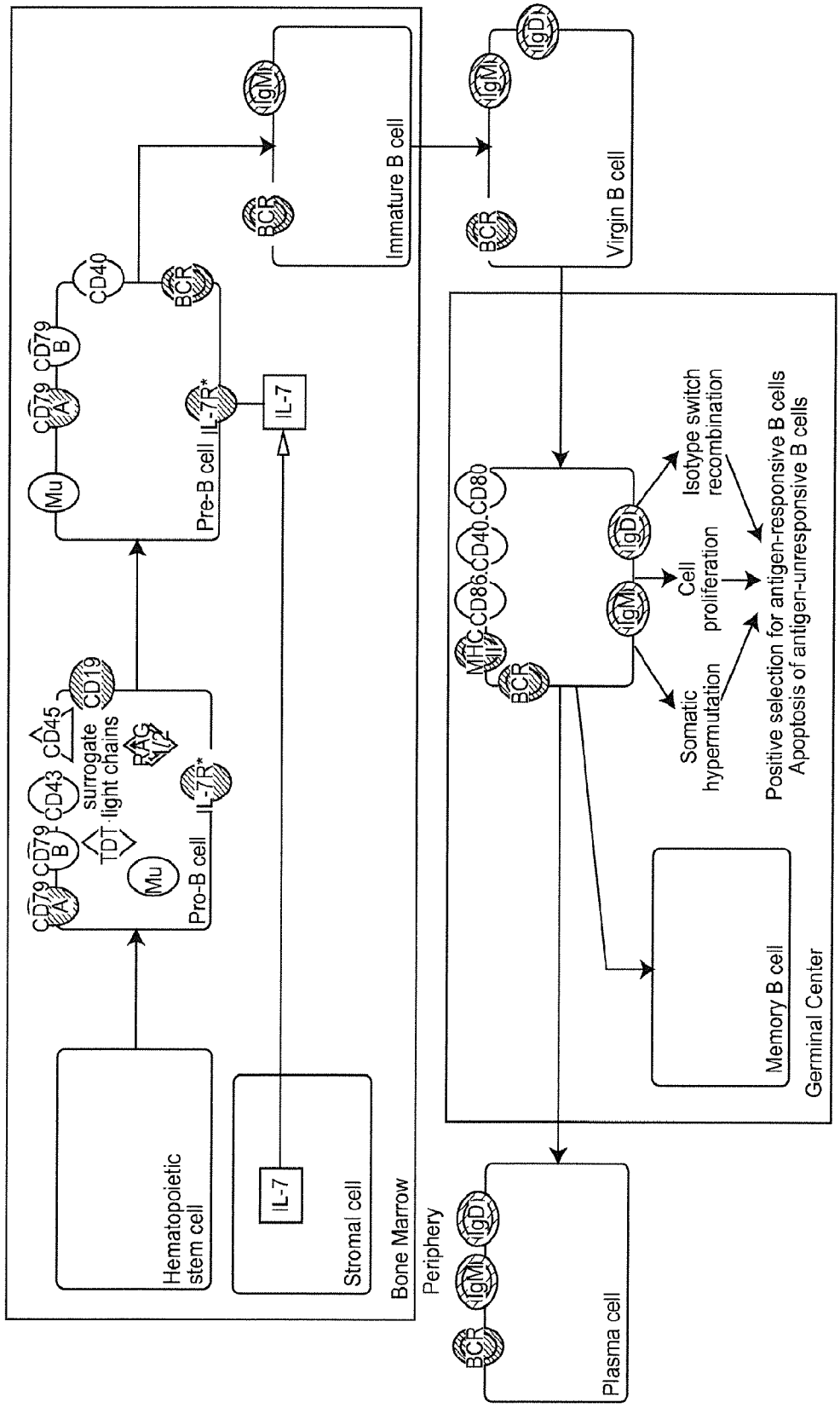

Bone Marrow
Periphery
Hematopoietic stem cell
Stromal cell
IL-7
Pro-B cell
Pre-B cell
surrogate light chains
TDT
RAG
IL-7R*
CD79A
CD79B
CD43
CD45
CD19
Mu
CD40
BCR
IgM
IgD
Immature B cell
Virgin B cell
Germinal Center
MHC
CD86
CD80
Somatic hypermutation
Cell proliferation
Isotype switch recombination
Positive selection for antigen-responsive B cells
Apoptosis of antigen-unresponsive B cells
Memory B cell
Plasma cell

Figure 12A

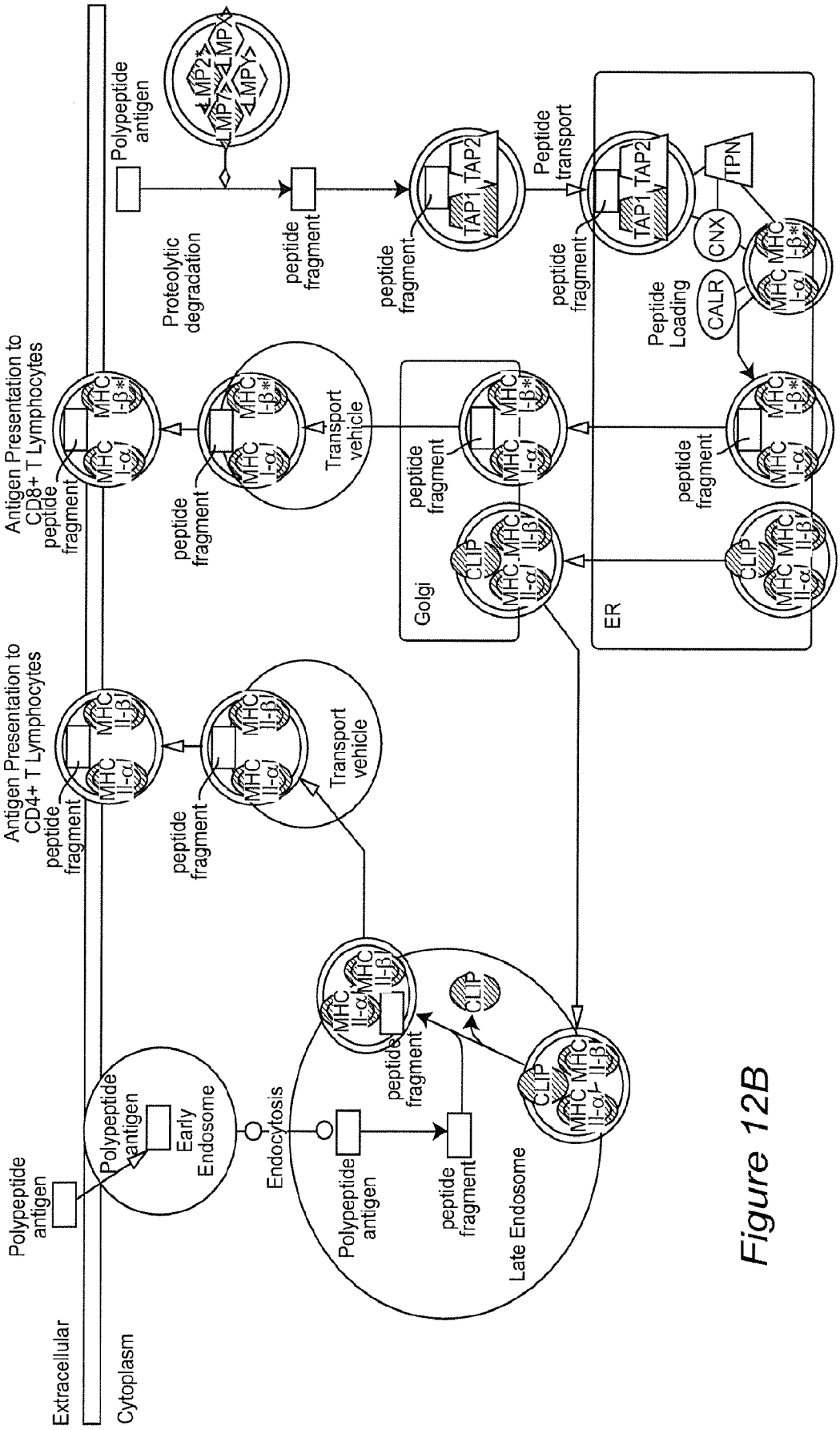

Extracellular
Cytoplasm
Polypeptide antigen
Polypeptide antigen
Early Endosome
Endocytosis
Polypeptide antigen
peptide fragment
Late Endosome
CLIP
MHC II-α MHC II-β
Antigen Presentation to CD4+ T Lymphocytes
peptide fragment
MHC II-α MHC II-β
peptide fragment
MHC II-α MHC II-β
Transport vehicle
Antigen Presentation to CD8+ T Lymphocytes
peptide fragment
MHC I-α MHC I-β*
peptide fragment
MHC I-α MHC I-β*
Transport vehicle
Golgi
peptide fragment
CLIP
MHC II-α MHC II-β
peptide fragment
MHC I-α MHC I-β*
ER
CLIP
MHC II-α MHC II-β
peptide fragment
MHC I-α MHC I-β*
Peptide Loading
CALR
CNX
TPN
MHC I-α MHC I-β*
peptide fragment
TAP1 TAP2
Peptide transport
peptide fragment
TAP1 TAP2
Proteolytic degradation
peptide fragment
LMP2 LMP7 LMP10 LMP5
Polypeptide antigen

Figure 12B

Figure 12C

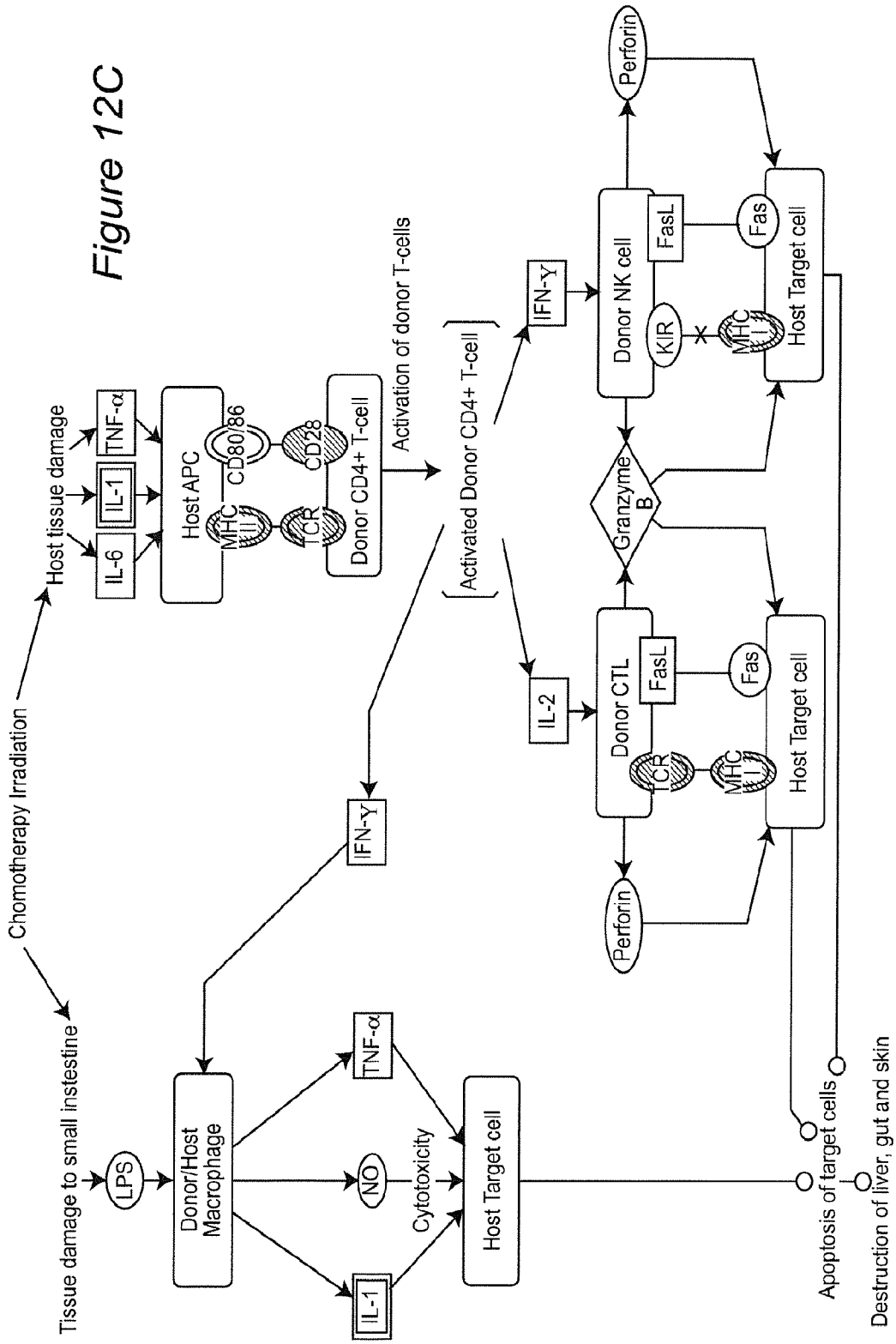
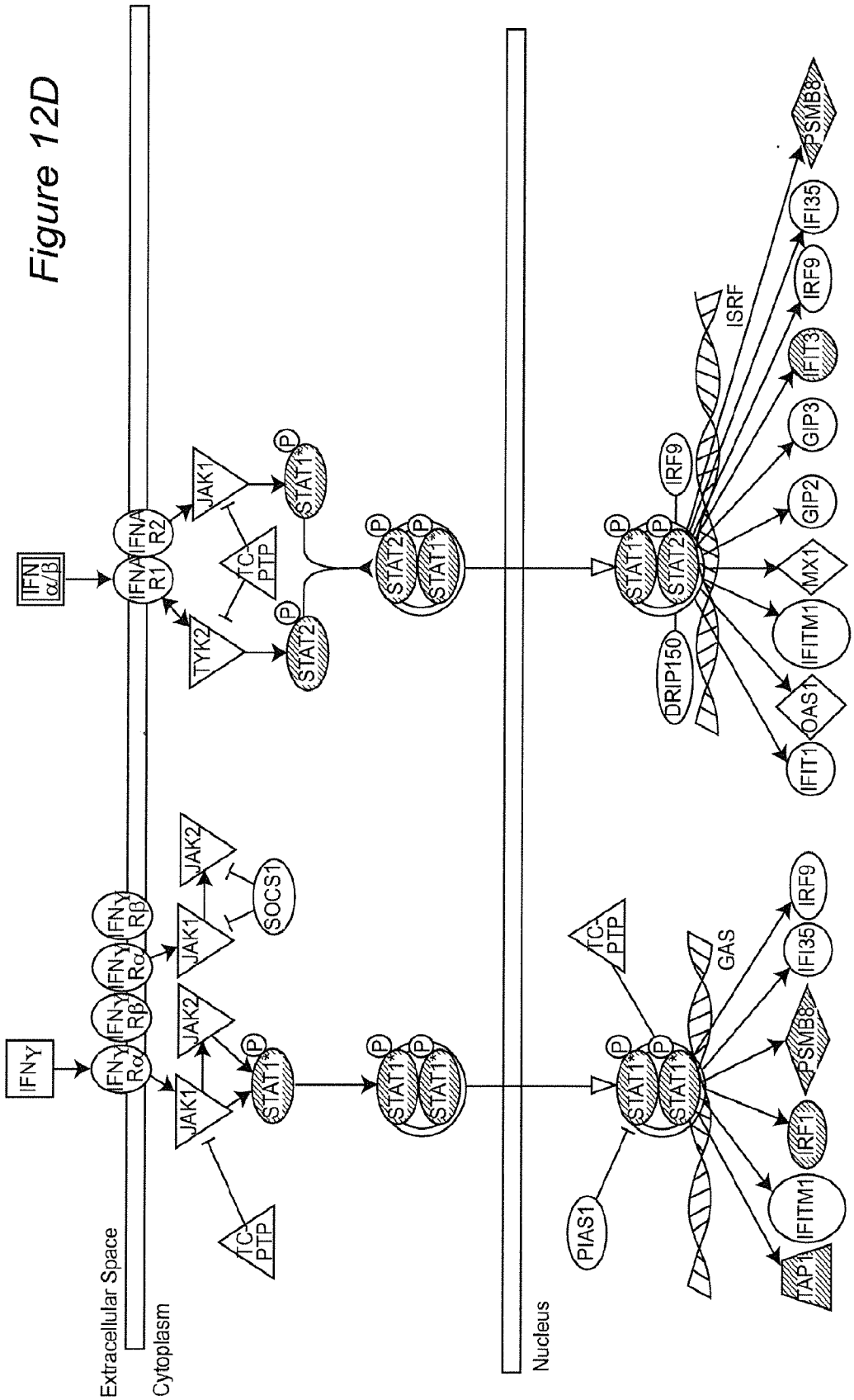

