US 20100256001A1

(19) **United States**

(12) **Patent Application Publication**

Niculescu et al.

(10) **Pub. No.: US 2010/0256001 A1**

(43) **Pub. Date: Oct. 7, 2010**

(54) **BLOOD BIOMARKERS FOR MOOD DISORDERS**

(75) Inventors: **Alexander B. Niculescu**, Indianapolis, IN (US); **John I. Nurnberger**, Indianapolis, IN (US); **Daniel R. Salomon**, San Diego, CA (US)

Correspondence Address:
BARNES & THORNBURG LLP
P.O. BOX 2786
CHICAGO, IL 60690-2786 (US)

(73) Assignee: **THE SCRIPPS RESEARCH INSTITUTE**, La Jolla, CA (US)

(21) Appl. No.: **12/594,378**

(22) PCT Filed: **Apr. 2, 2008**

(86) PCT No.: **PCT/US08/59120**

§ 371 (c)(1),
(2), (4) Date: **Jun. 24, 2010**

Related U.S. Application Data

(60) Provisional application No. 60/909,859, filed on Apr. 3, 2007.

Publication Classification

(51) **Int. Cl.**

C40B 30/00	(2006.01)
G01N 33/68	(2006.01)
C12Q 1/48	(2006.01)
C12Q 1/44	(2006.01)
C12Q 1/42	(2006.01)
C12Q 1/02	(2006.01)
C12Q 1/68	(2006.01)
G01N 33/53	(2006.01)

(52) **U.S. Cl.** **506/7**; 436/86; 436/87; 435/15; 435/19; 435/21; 435/29; 435/6; 435/7.92

(57) **ABSTRACT**

A plurality of markers determine the diagnosis of a mood disorder based on their expression in a sample such as blood. Subsets of biomarkers predict the diagnosis of high or low mood disorders. The biomarkers are identified using a convergent functional genomics approach based on animal and human data. Methods and compositions for clinical diagnosis of mood disorders are provided.

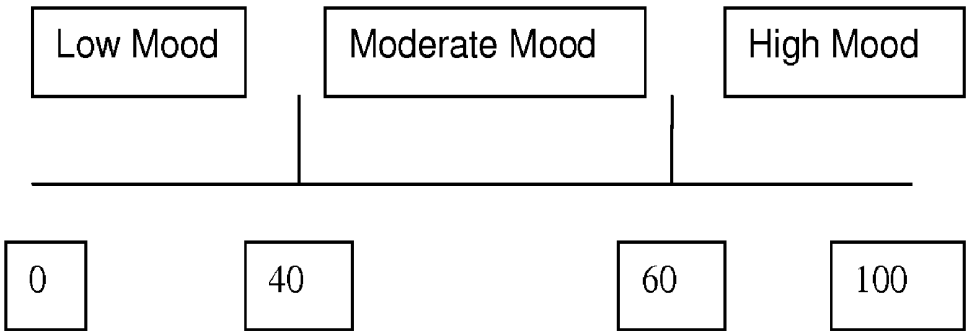


FIG. 1

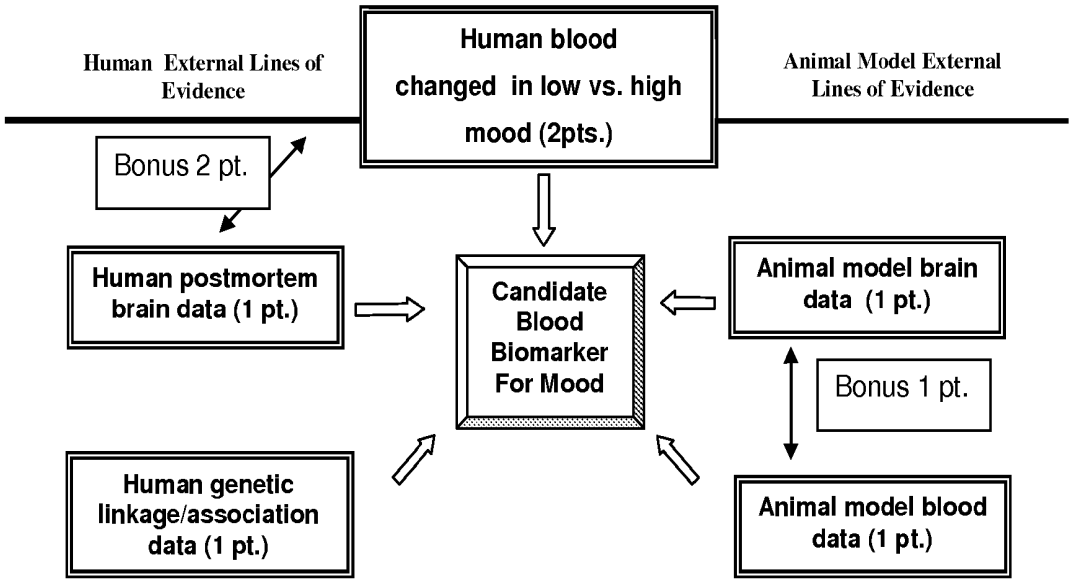

Human External Lines of Evidence
Human blood changed in low vs. high mood (2pts.)
Animal Model External Lines of Evidence
Bonus 2 pt.
Human postmortem brain data (1 pt.)
Candidate Blood Biomarker For Mood
Animal model brain data (1 pt.)
Bonus 1 pt.
Human genetic linkage/association data (1 pt.)
Animal model blood data (1 pt.)

FIG. 2A

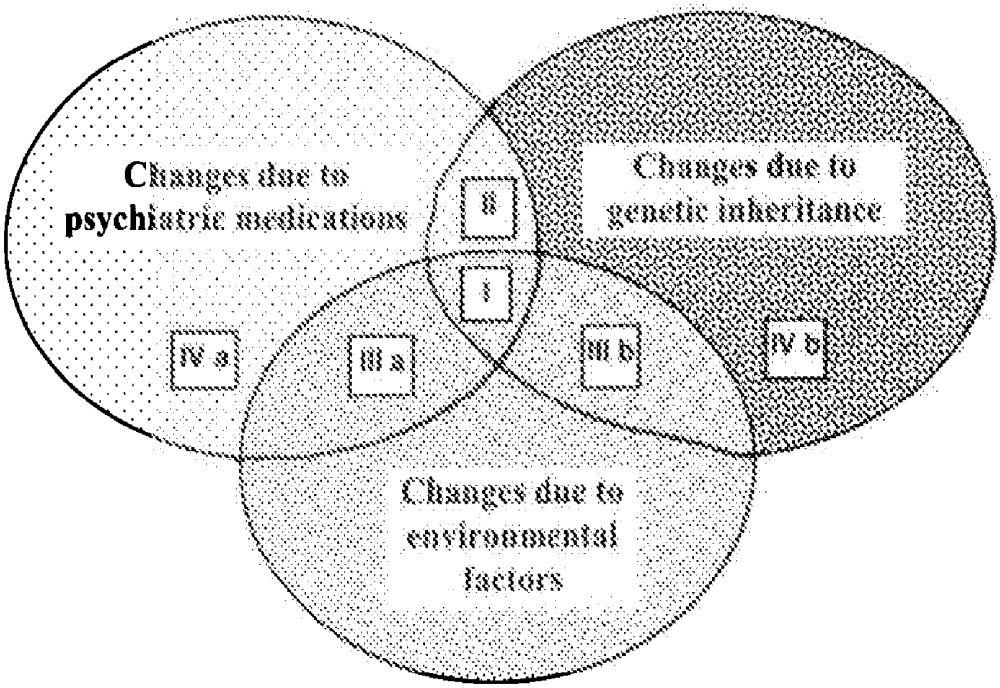

Changes due to psychiatric medications
Changes due to genetic inheritance
Changes due to environmental factors
II
I
IV a
III a
III b
IV b

FIG. 2B

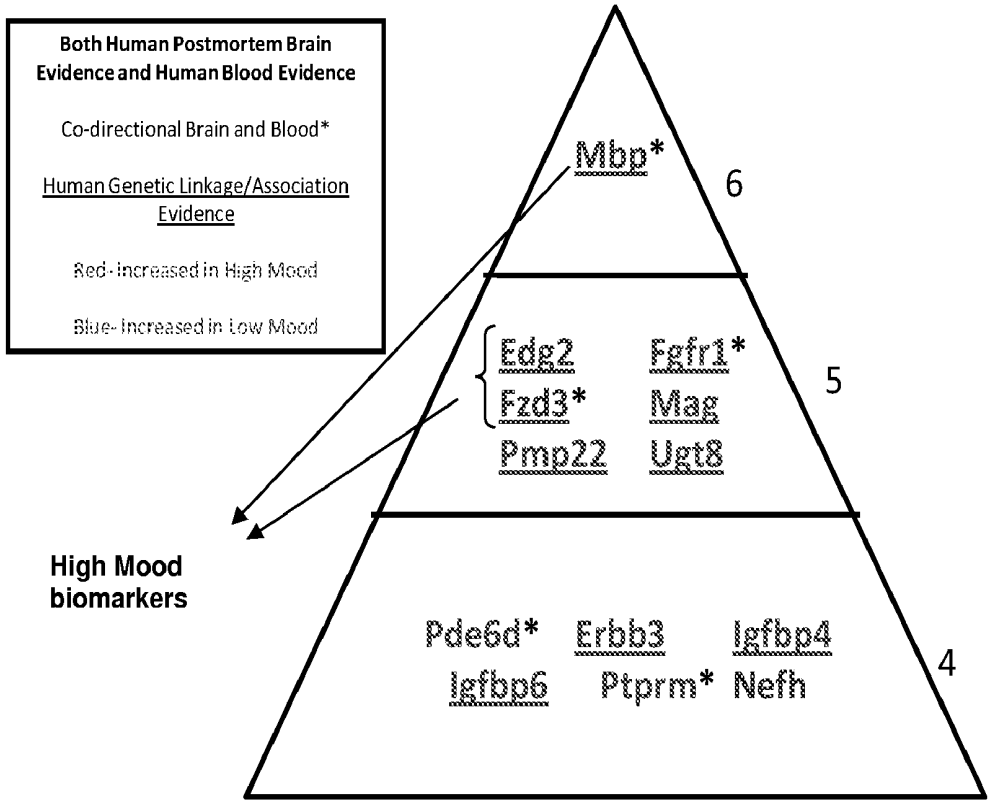
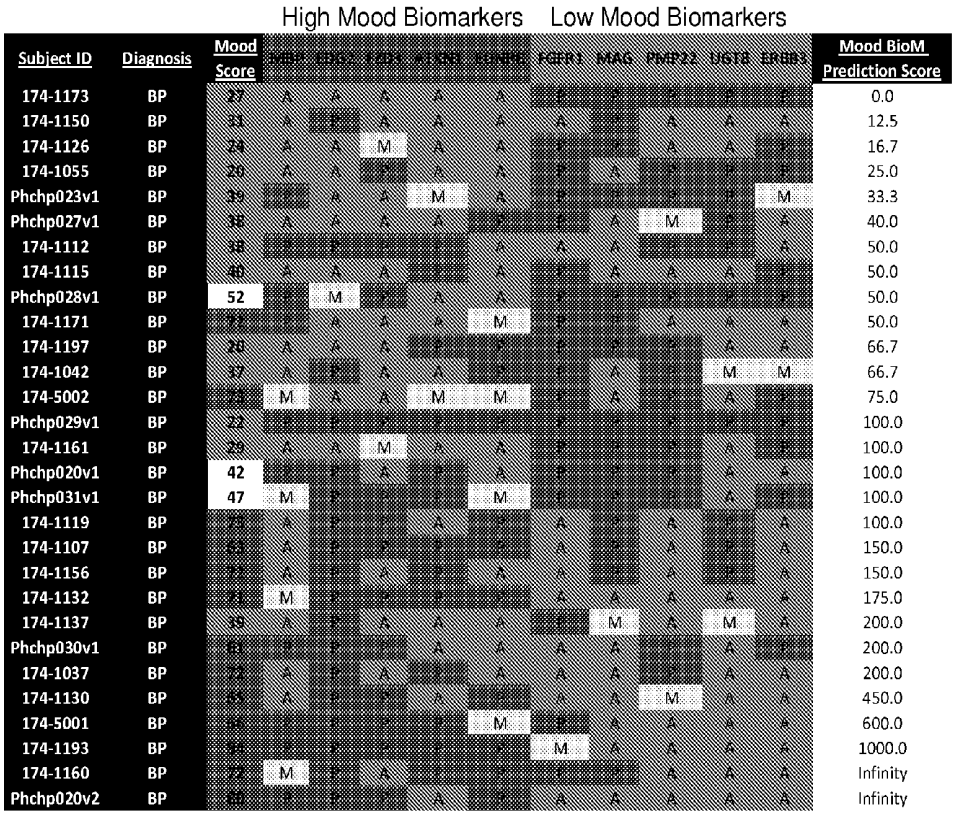

Both Human Postmortem Brain Evidence and Human Blood Evidence
Co-directional Brain and Blood*
Human Genetic Linkage/Association Evidence
Red- increased in High Mood
Blue- increased in Low Mood
Mbp*
6
Edg2
Fgfr1*
Fzd3*
Mag
Pmp22
Ugt8
5
High Mood biomarkers
Pde6d*
Erbb3
Igfbp4
Igfbp6
Ptprm*
Nefh
4

FIG. 3



Subject ID	Diagnosis	Mood Score	High Mood Biomarkers					Low Mood Biomarkers					Mood BioM Prediction Score
			PHB1	EDG2	1E24	ATSG1	EGFRE	FGFR1	MAG	PMP22	UGT8	ERBB3	
174-1173	BP	27	A	A	A	A	A	P	P	P	P	P	0.0
174-1150	BP	31	A	P	A	A	A	A	P	A	A	A	12.5
174-1126	BP	24	A	A	M	A	A	P	P	A	A	P	16.7
174-1055	BP	20	A	A	P	A	A	P	A	P	P	P	25.0
Phchp023v1	BP	39	P	A	A	M	A	P	P	P	P	M	33.3
Phchp027v1	BP	38	A	A	A	A	P	P	A	M	P	A	40.0
174-1112	BP	38	P	P	P	P	A	A	A	P	P	A	50.0
174-1115	BP	40	A	A	A	P	A	P	A	A	A	P	50.0
Phchp028v1	BP	52	P	M	P	A	A	P	P	P	P	P	50.0
174-1171	BP	71	P	A	A	A	M	P	P	A	A	A	50.0
174-1197	BP	20	A	A	A	P	P	P	P	P	A	A	66.7
174-1042	BP	47	A	P	A	A	P	P	A	P	M	M	66.7
174-5002	BP	73	M	A	A	M	M	P	A	P	A	P	75.0
Phchp029v1	BP	22	P	P	P	P	P	P	P	P	P	P	100.0
174-1161	BP	29	A	A	M	A	A	P	P	P	A	P	100.0
Phchp020v1	BP	42	P	P	A	P	A	P	P	P	A	A	100.0
Phchp031v1	BP	47	M	P	P	P	M	P	P	P	A	P	100.0
174-1119	BP	75	A	P	P	A	P	A	P	A	P	A	100.0
174-1107	BP	63	A	P	P	P	P	A	P	A	P	A	150.0
174-1156	BP	72	A	P	A	P	A	A	P	P	P	A	150.0
174-1132	BP	71	M	P	P	P	P	A	A	A	A	A	175.0
174-1137	BP	39	A	P	A	A	P	P	M	A	M	A	200.0
Phchp030v1	BP	61	P	P	P	A	A	A	A	P	A	P	200.0
174-1037	BP	72	A	P	A	P	A	A	A	P	A	A	200.0
174-1130	BP	65	A	P	P	A	P	A	A	M	A	A	450.0
174-5001	BP	66	P	P	P	P	M	P	A	A	A	A	600.0
174-1193	BP	84	P	P	P	P	P	M	A	A	A	A	1000.0
174-1160	BP	72	M	P	A	P	P	P	P	A	A	A	Infinity
Phchp020v2	BP	88	P	P	P	A	P	A	A	A	A	A	Infinity

FIG. 4

High Mood biomarkers Low Mood Biomarkers

	Diagnosis	Mood Score	MBP	KCNC1	KCC1	ATP6V1	EDNRB	FGFR1	MAG	PMP22	GGT6	ERBB3	Mood BioM Prediction Score
phchp015v1	SubPD	75	A	A	A	A	A	A	M	A	A	A	0.0
phchp033v1	SZA	38	A	A	A	A	A	A	A	A	A	A	25.0
phchp010v1	SZA	50	A	A	A	A	A	A	A	A	A	A	33.3
phchp025v1	SZ	30	A	A	A	A	A	A	A	A	A	A	50.0
phchp003v3	SZ	50	A	A	A	A	A	A	A	A	A	M	57.1
phchp024v1	SZA	50	A	A	A	A	A	A	A	A	A	A	60.0
phchp026v1	SZA	48	A	A	A	A	A	A	A	A	A	A	60.0
phchp006v2	SZA	33	M	A	A	A	A	A	A	A	A	A	62.5
phchp010v3	SZA	44	A	A	A	A	A	A	A	A	A	A	66.7
phchp021v1	SZA	39	A	A	M	A	A	A	A	A	A	A	70.0
phchp006v1	SZA	59	A	A	A	A	A	A	A	A	A	A	75.0
phchp019v1	SubPD	36	A	A	A	A	A	A	A	A	A	A	75.0
phchp022v2	SZ	45	A	A	A	A	A	A	A	A	A	A	75.0
phchp017v2	SZA	56	A	A	A	A	A	M	A	A	A	A	80.0
phchp004v1	SZA	70	A	A	A	A	A	A	A	A	A	A	100.0
phchp008v1	SZ	73	A	A	A	A	A	A	A	A	A	A	100.0
phchp009v1	SZ	55	A	A	A	A	A	A	A	A	A	A	100.0
phchp012v1	SZA	71	A	A	A	A	A	A	A	A	A	A	100.0
phchp013v1	SZA	58	A	A	A	A	A	A	A	A	A	A	100.0
phchp022v1	SZ	67	A	A	A	A	A	A	A	A	A	A	100.0
phchp014v1	SubPD	71	A	A	A	A	A	A	A	A	A	A	125.0
phchp005v1	SZA	34	A	A	A	A	A	A	A	A	A	M	133.3
phchp003v1	SZ	64	A	A	A	A	A	A	A	A	A	A	150.0
phchp010v2	SZA	56	A	A	A	A	A	A	A	A	A	A	150.0
phchp012v2	SZA	73	A	A	A	A	A	A	A	A	A	A	150.0
phchp018v1	SZA	74	A	A	A	A	A	A	A	A	A	A	200.0
phchp021v2	SZA	29	A	A	A	M	A	A	A	A	M	M	225.0
phchp005v2	SZA	38	A	A	A	A	A	A	A	A	A	A	250.0
phchp003v2	SZ	54	A	A	A	A	A	A	A	A	A	A	Infinity
phchp016v1	SZ	36	A	A	A	A	A	A	A	A	A	A	Infinity

FIG. 5

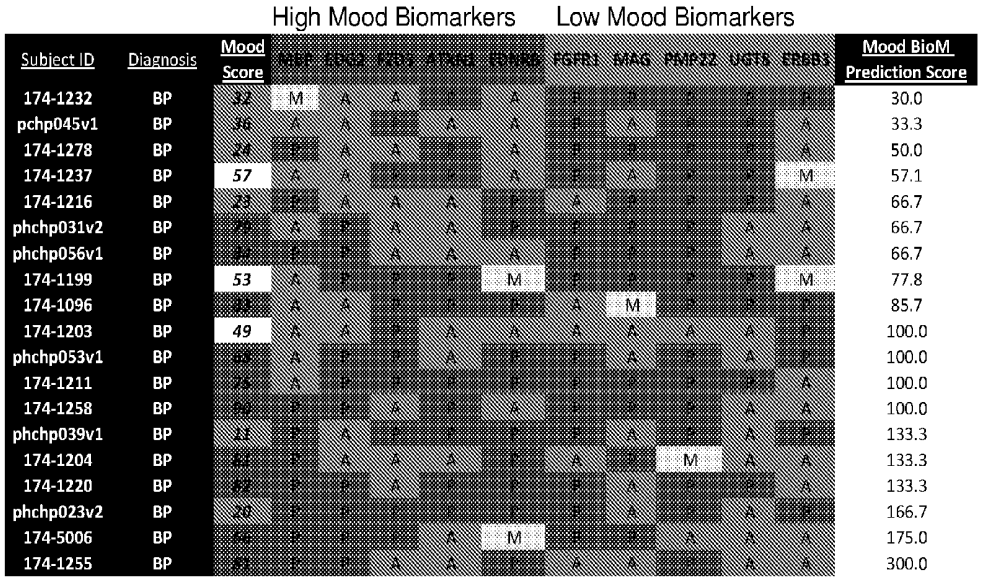

High Mood Biomarkers
Low Mood Biomarkers
Subject ID
Diagnosis
Mood Score
FGFR1
MAG
PMP22
UGT8
ERBB3
Mood BioM Prediction Score
174-1232 BP 12 M 30.0
pchp045v1 BP 36 33.3
174-1278 BP 24 50.0
174-1237 BP 57 57.1
174-1216 BP 29 66.7
phchp031v2 BP 75 66.7
phchp056v1 BP 66.7
174-1199 BP 53 77.8
174-1096 BP 85.7
174-1203 BP 49 100.0
phchp053v1 BP 100.0
174-1211 BP 100.0
174-1258 BP 100.0
phchp039v1 BP 133.3
174-1204 BP 133.3
174-1220 BP 133.3
phchp023v2 BP 166.7
174-5006 BP 175.0
174-1255 BP 300.0

FIG. 6

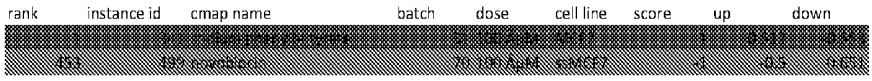

rank
instance id
cmap name
batch
dose
cell line
score
up
down

FIG. 7

BLOOD BIOMARKERS FOR MOOD DISORDERS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. provisional application Ser. No. 60/909,859, filed Apr. 2, 2007, the disclosure of which is hereby incorporated by reference in its entirety.

[0002] Part of the work during the development of this invention was made with government support from the National Institutes of Health under grant NIMH R01 MH071912-01. The U.S. Government has certain rights in the invention.

BACKGROUND

[0003] Research into the biological basis of mood disorders (e.g., bipolar disorders, depression) has been primarily focused in human and animal studies mostly independently. The two avenues of research have complementary strengths and weaknesses. In human genetic studies, for example, in samples of patients with mood disorders and their family members, positional cloning methods such as linkage analysis, linkage-disequilibrium mapping, and candidate-gene association analysis are narrowing the search for the chromosomal regions harboring risk genes for the illness and, in some cases, identifying plausible candidate genes and polymorphisms that will require further validation. Human postmortem brain gene expression studies have also been employed as a way of trying to identify candidate genes for mood and other neuropsychiatric disorders. In general, human studies suffer from issues of sensitivity—the signal is often difficult to detect due to the noise generated by the genetic heterogeneity of individuals and the effects of diverse environmental exposures on gene expression and phenotypic penetrance.

[0004] In animal studies, carried out in isogenic strains with controlled environmental exposure, the identification of putative neurobiological substrates of mood disorders is typically accomplished by modeling human mood disorders through pharmacological or genetic manipulations. Animal model studies suffer from issues of specificity-questions regarding direct relevance to the human disorder modeled. Each independent line of investigation (i.e., human and animal studies) is contributing to the incremental gains in knowledge of mood disorders etiology witnessed in the last decade.

[0005] However, a lack of integration between these two lines of investigation, hinders scientific understanding and slows the pace of discovery. Psychiatric phenotypes, as currently defined, are primarily the result of clinical consensus criteria rather than empirical determination. The present disclosure provides methods and compositions that empirically determine disease states for diagnosis and treatment.

[0006] Objective biomarkers of illness and treatment response would make a significant difference in the ability to diagnose and treat patients with psychiatric disorders, eliminating subjectivity and reliance of patient's self-report of symptoms. Blood gene expression profiling has emerged as a particularly interesting area of research in the search for peripheral biomarkers. Most of the studies to date have focused on human 1 lymphoblastoid cell lines (LCLs) gene expression profiling, comparison between illness groups and normal controls. They suffer from one of both of the following limitations: 1) the sample size used is often small. Given the genetic heterogeneity in human samples and the effects of illness state and environmental history, including medications and drugs, on gene expression, it may not be reliable to extract bona fide findings. 2) Use of lymphoblastoid cell lines—passaged lymphoblastoid cell lines provide a self-renewable source of material, and are purported to avoid the effects of environmental exposure of cells from fresh blood. Fresh blood, however, with phenotypic state information gathered at time of harvesting, may be more informative than immortalized lymphocytes, and may avoid some of the caveats of Epstein-Barr virus (EBV) immortalization and cell culture passaging.

[0007] The current state of the understanding of the genetic and neurobiological basis for mood disorders (such as bipolar disorder and depression) in general, and of peripheral molecular biomarkers of the illness in particular, is still inadequate. Almost all of the fundamental genetic, environmental, and biological elements needed to delineate the etiology and pathophysiology of mood disorders are yet to be completely identified, understood and validated. One of the rate-limiting steps has been the lack of concerted integration across disciplines and methodologies. The use of a multidisciplinary, integrative research framework as in the present disclosure provided herein, should lead to a reduction in the historically high rate of inferential errors committed in studies of complex diseases like bipolar disorder and depression.

[0008] Identification and validation of peripheral biomarkers for bipolar mood disorders has proven arduous, despite recent large-scale efforts. Human genomic studies are susceptible to the issue of being underpowered, due to genetic heterogeneity, the effect of variable environmental exposure on gene expression, and difficulty of accrual of large samples. Animal model gene expression studies, in a genetically homogeneous and experimentally tractable setting, can avoid artifacts and provide sensitivity of detection. Subsequent comparisons of the animal datasets with human genetic and genomic datasets can ensure cross-validatory power and specificity.

[0009] Convergent functional genomics (CFG), is an approach that translationally cross-matches animal model gene expression data with human genetic linkage data and human tissue data (blood, postmortem brain), as a Bayesian strategy of cross validating findings and identifying candidate genes, pathways and mechanisms for neuropsychiatric disorders. Predictive biomarkers for mood disorders are desired for clinical diagnosis and treatment purposes. The present disclosure provides several biomarkers that are predictive of mood disorders in clinical settings.

SUMMARY

[0010] Methods and compositions to clinically diagnose mood disorders using a panel of biomarkers are disclosed. A panel of biomarkers may include 1 to about 100 or more biomarkers. The panel of biomarkers includes one or more biomarkers for high and low mood disorders. Blood is a suitable sample for measuring the levels or presence of one or more of the biomarkers provided herein.

[0011] In an aspect, psychiatric symptoms measured in a quantitative fashion at time of blood draw in human subjects focus on all or nothing phenomena (genes turned on and off in low symptom states vs. high symptom states). Some of the biomarkers have cross-matched animal and human data,

using a convergent functional genomics approach including blood datasets from animal models.

[0012] Prioritized list of high probability blood biomarkers, provided herein, for mood state using cross-matching of animal and human data, provide a unique predictive power of the biomarkers, which have been experimentally tested.

[0013] The disclosure also provides various methods of assigning prediction scores for mood state based on the ratio of biomarkers for high mood vs. biomarkers for low mood in the blood of individual subjects, termed as BioM Mood Prediction Score. In an aspect, a panel of about 10 biomarkers, designated as BioM-10 Mood Panel, demonstrated good accuracy in predicting actual measured mood (high and low) in an enlarged cohort of subjects.

[0014] In an aspect, the present disclosure provides methods and compositions for developing clinical blood tests to quantify gene expression for diagnosis and quantitation of protein levels through immunological approaches such as enzyme-linked immunosorbent assays (ELISA).

[0015] A method of diagnosing a mood disorder in an individual includes the steps of:

[0016] (a) determining the level of a plurality of biomarkers for the mood disorder in an isolated sample from the individual, the plurality of biomarkers selected from the group of biomarkers listed in Tables 3 and 7; and

[0017] (b) diagnosing the mood state (high mood—mania, low mood—depression) in the individual based on the level of the plurality of biomarkers.

[0018] A plurality of biomarkers, in an aspect, includes a subset of about 10 markers designated as Mbp, Edg2, Fzd3, Atxn1, Ednrb for high mood and Fgfr1, Mag, Pmp22, Ugt8, Erbb3 for low mood.

[0019] A plurality of biomarkers, in an aspect, includes a subset of about 20 biomarkers designated as Mbp, Edg2, Fzd3, Atxn1, Ednrb, Pde9a, Plxnd1, Camk2d, Dio2, Lepr for high mood and Fgfr1, Mag, Pmp22, Ugt8, Erbb3, Igfbp4, Igfbp6, Pde6d, Ptprm, Nefh for low mood.

[0020] A mood disorder is a Bipolar disorder and the sample is a bodily fluid. A suitable sample is blood. The level of the biomarker may be determined in a blood sample of the individual.

[0021] In an aspect, the level of the biomarker is determined by analyzing the expression level of RNA transcripts. In an aspect, the expression level of the biomarker is determined by analyzing the level of protein or peptides or fragments thereof. Suitable detection techniques include microarray gene expression analysis, polymerase chain reaction (PCR), real-time PCR, quantitative PCR, immunohistochemistry, enzyme-linked immunosorbent assays (ELISA), and antibody arrays.

[0022] In an aspect, the determination of the level of the plurality of biomarkers is performed by an analysis of the presence or absence of the biomarkers.

[0023] A method of diagnosing mood disorder in an individual includes the steps of:

[0024] (a) performing a quantitative determination of the level of a panel of at least 10 biomarkers selected from Tables 3 and 7 in a bodily fluid sample isolated from the individual, wherein the panel comprises at least one biomarker for high mood disorder;

[0025] (b) assigning a predictive value or score to the level of the biomarkers; and

[0026] (c) diagnosing the mood disorder based on the assigned value or score.

[0027] A method of predicting the probable course and outcome (prognosis) of a mood disorder includes the steps of:

[0028] (a) obtaining a test sample from a subject, wherein the subject is suspected of having a mood disorder;

[0029] (b) analyzing the test sample for the presence or level of a plurality of biomarkers of the mood disorder, the markers selected from the group consisting of biomarkers listed in Tables 3 and 7; and

[0030] (c) determining the prognosis of the subject based on the presence or level of the biomarkers and one or more clinicopathological data to implement a particular treatment plan for the subject.

[0031] A treatment plan for a high mood disorder includes administering a pharmaceutical composition selected from a group that includes Depakote (divalproex), Lithobid (lithium), Lamictal (lamotrigene), Tegretol (carbamazepine), Topomax (topiramate).

[0032] A treatment plan for a low mood disorder includes administering a pharmaceutical composition selected from the group consisting of Prozac (fluoxetine), Zolof (sertraline), Celexa (citalopram), Cymbalta (duloxetine), Effexor (venlafaxine) or Wellbutrin (buproprion).

[0033] A clinicopathological data is selected from a group that includes patient age, previous personal and/or familial history of the mood disorder, previous personal and/or familial history of response to treatment, and any genetic or biochemical predisposition to psychiatric illness.

[0034] A suitable test sample includes fresh blood, stored blood, fixed, paraffin-embedded tissue, tissue biopsy, tissue microarray, fine needle aspirates, peritoneal fluid, ductal lavage and pleural fluid or a derivative thereof.

[0035] A method of predicting the likelihood of a successful treatment for a mood disorder in a patient includes the steps of:

[0036] (a) determining the expression level of at least 10 biomarkers, wherein the biomarkers comprise a subset of about 10 markers designated as Mbp, Edg2, Fzd3, Atxn1, Ednrb for high mood and Fgfr1, Mag, Pmp22, Ugt8, Erbb3 for low mood; and

[0037] (b) predicting the likelihood of successful treatment for the mood disorder by determining whether the sample from the patient expresses biomarkers for a high mood disorder or a low mood disorder.

[0038] A method of treating a patient suspected of suffering a mood disorder, the method includes the steps of:

[0039] (a) diagnosing whether the patient suffers from a high mood or a low mood disorder by determining the expression level of one or more of the biomarkers listed in Tables 3 and 7 in a sample obtained from the patient;

[0040] (b) selecting a treatment for the mood disorder based on the determination whether the patient suffers from a high mood or a low mood disorder; and

[0041] (c) administering to the patient a therapeutic agent capable of treating the high or the low mood disorder.

[0042] A treatment plan may be a personalized plan for the patient.

[0043] A method for clinical screening of agents capable of affecting a mood disorder, the method includes the steps of:

[0044] (a) administering a candidate agent to a population of individuals suspected of suffering from a mood disorder or induced to suffer a mood disorder;

[0045] (b) monitoring the expression profile of one or more of the biomarkers listed in Tables 3 and 7 in blood samples

obtained from the individuals receiving the candidate agent compared to a control group; and

[0046] (c) determining that the candidate agent is capable of affecting the mood disorder based on the expression profile of one or more of the biomarkers in the blood samples obtained from the individuals receiving the candidate drug compared to the control.

[0047] A mood disorder microarray includes a plurality of nucleic acid molecules representing genes selected from the group of genes listed in Tables 3 and 7.

[0048] A kit for diagnosing a mood disorder includes a component selected from the group consisting of (i) oligonucleotides for amplification of one or more genes listed in Tables 3 and 7, (ii) immunohistochemical agents capable of identifying the protein products of one or more biomarkers listed in Table 7, (iii) a microarray to detect the plurality of markers listed in Tables 3 and 7, and (iv) a biomarker expression index representing the genes listed in Tables 3 and 7 for correlation.

[0049] A diagnostic microarray includes a panel of about 10 biomarkers that are predictive of a mood disorder, wherein the microarray includes nucleic acid fragments representing biomarkers designated as Mbp, Edg2, Fzd3, Atxn1, Ednrb for high mood and Fgfr1, Mag, Pmp22, Ugt8, Erbb3 for low mood.

[0050] A diagnostic antibody array includes a plurality of antibodies that recognize one or more epitopes corresponding to the protein products of the biomarkers designated as Mbp, Edg2, Fzd3, Atxn1, Ednrb for high mood and Fgfr1, Mag, Pmp22, Ugt8, Erbb3 for low mood.

[0051] A diagnostic microarray consists essentially of the top candidate markers from tables 3 and 7.

BRIEF DESCRIPTION OF THE DRAWINGS

[0052] FIG. **1** shows Visual-Analog Mood Scale (VAS) scoring for some of the biomarkers used herein.

[0053] FIG. **2** shows prioritization (A) and conceptualization (B) of results. A: convergent functional genomics approach for candidate biomarker prioritization. Scoring of independent lines of evidence yields (maximum score=9 points). B: Conceptualization of blood candidate biomarker genes.

[0054] FIG. **3** illustrates some of the candidate biomarker genes for mood. Prioritization was based on integration of multiple lines of evidence. On the right side of the pyramid is their CFG score.

[0055] FIG. **4** is a comparison of BioM-10 Mood Prediction Score and actual mood scores in the primary cohort of bipolar subjects (n=29). BP—bipolar. Mood scores are based on subject self-report on mood VAS scale administered at time of blood draw. For biomarkers: A—called Absent by MASS analysis. P—called Present by MASS analysis. M—called Marginally Present by MASS analysis. A is scored as 0, M as 0.5 and P as 1. BioM Mood Prediction Score is based on the ratio of the sum of the scores for high mood biomarkers and sum of scores for low mood biomarkers, multiplied by 100. A cutoff score of 100 and above was used for high mood. inf—infinity-denominator is 0.

[0056] FIG. **5** is a comparison of BioM-10 Mood Prediction Score and actual mood scores in an independent cohort of psychotic disorders subjects (n=30). SZ—schizophrenia; SZA—schizoaffective disorder; SubPD—substance induced psychosis. Mood scores are based on subject self-report on mood VAS scale administered at time of blood draw. For biomarkers: A—called Absent by MASS analysis. P—called Present by MASS analysis. M—called Marginally Present by MASS analysis. A is scored as 0, M as 0.5 and P as 1. BioM Mood Prediction Score is based on the ratio of the sum of the scores for high mood biomarkers and sum of scores for low mood biomarkers, multiplied by 100. A cutoff score of 100 and above was used for high mood. inf—infinity-denominator is 0.

[0057] FIG. **6** shows Connectivity Map interrogation of drugs that have similar gene expression signatures to that of high mood. A score of 1 indicates a maximal similarity with the gene expression effects of high mood, and a score of −1 indicates a maximal opposite effects to high mood.

[0058] FIG. **7** shows Comparison of BioM-10 Mood Prediction Score and actual mood scores in a secondary independent cohort of bipolar subjects (n=19). BP—bipolar. Mood scores are based on subject self-report on mood VAS scale administered at time of blood draw. For biomarkers: A—called Absent by MASS analysis. P—called Present by MASS analysis. M—called Marginally Present by MASS analysis. A is scored as 0, M as 0.5 and P as 1. BioM Mood Prediction Score is based on the ratio of the sum of the scores for high mood biomarkers and sum of scores for low mood biomarkers, multiplied by 100. A cutoff score of 100 and above was used for high mood.

DETAILED DESCRIPTION

[0059] Patterns of changes in the blood that reflect whether a person has low mood (depression) or high mood (mania) are disclosed. In an embodiment, these changes are analyzed at the level of gene expression, and involved genes that generally are expressed in the brain.

[0060] Unlike cancer, in psychiatric disorders, one cannot perform a biopsy the target organ (brain). Therefore, implementing a readily accessible peripheral readout in blood or any other non-brain tissue is highly useful. Blood-based screening is clinically easier to perform than a cerebro-spinal fluid (CSF) analysis or a nasal epithelium bipopsy.

[0061] Relying on the patients' self-report of symptoms and the clinician's impression of how ill the patient is alone do not necessarily provide an accurate diagnosis. Because patients' mind itself is affected in a mood disorder, their reporting of symptoms of how they feel may not be accurate or may not predict the nature of disease outcome. For example, patients aren't sure how ill they really are, and neither is the clinician—sometimes dismissing their symptoms, sometimes overestimating them. Therefore, an objective test for disease state, disease severity, and to measure response to treatment is highly desirable.

[0062] For example, for depression in general, a patient gets started on an antidepressant, and it may take weeks or months before it is known if the medication is working or if something else needs to be tried. A blood test for mood state for the biomarkers disclosed herein is able to objectively reflect whether a treatment works.

[0063] For example, for someone diagnosed as depressed but is in reality bipolar (manic-depressive), an antidepressant medication may be started only initially, and if bipolar, will be flipped by the antidepressant into a mixed state or frank mood elevation—hypomania or mania. With a panel of mood state markers, such unclear patients are monitored by repeated lab tests after the antidepressant is started, and if the markers indicate a shift beyond normal mood, to high mood, then medications can be systematically changed, a mood stabilizer

added, and a potentially dangerous and certainly miserable episode for the patient averted. This approach is useful, especially in children and adolescents, who are hard to diagnose using traditional clinical criteria only, and in whom mood states rapidly change.

[0064] Sub-groups of biomarkers can be identified for different subpopulations, potential gender differences, age related differences, response to different medications. Biomarkers disclosed herein may be personalized and tailored to the individual, based on their biomarker profiles.

[0065] In an aspect, the biomarkers disclosed herein are (i) derived from fresh blood, not immortalized cell lines; (ii) capable of providing quantitative mood state information obtained at the time of the blood draw; (iii) were derived from comparisons of extremes of low mood and high mood in patients, as opposed to patients vs. normal controls (where the differences could be due to a lot of other environmental factors, medication (side) effects vs. no medications; (iv) scored based on an all or nothing (Absent/Present) call for gene expression changes, not incremental changes in expression—statistically more robust and avoids false positives; (v) based on integration of multiple independent lines of evidence that permits extraction of signal from noise (large lists of genes), and prioritization of top candidates; and (vi) used to form the basis of prediction score algorithm based.

[0066] Integration of animal model and human data was used as a way of reducing the false-positives inherent in each approach and helping identify true biomarker molecules. Gene expression differences were measured in fresh blood samples from patients with bipolar disorder (manic-depressive illness) that have low mood vs. those that have high mood at the time of the blood draw. Separately, changes in gene expression were measured in the brains and bloods of a mouse pharmacogenomic model of bipolar disorder. Human blood gene expression data was integrated with animal model gene expression data, human genetic linkage/association data, and human postmortem data for cross-validating and prioritizing findings.

[0067] Gene expression changes in specific brain regions and blood from a pharmacogenomic animal model were used as cross-validators to identify human blood biomarkers for mood disorders. Pharmacogenomic mouse model of relevance to bipolar disorder includes treatments with an agonist of the illness/bipolar disorder-mimicking drug (methamphetamine) and an antagonist of the illness/bipolar disorder-treating drug (valproate). The pharmacogenomic approach is a tool for tagging genes that may have pathophysiological relevance. As an added advantage, some of these genes may be involved in potential medication effects present in human blood data (FIG. **2**).

[0068] In an aspect, human whole blood gene expression studies were initially performed in a primary cohort of bipolar subjects. Whole blood was used as a way of minimizing potential artifacts related to sample handling and separation of individual cell types, and also as a way of having a streamlined approach that lends itself well to scalability, future large scale studies in the field, and easy applicability in clinical laboratory settings and doctor's offices. Genes that were differentially expressed in low mood vs. high mood subjects were compared with: 1) the results of animal model brain and blood data, as well as 2) human genetic linkage/association data, and 3) human postmortem brain data, as a way of cross-validating the findings, prioritizing them, and identifying a short list of high probability biomarker genes (FIGS. **2**A and **3**).

[0069] A focused approach was used to analyze discrete quantitative phenotypic item (phene)—a Visual-Analog Scale (VAS) for mood. This approach avoids the issue of corrections for multiple comparisons that would arise if one were to look in a discovery fashion at multiple phenes in a comprehensive phenotypic battery (PhenoChipping) changed in relationship with all genes on a GeneChip microarray. Larger sample cohorts would be needed for the latter approach.

[0070] A panel of a subset of top candidate biomarker genes for mood state identified by the approach described herein was then used to generate a prediction score for mood state (low mood vs. high mood). This prediction score was compared to the actual self-reported mood scores from bipolar subjects in the primary cohort (FIG. **4**). This panel of mood biomarkers and prediction score were also examined in a separate independent cohort of psychotic disorders patients for which gene expression data and mood state data (FIG. **5**) were obtained, as well as in a second independent bipolar cohort (FIG. **6**).

[0071] Sample size for human subjects (n=29 for the primary bipolar cohort, n=30 for the psychotic disorders cohort, n=19 for the secondary bipolar cohort) is comparable to the size of cohorts for human postmortem brain gene expression studies in the field. Live donor blood samples instead of postmortem donor brains were studied, with the advantage of better phenotypic characterization, more quantitative state information, and less technical variability. This approach also permits repeated intra-subject measures when the subject is in different mood states.

[0072] The experimental approach for detecting gene expression changes relies on a well-established methodology—oligonucleotide microarrays. To avoid the possibility that some of the gene expression changes detected from a single biological experiment are biological or technical artifacts, the analyses described herein were designed to minimize the likelihood of having false positives, even at the expense of potentially having false negatives, due to the high cost in time and resources of pursuing false leads.

[0073] For the animal model work, using isogenic mouse strain affords a suitable control baseline of saline injected animals for the drug-injected animals. Three independent de novo biological experiments were performed, at different times, with different batches of mice. This overall design is geared to factor out both biological and technical variability. It is to be noted that the concordance between reproducible microarray experiments using the latest generations of oligonucleotide microarrays and other methodologies such as quantitative PCR, with their own attendant technical limitations, is estimated to be over 90%. For the human blood samples differential gene expression analyses, which are the results of single biological experiments, it has to be noted that the approach described herein used a very restrictive and technically robust, all or nothing induction of gene expression (change from Absent Call (A) to Present Call (P)). It is possible that not all biomarker genes for mood will show this complete induction related to state, but rather some will show modulation in gene expression levels, and are thus missed by a stringent filtering approach. Moreover, given the genetic

heterogeneity and variable environmental exposure, it is possible, indeed likely, that not all subjects will show changes in all the biomarker genes.

[0074] To identify candidate biomarker genes, two stringency thresholds were used: changes in 75% of subjects, and in 60% of subjects with low mood vs. high mood. Moreover, the approach, as described herein, is predicated on the existence of multiple cross-validators for each gene that is called a candidate biomarker (FIG. **2**A): 1) is it changed in human blood, 2) is it changed in animal model brain, 3) is it changed in animal model blood, 4) is it changed in postmortem human brain, and 5) does it map to a human genetic linkage locus. All these lines of evidence are the result of independent experiments. The virtues of this networked Bayesian approach are that, if one or another of the nodes (lines of evidence) becomes questionable/non-functional upon further evidence in the field, the network is resilient and maintains functionality. Additional lines of evidence may move certain genes in the prioritization scoring. Using approaches described herein, a small number of genes were identified and prioritized as top biomarkers, out of the over 40,000 transcripts (about half of which are detected as Present in each subject) measured by the microarrays that were used.

[0075] A validation of the novel and non-obvious approach described herein is the fact that the biomarker panel showed sensitivity and specificity, of a comparable nature, in both independent replication cohorts (psychotic disorder cohort and secondary bipolar cohort). Thus, the approach of using a visual analog scale phene reflecting an internal subjective experience of well being or distress (as opposed to more complex and disease specific state/trait clinical instruments), and looking at extremes of state combined with robust differential expression based on A/P calls, and Convergent Functional Genomics prioritization, is able to identify state biomarkers for mood, that are, at least in part, independent of specific diagnoses or medications. Nevertheless, a comparison with existing clinical rating scales (FIG. **6**), actimetry and functional neuroimaging, as well as analysis of biomarker data using such instruments may be of interest, as a way of delineating state vs. trait issues, diagnostic boundaries or lack thereof, and informing the design of clinical trials that may incorporate clinical and biomarker measures of response to treatment.

[0076] Human blood gene expression changes may be influenced by the presence or absence of both medications and drugs of abuse. While access to the subject's medical records was available and clinical information as part of the informed consent for the study, medication non-compliance, on the one hand, and substance abuse, on the other hand, are not infrequent occurrences in psychiatric patients. That medications and drugs of abuse may have effects on mood state and gene expression is in fact being partially modeled in the mouse pharmacogenomic model, with valproate and methamphetamine treatments respectively. The association of blood biomarkers with mood state is analyzed, regardless of the proximal causes, which could be diverse (see FIG. **2**B). The performance the biomarkers identified herein can also be analyzed at a protein level, in larger cohorts of both genders, in different age groups, and in theragnostic settings—measuring responses to specific treatments/medications.

[0077] A subset of top candidate biomarker genes include five genes involved in myelination (Mbp, Edg2, Mag, Pmp22 and Ugt8), six genes involved in growth factor signaling (Fgfr1, Fzd3, Erbb3, Igfbp4, Igfbp6), one gene involved in light transduction (PDE6D), and one gene involved in neurofilaments (Nefh). These genes were selected as having a line of evidence (CFG) score of 4 or higher (Table 3). That means, in addition to the human blood data, these genes have at least two other independent lines of evidence implicating them in mood disorders and/or concordance of expression in human brain and blood. Using this cutoff score, about 13 genes (FIG. **3**), all of which have evidence of differential expression in human postmortem brains from mood disorder patients.

[0078] It is intriguing that genes which have a well-established role in brain functioning may show changes in blood in relationship to psychiatric symptoms state (FIG. **3**, Table 3 and Table 7), and moreover that the direction of change may be concordant with that found in human postmortem brain studies. It is possible that trait promoter sequence mutations or epigenetic modifications influence expression in both tissues (brain and blood), and that state dependent transcription factor changes that modulate the expression of these genes may be contributory as well.

[0079] The data provided herein demonstrate that genes involved in brain infrastructure changes (myelin, growth factors) are prominent players in mood disorders, and are reflected in the blood profile. Myelin abnormalities have emerged as a common if perhaps non-specific denominator across a variety of neuropsychiatric disorders. For example, Mbp, is a top scoring candidate biomarker (FIG. **3**), associated with high mood state. The data provided herein regarding insulin growth factor signaling changes may provide an underpinning for the co-morbidity with diabetes and metabolic syndrome often encountered in mood disorder patients. These changes may be etiopathogenic, compensatory mechanisms, side-effects of medications, or results of illness—induced lifestyle changes (FIG. **2**B).

[0080] The fact that many of the biomarkers identified are associated with a low mood state (depression) as opposed to high mood state (FIG. **3** and Table 3) indicates that depression may have more of an impact on blood gene expression changes, perhaps through a neuro-endocrine-immunological axis, as part of a whole-body reaction to a perceived hostile environment.

[0081] Some of the candidate biomarker genes identified herein have no previous evidence for involvement in mood disorders (Tables 3 and 7). They merit further exploration in genetic candidate gene association studies, as well as comparison with emerging results from whole—genome association studies of bipolar disorder and depression. If needed, the composition of biomarker panels for mood can be refined or changed for different sub-populations, depending upon the availability of additional evidence. Panels containing different number of biomarkers and different biomarkers can be developed using the guidelines described herein and from the biomarkers identified herein. A large number of the biomarkers that would be of use in different panels and permutations are already present in the complete list of candidate biomarker genes identified (Tables 3 and 7).

[0082] An interrogation of a connectivity map with a signature query composed of the genes in a panel of top biomarkers for low mood and high mood revealed that sodium phenylbutyrate exerts the most similar effects to high mood, and novobiocin the most similar effects to low mood (FIG. **5**). Sodium phenylbutirate is a medication used to treat hyperammonemia that also has histone deacetylase (HDAC) properties, cell survival and anti-apoptotic effects. The mood sta-

bilizer drug valproate, also a HDAC inhibitor, as well as sodium phenylbutirate and another HDAC inhibitor, trichostatin A, were shown to induce alpha-synuclein in neurons through inhibition of HDAC and that this alpha-synuclein induction was critically involved in neuroprotection against glutamate excitotoxicity. Human postmortem brain studies, as well as animal model and clinical studies have implicated glutamate abnormalities and histone deacetylase modulation as therapeutic targets in mood disorders. Novobiocin is an antibiotic drug that also has anti-tumor activity and apoptosis-inducing properties, through Hsp90 inhibition of Akt kinase an effect opposite to that of the valproate, trichostatin A and sodium phenylbutyrate (Table 6).