Chomotherapy Irradiation
Tissue damage to small instestine
LPS
Donor/Host Macrophage
TNF-α
NO
IL-1
Cytotoxicity
Host Target cell
Apoptosis of target cells
Destruction of liver, gut and skin
Host tissue damage
IL-6
IL-1
TNF-α
Host APC
MHC II
CD80/86
TCR
CD28
Donor CD4+ T-cell
Activation of donor T-cells
Activated Donor CD4+ T-cell
IFN-γ
IL-2
Donor NK cell
KIR
MHC I
FasL
Fas
Perforin
Granzyme B
Donor CTL
TCR
MHC I
FasL
Fas
Perforin
Host Target cell
Host Target cell

Figure 12D


IFNγ
IFN α/β
IFNγ Rα
IFNγ Rβ
IFNA R1
IFNA R2
JAK1
JAK2
TYK2
TC-PTP
SOCS1
STAT1*
STAT2
P
IRF9
DRIP150
ISRF
GAS
PIAS1
MX1
GIP2
GIP3
IFIT3
IRF9
IFI35
PSMB8
IFIT1
OAS1
IFITM1
IRF1
TAP1
Extracellular Space
Cytoplasm
Nucleus

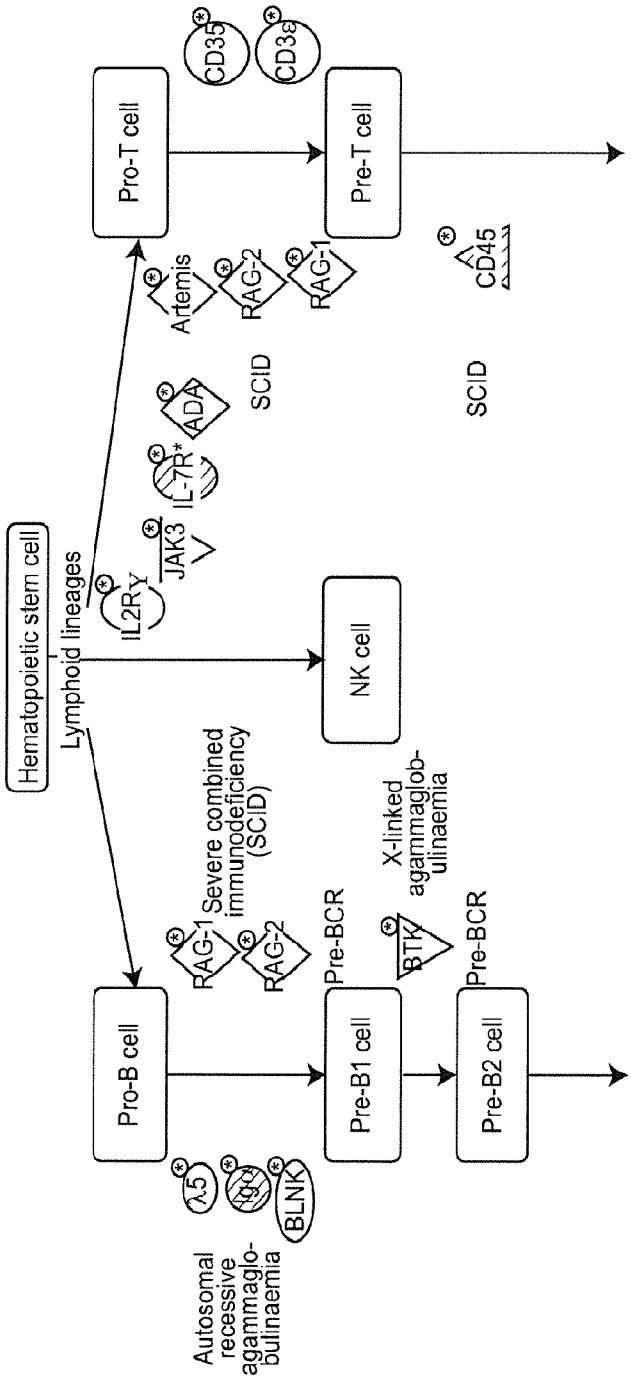

Hematopoietic stem cell
Lymphoid lineages
Pro-B cell
Pre-B1 cell
Pre-B2 cell
Pro-T cell
Pre-T cell
NK cell
Pre-BCR
Pre-BCR
λ5
BLNK
Autosomal recessive agammaglobulinaemia
RAG-1
RAG-2
Severe combined immunodeficiency (SCID)
BTK
X-linked agammaglobulinaemia
IL2RY
JAK3
IL-7R*
ADA
SCID
Artemis
RAG-2
RAG-1
SCID
CD35
CD3ε
CD45

Figure 12E-1

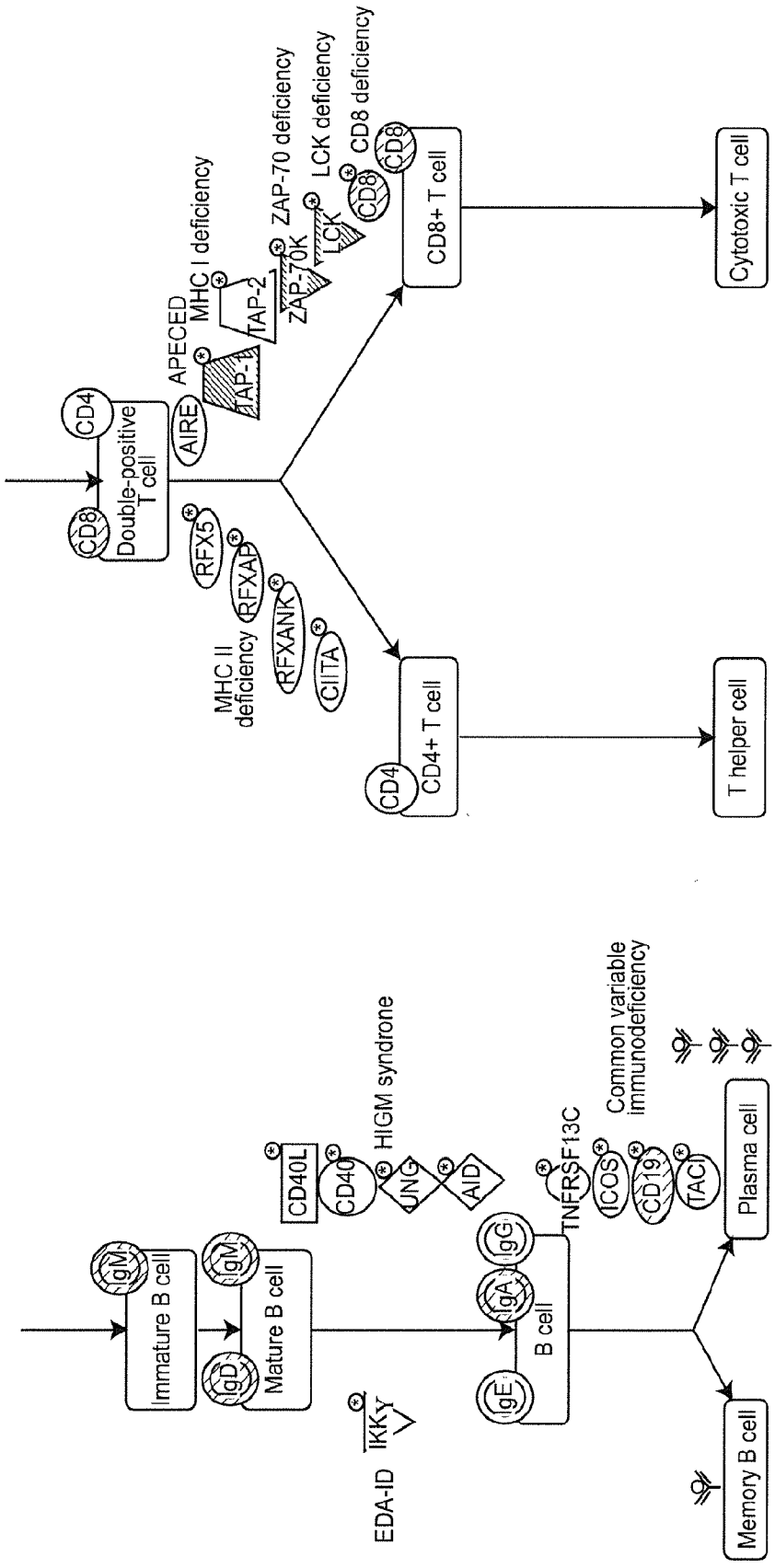

Double-positive T cell
CD4
CD8
AIRE
APECED
MHC I deficiency
TAP-1
TAP-2
ZAP-70K
ZAP-70 deficiency
LCK
LCK deficiency
CD8
CD8 deficiency
CD8+ T cell
Cytotoxic T cell
RFX5
RFXAP
RFXANK
CIITA
MHC II deficiency
CD4+ T cell
T helper cell
Immature B cell
IgM
IgD
Mature B cell
CD40L
CD40
HIGM syndrone
UNG
AID
IgM
IgA
IgG
IgE
B cell
TNFRSF13C
ICOS
CD19
TACI
Common variable immunodeficiency
Plasma cell
EDA-ID
IKKγ
Memory B cell

Figure 12E-2

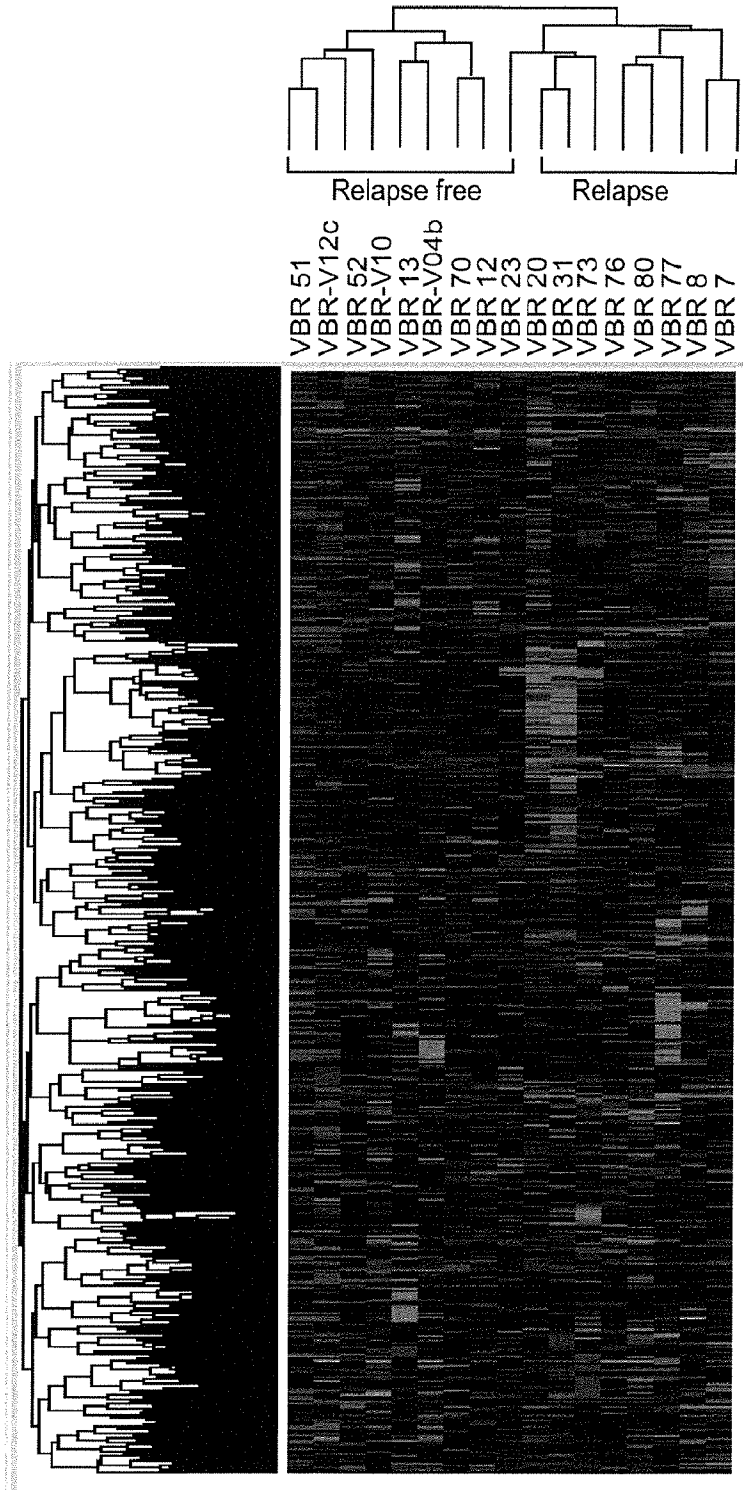

Relapse free
Relapse
VBR 51
VBR-V12c
VBR 52
VBR-V10
VBR 13
VBR-V04b
VBR 70
VBR 12
VBR 23
VBR 20
VBR 31
VBR 73
VBR 76
VBR 80
VBR 77
VBR 8
VBR 7

Figure 13A

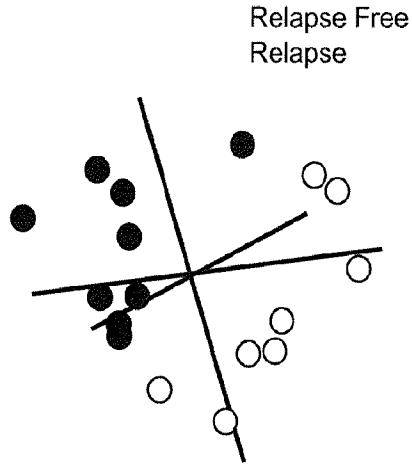


Figure 13B

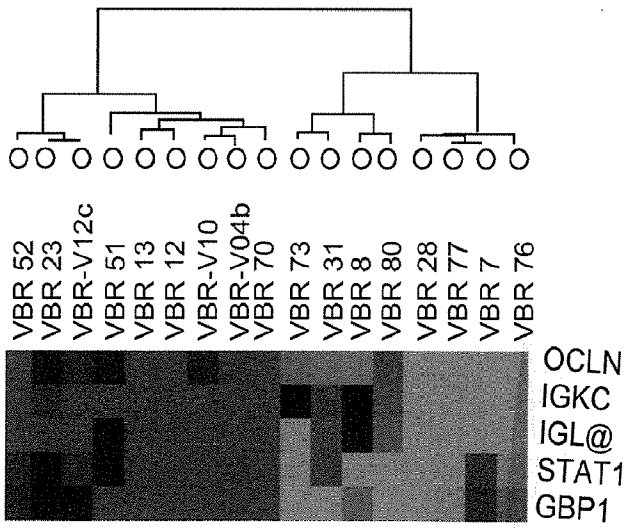


Figure 14

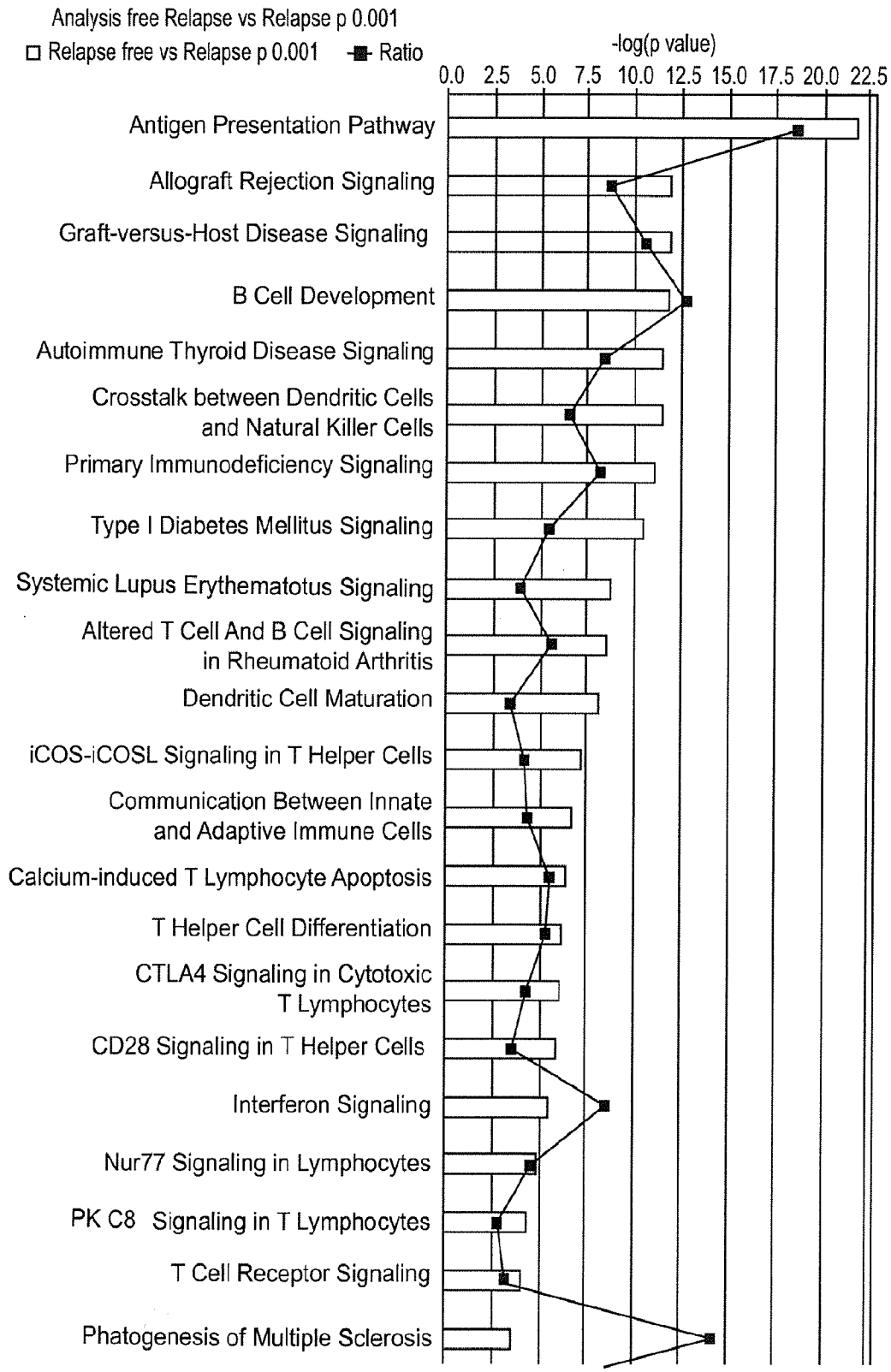

Analysis free Relapse vs Relapse p 0.001
Relapse free vs Relapse p 0.001
Ratio
-log(p value)
0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5 20.0 22.5
Antigen Presentation Pathway
Allograft Rejection Signaling
Graft-versus-Host Disease Signaling
B Cell Development
Autoimmune Thyroid Disease Signaling
Crosstalk between Dendritic Cells and Natural Killer Cells
Primary Immunodeficiency Signaling
Type I Diabetes Mellitus Signaling
Systemic Lupus Erythematotus Signaling
Altered T Cell And B Cell Signaling in Rheumatoid Arthritis
Dendritic Cell Maturation
iCOS-iCOSL Signaling in T Helper Cells
Communication Between Innate and Adaptive Immune Cells
Calcium-induced T Lymphocyte Apoptosis
T Helper Cell Differentiation
CTLA4 Signaling in Cytotoxic T Lymphocytes
CD28 Signaling in T Helper Cells
Interferon Signaling
Nur77 Signaling in Lymphocytes
PK C8 Signaling in T Lymphocytes
T Cell Receptor Signaling
Phatogenesis of Multiple Sclerosis

Figure 15

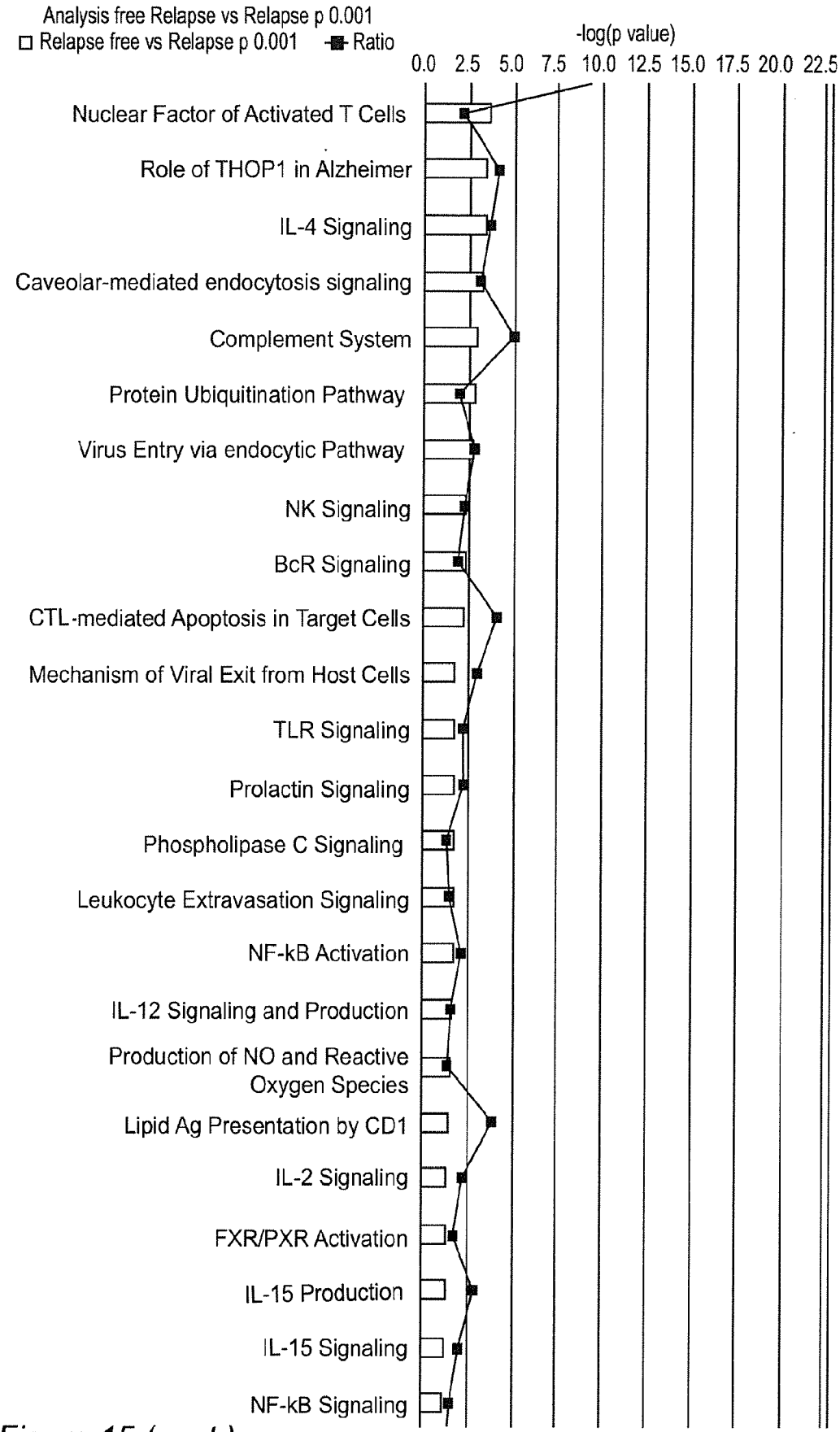


Figure 15 (cont.)

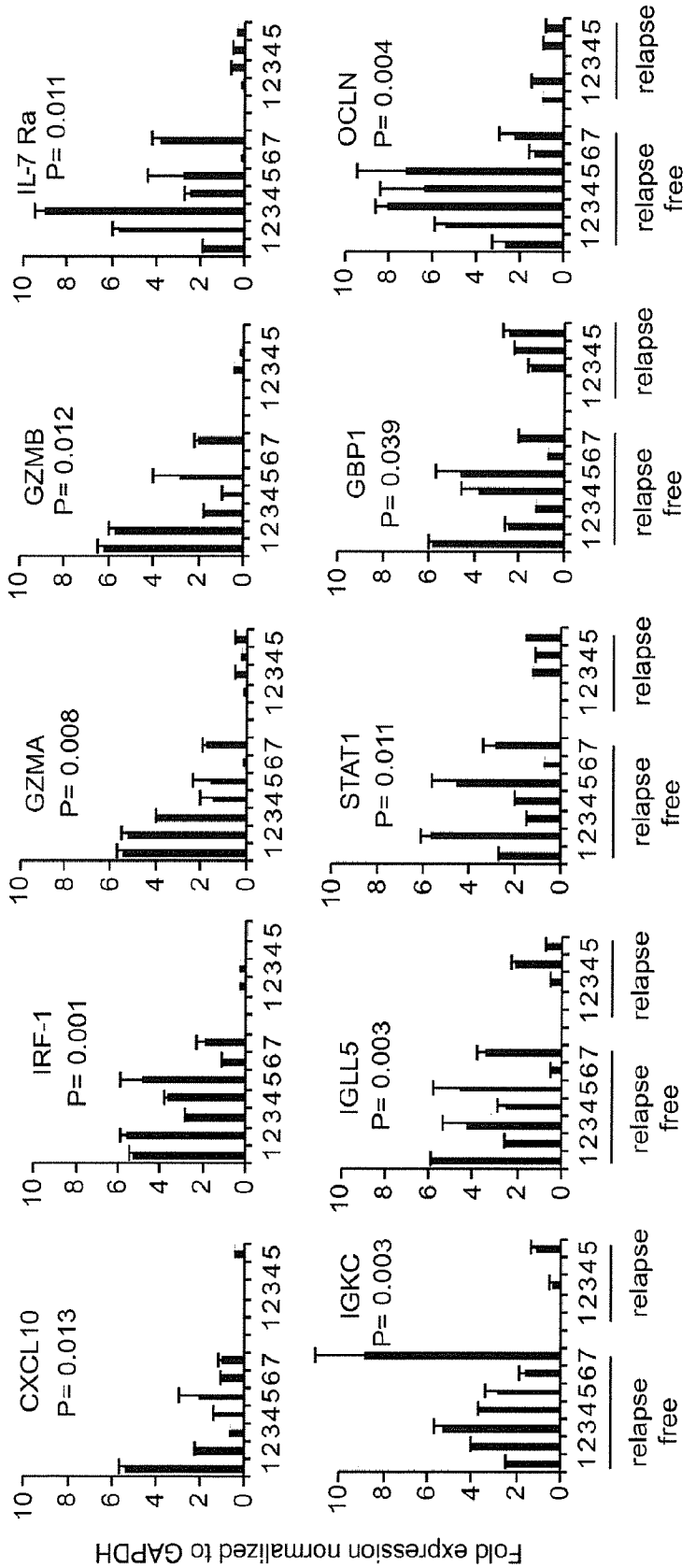

CXCL10
P= 0.013
IRF-1
P= 0.001
GZMA
P= 0.008
GZMB
P= 0.012
IL-7 Ra
P= 0.011
IGKC
P= 0.003
IGLL5
P= 0.003
STAT1
P= 0.011
GBP1
P= 0.039
OCLN
P= 0.004
Fold expression normalized to GAPDH
1234567 12345
relapse free
relapse relapse