[0083] This connectivity map analysis with a mood panel genes provides an interesting external biological cross-validation for the internal consistency of the biomarker approach, as well as illustrates the utility of the connectivity map for non-hypothesis driven identification of novel drug treatments and interventions.

[0084] The results provided herein are consistent with a trophicity model for genes involved in mood regulation: cell survival and proliferation associated with high mood, and cell shrinkage and death associated with low mood. This perspective is both of evolutionary interest and pragmatic clinical importance. Nature may have selected primitive cellular mechanisms for analogous higher organism level-functions: survival and expansion in favorable, mood-elevating environments, withdrawal and death (apoptosis) in unfavorable, depressogenic environments. In this view, suicide is the organismal equivalent of cellular apoptosis (programmed cell death). Pragmatically, the results point to an unappreciated molecular and therapeutic overlap between two broad areas of medicine: mood disorders and cancer. This overlap is relevant for the clinical co-morbidity of mood disorders and cancer, as well as for empirical studies to evaluate the use of mood-regulating drugs in cancer, and of cancer drugs in mood disorders.

[0085] In clinical practice there is a high degree of overlap and co-morbidity between mood disorders, psychosis and substance abuse. The data in bipolar and psychotic disorder cohorts point to the issue of heterogeneity, overlap and interdependence of major psychiatric syndromes as currently defined by DSM-IV, and the need for a move towards comprehensive empirical profiling and away from categorical diagnostic classifications.

[0086] There are to date no reliable clinical laboratory blood tests for mood disorders. A translational convergent approach is disclosed herein to identify and prioritize blood biomarkers of mood state. Data demonstrate that blood biomarkers are effective in offering an unexpectedly informative window into brain functioning and disease state. Panels of such biomarkers may serve as a basis for objective clinical laboratory tests, a longstanding unmet need for psychiatry. Biomarker-based tests are extremely valuable for early intervention and prevention efforts, as well as monitoring response to various treatments. In conjunction with other relevant clinical information, biomarker tests play a desirable part of personalizing treatment to increase effectiveness and avoid adverse reactions—personalized medicine in psychiatry. Moreover, the biomarkers identified herein are useful for identifying or screening new neuropsychiatric drugs, at both a pre-clinical and clinical (Phase I, II and III) stages of the process.

[0087] Because brain is a highly specialized organ, it is not expected that the genes expressed in the brain would be present in the blood. Expression products of genes (e.g., RNA and protein) are generally tissue specific and are not expected or predicted to be expressed in an unrelated tissue, e.g., blood. Therefore, the finding that certain markers are expressed in blood and are predictable for mood disorder in patients is surprising and non-obvious. Not all markers differentially expressed in blood and are associated with predicting/diagnosing mood disorder are expressed in the brain. Similarly, not all genes that are differentially expressed in brain are expressed in blood for predicting/diagnosing mood disorder.

[0088] Human postmortem brain gene expression studies are generally susceptible to the issue of being underpowered, due to uncertainty of diagnosis and difficulty of accrual of large well-characterized cohorts, as well as due to genetic heterogeneity and the effect of variable environmental exposure on gene expression. Moreover, postmortem work artifacts (agonal interval, pH and tissue degradation) may influence gene expression changes.

[0089] For example, the data presented herein has not found reliable blood evidence for some of the top candidate genes derived from postmortem work, such as: Gria1 (glutamate receptor, ionotropic, AMPA1 (alpha 1)), Grik1 (glutamate receptor, ionotropic, kainate 1), Gsk3b (glycogen synthase kinase 3 beta) and Arnt1 aryl hydrocarbon receptor nuclear translocator-like. Conversely, some of the top blood biomarkers identified herein do not appear to have reliable human postmortem brain evidence to date: Btg1 (B-cell translocation gene 1, anti-proliferative), Ednrb (endothelin receptor type B), Elovl5 (ELOVL family member 5, elongation of long chain fatty acids), and Trpc1 (transient receptor potential cation channel, subfamily C, member 1).

[0090] A plurality of high probability blood candidate biomarker genes for mood state is identified. In an aspect, a select panel of biomarkers include for example, five genes involved in myelination (Mbp, Edg2, Mag, Pmp22 and Ugt8), six genes involved in growth factor signaling (Fgfr1, Fzd3, Erbb3, Igfbp4, Igfbp6), one gene involved in light transduction (PDE6D), and one gene involved in neurofilaments (Nefh). These genes have evidence of differential expression in human postmortem brains from mood disorder patients.

[0091] A predictive score developed based on a panel of 10 top candidate biomarkers, designated herein as BioM-10 (5 for high mood, 5 for low mood) shows specificity and sensitivity for high mood and low mood states.

[0092] A parallel profiling of cognitive and affective state was performed to investigate: (i) relationships between phenotypic items ("phenes"), including with objective motor measures, and (ii) relationships between subjects. This approach is useful in advancing current diagnostic classifications, and indicates that a combinatorial building-block structure underlies many psychiatric syndromes. The adaptation of microarray-based informatic tools for phenotypic analysis facilitates direct integration with gene expression profiling of blood in the same individuals, a strategy for molecular biomarker identification. Empirically derived clusterings of (endo) phenotypes and of patients provide a basis for genetic, pharmacological, and imaging research, as well as clinical practice.

[0093] In an aspect, some of the candidate genes included in a panel of biomarkers used herein, have no previous evidence for involvement in mood disorders other than being mapped to bipolar genetic linkage loci (Table 3). These genes

constitute novel candidate genes for bipolar disorder and depression. The composition of biomarker panels for mood can be refined or changed for different sub-populations. Panels containing different number of biomarkers and different biomarkers can be developed using the guidelines described herein and from the complete list of more than 600 biomarkers identified (Tables 3 and 7).

[0094] Any number of biomarkers can be used as a panel for diagnosis. The panel may contain equal number of biomarkers for high and low mood, or different number of biomarkers associated with low mood than high mood. The panel may be tested as a microarray or as any form of diagnostic analysis.

[0095] In the present disclosure, gene expression changes in specific brain regions and blood of animal models developed were studied to identify one or more of the biomarkers disclosed herein. Data were obtained from a pharmacogenomic mouse model of bipolar (involving treatments with a stimulant—methamphetamine, and a mood stabilizer—valproate) as a discovery engine and cross-validator for the identification of potential peripheral blood biomarkers (see Ogden et al., (2004), Candidate genes, pathways and mechanisms for bipolar (manic-depressive) and related disorders: an expanded convergent functional genomics approach. *Mol Psychiatry* 9(11): 1007-29.). Data from other animal models of bipolar disorder, such as genetic models, can be used (see Le-Niculescu et al. (2008) Phenomic, convergent functional genomic and biomarker studies in a stress-reactive genetic animal model of bipolar disorder and co-morbid alcoholism. *American Journal of Medical Genetics Part B* (*Neuropsychiatric Genetics*) 147(2):134-66.

[0096] In an embodiment, a comprehensive analysis of: (i) fresh human blood gene expression data tied to illness state (quantitative measures of symptoms), (ii) cross-validation of blood gene expression profiling in conjunction with brain gene expression studies in animal models presenting key features of bipolar disorder, and (iii) integration of the results in the context of the available human genetic linkage/association and postmortem brain findings in the field is provided.

[0097] In an aspect, human blood gene expression studies were carried out in a primary group of bipolar subjects with low mood states and high mood states, as well as in a group of subjects with psychotic disorders (schizoaffective disorder, schizophrenia, and substance induced psychotic disorder), and in a second, independent, group of subjects with bipolar disorder. Genes that were differentially expressed in low mood vs. high mood subjects were compared with (i) the results of animal model data, (ii) human genetic linkage/association data and (iii) human postmortem brain data to cross-validate the results, prioritizing the genes, and identifying a short list of high probability candidate biomarker genes. A panel of candidate biomarker genes identified by this approach was then used to generate a prediction score for mood state (low mood/depression vs. high mood/mania). This prediction score was compared to the actual self-reported mood scores from human subjects. The prediction score developed by the analysis of convergent data provided a highly correlative basis for the diagnosis of mood state.

[0098] In an embodiment, a panel of biomarkers illustrated in Table 3 is suitable. These biomarkers include Mbp, Edg2, Fgfr1, Fzd3, Mag, Pmp22, Ugt8, Erbb3, Igfbp4, Igfbp6, Pde6d, Ptprm, Nefh, Atp2c1, Atxn1, Btg1, C6orf182, Dicer1, Dnajc6, Ednrb, Elovl5, Gnal, Klf5, Lin7a, Manea, Nupl1, Pde6b, Slc25a23, Synpo, Tgm2, Tjp3, Tpd52, Trpc1, Bclaf1, Gosr2, Rdx, Wdr34, Bic, C8orf42, Dock9, Hrasls, Ibrdc2, P2ry12, Specc1, Vil2.

[0099] In an embodiment, a panel of about 10 biomarkers, e.g., Mbp, Edg2, Fgfr1, Fzd3, Mag, Pmp22, Ugt8, Erbb3, Igfbp4, and Igfbp6, is suitable for diagnosing or predicting mood disorder.

[0100] In an embodiment, a panel of biomarkers include for example, Mbp, Edg2, Fzd3, Atxn1, and Ednrb that are increased in high mood (mania) condition.

[0101] An embodiment of a first sub-group of markers that are used for analysis include for example: Mbp, Edg2, Fzd3, Atxn1, Ednrb (markers for high mood) and Fgfr1, Mag, Pmp22, Ugt8, Erbb3 (markers for low mood). An embodiment includes a second sub-group e.g., Pde9a, Plxnd1, Camk2d, Dio2, Lepr (markers for high mood) and Igfbp4, Igfbp6, Pde6d, Ptprm, Nefh Atp2c1 (markers for low mood). An embodiment includes a third sub-group e.g., Myom2, Nfix, Nt5m, Or7e104p, Rrp1 (markers for high mood) and Atp2c1, Btg1, Elov5, Lrrc8b, Dicer1, Dnajc6 (markers for low mood). An embodiment includes a fourth sub-group e.g., Sept2, Sfrs4, Sla2, Tex261, Ube2i (markers for high mood) and Gnal, Klf5, Lin7a, Manea, Nupl1 (markers for low mood). An embodiment includes a fifth sub-group e.g., Usp7, Zdhhc4, Znf169, Cuedc1, Bivm (markers for high mood) and Pde6b, Slc25a23, Synpo, Tgm2, Tjp3 (markers for low mood). An embodiment includes a sixth sub-group e.g., Hla-dqa1, C20orf94, C21orf56, Flj10986, Loc91431 (markers for high mood), Tpd52, Trpc1, Phlda1, Znf502, Amn (markers for low mood) or a combination of one or more of the sub-groups 1-6 disclosed herein. Sub-groups 1-5 constitute a representative example and any number of sub-groups that has about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, or more markers selected from Table 7.

[0102] An embodiment of a first sub-group of markers that are used for analysis include for example: Mbp, Edg2, Fzd3, Atxn1, Ednrb (markers for high mood), Fgfr1, Mag, Pmp22, Ugt8, Erbb3 (markers for low mood); second subgroup includes for example: Pde9a, Plxnd1, Camk2d, Dio2, Lepr (markers for high mood), Igfbp4, Igfbp6, Pde6d, Ptprm, Nefh, (markers for low mood); third subgroup includes for example: Myom2, Nfix, Nt5m, Or7e104p, Rrp1 (markers for high mood), Btg1, Elov5, Lrrc8b, Dicer1, Atp2c1, (markers for low mood); fourth subgroup includes for example: Sept2, Sfrs4, Sla2, Tex261, Ube2i (markers for high mood), Gnal, Klf5, Lin7a, Manea, Dnajc6 (markers for low mood); fifth subgroup includes for example: Usp7, Zdhhc4, Znf169, Cuedc1, Bivm (markers for high mood), Pde6b, Slc25a23, Synpo, Tgm2, Nupl1 (markers for low mood); and sixty subgroup includes for example: Hla-dqa1, C20orf94, C21orf56, Vil2, Loc91431 (markers for high mood), Tpd52, Trpc1, Phlda1, Tjp3, Amn (markers for low mood) or a combination of one or more of the sub-groups 1-6 disclosed herein. Sub-groups 1-5 constitute a representative example and any number of sub-groups that has about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, or more markers selected from Tables 3 and 7.

[0103] A panel of 36 biomarkers, as illustrated in an example described herein, is a suitable subset that is useful in diagnosing a mood disorder. Larger subsets that includes for example, 150, 200, 250, 300, 350, 400, 450, 500, 600 or about 700 markers are also suitable. Smaller subsets that include high-value markers including about 2, 5, 10, 15, 20, 25, 50,

75, and 100 are also suitable. A variable quantitative scoring scheme can be designed using any standard algorithm, such as a variable selection or a subset feature selection algorithms can be used. Both statistical and machine learning algorithms are suitable in devising a frame work to identify, rank, and analyze association between marker data and phenotypic data (e.g., mood disorders).

[0104] In an embodiment, a prediction score for each subject is derived based on the presence or absence of e.g., 10 biomarkers of the panel in their blood. Each of the 10 biomarkers gets a score of 1 if it is detected as "present" (P) in the blood form that subject, 0.5 if it is detected as "marginally present" (M), and 0 if it is called "absent" (A). The ratio of the sum of the high mood biomarker scores divided by the sum of the low mood biomarker scores is multiplied by 100, and provides a prediction score. If the ratio of high biomarker genes to low mood biomarker genes is 1, i.e. the two sets of genes are equally represented, the mood prediction score is 1×100=100. The higher this score, the higher the predicted likelihood that the subject will have high mood. The predictive score was compared with actual self-reported mood scores in a primary cohort of subjects with a diagnosis of bipolar mood disorder. A prediction score of 100 and above had a 84.6% sensitivity and a 68.8% specificity for predicting high mood. A prediction score below 100 had a 76.9% sensitivity and 81.3% specificity for predicting low mood. The term "present" indicates that a particular biomarker is expressed to a detectable level, as determined by the technique used. For example, in an experiment involving a microarray or gene chip obtained from a commercial vendor Affymetrix (Santa Clara, Calif.), the embedded software rendered a "present" call for that biomarker. The term "present" refers to a detectable presence of the transcript or its translated protein/peptide and not necessarily reflects a relative comparison to for example, a sample from a normal subject. In other words, the mere presence or absence of a biomarker is assigned a value, e.g., 1 and a prediction score is calculated as described herein. The term "marginally present: refers to border line expression level that may be less intense than the "present" but statistically different from being marked as "absent" (above background noise), as determined by the methodology used.

[0105] In an embodiment, a prediction score based on differential expression (instead of "present", "absent") is used. For example, if a subject has a plurality of markers for high or low mood are differentially expressed, a prediction based on the differential expression of markers is determined. Differential expression of about 1.2 fold or 1.3 or 1.5 or 2 or 3 or 4 or 5-fold or higher for either increased or decreased levels can be used. Any standard statistical tool such as ANOVA is suitable for analysis of differential expression and association with high or low mood diagnosis or prediction.

[0106] A prediction based on the analysis of either high or low mood markers alone (instead of a ratio of high versus low mood markers) may also be practiced. If a plurality of high mood markers (e.g., about 6 out of 10 tested) are differentially expressed to a higher level compared to the low mood markers (e.g., 4 out of 10 tested), then a prediction or diagnosis of high mood status can be made by analyzing the expression levels of the high mood markers alone without factoring the expression levels of the low mood markers as a ratio.

[0107] In an embodiment, a detection algorithm uses probe pair intensities to generate a detection p-value and assign a Present, Marginal, or Absent call. Each probe pair in a probe set is considered as having a potential vote in determining whether the measured transcript is detected (Present) or not detected (Absent). The vote is described by a value called the Discrimination score [R]. The score is calculated for each probe pair and is compared to a predefined threshold Tau. Probe pairs with scores higher than Tau vote for the presence of the transcript. Probe pairs with scores lower than Tau vote for the absence of the transcript. The voting result is summarized as a p-value. The greater the number of discrimination scores calculated for a given probe set that are above Tau, the smaller the p-value and the more likely the given transcript is truly Present in the sample. The p-value associated with this test reflects the confidence of the Detection call.

[0108] Regarding detection p-value, a two-step procedure determines the Detection p-value for a given probe set. The Discrimination score [R] is calculated for each probe pair and the discrimination scores are tested against the user-definable threshold Tau. The detection Algorithm assesses probe pair saturation, calculates a Detection p-value, and assigns a Present, Marginal, or Absent call. In an embodiment, the default thresholds of the Affymetrix MAS 5 software were used.

[0109] In spiking experiments by the manufacturer to establish default thresholds (adding of known quantities of test transcripts to a mixture, to measure the sensitivity of the Affymetrix MAS 5 detection algorithm) 80% of spiked transcripts are called Present at a concentration of 1.5 pM. This concentration corresponds to approximately one transcript in 100,000 or 3.5 copies per cell. The false positive rate of making a Present call was roughly 10%, as noted by 90% of the transcripts being called Absent when not spiked into the sample (0 pM concentration).

[0110] The term "predictive" or the term "prognostic" does not imply 100% predictive ability. The use of these terms indicates that subjects with certain characteristics are more likely to experience a particular mood state or clinical outcome than subjects who do not have such characteristics. For example, characteristics that determine the prediction include one or more of the biomarkers for the mood disorder disclosed herein. The phrase "clinical outcome" refers to biological or biochemical or physical or physiological responses to treatments or therapeutic agents that are generally prescribed for that condition compared to a condition would occur in the absence of any treatment. A "clinically positive outcome" does not necessarily indicate a cure, but could indicate a lessening of symptoms experienced by a subject.

[0111] The terms "marker" and "biomarker" are synonymous and as used herein, refer to the presence or absence or the levels of nucleic acid sequences or proteins or polypeptides or fragments thereof to be used for associating or correlating a phenotypic state. A biomarker includes any indicia of the level of expression of an indicated marker gene. The indicia can be direct or indirect and measure over- or under-expression of the gene given the physiologic parameters and in comparison to an internal control, normal tissue or another phenotype. Nucleic acids or proteins or polypeptides or portions thereof used as markers are contemplated to include any fragments thereof, in particular, fragments that can specifically hybridize with their intended targets under stringent conditions and immunologically detectable fragments. One or more markers may be related. Marker may also refer to a gene or DNA sequence having a known location on a chromosome and associated with a particular gene or trait. Genetic markers associated with certain diseases or for pre-

disposing disease states can be detected in the blood and used to determine whether an individual is at risk for developing a disease. Levels of gene expression and protein levels are quantifiable and the variation in quantification or the mere presence or absence of the expression may also serve as markers. Using proteins/peptides as biomarkers can include any method known in the art including, without limitation, measuring amount, activity, modifications such as glycosylation, phosphorylation, ADP-ribosylation, ubiquitination, etc., immunohistochemistry (IHC).

[0112] As used herein, "array" or "microarray" refers to an array of distinct polynucleotides, oligonucleotides, polypeptides, or oligopeptides synthesized on a substrate, such as paper, nylon, or other type of membrane, filter, chip, glass slide, or any other suitable solid support. Arrays also include a plurality of antibodies immobilized on a support for detecting specific protein products. There are several microarrays that are commercially available. A microarray may include one or more biomarkers disclosed herein. A panel of about 20 biomarkers as nucleic acid fragments can be included in an array. The nucleic acid fragments may include oligonucleotides or amplified partial or complete nucleotide sequences of the biomarkers. The term "consisting essentially of" generally refers to a collection of markers that substantially affects the determination of the disorder and may include other components such as controls. For example, a microarray consists essentially of markers from Table 3.

[0113] In an embodiment, the microarray is prepared and used according to the methods described in U.S. Pat. No. 5,837,832, Chee et al.; PCT application WO95/11995, Chee et al.; Lockhart et al., 1996. *Nat Biotech.,* 14:1675-80; and Schena et al., 1996. *Proc. Natl. Acad. Sci.* 93:10614-619, all of which are herein incorporated by reference to the extent they relate to methods of making a microarray. Arrays can also be produced by the methods described in Brown et al., U.S. Pat. No. 5,807,522. Arrays and microarrays may be referred to as "DNA chips" or "protein chips."

[0114] A variety of clustering methods are available for microarray-based gene expression analysis. See for example, Shamir & Sharan (2002) Algorithmic approaches to clustering gene expression data. In *Current Topics In Computational Molecular Biology* (Edited by: Jiang T, Xu Y, Smith T). 2002, 269-300; Tamames et al., (2002): Bioinformatics methods for the analysis of expression arrays: data clustering and information extraction, *J Biotechnol,* 98:269-283.

[0115] "Therapeutic agent" means any agent or compound useful in the treatment, prevention or inhibition of mood disorder or a mood-related disorder.

[0116] The term "condition" refers to any disease, disorder or any biological or physiological effect that produces unwanted biological effects in a subject.

[0117] The term "subject" refers to an animal, or to one or more cells derived from an animal. The animal may be a mammal including humans. Cells may be in any form, including but not limited to cells retained in tissue, cell clusters, immortalized cells, transfected or transformed cells, and cells derived from an animal that have been physically or phenotypically altered.

[0118] Any body fluid of an animal can be used in the methods of the invention. Suitable body fluids include a blood sample (e.g., whole blood, serum or plasma), urine, saliva, cerebrospinal fluid, tears, semen, and vaginal secretions. Also, lavages, tissue homogenates and cell lysates can be utilized.

[0119] Many different methods can be used to determine the levels of markers. For example, protein arrays, protein chips, cDNA microarrays or RNA microarrays are suitable. More specifically, one of ordinary skill in the art will appreciate that in one example, a microarray may comprise the nucleic acid sequences representing genes listed in Table 1. For example, functionality, expression and activity levels may be determined by immunohistochemistry, a staining method based on immunoenzymatic reactions uses monoclonal or polyclonal antibodies to detect cells or specific proteins. Typically, immunohistochemistry protocols include detection systems that make the presence of markers visible (to either the human eye or an automated scanning system), for qualitative or quantitative analyses. Mass-spectrometry, chromatography, real-time PCR, quantitative PCR, probe hybridization, or any other analytical method to determine expression levels or protein levels of the markers are suitable. Such analysis can be quantitative and may also be performed in a high-through put fashion. Cellular imaging systems are commercially available that combine conventional light microscopes with digital image processing systems to perform quantitative analysis on cells and tissues, including immunostained samples. (See e.g. the CAS-200 System (Becton, Dickinson & Co.)). Some other examples of methods that can be used to determine the levels of markers include immunohistochemistry, automated systems, quantitative IHC, semi-quantitative IHC and manual methods. Other analytical systems include western blotting, immunoprecipitation, fluorescence in situ hybridization (FISH), and enzyme immunoassays.

[0120] The term "diagnosis", as used in this specification refers to evaluating the type of disease or condition from a set of marker values and/or patient symptoms where the subject is suspected of having a disorder. This is in contrast to disease predisposition, which relates to predicting the occurrence of disease before it occurs, and the term "prognosis", which is predicting disease progression in the future based on the marker levels/patterns.

[0121] The term "correlating," as used in this specification refers to a process by which one or more biomarkers are associated to a particular disease state, e.g., mood disorder. In general, identifying such correlation or association involves conducting analyses that establish a statistically significant association- and/or a statistically significant correlation between the presence (or a particular level) of a marker or a combination of markers and the phenotypic trait in the subject. An analysis that identifies a statistical association (e.g., a significant association) between the marker or combination of markers and the phenotype establishes a correlation between the presence of the marker or combination of markers in a subject and the particular phenotype being analyzed.

[0122] This relationship or association can be determined by comparing biomarker levels in a subject to levels obtained from a control population, e.g., positive control—diseased (with symptoms) population and negative control—disease-free (symptom-free) population. The biomarkers disclosed herein provide a statistically significant correlation to diagnosis at varying levels of probability. Subsets of markers, for example a panel of about 20 markers, each at a certain level range which are a simple threshold, are said to be correlative or associative with one of the disease states. Such a panel of correlated markers can be then be used for disease detection, diagnosis, prognosis and/or treatment outcome. Preferred methods of correlating markers is by performing marker

selection by any appropriate scoring method or by using a standard feature selection algorithm and classification by known mapping functions. A suitable probability level is a 5% chance, a 10% chance, a 20% chance, a 25% chance, a 30% chance, a 40% chance, a 50% chance, a 60% chance, a 70% chance, a 75% chance, a 80% chance, a 90% chance, a 95% chance, and a 100% chance. Each of these values of probability is plus or minus 2% or less. A suitable threshold level for markers of the present invention is about 25 pg/mL, about 50 pg/mL, about 75 pg/mL, about 100 pg/mL, about 150 pg/mL, about 200 pg/mL, about 400 pg/mL, about 500 pg/mL, about 750 pg/mL, about 1000 pg/mL, and about 2500 pg/mL.

[0123] Prognosis methods disclosed herein that improve the outcome of a disease by reducing the increased disposition for an adverse outcome associated with the diagnosis. Such methods may also be used to screen pharmacological compounds for agents capable of improving the patient's prognosis, e.g., test agents for mood disorders.

[0124] The analysis of a plurality of markers, for example, a panel of about 20 or 10 markers may be carried out separately or simultaneously with one test sample. Several markers may be combined into one test for efficient processing of a multiple of samples. In addition, one skilled in the art would recognize the value of testing multiple samples (for example, at successive time points) from the same individual. Such testing of serial samples may allow the identification of changes in marker levels over time, within a period of interest, or in response to a certain treatment.

[0125] In another embodiment, a kit for the analysis of markers includes for example, devises and reagents for the analysis of at least one test sample and instructions for performing the assay. Optionally, the kits may contain one or more means for using information obtained from marker assays performed for a marker panel to diagnose mood disorders. Probes for markers, marker antibodies or antigens may be incorporated into diagnostic assay kits depending upon which markers are being measured. A plurality of probes may be placed in to separate containers, or alternatively, a chip may contain immobilized probes. In an embodiment, another container may include a composition that includes an antigen or antibody preparation. Both antibody and antigen preparations may preferably be provided in a suitable titrated form, with antigen concentrations and/or antibody titers given for easy reference in quantitative applications.

[0126] The kits may also include a detection reagent or label for the detection of specific reaction between the probes provided in the array or the antibody in the preparation for immunodetection. Suitable detection reagents are well known in the art as exemplified by fluorescent, radioactive, enzymatic or otherwise chromogenic ligands, which are typically employed in association with the nucleic acid, antigen and/or antibody, or in association with a secondary antibody having specificity for first antibody. Thus, the reaction is detected or quantified by means of detecting or quantifying the label. Immunodetection reagents and processes suitable for application in connection with the novel methods of the present invention are generally well known in the art.

[0127] The reagents may also include ancillary agents such as buffering agents and protein stabilizing agents, e.g., polysaccharides and the like. The diagnostic kit may further include where necessary agents for reducing background interference in a test, agents for increasing signal, software and algorithms for combining and interpolating marker values to produce a prediction of clinical outcome of interest, apparatus for conducting a test, calibration curves and charts, standardization curves and charts, and the like.

[0128] In some embodiments, the methods of correlating biomarkers with treatment regimens can be carried out using a computer database. Computer-assisted methods of identifying a proposed treatment for mood disorders are suitable. The method involves the steps of (a) storing a database of biological data for a plurality of patients, the biological data that is being stored including for each of said plurality of patients (i) a treatment type, (ii) at least one marker associated with a mood disorder and (iii) at least one disease progression measure for the mood disorder from which treatment efficacy can be determined; and then (b) querying the database to determine the dependence on the marker of the effectiveness of a treatment type in treating the mood disorder, to thereby identify a proposed treatment as an effective treatment for a subject carrying the marker correlated with the mood disorder.

[0129] In an embodiment, treatment information for a patient is entered into the database (through any suitable means such as a window or text interface), marker information for that patient is entered into the database, and disease progression information is entered into the database. These steps are then repeated until the desired number of patients has been entered into the database. The database can then be queried to determine whether a particular treatment is effective for patients carrying a particular marker, not effective for patients carrying a particular marker, and the like. Such querying can be carried out prospectively or retrospectively on the database by any suitable means, but is generally done by statistical analysis in accordance with known techniques, as described herein.

EXAMPLES

[0130] The following examples are to be considered as exemplary and not restrictive or limiting in character and that all changes and modifications that come within the spirit of the disclosure are desired to be protected.

Example 1

Experimental Framework for Identification of Biomarkers Used in Diagnosis of Mood Disorders

[0131] Gene expression changes in specific brain regions and blood from a pharmacogenomic animal model were used as cross-validators for identification of potential human blood biomarkers. Pharmacogenomic mouse model of relevance to bipolar disorder consists of treatments with an agonist of the illness/bipolar disorder-mimicking drug (methamphetamine) and an antagonist of the illness/bipolar disorder-treating drug (valproate). The pharmacogenomic approach is a tool for tagging genes that may have pathophysiological relevance.

[0132] Human blood gene expression studies were carried out in a cohort of bipolar subjects. Genes that were differentially expressed in low mood vs. high mood subjects were compared with: 1) the results of the animal model brain and blood data, as well as 2) published human genetic linkage/association data, and 3) human postmortem brain data, as a way of cross-validating the findings, prioritizing them, and coming up with a short list of high probability candidate biomarker genes (FIGS. **2**A and **3**).

[0133] A Visual-Analog Scale (VAS) for mood was used for the scoring analysis. This approach avoids the issue of corrections for multiple comparisons that would arise if multiple symptom (phenotypic) scores (i.e. "phenes") were analyzed in a comprehensive phenotypic battery changed in relationship with all genes on a GeneChip microarray. Larger sample cohorts would be needed for the latter approach.

[0134] A panel of top candidate biomarker genes for mood state identified was then used to generate a prediction score for mood state (low mood vs. high mood). This prediction score was compared to the actual self-reported mood scores from bipolar subjects (FIG. **4**). This panel of mood biomarkers and prediction score were examined in a separate independent cohort of psychotic disorders patients for which gene expression data and mood state data is available (FIG. **5**), and in a second independent cohort of bipolar disorder subjects (FIG. **6**).

[0135] Sample size for human subjects (n=29 for the bipolar cohort, n=30 for the psychotic disorders cohort) is comparable to the size of cohorts for human postmortem brain gene expression studies. Live donor blood samples were used instead of postmortem donor brains, with the advantage of better phenotypic characterization, more quantitative state information, and less technical variability.

[0136] For the animal model work, isogenic mouse strain was used. Three independent de novo biological experiments were performed, at different times, with different batches of mice. This overall design is geared to factor out both biological and technical variability. Concordance between reproducible microarray experiments using the latest generations of oligonucleotide microarrays and other methodologies such as quantitative PCR, with their own attendant technical limitations, is estimated to be over 90%. For the human blood samples gene expression analyses, which are the results of single biological experiments, a very restrictive, all or nothing induction of gene expression (change from Absent Call to Present Call). It is possible that not all biomarker genes for mood may show this complete induction related to state, but rather some may show modulation in gene expression levels, and are thus missed by this filtering. Moreover, given the genetic heterogeneity and variable environmental exposure, it is possible, indeed likely, that not all subjects may show changes in all the biomarker genes. Hence two stringency thresholds were used: changes in 75% of subjects, and in 60% of subjects with low mood vs. high mood. This approach is predicated on the existence of multiple cross-validators for each gene that is called a candidate biomarker (FIG. **2**B): 1) is it changed in human blood, 2) is it changed in animal model brain, 3) is it changed in animal model blood, 4) is it changed in postmortem human brain, and 5) does it map to a human genetic linkage locus. All these lines of evidence are the result of independent experiments.

[0137] Human blood gene expression changes may be influenced by the presence or absence of both medications and drugs of abuse. That medications and drugs of abuse may have effects on mood state and gene expression is in fact being partially modeled in the mouse pharmacogenomic model, with valproate and methamphetamine treatments respectively. It is the association of blood biomarkers with mood state that is the primary purpose of this analysis, regardless of the proximal causes, which could be diverse (see FIG. **2**B).

[0138] The human subjects used in this example included those who were directly recruited, and data collected in other procedures/settings. Blood samples were collected.

[0139] Human data from three independent cohorts of patients are presented. One cohort consists of 29 subjects with bipolar I disorder. The second cohort consists of 30 subjects with psychotic disorders (schizophrenia, schizoaffective disorder, and substance induced psychosis). The third cohort consists of 19 subjects with bipolar I disorder. The diagnosis is established by a structured clinical interview—Diagnostic Interview for Genetic Studies (DIGS), which has details on the course of illness and phenomenology, and is the scale used by the Genetics Initiative Consortia for both Bipolar Disorder and Schizophrenia.

[0140] Subjects included men and women over 18 years of age. A demographic breakdown is shown in Table 1. Initial studies were focused primarily on a male population, due to the demographics of the catchment area (primarily male in a VA Medical Center), and to minimize any potential gender-related state effects on gene expression, which would have decreased the discriminative power of the analysis for the sample size used. Subjects were recruited from the general population, the patient population at the IU school of Medicine, the Indianapolis VA Medical Center, as well as various facilities that serve people with mental illnesses in Indiana. The subjects were recruited largely through referrals from care providers, the use of brochures left in plain sight in public places and mental health clinics, and through word of mouth. Subjects were excluded if they had significant medical or neurological illness or had evidence of active substance abuse or dependence. All subjects understood and signed informed consent forms before assessments began. All subjects signed an informed consent form detailing the research goals, procedure, caveats and safeguards. Subjects completed diagnostic assessments (DIGS), and then a visual-analog scale for mood (VAS Mood) at the time of blood draw.

[0141] Human Blood Gene Expression Experiments and Analysis:

[0142] RNA extraction: 2.5-5 ml of whole blood was collected into each PaxGene tube by routine venipuncture. PaxGene tubes contain proprietary reagents for the stabilization of RNA. The cells from whole blood will be concentrated by centrifugation, the pellet washed, resuspended and incubated in buffers containing Proteinase K for protein digestion. A second centrifugation step will be done to remove residual cell debris. After the addition of ethanol for an optimal binding condition the lysate is applied to a silica-gel membrane/column. The RNA bound to the membrane as the column is centrifuged and contaminants are removed in three wash steps. The RNA is then eluted using DEPC-treated water.

[0143] Globin reduction: To remove globin mRNA, total RNA from whole blood is mixed with a biotinylated Capture Oligo Mix that is specific for human globin mRNA. The mixture is then incubated for 15 min to allow the biotinylated oligonucleotides to hybridize with the globin mRNA. Streptavidin Magnetic Beads are then added, and the mixture is incubated for 30 min. During this incubation, streptavidin binds the biotinylated oligonucleotides, thereby capturing the globin mRNA on the magnetic beads. The Streptavidin Magnetic Beads are then pulled to the side of the tube with a magnet, and the RNA, depleted of the globin mRNA, is transferred to a fresh tube. The treated RNA is further purified using a rapid magnetic bead-based purification method. This

consists of adding an RNA Binding Bead suspension to the samples, and using magnetic capture to wash and elute the GLOBINclear RNA.

[0144] Sample Labeling: Sample labeling is performed using the Ambion MessageAmp II-BiotinEnhanced aRNA amplification kit. The procedure is briefly outlined herein and involves the following steps:

[0145] 1. Reverse Transcription to Synthesize First Strand cDNA is primed with the T7 Oligo(dT) Primer to synthesize cDNA containing a T7 promoter sequence.

[0146] 2. Second Strand cDNA Synthesis converts the single-stranded cDNA into a double-stranded DNA (dsDNA) template for transcription. The reaction employs DNA Polymerase and RNase H to simultaneously degrade the RNA and synthesize second strand cDNA.

[0147] 3. cDNA Purification removes RNA, primers, enzymes, and salts that would inhibit in vitro transcription.

[0148] 4. In Vitro Transcription to Synthesize aRNA with Biotin-NTP Mix generates multiple copies of biotin-modified aRNA from the double-stranded cDNA templates; this is the amplification step.

[0149] 5. aRNA Purification removes unincorporated NTPs, salts, enzymes, and inorganic phosphate to improve the stability of the biotin-modified aRNA.

[0150] Microarrays: Biotin labeled aRNA are hybridized to Affymetrix HG-U133 Plus 2.0 GeneChips according to manufacturer's protocols (Affymetrix Inc., Santa Clara, Calif.). All GAPDH 3'/5' ratios should be less than 2.0 and backgrounds under 50. Arrays are stained using standard Affymetrix protocols for antibody signal amplification and scanned on an Affymetrix GeneArray 2500 scanner with a target intensity set at 250. Present/Absent calls are determined using GCOS software with thresholds set at default values.

[0151] The human blood gene expression experiments and analysis was performed at two levels: (i) high threshold >75%; 3× enrichment, and (ii) low threshold (>60%; 1.5× enrichment). The animal model data included pharmacogenomic models that involved DBP KO mouse.

[0152] The cross-validation and integration of data from human blood gene expression, mouse models and other mouse and human data were processed through a convergent functional genomics approach.

[0153] For generating the animal model data, standard pharmacogenomic testing methodologies were adopted. All experiments were performed with male C57/BL6 mice, 8 to 12 weeks of age, obtained from Jackson Laboratories (Bar Harbor, Me.), and acclimated for at least two weeks in an animal facility prior to any experimental manipulation. The bipolar pharmacogenomic model included Methamphetamine and Valproate treatments in mice (see Ogden et al. (2004)). Briefly, mice were treated by intraperitoneal injection with either single-dose saline, methamphetamine (10 mg/kg), valproate (200 mg/kg), or a combination of methamphetamine and valproate (10 mg/kg and 200 mg/kg respectively). Three independent de novo biological experiments were performed at different times. Each experiment included three mice per treatment condition, for a total of 9 mice per condition across the three experiments.

[0154] RNA extraction and microarray analysis: Standard techniques were used to obtain total RNA (22 gauge syringe homogenization in RLT buffer) and to purify the RNA (RNeasy mini kit, Qiagen, Valencia, Calif.) from micro-dissected mouse brain regions. For the human and whole mouse blood RNA extraction, PAXgene blood RNA extraction kit (PreAnalytiX, a QIAGEN/BD company) was used, followed by GLOBINclear™-Human or GLOBINclear™-Mouse/Rat (Ambion/Applied Biosystems Inc., Austin, Tex.) to remove the globin mRNA. All the methods and procedures were carried out as per manufacturer's instructions. The quality of the total RNA was confirmed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, Calif.). The quantity and quality of total RNA was also independently assessed by 260 nm UV absorption and by 260/280 ratios, respectively (Nanodrop spectrophotometer). Starting material of total RNA labeling reactions was kept consistent within each independent microarray experiment.

[0155] For the mouse analysis, blood or brain tissue regions from 3 mice were pooled for each experimental condition, and equal amounts of total RNA extracted from tissue samples or blood was used for labeling and microarray assays. Mouse Genome 430 2.0 arrays (Affymetrix, Santa Clara, Calif.) were used. The GeneChip™ Mouse Genome 430 2.0 Array contain over 45,000 probe sets that analyze the expression level of over 39,000 transcripts and variants from over 34,000 well-characterized mouse genes. For the human analysis, Affymetrix Human Genome U133 Plus 2.0 GeneChip with over 40,000 genes and ESTs were used. Standard manufacturer's protocols were used to reverse transcribe the messenger RNA and generate biotinlylated cRNA. The amount of cRNA used to prepare the hybridization cocktail was kept constant intra-experiment. Samples were hybridized at 45° C. for 17 hours under constant rotation. Arrays were washed and stained using the Affymetrix Fluidics Station 400 and scanned using the Affymetrix Model 3000 Scanner controlled by GCOS software. All sample labeling, hybridization, staining and scanning procedures were carried out as per manufacturer's recommendations.

[0156] Quality control: All arrays were scaled to a target intensity of 1000 using Affymetrix MASv 5.0 array analysis software. Quality control measures including 3'/5' ratios for GAPDH and beta-actin, scaling factors, background, and Q values were within acceptable limits.

[0157] Microarray data analysis: Data analysis was performed using Affymetrix Microarray Suite 5.0 software (MAS v5.0). Default settings were used to define transcripts as present (P), marginal (M), or absent (A). For the pharmacogenomic mouse model, a comparison analysis was performed for each drug treatment, using its corresponding saline treatment as the baseline. "Signal," "Detection," "Signal Log Ratio," "Change," and "Change p-value," were obtained from this analysis. Only transcripts that were called Present in at least one of the two samples (saline or drug) intra-experiment, and that were reproducibly changed in the same direction in at least two out of three independent experiments, were analyzed further. For the DBP knock-out mice, a comparison analysis was performed for each KO saline and KO Meth mouse, using WT saline mice as the baseline "Signal," "Detection," "Signal Log Ratio," "Change," and "Change p-value," were obtained from this analysis. Only transcripts that were called Present in at least one of the two samples in a comparison pair, and that were reproducibly changed in the same direction in at least six out of 9 comparisons, were analyzed further.

Example 2

Analysis and Identification of Biomarkers

[0158] Gene expression profiling studies were performed with peripheral whole blood samples from a primary cohort

of 29 human subjects with bipolar I disorder (27 males, 2 females) (Table 1). 13 had low self-reported mood scores (below 40) on the Visual-Analog (VAS) Mood Scale (FIG. **1**), and 13 had high self-reported mood scores (above 60). 3 of them had intermediate mood scores (between 40 and 60). Their mood scores at time of blood collection were used as a way of narrowing the field and identifying candidate biomarker genes for mood. Only t all or nothing gene expression differences were identified by Absent (A) vs. Present (P) Calls in the Affymetrix MAS software. Genes whose expression is detected as Absent in the Low Mood subjects and detected as Present in the High Mood subjects were classified, as being candidate biomarker genes for elevated mood state (mania). Conversely, genes whose expression is detected as Present in the Low Mood subjects and Absent in the High Mood subjects are being classified as candidate biomarker genes for low mood state (depression) (Tables 2 and 3). It is possible that some of the genes associated with high mood state or low mood state may not necessarily be involved in the induction of that state, but rather in its suppression as part of a homeostatic regulatory networks or treatment response mechanisms (similar conceptually to oncogenes and tumor-suppressor genes).

[0159] Two thresholds for analysis of gene expression differences between low mood and high mood (Table 2) were undertaken. First, a high threshold was used, with at least 75% of subjects in a cohort showing a change in expression from Absent to Present between low and high mood (reflecting an at least 3 fold mood state related enrichment of the genes thus filtered) As psychiatric disorders are clinically and (likely) genetically heterogeneous, with different combinations of genes and biomarkers present in different subgroups, a low threshold was also used, with at least 60% of subjects in a cohort showing a change in expression from Absent to Present between low and high mood (reflecting an at least 1.5 fold mood state related enrichment of the genes thus filtered). The high threshold identified candidate biomarker genes that are more common for all subjects, with a lower risk of false positives, whereas the lower threshold identified genes that are present in more restricted subgroups of subjects, with a lower risk of false negatives. The high threshold candidate biomarker genes have, as an advantage, a higher degree of reliability, as they are present in at least 75% of all subjects with a certain mood state (high or low) tested. They may reflect common aspects related to mood disorders across a diverse subject population, but may also be a reflection of the effects of common medications used in the population tested, such as mood stabilizers. The low threshold genes may have lower reliability compared to the high threshold, being present in at least 60% of the subject population tested, but, nevertheless, captures more of the diversity of genes and biological mechanisms present in a genetically diverse human subject population.

[0160] By cross-validating with animal model and other human datasets (FIG. **2**A) using CFG, a shorter list of genes was identified for which there are external corroborating line of evidence (e.g., human genetic evidence, human postmortem brain data, animal model brain and blood data) linking them to mood disorders (bipolar disorder, depression), thus reducing the risk of false positives. This cross-validation identifies candidate biomarkers that are more likely directly related to the relevant neuropathology, as opposed to being potential artifactual effects or indirect effects of lifestyle, environment, etc.

[0161] Using the approach described herein, out of over 40,000 genes and ESTs on the Affymetrix Human Genome U133 Plus 2.0 GeneChip, by using the high threshold, in an embodiment, about 21 novel candidate biomarker genes (13 genes with known functions and 7 ESTs) (Table 3), of which 8 had at least one line of prior independent evidence for potential involvement in mood disorders (i.e. CFG score of 3 or above). In addition to the high threshold genes, by using the low threshold, a larger list totaling 661 genes (539 genes and 122 ESTs) (Table 7), of which an additional 24 had at least two lines of prior independent evidence for potential involvement in mood disorders (i.e. CFG score of 3 or above). Of interest, four of the low threshold candidate biomarker genes (Bclaf1 and Rdx8, Gosr2 and Wdr3413) are changed in expression in the same direction, in lymphoblastoid cell lines (LCLs) from bipolar subjects.

[0162] Making a combined list of all the high value candidate biomarker genes identified as described above—including the high threshold genes with at least on other external line of evidence (8) and of the additional low threshold genes with at least two other external lines of evidence (24), and the low threshold genes with prior LCL evidence (4), a list of 36 candidate biomarker genes for mood, prioritized based on CFG score (Table 3) was developed.

[0163] In an embodiment, selecting the 5 top scoring candidate biomarkers for high mood (MBP, EDG2, FZD3, ATXN1, EDNRB) and the 5 top scoring candidate biomarkers for low mood (FGFR1, MAG, PMP22, UGT8, ERBB3), a panel of 10 biomarkers for mood disorder was developed that has diagnostic and predictive value.

[0164] To test the predictive value of a panel (e.g, the BioM-10 Mood panel), a cohort of 29 bipolar disorder subjects, containing the 26 subjects (13 low mood, 13 high mood) from which the candidate biomarker data was derived, as well as 3 additional subjects with mood in the intermediate range (self-reported mood scores between 40 and 60) was used. A prediction score for each subject, based on the presence or absence of the 10 biomarkers of the panel in the blood GeneChip data. Each of the 10 biomarkers gets a score of 1 if it is detected as Present (P) in the blood form that subject, 0.5 if it is detected as Marginally Present (M), and 0 if it is called Absent (A). The ratio of the sum of the high mood biomarker scores divided by the sum of the low mood biomarker scores is multiplied by 100, and provides a prediction score. If the ratio of high biomarker genes to low mood biomarker genes is 1, i.e. the two sets of genes are equally represented, the mood prediction score is 1×100=100. The higher this score, the higher the predicted likelihood that the subject will have high mood. The predictive score was compared with actual self-reported mood scores in the primary cohort of subjects with a diagnosis of bipolar mood disorder (n=29). A prediction score of 100 and above had a 84.6% sensitivity and a 68.8% specificity for predicting high mood. A prediction score below 100 had a 76.9% sensitivity and 81.3% specificity for predicting low mood (Table 4A and FIG. **4**).

[0165] Table 5 shows a representative sample of biological roles based on ingenuity pathway analysis (IPA) of biological roles categories among the top blood candidate biomarker genes for mood.

[0166] Human blood gene expression analysis was conducted in an independent cohort consisting of 30 subjects with other psychotic disorders (schizophrenia, schizoaffective disorder, substance induced psychosis), who also had mood state scores obtained at the time of the blood draw. The

subjects in the psychosis cohort also had a distribution of low (n=9), intermediate (n=7) and high (n=14) mood scores. This cohort was used as a way to verify the predictive power of the mood state biomarker panel, independent of a bipolar disorder diagnosis.

[0167] In the psychotic disorders cohort (n=30), with various psychotic disorders diagnoses, a prediction score of 100 and above had a 71.4% sensitivity and a 62.5% specificity for predicting high mood. A prediction score below 100 had a 66.7% sensitivity and 61.9% specificity for predicting low mood (Table 4B and FIG. **5**).

[0168] Human blood gene expression analysis was also conducted in a second independent bipolar disorder cohort, subsequently collected, consisting of 19 subjects. The subjects in the secondary bipolar cohort had a distribution of low (n=6), intermediate (n=3) and high (n=10) mood scores. The second bipolar cohort was used as a replication cohort, to verify the predictive power of the mood state biomarker panel identified by analysis of data from the primary bipolar cohort.

[0169] In the second bipolar cohort (n=19), a prediction score of 100 and above had a 70.0% sensitivity and a 66.7% specificity for predicting high mood. A prediction score below 100 had a 66.7% sensitivity and 61.5% specificity for predicting low mood (Table 4C and FIG. **6**).

[0170] The primary and secondary bipolar mood disorder cohorts are apriori more related and germane to mood state biomarkers identification, but may have blood gene expression changes due at least in part to the common pharmacological agents used to treat bipolar mood disorders. The psychotic disorders cohort may have blood gene expression changes related to mood state irrespective of the diagnosis and the different medication classes subjects with different diagnoses are on (Table 1 and FIG. **2**B). The psychosis cohort was also notably different in terms of the ethnic distribution (see Table 1b).

[0171] The MIT/Broad Institute connectivity map was interrogated with a signature query composed of the genes in the BioM-10 Mood panel of top biomarkers for low mood and high mood (FIG. **5**). The effects of drugs in the Connectivity Map database and their effects on gene expression as the effects of high mood or low mood on gene expression. As such, as part of the signature query, the 5 biomarkers for high mood were considered as genes "Increased" by high mood, the 5 biomarkers for low mood were genes "Decreased" by high mood. The analysis revealed that sodium phenylbutyrate exerts the most similar effects to high mood, and novobiocin the most similar effects to low mood. Conventional gene expression analysis may not result in the same set of biomarkers.

[0172] By selecting 5 candidate biomarkers for high mood and 5 candidate biomarkers for low mood, a panel of 10 biomarkers for mood disorder that has diagnostic and predictive value was developed based on the scores of the biomarkers for low and high mood sections.