Figure 16

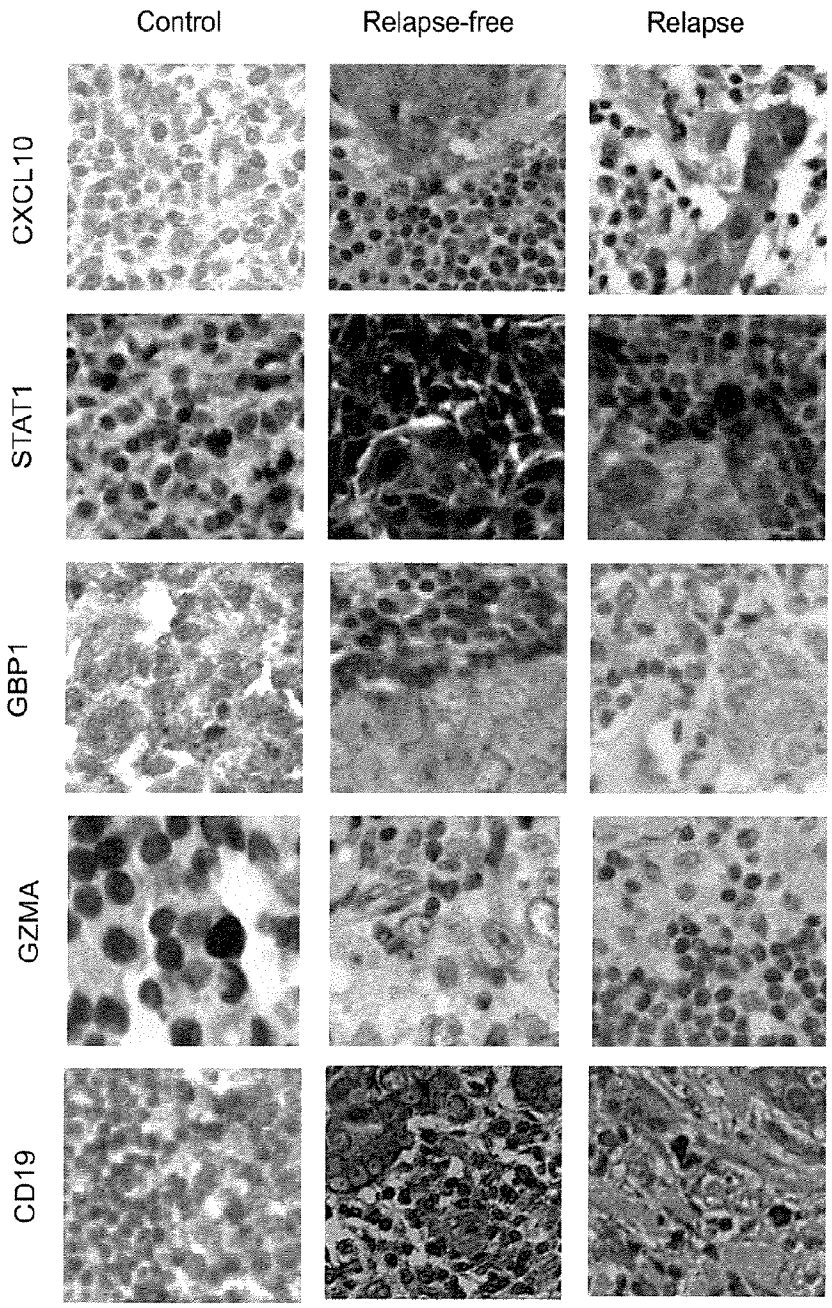


Figure 17A

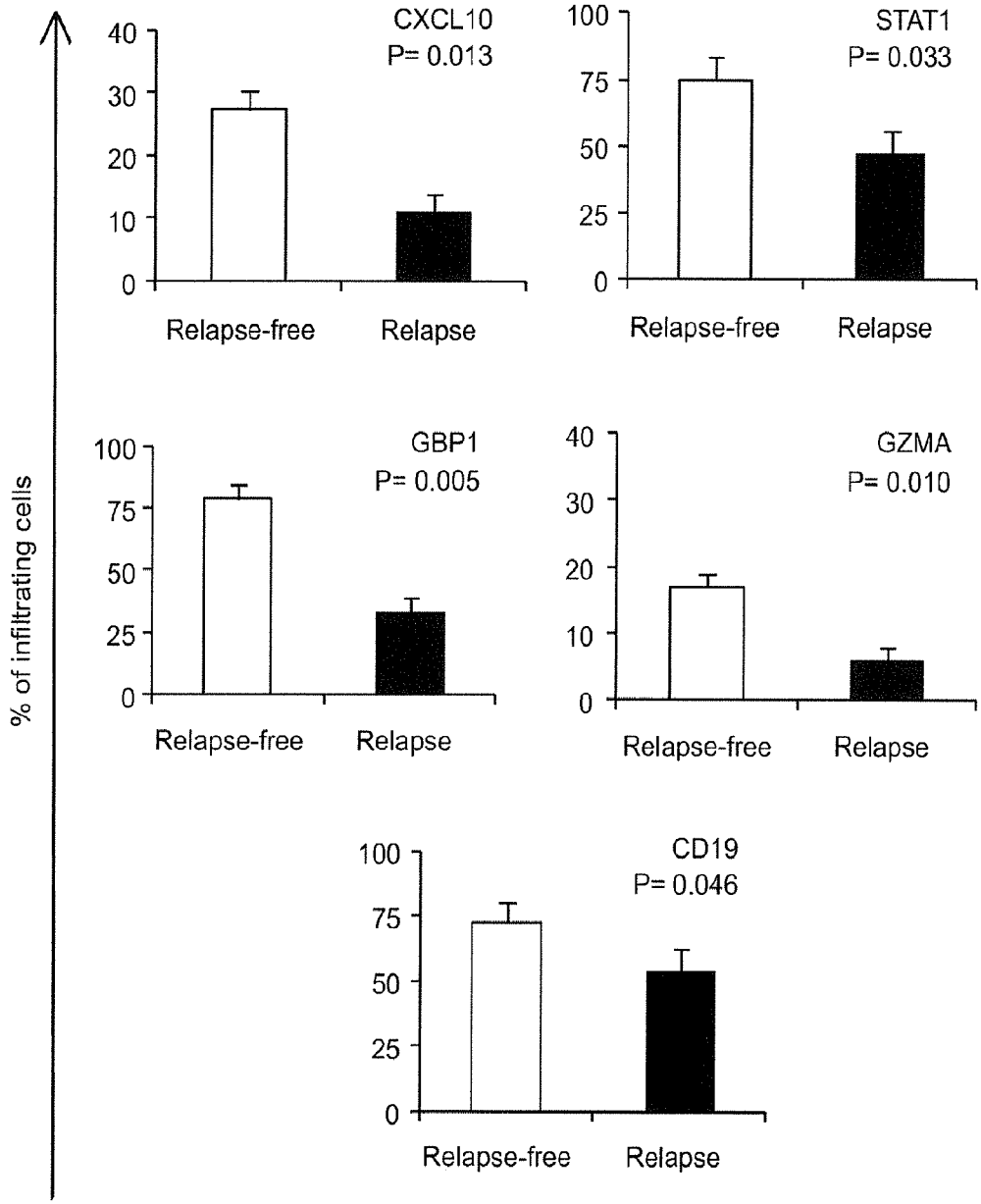

% of infiltrating cells
CXCL10
P= 0.013
40
30
20
10
0
Relapse-free
Relapse
STAT1
P= 0.033
100
75
50
25
0
Relapse-free
Relapse
GBP1
P= 0.005
100
75
50
25
0
Relapse-free
Relapse
GZMA
P= 0.010
40
30
20
10
0
Relapse-free
Relapse
CD19
P= 0.046
100
75
50
25
0
Relapse-free
Relapse

Figure 17B

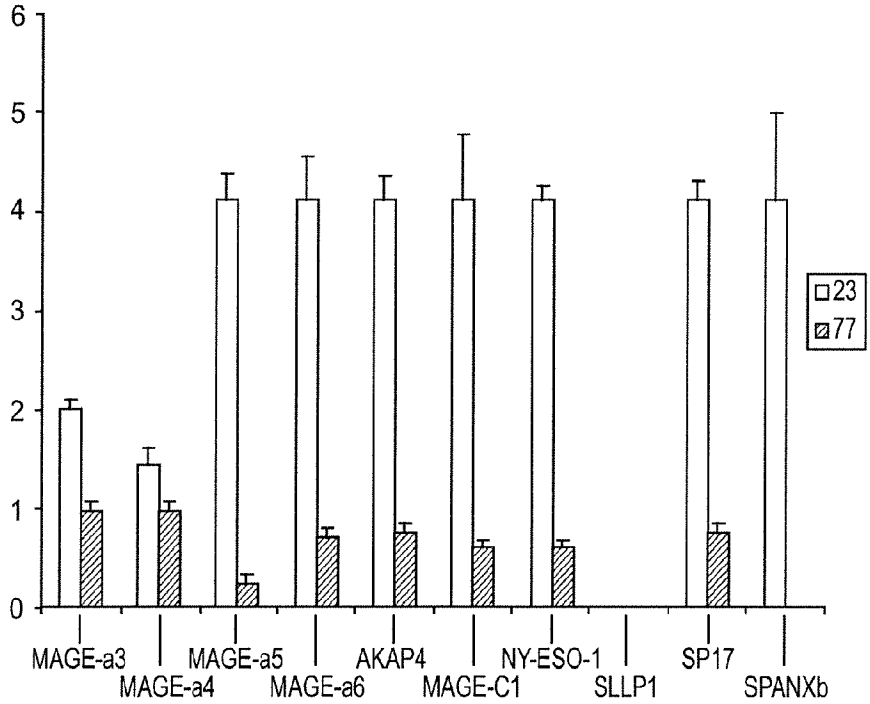


Figure 18A

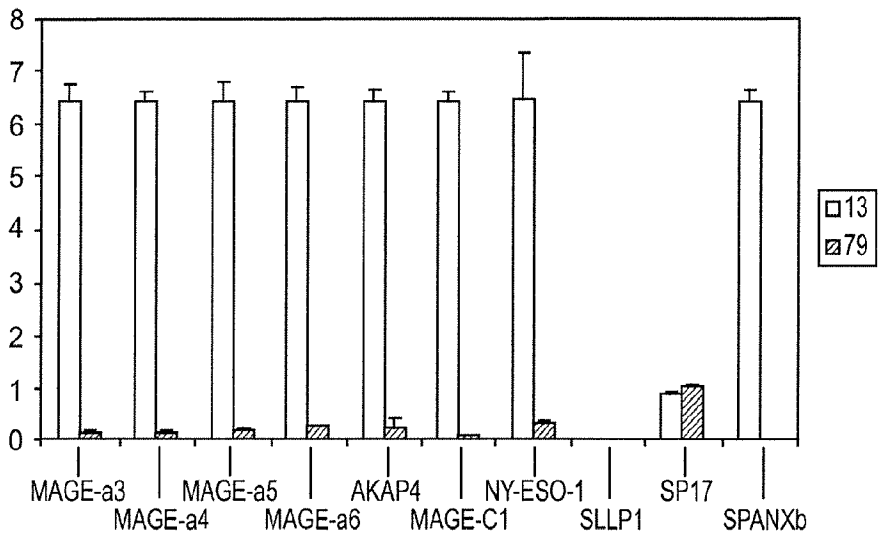


Figure 18B

1 gagattctcg ccctgagcaa cgagcgacgg cctgacgtcg gcggagggaa gccggcccag

61 gctcggtgag gaggcaaggt tctgagggga caggctgacc tggaggacca gaggcccccg

121 gaggagcact gaaggagaag atctgccagt gggtctccat tgcccagctc ctgcccacac

181 tcccgcctgt tgccctgacc agagtcatca tgcctcttga gcagaggagt cagcactgca

241 agcctgaaga aggccttgag gcccgaggag aggccctggg cctggtgggt gcgcaggctc

301 ctgctactga ggagcaggag gctgcctcct cctcttctac tctagttgaa gtcaccctgg

361 gggaggtgcc tgctgccgag tcaccagatc ctccccagag tcctcaggga gcctccagcc

421 tccccactac catgaactac cctctctgga gccaatccta tgaggactcc agcaaccaag

481 aagaggaggg gccaagcacc ttccctgacc tggagtccga gttccaagca gcactcagta

541 ggaaggtggc cgagttggtt cattttctgc tcctcaagta tcgagccagg gagccggtca

601 caaaggcaga aatgctgggg agtgtcgtcg gaaattggca gtatttcttt cctgtgatct

661 tcagcaaagc ttccagttcc ttgcagctgg tctttggcat cgagctgatg gaagtggacc

721 ccatcggcca cttgtacatc tttgccacct gcctgggcct ctcctacgat ggcctgctgg

781 gtgacaatca gatcatgccc aaggcaggcc tcctgataat cgtcctggcc ataatcgcaa

841 gagagggcga ctgtgccccc gaggagaaaa tctgggagga gctgagtgtg ttagaggtgt

901 ttgaggggag ggaagacagt atcttggggg atcccaagaa gctgctcacc caacatttcg

961 tgcaggaaaa ctacctggag taccggcagg tccccggcag tgatcctgca tgttatgaat

1021 tcctgtgggg tccaagggcc ctcgttgaaa ccagctatgt gaaagtcctg caccatatgg

1081 taaagatcag tggaggacct cacatttcct acccaccccc gcatgagtgg gttttgagag

1141 agggggaaga gtgagtctga gcacgagttg cagccagggc cagtgggagg gggtctgggc

1201 cagtgcacct tccggggccg catcccttag tttccactgc ctcctgtgac gtgaggccca

1261 ttcttcactc tttgaagcga gcagtcagca ttcttagtag tgggtttctg ttctgttgga

1321 tgactttgag attattcttt gtttcctgtt ggagttgttc aaatgttcct tttaacggat

1381 ggttgaatga gcgtcagcat ccaggtttat gaatgacagt agtcacacat agtgctgttt

1441 atatagttta ggagtaagag tcttgttttt tactcaaatt gggaaatcca ttccattttg

1501 tgaattgtga cataataata gcagtggtaa aagtatttgc ttaaaattgt gagcgaatta

Figure 19

1561 gcaataacat acatgagata actcaagaaa tcaaaagata gttgattctt gccttgtacc

1621 tcaatctatt ctgtaaaatt aaacaaatat gcaaaccagg atttccttga cttctttgag

1681 aatgcaagcg aaattaaatc tgaataaata attcttcctc ttcaaaaaaa aaaaaaaaaa

1741 aaaaaaaaaa aaa (SEQ ID NO: 1)

Figure 19 (continued)

1 ttagagagaa gcgagctgct ctgaccagcc gcttgggatt ggcggaggga agcgggccag

61 gccctgtgag gagtcaaggt tctgagcaga caggccaacc ggaggacagg attccctgga

121 ggccacagag gagcaccaag gagaagatct gcctgtgggt ccccattgcc cagcttttgc

181 ctgcactctt gcctgctgcc ctgaccagag tcatcatgtc ttctgagcag aagagtcagc

241 actgcaagcc tgaggaaggc gttgaggccc aagaagaggc cctgggcctg gtgggtgcac

301 aggctcctac tactgaggag caggaggctg ctgtctcctc ctcctctcct ctggtccctg

361 gcaccctgga ggaagtgcct gctgctgagt cagcaggtcc tccccagagt cctcagggag

421 cctctgcctt acccactacc atcagcttca cttgctggag gcaacccaat gagggttcca

481 gcagccaaga agaggagggg ccaagcacct cgcctgacgc agagtccttg ttccgagaag

541 cactcagtaa caaggtggat gagttggctc attttctgct ccgcaagtat cgagccaagg

601 agctggtcac aaaggcagaa atgctggaga gagtcatcaa aaattacaag cgctgctttc

661 ctgtgatctt cggcaaagcc tccgagtccc tgaagatgat ctttggcatt gacgtgaagg

721 aagtggaccc cgccagcaac acctacaccc ttgtcacctg cctgggcctt tcctatgatg

781 gcctgctggg taataatcag atctttccca agacaggcct tctgataatc gtcctgggca

841 caattgcaat ggagggcgac agcgcctctg aggaggaaat ctgggaggag ctgggtgtga

901 tgggggtgta tgatgggagg gagcacactg tctatgggga gcccaggaaa ctgctcaccc

961 aagattgggt gcaggaaaac tacctggagt accggcaggt acccggcagt aatcctgcgc

1021 gctatgagtt cctgtggggt ccaagggctc tggctgaaac cagctatgtg aaagtcctgg

1081 agcatgtggt cagggtcaat gcaagagttc gcattgccta cccatccctg cgtgaagcag

1141 ctttgttaga ggaggaagag ggagtctgag catgagttgc agccagggct gtggggaagg

1201 ggcagggctg ggccagtgca tctaacagcc ctgtgcagca gcttcccttg cctcgtgtaa

1261 catgaggccc attcttcact ctgtttgaag aaaatagtca gtgttcttag tagtgggttt

1321 ctattttgtt ggatgacttg gagatttatc tctgtttcct tttacaattg ttgaaatgtt

Figure 20

1381 ccttttaatg gatggttgaa ttaacttcag catccaagtt tatgaatcgt agttaacgta

1441 tattgctgtt aatatagttt aggagtaaga gtcttgtttt ttattcagat tgggaaatcc

1501 gttctatttt gtgaatttgg gacataataa cagcagtgga gtaagtattt agaagtgtga

1561 attcaccgtg aaataggtga gataaattaa aagatactta attcccgcct tatgcctcag

1621 tctattctgt aaaatttaaa aaatatatat gcatacctgg atttccttgg cttcgtgaat

1681 gtaagagaaa ttaaatctga ataaataatt ctttctgtta a (SEQ ID NO: 2)

Figure 20 (continued)

1 agccgactga gatcagcaga ggggaatggg cccgggctct gtgaggaggc aaggttctca

61 ggggacaggc tgaccaggat caccaggaag ctccagagga tccccaggag gccctagagg

121 agcaccaaag gagaagatct gccagtgggt ctccattgcc cagctcctgc ccacactcct

181 gcctgttgcg gtgaccagag tcgtcatgtc tcttgagcag aagagtcagc actgcaagcc

241 tgaggaaggc cttgacaccc aagaagaggc cctgggcctg gtgggtgtgc aggctgccac

301 tactgaggag caggaggctg tgtcctcctc ctctcctctg gtcccaggca ccctgggggа

361 ggtgcctgct gctgggtcac caggtcctct caagagtcct cagggagcct ccgccatccc

421 cactgccatc gatttcactc tatggaggca atccattaag ggctccagca accaagaaga

481 ggaggggcca agcacctccc ctgacccaga gtctgtgttc cgagcagcac tcagtaagaa

541 ggtggctgac ttgattcatt ttctgctcct caagtattaa gtcaaggagc cggtcacaaa

601 ggcagaaatg ctggagagag tcatcaaaaa ttacaagcgc tgctttcctg tgatcttcgg

661 caaagcctcc gagtccttgc agctggtctt tggcattgac gtgaaggaag cggaccccac

721 cagcaacacc tacacccttg tcacctgcct gggactccta tgatggcctg gtggttgata

781 ataatcagat catgcccaag acgggcctcc tgataatcgt cttgggcatg attgcaatgg

841 agggcaaatg cgtccctgag gagaaaatct gggaggagct gagtgtgatg aaggtgtatg

901 ttgggaggga gcacagtgtc tgtggggagc ccaggaagct gctcacccaa gatttggtgc

961 aggaaaacta cctggagtac cggcaggtgc ccagcagtga tcccatatgc tatgagttac

1021 tgtggggtcc aagggcactc gctgcttgaa agtactggag cacgtggtca gggtcaatgc

1081 aagagttctc atttcctacc catccctgca tgaagcagct ttgagagagg aggaagaggg

1141 agtctgagca tgagctgcag ccagggccac tgcgaggggg gctgggccag tgcaccttcc

1201 agggctccgt ccactagttt cccctgcctt aatgtgacat gaggcccatt cttctctctt

1261 tgaagagagc agtcaacatt cttagtagtg ggtttctgtt ctattggatg actttgagat

1321 ttgtctttgt ttccttttgg aattgttcaa atgttccttt taatgggtgg ttgaatgaac

1381 ttcagcattc aaatttatga atgacagtag tcacacatag tgctgtttat atagtttagg

1441 agtaagagtc ttgtttttta ttcagattgg gaaatccatt ccattttgtg aattgggaca

Figure 21

1501 tagttacagc agtggaataa gtattcattt agaaatgtga atgagcagta aaactgatga

1561 cataaagaaa ttaaaagata tttaattctt gccttatact cagtctactc ggtaaaattt

1621 tttttaaaaa atgtgcatac ctggatttcc ttggcttctt tgagaatgta agacaaatta

1681 aatctgaata aa (SEQ ID NO: 3)

Figure 21 (continued)

1 agcaacgagc gacggcctga cgtcggcgga gggaagccgg cccaggctcg gtgaggaggc

61 aagtcctgca gcctcagcat gcgctggccg gatgtaccct gaggtgccct ctcacttcct

121 ccttcaggtt ctgaggggac aggctgacct ggaggaccag aggcccccgg aggagcactg

181 aaggagaaga tctgccagtg ggtctccatt gcccagctcc tgcccacact cccgcctgtt

241 gccctgacca gagtcatcat gcctcttgag cagaggagtc agcactgcaa gcctgaagaa

301 ggccttgagg cccgaggaga ggccctgggc ctggtgggtg cgcaggctcc tgctactgag

361 gagcaggagg ctgcctcctc ctcttctact ctagttgaag tcaccctggg ggaggtgcct

421 gctgccgagt caccagatcc tccccagagt cctcagggag cctccagcct ccccactacc

481 atgaactacc ctctctggag ccaatcctat gaggactcca gcaaccaaga agaggagggg

541 ccaagcacct tccctgacct ggagtctgag ttccaagcag cactcagtag gaaggtggcc

601 aagttggttc attttctgct cctcaagtat cgagccaggg agccggtcac aaaggcagaa

661 atgctgggga gtgtcgtcgg aaattggcag tacttctttc ctgtgatctt cagcaaagct

721 tccgattcct tgcagctggt ctttggcatc gagctgatgg aagtggaccc catcggccac

781 gtgtacatct ttgccacctg cctgggcctc tcctacgatg gcctgctggg tgacaatcag

841 atcatgccca agacaggctt cctgataatc atcctggcca taatcgcaaa agagggcgac

901 tgtgcccctg aggagaaaat ctgggaggag ctgagtgtgt tagaggtgtt tgaggggagg

961 gaagacagta tcttcgggga tcccaagaag ctgctcaccc aatatttcgt gcaggaaaac

1021 tacctggagt accggcaggt ccccggcagt gatcctgcat gctatgagtt cctgtggggt

1081 ccaagggccc tcattgaaac cagctatgtg aaagtcctgc accatatggt aaagatcagt

1141 ggaggacctc gcatttccta cccactcctg catgagtggg ctttgagaga gggggaagag

1201 tgagtctgag cacgagttgc agccagggcc agtgggaggg ggtttgggcc agtgcacctt

1261 ccggggcccc atcccttagt ttccactgcc tcctgtgacg tgaggcccat tcttcactct

1321 ttgaagcgag cagtcagcat tcttagtagt gggtttctgt tctgttggat gactttgaga

Figure 22

1381 ttattctttg tttcctgttg gagttgttca aatgttcctt ttaacggatg gttgaatgag

1441 cgtcagcatc caggtttatg aatgacagta gtcacacata gtgctgttta tatagtttag

1501 gagtaagagt cttgtttttt attcagattg ggaaatccat tccattttgt gaattgtgac

1561 ataataatag cagtggtaaa agtatttgct taaaattgtg agcgaattag caataacata

1621 catgagataa ctcaagaaat caaaagatag ttgattcttg ccttgtacct caatctattc

1681 tgtaaaatta aacaaatatg caaaccagga tttccttgac ttctttgaga atgcaagcga

1741 aattaaatct gaataaataa ttaaaaaaaa aaaaaaaaaa aaaaaaa (SEQ ID NO: 4)

Figure 22 (continued)

1 gctttgccgg atgtgctttc ccggcggcca tcttgggagt ctgaaggacc tgaggcattt

61 tgtgacgagg atcgtctcag gtcagcggag ggaggagact tatagaccta tccagtcttc

121 aaggtgctcc agaaagcagg agttgaagac ctgggtgtga gggacacata catcctaaaa

181 gcaccacagc agaggaggcc caggcagtgc caggagtcaa ggttcccaga agacaaaccc

241 cctaggaaga caggcgacct gtgaggccct agagcaccac cttaagagaa gaagagctgt

301 aagccggcct ttgtcagagc catcatgggg gacaaggata tgcctactgc tgggatgccg

361 agtcttctcc agagttcctc tgagagtcct cagagttgtc ctgaggggga ggactcccag

421 tctcctctcc agattcccca gagttctcct gagagcgacg acaccctgta tcctctccag

481 agtcctcaga gtcgttctga gggggaggac tcctcggatc ctctccagag acctcctgag

541 gggaaggact cccagtctcc tctccagatt ccccagagtt ctcctgaggg cgacgacacc

601 cagtctcctc tccagaattc tcagagttct cctgagggga aggactccct gtctcctcta

661 gagatttctc agagccctcc tgagggtgag gatgtccagt ctcctctgca gaatcctgcg

721 agttccttct tctcctctgc tttattgagt attttccaga gttcccctga gagtactcaa

781 agtccttttg agggttttcc ccagtctgtt ctccagattc ctgtgagcgc cgcctcctcc

841 tccactttag tgagtatttt ccagagttcc cctgagagta ctcaaagtcc ttttgagggt

901 tttccccagt ctccactcca gattcctgtg agccgctcct tctcctccac tttattgagt

961 attttccaga gttcccctga gagaactcag agtacttttg agggttttgc ccagtctcct

1021 ctccagattc ctgtgagccc ctcctcctcc tccactttac tgagtctttt ccagagtttc

1081 tctgagagaa ctcagagtac ttttgagggt tttgcccagt cttctctcca gattcctgtg

1141 agcccctcct tctcctccac tttagtgagt cttttccaga gttcccctga gagaactcag

1201 agtacttttg agggttttcc ccagtctcct ctccagattc ctgtgagctc ctcctcctcc

1261 tccactttat tgagtctttt ccagagttcc cctgagagaa ctcacagtac ttttgagggt

1321 tttccccagt ctcttctcca gattcctatg acctcctcct tctcctctac tttattgagt

1381 attttccaga gttctcctga gagtgctcaa agtacttttg agggttttcc ccagtctcct

1441 ctccagattc ctgggagccc ctccttctcc tccactttac tgagtctttt ccagagttcc

1501 cctgagagaa ctcacagtac ttttgagggt tttccccagt ctcctctcca gattcctatg

Figure 23

1561 acctcctcct tctcctctac tttattgagt attttacaga gttctcctga gagtgctcaa

1621 agtgcttttg agggttttcc ccagtctcct ctccagattc ctgtgagctc ctctttctcc

1681 tacactttat tgagtctttt ccagagttcc cctgagagaa ctcacagtac ttttgagggt

1741 tttccccagt ctcctctcca gattcctgtg agctcctcct cctcctcctc cactttattg

1801 agtcttttcc agagttcccc tgagtgtact caaagtactt ttgagggttt tccccagtct

1861 cctctccaga ttcctcagag tcctcctgaa ggggagaata cccattctcc tctccagatt

1921 gttccaagtc ttcctgagtg ggaggactcc ctgtctcctc actactttcc tcagagccct

1981 cctcaggggg aggactccct atctcctcac tactttcctc agagccctcc tcagggggag

2041 gactccctgt ctcctcacta ctttcctcag agccctcagg gggaggactc cctgtctcct

2101 cactactttc ctcagagccc tcctcagggg gaggactcca tgtctcctct ctactttcct

2161 cagagtcctc ttcaggggga ggaattccag tcttctctcc agagccctgt gagcatctgc

2221 tcctcctcca ctccatccag tcttccccag agtttccctg agagttctca gagtcctcct

2281 gaggggcctg tccagtctcc tctccatagt cctcagagcc ctcctgaggg gatgcactcc

2341 caatctcctc tccagagtcc tgagagtgct cctgaggggg aggattccct gtctcctctc

2401 caaattcctc agagtcctct tgagggagag gactccctgt cttctctcca ttttcctcag

2461 agtcctcctg agtgggagga ctccctctct cctctccact ttcctcagtt tcctcctcag

2521 ggggaggact tccagtcttc tctccagagt cctgtgagta tctgctcctc ctccacttct

2581 ttgagtcttc cccagagttt ccctgagagt cctcagagtc ctcctgaggg gcctgctcag

2641 tctcctctcc agagacctgt cagctccttc ttctcctaca ctttagcgag tcttctccaa

2701 agttcccatg agagtcctca gagtcctcct gaggggcctg cccagtctcc tctccagagt

2761 cctgtgagct ccttcccctc ctccacttca tcgagtcttt cccagagttc tcctgtgagc

2821 tccttcccct cctccacttc atcgagtctt tccaagagtt cccctgagag tcctctccag

2881 agtcctgtga tctccttctc ctcctccact tcattgagcc cattcagtga agagtccagc

2941 agcccagtag atgaatatac aagttcctca gacaccttgc tagagagtga ttccttgaca

3001 gacagcgagt ccttgataga gagcgagccc ttgttcactt atacactgga tgaaaaggtg

Figure 23 (con't)