[0173] Thus, the biomarkers identified herein provide quantitative tools for predicting disease states/conditions in subjects suspected of having a mood disorder or in any individual for psychiatric evaluation.

[0174] A meta-analysis of the two bipolar subject cohorts was also conducted. A panel of 10 top biomarkers identified by the meta-analysis was tested for sensitivity and specificity for low and high mood in the two bipolar cohorts (Table 4D). The panel included Edg2, Ednrb, Vil2, Bivm, Camk2d (high mood markers) and Trpc1, Elovl5, Ugt8, Btg1, Nefh (low mood markers). A number of new biomarker genes revealed only in the meta-analysis were identified (see Table 3).

Example 3

Cross-Validation and Integration Using Convergent Functional Genomics Approaches to Identify and Prioritize Biomarkers for Mood Disorders

[0175] The identities of transcripts were established using NetAFFX™ to correlate the GeneChip® array results with array design and annotation information (Affymetrix, Santa Clara, Calif.), and confirmed by cross-checking the target mRNA sequences that had been used for probe design in the Mouse Genome 430 2.0 Array GeneChip® or the Affymetrix Human Genome U133 Plus 2.0 GeneChip® with the GenBank database. Where possible, identities of ESTs were established by BLAST searches of the nucleotide database. A National Center for Biotechnology Information (NCBI) (Bethesda, Md.) BLAST analysis of the accession number of each probe-set was done to identify each gene name. BLAST analysis identified the closest (most similar) known gene existing in the database (the highest known gene at the top of the BLAST list of homologues) which then could be used to search the GeneCards database (Weizmann Institute, Rehovot, Israel). Probe-sets that did not have a known gene were labeled "EST" and their accession numbers were kept as identifiers.

[0176] Human Postmortem Brain Convergence: Information about the candidate genes was obtained using GeneCards, the Online Mendelian Inheritance of Man database at the NCBI database, as well as database searches using PubMed and various combinations of keywords (gene name, bipolar, depression, psychosis, schizophrenia, alcoholism, suicide, dementia, Alzheimer, opiates, cocaine, marijuana, hallucinogens, amphetamines, benzodiazepines, human, brain, postmortem, lymphocytes, fibroblasts). Postmortem convergence was deemed to occur for a gene (or a biomarker) if there were human postmortem data showing changes in expression of that gene in brains from patients with mood disorders (bipolar disorder, depression), or secondarily of other major neuropsychiatric disorders that impact mood (schizophrenia, anxiety, alcoholism).

[0177] Human Genetic Data Convergence: To designate convergence for a particular gene, the gene may have positive reports from candidate gene association studies, or map within 10 cM of a microsatellite marker for which at least one study demonstrated evidence for genetic linkage to mood disorders (bipolar disorder or depression) or secondarily to another neuropsychiatric disorder. The University of Southampton's sequence-based integrated map of the human genome (The Genetic Epidemiological Group, Human Genetics Division, was used to obtain cM locations for both genes and markers. The sex-averaged cM value was calculated and used to determine convergence to a particular marker. For markers that were not present in the Southampton database, the Marshfield database (Center for Medical Genetics, Marshfield, Wis., USA) was used with the NCBI Map Viewer web-site to evaluate linkage convergence.

[0178] Gene Ontology (GO) analysis: The NetAffx™ Gene Ontology Mining Tool (Affymetrix, Santa Clara, Calif.) was employed to categorize the genes in the datasets into functional categories, using the Biological Process ontology branch.

[0179] Ingenuity analysis: Ingenuity Pathway Analysis 5.1 software(Ingenuity Systems, Redwood City, Calif.) was used to analyze the direct interactions of the top candidate genes resulting from the CFG analysis, biological roles, as well as employed to identify genes in the datasets that are the target of existing drugs.

[0180] Convergent Functional Genomics (CFG) Analysis Scoring (see FIG. **2**A) is presented as follows:

[0181] (i) Biomarkers were given the maximum score of 2 points if changed in the human blood samples with high threshold analysis, and only 1 point if changed with low threshold.

[0182] (ii) Biomarkers received 1 point for each external cross-validating line of evidence (human postmortem brain data, human genetic data, animal model brain data, and animal model blood data).

[0183] (iii) Biomarkers received additional bonus points if changed in human brain and blood, as follows:

[0184] (a) 2 points if changed in the same direction;

[0185] (b) 1 point if changed in opposite direction;

[0186] (iv) Biomarkers also received additional bonus points if changed in brain and blood of the animal model, as follows:

[0187] (a) 1 point if changed in the same direction in the brain and blood;

[0188] (b) 0.5 points if changed in opposite direction.

[0189] Thus the total maximum CFG score that a candidate biomarker gene can have is 9 (2+4+2+1). To identify blood biomarkers the scoring pattern described herein is biased more towards awarding additional points for live subject human blood data (if it made the high threshold cut) than literature-derived human postmortem brain data, human genetic data, or the animal model data. The human blood-brain concordance is weighted more favorably than the animal model blood-brain concordance. The scoring analysis presented herein is just one example of assigning quantifiable values to prioritizing biomarkers for mood disorder analysis. Other ways of weighing the scores of line of evidence may give slightly different results in terms of prioritization, if not in terms of the list of genes per se.

[0190] The weightage given to a particular evidence, e.g., post-mortem data or blood expression may be varied. Additional scoring matrices may also be included to account for additional variables. One such example would be the temporal aspect—how long a particular biomarker is turned on.

Example 4

Clinical Applications

[0191] A sample, such as, 5-10 ml of blood is obtained from a patient suspected of having a mood disorder. RNA is isolated from the blood using standard protocols, for example with PAXgene blood RNA extraction kit (PreAnalytiX, a QIAGEN/BD company), followed by GLOBINclear™-Human or GLOBINclear™-Mouse/Rat (Ambion/Applied Biosystems Inc., Austin, Tex.) to remove the globin mRNA. Isolated RNA is labeled using any suitable detectable label if necessary for the gene expression analysis.

[0192] The labeled RNA is then quantified for the presence of one or more of the biomarkers disclosed herein. For example, gene expression analysis is performed using a panel of about 10 biomarkers (e.g., BioM 10 panel) by any standard technique, for example microarray analysis or quantitative PCR or an equivalent thereof. The gene expression levels are analyzed and the absent/present state or fold changes (either increased, decreased, or no change) are determined and a score is established

[0193] Applications of biomarkers for mood disorders: There are no reliable clinical laboratory blood tests for mood disorders. Given the complex nature of mood disorders, the current reliance on patient self-report of symptoms and the clinician's impression on interview of patient is a rate limiting step in delivering the best possible care with existing treatment modalities, as well as in developing new and improved treatment approaches, including new medications.

[0194] Biomarkers disclosed herein are used in the form of panels of biomarkers, as exemplified by a BioM-10 Mood panel, for clinical laboratory tests for mood disorders. Such tests can be: 1) at an mRNA level, quantitation of gene expression through polymerase chain reaction, 2) at a protein level, quantitation of protein levels through immunological approaches such as enzyme-linked immunosorbent assays (ELISA).

[0195] In conjunction with other clinical information, biomarker testing of blood and other fluids (CSF, urine) may play a desirable part of personalizing treatment to increase effectiveness and avoid adverse reactions—personalized medicine in psychiatry.

[0196] Biomarker-based tests for mood disorders help: 1) Diagnosis, early intervention and prevention efforts; 2) Prognosis and monitoring response to various treatments; 3) New neuropsychiatric drug development efforts by pharmaceutical companies, at both a pre-clinical and clinical (Phase I, II and III) stages of the process; 4) Identifying vulnerability to mood problems for people in high stress occupations (for example, military, police, homeland security).

Example 4A

[0197] Diagnosis, early intervention and prevention efforts. A patient with no previous history of mood disorders presents to a primary care doctor or internist complaining of non-specific symptoms: low energy, fatigue, general malaise, aches and pains. Such symptoms are reported in conditions such as stress after a job loss, bereavement, mononucleosis, fibromyalgia, and postpartum in the general population, as well as Gulf War syndrome in veterans. A panel of mood biomarkers can substantiate that the patient is showing objective changes in the blood consistent with a low mood/depressive state. This will direct treatment towards, and substantiate the need to use, anti-depressant medications such as Prozac (fluoxetine), Zolof (sertraline), Celexa (citalopram), Cymbalta (duloxetine), Effexor (venlafaxine) or Wellbutrin (buproprion).

Example 4B

[0198] Clinical diagnosis of a young patient. A young patient (child, adolescent, young adult) with no previous history of mood disorders, but coming from a family where one or more blood relatives suffer from depression may be monitored with regular laboratory tests by their primary care doctor/pediatrician using panels of mood biomarkers. These tests may detect early on a change towards decreased mood (depression) or towards increased mood (mania). This indicates and substantiates the need for initiation of medication treatment with anti-depressants (mentioned above), or mood stabilizers for mania—medication such as Depakote (divalproex), Lithobid (lithium), Lamictal (lamotrigene), Tegretol

(carbamazepine), Topomax (topiramate). This early intervention may be helpful to prevent full-blown illness and hospitalizations, with their attendant negative medical and social consequences. The decision to start medications in children and adolescents is particularly difficult without objective proof, due to the potential side-effects of medications in that age group (agitation, suicidality, weigh-gain, sexual side-effects).

Example 4C

[0199] Monitoring mood biomarkers over an extended period. Many patients with bipolar disorder may present initially with a depressive episode to their primary care doctor or psychiatrist. Monitoring mood biomarkers over time may also help to differentiate depression vs. bipolar disorder (manic-depression). This distinction is helpful because the first-line treatments for the two disorders are different: anti-depressants for depression, mood stabilizers for bipolar. If patients are miss-diagnosed as depressives instead of bipolars, and started on anti-depressant medications only, this can lead to activation and flip into manic states. If prior objective substantiation through biomarker testing of mood cyclicity (going up and down) existed, or early detection of mania in patients put on anti-depressants by seeing a change in biomarker profile towards a high mood state profile before full blown illness and clinical symptoms, an appropriate addition or change to a mood stabilizer medication can be implemented, preventing clinical decompensation, needles suffering and socio-economic loss (employment, relationships).

Example 4D

[0200] Prognosis and monitoring response to various treatments. In depression, initiating medication treatment with current anti-depressants medications is a trial-and-error endeavor. It takes up to 6-8 weeks to see if a medication truly works. By doing a baseline biomarker panel test, and then a repeat test early one in treatment (after 1 week, for example), there would be an early objective indication if an anti-depressant is starting to work or not, and if a switch to another medication is indicated. This would save time and avoid needles suffering for patients, with the attendant socio-economic losses.

Example 4E

[0201] Detecting loss of efficacy of an existing treatment. When a patient has been stable for a while on a medication for depression or bipolar disorder, regular biomarker testing may detect early loss of efficacy of the medication or recurrence of the illness, which would indicate the dose needs to be increased, medication changed, or another medication added, to prevent full blown clinical symptoms.

Example 4F

[0202] Determining adequacy of treatment plan. Objective monitoring with blood biomarker panels of the effect of less reliable or evidence-based interventions: psychotherapy, lifestyle changes, diet and exercise programs for improving mood health. This will show whether the particular intervention works, is sufficient, or medications may need to be added to the regimen.

Example 4G

[0203] Identifying vulnerability to developing mood problems for people in high stress occupations. Military personnel (recruits in boot-camp, active duty soldiers), other people in high-stress jobs (police, homeland security, astronauts), can be monitored on a regular basis to detect early objective changes in mood biomarker profile that would indicate the need for preventive intervention and/or the temporary removal from a high-stress environment.

Example 5

New Neuropsychiatric Drug Development

[0204] Early-stage pre-clinical work and clinical trials of new neuropsychiatric medications for treating mood disorders may benefit from biomarker monitoring to help make a decision early on whether the compound is working. This will speed up the drug-development process and avoid unnecessary costs. Depending on the expression profile of the biomarkers, the results of clinical trials may be obtained earlier than usual.

[0205] In later-stage large clinical trials, a new compound being tested may show an overall statistically non-significant positive effect, despite working well in a sub-group of people in the study. Biomarker testing may provide an objective signature of the genetic and biological make-up of the responders, which can inform recruitment for subsequent validatory clinical trials with higher likelihood of success, as well as inform which patients should be getting the medication, once it is FDA approved and on the market.

TABLE 1

Demographics: (a) individual (b) aggregate
Diagnosis established by DIGS comprehensive structured clinical interview.

(a) Individual demographic data.

Subject ID	Diagnosis	Age	Gender	Ethnicity	VAS Mood (0-100)
Primary Bipolar Cohort					
174-1197-001	BP	37	Male	Caucasian	20
174-1055-001	BP	46	Male	Caucasian	20
phchp029v1	BP	56	Male	Caucasian	22
174-1126-001	BP	33	Male	Caucasian	24
174-1173-001	BP	56	Male	Caucasian	27
174-1161-001	BP	46	Male	Caucasian	29
174-1150-001	BP	52	Male	Caucasian	31
174-1042-001	BP	58	Male	Caucasian	37

TABLE 1-continued

Demographics: (a) individual (b) aggregate
Diagnosis established by DIGS comprehensive structured clinical interview.

174-1112-001	BP	24	Male	Caucasian	38
phchp027v1	BP	40	Male	Caucasian	38
174-1137-001	BP	48	Male	African American	39
phchp023v1	BP	52	Male	Caucasian	39
174-1115-001	BP	42	Male	American Indian	40
phchp020v1	BP	62	Male	Caucasian	42
phchp031v1	BP	51	Male	Caucasian	47
phchp028v1	BP	50	Female	Asian	52
phchp030v1	BP	49	Male	Caucasian	61
174-1107-001	BP	39	Male	Caucasian	63
174-1130-001	BP	21	Male	African American	65
174-5001-001	BP	23	Male	Caucasian	66
174-1132-001	BP	22	Male	African American	71
174-1160-001	BP	52	Male	Caucasian	72
174-1171-001	BP	56	Female	Caucasian	72
174-1156-001	BP	57	Male	Caucasian	72
174-1037-001	BP	54	Male	Caucasian	72
174-5002-001	BP	26	Male	Caucasian	73
174-1119-001	BP	38	Male	Caucasian	73
phchp020v2	BP	62	Male	Caucasian	80
174-1193-001	BP	53	Male	African American	84
			Psychosis Cohort		
phchp022v2	SZ	48	Male	Caucasian	15
phchp005v2	SZA	45	Male	Caucasian	19
phchp025v1	SZ	42	Male	Caucasian	29
phchp021v2	SZA	49	Male	Hispanic	29
phchp006v2	SZA	52	Male	African American	33
phchp033v1	SZA	48	Male	Caucasian	35
phchp016v1	SZ	54	Male	African American	38
phchp021v1	SZA	48	Male	Hispanic	39
phchp019v1	SubPD	50	Male	African-American	41
phchp003v3	SZ	50	Male	African American	47
phchp010v1	SZA	45	Male	Caucasian	48
phchp024v1	SZA	49	Male	African American	49
phchp003v2	SZ	50	Male	African American	53
phchp009v1	SZ	55	Male	African American	54
phchp010v2	SZA	45	Male	Caucasian	55
phchp006v1	SZA	52	Male	African American	57
phchp026v1	SZA	49	Male	African-American	64
phchp022v1	SZ	48	Male	Caucasian	65
phchp010v3	SZA	45	Male	Caucasian	65
phchp014v1	SubPD	55	Male	African American	69
phchp004v1	SZA	55	Male	African American	69
phchp012v1	SZA	55	Male	Caucasian	70
phchp012v2	SZA	55	Male	Caucasian	71
phchp018v1	SZA	54	Female	Caucasian	73
phchp015v1	SubPD	48	Male	African American	76
phchp008v1	SZ	47	Male	African American	76
phchp005v1	SZA	45	Male	Caucasian	81
phchp017v2	SZA	53	Male	African American	84
phchp013v1	SZA	53	Male	African American	89
phchp003v1	SZ	50	Male	African American	93
			Secondary Bipolar Cohort		
phchp039v1	BP	52	Male	Caucasian	11
phchp023v2	BP	52	Male	Caucasian	20
174-1216-001	BP	60	Male	Caucasian	23
174-1278-001	BP	22	Male	Caucasian	24
174-1232-001	BP	45	Male	Caucasian	32
phchp045v1	BP	36	Male	Caucasian	36
174-1203-001	BP	39	Male	African American	49
174-1199-001	BP	41	Male	Caucasian	53
174-1237-001	BP	36	Male	Caucasian	57
174-5006-001	BP	60	Male	Caucasian	66
phchp053v1	BP	58	Male	Caucasian	68
174-1211-001	BP	27	Male	Caucasian	75
phchp031v2	BP	51	Male	Caucasian	79
174-1204-001	BP	52	Male	Caucasian	81
174-1255-001	BP	50	Male	Caucasian	81

TABLE 1-continued

Demographics: (a) individual (b) aggregate
Diagnosis established by DIGS comprehensive structured clinical interview.

174-1220-001	BP	68	Male	Caucasian	82
174-1096-001	BP	50	Male	Caucasian	83
phchp056v1	BP	36	Male	Caucasian	84
174-1258-001	BP	36	Male	Caucasian	90

(b) Aggregate demographic data

	Primary Bipolar Cohort			Psychosis Cohort			Secondary Bipolar Cohort		
	BP Low Mood	BP High Mood	BP Overall	SZA	SZ	Substance induced psychotic disorder	BP Low Mood	BP High Mood	BP Overall
Number of Subjects	13	13	29	18	9	3	6	10	19
Gender (males:females)	13:0	12:1	27:2	17:1	9:0	3:0	6:0	10:0	19:0
Age mean years (SD) range	45.4 (10.0) 24 to 58	41.9 (15.6) 21 to 62	45.0 (12.5) 21 to 62	49.8 (3.9) 45 to 55	49.3 (3.8) 42 to 55	51.0 (3.6) 48 to 55	44.5 (13.6) 22 to 60	48.8 (12.5) 27 to 68	45.8 (12.0) 22 to 68
Duration of Illness mean years (SD) range	22.8 (10.2) 5 to 40	20.4 (17.1) 2 to 49	21.6 (13.7) 2 to 49	31.2 (6.3) 17 to 42	26.7 (4.7) 20 to 26	25.0 (6.0) 20 to 32	25.8 (16.1) 7 to 53	27.2 (6.6) 19 to 38	25.8 (10.5) 7 to 53
Ethnicity (Caucasian/ Other)	11/2	10/3	23/6	9/9	3/6	0/3	6/0	10/0	18/1

BP—bipolar,

SubPD—substance induced psychosis,

SZ—schizophrenia,

SZA—schizoaffective disorder.

VAS Mood score at time of blood draw, on a scale 0 (lowest mood) to 100 (highest mood).

TABLE 2

High threshold and low threshold analysis in primary bipolar cohort and in meta-analysis of both bipolar cohorts. Genes are considered candidate biomarkers for high mood if they are called by the Affymetrix MAS5 software as Absent (A) in the blood of low mood subjects and detected as Present (P) in the blood of high mood subjects. Conversely, genes are considered candidate biomarkers for low mood if they are detected as Present (P) in low mood subjects and Absent (A) in high mood subjects.

Primary Cohort Analysis	Bipolar Subjects (n = 29) 13 Low Mood and 13 High Mood
High Threshold Candidate Biomarker Genes (changed in greater than or equal to 75% subjects; i.e. at least 3-fold enrichment)	10/13 Low Mood vs 10/13 High Mood A/P and P/A analysis
Low Threshold Candidate Biomarker Genes (changed in greater than or equal to 60% subjects; i.e. at least 1.5-fold enrichment)	8/13 Low Mood vs 8/13 High Mood A/P and P/A analysis
	Bipolar Subjects (n = 42)
Meta-analysis	19 Low Mood and 23 High Mood
High Threshold Candidate Biomarker Genes (changed in greater than or equal to 75% subjects; i.e. at least 3-fold enrichment)	15/19 Low Mood vs 18/23 High Mood
Low Threshold Candidate Biomarker Genes (changed in greater than or equal to 60% subjects; i.e. at least 1.5-fold enrichment)	12/19 Low Mood vs 14/23 High Mood

TABLE 3

Top candidate biomarker genes for mood prioritized by CFG score for multiple independent lines of evidence.

Gene Symbol/Name	Entrez Gene ID	Hu. Bl Data	Hu. Postmortem Brain, LCL	Hu. Br. and Bl. Concordance/ Co-Directionality	Hu. Genetic Linkage/ Association	BP Mouse Model Brain[2]	BP Mouse Model Blood	CFG Score
Mbp myelin basic protein	4155	I	Up (male) BP Down (female) BP Down Bipolar	Yes/Yes	18q23 BP'		Cat-IV Meth (I)	6
Edg2 Endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 2	1902	I	Down MDD Down (PFC) BP Up (Parietal Cortex) BP	Yes/Yes	9q31.3 BP			5
Fgfr1 fibroblast growth factor receptor 1	2260	D	Up MDD	Yes/Yes	8p12 BP			5
Fzd3 frizzled homolog 3 (*Drosophila*)	7976	I	Down BP	Yes/Yes	8p21.1 BP			5
Mag myelin-associated glycoprotein	4099	D	Down MDD	Yes/No	19q13.12 Depression	CP Cat-IV Meth (I)		5
Pmp22 peripheral myelin protein 22	5376	D	Down MDD	Yes/No	17p12 BP	CP Cat-IV Meth (I)		5
Ugt8 UDP glycosyltransferase 8 (UDP-galactose ceramide galactosyltransferase)	7368	D	Down MDD	Yes/No	4q26 BP	CP Cat-II (I)		5
Erbb3 Neuregulin receptor (v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian))	2065 (	D	Down MDD Down BP	Yes/No	12q13.2 Depression			4
Igfbp4 insulin-like growth factor binding protein 4	3487	D	Down BP	Yes/No	17q21.2 Depression			4
Igfbp6 insulin-like growth factor binding protein 6	3489	D	Down BP	Yes/No	12q13 Depression			4
Pde6d phosphodiesterase 6D, cGMP-specific, rod, delta	5147	D	Up BP	Yes/Yes	2q37.1			4
Ptprm protein tyrosine phosphatase, receptor type, M	5797	D	Up BP	Yes/Yes	18p11.23			4
Nefh neurofilament, heavy polypeptide 200 kDa	4744	D (MA)	DownBP	Yes/No	22q12.2			4
Atp2c1 ATPase, Ca++-sequestering	27032	D			3q21.3 BP			3
Atxn1 Ataxin 1	6310	I			6p22.3 BP	CP Cat-IV Meth (D)		3
Btg1 B-cell translocation gene 1, anti-proliferative	694	D			12q21.33 BP		Cat-III Meth (D)	3
C6orf182 chromosome 6 open reading frame 182	285753	D			6q21 BP			3
Dicer1 Dicer1, Dcr-1 homolog (*Drosophila*)	23405	D	Down MDD	Yes/No	14q32.13			3

TABLE 3-continued

Top candidate biomarker genes for mood prioritized by CFG score for multiple independent lines of evidence.

Gene Symbol/Name	Entrez Gene ID	Hu. Bl Data	Hu. Postmortem Brain, LCL	Hu. Br. and Bl. Concordance/ Co-Directionality	Hu. Genetic Linkage/ Association	BP Mouse Model Brain[2]	BP Mouse Model Blood	CFG Score
Dnajc6 DnaJ (Hsp40) homolog, subfamily C, member 6	9829	D (HT)			1p31.3 BP Depression			3
Ednrb endothelin receptor type B	1910	I			13q22.3 BP	CP Cat-III Meth (I)		3
Elovl5 ELOVL family member 5, elongation of long chain fatty acids (yeast)	60481	D			6p12.1 BP		Cat-IV VPA (D)	3
Gnal guanine nucleotide binding protein, alpha stimulating, olfactory type	2774	D (HT)			18p11.21 BP			3
Klf5 Kruppel-like factor 5	688	D (HT)			13q22.1 BP			3
Lin7a lin 7 homolog a (*C. elegans*)	8825	D			12q21.31 BP	Increased (Rat Brain)		3
Manea mannosidase, endo-alpha	79694	D (HT)			6q16.1 BP Depression			3
Nupl1 nucleoporin like 1	9818	D (HT)			13q12.13 BP			3
Pde6b phosphodiesterase 6B, cGMP-specific, rod, beta (congenital stationary night blindness 3, autosomal dominant)	5158	D	Down MDD	Yes/No	4p16.3			3
Slc25a23 solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 23	79085	D (HT)			19p13.3	CP Cat-IV VPA (I)		3
Synpo synaptopodin	11346	D			5q33.1 BP	PFC Cat-III Meth (D)		3
Tgm2 transglutaminase 2, C polypeptide	7052	D			20q11.23 BP		Cat-III Meth (D)	3
Tjp3 tight junction protein 3 (zona occludens 3)	27134	D (HT)			19p13.3 BP			3
Tpd52 tumor protein D52	7163	D (HT)			8q21.13 BP			3
Trpc1 transient receptor potential cation channel, subfamily C, member 1	7220	D			3q23 BP	CP Cat-IV VPA (I)		3
Bclaf1 BCL2-associated transcription factor 1	9774	D	Down (Lymphocytes) BP		6q23.3			2
Gosr2 golgi SNAP receptor complex member 2	9570	D	Down (Lymphocytes) BP		17q21.32			2
Rdx radixin	5962	D	Down (Lymphocytes) BP		11q22.3			2
Wdr34 WD repeat domain 34	89891	D	Down (Lymphocytes) BP		9q34.11			2

TABLE 3-continued

Top candidate biomarker genes for mood prioritized by CFG score for multiple independent lines of evidence.

Gene Symbol/Name	Entrez Gene ID	Hu. Bl Data	Hu. Postmortem Brain, LCL	Hu. Br. and Bl. Concordance/ Co-Directionality	Hu. Genetic Linkage/ Association	BP Mouse Model Brain[2]	BP Mouse Model Blood	CFG Score
Bic BIC transcript (microRNA 155)	114614	D (MA)			21q21.3			2
C8orf42 chromosome 8 open reading frame 42	157695	D (MA)			8p23.3			2
Dock9 Dedicator of cytokinesis 9	23348	D (MA)			13q32.3		Cat-III-Val (D)	2
Hrasls HRAS-like suppressor	57110	D (MA)			3q29			2
Ibrdc2 IBR domain containing 3 (Rnf144b)	255488	D (MA)			6p22.3			2
P2ry12 purinergic receptor P2Y, G-protein coupled 12	64805	D (MA)			3q25.1			2
Specc1 spectrin domain with coiled-coils 1	92521	D (MA)			17p11.2			2
Vil2 villin 2 (ezrin)	7430	I (MA)			6q25.3			2
C20orf7 chromosome 20 open reading frame 7	79133	D (MA)			20p12.1			1
Chrnb3 cholinergic receptor, nicotinic, beta 3	1142	I (MA)			8p11.21			1
Eif4a2 eukaryotic translation initiation factor 4A, isoform 2	1974	I (MA)			3q27.3			1
Gins4 GINS complex subunit 4 (Sld5 homolog)	84296	I (MA)			8p11.21			1
Grhl1 grainyhead-like 1 (*Drosophila*)	29841	D (MA)			2p25.1			1
Gtpbp8 GTP-binding protein 8 (putative)	29083	D (MA)			3q13.2			1
Heatr6 HEAT repeat containing 6	63897	I (MA)			17q23.2			1
Igl@ immunoglobulin lambda chain, variable 1	3535	I (MA)			22q11.1-q11.2			1
Il17rc interleukin 17 receptor C	84818	I (MA)			3p25.3			1
Itfg1 integrin alpha 2b	81533	D (MA)			16q12.1			1
Loc388692 hypothetical gene supported by AK123662	388692	D (MA)			1q21.2			1
Loc654342 Similar to lymphocyte-specific protein 1	654342	I (MA)			2p11.1			1
Lrrc37a leucine rich repeat containing 37A	9884	D (MA)			17q21.31			1
Pbrm1 polybromo 1	55193	I (MA)			3p21.1			1

TABLE 3-continued

Top candidate biomarker genes for mood prioritized by CFG score for multiple independent lines of evidence.

Gene Symbol/Name	Entrez Gene ID	Hu. Bl Data	Hu. Postmortem Brain, LCL	Hu. Br. and Bl. Concordance/ Co-Directionality	Hu. Genetic Linkage/ Association	BP Mouse Model Brain[2]	BP Mouse Model Blood	CFG Score
Pex13 peroxisome biogenesis factor 13	5194	D (MA)			2p16.1			1
Pol3s polyserase 3	339105	I (MA)			16p11.2			1
Pparbp PPAR binding protein	5469	D (MA)			17q12			1
Prkd2 protein kinase D2	25865	D (MA)			19q13.32			1
Prr7 proline rich 7 (synaptic)	80758	D (MA)			5q35.3			1
Psph phosphoserine phosphatase	5723	D (MA)			7p11.2			1
Rfx3 regulatory factor X, 3 (influences HLA class II expression)	5991	I (MA)			9p24.2			1
Rps16 ribosomal protein S16	6217	D (MA)			19q13.2			1
Samd4a sterile alpha motif domain containing 4A	23034	I (MA)			14q22.2			1
Scamp1 secretory carrier membrane protein 1	9522	D (MA)			5q14.1			1
Scn11a sodium channel, voltage-gated, type XI, alpha	11280	I (MA)			3p22.2			1
Spa17 sperm autoantigenic protein 17	53340	D (MA)			11q24.2			1
Tcf7l2 transcription factor 7-like 2, T-cell specific, HMG-box	6934	I (MA)			10q25.3			1
Wbscr16 Williams-Beuren syndrome chromosome region 16	81554	I (MA)			7q11.23			1
Wdr55 WD repeat domain 55	54853	D (MA)			5q31.3			1
Znf492 zinc finger protein 492	57615	D (MA)			19p12			1
Znf576 zinc finger protein 576	79177	I (MA)			19q13.31			1

Top candidate biomarker genes for mood.

For human blood (Hu Bl.) data:

I—increased in high mood (mania);

D—decreased in high mood (mania)/increased in low mood (depression).

For human postmortem brain (Hu Br.) data:

Up—increased;

Down—decreased in expression.

For mouse data

METH—methamphetamine,

VPA—valproate.

MDD—major depressive disorder.

LCL—lymphoblastoid cell lines.

(HT) High threshold.

(MA) identified by meta-analysis only.

TABLE 4

BioM-10 Mood panel derived from primary bipolar cohort analysis: sensitivity and specificity for predicting mood state. Primary Bipolar cohort (A), Psychosis Cohort (B) and Secondary Bipolar cohort (C). Results with meta-analysis derived panel (D).

	Sensitivity	Specificity
A.		
Primary Bipolar Cohort		
High Mood	84.6%	68.8%
Low Mood	76.9%	81.3%
B.		
Other Psychotic Disorders Cohort		
High Mood	71.4%	62.5%
Low Mood	66.7%	61.9%
C.		
Secondary Bipolar Cohort		
High Mood	70%	66.7%
Low Mood	66.7%	61.5%

High Mood: Mbp, Edg2, Fzd3, Atxn1, Ednrb
Low Mood: Fgfr1, Mag, Pmp22, Ugt8, Erbb3

	Sensitivity	Specificity
D.		
Primary Bipolar Cohort		
High Mood	84.6%	80.0%
Low Mood	61.5%	87.5%

TABLE 4-continued

Secondary Bipolar Cohort		
High Mood	90%	88.9%
Low Mood	66.7%	92.3%

Meta-analysis derived BioM 10 Mood Panel

High Mood: Edg2, Ednrb, Vil2, Bivm, Camk2d

Low Mood: Trpc1, Elovl5, Ugt8, Btg1, Nefh

TABLE 5

Biological Roles. Ingenuity pathway analysis (IPA) of biological functions categories among our top blood candidate biomarker genes for mood. Genes from Table 3. Top categories, over-represented with a significance of $p < 0.05$, are shown.

Cell Death	1.43E−07-4.54E−02
Nervous System Development and Function	8.31E−07-4.63E−02
Cell Morphology	2.25E−05-4.63E−02
Cellular Assembly and Organization	4.48E−05-4.56E−02
Neurological Disease	7.46E−05-4.63E−02
Cellular Growth and Proliferation	1.11E−04-4.89E−02
Skeletal and Muscular System Development and Function	1.11E−04-3.83E−02
Tissue Morphology	1.12E−04-4.09E−02
Behavior	2.08E−04-4.63E−02
Digestive System Development and Function	2.08E−04-4.63E−02
Cellular Development	2.86E−04-4.63E−02
Cancer	5.50E−04-4.89E−02

TABLE 6

Targets of existing drugs. Complete list of the blood candidate biomarker genes for mood that are the direct target of existing drugs.

Genes	Gene Name	Drugs
ADA	adenosine deaminase	pentostatin
AGTR1	angiotensin II receptor, type 1	losartan/hydrochlorothiazide, valsartan/hydrochlorothiazide, candesartan cilexetil, olmesartan medoxomil, irbesartan, losartan potassium, telmisartan, eprosartan, candesartan cilexetil/hydrochlorothiazide, hydrochlorothiazide/irbesartan, eprosartan/hydro
COL6A2	collagen, type VI, alpha 2	collagenase
DHFR	dihydrofolate reductase	iclaprim, methotrexate, LY231514, PT 523
EDNRB	endothelin receptor type B	bosentan, sitaxsentan, atrasentan
GNRH1	gonadotropin-releasing hormone 1 (luteinizing-releasing hormone)	leuprolide, goserelin
GNRHR	gonadotropin-releasing hormone receptor	cetrorelix, triptorelin, abarelix
GUCY1A3	guanylate cyclase 1, soluble, alpha 3	nitroglycerin, isosorbide-5-mononitrate, isosorbide dinitrate, nitroprusside, isosorbide dinitrate/hydralazine
KCNMB4	potassium large conductance calcium-activated channel, subfamily M, beta member 4	tedisamil
PDE4D	phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 dunce homolog, *Drosophila*)	arofylline, tetomilast, anagrelide, cilomilast, milrinone, rolipram, L-826,141, roflumilast, caffeine
PDE5A	phosphodiesterase 5A, cGMP-specific	DA-8159, sildenafil, dipyridamole, aspirin/dipyridamole, vardenafil, tadalafil
POLE	polymerase (DNA directed), epsilon	nelarabine, gemcitabine, clofarabine, trifluridine
PPARA	peroxisome proliferative activated receptor, alpha	tesaglitazar, clofibrate, fenofibrate, gemfibrozil

TABLE 6-continued

Targets of existing drugs. Complete list of the blood candidate biomarker genes for mood that are the direct target of existing drugs.

Genes	Gene Name	Drugs
SLC18A2	solute carrier family 18 (vesicular monoamine), member 2	deserpidine/methyclothiazide, deserpidine, reserpine/trichlormethiazide, chlorothiazide/reserpine, chlorthalidone/reserpine, hydralazine/hydrochlorothiazide/reserpine, hydroflumethiazide/reserpine, polythiazide/reserpine, hydrochlorothiazide/reserpine, r
TLR9	toll-like receptor 9	PF-3512676

TABLE 7

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Abca11 ATP-binding cassette, sub-family A (ABC1), member 11 (pseudogene	10348	D
Abhd6 abhydrolase domain containing 6	57406	D
Acacb acetyl-Coenzyme A carboxylase beta	32	D
Adamts5 ADAM metallopeptidase with thrombospondin type 1 motif, 5 (aggrecanase-2)	11096	D
Agmat agmatine ureohydrolase (agmatinase)	79814	D
Agpat7 1-acylglycerol-3-phosphate O-acyltransferase 7 (lysophosphatidic acid acyltransferase, eta)	254531	D
Agrn agrin	375790	D
Agtr1 angiotensin II receptor, type 1	185	D
Amn Amnionless homolog (mouse)	81693	D (HT)
Anapc10 anaphase promoting complex subunit 10	10393	D
Ankdd1a ankyrin repeat and death domain containing 1A	348094	I
Ankrd13b ankyrin repeat domain 13B	124930	D
Ankrd22 ankyrin repeat domain 22	118932	I
Ankrd54 ankyrin repeat domain 54	129138	D
Ankrd57 ankyrin repeat domain 57	65124	D
Anubl1 AN1, ubiquitin-like, homolog (*Xenopus laevis*)	93550	D
Apobec4 apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 4 (putative)	403314	I
Arid1b AT rich interactive domain 1B (SWI1-like)	57492	D
Armc8 armadillo repeat containing 8	25852	D
Arsk arylsulfatase family, member K	153642	D
Atad2 ATPase family, AAA domain containing 2	29028	D
Atp2c1 ATPase, Ca++-sequestering	27032	D
Atp6v1e2 ATPase, H+ transporting, lysosomal 31 kDa, V1 subunit E2	90423	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Atp7b ATPase, Cu++ transporting, beta polypeptide	540	D
Atxn 1 Ataxin 1	6310	I
Azi2 5-azacytidine induced gene 2	64343	D
B3gnt1 UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 1	11041	D
Bcas4 breast carcinoma amplified sequence 4	55653	D
Bclaf1 BCL2-associated transcription factor 1	9774	D
Bet3l BET3 like (*S. cerevisiae*	221300	D
Bhlhb3 basic helix-loop-helix domain containing, class B, 3	79365	D
Bic BIC transcript	114614	D
Bivm basic, immunoglobulin-like variable motif containing	54841	I
Bmpr1a bone morphogenetic protein receptor, type 1A	657	D
Bnip1 BCL2/adenovirus E1B 19 kDa interacting protein 1	662	D
Brwd 1 bromodomain and WD repeat domain containing 1	54014	D
Btbd12 BTB (POZ) domain containing 12	84464	I
Btg1 B cell translocation gene 1, anti-proliferative	694	D
Btnl9 butyrophilin-like 9	153579	D
C10orf110 chromosome 10 open reading frame 110	55853	I
C10orf18 chromosome 10 open reading frame 18	54906	I
C11orf71 chromosome 11 open reading frame 71	54494	D
C11orf74 chromosome 11 open reading frame 74	119710	D
C12orf29 chromosome 12 open reading frame 29	91298	D
C12orf47 chromosome 12 open reading frame 47	51275	D
C14orf118 chromosome 14 open reading frame 118	55668	D
C14orf131 chromosome 14 open reading frame 131	55778	D
C14orf145 chromosome 14 open reading frame 145	145508	D
C14orf64 chromosome 14 open reading frame 64	388011	D
C16orf52 chromosome 16 open reading frame 52	146174	D
C18orf1 Chromosome 18 open reading frame 1	753	D
C18orf25 chromosome 18 open reading frame 25	147339	I
C18orf55 chromosome 18 open reading frame 55	29090	D
C19orf52 chromosome 19 open reading frame 52	90580	I
C1orf89 chromosome 1 open reading frame 89	79363	D
C20orf112 chromosome 20 open reading frame 112	140688	D
C20orf94 chromosome 20 open reading frame 94	128710	I

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
C21orf109 chromosome 21 open reading frame 109 /// similar to Protein C21orf109	193629	D
C21orf114 chromosome 21 open reading frame 114	193629	D
C21orf56 chromosome 21 open reading frame 56	84221	I
C2orf40 chromosome 2 open reading frame 40	84417	D
C3orf23 chromosome 3 open reading frame 23	285343	D
C6orf170 chromosome 6 open reading frame 170	221322	D
C6orf182 chromosome 6 open reading frame 182	285753	D
C6orf26 chromosome 6 open reading frame 26	401251	D
C6orf60 chromosome 6 open reading frame 60	79632	D
C7orf26 chromosome 7 open reading frame 26	79034	D
C7orf36 chromosome 7 open reading frame 36	57002	D
C8orf33 chromosome 8 open reading frame 33	65265	D
C9orf61 chromosome 9 open reading frame 61	9413	I
C9orf71 chromosome 9 open reading frame 71	169693	D
C9orf82 chromosome 9 open reading frame 82	79886	D
C9orf90 chromosome 9 open reading frame 90	203245	I
Cadm1 cell adhesion molecule 1	23705	D
Camk2d Calcium/calmodulin-dependent protein kinase (CaM kinase) II delta	817	I
Catsper2 cation channel, sperm associated 2	117155	D
Cbfb core binding factor beta	865	D
Cc2d2a/Kiaa1345 KIAA1345 protein	57545	D
Ccdc6 coiled-coil domain containing 6	8030	D
Ccdc65 coiled-coil domain containing 65	85478	D
Ccdc88a coiled-coil domain containing 88A	55704	D
Ccdc99 coiled-coil domain containing 99	54908	D
Ccne2 cyclin E2	9134	D
Cdc7 cell division cycle 7 (*S. cerevisiae*)	8317	D
Cdkn2b cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)	1030	D
Ceacam6 carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen)	4680	D
Cfl2 cofilin 2 (muscle)	1073	D
Clcf1 cardiotrophin-like cytokine factor 1	23529	D
Clcn3 chloride channel 3	1182	D
Cllu1 chronic lymphocytic leukemia up-regulated 1	574028	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Cmtm4 CKLF-like MARVEL transmembrane domain containing 4	146223	D
Cnnm1 cyclin M1	26507	D
Cnot2 CCR4-NOT transcription complex, subunit 2	4848	D
Col6a2 procollagen, type VI, alpha 2	1292	D
Coq3 coenzyme Q3 homolog, methyltransferase (*S. cerevisiae*)	51805	D
Cplx3 complexin 3	594855	I
Cpm carboxypeptidase M	1368	I
Cpvl Carboxypeptidase, vitellogenic-like	54504	D
Cr2 complement component (3d/Epstein Barr virus) receptor 2	1380	D
Csnk1a1 casein kinase 1, alpha 1	1452	D
Ctr9 Ctr9, Paf1/RNA polymerase II complex component, homolog (*S. cerevisiae*)	9646	D
Cuedc1 CUE domain containing 1	404093	I
Cwf19l2 CWF19-like 2, cell cycle control (*S. pombe*)	143884	I
Cxcl12 chemokine (C—X—C motif) ligand 12	6387	D
Cxcl6 chemokine (C—X—C motif) ligand 6 (granulocyte chemotactic protein 2)	6372	D
Cyp2c19 CYP2C9 cytochrome P450, family 2, subfamily C, polypeptide 19	1557	I
Cyp2c9 cytochrome P450, family 2, subfamily C, polypeptide 9	1559	D
Cyp2e1 cytochrome P450, family 2, subfamily E, polypeptide 1	1571	D
Cyp2u1 cytochrome P450, family 2, subfamily U, polypeptide 1	113612	D
Daam1 dishevelled associated activator of morphogenesis 1	23002	D
Dcbld1 discoidin, CUB and LCCL domain containing 1	285761	D
Depdc6 DEP domain containing 6	64798	D
Dhfr dihydrofolate reductase	1719	D
Dhx35 DEAH (Asp-Glu-Ala-His) box polypeptide 35	60625	D
Dicer1 Dicer1, Dcr-1 homolog (*Drosophila*)	23405	D
Dio2 deiodinase, iodothyronine, type II	1734	I
Dip2b DIP2 disco-interacting protein 2 homolog B (*Drosophila*)	57609	I
Disp1 dispatched homolog 1 (*Drosophila*)	84976	D
DKFZp564H213 hypothetical gene supported by AL049275	440432	I
Dnajb9 DnaJ (Hsp40) homolog, subfamily B, member 9	4189	D
Dnajc6 DnaJ (Hsp40) homolog, subfamily C, member 6	9829	D (HT)
Dock5 dedicator of cytokinesis 5	80005	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Dscaml1 Down syndrome cell adhesion molecule-like 1	57453	D
Dst dystonin	667	D
Dtwd1 DTW domain containing 1	56986	D
Dtx3 deltex 3 homolog (*Drosophila*)	196403	D
Dus4l dihydrouridine synthase 4-like (*S. cerevisiae*)	11062	D
Dynlrb1 dynein, light chain, roadblock-type 1	83658	D
E2f5 E2F transcription factor 5, p130-binding	1875	D
E2f7 E2F transcription factor 7	144455	D
Edg2 endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 2	1902	I
Ednrb endothelin receptor type B	1910	I
Egr1 early growth response 1	1958	D
Eid3 E1A-like inhibitor of differentiation 3	493861	D
Eif4g3 eukaryotic translation initiation factor 4 gamma, 3	8672	D
Elovl5 ELOVL family member 5, elongation of long chain fatty acids (yeast)	60481	D
Emid1 EMI domain containing 1	129080	D
Emilin2 elastin microfibril interfacer 2	84034	D
Eml5 echinoderm microtubule associated protein like 5	161436	D
Enpp4 ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)	22875	D
Entpd4 ectonucleoside triphosphate diphosphohydrolase 4	9583	D
Epb41l4a erythrocyte membrane protein band 4.1 like 4A	64097	D
Epn1 epsin 1	29924	D
Eps8 epidermal growth factor receptor pathway substrate 8	2059	D
Erbb3 Neuregulin receptor (v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	2065	D
Espn espin	83715	D
Ezh1 enhancer of zeste homolog 1 (*Drosophila*)	2145	I
Fa2h fatty acid 2-hydroxylase	79152	I
Faah2 fatty acid amide hydrolase 2	158584	D
Fam13a1 family with sequence similarity 13, member A1	10144	D
Fam55a family with sequence similarity 55, member A	120400	I
Fam63b family with sequence similarity 63, member B	54629	D
Fam98a family with sequence similarity 98, member A	25940	D
Fam108c1 family with sequence similarity 108, member C1	58489	D
Fam120a family with sequence similarity 120A	23196	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Fam139A/Flj40722 hypothetical protein FLJ40722	285966	D
Fastk Fas-activated serine/threonine kinase	10922	D
Fbxo15 F-box protein 15	201456	D
Fbxo31 F-box protein 31	79791	D
Fbxo5 F-box protein 5	26271	D
Fcer2 Fc fragment of IgE, low affinity II, receptor for (CD23)	2208	D
Fchsd1 FCH and double SH3 domains 1	89848	D
Fcrl2 Fc receptor-like 2	79368	D
Fer1l3 fer-1-like 3, myoferlin (*C. elegans*)	26509	D
Fgfr1 fibroblast growth factor receptor 1	2260	D
Fggy/Flj10986 hypothetical protein FLJ10986	55277	I
Flj13305 hypothetical protein FLJ13305	84140	D
Flj22167 hypothetical protein FLJ22167	79583	D
Flnc filamin C, gamma (actin binding protein 280	2318	D
Fndc3b fibronectin type III domain containing 3B	64778	I
Fosl2 FOS-like antigen 2	2355	D
Fsd1l Fibronectin type III and SPRY domain containing 1-like	83856	D
Fzd3 frizzled homolog 3 (*Drosophila*)	7976	I
G3bp1 GTPase activating protein (SH3 domain) binding protein 1	10146	D
Gata3 GATA binding protein 3	2625	D
Gins3 GINS complex subunit 3 (Psf3 homolog)	64785	D
Gm2a GM2 ganglioside activator	2760	D
Gmppb GDP-mannose pyrophosphorylase B	29925	D
Gnal guanine nucleotide binding protein, alpha stimulating, olfactory type	2774	D (HT)
Gng8 guanine nucleotide binding protein (G protein), gamma 8 subunit	94235	D
Gnrh1 gonadotropin-releasing hormone 1 (luteinizing-releasing hormone)	2796	D
Gnrhr gonadotropin-releasing hormone receptor	2798	D
Golga2l1 golgi autoantigen, golgin subfamily a, 2-like 1	55592	D
Golga3 golgi autoantigen, golgin subfamily a, 3	2802	D
Golga8g golgi autoantigen, golgin subfamily a, 8G	283768	D
Gosr2 golgi SNAP receptor complex member 2	9570	D
Gp1bb glycoprotein Ib (platelet), beta polypeptide	2812	I

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Gpatch2 G patch domain containing 2	55105	D
Gpd2 glycerol phosphate dehydrogenase 2, mitochondrial	2820	D
Gpr180 G protein-coupled receptor 180	160897	D
Gpr19 G protein-coupled receptor 19	2842	D
Gpsm1 G-protein signalling modulator 1 (AGS3-like, *C. elegans*)	26086	D
Gramd2 GRAM domain containing 2	196996	D
Grb2 growth factor receptor-bound protein 2	2885	D
Gtdc1 glycosyltransferase-like domain containing 1	79712	I
Habp4 hyaluronan binding protein 4	22927	D
Hells helicase, lymphoid-specific	3070	D
Hemk1 HemK methyltransferase family member 1	51409	D
Herpud2 HERPUD family member 2	64224	D
Hif1an hypoxia-inducible factor 1, alpha subunit inhibitor	55662	D
Hist1h3b histone cluster 1, H3b	8358	D
Hla-dqa1 major histocompatibility complex, class II, DQ alpha 1 /// major histocompatibility complex, class II, DQ alpha 1	3117	I
Hla-drb1 major histocompatibility complex, class II, DR beta 1	3123	I
Hps3 Hermansky-Pudlak syndrome 3	84343	D
Hrasls3 HRAS-like suppressor 3	11145	D
Huwe1 HECT, UBA and WWE domain containing 1	10075	D
Ica1 intestinal cell kinase	3382	D
Ift172 intraflagellar transport 80 homolog (*Chlamydomonas*)	26160	D
Igfbp4 insulin-like growth factor binding protein 6	3487	D
Igfbp6 Immunoglobulin heavy chain 1a (serum IgG2a)	3489	D
Ighg1 immunoglobulin heavy constant gamma 1 (G1m marker)	3500	D
Igsf22 immunoglobulin superfamily, member 3	283284	D
Il15 interleukin 17 receptor A	3600	D
Immp1l inner membrane protein, mitochondrial	196294	D
Insc insulin induced gene 1	387755	D
Insr insulin receptor	3643	D
Ints10 integrator complex subunit 7	55174	D
Ints7 integrator complex subunit 8	25896	D
Ipo11 Intracisternal A particle-promoted polypeptide	51194	D
Itch Integrin alpha FG-GAP repeat containing 1	83737	D
Itsn2 isovaleryl Coenzyme A dehydrogenase	50618	I

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Jag1 jagged 2	182	D
Jakmip2 jumonji, AT rich interactive domain 1B	9832	D
Jmjd5 junction-mediating and regulatory protein	79831	D
Josd2 Josephin domain containing 2	126119	D
Katnal1 katanin p60 subunit A-like 1	84056	D
Kbtbd3 kelch repeat and BTB (POZ) domain containing 3	143879	D
Kcnmb4 potassium large conductance calcium-activated channel, subfamily M, beta member 4	27345	D
Khk ketohexokinase (fructokinase) /// ketohexokinase (fructokinase)	3795	D
Kiaa0494 KIAA0494	9813	D
Kiaa1009 KIAA1009	22832	D
Kiaa1107 KIAA1107	23285	D
Kiaa1377 KIAA1377	57562	D
Kiaa1586 KIAA1586	57691	D
Kiaa1704 1200011I18Rik KIAA1704	55425	D
Kiaa1729 KIAA1729 protein	85460	D
Kif5c kinesin family member 5C	3800	D
Klf12 Kruppel-like factor 12	11278	D
Klf5 Kruppel-like factor 5	688	D (HT)
Klk7 kallikrein-related peptidase 7	5650	D
Krtap4-9 keratin associated protein 4-9	85286	I
L2hgdh L-2-hydroxyglutarate dehydrogenase	79944	D
Laptm4b lysosomal associated protein transmembrane 4 beta	55353	D
Larp4 La ribonucleoprotein domain family, member 4	113251	D
Lepr leptin receptor	3953	I
Lgals4 lectin, galactoside-binding, soluble, 4 (galectin 4)	3960	D
Lhx4 LIM homeobox 4	89884	I
Lims2 LIM and senescent cell antigen-like domains 2	55679	D
Lin7a lin 7 homolog a (*C. elegans*)	8825	D
Lin7b lin-7 homolog B (*C. elegans*)	64130	D
Lins1 lines homolog 1 (*Drosophila*)	55180	D
Loc144481 hypothetical protein LOC144481	144481	D
Loc144874 Hypothetical protein LOC144874	144874	D
Loc145783 hypothetical protein LOC145783	145783	D
Loc148709 actin pseudogene	148709	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Loc158863 hypothetical protein LOC158863	158863	D
Loc253012 hypothetical protein LOC253012	253012	D
Loc253039 hypothetical protein LOC253039	253039	D
Loc283140 hypothetical protein LOC283140	283140	I
Loc283481 hypothetical protein LOC283481	283481	D
Loc284373 hypothetical protein LOC284373	284373	D
Loc284749 hypothetical protein LOC284749	284749	D
Loc285014 hypothetical protein LOC285014	285014	D
Loc285378 hypothetical protein LOC285378	285378	D
Loc285535 hypothetical protein LOC285535	285535	D
Loc285813 hypothetical protein LOC285813	285813	D
Loc285831 hypothetical protein LOC285831	285831	D
Loc338653 hypothetical protein LOC338653	338653	I
Loc339803 hypothetical protein LOC339803	339803	D
Loc340544 hypothetical protein LOC340544	340544	D
Loc344405 hypothetical LOC344405	344405	D
Loc348180 hypothetical protein LOC348180, isoform 1	348180	D
Loc387647 hypothetical gene supported by BC014163	387647	D
Loc388692 hypothetical gene supported by AK123662	388692	D
Loc401913 hypothetical LOC401913	401913	I
Loc441383 hypothetical gene supported by AF086559; BC065734	441383	D
Loc442257 similar to 40S ribosomal protein S4, Y isoform 2	442257	D
Loc51035 SAPK substrate protein 1	51035	D
Loc51255 hypothetical protein LOC51255	51255	D
Loc554203 hypothetical LOC554203	554203	I
Loc554206 hypothetical LOC554206	554206	D
Loc56755 hypothetical protein LOC56755	56755	D
Loc619208 hypothetical protein LOC619208	619208	D
Loc645513 Similar to septin 7	645513	D
Loc730202 hypothetical protein LOC730202	730202	D
Loc91431 prematurely terminated mRNA decay factor-like	91431	I
Loh3cr2a loss of heterozygosity, 3, chromosomal region 2, gene A	29931	I
Lrp16 LRP16 protein	28992	D
Lrrc16 leucine rich repeat containing 16	55604	D
Lrrc8a leucine rich repeat containing 8 family, member A	56262	I

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Lrrc8b leucine rich repeat containing 8 family, member B	23507	D (HT)
Lrrcc1 leucine rich repeat and coiled-coil domain containing 1	85444	D
Lrrk1 leucine-rich repeat kinase 1	79705	I
Luzp1 leucine zipper protein 1	7798	D
Lyrm4 LYR motif containing 4	57128	D
Maf avian musculoaponeurotic fibrosarcoma (v-maf) AS42 oncogene homolog	4094	D
Mag myelin-associated glycoprotein	4099	D
Manea mannosidase, endo-alpha	79694	D (HT)
Mbd5 methyl-CpG binding domain protein 5	55777	D
Mbp myelin basic protein	4155	I
Mcf2l MCF.2 cell line derived transforming sequence-like	23263	D
Mcm3ap minichromosome maintenance complex component 3 associated protein	8888	I
Mcoln3 mucolipin 3	55283	D
Mds2 myelodysplastic syndrome 2 translocation associated	259283	D
Me3 malic enzyme 3, NADP(+)-dependent, mitochondrial	10873	D
Mfrp membrane frizzled-related protein	83552	D
Mgat4a mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase, isozyme A	11320	D
Mgc10997 MGC10997	84741	D
Mgc33556 hypothetical LOC339541	339541	I
Mgc39900 hypothetical protein MGC39900	286527	D
Mgc46336 hypothetical protein MGC46336	283933	D
Mia melanoma inhibitory activity	8190	D
Mical3 microtubule associated monoxygenase, calponin and LIM domain containing 3	57553	D
Mier3 mesoderm induction early response 1, family member 3	166968	D
Mki67 antigen identified by monoclonal antibody Ki-67	4288	D
Mks1 Meckel syndrome, type 1	54903	D
Mllt4 myeloid/lymphoid or mixed lineage-leukemia translocation to 4 homolog (*Drosophila*)	4301	I
Mmaa methylmalonic aciduria (cobalamin deficiency) cblA type	166785	D
Mmd2 monocyte to macrophage differentiation-associated 2	221938	I
Mocs2 molybdenum cofactor synthesis 2	4338	D
Morn3 MORN repeat containing 3	283385	D
Mrpl30 mitochondrial ribosomal protein L30	51263	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Mrps15 mitochondrial ribosomal protein S15	64960	I
Mta3 metastasis associated 1 family, member 3	57504	D
Mtap methylthioadenosine phosphorylase	4507	D
Mtfr1 mitochondrial fission regulator 1	9650	D
Mtmr3 myotubularin related protein 3	8897	D
Mustn1 musculoskeletal, embryonic nuclear protein 1	389125	D
Mxra8 matrix-remodelling associated 8	54587	D
Myef2 myelin expression factor 2	50804	D
Mylip myosin regulatory light chain interacting protein	29116	I
Myo6 myosin VI	4646	D
Myo1E myosin IE	4643	D
Myom2 myomesin (M-protein) 2, 165 kDa	9172	I
Naip similar to Occludin	4671	D
Nap1l3 nucleosome assembly protein 1-like 3	4675	D
Nedd1 neural precursor cell expressed, developmentally down-regulated 1	121441	D
Nenf neuron derived neurotrophic factor	29937	D
Nfe2l3 nuclear factor (erythroid-derived 2)-like 3	9603	D
Nfib nuclear factor I/B	4781	D
Nfix nuclear factor I/X	4784	I
Nfkbie nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon	4794	D
Nhedc1 Na+/H+ exchanger domain containing 1	150159	D
Nipsnap3b nipsnap homolog 3B (*C. elegans*)	55335	D
Nln neurolysin (metallopeptidase M3 family)	57486	I
Nr4a2 nuclear receptor subfamily 4, group A, member 2	4929	D
Nrbp2 nuclear receptor binding protein 2	340371	D
Nt5m 5',3'-nucleotidase, mitochondrial	56953	I
Nupl1 nucleoporin like 1	9818	D (HT)
Ocm oncomodulin	654231	I
Or7e104p olfactory receptor, family 7, subfamily E, member 104 pseudogene	81137	I
Orc4l origin recognition complex, subunit 4-like (*S. cerevisiae*)	5000	D
Osgepl1 Osialoglycoprotein endopeptidase-like 1	64172	D
Otud7b OTU domain containing 7B	56957	D
Pabpc1l2b/Rp11-493k23.2 similar to poly(A) binding protein, cytoplasmic 1	645974	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Pafah1b1 platelet-activating factor acetylhydrolase, isoform 1b, beta1 subunit	5048	D
Pard6b par-6 partitioning defective 6 homolog beta (*C. elegans*)	84612	D
Parp2 poly (ADP-ribose) polymerase family, member 2	10038	D
Pawr PRKC, apoptosis, WT1, regulator	5074	D
Pbrm1 polybromo 1	55193	I
Pcsk5 proprotein convertase subtilisin/kexin type 5	5125	D
Pde4dip phosphodiesterase 4D interacting protein (myomegalin)	9659	D
Pde6b phosphodiesterase 6B, cGMP-specific, rod, beta (congenital stationary night blindness 3, autosomal dominant)	50940	D
Pde6d phosphodiesterase 6D, cGMP-specific, rod, delta	5147	D
Pde9a phosphodiesterase 9A	5152	I
Pex11a Peroxisomal biogenesis factor 11A	8800	D
Pex6 peroxisomal biogenesis factor 6	5190	I
Pgap1 GPI deacylase	80055	D
Phlda2 pleckstrin homology-like domain, family A, member 2	7262	I
Phtf1 Putative homeodomain transcription factor 1	10745	D
Pik3ip1 phosphoinositide-3-kinase interacting protein 1	113791	I
Piwil4 piwi-like homolog 4 (*Drosophila*)	143689	D
Pknox2 PBX/knotted 1 homeobox 2	63876	D
Plce1 phospholipase C, epsilon 1	51196	D
Plekha8 Pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 8	84725	D
Plekhk1 pleckstrin homology domain containing, family K member 1	219790	D
Plxnd1 Plexin D1	23129	I
Pmp22 peripheral myelin protein 22	5376	D
Pms2l1 postmeiotic segregation increased 2-like 1	5379	D
Ppara peroxisome proliferator-activated receptor alpha	5465	I
Ppard peroxisome proliferator-activated receptor delta	5467	D
Ppp1r9a protein phosphatase 1, regulatory (inhibitor) subunit 9A	55607	D
Prr14 proline rich 14	78994	D
Prss23 protease, serine, 23	11098	D
Psg6 pregnancy specific beta-1-glycoprotein 6	5675	D
Psmd9 proteasome 26S subunit, non-ATPase, 9	5715	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Ptcd1 pentatricopeptide repeat domain 1	26024	D
Ptprm protein tyrosine phosphatase, receptor type, M	5797	D
Pus10/Flj32312 hypothetical protein FLJ32312	150962	D
Pus7l pseudouridylate synthase 7 homolog (*S. cerevisiae*)-like	83448	D
Qrsl1 Glutaminyl-tRNA synthase (glutamine-hydrolyzing)-like 1	55278	D
Rad18 RAD18 homolog (*S. cerevisiae*)	56852	D
Rad23b RAD23 homolog B (*S. cerevisiae*)	5887	D
Rad51c RAD51 homolog C (*S. cerevisiae*)	5889	D
Rad54b RAD54 homolog B (*S. cerevisiae*)	25788	D
Rad54l2 RAD54-like 2 (*S. cerevisiae*)	23132	D
Ralgps1 Ral GEF with PH domain and SH3 binding motif 1	9649	D
Ralgps2 Ral GEF with PH domain and SH3 binding motif 2	55103	D
Rap1gds1 RAP1, GTP-GDP dissociation stimulator 1	5910	D
Rasgrf2 Ras protein-specific guanine nucleotide-releasing factor 2	5924	D
Rbm45/Drb1 Developmentally regulated RNA-binding protein 1	129831	D
Rdx Radixin	5962	D
Rfpl2 ret finger protein-like 2	10739	D
Rgs9bp/Rgs9 regulator of G-protein signaling 9	388531	D
Rnf2 ring finger protein 2	6045	D
Rpap1 RNA polymerase II associated protein 1	26015	D
Rpl10 ribosomal protein, large, 10	6134	I
Rrp1 ribosomal RNA processing 1 homolog (*S. cerevisiae*)	8568	I
Rps3a ribosomal protein S3A	6189	D
Rps16 ribosomal protein S16	6217	D
Rttn rotatin	25914	D
Rundc2c RUN domain containing 2C	440352	D
Sccpdh saccharopine dehydrogenase (putative)	51097	D
Sclt1 sodium channel and clathrin linker 1	132320	D
Scoc short coiled-coil protein	60592	D
Sdccag8 serologically defined colon cancer antigen 8	10806	D
Sdhb succinate dehydrogenase complex, subunit B, iron sulfur (lp)	6390	D
Sec22a SEC22 vesicle trafficking protein homolog C (*S. cerevisiae*)	26984	D
Sec23ip SEC23 interacting protein	11196	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Sephs1 selenophosphate synthetase 1	22929	D
Sept2 septin 2	4735	I
Sept8 septin 8	23176	D
Setd4 SET domain containing 4	54093	D
Setd8 SET domain containing (lysine methyltransferase) 8	387893	D
Sfrs10 splicing factor, arginine/serine-rich 10 (transformer 2 homolog, *Drosophila*)	6434	D
Sfrs2ip splicing factor, arginine/serine-rich 2, interacting protein	9169	D
Sfrs4 splicing factor, arginine/serine-rich 4	6429	I
Sgta small glutamine-rich tetratricopeptide repeat (TPR)-containing, alpha	6449	D
Siah1 seven in absentia homolog 1 (*Drosophila*)	6477	I
Sipa1l3 signal-induced proliferation-associated 1 like 3	23094	D
Sla2 Src-like-adaptor 2	84174	I
Slc16a1 solute carrier family 16 (monocarboxylic acid transporters), member 1	6566	D
Slc18a2 solute carrier family 18 (vesicular monoamine), member 2	6571	D
Slc19a2 solute carrier family 19 (thiamine transporter), member 2	10560	D
Slc25a23 solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 23	79085	D (HT)
Slc2a13 solute carrier family 2 (facilitated glucose transporter), member 13	114134	D
Slc30a5 solute carrier family 30 (zinc transporter), member 5	64924	D
Slc39a8 solute carrier family 39 (zinc transporter), member 8	64116	D
Slc45a3 solute carrier family 45, member 3	85414	I
Smek2 AW011752 KIAA1387 protein	57223	D
Smg5 Smg-5 homolog, nonsense mediated mRNA decay factor (*C. elegans*)	23381	D
Sorbs3 sorbin and SH3 domain containing 3	10174	D
Spag10 sperm associated antigen 10	54740	I
Sphk2 sphingosine kinase 2	56848	D
Ssh3 slingshot homolog 3 (*Drosophila*)	54961	D
St3gal3 ST3 beta-galactoside alpha-2,3-sialyltransferase 3	6487	D
St8sia1 ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1	6489	D
Stag3 stromal antigen 3	10734	D
Steap3 STEAP family member 3	55240	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Strbp Spermatid perinuclear RNA binding protein	55342	D
Stx6 syntaxin 6	10228	D
Suhw2 suppressor of hairy wing homolog 2 (*Drosophila*)	140883	D
Sycp2 synaptonemal complex protein 2	10388	D
Syne1 synaptic nuclear envelope 1	23345	D
Synpo synaptopodin	11346	D
Tas2r14 Taste receptor, type 2, member 14	50840	D
Tbc1d24 TBC1 domain family, member 24	57465	D
Tc2n/Mtac2d1 membrane targeting (tandem) C2 domain containing 1	123036	I
Tdrkh tudor and KH domain containing	11022	D
Tex261 testis expressed sequence 261	113419	I
Tfec transcription factor EC	22797	D
Tgfb3 transforming growth factor, beta 3	7043	I
Tgm2 transglutaminase 2, C polypeptide	7052	D
Thap9 THAP domain containing 9	79725	D
Thbs1 thrombospondin 1	7057	I
Tigd7 tigger transposable element derived 7	91151	D
Tjp3 tight junction protein 3 (zona occludens 3)	27134	D (HT)
Tk1 thymidine kinase 1, soluble	7083	I
Tlr9 toll-like receptor 9	54106	D
Tmem126b transmembrane protein 126B	55863	D
Tmem169 transmembrane protein 169	92691	D
Tmem30b transmembrane protein 30B	161291	D
Tmem41a transmembrane protein 41a	90407	D
Tmprss6 transmembrane protease, serine 6	164656	D
Tmtc1 transmembrane and tetratricopeptide repeat containing 1	83857	D
Tnfrsf11A tumor necrosis factor receptor superfamily, member 11a, NFKB activator	8792	D
Tnk1 tyrosine kinase, non-receptor, 1	8711	D
Top1mt topoisomerase (DNA) I, mitochondrial	116447	D
Tpd52 tumor protein D52	7163	D (HT)
Tpp2 tripeptidyl peptidase II	7174	D
Trabd TaB domain containing	80305	D
Trim6 tripartite motif-containing 6	117854	D
Trip11 thyroid hormone receptor interactor 11	9321	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Trove2 TROVE domain family, member 2	6738	D
Trpc1 transient receptor potential cation channel, subfamily C, member 1	7220	D
Trpm7 transient receptor potential cation channel, subfamily M, member 7	54822	I
Trspap1 tRNA selenocysteine associated protein 1	54952	I
Tshz2 Teashirt family zinc finger 2	128553	D
Ttc18 tetratricopeptide repeat domain 18	118491	D
Ttc30b tetratricopeptide repeat domain 30B	150737	D (HT)
Txndc8 thioredoxin domain containing 8	255220	I
Ube2i ubiquitin-conjugating enzyme E2I	7329	I
Ugt8 UDP glycosyltransferase 8 (UDP-galactose ceramide galactosyltransferase)	7368	D
Urg4 up-regulated gene 4	55665	D
Usp6nl USP6 N-terminal like	9712	D
Usp7 Ubiquitin specific peptidase 7 (herpes virus-associated)	7874	I
Wdr20 WD repeat domain 20	91833	D
Wdr23 WD repeat domain 23	80344	D
Wdr34 WD repeat domain 34	89891	D
Wdr52 WD repeat domain 52	55779	D
Wdr71 WD repeat domain 71	80227	D
Wee1 WEE1 homolog (*S. pombe*)	7465	D
Xpr1 xenotropic and polytropic retrovirus receptor	9213	D
Zc3H12c zinc finger CCCH-type containing 12C	85463	D
Zdhhc11 zinc finger, DHHC-type containing 11	79844	D
Zdhhc14 zinc finger, DHHC domain containing 14	79683	D
Zdhhc21 zinc finger, DHHC domain containing 21	340481	D
Zdhhc24 zinc finger, DHHC-type containing 24	254359	D
Zdhhc4 zinc finger, DHHC-type containing 4	55146	I
Zmiz2 zinc finger, MIZ-type containing 2	83637	I
Zmym5 zinc finger, MYM-type 5	9205	D
Znf10 zinc finger protein 10	7556	D
Znf169 zinc finger protein 169	169841	I
Znf204 zinc finger protein 204	7754	D
Znf236 zinc finger protein 236	7776	D
Znf24 zinc finger protein 24	7572	D

TABLE 7-continued

Complete list of blood candidate biomarker genes for mood derived from primary bipolar cohort analysis.

Gene Symbol/Name	Entrez Gene ID	Human Blood Data
Znf318 zinc finger protein 318	24149	D
Znf492 zinc finger protein 492	57615	D
Znf502 zinc finger protein 502	91392	D (HT)
Znf540 zinc finger protein 540	163255	D
Znf557 zinc finger protein 557	79230	D
Znf569 zinc finger protein 569	148266	D
Znf585A zinc finger protein 585A	199704	D
Znf608 zinc finger protein 608	57507	D
Znf614 zinc finger protein 614	80110	D
Znf624 zinc finger protein 624	57547	D
Znf667 zinc finger protein 667	63934	D
Znf684 zinc finger protein 684	127396	D
Znf710 zinc finger protein 710	374655	D
Znf711 zinc finger protein 711	7552	D
Znf718 zinc finger protein 718	255403	D
Znf793 zinc finger protein 793	390927	I
ZNF91 zinc finger protein 91	7644	D
Zscan5 zinc finger and SCAN domain containing 5	79149	D

All-low and high threshold candidate genes.
Known genes shown only.
For human blood data:
I - increased in high mood (mania);
D - decreased in high mood (mania)/increased in low mood (depression).
(HT)—High threshold.

[0206] The nucleic acid sequences provided herein represent a region or a segment of the genes listed in one or more of the tables. The completed nucleic acid sequences for the genes listed in the tables are readily obtained from a public database (e.g., NCBI) using the gene identification (Gene ID) number and the gene names provided in the tables. Expression profiles of the genes listed in the tables are performed using either oligos, regions or segments of the genes or full or partial cDNA sequences, ESTs in a microarray format. Similarly, the presence or absence of the protein products or peptide fragments thereof of the genes listed in the tables are also analyzed for predictive and dignostic purposes. Antibodies to the protein or peptides are placed in an array format for serial or parallel expression profiling.

Mbp myelin basic protein Entrez Gene ID No. 4155 (SEQ ID NO: 1)

gaagataaccggctcattcacttcctcccagaagacgcgtggtagcgagt

aggcacaggcgtgcacctgctcccgaattactcaccgagacacacgggct

-continued

gagcagacggcccctgtgatggagacaaagagctcttctgaccatatcct

tcttaacacccgctggcatctcctttcgcgcctccctccctaacctactg

acccacctttgattttagcgcacctgtgattgataggccttccaaagag

tcccacgctggcatcaccctccccgaggacggagatgaggagtagtcagc

gtgatgccaaaacgcgtcttcttaatccaattctaattctgaatgtttcg

tgtgggcttaataccatgtctattaatatatagcctcgatgatgagagag

ttacaaagaacaaaactccagacacaaacctccaaatttttcagcagaag

cactctgcgtcgctgagctgaggtcggctctgcgatccatacgtggccgc

acccacacagcacgtgctgtgacgatggctgaac

-continued

Edg2 Endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 2 Entrez Gene ID No. 1902,

(SEQ ID NO: 2)

aatgagcgccacctttaggcagatcctctgctgccagcgcagtgagaacc

ccaccggccccacagaaagctcagaccgctcggcttcctccctcaaccac

accatcttggctggagttcacagcaatgaccactctgtggtttagaacgg

aaactgagatgaggaaccagccgtcctctcttggaggataaacagcctcc

ccctacccaattgccagggcaaggtggggtgtgagagaggagaaaagtca

actcatgtacttaaacactaaccaatgacagtatttgttcctggacccca

caagacttgatatatattgaaaattagcttatgtgacaaccctcatcttg

atccccatcccttctgaaagtaggaagttggagctcttgcaatggaattc

aagaacagactctggagtgtccattta

Fgfr1 fibroblast growth factor receptor 1 Entrez Gene ID No. 2260,

(SEQ ID NO: 3)

ctcctctccacctgctggtgagaggtgcaaagaggcagatctttgctgcc

agccacttcatcccctcccagatgttggaccaacacccctccctgccacc

aggcactgcctgagggcagggagtgggagccaatgaacaggcatgcaagt

gagagcttcctgagctttctcctgtcggtttggtctgttttgccttcacc

cataagcccctcgcactctggtggcaggtgcttgtcctcagggctacagc

agtagggaggtcagtgcttcgagccacgattgaaggtgacctctgcccca

gataggtggtgccagtggcttattaattccgatactagtttgctttgctg

accaaatgcctggtaccagaggatggtgaggcgaaggcaggttgggggca

gtgttgtggcctggggccagccaacactggggctctgtatatagctatga

agaaaacacaaagttgataaatctgagtatatatttacatgtctttttaa

aagggtcgttaccagagatttacccatcg

Fzd3 frizzled homolog 3 (*Drosophila*) Entrez Gene ID No. 7976

(SEQ ID NO: 4)

aatcctaaatgtgtggtgactgctttgtagtgaactttcatatactataa

actagttgtgagataacattctggtagctcagttaataaaacaatttcag

aattaaagaaattttctatgcaaggtttacttctcagatgaacagtagga

ctttgtagttttatttccactaagtgaaaaaagaactgtgtttttaaact

gtaggagaatttaataaatcagcaagggtattttagctaatagaataaaa

gtgcaacagaagaatttgattagtctatgaaaggttctcttaaaattcta

tcgaaataatcttcatgcagagatattcagggtttggattagcagtggaa

taaagagatgggcattgtttcccctataattgtgctgtttttataacttt

tgtaaatattacttttctggctgtgtttttataacttatccatatgcat

gatggaaaaattttaatttgtagccatcttttcccat

Mag myelin-associated glycoprotein Entrez Gene ID No. 4099

(SEQ ID NO: 5)

tttggcgtcgtcctcaagttatattagaatcgtgtcctcccggctttggc

caacttactattctaggacttgattccttcattcagtcacaatttattga

-continued

gcaccgactttgcatcaacctcttgctgaagataacagtgctgacaatat

acagccctgccctcagagcttatatagtagaggagaaaaagtgaacccat

aatatacagtcagtagcgagtatttactaagtactttctatttgcgaggc

cctgataaaagtactgtcctggccaggcgcggtggctcacgcctgtaatt

ccagcactttgggaggtcgaggtgggcagatcacctaaggtcaggagttc

gagatcagcctggctaacatggggaaaccccgtctctactaaaaatggaa

aaattagctgggcatggtggcgggcgcctgtaatcccagctactcgggag

gctgagacaggagaatgacttgaacccaggagttgcagtggccaagataa

gatagcgccattgtactcc

Pmp22 peripheral myelin protein Entrez Gene ID No. 225376

(SEQ ID NO: 6)

tgtgaagctttacgcgcacacggacaaaatgcccaaactggagcccttgc

aaaaacacggcttgtggcattggcatacttgcccttacaggtggagtatc

ttcgtcacacatctaaatgagaaatcagtgacaacaagtctttgaaatgg

tgctatggatttaccattccttattatcactaatcatctaaacaactcac

tggaaatccaattaacaattttacaacataagatagaatggagacctgaa

taattctgtgtaatataaatggtttataactgcttttgtacctagctagg

ctgctattattactataatgagtaaatcataaagccttcatcactcccac

atttttcttacggtcggagcatcagaacaagcgtctagactccttgggac

cgtgagttcctagagcttggctgggtctaggctgttctgtgcctccaagg

actgtctggcaatgacttgtattggccaccaactgtagatgtatatatgg

tgcccttctgatgctaagactccagacctttttgt

Ugt8 UDP glycosyltransferase 8 (UDP-galactose ceramide galactosyltransferase) Entrez Gene ID No. 7368

(SEQ ID NO: 7)

attccatgtcattctgtttacttagcacttgcactacccttgtnggttga

gtgtatgctttatttgtttctagtttgaaatcccacatctgatagctgag

agtaggcaaatacaacatttacctaatgtcattcactaacatggaagagt

tgtgaaaattctagagtgctgtaaatccttggcatacactatgacaaaca

acttcattactctcccaccaggagctgctctcctgcacttagaaataatg

tcacaagtagttttctaatgtacaatgcagacaaatgtactgctctctga

atacttgaagaaatggtattatacatacatagaaacttattagttatacc

ttttcacaatcttattacgatgttgccgttaaaagggaaaaaagacacag

gcaatgaatggtgggatagtaagaggacttagagtgtatgaatgagttga

ttttacttttttggaatttgattaagttgacagtaggcactgattggatg

attaaacataagttaatctccactgtgat

Erbb3 Neuregulin receptor (v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)) Entrez Gene ID No. 2065

(SEQ ID NO: 8)

cttatggtatgtagccagctgtgcactttcttctctttcccaaccccagg

aaaggttttccttattttgtgtgctttcccagtcccattcctcagcttct

-continued

tcacaggcactcctggagatatgaaggattactctccatatcccttcctc

tcaggctcttgactacttggaactaggctcttatgtgtgcctttgtttcc

catcagactgtcaagaagaggaaagggaggaaacctagcagaggaaagtg

taattttggtttatgactcttaacccctagaaagacagaagcttaaaat

ctgtgaagaaagaggttaggagtagatattgattactatcataattcagc

acttaactatgagccaggcatcatactaaacttcacctacattatctcac

t

Igfbp4 insulin-like growth factor binding protein Entrez Gene ID No. 43487

(SEQ ID NO: 9)

agagacatgtaccttgaccatcgtccttcctctcaagctagcccagaggg

tgggagcctaaggaagcgtggggtagcagatggagtaatggtcacgaggt

ccagacccactcccaaagctcagacttgccaggctccctttctcttcttc

cccaggtccttccttttaggtctggttgttgcaccatctgcttggttggct

ggcagctgagagccctgctgtgggagagcgaaggggtcaaaggaagact

tgaagcacagagggctagggaggtggggtacatttctctgagcagtcagg

gtgggaagaaagaatgcaagagtggactgaatgtgcctaatggagaagac

ccacgtgctaggggatgaggggcttcctgggtcctgttcccctaccccat

ttgtggtcacagccatgaagtcaccgggatgaacctatccttccagtggc

tcgctccctgtagctctgcctccctctccatatctccttcccctacacct

ccctccccacacctccctactcccctgggcatcttctggcttgactggat

gg

Igfbp6 insulin-like growth factor binding protein Entrez Gene ID No. 63489

(SEQ ID NO: 10)

gcgcgcctgctgttgcagaggagaatcctaaggagagtaaaccccaagca

ggcactgcccgcccacaggatgtgaaccgcagagaccaacagaggaatcc

aggcacctctaccacgccctcccagcccaattctgcgggtgtccaagaca

ctgagatgggcccatgccgtagacatctggactcagtgctgcagcaactc

cagactgaggtctaccgaggggctcaaacactctacgtgcccaattgtga

ccatcgaggcttctaccggaagcggcagtgccgctcctcccaggggcagc

gccgaggtccctgctggtgtgtggatcggatgggcaagtccctgccaggg

tctccagatggcaatggaagctcctcctgccccactgggagtagcggcta

aagctgggggatagaggggctgcagggccactggaaggaacatggagctg

tcatcactcaac

Pde6 dphosphodiesterase 6D, cGMP-specific, rod, delta Entrez Gene ID No. 5147

(SEQ IDNO: 11)

gaagcaacgttttcatctgattggaaataccgtattgctgaaaagaagaa

aggccttttaatggcttttgaacaaagcagaaaagtttgagcttctcac

cttcagtcttagctcttgaacctgttgagaaagaggataagagacaaata

cggaaaagagtttcagaaagcagaatctgtgtcagcccactggaaggaaa

agcgaatcaaccgattcagtgatgttagtgcatccagaaacaggcttttg

-continued

ggaaaagcttgacctgagctgattaaatcctgaagcacaanggaagcagc

cacatcaaaaagttagcatgagagcagtggcgtgctcatcctctggtagc

ctttactgggcatttgtggagtaagagagaaaagaaaagcaggaatgtta

agatatgctactaccttcaggaaa

Ptprm protein tyrosine phosphatase, receptor type, M Entrez Gene ID No. 5797

(SEQ ID NO: 12)

gagcagcgtagacagctggtaaactgaagagcacaactatattcttatga

aggaatttgtacctttggggtattattttgtggcccgtgaccctcgttat

tgttacagctgagtgtatgtttttgttctgtggagaatgctatctggcat

tatggtaatatattatttttaggtaatatttgtactttaacatgttgcata

atatatgcttatgtagctttccaggactaacagataaatgtgtaatgaac

aaagatatgttgtatgagtcgtcgtttctgtcagatttgtattgtttcca

agggaaaagcttgggggaggactcagttcacaaaatgcaaaactcaacga

tcagattcacggacccagagcttttccatgtgt

Atp2c1 ATPase, Ca++-sequestering Entrez Gene ID No. 27032

(SEQ ID NO: 13)

tttttctcttctggcttcataaatgccttgctgtataaattgaaatattg

atactgaactgtctttttaatgatgacctaactttattcaacccatcgga

atttactttttccctgaaataagatcttttccactggtctactacctgac

cataaacatgtctgcatttgaattctctaaaccctaaatctgtgtctatg

Atxn1 Ataxin Entrez Gene ID No. 16310

(SEQ ID NO: 14)

aaacagagcagccacgggctcgaaccgaatccccgccgtccttagaaacg

gatttttttttgttttgttttgttttctggcagagtctcgatcaccaccc

tactnccacccccactaaggttcttgctcaatctccctagaaaacctgaa

ttgtttcatccctttcagtcagcccctacgtggtctgaaacaaaatgaa

agcacaagccacggagtttaagaaggcagcctgaaggcgggggctgaag

aggggtcggggcgctgcagagtcagccaagtagccaaggaagggccccct

ccgcgtcgcgacggccctcgccccccgcccggcgcgcgcgcgcacaaata

cacacacacagtcactcacacactcactcacactcacgccgcgctccgac

accgcctcacctctcgctccgcccgtccggccccctccctgcccccctgc

gggaccggcctgcgcgcagcactggaaccacgtaggaggaggcggcgg

Btg1 B-cell translocation gene 1, anti-proliferative Entrez Gene ID No. 694

(SEQ ID NO: 15)

gaaggtatacagactgccaacattttagacttatctctcagtctctgcca

ctgaacttttatatatggctgctatcaaaatataaccggtttattttcat

atttggaactaatataacagtatcacaaaatctttttacagtaagatagtt

tgtaataccagccgacccagctgcttaactgagtccttaaatcatttaat

atatgggactgtaaatagagaaatctgtacattannagatctgatttctg

gttatgcctatagatctttattttctttatccctatagatcattttcttc

tgatgttaagtgttatatttttgaaatgctcctaaacaagtaagccttag

-continued

attgtattaatcctgaggattgataccatttcctcaacactttgtggagt

atgattgacaccagtttttttttgactgcaggtttaacttggcttatcac

ttttcgtattgctcagtatgtccaag

C6orf182 chromosome 6 open reading frame Entrez Gene ID No. 182285753

(SEQ ID NO: 16)

cacagttttgtgaagaagcctttaatccaagttttatgagacagtaggaa

cctatagctacttaatttttaggagacagtaattcagttgcagaagcttt

cctctgctattcccattctcttttacaaaactagtttttttttaaaaaatc

aatgatcaattttatcttactactttataacttttgctgacttttattct

ttgcattgtatgtaatgtccatcagtataaattgagactgtgaatttcta

ggacccacaaaagtggtatttttttttttgtacaagaaagtatagaggaag

aggaagctgggcattaaattacctcatccagcag

Dicer1 Dicer1, Dcr-1 homolog (*Drosophila*) Entrez Gene ID No. 23405

(SEQ ID NO: 17)

tatcactggtaaggagcccacgaccagacttcatttctgggaaaatgaga

ctttgtgttgatctcatcgtgttggccttgtaaaagtgatctatgcatgt

acagtgttcatgcttaatattcaagggatggggcggggaacaaaaggaat

agaaagaattcttttccttgttatttggggagcacgtattgctttataac

tttggttgttgggagtatggctatcatataccctcatcagtgtcatttta

tatctgcctaattagagaaattttaaccttagtattttgatgtgttttcc

ccattttatcctccgcaaatatctttctcttgcccattcagtgctgcttt

tggtttttgatttagttgtatattctggatgtatttccacagccttttat

tgttcttcc

Dnajc6 DnaJ (Hsp40) homolog, subfamily C, member Entrez Gene ID No. 69829

(SEQ ID NO: 18)

gatccagtatgtgcttgtcttatttaaaattggaatgtgagacatgttgc

tgtgacctgtttttctttctcattcacatttgtagatattgtgtgaacta

cagtatataatgataacaattaaaaggatattctgtggatgtcacgtatt

ttgaaatgatagaactacattagctttgtatcatgtttggataattcatc

aatgttcacagtttaaaacatcattaaacattatgtaattacaatgagaa

agaatcttacttaaatttggagattttccccccacatctcttttccggata

cattataattctggacccctatttatctcaaaactcttaatatatgcaga

ccaacaggtctttgcattccttttaaataactggttgtgacaaagcttgt

tgttgatcagattcactg

Ednrb endothelin receptor type B Entrez Gene ID No. 1910

(SEQ ID NO: 19)

aactgctttaagtcatgcttatgctgctggtgccagtcatttgaagaaaa

acagtccttggaggaaaagcagtcgtgcttaaagttcaaagctaatgatc

acggatatgacaacttccgttccagtaataaatacagctcatcttgaaa

-continued

Elovl5 ELOVL family member 5, elongation of long chain fatty acids (yeast) Entrez Gene ID No. 60481

(SEQ ID NO: 20)

tcgaggtatcagcagctctgtcctcagaatgggtcccccacttcacacag

ttgtaggatggctacagcagctccaagcagcacattcagaggaagaagaa

aaaatgtttcatttgtgtggttttaagcataaagaagttgtttcccagag

ctctttgccagctgtcatccctcaaaactcactggccacaattgcttcac

atgcccctgcctgaaccagccaccagaggggttaggaccaccacagatt

gactagatcctaaggattctctacctggggctggagttcgggtcaggtgc

ccgggaagcactgtgcggtgtgggaggatg

Gnal guanine nucleotide binding protein, alpha stimulating, olfactory type Entrez Gene ID No. 2774

(SEQ ID NO: 21)

tgttgttgtccctattgctggtttattacactgtacagaccacaaaatgt

aatattcttttgtataactactaaagaaaaatccttgtagancnnnnnnc

cttcaccatggctatctatacctgtacatgaaatgtgtttgtattgtgct

gaagngcttaatgtcaacattacctgctgnttactctgaaaaaaggaatg

aatggtagctntagaatttaggatattttatcaggttggcactttataaa

atactccctgatttaaaaaattgtaagttatacacgttaatcatccacat

tctatcgacaatgtaccaacatcacaagctgttgcaaccacctgnctgtt

acttctctgagctgttaaaancctggaacttcaatttcaggggggcacaa

at

Klf5 Kruppel-like factor Entrez Gene ID No. 5688

(SEQ ID NO: 22)

ttacagtgcagtttagttaatctattaatactgactcagtgtctgccttt

aaatataaatgatatgttgaaaacttaaggaagcaaatgctacatatatg

caatataaaatagtaatgtgatgctgatgctgttaaccaaagggcagaat

aaataagcaaaatgccaaaaggggtcttaattgaaatgaaaatttaatttt

tgtttttaaaatattgtttatctttatttatttgggggtaatattgtaag

ttttttagaagacaattttcataacttgataaattatagttttgtttgtt

agaaaagtagctcttaaaagatgtaaatagatgacaaacgatgtaaataa

ttttgtaagaggcttcaaaatgtttatacgtggaaacacacctacatgaa

aagcagaaatcggttgctgttttgcttcttttttccctcttatttttgtat

tgtggtcatttcctatgcaaataatggagcaaacagctgtatagttgtag

aat

Lin7a lin 7 homolog a (*C. elegans*) Entrez Gene ID No. 8825

(SEQ ID NO: 23)

aatggaccaattttagctcaacttttagtttgttagaagcaagtgtagga

actctagcactgtagtttttaattattgcttgtatctattattattaatt

ccaacagagtataatgtatatttattctataaaatatatattatcagagt

gcatttgttacaacttaggttctttttcttaccaagtattaagnaatctag

-continued

taagagnaatactagncaaaggacctagncctgtgnaacagnntctcgt

atgttatatacataaacccactctgc

Manea mannosidase, endo-alpha Entrez Gene ID No. 79694

(SEQ ID NO: 24)

gtgtttctgctttacagtgctgaattccatattttagaagctatgaaagt

ccttttatgaaaaagttactgattgcttctcagttattaggaaaacagtt

gtttcacaattattatgtagatatgatgcccaaatatcatttttagtata

tcttgtcgatctttaagttgttactattgtgttattcatgtctttaaatc

agataccaaatatttttttaggaaagaaaaatgttattactgtcattaggt

tgtcttttaatactttaagttattttgacgaaaagtaatagagaaaattt

acttagcattttagattctagagacatggaaatgaaaattatttttatgtc

tagagtaggtcctgaagtttggctttacattaagtttagcactgtatcag

aatgaagaaactaatattttacataaaaactaatactttcaatttttttat

atagtaatatccccattttgtaaatgttagacttttatcatacctgtaa

Nupl1 nucleoporin like 19818

(SEQ ID NO: 25)

aaaataggcattgcatacacatatgcacacgtatgtgcacgtgccacaca

ttttttgtataatgttgggtttgattataaaagtgttgtcaaatgtttta

tttatctgcatntagcagtggttggctttttttgaattgaaatttttgcgc

attgatgcattgaaataaggaaaattatttatctctgagcactaaactta

tttttgcatatttctgtaatattgcagtccccagatccagaacatgggaa

gttagggaaaatgtgtgattttgtgttttgaattactgtcagaattacat

acacaattacaacaaacttttttttaaaagacatttcattgtactgcaaaa

atctgaatatttatatttcttgttttttttctttatatgttttgcatttta

atatgttgagccactgg

Pde6b phosphodiesterase 6B, cGMP-specific, rod, beta (congenital stationary night blindness 3, autosomal dominant) Entrez Gene ID No. 5158

(SEQ ID NO: 26)

tcttttgaaaatgctggcctgctgcacctgttctcagaggggcatgagaa

aagaggattctcttcctagggctgtgggaagccccttgagatgttggaag

cagggcagggcaggagaacccagggagggcacagagctgtggacgagggc

tgggaaagccatcccgcctccccaggggtctccgaggagtgctgctgtg

gccaaaccaggggggccactttgtgctttctgtttaggtgatgggatgct

tctatctcctcagcaccccacaccaaatcccctgttattgccagatgatg

tgcctaatgatcaaattgctt

Slc25a23 solute carrier family 25 (mitochondrial carrier; phosphate carrier), member Entrez Gene ID No. 2379085

(SEQ ID NO: 27)

ctcccaccttctaggcgaatagtccccagagctgtgttcctccaaggggt

ccgaggaatcactcactcctggaggctggcaaggagacagtctgaggcca

gggacacatgaagggatgtccccaccccagcactatcagggcctccccag

gcttccagagttgaaagccaggagaaaatcggcaaagaccacccttccct

-continued

aaacccaagcacccaatgatgcaaaaaacaaaaacaaaaaaaaaccacca

aatccccaaattcattccagatctatttttctaccagagagaggagcaaa

gtcctcctcccctgcgcccttacattctgcacttcatagttggattctga

gcttaggatcatctggagaccccatggagggacttgga

Synpo synaptopodin Entrez Gene ID No. 11346

(SEQ ID NO: 28)

ggtgatgttagatctggaacccccaagtgaggctggagggagttaaggtc

agtatggaagatagggttgggacagggtgctttggaatgaaagagtgacc

ttagagggctccttgggcctcaggaatgctcctgctgctgtgaagatgag

aaggtgctcttactcagttaatgatgagtgactatatttaccaaagcccc

tacctgctgctgggtcccttgtagcacaggagactggggctaagggcccc

tcccagggaagggacaccatcaggcctctggctgaggcagtagcatagag

gatccatttctacctgcatttcccagaggactagcaggaggcagccttga

gaaaccggcagttcccaagccagcgcctggctgttctctcattgtcactg

ccctctccccaacctctcctctaacccactagagattgcctgtgtcctgc

c

Tgm2 transglutaminase 2, C polypeptide Entrez Gene ID No. 7052

(SEQ ID NO: 29)

ccttcagaccccagttctaggggagaagagccctggacacccctgctcta

cccatgagcctgcccgctgcaatgcctagacttcccaacagccttagctg

ccagtgctggtcactaaccaacaaggttggncaccccagctaccccttct

ttgcagggctaaggcccccaaacatagcccctgccccggaggaagcttgg

ggaacccatgagttgtcagctttgactttatctcctgctctttctacatg

actgggcctcccttgggctggaagaattggggattctctattggaggtga

gatcacagcctccagggccccccaaatcccagggaaggacttggagagaa

tcatgctgttgcatttagaactttctgctttgcacaggaaagagtcacac

aattaatcaacatgtatattttctctatacatagagctctatttctctac

ggtttta

Tjp3 tight junction protein 3 (*zona occludens* 3) Entrez Gene ID No. 27134

(SEQ ID NO: 30)

acacggatgtggatgatgagccccggctccagccctggcccggtcctcg

gagcccgtgcaggcagatgagtcccagagcccgagggatcgtgggagaat

ctcggctcatcaggggcccaggtggacagccgccacccccagggacagt

ggcgacaggacagcatgcgaacctatgaacgggaagccctgaagaaaaag

tttangcgagtncatgatgcggagtcctccgatgaag

Tpd52 tumor protein D Entrez Gene ID No. 527163

(SEQ ID NO: 31)

agatgtgtccaaagagatccttacattaaggtcatttgtcagaatgatgt

ttgngtttgtttagagttggctgacctccaactcctggggtcaaggaatc

ctactgccttagcttcccaaatagctaggactataggcatgcacccggcc

atgtgttttatttatagctcttaaagcccagatgaagaaatcacatttttt

gcccatagtgaagaaacatttggccattgattagtccttattttcagtga

-continued

ctgtcctgttttcattagattagagagaccctgtgtgggccacagttaat

ataaaccattatcacttttaagtaaacctgcacatcttagatttcataat

ttccttattgttctgactcaaaatgaactaagagcttttcacttttttgtt

tgtaagttctcagagagtcgggtctgcaagtgcttt

Trpc1 transient receptor potential cation channel, subfamily C, member Entrez Gene ID No. 17220

(SEQ ID NO: 32)

gggtgcaggtaacccttggtctgtaagcaccaccgatccagggatcattg

tctaaataggttactattgtttgtttcatcttgcttttgcatttttattt

tttaatttccaaatttaagtgttccctctttggggcaaattcttataaa

aatgtttattgtaaagttatatattttgtctacgatgggattatgcactt

cccaattgggattttacatctggatttttagtcattctaaaaaacaccta

attattaaaacatttatagagtgcctactgtatgcatgagttgagttgct

tctgaggtacattttga

Bclaf1 BCL2-associated transcription factor Entrez Gene ID No. 19774

(SEQ ID NO: 33)

gttactagactataccttggcctttaaaaatgaatctcactgattaaaga

gaacaggcattattaggatagcattccaccacaactagaaacattcaaat

aatgtgtcttaattgtaatctgtatataggaaaatttttcctatggatat

ttttggtgttttaccacagtgaactgatttgtagcacttatgaagtgcag

aaggtaatattcttgaaaatagaaaaaggttgggtgagcaggctttaatg

ccttttcccccaagaatatacatcgaatttttcttaatcttttggggttgg

ccagcttccagatttcattaataatgagctctgcctttaataaaagtaca

tgatcatagctacactgtatgtttaggtggtgtgaaatgatttataatca

cagcttgaactgtgtttgcttggtactgtca

Gosr2 golgi SNAP receptor complex member Entrez Gene ID No. 29570

(SEQ ID NO: 34)

ccctgtgcctcagtgacatgtagatgactgactgccaatacttgtcacca

ttccctggaagcagctacctaggggaaacaagatgtagtgctattgccga

taacaagtaagattttccacactanannnnggtgtttctcttttctaaag

tgaggccagtgttatttcccgggagtgttcagtcttgaccctagtcactg

atttttttctagttgttaatagagtggttggcttttaaggttcagagactg

tggcttggcacctgcgcccaggctttgtgggcctttgcccccttagaaagt

agctgtaggcaaagatttgtgattttccaattacagtctcagctctagtt

ttagtatctctaattctttggttcccttctcttccctgaaatatattagc

acctgccagccaggccctcattttgcccagccagtgtgggcagatcccac

cgtggagacatctgtagtgtgtatgtccttgtaacactctgttttcaggg

actacaacctttttccttctgtgaccagccccggattcaggctgtac

Rdx radixin Entrez Gene ID No. 5962

(SEQ ID NO: 35)

ttaacactaattatcacgtctgacaaatgtgtatgtgtggtttcagttct

gtgtacattttaaaggataatggtgaacattttaatgggtttcccttgcc

-continued

ctttccatatttaacctatttccacattctctctcactcacattttctca

gtgtgcccttctcttatctgccatgtccatagccataattccaccatcat

acagatcaggcagtgtttaaaatgatggtaggtagcacagtggacagtct

ttgatcatcatgtagaatatggctatgaatcaggaaagagattagaacat

ttaataatgtatgtacagctggtgcttagtttttttttaatctaaattta

attaccttattggatatttgatatttggttatttaatcacagtcatcttt

aacagcttacactgattggtgttttatctcctgtgatcctttgatggctt

tttttgcctaccatttcacagaggtt

Wdr34 WD repeat domain Entrez Gene ID No. 3489891

(SEQ ID NO: 36)

cgctgggactgacgggcatgtccacctgtactccatgctgcaggcccctc

ccttgacttcgctgcagctctccctcaagtatctgtttgctgtgcgctgg

tccccagtgcggcccttggtttttgcagctgcctctgggaaaggtgacgt

gcagctgtttgatctccagaaaagctcccagaaacccacagttttgatca

agcaaacccaggatgaaagccctgtctactgtctggagttcaacagccag

cagactcagctcttggctgcgggcgatgcccagggcacagtgaaggtgtg

gcagctgagcacagagttcacggaacaagggccccgggaagctgaggacc

tggactgcctggcagcagaggtggcggcctgaggggtcccgggaggcggg

tgcaagccttcgctgtgccgagccttgtgtttctgacgcaagcca

Nefh neurofilament, heavy polypeptide 200 kDa Entrez Gene ID No. 4744

(SEQ ID NO: 37)

ccccaggcgatggacaattatgatagcttatgtagctgaatgtgatacat

gccgaatgccacacgtaaacacttgactataaaaactgccccccctccttt

ccaaataagtgcatttattgcctctatgtgcaactgacagatgaccgcaa

taatgaatgagcagttagaaatacattatgcttgagatgtcttaacctat

tcccaaatgccttctgttttccaaaggagtggtcaagcccttgcccagag

ctctctattctggaagagcggtccaggtggggccgggcactggccactga

attatgccagggcgcactttccactggagttcactttcaattgcttctgt

gcaataaaaccaagtg

Bic BIC transcript

(SEQ ID NO: 38)

gggtaaataacatctgacagctaatgagatatttttttccatacaagataa

aaagatttaatcaaaaaatttcatatttgaaatgaagtcccaaatctang

ttcaagttcaatagcttagccacataatacggttgtgcgagcagagaatc

tacctttccacttctaagcctgtttcttcctccatnnnatggggataata

ctttacaaggttgttgtgaggcttagatgagatagagaattattccataa

gataatcaagtgctacattaatgttatagttagattaatccaagaactag

tcaccctactttattagagaagagaaaagctaatgatttgatttgcagaa

tatttaaggtttggatttctatgcagtttttctaaataaccatcacttac

aaatatgtaaccaaacgtaattgttagtatatttaatgtaaacttgtttt

-continued

aacaactcttctcaacattttgtccaggttattcactgtaaccaaataaa

tctcatgagtctttagttgattta

Bivm basic, immunoglobulin-like variable motif containing Entrez Gene ID No. 54841