3061 gacgagttgg cgcggtttct tctcctcaaa tatcaagtga agcagcctat cacaaaggca

3121 gagatgctga cgaatgtcat cagcaggtac acgggctact ttcctgtgat cttcaggaaa

3181 gcccgtgagt tcatagagat acttttggc atttccctga gagaagtgga ccctgatgac

3241 tcctatgtct ttgtaaacac attagacctc acctctgagg ggtgtctgag tgatgagcag

3301 ggcatgtccc agaaccgcct cctgattctt attctgagta tcatcttcat aaagggcacc

3361 tatgcctctg aggaggtcat ctgggatgtg ctgagtggaa taggggtgcg tgctgggagg

3421 gagcactttg cctttgggga gcccagggag ctcctcacta aagtttgggt gcaggaacat

3481 tacctagagt accgggaggt gcccaactct tctcctcctc gttacgaatt cctgtggggt

3541 ccaagagctc attcagaagt cattaagagg aaagtagtag agtttttggc catgctaaag

3601 aataccgtcc ctattacctt tccatcctct tacaaggatg ctttgaaaga tgtggaagag

3661 agagcccagg ccataattga caccacagat gattcgactg ccacagaaag tgcaagctcc

3721 agtgtcatgt cccccagctt ctcttctgag tgaagtctag ggcagattct tccctctgag

3781 tttgaagggg gcagtcgagt ttctacgtgg tggagggcct ggttgaggct ggagagaaca

3841 cagtgctatt tgcatttctg ttccatatgg gtagttatgg ggtttacctg ttttacttttt

3901 gggtattttt caaatgcttt tcctattaat aacaggttta aatagcttca gaatcctagt

3961 ttatgcacat gagtcgcaca tgtattgctg tttttctggt ttaagagtaa cagtttgata

4021 ttttgtaaaa acaaaaacac acccaaacac accacattgg gaaaaccttc tgcctcattt

4081 tgtgatgtgt cacaggttaa tgtggtgtta ctgtaggaat tttcttgaaa ctgtgaagga

4141 actctgcagt taaatagtgg aataaagtaa aggattgtta atgtttgcat ttcctcaggt

4201 cctttagtct gttgttcttg aaaactaaag atacatacct ggtttgcttg gcttacgtaa

4261 gaaagtagaa gaaagtaaac tgtaataaat aaaagtgtca gtgactcatt tatttgatga

4321 aaaaaaaaaa aaaaaaa (SEQ ID NO: 4)

Figure 23 (con't)

1 ccagctggca gtcaaggctg taggagggca tggagagttg aagaaaaaag cagtatcttg

61 aggcagactg gaagagtcat cacagcatcc aaatcaacaa gaaaacatca ttccagggtc

121 ctacatgatg gcgtactctg atactacaat gatgtctgat gatattgact ggttacgcag

181 ccacaggggt gtgtgcaagg tagatctcta caacccagaa ggacagcaag atcaggaccg

241 gaaagtgata tgctttgtcg atgtgtccac cctgaatgta gaagataaag attacaagga

301 tgctgctagt tccagctcag aaggcaactt aaacctggga agtctggaag aaaaagagat

361 tatcgtgatc aaggacactg agaagaaaga ccagtctaag acagagggat ctgtatgcct

421 tttcaaacaa gctccctctg atcctgtaag tgtcctcaac tggcttctca gtgatctcca

481 gaagtatgcc ttgggtttcc aacatgcact gagcccctca acctctacct gtaaacataa

541 agtaggagac acagagggcg aatatcacag agcatcctct gagaactgct acagtgtcta

601 tgccgatcaa gtgaacatag attatttgat gaacagacct caaaacctac gtctagaaat

661 gacagcagct aaaaacacca acaataatca aagtccttca gctcctccag ccaaacctcc

721 tagcactcag agagcagtca tttcccctga tggagaatgt tctatagatg acctttcctt

781 ctacgtcaac cgactatctt ctctggtaat ccagatggcc cataaggaaa tcaaggagaa

841 gttggaaggt aaaagcaaat gccttcatca ttcaatctgt ccatcccctg ggaacaaaga

901 gagaatcagt ccccgaactc ctgcgagcaa gattgcttct gaaatggcct atgaagctgt

961 ggaactgaca gctgcagaaa tgcgtggcac tggagaggag tccagggaag gtggccagaa

1021 aagctttcta tatagcgaat tatccaacaa gagcaaaagt ggagacaaac agatgtccca

1081 gagagagagc aaagaatttg cagattccat cagcaagggg ctcatggttt atgcaaatca

1141 ggtggcatct gacatgatgg tctctctcat gaagaccttg aaagtgcaca gctctgggaa

1201 gccaattcca gcatctgtgg tcctgaagag ggtgttgcta aggcacacca aggagattgt

1261 gtccgatttg attgattctt gcatgaagaa cctgcataat attactgggg tcctgatgac

1321 tgactcagac tttgtctcag ctgtcaagag aaatctgttc aaccagtgga aacaaaatgc

1381 tacagacatc atggaggcca tgctgaagcg cttggtcagt gcccttatag gtgaggagaa

1441 ggagactaag tctcagagtc tgtcatatgc atctttaaaa gctgggtccc atgatcccaa

1501 atgcaggaat cagagtcttg aattctccac catgaaagct gaaatgaaag agagggacaa

Figure 24

1561 aggcaaaatg aaatcagacc catgcaagtc actgactagt gctgagaaag tcggtgaaca

1621 cattctcaaa gagggcctaa ccatctggaa ccaaaagcaa ggaaactcat gcaaggtggc

1681 taccaaagca tgcagcaata aagatgagaa aggagaaaag atcaatgctt ccacagattc

1741 actggccaag gacctgattg tctctgccct taagctgatc cagtaccatc tgacccagca

1801 gactaagggc aaagatacat gtgaagaaga ctgtcctggt tccaccatgg gctatatggc

1861 tcagagtact caatatgaaa agtgtggagg tggccaaagt gccaaagcac tttcagtgaa

1921 acaactagaa tctcacagag cccctggacc atccacctgt caaaaggaga accaacacct

1981 ggactcccag aaaatggata tgtcaaacat cgttctaatg ctgattcaga aactgcttaa

2041 tgagaacccc ttcaaatgtg aggatccatg cgaaggtgag aacaagtgtt ctgagcccag

2101 ggcaagcaaa gcagcttcca tgtccaacag atctgacaaa gcggaagaac aatgccagga

2161 gcatcaagaa cttgactgta ccagtgggat gaagcaagcg aacgggcaat ttatagataa

2221 actagtagaa tctgtgatga agctctgcct tatcatggct aagtatagca acgatggggc

2281 agcccttgct gagttggaag aacaagcagc ctcggcaaat aagcccaatt tcaggggcac

2341 cagatgcatt cacagtggtg caatgccaca gaactatcaa gactctcttg gacatgaagt

2401 aattgtcaat aatcagtgct ctacaaatag cttgcagaag cagctccagg ctgtcctgca

2461 gtggattgca gcctcccagt ttaacgtgcc catgctctac ttcatgggag ataaggatgg

2521 acaactggaa aagcttcctc aggtttcagc taaagcagca gagaaggggt acagtgtagg

2581 aggtcttctt caagaggtca tgaagtttgc caaggaacgg caaccagatg aagctgtggg

2641 aaaggtggcc aggaaacagt tgctggactg gctgctcgct aacctgtgag ctgatccttg

2701 actcctcttc atcttagccc ccctagcagc attccatccc agccagagca cccccaccat

2761 caggccagtc aactgcacaa tacacaactg tatttcccaa tacacttgag cagttgcctg

2821 tgaatgtaag aggtgtcaac aaactgggaa ataaaataaa aaaaaataat aataaatgtg

2881 t (SEQ ID NO: 6)

Figure 24 (con't)

1 atcctcgtgg gccctgacct tctctctgag agccgggcag aggctccgga gccatgcagg

61 ccgaaggccg gggcacaggg ggttcgacgg gcgatgctga tggcccagga ggccctggca

121 ttcctgatgg cccaggggggc aatgctggcg gcccaggaga ggcgggtgcc acgggcggca

181 gaggtccccg gggcgcaggg gcagcaaggg cctcggggcc gggaggaggc gccccgcggg

241 gtccgcatgg cggcgcggct tcagggctga atggatgctg cagatgcggg gccaggggggc

301 cggagagccg cctgcttgag ttctacctcg ccatgccttt cgcgacaccc atggaagcag

361 agctggcccg caggagcctg gcccaggatg ccccaccgct tcccgtgcca ggggtgcttc

421 tgaaggagtt cactgtgtcc ggcaacatac tgactatccg actgactgct gcagaccacc

481 gccaactgca gctctccatc agctcctgtc tccagcagct ttccctgttg atgtggatca

541 cgcagtgctt tctgcccgtg tttttggctc agcctccctc agggcagagg cgctaagccc

601 agcctggcgc cccttcctag gtcatgcctc ctcccctagg gaatggtccc agcacgagtg

661 gccagttcat tgtgggggcc tgattgtttg tcgctggagg aggacggctt acatgtttgt

721 ttctgtagaa aataaaactg agctacgaaa aa (SEQ ID NO: 7)

Figure 25

1 gccctggcaa ggttgtgggg gacatcttga gctgaagcag ggttttgagc cactgctgct

61 gctgccattg tcaccatggt ctcagctctg cggggagcac ccctgatcag ggtgcactca

121 agccctgttt cttctccttc tgtgagtgga ccacggaggc tggtgagctg cctgtcatcc

181 caaagctcag ctctgagcca gagtggtggt ggctccacct ctgccgccgg catagaagcc

241 aggagcaggg ctctcagaag gcggtggtgc ccagctggga tcatgttgtt ggccctggtc

301 tgtctgctca gctgcctgct accctccagt gaggccaagc tctacggtcg ttgtgaactg

361 gccagagtgc tacatgactt cgggctggac ggataccggg gatacagcct ggctgactgg

421 gtctgccttg cttatttcac aagcggtttc aacgcagctg ctttggacta cgaggctgat

481 gggagcacca acaacgggat cttccagatc aacagccgga ggtggtgcag caacctcacc

541 ccgaacgtcc ccaacgtgtg ccggatgtac tgctcagatt tgttgaatcc taatctcaag

601 gataccgtta tctgtgccat gaagataacc caagagcctc agggtctggg ttactgggag

661 gcctggaggc atcactgcca gggaaaagac ctcactgaat gggtggatgg ctgtgacttc

721 taggatggac ggaaccatgc acagcaggct gggaaatgtg gtttggttcc tgacctaggc

781 ttgggaagac aagccagcga ataaaggatg gttgaacgt (SEQ ID NO: 8)

Figure 26

1 tcgccccttc tcggccgccg tagtttttt ttttttttaa agaaaaaacg gttacccagc

61 aactagaaaa acaaccggaa ccggcggcac cagctcggag agaaaggagg ttccataggc

121 agttcttacc aagaagatgt cgattccatt ctccaacacc cactaccgaa ttccacaagg

181 atttgggaat cttcttgaag ggctgacacg cgagattctg agagagcaac cggacaatat

241 accagctttt gcagcagcct attttgagag ccttctagag aaaagagaga aaaccaactt

301 tgatccagca gaatggggga gtaaggtaga agaccgcttc tataacaatc atgcattcga

361 ggagcaagaa ccacctgaga aaagtgatcc taaacaagaa gagtctcaga tatctgggaa

421 ggaggaagag acatcagtca ccatcttaga ctcttctgag gaagataagg aaaaagaaga

481 ggttgctgct gtcaaaatcc aagctgcctt ccggggacac atagccagag aggaggcaaa

541 gaaaatgaaa acaaatagtc ttcaaaatga ggaaaaagag gaaaacaagt gaggacactg

601 gttttacctc caggaaacat gaaaaataat ccaaatccat caaccttctt attaatgtca

661 tttcttcctg aggaaggaag atttgatgtt gtgaaataac attcgttact gttgtgaaaa

721 tctgtcatga gcatttgttt aataagcata ccattgaaac atgccacttg aagatttctc

781 tgagatcatg agtttgttta cacttgtctc aagcctatct atagagaccc ttggatttag

841 aattatagaa ctaaagtatc tgagattaca gagatctcag aggttatgtg ttctaactat

901 tatcaaatga ataaatcctc tctatcacat cccccaaaaa aaaaaaaaaa aaaaa (SEQ ID NO: 9)

Figure 27

1 gtcaccagga gggtatgcat agggagggca agagctctgg gccactgcga agattcaaaa

61 gctccaaaaa cctactgtag acatcgaaga accaatatat acaatgggcc aacaatccag

121 tgtccgcagg ctgaagagga gcgtcccctg tgaatccaac gaggccaacg aggccaatga

181 ggccaacaag acgatgccgg agaccccaac tggggactca gacccgcaac ctgctcctaa

241 aaaaatgaaa acatctgagt cctcgaccat actagtggtt cgctacagga ggaacgtgaa

301 aagaacatct ccagaggaac tggtgaatga ccacgcccga gagaacagaa tcaaccccga

361 ccaaatggag gaggaggaat tcatagaaat aacgactgaa agacctaaaa agtagcaaga

421 agctacatcc ctcaaacttc ggcaatgaaa ataaagtttg agaagctgaa aa (SEQ ID NO: 10)

Figure 28

1 atgcctgggt cagagccctg gggaggaact gctcagttag gacccagacg gaaccatgga

61 agccccagca cagcttctct tcctcctgct actctggctc ccagatacca cggagaaatt

121 gtaatgacac agtctccagc caccctgtct ttgtctccag gggaaagagc caccctctcc

181 tgcagggcca gtcagagtgt tagcagcagc tacttatcct ggtaccagca gaaacctggg

241 caggctccca ggctcctcat ctatggtgca tccaccaggg ccactggcat cccagccagg

301 ttcagtggca gtgggtctgg gacagacttc actctcacca tcagcagcct gcagcctgaa

361 gattttgcag tttattactg tcagcaggat tataacttac ctcccacagt gattcaacat

421 gaaacaaaaa cctcaacaag accatcagtg tttactagat tttaccagct gcttcctttа

481 cagacagcta atgtggtggc cactcagttt tag (SEQ ID NO: 11)