(SEQ ID NO: 39)

atcatcatgatcgctcgacatcgatnnnnnnnnnnnnnttttttttttt

ttttttttttttnttnnaagtagaaaacaaaactttatttgatgaaatct

ttttaaaagttccagtatgaantaacaaaatcaacaacctacaaatctct

ttcagtcctttgcatttcaagcaaaatattctcttcagaaaaatgaccat

ttcataatatatatccccttctgtcg

Bnip1BCL2/adenovirus E1B 19 kDa interacting protein 1 Entrez Gene ID No. 662

(SEQ ID NO: 40)

aaaccaccaaagagagcctggcccagacatccagtaccatcactgagagc

ctcatggggatcagcaggatgatggcccagcaggtccagcagagcgagga

ggccatgcagtctctagtcacttcttcacgaacgatcctggatgcaaatg

aagaatttaagtccatgtcgggcaccatccagctgggccggaagcttatc

acaaaatacaatcgccgggagctgacggacaagcttctcatcttccttgc

gctacgcctgtttcttgctacggtcctctatattgtgaaaaagcggctct

ttccatttttgtgagatcccaaaggtgccagttctggccctttcagctcc

tgtttcaggatctgtcctggttcctgagctctaggctgctaagctgagcc

acacacc

C8orf42 chromosome 8 open reading frame 42 Entrez Gene ID No. 157695

(SEQ ID NO: 41)

gaaactctggaaatcacgtgtgtggggagatggggacgcttcccatgttg

tggggagctctgtggctgtgatggctgcagttgccgtgcctctgttggaa

cgcnaagtgcctgcaactcacgtcaatcatagaattgtgacgcacagttg

gcaaaatagttctttatgctatttctcaaaantttgaggacaaacccaga

ttgggattggaatatgcactgtaaatcaaatttttcttatctacaaagac

taatgtaaaaatgatttttttcttctgtgcctgattaaattaactgtggtt

tttaatataaatatttattggtgtgctttgggagaaaaattatctttttct

tgaaagaanttatcaaagcaaatttattatcttcacaagttaatgggaga

atgtggttttgattctgggtgtttgaattgtgtaaacacacagcttcctt

gtg

Camk2d calcium/calmodulin-dependent protein kinase II, delta Entrez Gene ID No. 817

(SEQ ID NO: 42)

atttccctttacattcattatgcaaattcacnttctattcntttctcac

acactactagccagcctccccaaanaaggaaaagggaaaaaagtaagaaa

agaatggaacaaaagaaaaataagaaagcaaacgaaaggaacaaagaaac

aggataaagaaaagagatcacagatttgagaaagaaaaacaattcaattc

agcaaattcaccaaaacaatgtgaatatatcctaaagtgattaaactcag

aaatgatgtgaatttttccagtttacacagtttgaccaaaaacagcatgg

-continued

ctttatgtggtagcaaaccaactgattcttgcttctactttcataagtga

ttttgcccacatatcatcccactttaattgttaatcagcaaaactttcaa

tgaaaaatcatccattttaaccaggatcacacca

Dock9 Dedicator of cytokinesis 9 Entrez Gene ID No. 23348

(SEQ ID NO: 43)

tctttgatcactgcctcttgatttttcctggatcattaagaggcttgaa

gaatactatgtagttgaaccagaggagtagtgtatgtcacatcctcactt

actccttaagccctttctcatggtcttggccctaaaacatattttcaggg

cttgtgacccagtgatcagtggtcacccttaaagtattacagatacgtgc

ctgttttacatgagaggtaactgtttatgtgtataagtcatcttaataaa

ataacatgaaatttattagctgaattgggtagatactgcttttctaagtt

gaacctaacttaagctgatgcagaaactgagtcagaaaagttgctataat

tttaaaatataagaagtaaaagtgaaatcttatgtagcatctttatctca

ttttggtttgtcagtataagtttctgatttccttttaagctctttacttttt

agaaacgtgaatttacaatcccttatccaaaactgctg

Fam13a1 family with sequence similarity 13, member A1 Entrez Gene ID No. 10144

(SEQ ID NO: 44)

gttagtggagttttactgttaatatcatcatgtccccctttgtgtttact

actgtctgaaattactgggatgtagaagcatatttcagtctgaaaattca

gccagcttattttggagaagttgtatcttgttcttgggcatgttagcctt

gtttttcatcccaatttga

Hist1h3b histone cluster 1, H3b Entrez Gene ID No. 8358

(SEQ ID NO: 45)

atggctcgtactaaacagacagctcggaaatccaccggcggtaaagcgcc

acgcaagcagctggctaccaaggctgctcgcaagagcgcgccggctaccg

gcggtgtgaaaaagcctcaccgttaccgtccgggtactgtggctctgcgt

gagatccgccgctaccaaaagtcgaccgagttgctgattcggaagctgcc

gttccagcgcctggtgcgagaaatcgcccaagacttcaagaccgatcttc

gcttccagagctctgcggtaatggcgctgcaggaggcttgtgaggcctac

ttggtagggctctttgaggacacaaacctttgcgccatccatgctaagcg

agtgactattatgcccaaagaca

Hrasls HRAS-like suppressor Entrez Gene ID No. 57110

(SEQ ID NO: 46)

agagcaggccaaccgagcgataagtaccgttgagtttgtgacagctgctg

ttggtgtcttctcattcctgggcttgtttccaaaaggacaaagagcaaaa

tactattaacaatttaccaaagagatattgatattgaaggaatttgggag

gaggaaaagaaacctggggtgaatacttattttcagtgcatcattactgt

tccagattcctatgatggatggc

Ibrdc2 IBR domain containing 3

(SEQ ID NO: 47)

cagccagtggctgtggtctacagaattgtttcatataaaatacgggtaga

gtggtagagtttcaaaactttcgtcatagatatctgggacctttctcagg

-continued

atctgtgttcacacagccaatagatttggaatcaggcctaagagtacaca

tggagggtaaatattaaagtgcgtattatgtacatctagaatccatgtga

cttgcagcctacctgtaatttctatccattgagcatgcatggatataccc

aatagtacacacaaaataaatgtttacttaagagccattctaaaaaannn

nnannnaaatggtttattgtaaatctgcctaaagatttttgcatattat

atatgtgaattttggttgtaagttcataacttacccaagggtatagactc

ataactcttttaaaacagtgcttagtacaatatcctgccatctctgtaaa

aacgctaattgataaccgagtcatttacatgttttcgaacacagaatagc

tcttttctcagcatcattattgctctttcagcatc

Kiaa1729 KIAA1729 protein Entrez Gene ID No. 85460
(SEQ ID NO: 48)

gattccagaatctctacctttaaacactatgttaccacttacttctcttc

aaattttattgagcattagatgtttccagtatttagaagtcaaatgcttc

gtttttaataggaacttacacagtcttttatgttttttatagccctcaa

tgtcactgatgtggattctcccaaactcgatactttgtttgtttttatgt

ccccataataagtctttaagaaaacagggcaagtgagctcaaaatcaaaa

gaaaacccaccaacagtgaatgcattcagggctatttccaggtctttctt

ttgaagaaagataagactcagtccagagagcacatctgtgacacaccgtg

cctcttgcctttggtgcgtggcagtcatctttggctcatgctgtacatta

ttctac

Klf12 Kruppel-like factor 12 Entrez Gene ID No. 11278
(SEQ ID NO: 49)

gtttcgattctgttttgttcatctgttcgagcagaggggcagttgaagtc

tcgtcctggtctctgccctggcatggactggcacagaggtgttctgtagt

tgaataggaagagcctgtctaaaaaactactgccccacttcaaattgcag

tgttctgtcacctaggcatcatctcttcctgcccctagtatttgattaca

aggaaccaggggaaaaaaactttcttagacacactggcaccaaggtaaga

ggtggggctgcccaggcaaagtcagtgaacatgaaaactcagacaaagca

gagatggaaataatgcgcctcttgaggagaaaagcaataatgaataaaag

gactttcctacaataacttcactgaggactcacgttaccaattttcatac

ttactaaagggattgtaaaaaacaccccagcattttaggtgtcttggttc

catttacagcactgaggtaatctttctgctgtttgttgtcctgcttggtt

gagtacc

Loc253012 hypothetical protein LOC253012 Entrez Gene ID No. 253012
(SEQ ID NO: 50)

ctcattattcctttacatgcagaatagaggcatttatgcaaattgaactg

cagtttttcagcatatacacaatgtcttgtgcaacagaaaaacatgttgg

ggaaatattcctcagtggagagtcgttctcatgctgacggggagaacgaa

-continued

agtgacaggggtttcctcataagttttgtatgaaatatctctacaaacct

caattagttctactctacactttcactatcatcaacactgagactatcct

gtctcacctacaaatgtggaaactttacattgttcgatttttcagcagac

t

Loc253039 hypothetical protein LOC253039 Entrez Gene ID No. 253039
(SEQ ID NO: 51)

gaacttaagttcacacacccttgtactgcaggacggggaatggaacctag

gtcttcttatttttggttcagtgttaactcccattctctaagcagactgg

gcctgttattcaaactgccttcccataggtgcttccctgcttctctcctc

acccagagaaggacttacaaacagcttatcttncagaggttttgtgcctg

atagttatggaatgtgctggtttgagcagggaggatgtaaggggagggaa

tgctaaaaggctgtctacttagagtcaggtttcctgggtaagtccctgga

accccatccccttcccctttcttgagaccccaggacttgctccagtaact

gccaccctgtgcctttgcttcagggccatgctggataaggagctggctgc

ctctgtgaacatcctactcaaggcatcttcactgtgagttttgctgttgc

cattggaggggnngtgggggggagtgtggggagtgctagggtcaggtcctg

gctggtgtaaagaac

Loc91431 prematurely terminated mRNA decay factor-like Entrez Gene ID No. 91431
(SEQ ID NO: 52)

gactttcaccatcctgatattaaaactgtgcaggtgtccacagtagatgc

ttttcagggagctgaaaaggagatcattattctgtcctgtgtaaggacaa

gacaagtaggattcattgattcagaaaaaagaatgaatgttgcattgact

agaggaaagaggcatttgttgattgtgggaaatttagcctgtttgaggaa

aaatcaactttggggacgagtgatccaacactgcgaaggaagggaagatg

gattgcaacatgcaaaccagtatgaaccacagctgaaccatctccttaaa

gattatttt

Lrp16 LRP16 protein Entrez Gene ID No. 28992
(SEQ ID NO: 53)

gtaagactggcaaggccaagatcaccggcggctatcggctcccggccaag

tacgtcatccacacagtggggcccatcgcctacggggagcccagcgccag

ccaggctgccgagctccgcagctgctacctgagcagtctggacctgctgc

tggagcaccggctccgctcggtggcgttcccctgcatctccaccggcgtg

tttggctacccctgtgaggcggccgccgagatcgtgctggccacgctgcg

agagtggctggagcagcacaaggacaaggtggaccggctgatcatctgcg

tgttcctcgagaaggacgaggacatctaccggagccggctcccccactac

ttccccgtggcctgaggctcccgcagcccaccctgaccgggactgacccg

ccttcgggaccccgctcccagctctgagaggtcgccaaagcctgcagcct

ggcctgggcctggccacccctttctttccctccgcgccccgccccgagga

gcctaataaagatctcgttgtggcaaaaa

-continued

Mical3 microtubule associated monoxygenase, calponin and LIM domain containing 3 Entrez Gene ID No. 57553

(SEQ ID NO: 54)

gggccccaagagcagactaggaacgcagggggctgctgctgccaggacnc

cacggagagccgggcacccgcctcacatgtctcctgtctggctccactga

gttagccgtttgagcccactcctatcttttggtggttagtgcatcttcag

ctcttttctgcaagacactggaacattcctaggctgtcccaaaaggagtt

ccaccatagcctttaaggtccgagcagggcaccaaggggttcacttttct

cccgagccattcagcttggggtgcctgcgggaggggcggacagccnagcc

ggcttcccggcggcggtacgagagcccaacaggagaggattagctgtgcc

aaggaacacgccactgctgcctgtctactgcccgccttctctccacttcc

atttttgcctttgtttttaacttgtgctcttgtgagttcttggtgtgttt

ctttgt

Mtap methylthioadenosine phosphorylase Entrez Gene ID No. 4507

(SEQ ID NO: 55)

acaggactatttgccacgacatttcaaaggattccaagagagaatattgg

tgtccatgctgtgatgattcctcagctcctctcatctgatctccgtcctg

gcccccatgactttctttgcggtagttagggtgtggtatgtgccactgag

gcccacacctattggcaatttatagcactgatctgtcatcaataccactt

gctgtcttggatgtgaagatgatttttcctgcagggattccctctacaaa

attaaaaaacactgggcatgtggaaataatattcacgctttaaattgtctt

ttctattcactacaccaggggtccccgacccctaggcaacagactgtggc

cctagtgtagtgaatagaaaagacaatttaaagcatgaatattatttcct

catgcccagtgtttttaattttggtactggtctgtggcttgttagaaa

Mtmr3 myotubularin related protein 3 Entrez Gene ID No. 8897

(SEQ ID NO: 56)

agccgtcagctgtctgctatgagctgcagctctgcccacttacactcaag

gaacttgcaccacaagtggctgcatagccactcaggaaggccatctgcaa

ccagcagccccgaccagccttcccgcagccacctggacgatgatggcatg

tcagtgtacacagacacgatccaacagcgcctgcgtcagattgagtcagg

ccaccagcaggaagtagaaactttgaagaaacaagtccaggagctgaaga

gtcgcctggagagccagtacctgaccagctccctacactttaatggagac

tttggggatgaggtgatgacccgttggcttcctgaccacctggccgccca

ctgctatgcgtgcgacagtgccttctggcttgccagcaggaagcaccact

gcaggaattgtgggaacgtattctgctccagttgttgtaaccagaaggtt

ccagttcccagccagcagctctttgaacccagtcgagtatgcaagtcttg

ctatagcagcctacatcccacaagctccagcattgaccttgaactg

P2ry12 purinergic receptor P2Y, G-protein coupled Entrez Gene ID No. 12 64805

(SEQ ID NO: 57)

aaatgtatatatatcctagtcccctaaccaaatcctgacctattgggata

cttataaaaatttaagtaagtgggatacacaaagaataataactattaac

-continued

ttttcattattagcaaaaacctaagggatttaaactaattgaaactgtat

ttgattggacttaatttttatgtttatttagaagataaagatttaaaga

agacctttacaataaagagaagaaatatcgaagtcattaaaataaggaga

cttacttttatgacattctaatactaaaaaatatagaaatatttccttaa

ttctagagaaactagttttactaatttttacaacttcaataataccatc

actgacacttacctttattaattagcttctagaaaatagctgctaattag

gttaatgaacattttaccttagtgnaaaaaaattaattaaatatgattac

aaagttgcacagcataactactgagaggaaagtgattgatctgtttgtaa

ttacttgt

Rad54b RAD54 homolog B (*S. cerevisiae*) Entrez Gene ID No. 25788

(SEQ ID NO: 58)

gtaagcaaggtctttgtggggcagttgtcgacctcaccaagacatctgaa

catattcagttttcagtagaagaacttaaaaatttgttcacattacatga

aagttcagattgtgttactcatgatctgcttgactgtgagtgtacaggag

aagaagttcatacaggtgattcgttggaaaaattcattgtctctagagat

tgtcagcttggtccacatcaccagaaatctaactccctgaaacctctttc

tatgtcccagctgaagcaatggaaacatttttctggagatcatttaaatc

ttacagatccttttcttgaaagaataacagaaaatgtgtcattcatttttt

cagaatataaccactcaagctactggcacatagtgaaagattacttctga

cattcca

Ralgps2 Ral GEF with PH domain and SH3 binding motif 2 Entrez Gene ID No. 55103

(SEQ ID NO: 59)

acaataattagatctttttccaagttaattgggttttcccttctcccagt

cataggtggtttttatcatcaagacagactgatattttgtcaggatattt

tcttttacagtgtttgatgtgcataatgccagagttatttttttattatt

cattttctctctttttgttcaatatgagattcaggatcatatttgtttaa

aaggtaacacatagagatgtatgtatatattttgttataagacatacaaa

ataattttaagagggataaaggtgaaaatatcagattctggaaattttaa

gtatctaaactttatacttgtatgatttaccataaacataccaaaacatt

tttctgaaaatttactgtcggtctctgacatgaaaccgtattttgtcagt

agttgaccaagcagttttatgagaactcttctatgcaatgatgca

Rttn rotatin Entrez Gene ID No. 25914

(SEQ ID NO: 60)

cttgctgcctgtctggaaagtgagaatcaaaatgctcagaggattggagc

agctgcccttgggctctgatttacaattatcagaaggcaaaaacagcttt

gaaaagcccatcagtaaaaagaagagtggatgaagcatactccttagcaa

agaaaactttcccaaactcagaagcaaaccctctaaatgcctattatttg

aaatgtcttgaaaacctcgtgcagctccttaattcttcctgagtgccatg

ggatgctacaccttgaagctgacagtcatcaacaggggagctaaagttga

agccagctgtgtgtagcagctgttacctgaagacgtgctacctctctaca

-continued

aagtgttgatccccttctttcccatgagagagagaactggtgatactcca

acaccgtccagttgtggcagc

Scoc short coiled-coil protein Entrez Gene ID No. 60592

(SEQ ID NO: 61)

caagagtagatgcagttaaggaagaaaatctgaagctaaaatcagaaaac

caagttcttggacaatatatagaaaatctcatgtcagcttctagtgtttt

tcaaacaactgacacaaaaagcaaaagaaagtaagggattgacacccttc

tgttttatggaattgctgctgatcatttttttctttaaaacttggatagat

tccaaaagttacagtacctttgtggcttcattgaatatttatgaagataa

tgtcagatgtagacaaaaataacacaataacaggagacttccataagttt

gtgtattatgttagtctatgaaaacgtgcaaatgtattgtagagacttta

tg

Specc1 spectrin domain with coiled-coils 1 Entrez Gene ID No. 92521

(SEQ ID NO: 62)

agctgaagactctgaccaagcagatgaaggaggagaccgaggaatggagg

cggttccaggcggatctgcagaccgcagtggtggtggccaatgacatcaa

gtgtgaggcccagcaggagctgcgcaccgtgaagaggaaactgctggagg

aggaggagaagaatgcccggttgcagaaggagctgggggatgtgcagggc

cacggcagggtggtcaccagcagagccgcccctccctccctgggctctgt

cagctagcagagcatttggtggaagaaagacagcccagctcttgccatga

ttgggagccgcagcccatctctagatgaaaggggaatgtgtagaggaga

aattgcctctttataaagagcccagttgtctccttgtgacattctctgtt

ctcagagtcattgccgtcgagtctctgctttttgtccacattttgggatc

agcttactgca

Tpp2 tripeptidyl peptidase II Entrez Gene ID No. 7174

(SEQ ID NO: 63)

gaagagtgcttaaggttgaagtacaatggcacaatctcggctcacctcga

cctctgcttcctgtattcaagtgattcttctgcctcagcctcccgagtag

ctgggattacaggcatgcgccatgacacttggctaatttttttngtgtgtt

tttagtagagatggaattttgccatgttggccaggctggtctcaaactcc

tgacctcaagtgatccacccacctcgacctcccaaagtgctggtattata

ggcatgagccaccatgcctggcctcatttatttttaaatagctgcagtaa

tcccggctttagatnaaancacatgaactaataatatcactagtgttca

Vil2 villin 2 (ezrin) Entrez Gene ID No. 7430

(SEQ ID NO: 64)

agagctgagtcatcctagagcaaacctctggagtggagagcgaactactt

cattcccctcccttagcctgggccagagagactccagctctgccttctcc

agccaaaaaatcaaaggcagatgggagaacagccttcagctttggataac

gatgaaatatctggcaccactgatgaatattaaactttctataacc

-continued

Znf204 zinc finger protein 204 Entrez Gene ID No. 7754

(SEQ ID NO: 65)

gagcacatatcttacaaaacaccaaaaaattcatagtgaagagaaatcaa

atatacatactgagtgtggggaanccattagacaaaactcttctttttna

caacaataaaancctcacactggagagnttctctgaatgccttaagaatt

tggttaatatggagacccttcccagggaaacagaaggaggatcgtgaaaa

ctgttgactacttagaatgatcacatggtttagtggagagagcatgattc

tgggttttaaaagtcatggatctcaatctcagctcctattactaactaga

tcttttactttggggtaagtcacttcatatctttaggccttaatttcctc

atctgaaaaactggaaggcctgacttgttgagcttta

Znf24 zinc finger protein 24 Entrez Gene ID No. 7572

(SEQ ID NO: 66)

ggcactgtgtaatcattccttgaagtagttggagatggtgctggtatgcc

actgaatgaggtctgagcaggttttcttcacatctgaggggacagtgcca

gccagtcaacttttggggtggggctgaagtctgctgaaaatctgcagttt

tacatgtttcatgggacattcttctgtgcaataaagtttgaga

Amn amnionless Entrez Gene ID No. 81693

(SEQ ID NO: 67)

tactggtgacacttcatggctgcgacccagaatgaacttaatgcacacag

ggacgcagggtgtcactggtcctgggcctttgtccatgactaggtggtca

gcaggacttctgcagctgactgtgcaatggctaaatgaaaagaaggccac

agactaacctccactttcctgtcttcaaaattctagtgacactgggaatg

ctataggacctcctactattctcttaaggtcctaggaaagtttcaggaac

tagggaaaagactgggtactgaggctgtgtccccagatgtctgcttccga

agcagccgcgtcatgacgggtttctgctgaggaagtggtgttggcagggc

cccatatgccctctcgggttgtcaggggtgggagacaggctgtatggggg

tccttcatgtgcagatggaacagcatcgcctcacagctgtgcagacgaac

agatgtggtctactgccacgaacaatgcgg

Ankrd13 bankyrin repeat domain 13B Entrez Gene ID No. 124930

(SEQ ID NO: 68)

cctcatggtgcctcggagagtggggagcatattgggctgnggtaagcact

agacccaagtagactggacacaaagggctcgcccagggccntggcgccac

ccccacccctcccaccagctgctgctagcctctgtggttgtacatccca

cttgcccccacacggagactgactctaaaacccttcatccaatggtgnta

acccccggctntcccctgccccacctcacccacccagagaagcacagacc

ccgccaggggcaggggcccaccgcacacccttgtcccgggcctgtctggg

actggccttcccggntcagcnagnnnnnnncagaagggacacaaagaggg

atggaagaaaagaacaaagagaaactgttcctcccacccccttccctgat

gccaggggcaccagactgattctgagg

-continued

Ankrd57 ankyrin repeat domain 57 Entrez Gene ID No. 65124

(SEQ ID NO: 69)

gtcatttaccatggttttcccactgaaggctttaacttttctgataaaat

aatattttaaattttcaaaaacccattcctgaggagaactacttctagca

ttccttttcatgatgtgcttttgtgcagtaagtagcattttcggctactt

aactttacattcctcttatttttcagtttccagtcaagattataaaaagc

aaatgattgatataatttgatattcatagagttgtgcctacctttaatgg

aaaaatacatgtcagatacttagatgtttattgatatgagactatgtggt

taaaaaacccaagtatgtccatgtgtttcttataaggtacacttgaaact

agtgagtgtttgtcacatttcactttcatggtatataaaatgcagtttgc

atatataacttgaatatctggtactagttttttcacgcctgcaatcttgg

agtctaggttgccttgtctct

Apobec4 apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 4 (putative) Entrez Gene ID No. 403314

(SEQ ID NO: 70)

aaatccaagttcctcttactggctttcaaggatcctcctttaacttcctg

ttgccttgtctcatcagcaagatcataagtacctacaggtcaagcactgt

ctctcccttctctttagcttttcccttaggatctagcacattacccagca

aaatgtgagtagcaaggctgaaatgacatctcaataacttcaccaatgat

tgtaactcagcatcccttctccatcccagctgaaagcctgcttcaccatc

ctgcaagagatgttttttctttttgttagcatccattcccctttctaatg

cagctcccaatgcataatgtgtg

Btnl9 butyrophilin-like 9 Entrez Gene ID No. 153579

(SEQ ID NO: 71)

ggtcatcgaatctgcatgcatccctcatacatctggagacttcgtgaagg

ttccagagttactgactgagatttctgagctttttttccccttttctgttt

ggtttcagagtggagagcagcccaaaaatatgcaggtaactgaagccagg

aaactgatttgtgttttggtttggnccggatncttnaaacagaagggagg

tggagagatctgagattagaggacggggctttataggagnccaagtatgg

ggcctgcacacacaagacacacacgcacacttgcaaacacgccacacgac

acatatgcctgcatgtgtatgcacacacatgcacacgtgagctcccaaac

acatcgctccttggggttacactaggtttgtttccatctggcttgaggct

atttgcaggcgagagtgcagagtctgtaatgaacctcccagattctctga

cgaaggggtcccc

C11orf71 chromosome 11 open reading frame 71 Entrez Gene ID No. 54494

(SEQ ID NO: 72)

ttcctgtgtgcttagtcatcgggtccgacacaggctggactgatctgggg

agccgcgaagggcctgccttcacaaagggacgtaacgcaagtactgcggg

cagtgtttgaatatggccctgaacaatgtgtccctgtcctccggtgatca

gaggagcagggtggcctaccgctcttcccatggcgacctcagaccgcggg

cgtcagcgttggcgatggtctccggagacggcttcctcgtttccaggcct

-continued

gaagcgattcatctaggacctcggcaggcggtgcgaccaagcgttcgggc

cgagagccgtcgagtggatggtggcggccggagcccaagggaaccagatg

gccggggccggagccgccaagccagattctcaccttacccaatccctgcc

gttgaacccgatctcctaagaagtgtgctgcaacagcgtttgattgcatt

aggaggtgttatcgcagctcgaatttcagtttaaacgaacacctttcctc

tggccctcacttagcttgtgaacag

C14orf145 chromosome 14 open reading frame 145 Entrez Gene ID No. 145508

(SEQ ID NO: 73)

gaaatgcatgggacaggtgactgacatgactgcatcgtggtttagatgta

tagataacacggggaggtgctttacattttaagactttgttcataattct

tttatttatggtttctctgaatcattcttttggaacattctaaaagagcc

agagga

Clorf89 chromosome 1 open reading frame Entrez Gene ID No. 8979363

(SEQ ID NO: 74)

cgggtgaagagtgtgccggggcggcggctggctgatgggcgcacactgga

cgggcgggctgggctggccgacgttgcccacatactcaatggccttgctg

agcagctgtggcaccaggaccaggtggcggctggcctgcttcccaacccc

ccagagagtgctcctgaatgagtcacgagtggttgcctgtgatcccaccc

ccaaccctcaggtctcgacatagggctggaggctggggcaggaacatgga

tcctatctggaggactggccagcatggcctgatcagggaggatgtggcca

gagaaggcccacccgcgagcagcgctttccttgcagaattcatggcaggg

aggtggggaccaaggccctgagctcgaacatctcccgtggcctttccccc

tttggcagcaccgatggaggatgactgggagaggggtgcctctcaagtt

acttcaatcaagaacctgtattggttgaggtgacaccatctgttgtaaca

g

C20orf7 chromosome 20 open reading frame Entrez Gene ID No. 779133

(SEQ ID NO: 75)

tgaatgaccttcctagagcacttgagcagattcattatattttaaaacca

gatggagtgtttatcggtgcaatgtttggaggcgacacactctatgaact

tcggtgttccttacagttagcggaaacggaaagggaaggaggattttctc

cacacatttctcctttcactgctgtcaatgacctgggacatctgcttggg

agagctggctttaatactctgactgtggacactgatgaaattcaagttaa

ctatcctggaatgtttgaattgatggaagatttacaaggtatgggtgaga

gtaactgtgcttggaatagaaaagccctgctgcatcgagacacaatgctg

gcagctgcggcagtgtacagagaaatgtacagaaatgaagatggttcagt

acctgctacataccagatctattacatgataggatggaaatatcatgagt

cacaggcaagaccagctgaaagaggttccgcaactgtgtcattt

Ccdc88a coiled-coil domain containing 88A Entrez Gene ID No. 55704

(SEQ ID NO: 76)

acatatgtacagtatcagtagggaaaatgtaaaaagatgttgttttcttt

tgtcatttaattaggccatctgtcctgttttaaagaaatagttaataatt

-continued

caacactttatataacaaatattaactaatacccatatttataaaacatt

tttcagatttaaaagattgttaatacttataaacttagtgttattcttag

aaaaccccatcaaatttaaatgtgatttacacagtgactaggaacatttg

tatttattgtttcttctctgcacttttcatcatctgataaatacaagagc

tcaagtaactgtctttttcttcaagatggcttctatacttgaaatcagtta

atacaatagtttttccagt

Ccne2 cyclin E2 Entrez Gene ID No. 9134

(SEQ ID NO: 77)

gaggaagtcactttactactctaagatatccctaaggaatttttttttttt

aatttagtgtgactaaggctttatttatgtttgtgaaactgttaaggtcc

tttctaaattcctccattgtgagataaggacagtgtcaaagtgataaagc

ttaacacttgacctaaacttctatttttcttaaggaagaagagtattaaat

atatactgactcctagaaatctatttattaaaaaaagacatgaaaacttg

ctgtacataggctagctatttctaaatattttaaattagcttttctaaaa

aaaaaatccagcctcataaagtagattagaaaactagattgctagtttat

tttgttatcagatatgtgaatctcttctccctttgaagaaactatacatt

tattgttacggtatgaagtcttctgtatagtttgtttttaaactaatatt

tgtttcagtattttgtctgaaaagaaaacaccactaattgtgtacatatg

tattatataaacttaacctttttaatactgtttatttttagcccattgtt

Ceacam6 carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen) Entrez Gene ID No. 4680

(SEQ ID NO: 78)

gttttaattcaacccagccatgcaatgccaaataatagaattgctcccta

ccagctgaacagggaggagtctgtgcagtttctgacacttgttgttgaac

atggctaaatacaatgggtatcgctgagactaagttgtagaaattaacaa

atgtgctgcttggttaaaatggctacactcatctgactcattctttattc

tattttagttggtttgtatcttgcctaaggtgcgtagtccaactcttggt

attaccctcctaatagtcatactagtagtcatactccctggtgtagtgta

ttctctaaaagctttaaatgtctgcatgcagccagccatcaaatagtgaa

tggtctctctttggctggaattacaaaactcagagaaatgtgtcatcagg

agaacatcataacccatgaaggataaaagccccaaatggtggtaactgat

aatagcactaatgctttaagatttggtcacactctcacctaggtgagcgc

attgagccagtggtg

Chrnb3 cholinergic receptor, nicotinic, beta 3 Entrez Gene ID No. 1142

(SEQ ID NO: 79)

tgatttttgtgaccctgtccatcattgttaccgtgtttgtcattaacgtt

caccacagatcttcttccacgtaccaccccatggccccctgggttaagag

gctctttctgcagaaacttccaaaattactttgcatgaaagatcatgtgg

atcgctactcatccccagagaaagaggagagtcaaccagtagtgaaaggc

aaagtcctcgaaaaaaagaaacagaaacagcttagtgatggagaaaaagt

tctagttgctttttttggaaaaagctgctgattccattagatacatttcga

-continued

gacatgtgaagaaagaacattttatcagccaggtagtacaagactggaaa

tttgtagctcaagttcttgaccgaatcttcctgtggctctttctgatagt

gtcagtaacaggctcggttctgatttttaccccctgctttgaagatgtggc

tacatagttaccattaggaatttaaaagacataagactaaattacacctt

agacctgacatctggcta

Cllu1 chronic lymphocytic leukemia up-regulated 1 Entrez Gene ID No. 574028

(SEQ ID NO: 80)

cagacacatatacaggaaagaccatgcgaagaagacacagagaggaaacg

gccatctacaatccaaggagagaggcctcagaacaagccaaccctgcaga

taccttgatctaggcatccagaattgtatgaaaatac

Cnnm1 cyclin M1 Entrez Gene ID No. 26507

(SEQ ID NO: 81)

agacaggagggccacagcgtgtggaagtaaagactttggagctagagatg

ccttttccagcaatgattattgacttcaccacacccctttgcctggcctgg

cctgaggctcagcagtgcatgacttctcgtagataacttcacagtcatcc

agtcccaacacctgctcttgcctggtaggaacaggcgaagtgtcagccct

caatgttgggtacttagacccaaaccaataaatggtgagttttgaacaag

aactaccatcatgcaggcttcttgcccagctgaccactggccccggggtg

cctgcctggctggtcttcatcacctgaggccaccaggctcaagccactgc

tgttgcattacacccatccctttgcaaaatccctatggagcctgtcacca

ctcccctccctatataccccccaccccacaaagattttcttcag

Cnot2 CCR4-NOT transcription complex, subunit 2 Entrez Gene ID No. 4848

(SEQ ID NO: 82)

gtgacgttggtaggaaagcatttctttacatggaggtttattttgtggga

cattacctcctctggatgttacttcctcagtttacaagtagtgtaaatnc

tcattgattctnttatgaattgtaanggatttnctcttagcttttgagaa

tttagaatctganatttgaataaaaagtaaatatattcagtataattntn

caaaatgctctagtttcaagatanttaaaanattatgtggaatttacata

attaattcnaaatatagtttgtatttagttccnttatcaaataatgcaaa

tagttggnagatacctcaatttcttttgagtgttaagnaagnagncaagn

aaaggnagtgaagtttttgcaacacattgtgtctttatttggtctgccta

tgtttttatcacattgcttataaaacttttaaaatccttgtttgtataaa

aagtttctttagttaaataaaagtgtgtgtattaattagtgtgccttctg

gacaaattaagaaatatttttttctatatttcaatgcggttgtatt

Dhfr dihydrofolate reductase Entrez Gene ID No. 1719

(SEQ ID NO: 83)

tcagatatttccagagaatgaccacaacctcttcagtagaaggtaaacag

aatctggtgattatgggtaagaagacctggttctccattcctgagaagaa

tcgacctttaaagggtagaattaatttagttctcagcagagaactcaagg

aacctccacaaggagctcattttctttccagaagtctagatgatgccttа

aaacttactgaacaaccagaattagcaaataaagtagacatggtctggat

-continued

agttggtggcagttctgtttataaggaagccatgaatcacccaggccatc

ttaaactatttgtgacaaggatcatgcaagactttgaaagtgacacgttt

tttccagaaattgatttggagaaatataaacttctgccagaatacccagg

tgttctctctgatgtccaggagga

Eif4a2 eukaryotic translation initiation factor 4A, isoform 2 Entrez Gene ID No. 1974

(SEQ ID NO: 84)

ataccacttagtatagttcgctactattttgtggcctacatgacaggtgt

caagnttttttgaatcaatttttaaaacatgccattgtgtttcaggctc

gcgggattgatgtgcaacaagtgtctttggttataaattatnatcnnnct

ancagtcgttgaaaactatanttcacaggtgnagaagccagcatcttggc

tgtattgaaaaaacttcatacgtttttctactgtgatttgtatgaaaggt

aacatcaaatcaaggaatagattcagtaaagtcagtagtgttcagtaaga

tgatgtaattaaatttgtactagggaaggttgatgagaacaaagtgggaa

aacttgtaaacattgcccagattgtggacatagggttnttttccacnaat

tgttggtcttaccttatgcttgagcttttagtgatgttcttgtgtccatg

tgtttttcttggtgatttttttnctatangttgggattttcnttggtgtcg

nctggtagnnnnnnnnantgaaccctggtttagttatagtggctttatccc

taaata

Fa2h fatty acid 2-hydroxylase Entrez Gene ID No. 79152

(SEQ ID NO: 85)

ctactcgggcgctcccagaaggagccacctctcagtgcctcacctccccc

tgcctcccagcctccgcagatgaggttcctgccccttcctcctcgtaacc

aaaaccctcactgctcccaggacggtcttatttataaaccagatacatgt

tcttagtctggtcccagaccaaggagctggtcagacggccctttctaatc

ctacatgttgagcttatgtaaaaaatgttgtttcctcctgtttttggttc

ctttcttacccacaaaccattactacttgaaacttaaaaaactcgccaag

tgtaaaggctaaagagaagcagtttgacggaccttgtgatttgtactgtt

tgctgcggagctattta

Fbxo15 F-box protein 15 Entrez Gene ID No. 201456

(SEQ ID NO: 86)

gagaacacctacctcttattggaaaagttggcctctcgtggaaaactgat

atttttgatggctgtataaagagttgttccatgatggacgtaactctttt

ggatgaacatgggaaacccttttggtgtttcagttccccggtgtgcctga

gatcgcctgccacaccctctgacagctctagcttcttgggacagacatac

aacgtggactacgttgatgcggaaggaagagtgcacgtggagctggtgtg

gatcagagagaccgaagaatacctttattgtcaacctggtcctttatctta

gtatcgcaaaaatcaaccattggtttgggactgaatattagcagtaggtg

gcaaattattgttgttatttagttgtttatttttgactggctttgttctt

g

-continued

Gins4 GINS complex subunit 4 (Sld5 homolog) Entrez Gene ID No. 84296

(SEQ ID NO: 87)

gccaggctttgtggtatgtacctttagtcccagctactctggaggctgag

gcaggaggatcacttaagccttggaggtcaagaatgcagtgagccattat

catgccactgtgtgaccagaaaccagatgtagccatttcaagcataaaac

atgatatttttgttttccttggactgaaacatagtctgggtcctcaacgt

tgccggtgatgatggttgaacatcatgttttttataaaccttaatttctc

atttaataggaagaaaatctcaggagagccaaaagggaggacctgaaggt

cagcatccaccaaatggagatggagaggatccgctacgtcctcagcagct

acttgcggtgtcgcctcatgaaggtttgacgtggagatacctcaaagtct

ccgacct

Grhl1 grainyhead-like 1 (*Drosophila*) Entrez Gene ID No. 29841

(SEQ ID NO: 88)

aaaacactactgcaatcacgtctttngttatgctagtatcagtcagatgc

acttagagtgaagaaacactgtaattacagcacacagattgcaagtattg

cgtaccaagtgatacaactcgaaatgcagttctcatcttcctgttttgag

aaatgattattttatcacgcatcagagccttcgtgctttgattatcttgt

atgttaacaattctagaaaacattcatgaattcacnaaaatangttacta

tggcaggggaacattttgtacacatttaagtatataaaaatactaaaata

tgtaattttataacaaagtcacgggtatctttaggttcagggaactagac

taggtcattcgtgtaaatggactggtagttacagtcttaggttaaagtat

tctaatgaagtatgggaactaaattgctggttttctaag

Gtpbp8 GTP-binding protein 8 (putative) Entrez Gene ID No. 29083

(SEQ ID NO: 89)

ttctcttgccctttacaaattagttttctcacttaaagctatttttgttt

tttgctttttcactagtgaaaatgttacttccccccatgannnnannntnnn

ntntnntanatctacctgtgaattttgctatatttctttgttggtttggc

tttacaaatatgtagctgtcttctcacatgttacctgctgtaaaaccatt

catttgaatcttaatagctttcacgtttacagtaacaaatgaatttccga

gaaatcaagtaagttgcccaagatccaacatatataataaacatcagaac

tagaacttgaattctgttctttcggttgtttccaacatggactaacacat

tttatcaagaatgttttcaatattcaaataaggactcgaaaaaatagget

tacatagtaacttttatccatcaacttacctatcgatgct

Heatr6 HEAT repeat containing Entrez Gene ID No. 663897

(SEQ ID NO: 90)

agggagatgacactggagcaccccacagcccacaggaaagagaccagatg

gtcagaatggcccttaaacacatgggcagcatccaggcaccaactggaga

cacagccagaagggccatcatgggctttttagaagagatcctggccgttt

gttttgactcatctggatcacaaggggcactcccagggttaacaaatcag

tgaagatcccaccatactttctagatgtcgaaggcggcagtaggaagacc

tgagcttgagcataagatctgtgggatttcatcttaggggcagaaacaat

-continued

ccgttcactatttatttagaatgacttagcagccatttaaatttttcacag

agggctcaaccacctttggagtgactccatagcactggccatggtcaggg

ttgttggaacatctgacctgtgcatccaggagccgaggagtcaggttgta

atacaggccaagcagacgggctttgagggcattta

Herpud2 HERPUD family member 2 Entrez Gene ID No. 64224

(SEQ ID NO: 91)

aaactaaacatcatatgttctcatatgtccctaagctatgaggatgcaaa

ggcataagaatgatacaatggactttggggactttcagggaaagggtgag

aagggcgtaagggataaaagactacaaattgggttcagtatatactgctc

gggtgatgggtgcnccaaaatcttaaaaatcgccaaagaacttatgtaac

taaataccnnctgttccccaaaaaactatggaaattaaaaattaaaaaat

aagtataatttctgctttagcgatattaactattcagtacncaataagtg

agtttagcaattcagtgatt

Ica1 intestinal cell kinase Entrez Gene ID No. 3382

(SEQ ID NO: 92)

catactgcatgctcaggacccatagatgaactattagacatgaaatctga

ggaaggtgcttgcctgggaccagtggcagggaccccggaacctgaaggtg

ctgacaaagatgacctgctgctgttgagtgagatcttcaatgcttcctcc

ttggaagagggcgagttcagcaaagagtgggccgctgtgtttggagacgg

ccaagtgaaggagccagtgcccactatggccctgggagagccagacccca

aggcccagacaggctcaggtttccttccttcgcagcttttagaccaaaat

atgaaagacttacaggcctcgctacaagaacctgctaaggctgcctcaga

cctgactgcctggttcagcctcttcgctgacctcgacccactctcaaatc

ctgatgctgttgggaaaaccgataaagaacacgaattgctcaatgcatga

atctgtacccttcggagggcactcacat

Igl immunoglobulin lambda chain, variable 1 Entrez Gene ID No. 3535

(SEQ ID NO: 93)

tctggatccaaagacgcttcggccaatgcagggattttactcatctctgg

cctccagtctgaggatgaggctgactattactgtatgatttggcgcggca

ccgctgtggtatttggcggagggac

Il15 interleukin 17 receptor A Entrez Gene ID No. 3600

(SEQ ID NO: 94)

gaagatcttattcaatctatgcatattgatgctactttatatacggaaag

tgatgttcaccccagttgcaaagtaacagcaatgaagtgctttctcttgg

agttacaagttatttcacttgagtccggagatgcaagtattcatgataca

gtagaaaatctgatcatcctagcaaacaacagtttgtcttctaatgggaa

tgtaacagaatctggatgcaaagaatgtgaggaactggaggaaaaaaata

ttaaagaatttttgcagagttttgtacata

-continued

Il17rc interleukin 17 receptor C Entrez Gene ID No. 84818

(SEQ ID NO: 95)

gctgcccgcagaagcgcacatgtgcnnnnnggctncnggtnggacccct

nncannncantgnnccnnncgctttcctgggagaangtcactgtggacnag

gttctcgagttnccattgntgaaaggncnccnnnnnnnnntntnnnnca

ggtgaannnnnnnnnnnnnnngcagctgcaggagtgcttgtgggctgact

ccctggggcctctcaaagacgatgtgctactgttggagacacgaggcccc

caggacaacagatccctctgtgccttggaacccagtggctgtacttcact

acccagcaaagcctccacgagggcagctcgccttggagagtacttactac

aagacctgcagtcaggccagtgtctgcagctatgggacgatgacttggga

gcgctatgggcctgccccatggacaaatacatccacaagcgctgggccct

cgtgtggctggcctgcctactctttnnngctgcgctttccctcatcctcc

ttctcaaaaaggatcacgcgaaagggtggctgaggctcttgaaacaggac

Insc insulin induced gene 1 Entrez Gene ID No. 387755

(SEQ ID NO: 96)

tccttcttactctgcagcaacatggaggagagttttgtgtagtgagtgtg

ggngaagaaatacatttggctgttctcacacccccctctgactatgcacca

gtgaacacatctgagtacataccagctctcctcatcttcttatttatact

taacttatttttgtgtgaaataaatggaggacgaaatcttagagcaacat

catcaaacagtctttggtccttgagaatcttctttgtgttttatttttttg

atttctgtagcttttcagttgcagatgttgaaattcgtaatgacaaatat

gacaaattgtcatgggtgattccacttcatcttatttttttctactctcac

tatacaatcttgcctcatttttttaaaactttggaaccagaggatttcaac

tgcctagca

Ipo11 Intracisternal A particle-promoted polypeptide Entrez Gene ID No. 51194

(SEQ ID NO: 97)

gagactacagcagtgttacctgtgcaaatacaacttactacttctgttac

cttgaacttggaaaaaaacagtgctctaccgaatgatgctgcttcaatgt

cagggaaaacatctctaatttgtacacaagaagttgagaagttgaatgag

gcntttgacattttgctagctttttttcatcttagcttgtgttttaatcan

ttttttgatctacaaagttgttcagtttaaacaaaaactaaaggcatcag

aaaactcaagggaaaatagacttgaatactacagcttttatcagtcagca

aggtataatgtaactgcctcaatttgtaacacttccccaaattctctaga

aagtcctggcttggagcagattcgacttcataaacaaantgttcctgaaa

atgaggcacaggtcattctttttgaacattctgctttataactcaactaa

atattgtctataagaaacttcagtgccatggacatgatt

Itch Integrin alpha FG-GAP repeat containing 1 Entrez Gene ID No. 83737

(SEQ ID NO: 98)

gatcattggtatgtcaatctcttgatgaaaaatcagtacctgaatatgtc

tttttgttttttttaagagacagggtcttgctatgttgcccaggcaggatt

tgaactcctgggatcctcccacctcagcctcccgagtacaatacctgaat

-continued

tttaaatagagttattgtaagtcttatgaaatgagattttgctgcactct

gacataagataataaaagacagagcaggaattcattattatgagctgctt

gatcagttttaaaccactccatttgatgaaacaagtgaggtccttccctc

ctgaccaggctgtggaatgctgtcttccccaaccccaccccctgcaaaa

gagcagaacaataaggcaattgctcatttt

Itfg1 integrin alpha 2b Entrez Gene ID No. 81533

(SEQ ID NO: 99)

gcagggaaaacacagtccactccccaccaccactccgctccttccacagc

aaatggggactgctcagaaaaccctgtcctttcttttcctctcttcaaag

gcaggcatgtgattgaggatgtgatggtgacttctggcctgttttatttt

gtggtaacttcactttagtcagggaaataattggaatatctttaatctga

tgtagtttgacctttagaaattgaaagtgaaacagctattgttgataatt

cacaaagtattaataaaacttctattacctgtaaaaannnnnnacttaat

ctgcgagaaaactatttagaatattatggaattgtgccatagcttcttta

tttgttcttaattctatattagatttttttttctctgcttcatgaacaag

ttcagattttaaaacattatgcctgaagacatgtccaactttatttttta

tatgttatattctggtccaatatactagagtaataatactctggtattta

ggatatctgtcattgaacctcagttacaaattaaacaatagtagctacca

tatctaagtgat

Kcnmb4 potassium large conductance calcium-activated channel, subfamily M, beta member 4 Entrez Gene ID No. 27345

(SEQ ID NO: 100)

agatgagattggttcccagccatttacttgctattttaatcaacatcaaa

gaccagatgatgtgcttctgcatcgcactcatgatgagattgtcctcctg

cattgcttcctctggcccctggtgacatttgtggtgggcgttctcattgt

ggtcctgaccatctgtgccaagagcttggcgatcaaggcggaagccatga

agaagcgcaagttctcttaaaggggaaggaggcttgtagaaagcaaagta

cagaagctgtactcatcggcacgcgtccacctgcggaacctgtgtttcct

ggcgcaggagatggacagggccacgacagggctctgagaggctcatccct

cagtggcaacagaaacaggcacaactggaagacttggaacctcaaagctt

gtattccatctgctgtagcaatggc

Kif14 kinesin family member 14 Entrez Gene ID No. 9928

(SEQ ID NO: 101)

ggatgtttatggtgaaatggcctgtacaagtttaactaagacaacttaac

ttgcattgttaatcaaaaattcttttctcaaagggttaactggttgccat

tttgaatagtatgttcaagggtgtagcttcctgtttctttccaaattata

agtagctacctaaatatagtataattatatattaataatatggcttgctg

gcacagtagtttaccctgttatctgtgtttcataatgggggctgtatgaa

tattatttaaaactaataaaatgttgccagaattatactaaactgttgga

tgagattaggagatcagaggctggaccttctcttgataatgcttgttttg

ttaaaggtataatgaaataatttgtatatgatttgatgaagattaaagac

-continued

ccttattttccacagctttaaaaaaaaaccttttatttatgatcaagtaat

aaagataatattctacttgtgggatcttacattatggaaatagtttgacg

tttttgacctcaagagtatgtataatttgaagagatactttgtaactatg

cttgggtg

Loc388692 hypothetical gene supported by AK Entrez Gene ID No. 123662 388692

(SEQ ID NO: 102)

ctcaacatcataaggaatcagacggatgcggaaaccnagncgggntggat

agnaaantctttccaggaaggctccggggcactcaactggtctccaaccn

tcccntgcaacntgtgacgcctgccatnttnccctatnttaggcgantgg

caacgcaanccctccgtttgctctgggcaaaacttcgagagttccctctg

aagctggagctttttcctcagatccaagatccaattggtcaccaattcgt

gatttc

Loc401913 hypothetical LOC401913 Entrez Gene ID No. 401913

(SEQ ID NO: 103)

caagcccattttcctgagggtgaaggggacttttattatataggggcctt

atttggtgggtcagtgctggaggtttacaggctgatcatggcctgtcacc

aggtgatgatgattgaccaagccaaccacatcgaggccctgtggcatgac

gaaagcctcttaaacaagtacctgcttaaccacaaacccactctcccttg

agtacatgtgggattaaaaagtcgatggagtatacgttggatgaatacct

ggtgggtttgtgcaccatgataaactgcaagagatctgtggtcttggtaa

ataacaatgnggaaatgtgaactgatgagggaagcttccagaaagagacc

agagaggggggtgattgccagtcagcccgnatcttcctcctgaaatgctac

cctgattta

Loc619208 Entrez Gene ID No. 619208

(SEQ ID NO: 104)

ccttgtcaacatcttcgagcatcggcagctccggangccggggtaactgg

cagcaggtaggaaactatgtgaaagaatctcctgatgtcataatttccgg

gtgtcaccggaacatttgatcatcattcctttggcaattccagccttctg

tggaaaggccagtagaaagcattgatttattcacctctacaggaatcaga

ctcagcctcttttggttttcagtgaagtatgccttttcaatttggaaccc

agccaaggaggtttccagtggaaggaggagattcttcaattgagctggaa

cctgggctgagctccagtgctgcctgtaatgggaaggagatgtcaccaac

caggcaactccggaggtgccctggaagtcattgcctgacaataactgatg

ttcccgtcactgtttatgcaacaacgagaaagccacctgcacaa

Loc645513 Similar to septin 7 Entrez Gene ID No. 645513

(SEQ ID NO: 105)

tgaacagagtaacaggactatatttctgaaaaaggatgaacttactggaa

acaggattgcgtgcctgaaatatacccaaatggatataaatgtcaactca

gttctgggctggagatatagatttggaatgcaccaacaatgcagatggta

atcctggcatgcgagtag

-continued

Loc654342 Similar to lymphocyte-specific protein 1 Entrez Gene ID No. 654342

(SEQ ID NO: 106)

ttctagaatctcagcttcttaaatcaagaaattgtgtcttgttcatgtct

gaattccccaagtgaacacagtgagtgggtgcttgacaaatctttgntgg

tnangngcnaaaaaaggggattctgtgcccaataccatgaaatcaatgca

cagaagattcantcaatcaagaaaggtgcacagacactggccacacacac

tgacatttgtttgcagatgttccagtccccctgacttccaccaatccatt

cattcattccacaagcatttttgctggggtggaagcagggccatacaggg

tgtgtacatcacagatagggtgggtttgtataatgagtataaaaaaacttc

tagcagaagatgacaaagtatatcaagaaagggctctcttcgaaatcaca

Lrrc37a leucine rich repeat containing 37A Entrez Gene ID No. 9884

(SEQ ID NO: 107)

ggctttgggagtgagcagctagacaccaatgacgagagtgatgttatcag

tgcactaagttacatttgccatatttctcagcagtaaacctagatgtgga

atcaatgttactaccgttcattaaactgccaaccacaggaaacagcctgg

caaagattcaaactgtaggccaaaaccggcaaaaagtgaatagagtcctc

atgggcccaatgagcatccagaaaaggcacttcaaagaggtgggaaggca

gagcatcaggagggaacagggtgcccaggcatctgtggagaacgctgccg

aagaaaaaaggctcgggagtccagccccaagggagctggaacagcctcac

acacagcaggggcctgagaagttagcgggaaacgccatctacaccaagcc

ttcgttcagccaagagcataaggcagcagtctctgtgctgacacccttct

ccaagggcgcgccttctacctccagccctgcaaaagccctaccacaggtg

agagacag

Mrp130 mitochondrial ribosomal protein L30 Entrez Gene ID No. 51263

(SEQ ID NO: 108)

ttatggctgggattttgcgcttagtagttcaatggcccccaggcagacta

cagaccgtgacaaaaggtgtggagtctcttatttgtacagattggattcg

tcacaattcaccagatcaagaattccagaaaaagtgtttcaggcctcacc

tgaagatcatgaaaaatacggtggggatccacagaaccctcataaactgc

atattgttaccagaataaaaagtacaagaagacgtccatattgggaaaaa

gatataataaagatgcttggattagaaaaagcacataccctcaagttca

caagaatatcccttcagtg

Nipsnap3 bnipsnap homolog 3B (*C. elegans*) Entrez Gene ID No. 55335

(SEQ ID NO: 109)

gttaatttgctgtgcttcttgcatttttgaaagttacatattctccactg

ctttaagaaataattcagttcactttcaccttggcatttcagtatctgtt

acacattagaagtagttgtcactatttcatc

Parp2 poly (ADP-ribose) polymerase family, member 2 Entrez Gene ID No. 10038

(SEQ ID NO: 110)

gccatgggcttcgaattgcccaccctgaagctcccatcacaggttacatg

tttgggaaaggaatctactttgctgacatgtcttccaagagtgccaatta

-continued

ctgctttgcctctcgcctaaagaatacaggactgctgctcttatcagagg

tagctctaggtcagtgtaatgaactactagaggccaatcctaaggccgaa

ggattgcttcaaggtaaacatagcaccaaggggctgggcaagatggctcc

cagttctgcccacttcgtcaccctgaatgggagtacagtgccattaggac

cagcaagtgacacaggaattctgaatccagatggttataccctcaactac

aatgaatatattgtatataaccccaaccaggtccgtatgcggtacctt

Pbrm1 polybromo 1 Entrez Gene ID No. 55193

(SEQ ID NO: 111)

gtcagcagtcagcaaattaacatcatcatactcttccatttttagtttct

gttggattttcatcaagtcaatgggctgagaaaccacttcataatagtct

ggttgatttcttcgctttggtgccctaatgaagagctcacagagaagtct

gccctgttcatccttatagtctcggatggtattatagagttcatggcaca

cggcaataggatctacagttggaagattggaaagtctcctccttttcctg

cttgggcctggtgttgacacagaatggtgcccatcatcaaagtccccgct

gacactgctggaaggggaggtagctcttcttctct

Pex13 peroxisome biogenesis factor 13 Entrez Gene ID No. 5194

(SEQ ID NO: 112)

ggatgaccatgtagttgccagagcagaatatgattttgctgccgtatctg

aagaagaaatttctttccgggctggtgatatgctgaacttagctctcaaa

gaacaacaacccaaagtgcgtggttggcttctggctagccttgatggcca

aacaacaggacttatacctgcgaattatgtcaaaattcttggcaaaagaa

aaggtaggaaaacggtggaatcaagtaaagtttccaagcagcaacaatct

tttaccaacccaacactaactaaaggagccacggttgctgattctttgga

tgaacaggaagctgcctttgaatctgtttttgttgaaactaataaggttc

cagttgcacctgattccattgggaaagatggagaaaagcaagatctttga

tatctttcatgtttgcctgc

Phlda1 pleckstrin homology-like domain, family A, member Entrez Gene ID No. 122822

(SEQ ID NO: 113)

gaagtgggacgagcacatttctattgtcttcacttggatcaaaagcaaaa

cagtctctccgccccgcaccagatcaagtagtttggacatcaccctactg

aaaacttgcgattcttcttagttttctgcatacttttcatcacgatgcag

gaaacgatttcgagtcaagaagacttttatttatgaacctttgaaaggat

cgtcttgtatggtgaattttctaggagcgatgatgtactgtaattttatt

ttaatgtattttgatttatgattatttattagtttttttttaaatgcttgt

tctaagacatttctgaatgtagaccattttccaaaaaggaaactttattt

tcaaaaacctaatccgtagtaattcctaatcttggagaataaaaaagggc

ggtggaggggaaaacattaagaatttattcattatttctcgagtactttc

agaaagtctgacactttcattgttgtgccagctggtt

Pol3s polyserase 3 Entrez Gene ID No. 339105

(SEQ ID NO: 114)

cagaccctgttccttcgaggaatggggagggagggagggaccaaagccgt

gaggatgaggacaactccaccctccttccttccccacaggccaaccaacc

-continued

agctgctgacaggggacctggccattctcaggacaagagaatgcaggcag

gcaaannngcattactgcccctgtcctnccccaccctgtcatgtgtgattc

caggcaccagggcaggcccagaagcccagcagctgtgggaaggaacctgc

ctggggccacaggtgcccactccccaccctgcaggacaggggtgtctgtg

gacactcccacacccaactctgctaccaagcaggcgtctcagctttcctc

ctcctttacccttcagatacaatcacgccagccacgttgttttgaaa

Pparbp PPAR binding protein Entrez Gene ID No. 5469

(SEQ ID NO: 115)

ggcttaggcctcaaatggcttcttctaaaaactatggctctccactcatc

agtggttccactccaaagcatgagcgtggctctcccagccatagtaagtc

accagcatatacccccagaatctggacagtgaaagtgagtcaggctcct

ccatagcagagaaatcttatcagaatagtcccagctcagacgatggtatc

cgaccacttccagaatacagcacagagaaacataagaagcacaaaaagga

aaagaagaaagtaaaagacaaagatagggaccgagaccgggacaaagacc

gagacaagaaaaaatctcatagcatcaagccagagagttggtccaaatca

cccatctcttcagaccagtccttgtctatgacaagtaacacaatcttatc

tgcagacagaccctcaaggctcagcccagactttatgatt

Prkd2 protein kinase D2 Entrez Gene ID No. 25865

(SEQ ID NO: 116)

gggagagggaggagtaatggaggaggagttggaaactggggagagatgga

aggaatgtgactggagggtagagaacttggagaa

Prr7 proline rich 7 (synaptic) Entrez Gene ID No. 80758

(SEQ ID NO: 117)

gaatcggacatgtccaaaccaccgtgttacgaagaggcggtgctgatggc

agagccgccgccgccctatagcgaggtgctcacggacacgcgcggcctct

accgcaagatcgtcacgcccttcctgagtcgccgcgacagcgcggagaag

caggagcagccgcctcccagctacaagccgctcttcctggaccggggcta

cacctcggcgctgcacctgcccagcgcccctcggcccgcgccgccctgcc

cagccctctgcctgcaggccgaccgtggccgccgggtcttccccagctgg

accgactcagagctcagcagccgcgagcccctggagcacggagcttggcg

tctgccggtctccatccccttgttcgggaggactacagccgtatagaggg

gcgcccggcgcccccgggccccaccggcggactcctggcctgactgcgggg

ctttttaaatgcttccctggactgcggggaggggcgggggggagggaggga

tttcttatcccgtttgttacatt

Psph phosphoserine phosphatase Entrez Gene ID No. 5723

(SEQ ID NO: 118)

ttttctactcagcagatgctgtgtgttttgatgttgacagcacggtcatc

agtgaagaaggaatcggatgctttcattggatttggaggaaatgtgatca

ggcaacaagtcaaggataacgccaaatggtatatcactgattttgtagag

ctgctgggagaaccggaagaataacatccattgtcatacagctccaaaca

-continued

acttcagatgaatttttacaagttacacagattgatactgtttgcttaca

attgcctattacaacttgctataaaaagttggtacagatgatctgcactg

tcaagtaaactacagttaggaatcctcaaagattggtttgtttgttttta

actgtagttccagtattatatgatcactatcgatttcctggagagttttg

taatctgaattctttatgtatattcctagctatatttcatacaaagtgtt

ttaagagtggagagtcaattaaacacctttactcttaggaatatagattc

ggcagccttcagtgaatatt

Rfx3 regulatory factor X, 3 (influences HLA class II expression) Entrez Gene ID No. 5991

(SEQ ID NO: 119)

tagagctgaatattacttgattacaaatcagattgcttaagggtgtggaa

tagcaggctagttttaataccaacttgttaacataaaatcatatatgttt

taganccattcttatttagttacaattttagaaagttaacaaagtaagca

ggtacttatcgaagtgcatcttttcagtctaaatgtttgtctgtgtgtct

aggtgctggtgagtccacatggacacatgnagnnnccatggggcaggagt

ctgctataaagtcagaaggtgagatcctagagagttacacccagccccat

tttaatttgcatgaaaagccaaggttcttttaagcactcaaattatttaa

tgnttaaaacacaagaaaggcacatctgttcatttaaat

Rps16 ribosomal protein S16 Entrez Gene ID No. 6217

(SEQ ID NO: 120)

agagggccttggagtgacaccctgacccccatccactagtacttganggc

cagtggtggcagaagccacagaaacaagaagcccagtgagatggctaagc

tgcccagcatgtaacttaaatccctgttcattccccattcctttagctgc

tggagccagttctgcttctcggcnaggagcgatttgctggtgtagacatc

cgtgtccgtgtaaagggtggtggtcacgtgggcccngatttatnnnagtc

ccanaactgggcgcatggaggaggtggctctgggagggaggccttcacag

cgctcctgtacccttaattgtgtgtctttctcacagctatcc

Samd4 asterile alpha motif domain containing 4A Entrez Gene ID No. 23034

(SEQ ID NO: 121)

ggtgaatgtgtattcctctgggaggaataggaagaaaacaggaatgttaa

taatgtcgaacagaaaacttcctcccttattaatatataatcctcatgta

tttatgcctaatgtaagctgacttttaaaaagctttcttttgttgcatgc

cctgtgcaggcatctgtattgtacatgcatgcctttcgtcctgttttcct

gtataaagttagtgaacaaagaaatatttttgcctagttcatgttgccaa

gcaatgcatatttttttaaatttgtcatatatggaaagagcatgtttgtta

catgtaaaagctttactgatatacagatatactaatgtttgaagatgctg

ttctttgcaagtgtacagttttcaaatgttgttaccagtgaaacaccctt

gtggtttaaacttgctacaatgtatttattattcatttcctcccatgtaa

ctaagaa

-continued

Scamp1 secretory carrier membrane protein 1 Entrez Gene ID No. 9522

(SEQ ID NO: 122)

tcagttcatatctttctgggcttgacatggctgatggtgtagctgaaacc

ctcctaacactaaaagccatttaatcttttctgtaataggagcagaaaat

agttaatcatccacctagtaatataagattactgtgaatattatcttcta

tacattaaaacagttctagtttgtagaataataccatacaagttttattt

ttaaattctagttattttcagtgcttacttaaatgtaattctagaattcc

tccacaacttttaatattttgtatgccagtgattctcaagataaatcatg

attgtagtagttgttactgttggcagtttgtagtagtattcaggtatttt

ggggatgggggaaaacaccaaaaatcagtgtcttttatctggtgatcact

gtggtatctacagtattctagtctcctgcacaaaaactgaacccactggg

cc

Scn11a sodium channel, voltage-gated, type XI, alpha Entrez Gene ID No. 11280

(SEQ ID NO: 123)

ccaaccatgatgaaactccgtctctactaaaatacaaaaattagctgggc

atggttgcgtgcgcctgtagtcccagctacttgggaggctgaggcaggag

aatcgcttaaacctgggagacggaggttgcagtgagccaagatcgtgtca

ctgcactccagcctggtgacagagtgagactctgtttcaaaaaagaaaag

aaaagaaacatggttcaaattatatctaaacaaaaaagaataagaaacaa

aaaacacattaaaattttaagttgtattttctatgtttctagatacatca

tttttgtttgatattttcctgatgcaagtatgtggtttatcacatgtagc

tcttttgcatgctaaatgaaaattcaagacttgccaataaatgaatagct

tattgcagacatttttttaccaacattaattattttgggtttgtttaaaac

ctagaggcacaatcttgacttgtcaattactacccttttcacaagctacca

tctcagatatatatatatatataaattcaataaagctttctgtttgtgtt

c

Sdccag8 serologically defined colon cancer antigen 8 Entrez Gene ID No. 10806

(SEQ ID NO: 124)

cttcacaatagcaaacgtaaacgatggaattgatggaatcaaccgaaatt

gacggaatcaatctaaatgttcatcactgacagattgtgtaaagaaaatg

tggaacatggacaccatggaatagtatgcagccataaaaagaatgagatc

cgatcttttgcaggaacatgcatggagccggagacagttatccttagcaa

actaacgcaggaagagaaagccaaatactgcatattcttacgtataagtg

ggagctaaatgataagaacttatgagcacaaagtaggaaaccacagacag

tggcatctccttgaggatatagggtgggagcagggagaggagcagaagag

atcactattgggtactgggcttaatacctgggtgataaaataatctgtat

aacaaaaccccgtgacatga

Sephs1 selenophosphate synthetase 1 Entrez Gene ID No. 22929

(SEQ ID NO: 125)

aaaggtgttctctgtgttatgtaaagtggaggcttccttatattttaacc

tactaagcaatgaggagggattcctgtcattaagcacaagggcgctggat

-continued

cctcaagtgcccatcttcgtgagagaaaaagcagcacatcctgcccattt

ctggtgctttctgctcacaggcaccaaagctgcacatgtaaactgacttc

ttgccaaaggaaatgacccctgggaagttcaagctcctggaagaggcttt

aactcggacgcgccctcctccaggaaccagtgggcagggcagccttcatg

catgtgtaactggacctccagccataagcatggtgtgcagtatggaagag

cctgctacggaactgaaagtgattggacattttataggaattgatagaga

tgttggtcctcaaaagctaca

Slc2a13 solute carrier family 2 (facilitated glucose transporter), member 13 Entrez Gene ID No. 114134

(SEQ ID NO: 126)

aacaacattattccatctcatttaaaggttnaaaaagaagagacaactct

agccnaagtagaaatttatattctacacgtccaaactgtctcctagcagc

ttttggactatatatcacttgatgttaaagtatcttttatttgtaataaa

tattcaaatttctatttagaagctctaatgtatacctagattaaatcaaa

tcacagttttatgcttttaaaatatatgtatttcaaactgtatattttaa

tttctgagtgcatgttatatagtatttaatacttcagatgtcttggcaaa

ttcaatataagtatttattcccacaagcgatatatgggatatctcttaaa

aattatgaatatgtaccatttccttcaaagtcatcctagcctatgctgta

tcaaaagtattgtatattttatggagatttagtgatatacatgtaaatgt

tttttaagttattttattgaagttcaatctttacataaaattaaaatctt

tttttaaaaaaagtgtcagtgccagaactgtaa

Slc30a5 solute carrier family 30 (zinc transporter), member Entrez Gene ID No. 564924

(SEQ ID NO: 127)

tgttcataaacatttgagcaccatgaaatcaaaatacccctataactactt

tctatagtcatatctaatttatatttttttcatttccanttgtaactaga

tatgtagtaaagtctgaaaagactttaccatagacaataacatgcagttt

tatcagcaccaaagaatgttgtccaaaagaaacttttttaatacctgtctt

tctatttataacatctgaatattttcattcttatattaagaattttgata

agtagattgaatttagtatgagtactatttttcttatatataccacaatgg

caaacatgtattataaatcatatttttgtcttaccaattttaatatatga

ggggttttagaaatttgttgtaagttatttttatattccttgtcttttgc

atatttttttggccaaaatcttcaatacat

Spa17 sperm autoantigenic protein 17 Entrez Gene ID No. 53340

(SEQ ID NO: 128)

ttcgaggagcaagaaccacctgagaaaagtgatcctaaacaagaagagtc

tcagatatctgggaaggaggaagagacatcagtcaccatcttagactctt

ctgaggaagataaggaaaaagaagaggttgctgctgtcaaaatccaagct

gccttccggggacacatagccagagaggaggcaaagaaaatgaaaacaaa

tagtcttcaaaatgaggaaaaagaggaaaacaagtgaggacactggtttt

acctccaggaaacatgaaaaataatccaaatccatcaaccttcttattaa

-continued

tgtcatttctccttgaggaaggaagatttgatgttgtgaaataacattcg

ttactgttgtgaaaatctgtcatgagcatttgtttaataagcataccatt

gaaacatgccacttgaagatttctctgagatcatgagtttgtttacactt

gtctcaagcctatctatagagacccttggatttagaattatagaactaaa

gtatctgagattacagagatctcagaggttatgtgttctaactattatc

Tcf7l2 transcription factor 7-like 2, T-cell specific, HMG-box Entrez Gene ID No. 6934

(SEQ ID NO: 129)

gaaatggccactgcttgatgtccaggcagggagcctccagagtagacaag

ccctcaaggatgcccggtccccatcaccggcacacattgtctctaacaaa

gtgccagtggtgcagcaccctcaccatgtccacccccctcacgcctcttat

cacgtacagcaatgaacacttcacgccgggaaacccacctccacacttac

cagccgacgtagaccccaaaacaggaatcccacggcctccgcaccctcca

gatatatccccgtattacccactatcgcctggcaccgtaggacaaatccc

ccatccgctaggatggttagtaccacagcaaggtcaaccagtgtacccaa

tcacgacaggaggattcagacacccctaccccacagctctgaccgtcaat

gcttccatgtccaggttccctccccatatggtcccaccacatcatacgct

acacacgacgggcattccgcatccggccatagtcacaccaacagtcaaac

aggaatcgtcccagagtgatgtcggc

Tmem30b transmembrane protein 30B Entrez Gene ID No. 161291

(SEQ ID NO: 130)

tatactcactcaaggcagtgcaagatcttgaagtacttttagcagttaa

gtaatattgaattgtattgaatagtttacatagtttattctagtctttga

aaattactgaacatggacaatgtgcatgtcattgacatctgccttagaac

ttctgggacaatcctgattcgagagattctatcccattatttacatatac

caaaaatactttgttaatttaatgtgttggcttcccaactcctgaacacg

acacaattttattattagattttgtatggtgattttaggctatgaaaaca

tgatcattatatgtatatagatacatttttatttgttacaaatgtttgag

cagctcactagcccacccctcctctattttgggtaagagaatttactacc

ttttttaactatgtagttgagagcaacatgtattttgttatttttagaat

ggtcagtatattgctataaaatttaaatgagactatgaaagttaaagta

ttctgattctggttaaattaacgaatatggttccaggccctgt

Trpm7 transient receptor potential cation channel, subfamily M, member 7 Entrez Gene ID No. 54822

(SEQ ID NO: 131)

gttgcagtgatgacttttgtgaaacaaaaacttatgtatcattttagtga

tactcttaagattatttgtttttttggcagtaaatgtgaaaattctttgtt

gttctactttatgaatagaacttaaggaaataactccaaaacaatgtaat

ttgtataagaaggttcataaaaatcctgtaaggtttaaattagtttagaa

gaaaaataatagtttgctgtaacttttttctccctaaagaaacaaggtcca

actaatccaatgctgtttcatcttgttcgagacgtcaaacaggtaagaga

ttatttttttgcttttga

-continued

Wbscr16 Williams-Beuren syndrome chromosome region 16 Entrez Gene ID No. 81554

(SEQ ID NO: 132)

agagctgagtcatcctagagcaaacctctggagtggagagcgaactactt

cattcccctcccttagcctgggccagagagactccagctctgccttctcc

agccaaaaaatcaaaggcagatgggagaacagccttcagctttggataac

gatgaaatatctggcaccactgatgaatattaaactttctataacc

Wdr20 WD repeat domain 20 Entrez Gene ID No. 91833

(SEQ ID NO: 133)

ggagactgtctcactgatgttgatttctttattcatttccgcatctgtta

cacgaacttcgtgtcataaattgctatcctttcatttgaaagtgtaaaaa

atttcctgcatttttatcatttctgtatacttgagtttattagagattgt

tatgttaggcgacactgtataaaattgtatggatattttgagtgaaaatc

aaaagtaaaattcacatgtatttccttttttatattttcatccaatttct

tgacaacttgaataaatttcataaagagccttcctaa

Wdr55 WD repeat domain 55 Entrez Gene ID No. 54853

(SEQ ID NO: 134)

gcaagctctcattggctctgagcgcgaccccgcctcccaggggggtggag

gtatccactgcacgtgcgccgcccgggcttcgctcagaccttcaggtgaa

agctgcaaagtcgcgggtgcgtatgtacgggggctgcctcccgaggagga

gctcccaagccgcagggtggacgctggagacaagaacctcagggtcacaa

gtttactgtttttctccctttccatccctacattggtctgctggggaag

gcggggctaggcatcactgacacacgcagactccgtggttgaggcatttt

attggacctttggcaattggtggtggggaggcatctgctccaactggtgc

ggggccctgcagatgggaccatctcaggctgggtccttgtagcccaggag

cacagactggactaagcctcctgggccttgtatgaaaaaggtgttgtacc

tggccgtttttgccagt

Znf492 zinc finger protein 492 Entrez Gene ID No. 57615

(SEQ ID NO: 135)

actccgtcctgggtgacaaagtgagactccctctcaaaaantaaatangt

aaataaantaaatggtggtaacnatacnctatttggtaaannnnnncnct

aacatctgtagtactaatctttttttccagtggctttaaactgcaaataag

gaatgttgtttctgtaggtaaaatttttatttatttttttcccatttaaat

ttacttttgttagtttttttcaggcatataatatttatgttatatatggca

tattctgataagaggcatacaatatgtaataatcacattagggtaaatga

ggtatccatcacctttagaatttatttttttgtattatgaacagttcaatt

gtacagttttagttttttttaaaaatatacgattgttattgactacagggt

Znf502 zinc finger protein 502 Entrez Gene ID No. 91392

(SEQ ID NO: 136)

tgataagactcagaacaggaagcctgcatgtgactgagcaagtcacctac

ataaccctgcctgtactaaggtgtacacctgtctattgtaagtttgccta

ggctgttggtgtacagagaccagaggagagagacacactaggactaacaa

tgtcctaacaaaatggtacttagtttgttggtctttaggagaaagcatta

-continued

gtaatgaaagaagaaagaattttcacttggttggacattggggctgctta

agaaagttgacatttgtcgtggaatgactttggaaagacttctaaaagaa

tctttttcaaatccctgaaaatcaggatagcacattttgctactgactgt

gacagtgttttattcttttgagagaaatgacatagttttccctttatttc

ccaaattcctttcatgttcttaactgctacccagaaattgagcttcagaa

gattgaggatagcctttgattggtattta

Znf557 zinc finger protein 557 Entrez Gene ID No. 79230

(SEQ ID NO: 137)

agtagatcttaccttgctgttcataagagaatccacaatggggagaaacc

ctatgaatgcaatgactgtgggaaaaccttcagcagcagatcttacctta

ctgttcataagagaatccacaatggggagaaaccctacgaatgcagtgac

tgtgggaaaaccttcagcaattcctcatacctcagaccgcacttgagaat

tcacactggagaaaaaccgtacaaatgtaaccagtgttttcgtgagttcc

gcactcagtcaatcttcacaaggcacaagagagttcatacgggggagggt

cattatgtatgtaatcagtgtggaaaggctttcggcacgaggtcatctct

-continued

ttcttcgcactatagcattcatacaggggagtacccttacgaatgccacg

attgtgggagaaccttcaggaggaggtcgaatctgacacagcacataaga

actcatact

Znf576 zinc finger protein 576 Entrez Gene ID No. 79177

(SEQ ID NO: 138)

aagctgttgacagggctgcttttcttttggaggctctaggggagcgtct

ttctttgcccttccagcttctagaagctgcccaaattctgtggtttgggg

cctcctttcaaaaccagcaatggccaatcagtcttacatcactcaaacac

ttgagtgttctgtctccctcttccatgtttgaggacccttgtgattacac

tgtgaaaacccagataagccaggataatctccctatcttattatgaggca

agtatgttaagattttattctataatcagagaatcttatgctatgattgt

tatatgtgagcattatagatgctcttgaaatgttaaaatcacatcagcac

tggaaaataactcctaaatgtccaaaaagaacatgagatttatggtgctt

gaaatgttgctaaacgtaaatttgtatctattctgaaattatataaatta

acctacctggccaggca

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 138

<210> SEQ ID NO 1
<211> LENGTH: 534
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

gaagataacc ggctcattca cttcctccca gaagacgcgt ggtagcgagt aggcacaggc 60

gtgcacctgc tcccgaatta ctcaccgaga cacacgggct gagcagacgg cccctgtgat 120

ggagacaaag agctcttctg accatatcct tcttaacacc cgctggcatc tcctttcgcg 180

cctccctccc taacctactg acccaccttt tgattttagc gcacctgtga ttgataggcc 240

ttccaaagag tcccacgctg gcatcaccct ccccgaggac ggagatgagg agtagtcagc 300

gtgatgccaa aacgcgtctt cttaatccaa ttctaattct gaatgtttcg tgtgggctta 360

ataccatgtc tattaatata tagcctcgat gatgagagag ttacaaagaa caaaactcca 420

gacacaaacc tccaaatttt tcagcagaag cactctgcgt cgctgagctg aggtcggctc 480

tgcgatccat acgtggccgc acccacacag cacgtgctgt gacgatggct gaac 534

<210> SEQ ID NO 2
<211> LENGTH: 427
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

aatgagcgcc acctttaggc agatcctctg ctgccagcgc agtgagaacc ccaccggccc 60

cacagaaagc tcagaccgct cggcttcctc cctcaaccac accatcttgg ctggagttca 120

-continued

cagcaatgac cactctgtgg tttagaacgg aaactgagat gaggaaccag ccgtcctctc 180

ttggaggata aacagcctcc ccctacccaa ttgccagggc aaggtggggt gtgagagagg 240

agaaaagtca actcatgtac ttaaacacta accaatgaca gtatttgttc ctggacccca 300

caagacttga tatatattga aaattagctt atgtgacaac cctcatcttg atccccatcc 360

cttctgaaag taggaagttg gagctcttgc aatggaattc aagaacagac tctggagtgt 420

ccattta 427

<210> SEQ ID NO 3
<211> LENGTH: 529
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

ctcctctcca cctgctggtg agaggtgcaa agaggcagat ctttgctgcc agccacttca 60

tcccctccca gatgttggac caacacccct ccctgccacc aggcactgcc tgagggcagg 120

gagtgggagc caatgaacag gcatgcaagt gagagcttcc tgagctttct cctgtcggtt 180

tggtctgttt tgccttcacc cataagcccc tcgcactctg gtggcaggtg cttgtcctca 240

gggctacagc agtagggagg tcagtgcttc gagccacgat tgaaggtgac ctctgcccca 300

gataggtggt gccagtggct tattaattcc gatactagtt tgctttgctg accaaatgcc 360

tggtaccaga ggatggtgag gcgaaggcag gttgggggca gtgttgtggc ctggggccag 420

ccaacactgg ggctctgtat atagctatga agaaaacaca aagttgataa atctgagtat 480

atatttacat gtctttttaa aagggtcgtt accagagatt tacccatcg 529

<210> SEQ ID NO 4
<211> LENGTH: 487
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

aatcctaaat gtgtggtgac tgctttgtag tgaactttca tatactataa actagttgtg 60

agataacatt ctggtagctc agttaataaa acaatttcag aattaaagaa attttctatg 120

caaggtttac ttctcagatg aacagtagga ctttgtagtt ttatttccac taagtgaaaa 180

aagaactgtg tttttaaact gtaggagaat ttaataaatc agcaagggta ttttagctaa 240

tagaataaaa gtgcaacaga agaatttgat tagtctatga aaggttctct taaaattcta 300

tcgaaataat cttcatgcag agatattcag ggtttggatt agcagtggaa taaagagatg 360

ggcattgttt cccctataat tgtgctgttt ttataacttt tgtaaatatt acttttttctg 420

gctgtgtttt tataacttat ccatatgcat gatggaaaaa ttttaatttg tagccatctt 480

ttcccat 487

<210> SEQ ID NO 5
<211> LENGTH: 519
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

tttggcgtcg tcctcaagtt atattagaat cgtgtcctcc cggctttggc caacttacta 60

ttctaggact tgattccttc attcagtcac aatttattga gcaccgactt tgcatcaacc 120

-continued

tcttgctgaa gataacagtg ctgacaatat acagccctgc cctcagagct tatatagtag 180

aggagaaaaa gtgaacccat aatatacagt cagtagcgag tatttactaa gtactttcta 240

tttgcgaggc cctgataaaa gtactgtcct ggccaggcgc ggtggctcac gcctgtaatt 300

ccagcacttt gggaggtcga ggtgggcaga tcacctaagg tcaggagttc gagatcagcc 360

tggctaacat ggggaaaccc cgtctctact aaaaatggaa aaattagctg ggcatggtgg 420

cgggcgcctg taatcccagc tactcgggag gctgagacag gagaatgact tgaacccagg 480

agttgcagtg gccaagataa gatagcgcca ttgtactcc 519

<210> SEQ ID NO 6
<211> LENGTH: 534
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 6

tgtgaagctt tacgcgcaca cggacaaaat gcccaaactg gagcccttgc aaaaacacgg 60

cttgtggcat tggcatactt gcccttacag gtggagtatc ttcgtcacac atctaaatga 120

gaaatcagtg acaacaagtc tttgaaatgg tgctatggat ttaccattcc ttattatcac 180

taatcatcta aacaactcac tggaaatcca attaacaatt ttacaacata agatagaatg 240

gagacctgaa taattctgtg taatataaat ggtttataac tgcttttgta cctagctagg 300

ctgctattat tactataatg agtaaatcat aaagccttca tcactcccac atttttctta 360

cggtcggagc atcagaacaa gcgtctagac tccttgggac cgtgagttcc tagagcttgg 420

ctgggtctag gctgttctgt gcctccaagg actgtctggc aatgacttgt attggccacc 480

aactgtagat gtatatatgg tgcccttctg atgctaagac tccagacctt ttgt 534

<210> SEQ ID NO 7
<211> LENGTH: 529
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (44)..(44)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 7

attccatgtc attctgttta cttagcactt gcactaccct tgtnggttga gtgtatgctt 60

tatttgtttc tagtttgaaa tcccacatct gatagctgag agtaggcaaa tacaacattt 120

acctaatgtc attcactaac atggaagagt tgtgaaaatt ctagagtgct gtaaatcctt 180

ggcatacact atgacaaaca acttcattac tctcccacca ggagctgctc tcctgcactt 240

agaaataatg tcacaagtag ttttctaatg tacaatgcag acaaatgtac tgctctctga 300

atacttgaag aaatggtatt atacatacat agaaacttat tagttatacc ttttcacaat 360

cttattacga tgttgccgtt aaaagggaaa aaagacacag gcaatgaatg gtgggatagt 420

aagaggactt agagtgtatg aatgagttga ttttactttt ttggaatttg attaagttga 480

cagtaggcac tgattggatg attaaacata agttaatctc cactgtgat 529

<210> SEQ ID NO 8
<211> LENGTH: 401
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 8

cttatggtat gtagccagct gtgcactttc ttctctttcc caaccccagg aaaggttttc 60

cttattttgt gtgctttccc agtcccattc ctcagcttct tcacaggcac tcctggagat 120

atgaaggatt actctccata tcccttcctc tcaggctctt gactacttgg aactaggctc 180

ttatgtgtgc ctttgtttcc catcagactg tcaagaagag gaaagggagg aaacctagca 240

gaggaaagtg taattttggt ttatgactct taacccccta gaaagacaga agcttaaaat 300

ctgtgaagaa agaggttagg agtagatatt gattactatc ataattcagc acttaactat 360

gagccaggca tcatactaaa cttcacctac attatctcac t 401

<210> SEQ ID NO 9
<211> LENGTH: 552
<212> TYPE: DNA
<213> ORGANISM: Drosophila melanogaster

<400> SEQUENCE: 9

agagacatgt accttgacca tcgtccttcc tctcaagcta gcccagaggg tgggagccta 60

aggaagcgtg gggtagcaga tggagtaatg gtcacgaggt ccagacccac tcccaaagct 120

cagacttgcc aggctccctt tctcttcttc cccaggtcct tcctttaggt ctggttgttg 180

caccatctgc ttggttggct ggcagctgag agccctgctg tgggagagcg aaggggtca 240

aaggaagact tgaagcacag agggctaggg aggtggggta catttctctg agcagtcagg 300

gtgggaagaa agaatgcaag agtggactga atgtgcctaa tggagaagac ccacgtgcta 360

ggggatgagg ggcttcctgg gtcctgttcc cctaccccat ttgtggtcac agccatgaag 420

tcaccgggat gaacctatcc ttccagtggc tcgctccctg tagctctgcc tccctctcca 480

tatctccttc ccctacacct ccctccccac acctccctac tcccctgggc atcttctggc 540

ttgactggat gg 552

<210> SEQ ID NO 10
<211> LENGTH: 462
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10

gcgcgcctgc tgttgcagag gagaatccta aggagagtaa accccaagca ggcactgccc 60

gcccacagga tgtgaaccgc agagaccaac agaggaatcc aggcacctct accacgccct 120

cccagcccaa ttctgcgggt gtccaagaca ctgagatggg cccatgccgt agacatctgg 180

actcagtgct gcagcaactc cagactgagg tctaccgagg ggctcaaaca ctctacgtgc 240

ccaattgtga ccatcgaggc ttctaccgga agcggcagtg ccgctcctcc caggggcagc 300

gccgaggtcc ctgctggtgt gtggatcgga tgggcaagtc cctgccaggg tctccagatg 360

gcaatggaag ctcctcctgc cccactggga gtagcggcta aagctggggg atagaggggc 420

tgcagggcca ctggaaggaa catggagctg tcatcactca ac 462

<210> SEQ ID NO 11
<211> LENGTH: 424
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (291)..(291)

-continued

<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 11

gaagcaacgt tttcatctga ttggaaatac cgtattgctg aaaagaagaa aggccttttt 60

aatggctttt gaacaaagca gaaaagtttg agcttctcac cttcagtctt agctcttgaa 120

cctgttgaga aagaggataa gagacaaata cggaaaagag tttcagaaag cagaatctgt 180

gtcagcccac tggaaggaaa agcgaatcaa ccgattcagt gatgttagtg catccagaaa 240

caggcttttg ggaaaagctt gacctgagct gattaaatcc tgaagcacaa nggaagcagc 300

cacatcaaaa agttagcatg agagcagtgg cgtgctcatc ctctggtagc ctttactggg 360

catttgtgga gtaagagaga aaagaaaagc aggaatgtta agatatgcta ctaccttcag 420

gaaa 424

<210> SEQ ID NO 12
<211> LENGTH: 383
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

gagcagcgta gacagctggt aaactgaaga gcacaactat attcttatga aggaatttgt 60

acctttgggg tattattttg tggcccgtga ccctcgttat tgttacagct gagtgtatgt 120

ttttgttctg tggagaatgc tatctggcat tatggtaata tattatttta ggtaatattt 180

gtactttaac atgttgcata atatatgctt atgtagcttt ccaggactaa cagataaatg 240

tgtaatgaac aaagatatgt tgtatgagtc gtcgtttctg tcagatttgt attgtttcca 300

agggaaaagc ttgggggagg actcagttca caaaatgcaa aactcaacga tcagattcac 360

ggacccagag cttttccatg tgt 383

<210> SEQ ID NO 13
<211> LENGTH: 200
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

tttttctctt ctggcttcat aaatgccttg ctgtataaat tgaaatattg atactgaact 60

gtctttttaa tgatgaccta actttattca acccatcgga atttacttttt tccctgaaat 120

aagatctttt ccactggtct actacctgac cataaacatg tctgcatttg aattctctaa 180

accctaaatc tgtgtctatg 200

<210> SEQ ID NO 14
<211> LENGTH: 498
<212> TYPE: DNA
<213> ORGANISM: Mus musculus
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (105)..(105)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 14

aaacagagca gccacgggct cgaaccgaat ccccgccgtc cttagaaacg gatttttttt 60

tgttttgttt tgttttctgg cagagtctcg atcaccaccc tactnccacc cccactaagg 120

ttcttgctca atctccctag aaaacctgaa ttgtttcatc cctttcagtc agcccctac 180

-continued

gtggtctgaa acaaaatgaa agcacaagcc acggagttta agaaggcagc ctgaaggcgg 240

ggggctgaag aggggtcggg gcgctgcaga gtcagccaag tagccaagga agggcccccc 300

ccgcgtcgcg acggccctcg ccccccgccc ggcgcgcgcg cgcacaaata cacacacaca 360

gtcactcaca cactcactca cactcacgcc gcgctccgac accgcctcac ctctcgctcc 420

gcccgtccgg ccccttccc tgccccctgc gggaccggcc tgcgcgcagc actggaacca 480

cgtaggagga ggcggcgg 498

<210> SEQ ID NO 15
<211> LENGTH: 476
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (235)..(236)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 15

gaaggtatac agactgccaa cattttagac ttatctctca gtctctgcca ctgaactttt 60

atatatggct gctatcaaaa tataaccggt ttattttcat atttggaact aatataacag 120

tatcacaaaa tcttttacag taagatagtt tgtaatacca gccgacccag ctgcttaact 180

gagtccttaa atcatttaat atatgggact gtaaatagag aaatctgtac attannagat 240

ctgatttctg gttatgccta tagatcttta ttttctttat ccctatagat cattttcttc 300

tgatgttaag tgttatattt ttgaaatgct cctaaacaag taagccttag attgtattaa 360

tcctgaggat tgataccatt tcctcaacac tttgtggagt atgattgaca ccagtttttt 420

tttgactgca ggtttaactt ggcttatcac ttttcgtatt gctcagtatg tccaag 476

<210> SEQ ID NO 16
<211> LENGTH: 334
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

cacagttttg tgaagaagcc tttaatccaa gttttatgag acagtaggaa cctatagcta 60

cttaattttt aggagacagt aattcagttg cagaagcttt cctctgctat tcccattctc 120

ttttacaaaa ctagtttttt ttaaaaaatc aatgatcaat tttatcttac tactttataa 180

cttttgctga cttttattct ttgcattgta tgtaatgtcc atcagtataa attgagactg 240

tgaatttcta ggacccacaa aagtggtatt tttttttgt acaagaaagt atagaggaag 300

aggaagctgg gcattaaatt acctcatcca gcag 334

<210> SEQ ID NO 17
<211> LENGTH: 409
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

tatcactggt aaggagccca cgaccagact tcatttctgg gaaaatgaga ctttgtgttg 60

atctcatcgt gttggccttg taaaagtgat ctatgcatgt acagtgttca tgcttaatat 120

tcaagggatg gggcggggaa caaaaggaat agaaagaatt cttttccttg ttatttgggg 180

agcacgtatt gctttataac tttggttgtt gggagtatgg ctatcatata ccctcatcag 240

tgtcatttta tatctgccta attagagaaa ttttaaccct agtattttga tgtgttttcc 300

ccattttatc ctccgcaaat atctttctct tgcccattca gtgctgcttt tggtttttga 360

-continued

tttagttgta tattctggat gtatttccac agccttttat tgttcttcc 409

<210> SEQ ID NO 18
<211> LENGTH: 418
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 18

gatccagtat gtgcttgtct tatttaaaat tggaatgtga gacatgttgc tgtgacctgt 60

ttttctttct cattcacatt tgtagatatt gtgtgaacta cagtatataa tgataacaat 120

taaaaggata ttctgtggat gtcacgtatt ttgaaatgat agaactacat tagctttgta 180

tcatgtttgg ataattcatc aatgttcaca gtttaaaaca tcattaaaca ttatgtaatt 240

acaatgagaa agaatcttac ttaaatttgg agattttccc ccacatctct tttccggata 300

cattataatt ctggacccct atttatctca aaactcttaa tatatgcaga ccaacaggtc 360

tttgcattcc ttttaaataa ctggttgtga caaagcttgt tgttgatcag attcactg 418

<210> SEQ ID NO 19
<211> LENGTH: 149
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 19

aactgcttta agtcatgctt atgctgctgg tgccagtcat ttgaagaaaa acagtccttg 60

gaggaaaagc agtcgtgctt aaagttcaaa gctaatgatc acggatatga caacttccgt 120

tccagtaata aatacagctc atcttgaaa 149

<210> SEQ ID NO 20
<211> LENGTH: 330
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

tcgaggtatc agcagctctg tcctcagaat gggtccccca cttcacacag ttgtaggatg 60

gctacagcag ctccaagcag cacattcaga ggaagaagaa aaaatgtttc atttgtgtgg 120

ttttaagcat aaagaagttg tttcccagag ctctttgcca gctgtcatcc ctcaaaactc 180

actggccaca attgcttcac atgcccctgc ctgaaccagc caccagaggg ggttaggacc 240

accacagatt gactagatcc taaggattct ctacctgggg ctggagttcg ggtcaggtgc 300

ccgggaagca ctgtgcggtg tgggaggatg 330

<210> SEQ ID NO 21
<211> LENGTH: 402
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (92)..(92)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (94)..(99)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (155)..(155)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

-continued

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (180)..(180)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (211)..(211)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (345)..(345)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (371)..(371)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 21

tgttgttgtc cctattgctg gtttattaca ctgtacagac cacaaaatgt aatattcttt 60

tgtataacta ctaaagaaaa atccttgtag ancnnnnnnc cttcaccatg gctatctata 120

cctgtacatg aaatgtgttt gtattgtgct gaagngctta atgtcaacat tacctgctgn 180

ttactctgaa aaaaggaatg aatggtagct ntagaattta ggatatttta tcaggttggc 240

actttataaa atactccctg atttaaaaaa ttgtaagtta tacacgttaa tcatccacat 300

tctatcgaca atgtaccaac atcacaagct gttgcaacca cctgnctgtt acttctctga 360

gctgttaaaa ncctggaact tcaatttcag gggggcacaa at 402

<210> SEQ ID NO 22
<211> LENGTH: 503
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 22

ttacagtgca gtttagttaa tctattaata ctgactcagt gtctgccttt aaatataaat 60

gatatgttga aaacttaagg aagcaaatgc tacatatatg caatataaaa tagtaatgtg 120

atgctgatgc tgttaaccaa agggcagaat aaataagcaa aatgccaaaa ggggtcttaa 180

ttgaaatgaa aatttaattt tgtttttaaa atattgttta tctttattta tttgggggta 240

atattgtaag ttttttagaa gacaattttc ataacttgat aaattatagt tttgtttgtt 300

agaaaagtag ctcttaaaag atgtaaatag atgacaaacg atgtaaataa ttttgtaaga 360

ggcttcaaaa tgtttatacg tggaaacaca cctacatgaa aagcagaaat cggttgctgt 420

tttgcttctt tttccctctt atttttgtat tgtggtcatt tcctatgcaa ataatggagc 480

aaacagctgt atagttgtag aat 503

<210> SEQ ID NO 23
<211> LENGTH: 276
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (193)..(193)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (207)..(207)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (216)..(216)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

-continued

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (229)..(229)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (237)..(237)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (243)..(244)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 23

aatggaccaa ttttagctca acttttagtt tgttagaagc aagtgtagga actctagcac 60

tgtagttttt aattattgct tgtatctatt attattaatt ccaacagagt ataatgtata 120

tttattctat aaaatatata ttatcagagt gcatttgtta caacttaggt tcttttctta 180

ccaagtatta agnaatctag taagagnaat actagncaaa ggacctagnc cctgtgnaac 240

agnntctcgt atgttatata cataaaccca ctctgc 276

<210> SEQ ID NO 24
<211> LENGTH: 499
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 24

gtgtttctgc tttacagtgc tgaattccat attttagaag ctatgaaagt ccttttatga 60

aaaagttact gattgcttct cagttattag gaaaacagtt gtttcacaat tattatgtag 120

atatgatgcc caaatatcat ttttagtata tcttgtcgat ctttaagttg ttactattgt 180

gttattcatg tctttaaatc agataccaaa tatttttag gaaagaaaaa tgttattact 240

gtcattaggt tgtcttttaa tactttaagt tattttgacg aaaagtaata gagaaaattt 300

acttagcatt ttagattcta gagacatgga aatgaaaatt attttatgtc tagagtaggt 360

cctgaagttt ggctttacat taagtttagc actgtatcag aatgaagaaa ctaatatttt 420

acataaaaac taatactttc aatttttat atagtaatat ccccattttg taaatgttag 480

acttttatca tacctgtaa 499

<210> SEQ ID NO 25
<211> LENGTH: 417
<212> TYPE: DNA
<213> ORGANISM: Mus musculus
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (112)..(112)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 25

aaaataggca ttgcatacac atatgcacac gtatgtgcac gtgccacaca ttttttgtat 60

aatgttgggt ttgattataa aagtgttgtc aaatgtttta tttatctgca tntagcagtg 120

gttggctttt ttgaattgaa atttttgcgc attgatgcat tgaaataagg aaaattattt 180

atctctgagc actaaactta tttttgcata tttctgtaat attgcagtcc ccagatccag 240

aacatgggaa gttagggaaa atgtgtgatt ttgtgttttg aattactgtc agaattacat 300

acacaattac aacaaacttt ttttaaaaga catttcattg tactgcaaaa atctgaatat 360

ttatatttct tgtttttttc tttatatgtt ttgcatttta atatgttgag ccactgg 417

-continued

<210> SEQ ID NO 26
<211> LENGTH: 321
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 26

tcttttgaaa atgctggcct gctgcacctg ttctcagagg ggcatgagaa aagaggattc 60

tcttcctagg gctgtgggaa gccccttgag atgttggaag cagggcaggg caggagaacc 120

cagggagggc acagagctgt ggacgagggc tgggaaagcc atcccgcctc cccagggggt 180

ctccgaggag tgctgctgtg gccaaaccag gggggccact ttgtgctttc tgtttaggtg 240

atgggatgct tctatctcct cagcacccca caccaaatcc cctgttattg ccagatgatg 300

tgcctaatga tcaaattgct t 321

<210> SEQ ID NO 27
<211> LENGTH: 388
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 27

ctcccacctt ctaggcgaat agtccccaga gctgtgttcc tccaaggggt ccgaggaatc 60

actcactcct ggaggctggc aaggagacag tctgaggcca gggacacatg aagggatgtc 120

cccaccccag cactatcagg gcctccccag gcttccagag ttgaaagcca ggagaaaatc 180

ggcaaagacc acccttccct aaacccaagc acccaatgat gcaaaaaaca aaaacaaaaa 240

aaaaccacca aatccccaaa ttcattccag atctatttttt ctaccagaga gaggagcaaa 300

gtcctcctcc cctgcgccct tacattctgc acttcatagt tggattctga gcttaggatc 360

atctggagac cccatggagg gacttgga 388

<210> SEQ ID NO 28
<211> LENGTH: 451
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 28

ggtgatgtta gatctggaac ccccaagtga ggctggaggg agttaaggtc agtatggaag 60

atagggttgg gacagggtgc tttggaatga aagagtgacc ttagagggct ccttgggcct 120

caggaatgct cctgctgctg tgaagatgag aaggtgctct tactcagtta atgatgagtg 180

actatattta ccaaagcccc tacctgctgc tgggtccctt gtagcacagg agactggggc 240

taagggcccc tcccagggaa gggacaccat caggcctctg gctgaggcag tagcatagag 300

gatccatttc tacctgcatt tcccagagga ctagcaggag gcagccttga gaaaccggca 360

gttcccaagc cagcgcctgg ctgttctctc attgtcactg ccctctcccc aacctctcct 420

ctaacccact agagattgcc tgtgtcctgc c 451

<210> SEQ ID NO 29
<211> LENGTH: 457
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (131)..(131)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