Figure 29

1 gtcaccagga gggtatgcat agggagggca agagctctgg gccactgcga agattcaaaa

61 gctccaaaaa cctactgtag acatcgaaga accaatatat acaatgggcc aacaatccag

121 tgtccgcagg ctgaagagga gcgtcccctg tgaatccaac gaggccaacg aggccaatga

181 ggccaacaag acgatgccgg agaccccaac tggggactca gacccgcaac ctgctcctaa

241 aaaaatgaaa acatctgagt cctcgaccat actagtggtt cgctacagga ggaacgtgaa

301 aagaacatct ccagaggaac tggtgaatga ccacgcccga gagaacagaa tcaaccccga

361 ccaaatggag gaggaggaat tcatagaaat aacgactgaa agacctaaaa agtagcaaga

421 agctacatcc ctcaaacttc ggcaatgaaa ataaagtttg agaagctgaa aa (SEQ ID NO: 10)

Figure 28

1 atgcctgggt cagagccctg gggaggaact gctcagttag gacccagacg gaaccatgga

61 agccccagca cagcttctct tcctcctgct actctggctc ccagatacca cggagaaatt

121 gtaatgacac agtctccagc caccctgtct ttgtctccag gggaaagagc caccctctcc

181 tgcagggcca gtcagagtgt tagcagcagc tacttatcct ggtaccagca gaaacctggg

241 caggctccca ggctcctcat ctatggtgca tccaccaggg ccactggcat cccagccagg

301 ttcagtggca gtgggtctgg gacagacttc actctcacca tcagcagcct gcagcctgaa

361 gattttgcag tttattactg tcagcaggat tataacttac ctcccacagt gattcaacat

421 gaaacaaaaa cctcaacaag accatcagtg tttactagat tttaccagct gcttcctttа

481 cagacagcta atgtggtggc cactcagttt tag (SEQ ID NO: 11)

Figure 29

1 agccggctgt ggtgggaggc agttgagccc tggattgtga ccgcttcagg gcagttggta

61 gatgcccctc tgggagagat ccccaggggt gacagccatg gaccctggaa gggcctgggc

121 tagggacagg gaccagagcc agtccaggga gaggacagag ccaatggact ggggtgtact

181 gtaacagccc tgctggcgag agggaccagg gcaccgtcct ccagggagcc catgctgcaa

241 gtcgggccag aggtgcccct gaacctgaag gccaatgaga cccaagacag gccaagtggg

301 ttgtgagacc cctgaggagc tgggccctgg tcccaggcag cgctggcccc tgctgctgct

361 gggtctggcc atggtcgccc atggcctgct gcgcccaatg gttgcaccgc aaagcgggga

421 cccagaccct ggagcctcag ttggaagcag ccgatccagc ctgcggagcc tgtggggcag

481 gctcctgctc cagcccagcc cccagagagc agaccccagg tgctggcccc gggggttttg

541 gtctgagcct cagtcactgt gttatgtctt cggaactggg accaaggtca ccgtcctagg

601 tcagcccaag gccaacccca ctgtcactct gttcccgccc tcctctgagg agctccaagc

661 caacaaggcc acactagtgt gtctgatcag tgacttctac ccgggagctg tgacagtggc

721 ctggaaggca gatggcagcc ccgtcaaggc gggagtggag accaccaaac cctccaaaca

781 gagcaacaac aagtacgcgg ccagcagcta cctgagcctg acgcccgagc agtggaagtc

841 ccacagaagc tacagctgcc aggtcacgca tgaagggagc accgtggaga agacagtggc

901 ccctacagaa tgttcatagg ttcccaactc taaccccacc cacgggagcc tggagctgca

961 ggatcccagg ggaggggtct ctctccccat cccaagtcat ccagcccttc tccctgcact

1021 catgaaaccc caataaatat cctcattgac aaccagaaaa aaaaaaaaaa aaaa (SEQ ID NO: 12)

Figure 30

1 gcctctctcc atcagacacc ccaaggttcc atccgaagca ggcggagcac cgaacgcacc

61 ccggggtggt cagggacccc catccgtgct gccccctagg agcccgcgcc tctcctctgc

121 gccccgcctc tcgggccgca acgtcgcgcg gttcctttaa cagcgcgctg gcagggtgtg

181 ggaagcagga ccgcgtcctc ccgccccctc ccatccgagt ttcaggtgaa ttggtcaccg

241 agggaggagg ccgacacacc acacctacac tcccgcgtcc acctctccct ccctgcttcc

301 tctggcggag gcggcaggaa ccgagagcca ggtccagagc gccgaggagc cggtctagga

361 cgcagcagat tggtttatct tggaagctaa agggcattgc tcatcctgaa gatcagctga

421 ccattgacaa tcagccatgt catccaggcc tcttgaaagt ccacctcctt acaggcctga

481 tgaattcaaa ccgaatcatt atgcaccaag caatgacata tatggtggag agatgcatgt

541 tcgaccaatg ctctctcagc cagcctactc tttttaccca gaagatgaaa ttcttcactt

601 ctacaaatgg acctctcctc caggagtgat tcggatcctg tctatgctca ttattgtgat

661 gtgcattgcc atctttgcct gtgtggcctc cacgcttgcc tgggacagag gctatggaac

721 ttccctttta ggaggtagtg taggctaccc ttatggagga agtggctttg gtagctacgg

781 aagtggctat ggctatggct atggttatgg ctatggctac ggaggctata cagacccaag

841 agcagcaaag ggcttcatgt tggccatggc tgccttttgt ttcattgccg cgttggtgat

901 ctttgttacc agtgttataa gatctgaaat gtccagaaca agaagatact acttaagtgt

961 gataatagtg agtgctatcc tgggcatcat ggtgtttatt gccacaattg tctatataat

1021 gggagtgaac ccaactgctc agtcttctgg atctctatat ggttcacaaa tatatgccct

1081 ctgcaaccaa ttttatacac ctgcagctac tggactctac gtggatcagt atttgtatca

1141 ctactgtgtt gtggatcccc aggaggccat tgccattgta ctggggttca tgattattgt

1201 ggcttttgct ttaataattt tctttgctgt gaaaactcga agaaagatgg acaggtatga

1261 caagtccaat attttgtggg acaaggaaca catttatgat gagcagcccc ccaatgtcga

1321 ggagtgggtt aaaaatgtgt ctgcaggcac acaggacgtg ccttcacccc catctgacta

1381 tgtggaaaga gttgacagtc ccatggcata ctcttccaat ggcaaagtga atgacaagcg

1441 gttttatcca gagtcttcct ataaatccac gccggttcct gaagtggttc aggagcttcc

Figure 31

1501 attaacttcg cctgtggatg acttcaggca gcctcgttac agcagcggtg gtaactttga

1561 gacaccttca aaaagagcac ctgcaaaggg aagagcagga aggtcaaaga gaacagagca

1621 agatcactat gagacagact acacaactgg cggcgagtcc tgtgatgagc tggaggagga

1681 ctggatcagg gaatatccac ctatcacttc agatcaacaa agacaactgt acaagaggaa

1741 ttttgacact ggcctacagg aatacaagag cttacaatca gaacttgatg agatcaataa

1801 agaactctcc cgtttggata aagaattgga tgactataga gaagaaagtg aagagtacat

1861 ggctgctgct gatgaataca atagactgaa gcaagtgaag ggatctgcag attacaaaag

1921 taagaagaat cattgcaagc agttaaagag caaattgtca cacatcaaga agatggttgg

1981 agactatgat agacagaaaa catagaaggc tgatgccaag ttgtttgaga aattaagtat

2041 ctgacatctc tgcaatcttc tcagaaggca aatgactttg gaccataacc ccggaagcca

2101 aacctctgtg agcatcacaa agttttggtt gctttaacat catcagtatt gaagcatttt

2161 ataaatcgct tttgataatc aactgggctg aacactccaa ttaaggattt tatgctttaa

2221 acattggttc ttgtattaag aatgaaatac tgtttgaggt ttttaagcct taaaggaagg

2281 ttctggtgtg aactaaactt tcacacccca gacgatgtct tcatacctac atgtatttgt

2341 ttgcataggt gatctcattt aatcctctca accacctttc agataactgt tatttataat

2401 cacttttttc cacataagga aactgggttc ctgcaatgaa gtctctgaag tgaaactgct

2461 tgtttcctag cacacacttt tggttaagtc tgttttatga cttcattaat aataaattcc

2521 ctggcctttc atattttagc tactatatat gtgatgatct accagcctcc ctatttttt

2581 tctgttatat aaatggttaa aagaggtttt tcttaaataa taaagatcat gtaaaagtaa

2641 caaatgtgtg aaatttaaag attgtaaata tatatttact tttttaagat caaagtttaa

2701 accccgtggt tagaattttg tgtgttttta aatacttttt atctttttgc atgccttttt

2761 taaaaaacca actagaactt ttcattatat cagaatatct gattacattt ataattcaat

2821 tgtgacttga actgtatctt acaggaatgt tcaatttcta tacatatttt ataaggtatt

2881 aaacctggtg ttttctttcc ataataacct gtttgatgtt attagtgctg ttaacataca

2941 gcaatggaaa accacactca ggagttgtat ctgttgttgt ttatactcct ttggatgctg

Figure 31 (con't)

3001 tgctggttag tcgtttccca ttcctttggc tgtaagaatg ctgatatgtc tgggaataga

3061 atgctatacc acgaaatacc aaataatttc aaatggtgcc cttaaattgt atcacttttt

3121 taaaaattca gattcttatt agtaaaatta gttgatagca ctgtgctgac caagttgatt

3181 gtgatcatcc cagcttagac ttttctaaaa actttttttt agaataatct ataaactgaa

3241 ctttagtatg catttcagat atttaggtat ataatttttt ttttttttg agacagagtc

3301 tcactctcac ccaggctgga atgcagtggt gctatcttgg ctcactgcaa cctccacctc

3361 ccgggttcaa gcaattctcc tgcctcagcc tctcgagtag ttgagactac aggtgcccat

3421 caccatgcgt ggctaatttt tgtattttta atagagacgg ggttttacca tagtggccag

3481 gttggtcttg aactcctgac cttgtggtct gcctgcctcg gcctcccaaa gtgctgggat

3541 tacaggcgtg agccaccatg cctggcctaa gtgtgtgtgt gtgtgtgtgt gtgtattttt

3601 ttttttttt ttttgagatg gagttttgct cttgttgaac aggctggagt gcaatgtcgc

3661 gatctcagct caccacaacc tccgcctccc aggttcaaac aattctcctg cctcggcctc

3721 ccgagtagct gggattacag gcatgcgcca ccacacctgg ctaatttttt ttttgtattt

3781 ttagtagaga tggggtttct ccatgttggt caggctggtc tcgaactcct gacctcaggt

3841 gatccatcca cctcggcctc ccaaagtgct gggattagag gcgtgagcca ctgtgcccgg

3901 cctataattt ttgatagatg attttgaatt attttccaga gataaaattt taaatgtttc

3961 cattatatca ctgatttatt tctgcaaatt gaataaattc ttaattttct gcatgcacat

4021 aatacaaaag gtattttcat agttttggat ttataccaaa tgaaaaggat tctcttgatg

4081 agcaccttta actgattttt ctgttaaagt tttaacaatt tgttcttgga agtcagttcg

4141 tgaaggcaag tttgtcagta ttttcacaaa actattcagc tgaatccaga aagtgaaaca

4201 gcaagaattt gcattgtaaa attgtgttat aaaattggac tttgaaattt caaaaataag

4261 aaaaattttc atgtgtattt atactaaata ccgttttagg aaactaggat cagggtgttt

4321 ctgttggcgt tggcattaac tagctggatg taaatttgaa aagccactca agcagcttcc

4381 tagtctagaa agtcagaggt ttagattaga tttccgacat cccttccatt tctgacctgt

Figure 31 (con't)

4441 agttcttgtc tggaattctg ctttgttata aactattgtt ctaaggagtt tgttgtgata

4501 gcacatagtt cattttgtaa agattccctg cgtataaagt gatgccctac atatgtgatt

4561 ttgtattaaa agtatatagg atcattattt tattttgaaa aatttaaata cagaaaagta

4621 taaaatataa gtaccatccg cccagaaata acatgtgtta atgttttgtc atatgtgctt

4681 tatatttttt gaaataaagt gaagtcaact agtatttata gtaaataagt tacatacaca

4741 taagtacata tatgatattt aatcctcaca acgatctttt gacatgtgac catttcttat

4801 tcttctttta tagacaagga actaatgata tgatagatta actggctgtt gtcacactag

4861 caagtggcaa aacaagggat taggatctta gtctcttcaa ctgttagatt ctatacttcc

4921 atcctgtgtt gactttgtta atggattgga taatgtgaga tcactctgat gtaaataaag

4981 tatcctatat taatttcgag tgcattttaa gtacttgtaa cataaatgct tcctgtgaaa

5041 tatctgtaaa gacctgaatg ggtacatgtg tgtaaagaag aatcagggca gaaaagtgct

5101 tttatcatgg ctccggggac cttagcttca gttggtgttg tgagaattcc tcacacaagg

5161 acattctcct tgcttcagca tcaggatgga agtgtttctc atctggactt tttcaaagac

5221 tcagctggag gaatcagaat tcataatttc ctggcagctc atgattctgc tacactacac

5281 catgccatct cttgtgtgaa aggacagatt tgatggagga ctatgtcatc cctcatgcgt

5341 ttcttattgt ctacatttat tctaatggga agaagtgagc aaaaacacct caataatttg

5401 ggtagttttt agaaaacctt gttagtaaat tagaatagtg ccactttggc attatgagaa

5461 agaagcatgg atacataact agggttttgt gtatgactac aacgaaatgc agaatggtgt

5521 ctccaaaagg tttccaattg ctgccacaag aactgcttgg tattgcctac atgtgttgtc

5581 ctatttttgc tttgcccttc tgcagttact tgctgtggga ccttggagaa attaacttag

5641 cctctctgta cttcagtttt ttgtatttgt aaaatatatt tgtaataatc tcatagttaa

5701 gaaggtagtt aatgtgtgac tcagtccttg tctaaaagta aatatgccta gctaccccca

5761 tcttccaaag ccagaaggtg aaactttaac aagttttcta aaagcaaatt gtgtttttta

5821 aaagtgcatg tgtcatccaa tcccatatga ttgatctgtg ctgggtgcag ccttagaatg

5881 taaattcttt tgaattctag gcagagaatg caggattggc attctaaata tttgtacatg

Figure 31 (con't)