-continued

<400> SEQUENCE: 29

ccttcagacc ccagttctag gggagaagag ccctggacac ccctgctcta cccatgagcc 60

tgcccgctgc aatgcctaga cttcccaaca gccttagctg ccagtgctgg tcactaacca 120

acaaggttgg ncaccccagc taccccttct ttgcagggct aaggccccca aacatagccc 180

ctgccccgga ggaagcttgg ggaacccatg agttgtcagc tttgacttta tctcctgctc 240

tttctacatg actgggcctc ccttgggctg gaagaattgg ggattctcta ttggaggtga 300

gatcacagcc tccagggccc cccaaatccc agggaaggac ttggagagaa tcatgctgtt 360

gcatttagaa ctttctgctt tgcacaggaa agagtcacac aattaatcaa catgtatatt 420

ttctctatac atagagctct atttctctac ggtttta 457

<210> SEQ ID NO 30
<211> LENGTH: 237
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (205)..(205)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (212)..(212)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 30

acacggatgt ggatgatgag cccccggctc cagccctggc ccggtcctcg gagcccgtgc 60

aggcagatga gtcccagagc ccgagggatc gtgggagaat ctcggctcat caggggccc 120

aggtggacag ccgccacccc cagggacagt ggcgacagga cagcatgcga acctatgaac 180

gggaagccct gaagaaaaag tttangcgag tncatgatgc ggagtcctcc gatgaag 237

<210> SEQ ID NO 31
<211> LENGTH: 436
<212> TYPE: DNA
<213> ORGANISM: Bos taurus
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (54)..(54)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 31

agatgtgtcc aaagagatcc ttacattaag gtcatttgtc agaatgatgt ttgngtttgt 60

ttagagttgg ctgacctcca actcctgggg tcaaggaatc ctactgcctt agcttcccaa 120

atagctagga ctataggcat gcacccggcc atgtgtttta tttatagctc ttaaagccca 180

gatgaagaaa tcacattttt gcccatagtg aagaaacatt tggccattga ttagtcctta 240

ttttcagtga ctgtcctgtt ttcattagat tagagagacc ctgtgtgggc cacagttaat 300

ataaaccatt atcactttta agtaaacctg cacatcttag atttcataat ttccttattg 360

ttctgactca aaatgaacta agagcttttc actttttgtt tgtaagttct cagagagtcg 420

ggtctgcaag tgcttt 436

<210> SEQ ID NO 32
<211> LENGTH: 317
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

-continued

<400> SEQUENCE: 32

gggtgcaggt aacccttggt ctgtaagcac caccgatcca gggatcattg tctaaatagg 60

ttactattgt ttgtttcatc ttgcttttgc atttttattt tttaatttcc aaatttaag 120

tgttccctct ttggggcaaa ttcttataaa aatgtttatt gtaaagttat atattttgtc 180

tacgatggga ttatgcactt cccaattggg attttacatc tggattttta gtcattctaa 240

aaaacaccta attattaaaa catttataga gtgcctactg tatgcatgag ttgagttgct 300

tctgaggtac attttga 317

<210> SEQ ID NO 33
<211> LENGTH: 431
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 33

gttactagac tataccttgg cctttaaaaa tgaatctcac tgattaaaga gaacaggcat 60

tattaggata gcattccacc acaactagaa acattcaaat aatgtgtctt aattgtaatc 120

tgtatatagg aaaatttttc ctatggatat ttttggtgtt ttaccacagt gaactgattt 180

gtagcactta tgaagtgcag aaggtaatat tcttgaaaat agaaaaaggt tgggtgagca 240

ggctttaatg cctttccccc aagaatatac atcgaatttt tcttaatctt ttggggttgg 300

ccagcttcca gatttcatta ataatgagct ctgcctttaa taaaagtaca tgatcatagc 360

tacactgtat gtttaggtgg tgtgaaatga tttataatca cagcttgaac tgtgtttgct 420

tggtactgtc a 431

<210> SEQ ID NO 34
<211> LENGTH: 547
<212> TYPE: DNA
<213> ORGANISM: Rattus norvegicus
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (125)..(125)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (127)..(130)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 34

ccctgtgcct cagtgacatg tagatgactg actgccaata cttgtcacca ttccctggaa 60

gcagctacct aggggaaaca agatgtagtg ctattgccga taacaagtaa gattttccac 120

actanannnn ggtgtttctc ttttctaaag tgaggccagt gttatttccc gggagtgttc 180

agtcttgacc ctagtcactg atttttttcta gttgttaata gagtggttgg cttttaaggt 240

tcagagactg tggcttggca cctgcgccca ggctttgtgg gcctttgccc cttagaaagt 300

agctgtaggc aaagatttgt gattttccaa ttacagtctc agctctagtt ttagtatctc 360

taattctttg gttcccttct cttccctgaa atatattagc acctgccagc caggccctca 420

ttttgcccag ccagtgtggg cagatcccac cgtggagaca tctgtagtgt gtatgtcctt 480

gtaacactct gttttcaggg actacaacct ttttccttct gtgaccagcc ccggattcag 540

gctgtac 547

-continued

<210> SEQ ID NO 35
<211> LENGTH: 476
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 35

ttaacactaa ttatcacgtc tgacaaatgt gtatgtgtgg tttcagttct gtgtacattt 60

taaaggataa tggtgaacat tttaatgggt ttcccttgcc ctttccatat ttaacctatt 120

tccacattct ctctcactca cattttctca gtgtgccctt ctcttatctg ccatgtccat 180

agccataatt ccaccatcat acagatcagg cagtgtttaa aatgatggta ggtagcacag 240

tggacagtct ttgatcatca tgtagaatat ggctatgaat caggaaagag attagaacat 300

ttaataatgt atgtacagct ggtgcttagt tttttttaa tctaaattta attaccttat 360

tggatatttg atatttggtt atttaatcac agtcatcttt aacagcttac actgattggt 420

gttttatctc ctgtgatcct ttgatggctt tttttgccta ccatttcaca gaggtt 476

<210> SEQ ID NO 36
<211> LENGTH: 445
<212> TYPE: DNA
<213> ORGANISM: Plasmodium chabaudi

<400> SEQUENCE: 36

cgctgggact gacgggcatg tccacctgta ctccatgctg caggcccctc ccttgacttc 60

gctgcagctc tccctcaagt atctgtttgc tgtgcgctgg tccccagtgc ggcccttggt 120

ttttgcagct gcctctggga aaggtgacgt gcagctgttt gatctccaga aaagctccca 180

gaaacccaca gttttgatca agcaaaccca ggatgaaagc cctgtctact gtctggagtt 240

caacagccag cagactcagc tcttggctgc gggcgatgcc cagggcacag tgaaggtgtg 300

gcagctgagc acagagttca cggaacaagg gccccgggaa gctgaggacc tggactgcct 360

ggcagcagag gtggcggcct gaggggtccc gggaggcggg tgcaagcctt cgctgtgccg 420

agccttgtgt ttctgacgca agcca 445

<210> SEQ ID NO 37
<211> LENGTH: 366
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 37

ccccaggcga tggacaatta tgatagctta tgtagctgaa tgtgatacat gccgaatgcc 60

acacgtaaac acttgactat aaaaactgcc cccctccttt ccaaataagt gcatttattg 120

cctctatgtg caactgacag atgaccgcaa taatgaatga gcagttagaa atacattatg 180

cttgagatgt cttaacctat tcccaaatgc cttctgtttt ccaaaggagt ggtcaagccc 240

ttgcccagag ctctctattc tggaagagcg gtccaggtgg ggccgggcac tggccactga 300

attatgccag ggcgcacttt ccactggagt tcactttcaa ttgcttctgt gcaataaaac 360

caagtg 366

<210> SEQ ID NO 38
<211> LENGTH: 524
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base

-continued

<222> LOCATION: (99)..(99)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (186)..(188)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 38

gggtaaataa catctgacag ctaatgagat atttttttcca tacaagataa aaagatttaa 60

tcaaaaaatt tcatatttga aatgaagtcc caaatctang ttcaagttca atagcttagc 120

cacataatac ggttgtgcga gcagagaatc tacctttcca cttctaagcc tgtttcttcc 180

tccatnnnat ggggataata ctttacaagg ttgttgtgag gcttagatga gatagagaat 240

tattccataa gataatcaag tgctacatta atgttatagt tagattaatc caagaactag 300

tcaccctact ttattagaga agagaaaagc taatgatttg atttgcagaa tatttaaggt 360

ttggatttct atgcagtttt tctaaataac catcacttac aaatatgtaa ccaaacgtaa 420

ttgttagtat atttaatgta aacttgtttt aacaactctt ctcaacattt tgtccaggtt 480

attcactgta accaaataaa tctcatgagt ctttagttga ttta 524

<210> SEQ ID NO 39
<211> LENGTH: 226
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (26)..(38)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (63)..(63)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (66)..(67)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (122)..(122)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 39

atcatcatga tcgctcgaca tcgatnnnnn nnnnnnnntt ttttttttttt ttttttttttt 60

ttnttnnaag tagaaaacaa aactttattt gatgaaatct ttttaaaagt tccagtatga 120

antaacaaaa tcaacaacct acaaatctct ttcagtcctt tgcatttcaa gcaaaatatt 180

ctcttcagaa aaatgaccat ttcataatat atatccccctt ctgtcg 226

<210> SEQ ID NO 40
<211> LENGTH: 407
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 40

aaaccaccaa agagagcctg gcccagacat ccagtaccat cactgagagc ctcatgggga 60

tcagcaggat gatggcccag caggtccagc agagcgagga ggccatgcag tctctagtca 120

cttcttcacg aacgatcctg gatgcaaatg aagaatttaa gtccatgtcg ggcaccatcc 180

agctgggccg gaagcttatc acaaaataca atcgccggga gctgacggac aagcttctca 240

tcttccttgc gctacgcctg tttcttgcta cggtcctcta tattgtgaaa aagcggctct 300

-continued

ttccattttt gtgagatccc aaaggtgcca gttctggccc tttcagctcc tgtttcagga 360

tctgtcctgg ttcctgagct ctaggctgct aagctgagcc acacacc 407

<210> SEQ ID NO 41
<211> LENGTH: 453
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (104)..(104)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (182)..(182)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (359)..(359)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 41

gaaactctgg aaatcacgtg tgtggggaga tggggacgct tcccatgttg tggggagctc 60

tgtggctgtg atggctgcag ttgccgtgcc tctgttggaa cgcnaagtgc ctgcaactca 120

cgtcaatcat agaattgtga cgcacagttg gcaaaatagt tctttatgct atttctcaaa 180

antttgagga caaacccaga ttgggattgg aatatgcact gtaaatcaaa ttttttcttat 240

ctacaaagac taatgtaaaa atgatttttt cttctgtgcc tgattaaatt aactgtggtt 300

tttaatataa atatttattg gtgtgctttg ggagaaaaat tatcttttct tgaaagaant 360

tatcaaagca aatttattat cttcacaagt taatgggaga atgtggtttt gattctgggt 420

gtttgaattg tgtaaacaca cagcttcctt gtg 453

<210> SEQ ID NO 42
<211> LENGTH: 434
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (33)..(33)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (42)..(42)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (75)..(75)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 42

atttcccttt tacattcatt atgcaaattc acnttctatt cntttctcac acactactag 60

ccagcctccc caaanaagga aaagggaaaa aagtaagaaa agaatggaac aaaagaaaaa 120

taagaaagca aacgaaagga acaaagaaac aggataaaga aaagagatca cagatttgag 180

aaagaaaaac aattcaattc agcaaattca ccaaaacaat gtgaatatat cctaaagtga 240

ttaaactcag aaatgatgtg aatttttcca gtttacacag tttgaccaaa aacagcatgg 300

ctttatgtgg tagcaaacca actgattctt gcttctactt tcataagtga ttttgcccac 360

atatcatccc actttaattg ttaatcagca aaactttcaa tgaaaaatca tccattttaa 420

ccaggatcac acca 434

-continued

<210> SEQ ID NO 43
<211> LENGTH: 488
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 43

tctttgatca ctgcctcttg atttttttcct ggatcattaa gaggcttgaa gaatactatg 60

tagttgaacc agaggagtag tgtatgtcac atcctcactt actccttaag ccctttctca 120

tggtcttggc cctaaaacat attttcaggg cttgtgaccc agtgatcagt ggtcaccctt 180

aaagtattac agatacgtgc ctgttttaca tgagaggtaa ctgtttatgt gtataagtca 240

tcttaataaa ataacatgaa atttattagc tgaattgggt agatactgct tttctaagtt 300

gaacctaact taagctgatg cagaaactga gtcagaaaag ttgctataat tttaaaatat 360

aagaagtaaa agtgaaatct tatgtagcat ctttatctca ttttggtttg tcagtataag 420

tttctgattt cctttaagct ctttactttt agaaacgtga atttacaatc ccttatccaa 480

aactgctg 488

<210> SEQ ID NO 44
<211> LENGTH: 169
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 44

gttagtggag ttttactgtt aatatcatca tgtccccctt tgtgtttact actgtctgaa 60

attactggga tgtagaagca tatttcagtc tgaaaattca gccagcttat tttggagaag 120

ttgtatcttg ttcttgggca tgttagcctt gtttttcatc ccaatttga 169

<210> SEQ ID NO 45
<211> LENGTH: 373
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 45

atggctcgta ctaaacagac agctcggaaa tccaccggcg gtaaagcgcc acgcaagcag 60

ctggctacca aggctgctcg caagagcgcg ccggctaccg gcggtgtgaa aaagcctcac 120

cgttaccgtc cgggtactgt ggctctgcgt gagatccgcc gctaccaaaa gtcgaccgag 180

ttgctgattc ggaagctgcc gttccagcgc ctggtgcgag aaatcgccca agacttcaag 240

accgatcttc gcttccagag ctctgcggta atggcgctgc aggaggcttg tgaggcctac 300

ttggtagggc tctttgagga cacaaacctt tgcgccatcc atgctaagcg agtgactatt 360

atgcccaaag aca 373

<210> SEQ ID NO 46
<211> LENGTH: 223
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 46

agagcaggcc aaccgagcga taagtaccgt tgagtttgtg acagctgctg ttggtgtctt 60

ctcattcctg ggcttgtttc caaaaggaca aagagcaaaa tactattaac aatttaccaa 120

agagatattg atattgaagg aatttgggag gaggaaaaga aacctggggt gaatacttat 180

-continued

tttcagtgca tcattactgt tccagattcc tatgatggat ggc 223

<210> SEQ ID NO 47
<211> LENGTH: 535
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (298)..(302)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (304)..(306)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 47

cagccagtgg ctgtggtcta cagaattgtt tcatataaaa tacgggtaga gtggtagagt 60

ttcaaaactt tcgtcataga tatctgggac ctttctcagg atctgtgttc acacagccaa 120

tagatttgga atcaggccta agagtacaca tggagggtaa atattaaagt gcgtattatg 180

tacatctaga atccatgtga cttgcagcct acctgtaatt tctatccatt gagcatgcat 240

ggatataccc aatagtacac acaaaataaa tgtttactta agagccattc taaaaaannn 300

nnannnaaat ggtttattgt aaatctgcct aaagattttt tgcatattat atatgtgaat 360

tttggttgta agttcataac ttacccaagg gtatagactc ataactcttt taaaacagtg 420

cttagtacaa tatcctgcca tctctgtaaa aacgctaatt gataaccgag tcatttacat 480

gttttcgaac acagaatagc tcttttctca gcatcattat tgctctttca gcatc 535

<210> SEQ ID NO 48
<211> LENGTH: 406
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 48

gattccagaa tctctacctt taaacactat gttaccactt acttctcttc aaattttatt 60

gagcattaga tgtttccagt atttagaagt caaatgcttc gtttttaata ggaacttaca 120

cagtctttta tgtttttttta tagccctcaa tgtcactgat gtggattctc ccaaactcga 180

tactttgttt gtttttatgt ccccataata agtctttaag aaaacagggc aagtgagctc 240

aaaatcaaaa gaaaacccac caacagtgaa tgcattcagg gctatttcca ggtctttctt 300

ttgaagaaag ataagactca gtccagagag cacatctgtg acacaccgtg cctcttgcct 360

ttggtgcgtg gcagtcatct ttggctcatg ctgtacatta ttctac 406

<210> SEQ ID NO 49
<211> LENGTH: 507
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 49

gtttcgattc tgttttgttc atctgttcga gcagaggggc agttgaagtc tcgtcctggt 60

ctctgccctg gcatggactg gcacagaggt gttctgtagt tgaataggaa gagcctgtct 120

aaaaaactac tgccccactt caaattgcag tgttctgtca cctaggcatc atctcttcct 180

gcccctagta tttgattaca aggaaccagg ggaaaaaaac tttcttagac acactggcac 240

caaggtaaga ggtggggctg cccaggcaaa gtcagtgaac atgaaaactc agacaaagca 300

-continued

gagatggaaa taatgcgcct cttgaggaga aaagcaataa tgaataaaag gactttccta 360

caataacttc actgaggact cacgttacca attttcatac ttactaaagg gattgtaaaa 420

aacaccccag cattttaggt gtcttggttc catttacagc actgaggtaa tctttctgct 480

gtttgttgtc ctgcttggtt gagtacc 507

<210> SEQ ID NO 50
<211> LENGTH: 301
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 50

ctcattattc ctttacatgc agaatagagg catttatgca aattgaactg cagtttttca 60

gcatatacac aatgtcttgt gcaacagaaa aacatgttgg ggaaatattc ctcagtggag 120

agtcgttctc atgctgacgg ggagaacgaa agtgacaggg gtttcctcat aagttttgta 180

tgaaatatct ctacaaacct caattagttc tactctacac tttcactatc atcaacactg 240

agactatcct gtctcaccta caaatgtgga aactttacat tgttcgattt ttcagcagac 300

t 301

<210> SEQ ID NO 51
<211> LENGTH: 515
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (183)..(183)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (462)..(463)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 51

gaacttaagt tcacacaccc ttgtactgca ggacggggaa tggaacctag gtcttcttat 60

ttttggttca gtgttaactc ccattctcta agcagactgg gcctgttatt caaactgcct 120

tcccataggt gcttccctgc ttctctcctc acccagagaa ggacttacaa acagcttatc 180

ttncagaggt tttgtgcctg atagttatgg aatgtgctgg tttgagcagg gaggatgtaa 240

ggggagggaa tgctaaaagg ctgtctactt agagtcaggt ttcctgggta agtccctgga 300

accccatccc cttccccttt cttgagaccc caggacttgc tccagtaact gccaccctgt 360

gcctttgctt cagggccatg ctggataagg agctggctgc ctctgtgaac atcctactca 420

aggcatcttc actgtgagtt ttgctgttgc cattggaggg gnngtggggg gagtgtgggg 480

agtgctaggg tcaggtcctg gctggtgtaa agaac 515

<210> SEQ ID NO 52
<211> LENGTH: 309
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 52

gactttcacc atcctgatat taaaactgtg caggtgtcca cagtagatgc ttttcaggga 60

gctgaaaagg agatcattat tctgtcctgt gtaaggacaa gacaagtagg attcattgat 120

tcagaaaaaa gaatgaatgt tgcattgact agaggaaaga ggcatttgtt gattgtggga 180

-continued

aatttagcct gtttgaggaa aaatcaactt tggggacgag tgatccaaca ctgcgaagga 240

agggaagatg gattgcaaca tgcaaaccag tatgaaccac agctgaacca tctccttaaa 300

gattatttt 309

<210> SEQ ID NO 53
<211> LENGTH: 529
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 53

gtaagactgg caaggccaag atcaccggcg gctatcggct cccggccaag tacgtcatcc 60

acacagtggg gcccatcgcc tacggggagc ccagcgccag ccaggctgcc gagctccgca 120

gctgctacct gagcagtctg gacctgctgc tggagcaccg gctccgctcg gtggcgttcc 180

cctgcatctc caccggcgtg tttggctacc cctgtgaggc ggccgccgag atcgtgctgg 240

ccacgctgcg agagtggctg gagcagcaca aggacaaggt ggaccggctg atcatctgcg 300

tgttcctcga gaaggacgag gacatctacc ggagccggct cccccactac ttccccgtgg 360

cctgaggctc ccgcagccca ccctgaccgg gactgacccg ccttcgggac cccgctccca 420

gctctgagag gtcgccaaag cctgcagcct ggcctgggcc tggccacccc ttctttccct 480

ccgcgccccg cccccgagga gcctaataaa gatctcgttg tggcaaaaa 529

<210> SEQ ID NO 54
<211> LENGTH: 456
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (49)..(49)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (296)..(296)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 54

gggccccaag agcagactag gaacgcaggg ggctgctgct gccaggacnc cacggagagc 60

cgggcacccg cctcacatgt ctcctgtctg gctccactga gttagccgtt tgagcccact 120

cctatctttt ggtggttagt gcatcttcag ctcttttctg caagacactg gaacattcct 180

aggctgtccc aaaaggagtt ccaccatagc ctttaaggtc cgagcagggc accaaggggt 240

tcacttttct cccgagccat tcagcttggg gtgcctgcgg gaggggcgga cagccnagcc 300

ggcttcccgg cggcggtacg agagcccaac aggagaggat tagctgtgcc aaggaacacg 360

ccactgctgc ctgtctactg cccgccttct ctccacttcc atttttgcct ttgtttttaa 420

cttgtgctct tgtgagttct tggtgtgttt ctttgt 456

<210> SEQ ID NO 55
<211> LENGTH: 448
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 55

acaggactat ttgccacgac atttcaaagg attccaagag agaatattgg tgtccatgct 60

gtgatgattc ctcagctcct ctcatctgat ctccgtcctg gcccccatga ctttctttgc 120

-continued

ggtagttagg gtgtggtatg tgccactgag gcccacacct attggcaatt tatagcactg 180

atctgtcatc aataccactt gctgtcttgg atgtgaagat gatttttcct gcagggattc 240

cctctacaaa attaaaaaca ctgggcatgt ggaaataata ttcacgcttt aaattgtctt 300

ttctattcac tacaccaggg gtccccgacc cctaggcaac agactgtggc cctagtgtag 360

tgaatagaaa agacaattta aagcatgaat attatttcct catgcccagt gtttttaatt 420

ttggtactgg tctgtggctt gttagaaa 448

<210> SEQ ID NO 56
<211> LENGTH: 546
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 56

agccgtcagc tgtctgctat gagctgcagc tctgcccact tacactcaag gaacttgcac 60

cacaagtggc tgcatagcca ctcaggaagg ccatctgcaa ccagcagccc cgaccagcct 120

tcccgcagcc acctggacga tgatggcatg tcagtgtaca cagacacgat ccaacagcgc 180

ctgcgtcaga ttgagtcagg ccaccagcag gaagtagaaa ctttgaagaa acaagtccag 240

gagctgaaga gtcgcctgga gagccagtac ctgaccagct ccctacactt taatggagac 300

tttggggatg aggtgatgac ccgttggctt cctgaccacc tggccgccca ctgctatgcg 360

tgcgacagtg ccttctggct tgccagcagg aagcaccact gcaggaattg tgggaacgta 420

ttctgctcca gttgttgtaa ccagaaggtt ccagttccca gccagcagct ctttgaaccc 480

agtcgagtat gcaagtcttg ctatagcagc ctacatccca caagctccag cattgacctt 540

gaactg 546

<210> SEQ ID NO 57
<211> LENGTH: 508
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (425)..(425)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 57

aaatgtatat atatcctagt cccctaacca aatcctgacc tattgggata cttataaaaa 60

tttaagtaag tgggatacac aaagaataat aactattaac ttttcattat tagcaaaaac 120

ctaagggatt taaactaatt gaaactgtat ttgattggac ttaatttttt atgtttattt 180

agaagataaa gatttaaaga agacctttac aataaagaga agaaatatcg aagtcattaa 240

aataaggaga cttactttta tgacattcta atactaaaaa atatagaaat atttccttaa 300

ttctagagaa actagtttta ctaatttttt acaacttcaa taataccatc actgacactt 360

acctttatta attagcttct agaaaatagc tgctaattag gttaatgaac attttacctt 420

agtgnaaaaa aattaattaa atatgattac aaagttgcac agcataacta ctgagaggaa 480

agtgattgat ctgtttgtaa ttacttgt 508

<210> SEQ ID NO 58
<211> LENGTH: 407
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 58

gtaagcaagg tctttgtggg gcagttgtcg acctcaccaa gacatctgaa catattcagt 60

tttcagtaga agaacttaaa aatttgttca cattacatga aagttcagat tgtgttactc 120

atgatctgct tgactgtgag tgtacaggag aagaagttca tacaggtgat tcgttggaaa 180

aattcattgt ctctagagat tgtcagcttg gtccacatca ccagaaatct aactccctga 240

aacctctttc tatgtcccag ctgaagcaat ggaaacattt ttctggagat catttaaatc 300

ttacagatcc ttttcttgaa agaataacag aaaatgtgtc attcattttt cagaatataa 360

ccactcaagc tactggcaca tagtgaaaga ttacttctga cattcca 407

<210> SEQ ID NO 59
<211> LENGTH: 445
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 59

acaataatta gatctttttc caagttaatt gggttttccc ttctcccagt cataggtggt 60

ttttatcatc aagacagact gatattttgt caggatattt tcttttacag tgtttgatgt 120

gcataatgcc agagttattt ttttattatt cattttctct ctttttgttc aatatgagat 180

tcaggatcat atttgtttaa aaggtaacac atagagatgt atgtatatat tttgttataa 240

gacatacaaa ataattttaa gagggataaa ggtgaaaata tcagattctg gaaattttaa 300

gtatctaaac tttatacttg tatgatttac cataaacata ccaaaacatt tttctgaaaa 360

tttactgtcg gtctctgaca tgaaaccgta ttttgtcagt agttgaccaa gcagttttat 420

gagaactctt ctatgcaatg atgca 445

<210> SEQ ID NO 60
<211> LENGTH: 421
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 60

cttgctgcct gtctggaaag tgagaatcaa aatgctcaga ggattggagc agctgccctt 60

gggctctgat ttacaattat cagaaggcaa aaacagcttt gaaaagccca tcagtaaaaa 120

gaagagtgga tgaagcatac tccttagcaa agaaaacttt cccaaactca gaagcaaacc 180

ctctaaatgc ctattatttg aaatgtcttg aaaacctcgt gcagctcctt aattcttcct 240

gagtgccatg ggatgctaca ccttgaagct gacagtcatc aacaggggag ctaaagttga 300

agccagctgt gtgtagcagc tgttacctga agacgtgcta cctctctaca aagtgttgat 360

ccccttcttt cccatgagag agagaactgg tgatactcca acaccgtcca gttgtggcag 420

c 421

<210> SEQ ID NO 61
<211> LENGTH: 352
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 61

caagagtaga tgcagttaag gaagaaaatc tgaagctaaa atcagaaaac caagttcttg 60

gacaatatat agaaaatctc atgtcagctt ctagtgtttt tcaaacaact gacacaaaaa 120

-continued

gcaaaagaaa gtaagggatt gacacccttc tgttttatgg aattgctgct gatcattttt 180

tctttaaaac ttggatagat tccaaaagtt acagtacctt tgtggcttca ttgaatattt 240

atgaagataa tgtcagatgt agacaaaaat aacacaataa caggagactt ccataagttt 300

gtgtattatg ttagtctatg aaaacgtgca aatgtattgt agagacttta tg 352

<210> SEQ ID NO 62
<211> LENGTH: 461
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 62

agctgaagac tctgaccaag cagatgaagg aggagaccga ggaatggagg cggttccagg 60

cggatctgca gaccgcagtg gtggtggcca atgacatcaa gtgtgaggcc cagcaggagc 120

tgcgcaccgt gaagaggaaa ctgctggagg aggaggagaa gaatgcccgg ttgcagaagg 180

agctggggga tgtgcagggc cacggcaggg tggtcaccag cagagccgcc cctccctccc 240

tgggctctgt cagctagcag agcatttggt ggaagaaaga cagcccagct cttgccatga 300

ttgggagccg cagcccatct ctagatgaaa gggggaatgt gtagaggaga aattgcctct 360

ttataaagag cccagttgtc tccttgtgac attctctgtt ctcagagtca ttgccgtcga 420

gtctctgctt tttgtccaca ttttgggatc agcttactgc a 461

<210> SEQ ID NO 63
<211> LENGTH: 349
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (143)..(143)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (315)..(315)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (319)..(319)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 63

gaagagtgct taaggttgaa gtacaatggc acaatctcgg ctcacctcga cctctgcttc 60

ctgtattcaa gtgattcttc tgcctcagcc tcccgagtag ctgggattac aggcatgcgc 120

catgacactt ggctaatttt ttngtgtgtt tttagtagag atggaatttt gccatgttgg 180

ccaggctggt ctcaaactcc tgacctcaag tgatccaccc acctcgacct cccaaagtgc 240

tggtattata ggcatgagcc accatgcctg gcctcattta tttttaaata gctgcagtaa 300

tcccggcttt agatnaaanc acatgaacta ataatatcac tagtgttca 349

<210> SEQ ID NO 64
<211> LENGTH: 196
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 64

agagctgagt catcctagag caaacctctg gagtggagag cgaactactt cattcccctc 60

ccttagcctg ggccagagag actccagctc tgccttctcc agccaaaaaa tcaaaggcag 120

-continued

atgggagaac agccttcagc tttggataac gatgaaatat ctggcaccac tgatgaatat 180

taaactttct ataacc 196

<210> SEQ ID NO 65
<211> LENGTH: 387
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (74)..(74)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (99)..(99)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (112)..(112)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (128)..(128)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 65

gagcacatat cttacaaaac accaaaaaat tcatagtgaa gagaaatcaa atatacatac 60

tgagtgtggg gaanccatta gacaaaactc ttctttttna caacaataaa ancctcacac 120

tggagagntt ctctgaatgc cttaagaatt tggttaatat ggagaccctt cccagggaaa 180

cagaaggagg atcgtgaaaa ctgttgacta cttagaatga tcacatggtt tagtggagag 240

agcatgattc tgggttttaa aagtcatgga tctcaatctc agctcctatt actaactaga 300

tcttttactt tggggtaagt cacttcatat ctttaggcct taatttcctc atctgaaaaa 360

ctggaaggcc tgacttgttg agcttta 387

<210> SEQ ID NO 66
<211> LENGTH: 193
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 66

ggcactgtgt aatcattcct tgaagtagtt ggagatggtg ctggtatgcc actgaatgag 60

gtctgagcag gttttcttca catctgaggg gacagtgcca gccagtcaac ttttggggtg 120

gggctgaagt ctgctgaaaa tctgcagttt tacatgtttc atgggacatt cttctgtgca 180

ataaagtttg aga 193

<210> SEQ ID NO 67
<211> LENGTH: 480
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 67

tactggtgac acttcatggc tgcgacccag aatgaactta atgcacacag ggacgcaggg 60

tgtcactggt cctgggcctt tgtccatgac taggtggtca gcaggacttc tgcagctgac 120

tgtgcaatgg ctaaatgaaa agaaggccac agactaacct ccactttcct gtcttcaaaa 180

ttctagtgac actgggaatg ctataggacc tcctactatt ctcttaaggt cctaggaaag 240

tttcaggaac tagggaaaag actgggtact gaggctgtgt ccccagatgt ctgcttccga 300

-continued

agcagccgcg tcatgacggg tttctgctga ggaagtggtg ttggcagggc cccatatgcc 360

ctctcgggtt gtcaggggtg ggagacaggc tgtatggggg tccttcatgt gcagatggaa 420

cagcatcgcc tcacagctgt gcagacgaac agatgtggtc tactgccacg aacaatgcgg 480

<210> SEQ ID NO 68
<211> LENGTH: 427
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (40)..(40)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (91)..(91)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (198)..(198)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (211)..(211)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (315)..(315)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (321)..(321)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (324)..(330)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 68

cctcatggtg cctcggagag tggggagcat attgggctgn ggtaagcact agacccaagt 60

agactggaca caaagggctc gcccagggcc ntggcgccac ccccaccccc tcccaccagc 120

tgctgctagc ctctgtggtt gtacatccca cttgccccca cacggagact gactctaaaa 180

cccttcatcc aatggtgnta acccccggct ntcccctgcc ccacctcacc cacccagaga 240

agcacagacc ccgccagggg caggggccca ccgcacaccc ttgtcccggg cctgtctggg 300

actggccttc ccggntcagc nagnnnnnnn cagaagggac acaaagaggg atggaagaaa 360

agaacaaaga gaaactgttc ctcccacccc cttccctgat gccaggggca ccagactgat 420

tctgagg 427

<210> SEQ ID NO 69
<211> LENGTH: 471
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 69

gtcatttacc atggttttcc cactgaaggc tttaactttt ctgataaaat aatattttaa 60

attttcaaaa acccattcct gaggagaact acttctagca ttccttttca tgatgtgctt 120

ttgtgcagta agtagcattt tcggctactt aactttacat tcctcttatt tttcagtttc 180

cagtcaagat tataaaaagc aaatgattga tataatttga tattcataga gttgtgccta 240

cctttaatgg aaaaatacat gtcagatact tagatgttta ttgatatgag actatgtggt 300

-continued

taaaaaaccc aagtatgtcc atgtgtttct tataaggtac acttgaaact agtgagtgtt 360

tgtcacattt cactttcatg gtatataaaa tgcagtttgc atatataact tgaatatctg 420

gtactagttt tttcacgcct gcaatcttgg agtctaggtt gccttgtctc t 471

<210> SEQ ID NO 70
<211> LENGTH: 323
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 70

aaatccaagt tcctcttact ggctttcaag gatcctcctt taacttcctg ttgccttgtc 60

tcatcagcaa gatcataagt acctacaggt caagcactgt ctctcccttc tctttagctt 120

ttcccttagg atctagcaca ttacccagca aaatgtgagt agcaaggctg aaatgacatc 180

tcaataactt caccaatgat tgtaactcag catcccttct ccatcccagc tgaaagcctg 240

cttcaccatc ctgcaagaga tgtttttttct ttttgttagc atccattccc ctttctaatg 300

cagctcccaa tgcataatgt gtg 323

<210> SEQ ID NO 71
<211> LENGTH: 463
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (175)..(175)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (182)..(182)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (186)..(186)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (240)..(240)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 71

ggtcatcgaa tctgcatgca tccctcatac atctggagac ttcgtgaagg ttccagagtt 60

actgactgag atttctgagc ttttttcccc ttttctgttt ggtttcagag tggagagcag 120

cccaaaaata tgcaggtaac tgaagccagg aaactgattt gtgttttggt ttggnccgga 180

tncttnaaac agaagggagg tggagagatc tgagattaga ggacggggct ttataggagn 240

ccaagtatgg ggcctgcaca cacaagacac acacgcacac ttgcaaacac gccacacgac 300

acatatgcct gcatgtgtat gcacacacat gcacacgtga gctcccaaac acatcgctcc 360

ttggggttac actaggtttg tttccatctg gcttgaggct atttgcaggc gagagtgcag 420

agtctgtaat gaacctccca gattctctga cgaaggggtc ccc 463

<210> SEQ ID NO 72
<211> LENGTH: 525
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 72

ttcctgtgtg cttagtcatc gggtccgaca caggctggac tgatctgggg agccgcgaag 60

-continued

ggcctgcctt cacaaaggga cgtaacgcaa gtactgcggg cagtgtttga atatggccct 120

gaacaatgtg tccctgtcct ccggtgatca gaggagcagg gtggcctacc gctcttccca 180

tggcgacctc agaccgcggg cgtcagcgtt ggcgatggtc tccggagacg gcttcctcgt 240

ttccaggcct gaagcgattc atctaggacc tcggcaggcg gtgcgaccaa gcgttcgggc 300

cgagagccgt cgagtggatg gtggcggccg gagcccaagg gaaccagatg gccggggccg 360

gagccgccaa gccagattct caccttaccc aatccctgcc gttgaacccg atctcctaag 420

aagtgtgctg caacagcgtt tgattgcatt aggaggtgtt atcgcagctc gaatttcagt 480

ttaaacgaac acctttcctc tggccctcac ttagcttgtg aacag 525

<210> SEQ ID NO 73
<211> LENGTH: 156
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 73

gaaatgcatg ggacaggtga ctgacatgac tgcatcgtgg tttagatgta tagataacac 60

ggggaggtgc tttacatttt aagactttgt tcataattct tttatttatg gtttctctga 120

atcattcttt tggaacattc taaaagagcc agagga 156

<210> SEQ ID NO 74
<211> LENGTH: 501
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 74

cgggtgaaga gtgtgccggg gcggcggctg gctgatgggc gcacactgga cgggcgggct 60

gggctggccg acgttgccca catactcaat ggccttgctg agcagctgtg gcaccaggac 120

caggtggcgg ctggcctgct tcccaacccc ccagagagtg ctcctgaatg agtcacgagt 180

ggttgcctgt gatcccaccc ccaaccctca ggtctcgaca tagggctgga ggctggggca 240

ggaacatgga tcctatctgg aggactggcc agcatggcct gatcagggag gatgtggcca 300

gagaaggccc acccgcgagc agcgctttcc ttgcagaatt catggcaggg aggtggggac 360

caaggccctg agctcgaaca tctcccgtgg cctttccccc tttggcagca ccgatggagg 420

atgactggga gaggggtgc ctctcaagtt acttcaatca agaacctgta ttggttgagg 480

tgacaccatc tgttgtaaca g 501

<210> SEQ ID NO 75
<211> LENGTH: 494
<212> TYPE: DNA
<213> ORGANISM: Xenopus laevis

<400> SEQUENCE: 75

tgaatgacct tcctagagca cttgagcaga ttcattatat tttaaaacca gatggagtgt 60

ttatcggtgc aatgtttgga ggcgacacac tctatgaact tcggtgttcc ttacagttag 120

cggaaacgga aagggaagga ggattttctc cacacatttc tcctttcact gctgtcaatg 180

acctgggaca tctgcttggg agagctggct ttaatactct gactgtggac actgatgaaa 240

ttcaagttaa ctatcctgga atgtttgaat tgatggaaga tttacaaggt atgggtgaga 300

gtaactgtgc ttggaataga aaagccctgc tgcatcgaga cacaatgctg gcagctgcgg 360

-continued

cagtgtacag agaaatgtac agaaatgaag atggttcagt acctgctaca taccagatct 420

attacatgat aggatggaaa tatcatgagt cacaggcaag accagctgaa agaggttccg 480

caactgtgtc attt 494

<210> SEQ ID NO 76
<211> LENGTH: 369
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 76

acatatgtac agtatcagta gggaaaatgt aaaaagatgt tgttttcttt tgtcatttaa 60

ttaggccatc tgtcctgttt taaagaaata gttaataatt caacacttta tataacaaat 120

attaactaat acccatattt ataaaacatt tttcagattt aaaagattgt taatacttat 180

aaacttagtg ttattcttag aaaaccccat caaatttaaa tgtgatttac acagtgacta 240

ggaacatttg tatttattgt ttcttctctg cacttttcat catctgataa atacaagagc 300

tcaagtaact gtcttttctt caagatggct tctatacttg aaatcagtta atacaatagt 360

ttttccagt 369

<210> SEQ ID NO 77
<211> LENGTH: 549
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 77

gaggaagtca ctttactact ctaagatatc cctaaggaat ttttttttt aatttagtgt 60

gactaaggct ttatttatgt ttgtgaaact gttaaggtcc tttctaaatt cctccattgt 120

gagataagga cagtgtcaaa gtgataaagc ttaacacttg acctaaactt ctattttctt 180

aaggaagaag agtattaaat atatactgac tcctagaaat ctatttatta aaaaaagaca 240

tgaaaacttg ctgtacatag gctagctatt tctaaatatt ttaaattagc ttttctaaaa 300

aaaaaatcca gcctcataaa gtagattaga aaactagatt gctagtttat tttgttatca 360

gatatgtgaa tctcttctcc ctttgaagaa actatacatt tattgttacg gtatgaagtc 420

ttctgtatag tttgtttttа aactaatatt tgtttcagta ttttgtctga aaagaaaaca 480

ccactaattg tgtacatatg tattatataa acttaacctt ttaatactgt ttatttttag 540

cccattgtt 549

<210> SEQ ID NO 78
<211> LENGTH: 515
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 78

gttttaattc aacccagcca tgcaatgcca aataatagaa ttgctcccta ccagctgaac 60

agggaggagt ctgtgcagtt tctgacactt gttgttgaac atggctaaat acaatgggta 120

tcgctgagac taagttgtag aaattaacaa atgtgctgct tggttaaaat ggctacactc 180

atctgactca ttctttattc tattttagtt ggtttgtatc ttgcctaagg tgcgtagtcc 240

aactcttggt attaccctcc taatagtcat actagtagtc atactccctg gtgtagtgta 300

ttctctaaaa gctttaaatg tctgcatgca gccagccatc aaatagtgaa tggtctctct 360

-continued

ttggctggaa ttacaaaact cagagaaatg tgtcatcagg agaacatcat aacccatgaa 420

ggataaaagc cccaaatggt ggtaactgat aatagcacta atgctttaag atttggtcac 480

actctcacct aggtgagcgc attgagccag tggtg 515

<210> SEQ ID NO 79
<211> LENGTH: 518
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 79

tgatttttgt gaccctgtcc atcattgtta ccgtgtttgt cattaacgtt caccacagat 60

cttcttccac gtaccacccc atggcccccc gggttaagag gctctttctg cagaaacttc 120

caaaattact ttgcatgaaa gatcatgtgg atcgctactc atccccagag aaagaggaga 180

gtcaaccagt agtgaaaggc aaagtcctcg aaaaaaagaa acagaaacag cttagtgatg 240

gagaaaaagt tctagttgct tttttggaaa aagctgctga ttccattaga tacatttcga 300

gacatgtgaa gaaagaacat tttatcagcc aggtagtaca agactggaaa tttgtagctc 360

aagttcttga ccgaatcttc ctgtggctct ttctgatagt gtcagtaaca ggctcggttc 420

tgatttttac ccctgctttg aagatgtggc tacatagtta ccattaggaa tttaaaagac 480

ataagactaa attacacctt agacctgaca tctggcta 518

<210> SEQ ID NO 80
<211> LENGTH: 137
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 80

cagacacata tacaggaaag accatgcgaa gaagacacag agaggaaacg gccatctaca 60

atccaaggag agaggcctca gaacaagcca accctgcaga taccttgatc taggcatcca 120

gaattgtatg aaaatac 137

<210> SEQ ID NO 81
<211> LENGTH: 443
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 81

agacaggagg gccacagcgt gtggaagtaa agactttgga gctagagatg ccttttccag 60

caatgattat tgacttcacc acaccccttg cctggcctgg cctgaggctc agcagtgcat 120

gacttctcgt agataacttc acagtcatcc agtcccaaca cctgctcttg cctggtagga 180

acaggcgaag tgtcagccct caatgttggg tacttagacc caaaccaata aatggtgagt 240

tttgaacaag aactaccatc atgcaggctt cttgcccagc tgaccactgg ccccggggtg 300

cctgcctggc tggtcttcat cacctgaggc caccaggctc aagccactgc tgttgcatta 360

cacccatccc tttgcaaaat ccctatggag cctgtcacca ctcccctccc tatatacccc 420

caccccacaa agattttctt cag 443

<210> SEQ ID NO 82
<211> LENGTH: 545
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

-continued

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (99)..(99)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (112)..(112)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (126)..(126)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (133)..(133)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (163)..(163)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (198)..(198)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (200)..(200)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (224)..(224)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (231)..(231)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (259)..(259)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (283)..(283)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (308)..(308)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (338)..(338)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (342)..(342)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (345)..(345)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (350)..(350)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (356)..(356)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 82

gtgacgttgg taggaaagca tttctttaca tggaggttta ttttgtggga cattacctcc 60

tctggatgtt acttcctcag tttacaagta gtgtaaatnc tcattgattc tnttatgaat 120

tgtaanggat ttnctcttag cttttgagaa tttagaatct ganatttgaa taaaaagtaa 180

-continued

atatattcag tataattntn caaaatgctc tagtttcaag atanttaaaa nattatgtgg 240

aatttacata attaattcna aatatagttt gtatttagtt ccnttatcaa ataatgcaaa 300

tagttggnag atacctcaat ttcttttgag tgttaagnaa gnagncaagn aaaggnagtg 360

aagtttttgc aacacattgt gtctttattt ggtctgccta tgtttttatc acattgctta 420

taaaactttt aaaatccttg tttgtataaa aagtttcttt agttaaataa aagtgtgtgt 480

attaattagt gtgccttctg gacaaattaa gaaatatttt ttctatattt caatgcggtt 540

gtatt 545

<210> SEQ ID NO 83
<211> LENGTH: 424
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 83

tcagatattt ccagagaatg accacaacct cttcagtaga aggtaaacag aatctggtga 60

ttatgggtaa gaagacctgg ttctccattc ctgagaagaa tcgaccttta aagggtagaa 120

ttaatttagt tctcagcaga gaactcaagg aacctccaca aggagctcat tttctttcca 180

gaagtctaga tgatgcctta aaacttactg aacaaccaga attagcaaat aaagtagaca 240

tggtctggat agttggtggc agttctgttt ataaggaagc catgaatcac ccaggccatc 300

ttaaactatt tgtgacaagg atcatgcaag actttgaaag tgacacgttt tttccagaaa 360

ttgatttgga gaaatataaa cttctgccag aatacccagg tgttctctct gatgtccagg 420

agga 424

<210> SEQ ID NO 84
<211> LENGTH: 556
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (55)..(55)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (142)..(142)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (146)..(148)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (152)..(152)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (171)..(171)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (182)..(182)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (388)..(388)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (397)..(397)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

-continued

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (472)..(472)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (478)..(478)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (491)..(491)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (501)..(501)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (509)..(515)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (517)..(517)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 84

ataccactta gtatagttcg ctactatttt gtggcctaca tgacaggtgt caagnttttt 60

ttgaatcaat ttttaaaaca tgccattgtg tttcaggctc gcgggattga tgtgcaacaa 120

gtgtctttgg ttataaatta tnatcnnnct ancagtcgtt gaaaactata nttcacaggt 180

gnagaagcca gcatcttggc tgtattgaaa aaacttcata cgtttttcta ctgtgatttg 240

tatgaaaggt aacatcaaat caaggaatag attcagtaaa gtcagtagtg ttcagtaaga 300

tgatgtaatt aaatttgtac tagggaaggt tgatgagaac aaagtgggaa aacttgtaaa 360

cattgcccag attgtggaca tagggttntt ttccacnaat tgttggtctt accttatgct 420

tgagctttta gtgatgttct tgtgtccatg tgtttttctt ggtgattttt tnctatangt 480

tgggattttc nttggtgtcg nctggtagnn nnnnnantga accctggttt agttatagtg 540

gctttatccc taaata 556

<210> SEQ ID NO 85
<211> LENGTH: 367
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 85

ctactcgggc gctcccagaa ggagccacct ctcagtgcct cacctccccc tgcctcccag 60

cctccgcaga tgaggttcct gccccttcct cctcgtaacc aaaaccctca ctgctcccag 120

gacggtctta tttataaacc agatacatgt tcttagtctg gtcccagacc aaggagctgg 180

tcagacggcc ctttctaatc ctacatgttg agcttatgta aaaaatgttg tttcctcctg 240

tttttggttc ctttcttacc cacaaaccat tactacttga aacttaaaaa actcgccaag 300

tgtaaaggct aaagagaagc agtttgacgg accttgtgat ttgtactgtt tgctgcggag 360

ctattta 367

<210> SEQ ID NO 86
<211> LENGTH: 401
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 86

-continued

gagaacacct acctcttatt ggaaaagttg gcctctcgtg gaaaactgat atttttgatg 60

gctgtataaa gagttgttcc atgatggacg taactctttt ggatgaacat gggaaaccct 120

tttggtgttt cagttccccg gtgtgcctga gatcgcctgc cacaccctct gacagctcta 180

gcttcttggg acagacatac aacgtggact acgttgatgc ggaaggaaga gtgcacgtgg 240

agctggtgtg gatcagagag accgaagaat accttattgt caacctggtc ctttatctta 300

gtatcgcaaa aatcaaccat tggtttggga ctgaatatta gcagtaggtg gcaaattatt 360

gttgttattt agttgtttat ttttgactgg ctttgttctt g 401

<210> SEQ ID NO 87
<211> LENGTH: 407
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 87

gccaggcttt gtggtatgta cctttagtcc cagctactct ggaggctgag gcaggaggat 60

cacttaagcc ttggaggtca agaatgcagt gagccattat catgccactg tgtgaccaga 120

aaccagatgt agccatttca agcataaaac atgatatttt tgttttcctt ggactgaaac 180

atagtctggg tcctcaacgt tgccggtgat gatggttgaa catcatgttt tttataaacc 240

ttaatttctc atttaatagg aagaaaatct caggagagcc aaaagggagg acctgaaggt 300

cagcatccac caaatggaga tggagaggat ccgctacgtc ctcagcagct acttgcggtg 360

tcgcctcatg aaggtttgac gtggagatac ctcaaagtct ccgacct 407

<210> SEQ ID NO 88
<211> LENGTH: 439
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (26)..(26)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (236)..(236)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (243)..(243)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 88

aaaacactac tgcaatcacg tctttngtta tgctagtatc agtcagatgc acttagagtg 60

aagaaacact gtaattacag cacacagatt gcaagtattg cgtaccaagt gatacaactc 120

gaaatgcagt tctcatcttc ctgttttgag aaatgattat tttatcacgc atcagagcct 180

tcgtgctttg attatcttgt atgttaacaa ttctagaaaa cattcatgaa ttcacnaaaa 240

tangttacta tggcagggga acattttgta cacatttaag tatataaaaa tactaaaata 300

tgtaatttta taacaaagtc acgggtatct ttaggttcag ggaactagac taggtcattc 360

gtgtaaatgg actggtagtt acagtcttag gttaaagtat tctaatgaag tatgggaact 420

aaattgctgg ttttctaag 439

<210> SEQ ID NO 89
<211> LENGTH: 440

-continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (90)..(93)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (95)..(96)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (98)..(101)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (103)..(103)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (105)..(106)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (109)..(109)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 89

ttctcttgcc ctttacaaat tagttttctc acttaaagct atttttgttt tttgcttttt 60

cactagtgaa aatgttactt cccccatgan nnnanntnnn ntntnntana tctacctgtg 120

aattttgcta tatttctttg ttggtttggc tttacaaata tgtagctgtc ttctcacatg 180

ttacctgctg taaaaccatt catttgaatc ttaatagctt tcacgtttac agtaacaaat 240

gaatttccga gaaatcaagt aagttgccca agatccaaca tatataataa acatcagaac 300

tagaacttga attctgttct ttcggttgtt tccaacatgg actaacacat tttatcaaga 360

atgttttcaa tattcaaata aggactcgaa aaaataggct tacatagtaa cttttatcca 420

tcaacttacc tatcgatgct 440

<210> SEQ ID NO 90
<211> LENGTH: 485
<212> TYPE: DNA
<213> ORGANISM: Tribolium castaneum

<400> SEQUENCE: 90

agggagatga cactggagca ccccacagcc cacaggaaag agaccagatg gtcagaatgg 60

cccttaaaca catgggcagc atccaggcac caactggaga cacagccaga agggccatca 120

tgggcttttt agaagagatc ctggccgttt gttttgactc atctggatca caaggggcac 180

tcccagggtt aacaaatcag tgaagatccc accatacttt ctagatgtcg aaggcggcag 240

taggaagacc tgagcttgag cataagatct gtgggatttc atcttagggg cagaaacaat 300

ccgttcacta tttatttaga atgacttagc agccatttaa attttcacag agggctcaac 360

caccttgga gtgactccat agcactggcc atggtcaggg ttgttggaac atctgacctg 420

tgcatccagg agccgaggag tcaggttgta atacaggcca agcagacggg ctttgagggc 480

attta 485

<210> SEQ ID NO 91
<211> LENGTH: 320
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

-continued

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (164)..(164)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (209)..(209)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (291)..(291)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 91

aaactaaaca tcatatgttc tcatatgtcc ctaagctatg aggatgcaaa ggcataagaa 60

tgatacaatg gactttgggg actttcaggg aaagggtgag aagggcgtaa gggataaaag 120

actacaaatt gggttcagta tatactgctc gggtgatggg tgcnccaaaa tcttaaaaat 180

cgccaaagaa cttatgtaac taaataccnc ctgttcccca aaaaactatg gaaattaaaa 240

attaaaaaat aagtataatt tctgctttag cgatattaac tattcagtac ncaataagtg 300

agtttagcaa ttcagtgatt 320

<210> SEQ ID NO 92
<211> LENGTH: 478
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 92

catactgcat gctcaggacc catagatgaa ctattagaca tgaaatctga ggaaggtgct 60

tgcctgggac cagtggcagg gaccccggaa cctgaaggtg ctgacaaaga tgacctgctg 120

ctgttgagtg agatcttcaa tgcttcctcc ttggaagagg gcgagttcag caaagagtgg 180

gccgctgtgt ttggagacgg ccaagtgaag gagccagtgc ccactatggc cctgggagag 240

ccagacccca aggcccagac aggctcaggt ttccttcctt cgcagctttt agaccaaaat 300

atgaaagact tacaggcctc gctacaagaa cctgctaagg ctgcctcaga cctgactgcc 360

tggttcagcc tcttcgctga cctcgaccca ctctcaaatc ctgatgctgt tgggaaaacc 420

gataaagaac acgaattgct caatgcatga atctgtaccc ttcggagggc actcacat 478

<210> SEQ ID NO 93
<211> LENGTH: 125
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 93

tctggatcca aagacgcttc ggccaatgca gggattttac tcatctctgg cctccagtct 60

gaggatgagg ctgactatta ctgtatgatt tggcgcggca ccgctgtggt atttggcgga 120

gggac 125

<210> SEQ ID NO 94
<211> LENGTH: 280
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 94

gaagatctta ttcaatctat gcatattgat gctactttat atacggaaag tgatgttcac 60

cccagttgca aagtaacagc aatgaagtgc tttctcttgg agttacaagt tatttcactt 120

-continued

gagtccggag atgcaagtat tcatgataca gtagaaaatc tgatcatcct agcaaacaac 180

agtttgtctt ctaatgggaa tgtaacagaa tctggatgca aagaatgtga ggaactggag 240

gaaaaaaata ttaaagaatt tttgcagagt tttgtacata 280

<210> SEQ ID NO 95
<211> LENGTH: 550
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (26)..(30)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (35)..(35)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (37)..(37)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (41)..(41)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (51)..(52)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (55)..(56)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (59)..(59)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (62)..(63)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (66)..(68)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (85)..(85)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (98)..(98)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (112)..(112)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (119)..(119)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (127)..(127)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (129)..(129)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (133)..(141)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

-continued

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (143)..(143)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (145)..(148)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (157)..(171)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (476)..(478)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 95

gctgcccgca gaagcgcaca tgtgcnnnnn ggctncnggt ngggacccct nncanncant 60

gnnccnnncg ctttcctggg agaangtcac tgtggacnag gttctcgagt tnccattgnt 120

gaaaggncnc ccnnnnnnnn ntntnnnnca ggtgaannnn nnnnnnnnnn ngcagctgca 180

ggagtgcttg tgggctgact ccctggggcc tctcaaagac gatgtgctac tgttggagac 240

acgaggcccc caggacaaca gatccctctg tgccttggaa cccagtggct gtacttcact 300

acccagcaaa gcctccacga gggcagctcg ccttggagag tacttactac aagacctgca 360

gtcaggccag tgtctgcagc tatgggacga tgacttggga gcgctatggg cctgccccat 420

ggacaaatac atccacaagc gctgggccct cgtgtggctg gcctgcctac tctttnnngc 480

tgcgctttcc ctcatcctcc ttctcaaaaa ggatcacgcg aaagggtggc tgaggctctt 540

gaaacaggac 550

<210> SEQ ID NO 96
<211> LENGTH: 409
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (53)..(53)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 96

tccttcttac tctgcagcaa catggaggag agttttgtgt agtgagtgtg ggngaagaaa 60

tacatttggc tgttctcaca ccccctctga ctatgcacca gtgaacacat ctgagtacat 120

accagctctc ctcatcttct tatttatact taacttattt ttgtgtgaaa taaatggagg 180

acgaaatctt agagcaacat catcaaacag tctttggtcc ttgagaatct tctttgtgtt 240

ttatttttg atttctgtag cttttcagtt gcagatgttg aaattcgtaa tgacaaatat 300

gacaaattgt catgggtgat tccacttcat cttatttttt ctactctcac tatacaatct 360

tgcctcattt tttaaaactt tggaaccaga ggatttcaac tgcctagca 409

<210> SEQ ID NO 97
<211> LENGTH: 489
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (153)..(153)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base

-continued

<222> LOCATION: (200)..(200)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (389)..(389)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 97

gagactacag cagtgttacc tgtgcaaata caacttacta cttctgttac cttgaacttg 60

gaaaaaaaca gtgctctacc gaatgatgct gcttcaatgt cagggaaaac atctctaatt 120

tgtacacaag aagttgagaa gttgaatgag gcntttgaca ttttgctagc ttttttcatc 180

ttagcttgtg ttttaatcan ttttttgatc tacaaagttg ttcagtttaa acaaaaacta 240

aaggcatcag aaaactcaag ggaaaatага cttgaatact acagctttta tcagtcagca 300

aggtataatg taactgcctc aatttgtaac acttccccaa attctctaga aagtcctggc 360

ttggagcaga ttcgacttca taaacaaant gttcctgaaa atgaggcaca ggtcattctt 420

tttgaacatt ctgctttata actcaactaa atattgtcta taagaaactt cagtgccatg 480

gacatgatt 489

<210> SEQ ID NO 98
<211> LENGTH: 380
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 98

gatcattggt atgtcaatct cttgatgaaa aatcagtacc tgaatatgtc tttttgtttt 60

tttaagagac agggtcttgc tatgttgccc aggcaggatt tgaactcctg ggatcctccc 120

acctcagcct cccgagtaca atacctgaat tttaaataga gttattgtaa gtcttatgaa 180

atgagatttt gctgcactct gacataagat aataaaagac agagcaggaa ttcattatta 240

tgagctgctt gatcagtttt aaaccactcc atttgatgaa acaagtgagg tccttccctc 300

ctgaccaggc tgtggaatgc tgtcttcccc aacccccacc ccctgcaaaa gagcagaaca 360

ataaggcaat tgctcatttt 380

<210> SEQ ID NO 99
<211> LENGTH: 562
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (288)..(293)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 99

gcagggaaaa cacagtccac tccccaccac cactccgctc cttccacagc aaatggggac 60

tgctcagaaa accctgtcct ttcttttcct ctcttcaaag gcaggcatgt gattgaggat 120

gtgatggtga cttctggcct gttttatttt gtggtaactt cactttagtc agggaaataa 180

ttggaatatc tttaatctga tgtagtttga cctttagaaa ttgaaagtga aacagctatt 240

gttgataatt cacaaagtat taataaaact tctattacct gtaaaaannn nnnacttaat 300

ctgcgagaaa actatttaga atattatgga attgtgccat agcttcttta tttgttctta 360

attctatatt agatttttt ttctctgctt catgaacaag ttcagatttt aaaacattat 420

gcctgaagac atgtccaact ttatttttta tatgttatat tctggtccaa tatactagag 480

-continued

taataatact ctggtattta ggatatctgt cattgaacct cagttacaaa ttaaacaata 540

gtagctacca tatctaagtg at 562

<210> SEQ ID NO 100
<211> LENGTH: 425
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 100

agatgagatt ggttcccagc catttacttg ctattttaat caacatcaaa gaccagatga 60

tgtgcttctg catcgcactc atgatgagat tgtcctcctg cattgcttcc tctggcccct 120

ggtgacattt gtggtgggcg ttctcattgt ggtcctgacc atctgtgcca agagcttggc 180

gatcaaggcg gaagccatga agaagcgcaa gttctcttaa aggggaagga ggcttgtaga 240

aagcaaagta cagaagctgt actcatcggc acgcgtccac ctgcggaacc tgtgtttcct 300

ggcgcaggag atggacaggg ccacgacagg gctctgagag gctcatccct cagtggcaac 360

agaaacaggc acaactggaa gacttggaac ctcaaagctt gtattccatc tgctgtagca 420

atggc 425

<210> SEQ ID NO 101
<211> LENGTH: 558
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 101

ggatgtttat ggtgaaatgg cctgtacaag tttaactaag acaacttaac ttgcattgtt 60

aatcaaaaat tcttttctca aagggttaac tggttgccat tttgaatagt atgttcaagg 120

gtgtagcttc ctgtttcttt ccaaattata agtagctacc taaatatagt ataattatat 180

attaataata tggcttgctg gcacagtagt ttaccctgtt atctgtgttt cataatgggg 240

gctgtatgaa tattatttaa aactaataaa atgttgccag aattatacta aactgttgga 300

tgagattagg agatcagagg ctggaccttc tcttgataat gcttgttttg ttaaaggtat 360

aatgaaataa tttgtatatg atttgatgaa gattaaagac ccttattttc cacagcttta 420

aaaaaaaacc tttatttatg atcaagtaat aaagataata ttctacttgt gggatcttac 480

attatggaaa tagtttgacg tttttgacct caagagtatg tataatttga agagatactt 540

tgtaactatg cttgggtg 558

<210> SEQ ID NO 102
<211> LENGTH: 256
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (37)..(37)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (40)..(40)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (45)..(45)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base

-continued

<222> LOCATION: (53)..(53)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (57)..(57)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (100)..(100)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (105)..(105)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (112)..(112)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (128)..(128)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (131)..(131)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (137)..(137)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (147)..(147)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (159)..(159)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 102

ctcaacatca taaggaatca gacggatgcg gaaaccnagn cgggntggat agnaaantct 60

ttccaggaag gctccggggc actcaactgg tctccaaccn tcccntgcaa cntgtgacgc 120

ctgccatntt ncccatnttt aggcgantgg caacgcaanc cctccgtttg ctctgggcaa 180

aacttcgaga gttccctctg aagctggagc tttttcctca gatccaagat ccaattggtc 240

accaattcgt gatttc 256

<210> SEQ ID NO 103
<211> LENGTH: 409
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (310)..(310)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (379)..(379)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 103

caagcccatt ttcctgaggg tgaaggggac ttttattata taggggcctt atttggtggg 60

tcagtgctgg aggtttacag gctgatcatg gcctgtcacc aggtgatgat gattgaccaa 120

gccaaccaca tcgaggccct gtggcatgac gaaagcctct taaacaagta cctgcttaac 180

cacaaaccca ctctcccttg agtacatgtg ggattaaaaa gtcgatggag tatacgttgg 240

atgaatacct ggtgggtttg tgcaccatga taaactgcaa gagatctgtg gtcttggtaa 300

-continued

ataacaatgn ggaaatgtga actgatgagg gaagcttcca gaaagagacc agagaggggg 360

tgattgccag tcagcccgna tcttcctcct gaaatgctac cctgattta 409

<210> SEQ ID NO 104
<211> LENGTH: 444
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (36)..(36)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 104

ccttgtcaac atcttcgagc atcggcagct ccggangccg gggtaactgg cagcaggtag 60

gaaactatgt gaaagaatct cctgatgtca taatttccgg gtgtcaccgg aacatttgat 120

catcattcct ttggcaattc cagccttctg tggaaaggcc agtagaaagc attgatttat 180

tcacctctac aggaatcaga ctcagcctct tttggttttc agtgaagtat gccttttcaa 240

tttggaaccc agccaaggag gtttccagtg gaaggaggag attcttcaat tgagctggaa 300

cctgggctga gctccagtgc tgcctgtaat gggaaggaga tgtcaccaac caggcaactc 360

cggaggtgcc ctggaagtca ttgcctgaca ataactgatg ttcccgtcac tgtttatgca 420

acaacgagaa agccacctgc acaa 444

<210> SEQ ID NO 105
<211> LENGTH: 168
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 105

tgaacagagt aacaggacta tatttctgaa aaaggatgaa cttactggaa acaggattgc 60

gtgcctgaaa tatacccaaa tggatataaa tgtcaactca gttctgggct ggagatatag 120

atttggaatg caccaacaat gcagatggta atcctggcat gcgagtag 168

<210> SEQ ID NO 106
<211> LENGTH: 400
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (97)..(97)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (102)..(102)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (104)..(104)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (106)..(106)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (109)..(109)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (162)..(162)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

-continued

<400> SEQUENCE: 106

ttctagaatc tcagcttctt aaatcaagaa attgtgtctt gttcatgtct gaattcccca 60

agtgaacaca gtgagtgggt gcttgacaaa tctttgntgg tnangngcna aaaaagggga 120

ttctgtgccc aataccatga aatcaatgca cagaagattc antcaatcaa gaaaggtgca 180

cagacactgg ccacacacac tgacatttgt ttgcagatgt tccagtcccc ctgacttcca 240

ccaatccatt cattcattcc acaagcattt ttgctggggt ggaagcaggg ccatacaggg 300

tgtgtacatc acagataggg tgggtttgta taatgagtat aaaaaacttc tagcagaaga 360

tgacaaagta tatcaagaaa gggctctctt cgaaatcaca 400

<210> SEQ ID NO 107
<211> LENGTH: 508
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 107

ggctttggga gtgagcagct agacaccaat gacgagagtg atgttatcag tgcactaagt 60

tacatttgcc atatttctca gcagtaaacc tagatgtgga atcaatgtta ctaccgttca 120

ttaaactgcc aaccacagga aacagcctgg caaagattca aactgtaggc caaaaccggc 180

aaaaagtgaa tagagtcctc atgggcccaa tgagcatcca gaaaaggcac ttcaaagagg 240

tgggaaggca gagcatcagg agggaacagg gtgcccaggc atctgtggag aacgctgccg 300

aagaaaaaag gctcgggagt ccagccccaa gggagctgga acagcctcac acacagcagg 360

ggcctgagaa gttagcggga aacgccatct acaccaagcc ttcgttcagc caagagcata 420

aggcagcagt ctctgtgctg acacccttct ccaagggcgc gccttctacc tccagccctg 480

caaaagccct accacaggtg agagacag 508

<210> SEQ ID NO 108
<211> LENGTH: 319
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 108

ttatggctgg gattttgcgc ttagtagttc aatggccccc aggcagacta cagaccgtga 60

caaaaggtgt ggagtctctt atttgtacag attggattcg tcacaattca ccagatcaag 120

aattccagaa aaagtgtttc aggcctcacc tgaagatcat gaaaaatacg gtggggatcc 180

acagaaccct cataaactgc atattgttac cagaataaaa agtacaagaa gacgtccata 240

ttgggaaaaa gatataataa agatgcttgg attagaaaaa gcacataccc ctcaagttca 300

caagaatatc ccttcagtg 319

<210> SEQ ID NO 109
<211> LENGTH: 131
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 109

gttaatttgc tgtgcttctt gcatttttga aagttacata ttctccactg ctttaagaaa 60

taattcagtt cactttcacc ttggcatttc agtatctgtt acacattaga agtagttgtc 120

actatttcat c 131

-continued

<210> SEQ ID NO 110
<211> LENGTH: 398
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 110

gccatgggct tcgaattgcc caccctgaag ctcccatcac aggttacatg tttgggaaag 60

gaatctactt tgctgacatg tcttccaaga gtgccaatta ctgctttgcc tctcgcctaa 120

agaatacagg actgctgctc ttatcagagg tagctctagg tcagtgtaat gaactactag 180

aggccaatcc taaggccgaa ggattgcttc aaggtaaaca tagcaccaag gggctgggca 240

agatggctcc cagttctgcc cacttcgtca ccctgaatgg gagtacagtg ccattaggac 300

cagcaagtga cacaggaatt ctgaatccag atggttatac cctcaactac aatgaatata 360

ttgtatataa ccccaaccag gtccgtatgc ggtacctt 398

<210> SEQ ID NO 111
<211> LENGTH: 335
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 111

gtcagcagtc agcaaattaa catcatcata ctcttccatt tttagtttct gttggatttt 60

catcaagtca atgggctgag aaaccacttc ataatagtct ggttgatttc ttcgctttgg 120

tgccctaatg aagagctcac agagaagtct gccctgttca tccttatagt ctcggatggt 180

attatagagt tcatggcaca cggcaatagg atctacagtt ggaagattgg aaagtctcct 240

ccttttcctg cttgggcctg gtgttgacac agaatggtgc ccatcatcaa agtccccgct 300

gacactgctg gaaggggagg tagctcttct tctct 335

<210> SEQ ID NO 112
<211> LENGTH: 420
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 112

ggatgaccat gtagttgcca gagcagaata tgattttgct gccgtatctg aagaagaaat 60

ttctttccgg gctggtgata tgctgaactt agctctcaaa gaacaacaac ccaaagtgcg 120

tggttggctt ctggctagcc ttgatggcca aacaacagga cttatacctg cgaattatgt 180

caaaattctt ggcaaaagaa aaggtaggaa aacggtggaa tcaagtaaag tttccaagca 240

gcaacaatct tttaccaacc caacactaac taaaggagcc acggttgctg attctttgga 300

tgaacaggaa gctgcctttg aatctgtttt tgttgaaact aataaggttc cagttgcacc 360

tgattccatt gggaaagatg gagaaaagca agatctttga tatctttcat gtttgcctgc 420

<210> SEQ ID NO 113
<211> LENGTH: 487
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 113

gaagtgggac gagcacattt ctattgtctt cacttggatc aaaagcaaaa cagtctctcc 60

gccccgcacc agatcaagta gtttggacat caccctactg aaaacttgcg attcttctta 120

-continued

gttttctgca tacttttcat cacgatgcag gaaacgattt cgagtcaaga agacttttat 180

ttatgaacct ttgaaaggat cgtcttgtat ggtgaatttt ctaggagcga tgatgtactg 240

taattttatt ttaatgtatt ttgatttatg attatttatt agtttttttt aaatgcttgt 300

tctaagacat ttctgaatgt agaccatttt ccaaaaagga aactttattt tcaaaaacct 360

aatccgtagt aattcctaat cttggagaat aaaaaagggc ggtggagggg aaaacattaa 420

gaatttattc attatttctc gagtactttc agaaagtctg acactttcat tgttgtgcca 480

gctggtt 487

<210> SEQ ID NO 114
<211> LENGTH: 398
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (156)..(157)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (177)..(177)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 114

cagaccctgt tccttcgagg aatggggagg gagggaggga ccaaagccgt gaggatgagg 60

acaactccac cctccttcct tccccacagg ccaaccaacc agctgctgac aggggacctg 120

gccattctca ggacaagaga atgcaggcag gcaaanngca ttactgcccc tgtcctnccc 180

caccctgtca tgtgtgattc caggcaccag ggcaggccca gaagcccagc agctgtggga 240

aggaacctgc ctggggccac aggtgcccac tccccaccct gcaggacagg ggtgtctgtg 300

gacactccca cacccaactc tgctaccaag caggcgtctc agctttcctc ctcctttacc 360

ctttcagata caatcacgcc agccacgttg ttttgaaa 398

<210> SEQ ID NO 115
<211> LENGTH: 440
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 115

ggcttaggcc tcaaatggct tcttctaaaa actatggctc tccactcatc agtggttcca 60

ctccaaagca tgagcgtggc tctcccagcc atagtaagtc accagcatat acccccccaga 120

atctggacag tgaaagtgag tcaggctcct ccatagcaga gaaatcttat cagaatagtc 180

ccagctcaga cgatggtatc cgaccacttc cagaatacag cacagagaaa cataagaagc 240

acaaaaagga aaagaagaaa gtaaaagaca aagataggga ccgagaccgg gacaaagacc 300

gagacaagaa aaaatctcat agcatcaagc cagagagttg gtccaaatca cccatctctt 360

cagaccagtc cttgtctatg acaagtaaca caatcttatc tgcagacaga ccctcaaggc 420

tcagcccaga ctttatgatt 440

<210> SEQ ID NO 116
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 116

-continued

gggagaggga ggagtaatgg aggaggagtt ggaaactggg gagagatgga aggaatgtga 60

ctggagggta gagaacttgg agaa 84

<210> SEQ ID NO 117
<211> LENGTH: 523
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 117

gaatcggaca tgtccaaacc accgtgttac gaagaggcgg tgctgatggc agagccgccg 60

ccgccctata gcgaggtgct cacggacacg cgcggcctct accgcaagat cgtcacgccc 120

ttcctgagtc gccgcgacag cgcggagaag caggagcagc cgcctcccag ctacaagccg 180

ctcttcctgg accggggcta cacctcggcg ctgcacctgc ccagcgcccc tcggcccgcg 240

ccgccctgcc cagccctctg cctgcaggcc gaccgtggcc gccgggtctt ccccagctgg 300

accgactcag agctcagcag ccgcgagccc ctggagcacg gagcttggcg tctgccggtc 360

tccatcccct tgttcgggag gactacagcc gtatagaggg gcgcccggcg ccccgggccc 420

caccggcgga ctcctggcct gactgcgggg ctttttaaat gcttccctgg actgcgggga 480

ggggcggggg gagggaggga tttcttatcc cgtttgttac att 523

<210> SEQ ID NO 118
<211> LENGTH: 520
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 118

ttttctactc agcagatgct gtgtgttttg atgttgacag cacggtcatc agtgaagaag 60

gaatcggatg ctttcattgg atttggagga aatgtgatca ggcaacaagt caaggataac 120

gccaaatggt atatcactga ttttgtagag ctgctgggag aaccggaaga ataacatcca 180

ttgtcataca gctccaaaca acttcagatg aatttttaca agttacacag attgatactg 240

tttgcttaca attgcctatt acaacttgct ataaaaagtt ggtacagatg atctgcactg 300

tcaagtaaac tacagttagg aatcctcaaa gattggtttg tttgttttta actgtagttc 360

cagtattata tgatcactat cgatttcctg gagagttttg taatctgaat tctttatgta 420

tattcctagc tatatttcat acaaagtgtt ttaagagtgg agagtcaatt aaacaccttt 480

actcttagga atatagattc ggcagccttc agtgaatatt 520

<210> SEQ ID NO 119
<211> LENGTH: 389
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (105)..(105)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (230)..(230)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (233)..(235)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base

-continued

<222> LOCATION: (353)..(353)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 119

tagagctgaa tattacttga ttacaaatca gattgcttaa gggtgtggaa tagcaggcta 60

gttttaatac caacttgtta acataaaatc atatatgttt taganccatt cttatttagt 120

tacaatttta gaaagttaac aaagtaagca ggtacttatc gaagtgcatc ttttcagtct 180

aaatgtttgt ctgtgtgtct aggtgctggt gagtccacat ggacacatgn agnnnccatg 240

gggcaggagt ctgctataaa gtcagaaggt gagatcctag agagttacac ccagccccat 300

tttaatttgc atgaaaagcc aaggttcttt taagcactca aattatttaa tgnttaaaac 360

acaagaaagg cacatctgtt catttaaat 389

<210> SEQ ID NO 120
<211> LENGTH: 343
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (47)..(47)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (174)..(174)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (236)..(236)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (244)..(246)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (254)..(254)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 120

agagggcctt ggagtgacac cctgaccccc atccactagt acttganggc cagtggtggc 60

agaagccaca gaaacaagaa gcccagtgag atggctaagc tgcccagcat gtaacttaaa 120

tccctgttca ttccccattc ctttagctgc tggagccagt tctgcttctc ggcnaggagc 180

gatttgctgg tgtagacatc cgtgtccgtg taaagggtgg tggtcacgtg ggcccngatt 240

tatnnnagtc ccanaactgg gcgcatggag gaggtggctc tgggagggag gccttcacag 300

cgctcctgta ccctttaatt gtgtgtcttt ctcacagcta tcc 343

<210> SEQ ID NO 121
<211> LENGTH: 457
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 121

ggtgaatgtg tattcctctg ggaggaatag gaagaaaaca ggaatgttaa taatgtcgaa 60

cagaaaactt cctcccttat taatatataa tcctcatgta tttatgccta atgtaagctg 120

acttttaaaa agctttcttt tgttgcatgc cctgtgcagg catctgtatt gtacatgcat 180

gcctttcgtc ctgttttcct gtataaagtt agtgaacaaa gaaatatttt tgcctagttc 240

atgttgccaa gcaatgcata ttttttaaat ttgtcatata tggaaagagc atgtttgtta 300

-continued

catgtaaaag ctttactgat atacagatat actaatgttt gaagatgctg ttctttgcaa 360

gtgtacagtt ttcaaatgtt gttaccagtg aaacaccctt gtggtttaaa cttgctacaa 420

tgtatttatt attcatttcc tcccatgtaa ctaagaa 457

<210> SEQ ID NO 122
<211> LENGTH: 452
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 122

tcagttcata tctttctggg cttgacatgg ctgatggtgt agctgaaacc ctcctaacac 60

taaaagccat ttaatctttt ctgtaatagg agcagaaaat agttaatcat ccacctagta 120

atataagatt actgtgaata ttatcttcta tacattaaaa cagttctagt ttgtagaata 180

ataccataca agttttattt ttaaattcta gttattttca gtgcttactt aaatgtaatt 240

ctagaattcc tccacaactt ttaatatttt gtatgccagt gattctcaag ataaatcatg 300

attgtagtag ttgttactgt tggcagtttg tagtagtatt caggtatttt ggggatgggg 360

gaaaacacca aaaatcagtg tcttttatct ggtgatcact gtggtatcta cagtattcta 420

gtctcctgca caaaaactga acccactggg cc 452

<210> SEQ ID NO 123
<211> LENGTH: 551
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 123

ccaaccatga tgaaactccg tctctactaa aatacaaaaa ttagctgggc atggttgcgt 60

gcgcctgtag tcccagctac ttgggaggct gaggcaggag aatcgcttaa acctgggaga 120

cggaggttgc agtgagccaa gatcgtgtca ctgcactcca gcctggtgac agagtgagac 180

tctgtttcaa aaaagaaaag aaaagaaaca tggttcaaat tatatctaaa caaaaaagaa 240

taagaaacaa aaaacacatt aaaattttaa gttgtatttt ctatgtttct agatacatca 300

tttttgtttg atattttcct gatgcaagta tgtggtttat cacatgtagc tcttttgcat 360

gctaaatgaa aattcaagac ttgccaataa atgaatagct tattgcagac attttttacc 420

aacattaatt attttgggtt tgtttaaaac ctagaggcac aatcttgact tgtcaattac 480

taccctttca caagctacca tctcagatat atatatatat ataaattcaa taaagctttc 540

tgtttgtgtt c 551

<210> SEQ ID NO 124
<211> LENGTH: 420
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 124

cttcacaata gcaaacgtaa acgatggaat tgatggaatc aaccgaaatt gacggaatca 60

atctaaatgt tcatcactga cagattgtgt aaagaaaatg tggaacatgg acaccatgga 120

atagtatgca gccataaaaa gaatgagatc cgatcttttg caggaacatg catggagccg 180

gagacagtta tccttagcaa actaacgcag gaagagaaag ccaaatactg catattctta 240

cgtataagtg ggagctaaat gataagaact tatgagcaca aagtaggaaa ccacagacag 300

-continued

tggcatctcc ttgaggatat agggtgggag cagggagagg agcagaagag atcactattg 360

ggtactgggc ttaatacctg ggtgataaaa taatctgtat aacaaaaccc cgtgacatga 420

<210> SEQ ID NO 125
<211> LENGTH: 421
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 125

aaaggtgttc tctgtgttat gtaaagtgga ggcttcctta tattttaacc tactaagcaa 60

tgaggaggga ttcctgtcat taagcacaag ggcgctggat cctcaagtgc ccatcttcgt 120

gagagaaaaa gcagcacatc ctgcccattt ctggtgcttt ctgctcacag gcaccaaagc 180

tgcacatgta aactgacttc ttgccaaagg aaatgacccc tgggaagttc aagctcctgg 240

aagaggcttt aactcggacg cgccctcctc caggaaccag tgggcagggc agccttcatg 300

catgtgtaac tggacctcca gccataagca tggtgtgcag tatggaagag cctgctacgg 360

aactgaaagt gattggacat tttataggaa ttgatagaga tgttggtcct caaaagctac 420

a 421

<210> SEQ ID NO 126
<211> LENGTH: 533
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (31)..(31)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (55)..(55)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 126

aacaacatta ttccatctca tttaaaggtt naaaaagaag agacaactct agccnaagta 60

gaaatttata ttctacacgt ccaaactgtc tcctagcagc ttttggacta tatatcactt 120

gatgttaaag tatcttttat ttgtaataaa tattcaaatt tctatttaga agctctaatg 180

tatacctaga ttaaatcaaa tcacagtttt atgcttttaa aatatatgta tttcaaactg 240

tatattttaa tttctgagtg catgttatat agtatttaat acttcagatg tcttggcaaa 300

ttcaatataa gtatttattc ccacaagcga tatatgggat atctcttaaa aattatgaat 360

atgtaccatt tccttcaaag tcatcctagc ctatgctgta tcaaaagtat tgtatatttt 420

atggagattt agtgatatac atgtaaatgt tttttaagtt attttattga agttcaatct 480

ttacataaaa ttaaaatctt tttttaaaaa aagtgtcagt gccagaactg taa 533

<210> SEQ ID NO 127
<211> LENGTH: 429
<212> TYPE: DNA
<213> ORGANISM: Danio rerio
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (89)..(89)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 127

tgttcataaa catttgagca ccatgaaatc aaaataccct ataactactt tctatagtca 60

-continued

tatctaattt atatttttt catttccant tgtaactaga tatgtagtaa agtctgaaaa 120

gactttacca tagacaataa catgcagttt tatcagcacc aaagaatgtt gtccaaaaga 180

aacttttaa tacctgtctt tctatttata acatctgaat attttcattc ttatattaag 240

aattttgata agtagattga atttagtatg agtactattt tcttatatat accacaatgg 300

caaacatgta ttataaatca tatttttgtc ttaccaattt taatatatga ggggttttag 360

aaatttgttg taagttattt ttatattcct tgtcttttgc atatttttg gccaaaatct 420

tcaatacat 429

<210> SEQ ID NO 128
<211> LENGTH: 549
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 128

ttcgaggagc aagaaccacc tgagaaaagt gatcctaaac aagaagagtc tcagatatct 60

gggaaggagg aagagacatc agtcaccatc ttagactctt ctgaggaaga taaggaaaaa 120

gaagaggttg ctgctgtcaa aatccaagct gccttccggg gacacatagc cagagaggag 180

gcaaagaaaa tgaaaacaaa tagtcttcaa aatgaggaaa aagaggaaaa caagtgagga 240

cactggtttt acctccagga aacatgaaaa ataatccaaa tccatcaacc ttcttattaa 300

tgtcatttct ccttgaggaa ggaagatttg atgttgtgaa ataacattcg ttactgttgt 360

gaaaatctgt catgagcatt tgtttaataa gcataccatt gaaacatgcc acttgaagat 420

ttctctgaga tcatgagttt gtttacactt gtctcaagcc tatctataga gacccttgga 480

tttagaatta tagaactaaa gtatctgaga ttacagagat ctcagaggtt atgtgttcta 540

actattatc 549

<210> SEQ ID NO 129
<211> LENGTH: 526
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 129

gaaatggcca ctgcttgatg tccaggcagg gagcctccag agtagacaag ccctcaagga 60

tgcccggtcc ccatcaccgg cacacattgt ctctaacaaa gtgccagtgg tgcagcaccc 120

tcaccatgtc caccccctca cgcctcttat cacgtacagc aatgaacact tcacgccggg 180

aaacccacct ccacacttac cagccgacgt agaccccaaa acaggaatcc cacggcctcc 240

gcaccctcca gatatatccc cgtattaccc actatcgcct ggcaccgtag gacaaatccc 300

ccatccgcta ggatggttag taccacagca aggtcaacca gtgtacccaa tcacgacagg 360

aggattcaga cacccctacc ccacagctct gaccgtcaat gcttccatgt ccaggttccc 420

tccccatatg gtcccaccac atcatacgct acacacgacg ggcattccgc atccggccat 480

agtcacacca acagtcaaac aggaatcgtc ccagagtgat gtcggc 526

<210> SEQ ID NO 130
<211> LENGTH: 543
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 130

-continued

tatactcact caaggcagtg caagatcttg aagtactttt tagcagttaa gtaatattga 60

attgtattga atagtttaca tagtttattc tagtctttga aaattactga acatggacaa 120

tgtgcatgtc attgacatct gccttagaac ttctgggaca atcctgattc gagagattct 180

atcccattat ttacatatac caaaaatact ttgttaattt aatgtgttgg cttcccaact 240

cctgaacacg acacaatttt attattagat tttgtatggt gattttaggc tatgaaaaca 300

tgatcattat atgtatatag atacattttt atttgttaca aatgtttgag cagctcacta 360

gcccacccct cctctatttt gggtaagaga atttactacc ttttttaact atgtagttga 420

gagcaacatg tattttgtta tttttagaat ggtcagtata ttgctataaa attttaaatg 480

agactatgaa agttaaagta ttctgattct ggttaaatta acgaatatgg ttccaggccc 540

tgt 543

<210> SEQ ID NO 131
<211> LENGTH: 317
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 131

gttgcagtga tgacttttgt gaaacaaaaa cttatgtatc attttagtga tactcttaag 60

attatttgtt ttttggcagt aaatgtgaaa attctttgtt gttctacttt atgaatagaa 120

cttaaggaaa taactccaaa acaatgtaat ttgtataaga aggttcataa aaatcctgta 180

aggtttaaat tagtttagaa gaaaaataat agtttgctgt aacttttttct ccctaaagaa 240

acaaggtcca actaatccaa tgctgtttca tcttgttcga gacgtcaaac aggtaagaga 300

ttatttttttg cttttga 317

<210> SEQ ID NO 132
<211> LENGTH: 196
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 132

agagctgagt catcctagag caaacctctg gagtggagag cgaactactt cattcccctc 60

ccttagcctg ggccagagag actccagctc tgccttctcc agccaaaaaa tcaaaggcag 120

atgggagaac agccttcagc tttggataac gatgaaatat ctggcaccac tgatgaatat 180

taaactttct ataacc 196

<210> SEQ ID NO 133
<211> LENGTH: 287
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 133

ggagactgtc tcactgatgt tgatttcttt attcatttcc gcatctgtta cacgaacttc 60

gtgtcataaa ttgctatcct ttcatttgaa agtgtaaaaa atttcctgca tttttatcat 120

ttctgtatac ttgagtttat tagagattgt tatgttaggc gacactgtat aaaattgtat 180

ggatattttg agtgaaaatc aaaagtaaaa ttcacatgta tttccttttt tatattttca 240

tccaatttct tgacaacttg aataaatttc ataaagagcc ttcctaa 287

-continued

<210> SEQ ID NO 134
<211> LENGTH: 467
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 134

gcaagctctc attggctctg agcgcgaccc cgcctcccag gggggtggag gtatccactg 60

cacgtgcgcc gcccgggctt cgctcagacc ttcaggtgaa agctgcaaag tcgcgggtgc 120

gtatgtacgg gggctgcctc ccgaggagga gctcccaagc cgcagggtgg acgctggaga 180

caagaacctc agggtcacaa gtttactgtt tttctccctt ttccatccct acattggtct 240

gctggggaag gcggggctag gcatcactga cacacgcaga ctccgtggtt gaggcatttt 300

attggacctt tggcaattgg tggtggggag gcatctgctc caactggtgc ggggccctgc 360

agatgggacc atctcaggct gggtccttgt agcccaggag cacagactgg actaagcctc 420

ctgggccttg tatgaaaaag gtgttgtacc tggccgtttt tgccagt 467

<210> SEQ ID NO 135
<211> LENGTH: 399
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (41)..(41)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (48)..(48)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (58)..(58)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (73)..(73)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (78)..(78)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (91)..(96)
<223> OTHER INFORMATION: a, c, g, t, unknown or other
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (98)..(98)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 135

actccgtcct gggtgacaaa gtgagactcc ctctcaaaaa ntaaatangt aaataaanta 60

aatggtggta acnatacnct atttggtaaa nnnnnncnct aacatctgta gtactaatct 120

tttttccagt ggctttaaac tgcaaataag gaatgttgtt tctgtaggta aaatttttat 180

ttatttttc ccatttaaat ttacttttgt tagtttttc aggcatataa tatttatgtt 240

atatatggca tattctgata agaggcatac aatatgtaat aatcacatta gggtaaatga 300

ggtatccatc acctttagaa tttatttttt gtattatgaa cagttcaatt gtacagtttt 360

agtttttta aaatatacga ttgttattga ctacagggt 399

<210> SEQ ID NO 136
<211> LENGTH: 479

-continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 136

tgataagact cagaacagga agcctgcatg tgactgagca agtcacctac ataaccctgc 60

ctgtactaag gtgtacacct gtctattgta agtttgccta ggctgttggt gtacagagac 120

cagaggagag agacacacta ggactaacaa tgtcctaaca aaatggtact tagtttgttg 180

gtctttagga gaaagcatta gtaatgaaag aagaaagaat tttcacttgg ttggacattg 240

gggctgctta agaaagttga catttgtcgt ggaatgactt tggaaagact tctaaaagaa 300

tctttttcaa atccctgaaa atcaggatag cacattttgc tactgactgt gacagtgttt 360

tattcttttg agagaaatga catagttttc cctttatttc ccaaattcct ttcatgttct 420

taactgctac ccagaaattg agcttcagaa gattgaggat agcctttgat tggtattta 479

<210> SEQ ID NO 137
<211> LENGTH: 459
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 137

agtagatctt accttgctgt tcataagaga atccacaatg gggagaaacc ctatgaatgc 60

aatgactgtg ggaaaacctt cagcagcaga tcttacctta ctgttcataa gagaatccac 120

aatggggaga aaccctacga atgcagtgac tgtgggaaaa ccttcagcaa ttcctcatac 180

ctcagaccgc acttgagaat tcacactgga gaaaaaccgt acaaatgtaa ccagtgtttt 240

cgtgagttcc gcactcagtc aatcttcaca aggcacaaga gagttcatac gggggagggt 300

cattatgtat gtaatcagtg tggaaaggct ttcggcacga ggtcatctct ttcttcgcac 360

tatagcattc atacagggga gtacccttac gaatgccacg attgtgggag aaccttcagg 420

aggaggtcga atctgacaca gcacataaga actcatact 459

<210> SEQ ID NO 138
<211> LENGTH: 467
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 138

aagctgttga cagggctgct tttctttttg gaggctctag gggagcgtct ttctttgccc 60

ttccagcttc tagaagctgc ccaaattctg tggtttgggg cctcctttca aaaccagcaa 120

tggccaatca gtcttacatc actcaaacac ttgagtgttc tgtctccctc ttccatgttt 180

gaggacccctt gtgattacac tgtgaaaacc cagataagcc aggataatct ccctatctta 240

ttatgaggca agtatgttaa gattttattc tataatcaga gaatcttatg ctatgattgt 300

tatatgtgag cattatagat gctcttgaaa tgttaaaatc acatcagcac tggaaaataa 360

ctcctaaatg tccaaaaaga acatgagatt tatggtgctt gaaatgttgc taaacgtaaa 420

tttgtatcta ttctgaaatt atataaatta acctacctgg ccaggca 467

1. A method of diagnosing a mood disorder, the method comprising:
(a) determining the expression of a plurality of biomarkers for the mood disorder in an isolated sample from the individual, the plurality of markers selected from the group of biomarkers listed in Tables 3 and 7; and
(b) diagnosing the presence or absence of the mood disorder based on the expression of the plurality of biomarkers.

2. The method of claim 1, wherein the plurality of biomarkers comprise a subset of about 10 markers designated as Mbp, Edg2, Fzd3, Atxn1, Ednrb for high mood and Fgfr1, Mag, Pmp22, Ugt8, Erbb3 for low mood.

3. The method of claim 1, wherein the plurality of biomarkers comprise a subset of about 10 markers designated as Edg2, Ednrb, Vil2, Bivm, Camk2d for high mood and Trpc1, Elovl5, Ugt8, Btg1, Nefh for low mood. This panel is derived from the meta-analysis.

4. The method of claim 1, wherein the plurality of markers comprise a subset of about 20 biomarkers designated as Mbp, Edg2, Fgfr1, Fzd3, Mag, Pmp22, Ugt8, Erbb3, Igfbp4, Igfbp6, Pde6d, Ptprm, Nefh, Atp2c1, Atxn1, Btg1, C6orf182, Dicer1, Dnajc6, and Ednrb.

5. The method of claim 1, wherein the plurality of markers comprise a subset of about 10 markers for high mood designated as Mbp, Edg2, Fzd3, Atxn1, Ednrb, Pde9a, Plxnd1, Camk2d, Dio2, Lepr, and a subset of about 10 markers for low mood designated as Fgfr1, Mag, Pmp22, Ugt8, Erbb3, Igfbp4, Igfbp6, Pde6d, Ptprm, and Nefh.

6. The method of claim 1, wherein the mood disorder is bipolar disorder or depression (major depressive disorder).

7. The method of claim 1, wherein the sample is a bodily fluid.

8. The method of claim 1, wherein the sample is blood.

9. The method of claim 1, wherein the level of the marker is determined in a tissue biopsy sample of the individual.

10. The method of claim 1, wherein the level of the marker is determined by analyzing the expression level of RNA transcripts.

11. The method of claim 1, wherein the expression level of the marker is determined by analyzing the level of protein or peptides or fragments thereof.

12. The method of claim 1, wherein the expression level is determined by an analytical technique selected from the group consisting of microarray gene expression analysis, polymerase chain reaction (PCR), real-time PCR, quantitative PCR, immunohistochemistry, enzyme-linked immunosorbent assays (ELISA), and antibody arrays.

13. The method of claim 1, wherein the determination of the level of the plurality of biomarkers is performed by an analysis of the presence or absence of the biomarkers.

14. (canceled)

15. (canceled)

16. A method of predicting the probable course and outcome (prognosis) of a mood disorder, the method comprising:
(b) analyzing the presence or level of a plurality of markers of the mood disorder in a test sample, the markers selected from the group consisting of markers listed in Tables 3 and 7; and
(c) determining the prognosis based on the presence or level of the markers and one or more clinicopathological data to implement a treatment plan.

17. The method of claim 16, wherein the treatment plan is for a high mood disorder if the molecular markers selected from the group consisting of Mbp, Edg2, Fzd3, Atxn1, and Ednrb are present.

18. The method of claim 16, wherein the treatment plan is for a low mood disorder if the molecular markers selected from the group consisting of Fgfr1, Mag, Pmp22, Ugt8, and Erbb3 are present.

19. The method of claim 16, wherein the treatment plan for a high mood disorder comprises administering a pharmaceutical composition selected from the group consisting of divalproex, lithium, lamotrigene, carbamazepine, topiramate.

20. The method of claim 16, wherein the treatment plan for a low mood disorder comprises administering a pharmaceutical composition selected from the group consisting of fluoxetine, sertraline, citalopram, duloxetine, venlafaxine and buproprion.

21. The method of claim 16, wherein the clinicopathological data is selected from the group consisting of patient age, previous personal and/or familial history of the mood disorder, previous personal and/or familial history of response to mood disorder, and any genetic or biochemical predisposition to psychiatric illness.

22. The method of claim 16, wherein the test sample from the subject is of a test sample selected from the group consisting of fresh blood, stored blood, fixed, paraffin-embedded tissue, tissue biopsy, tissue microarray, fine needle aspirates, peritoneal fluid, ductal lavage and pleural fluid or a derivative thereof.

23. (canceled)

24. (canceled)

25. The method of claim 16, wherein the treatment plan is a personalized plan for the patient.

26. (canceled)

27. (canceled)

28. (canceled)

29. (canceled)

30. (canceled)

31. (canceled)

32. (canceled)

33. (canceled)

34. (canceled)

35. A method of diagnosing bipolar mood disorder using blood biomarkers, the method comprising analyzing expression profile of a plurality of biomarkers selected from the group consisting biomarkers listed in Tables 3 and 7 whose expression levels in a blood sample is associated with an increased risk of bipolar disorder.

* * * * *

专利名称(译)	情绪障碍的血液生物标志物		
公开(公告)号	US20100256001A1	公开(公告)日	2010-10-07
申请号	US12/594378	申请日	2008-04-02
[标]申请(专利权)人(译)	斯克里普斯研究学院		
申请(专利权)人(译)	斯克里普斯研究所		
当前申请(专利权)人(译)	印第安纳大学研究与科技股份有限公司 斯克里普斯研究所		
[标]发明人	NICULESCU ALEXANDER B NURNBERGER JOHN I SALOMON DANIEL R		
发明人	NICULESCU, ALEXANDER B. NURNBERGER, JOHN I. SALOMON, DANIEL R.		
IPC分类号	C40B30/00 G01N33/68 C12Q1/48 C12Q1/44 C12Q1/42 C12Q1/02 C12Q1/68 G01N33/53		
CPC分类号	C12Q1/6883 G01N33/6896 G01N2800/304 C12Q2600/158 C12Q2600/106 C12Q2600/112 C12Q2600/136 G01N2800/52		
优先权	60/909859 2007-04-03 US		
外部链接	Espacenet USPTO		

摘要(译)

多种标志物基于它们在诸如血液的样品中的表达来确定情绪障碍的诊断。生物标志物的子集预测高或低情绪障碍的诊断。使用基于动物和人类数据的会聚功能基因组学方法鉴定生物标志物。提供了用于临床诊断情绪障碍的方法和组合物。

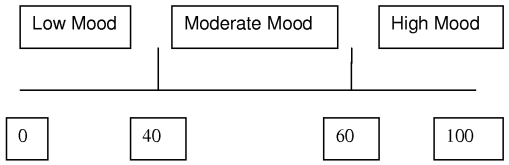


FIG. 1