5941 ataaacaaat gcttctttag gttagagcaa atagtttact tatcaagatc acaattgtta

6001 gatactgttg tcaattacag aggttttaga tgaggctttc tggaatgatt tagtttccct

6061 gtaagggagc ctgtctattg gaatagacag gttcacttct cccagtcttt caagttgcat

6121 gctttttata tctgattcca ctggctgagc tgattgtgaa tgtcctaacc ctgttgattg

6181 tgtctggcca ctcatgggca aagaacagat tatccattct ttatagttgt cttttagttt

6241 tacaagttga aaaaacatct gagtaggtta gataatttat tctaccactt tgtaaatgat

6301 tagaatatgt cagtcataat catgccaaga gattatggat ttatgcatat tttgttttgc

6361 tgtagtacca ttcctagttg aatcttaaca tccatgtcta aaatctatac agaacaaata

6421 ttacagttgg gaaaactgaa aaaaaaaaaa a (SEQ ID NO: 13)

Figure 31 (con't)

1 gctgagcgcg gagccgcccg gtgattggtg ggggcggaag ggggccgggc gccagcgctg

61 ccttttctcc tgccgggtag tttcgctttc ctgcgcagag tctgcggagg ggctcggctg

121 caccgggggg atcgcgcctg gcagacccca gaccgagcag aggcgaccca gcgcgctcgg

181 gagaggctgc accgccgcgc ccccgcctag cccttccgga tcctgcgcgc agaaaagttt

241 catttgctgt atgccatcct cgagagctgt ctaggttaac gttcgcactc tgtgtatata

301 acctcgacag tcttggcacc taacgtgctg tgcgtagctg ctcctttggt tgaatcccca

361 ggcccttgtt ggggcacaag gtggcaggat gtctcagtgg tacgaacttc agcagcttga

421 ctcaaaattc ctggagcagg ttcaccagct ttatgatgac agttttccca tggaaatcag

481 acagtacctg gcacagtggt tagaaaagca agactgggag cacgctgcca atgatgtttc

541 atttgccacc atccgttttc atgacctcct gtcacagctg gatgatcaat atagtcgctt

601 ttctttggag aataacttct tgctacagca taacataagg aaaagcaagc gtaatcttca

661 ggataatttt caggaagacc caatccagat gtctatgatc atttacagct gtctgaagga

721 agaaaggaaa attctggaaa acgcccagag atttaatcag gctcagtcgg ggaatattca

781 gagcacagtg atgttagaca aacagaaaga gcttgacagt aaagtcagaa atgtgaagga

841 caaggttatg tgtatagagc atgaaatcaa gagcctggaa gatttacaag atgaatatga

901 cttcaaatgc aaaaccttgc agaacagaga acacgagacc aatggtgtgg caaagagtga

961 tcagaaacaa gaacagctgt tactcaagaa gatgtattta atgcttgaca ataagagaaa

1021 ggaagtagtt cacaaaataa tagagttgct gaatgtcact gaacttaccc agaatgccct

1081 gattaatgat gaactagtgg agtggaagcg gagacagcag agcgcctgta ttgggggggcc

1141 gcccaatgct tgcttggatc agctgcagaa ctggttcact atagttgcgg agagtctgca

1201 gcaagttcgg cagcagctta aaaagttgga ggaattggaa cagaaataca cctacgaaca

1261 tgaccctatc acaaaaaaca aacaagtgtt atgggaccgc accttcagtc ttttccagca

1321 gctcattcag agctcgtttg tggtggaaag acagccctgc atgccaacgc accctcagag

1381 gccgctggtc ttgaagacag gggtccagtt cactgtgaag ttgagactgt tggtgaaatt

1441 gcaagagctg aattataatt tgaaagtcaa agtcttattt gataaagatg tgaatgagag

Figure 32

1501 aaatacagta aaaggattta ggaagttcaa cattttgggc acgcacacaa aagtgatgaa

1561 catggaggag tccaccaatg gcagtctggc ggctgaattt cggcacctgc aattgaaaga

1621 acagaaaaat gctggcacca gaacgaatga gggtcctctc atcgttactg aagagcttca

1681 ctcccttagt tttgaaaccc aattgtgcca gcctggtttg gtaattgacc tcgagacgac

1741 ctctctgccc gttgtggtga tctccaacgt cagccagctc ccgagcggtt gggcctccat

1801 cctttggtac aacatgctgg tggcggaacc caggaatctg tccttcttcc tgactccacc

1861 atgtgcacga tgggctcagc tttcagaagt gctgagttgg cagttttctt ctgtcaccaa

1921 aagaggtctc aatgtggacc agctgaacat gttgggagag aagcttcttg gtcctaacgc

1981 cagccccgat ggtctcattc cgtggacgag gttttgtaag gaaaatataa atgataaaaa

2041 ttttcccttc tggctttgga ttgaaagcat cctagaactc attaaaaaac acctgctccc

2101 tctctggaat gatgggtgca tcatgggctt catcagcaag gagcgagagc gtgccctgtt

2161 gaaggaccag cagccgggga ccttcctgct gcggttcagt gagagctccc gggaaggggc

2221 catcacattc acatgggtgg agcggtccca gaacggaggc gaacctgact tccatgcggt

2281 tgaaccctac acgaagaaag aactttctgc tgttactttc cctgacatca ttcgcaatta

2341 caaagtcatg gctgctgaga atattcctga gaatcccctg aagtatctgt atccaaatat

2401 tgacaaagac catgcctttg gaaagtatta ctccaggcca aaggaagcac cagagccaat

2461 ggaacttgat ggccctaaag gaactggata tatcaagact gagttgattt ctgtgtctga

2521 agttcaccct tctagacttc agaccacaga caacctgctc cccatgtctc ctgaggagtt

2581 tgacgaggtg tctcggatag tgggctctgt agaattcgac agtatgatga acacagtata

2641 gagcatgaat ttttttcatc ttctctggcg acagttttcc ttctcatctg tgattccctc

2701 ctgctactct gttccttcac atcctgtgtt tctagggaaa tgaaagaaag gccagcaaat

2761 tcgctgcaac ctgttgatag caagtgaatt tttctctaac tcagaaacat cagttactct

2821 gaagggcatc atgcatctta ctgaaggtaa aattgaaagg cattctctga agagtgggtt

2881 tcacaagtga aaaacatcca gatacaccca aagtatcagg acgagaatga gggtcctttg

2941 ggaaaggaga agttaagcaa catctagcaa atgttatgca taaagtcagt gcccaactgt

3001 tataggttgt tggataaatc agtggttatt tagggaactg cttgacgtag gaacggtaaa

Figure 32 (con't)

3061 tttctgtggg agaattctta catgttttct ttgctttaag tgtaactggc agttttccat

3121 tggtttacct gtgaaatagt tcaaagccaa gtttatatac aattatatca gtcctctttc

3181 aaaggtagcc atcatggatc tggtaggggg aaaatgtgta ttttattaca tctttcacat

3241 tggctattta aagacaaaga caaattctgt ttcttgagaa gagaatatta gctttactgt

3301 ttgttatggc ttaatgacac tagctaatat caatagaagg atgtacattt ccaaattcac

3361 aagttgtgtt tgatatccaa agctgaatac attctgcttt catcttggtc acatacaatt

3421 atttttacag ttctcccaag ggagttaggc tattcacaac cactcattca aaagttgaaa

3481 ttaaccatag atgtagataa actcagaaat ttaattcatg tttcttaaat gggctacttt

3541 gtccttttg ttattagggt ggtatttagt ctattagcca caaaattggg aaaggagtag

3601 aaaaagcagt aactgacaac ttgaataata caccagagat aatatgagaa tcagatcatt

3661 tcaaaactca tttcctatgt aactgcattg agaactgcat atgtttcgct gatatatgtg

3721 tttttcacat ttgcgaatgg ttccattctc tctcctgtac tttttccaga cacttttttg

3781 agtggatgat gtttcgtgaa gtatactgta ttttttacctt tttccttcct tatcactgac

3841 acaaaaagta gattaagaga tgggtttgac aaggttcttc ccttttacat actgctgtct

3901 atgtggctgt atcttgtttt tccactactg ctaccacaac tatattatca tgcaaatgct

3961 gtattcttct ttggtggaga taaagatttc ttgagttttg ttttaaaatt aaagctaaag

4021 tatctgtatt gcattaaata taatatgcac acagtgcttt ccgtggcact gcatacaatc

4081 tgaggcctcc tctctcagtt tttatataga tggcgagaac ctaagtttca gttgatttta

4141 caattgaaat gactaaaaaa caaagaagac aacattaaaa caatattgtt tctaattgct

4201 gaggtttagc tgtcagttct ttttgccctt tgggaattcg gcatggtttc attttactgc

4261 actagccaag agactttact tttaagaagt attaaaattc taaaattcaa aaaaaaaaaa

4321 aaaaaa (SEQ ID NO: 14)

Figure 32 (con't)

1 ggagtcagtg atttgaacga agtactttca gtttcatatt actctaaatc cattacaaat

61 ctgcttagct tctaaatatt tcatcaatga ggaaatccca gccctacaac ttcggaacag

121 tgaaatatta gtccagggat ccagtgagag acacagaagt gctagaagcc agtgctcgtg

181 aactaaggag aaaaagaaca gacaagggaa cagcctggac atggcatcag agatccacat

241 gacaggccca atgtgcctca ttgagaacac taatgggcga ctgatggcga atccagaagc

301 tctgaagatc ctttctgcca ttacacagcc tatggtggtg gtggcaattg tgggcctcta

361 ccgcacaggc aaatcctacc tgatgaacaa gctggctgga aagaaaaagg gcttctctct

421 gggctccacg gtgcagtctc acactaaagg aatctggatg tggtgtgtgc cccaccccaa

481 gaagccaggc cacatcctag ttctgctgga caccgagggt ctgggagatg tagagaaggg

541 tgacaaccag aatgactcct ggatcttcgc cctggccgtc ctcctgagca gcaccttcgt

601 gtacaatagc ataggaacca tcaaccagca ggctatggac caactgtact atgtgacaga

661 gctgacacat agaatccgat caaaatcctc acctgatgag aatgagaatg aggttgagga

721 ttcagctgac tttgtgagct tcttcccaga ctttgtgtgg acactgagag atttctccct

781 ggacttggaa gcagatggac aacccctcac accagatgag tacctgacat actccctgaa

841 gctgaagaaa ggtaccagtc aaaaagatga aacttttaac ctgcccagac tctgtatccg

901 gaaattcttc ccaaagaaaa aatgctttgt ctttgatcgg cccgttcacc gcaggaagct

961 tgcccagctc gagaaactac aagatgaaga gctggacccc gaatttgtgc aacaagtagc

1021 agacttctgt tcctacatct ttagtaattc caaaactaaa actctttcag gaggcatcca

1081 ggtcaacggg cctcgtctag agagcctggt gctgacctac gtcaatgcca tcagcagtgg

1141 ggatctgccg tgcatggaga acgcagtcct ggccttggcc cagatagaga actcagctgc

1201 agtgcaaaag gctattgccc actatgaaca gcagatgggc cagaaggtgc agctgcccac

1261 agaaaccctc caggagctgc tggacctgca cagggacagt gagagagagg ccattgaagt

1321 cttcatcagg agttccttca aagatgtgga ccatctattt caaaaggagt tagcggccca

1381 gctagaaaaa aagcgggatg acttttgtaa acagaatcag gaagcatcat cagatcgttg

1441 ctcagcttta cttcaggtca ttttcagtcc tctagaagaa gaagtgaagg cgggaattta

1501 ttcgaaacca gggggctatc gtctctttgt tcagaagcta caagacctga agaaaaagta

Figure 33

1561 ctatgaggaa ccgaggaagg ggatacaggc tgaagagatt ctgcagacat acttgaaatc

1621 caaggagtct atgactgatg caattctcca gacagaccag actctcacag aaaaagaaaa

1681 ggagattgaa gtggaacgtg tgaaagctga gtctgcacag gcttcagcaa aaatgttgca

1741 ggaaatgcaa agaaagaatg agcagatgat ggaacagaag gagaggagtt atcaggaaca

1801 cttgaaacaa ctgactgaga agatggagaa cgacagggtc cagttgctga aagagcaaga

1861 gaggaccctc gctcttaaac ttcaggaaca ggagcaacta ctaaaagagg gatttcaaaa

1921 agaaagcaga ataatgaaaa atgagataca ggatctccag acgaaaatga gacgacgaaa

1981 ggcatgtacc ataagctaaa gaccagagcc ttcctgtcac ccctaaccaa ggcataattg

2041 aaacaatttt agaatttgga acaagcgtca ctacatttga taataattag atcttgcatc

2101 ataacaccaa aagtttataa aggcatgtgg tacaatgatc aaaatcatgt tttttcttaa

2161 aaaaaaaaaa agactgtaaa ttgtgcaaca aagatgcatt tacctctgta tcaactcagg

2221 aaatctcata agctggtacc actcaggaga agtttattct tccagatgac cagcagtaga

2281 caaatggata ctgagcagag tcttaggtaa aagtcttggg aaatatttgg gcattggtct

2341 ggccaagtct acaatgtccc aatatcaagg acaaccaccc tagcttctta gtgaagacaa

2401 tgtacagtta tccgttagat caagactaca cggtctatga gcaataatgt gatttctgga

2461 cattgcccat gtataatcct cactgatgat ttcaagctaa agcaaaccac cttatacaga

2521 gatctagaat ctctttatgt tctccagagg aaggtggaag aaaccatggg caggagtagg

2581 aattgagtga taaacaattg ggctaatgaa gaaaacttct cttattgttc agttcatcca

2641 gattataact tcaatgggac actttagacc attagacaat tgacactgga ttaaacaaat

2701 tcacataatg ccaaatacac aatgtattta tagcaacgta taatttgcaa agatggactt

2761 taaaagatgc tgtgtaacta aactgaaata attcaattac ttattattta gaatgttaaa

2821 gcttatgata gtcttttcta actcttaaca ctcatacttg aaaactttct gagtttcccc

2881 agaagagaat atgggatttt ttttgacatt tttgactcat ttaataatgc tcttgtgttt

2941 acctagtata tgtagacttt gtcttatgtg tgaaaagtcc taggaaagtg gttgatgttt

3001 cttatagcaa ttaaaaatta tttttgaact gaaaatacaa tgtatttcac (SEQ ID NO: 15)

Figure 33 (con't)