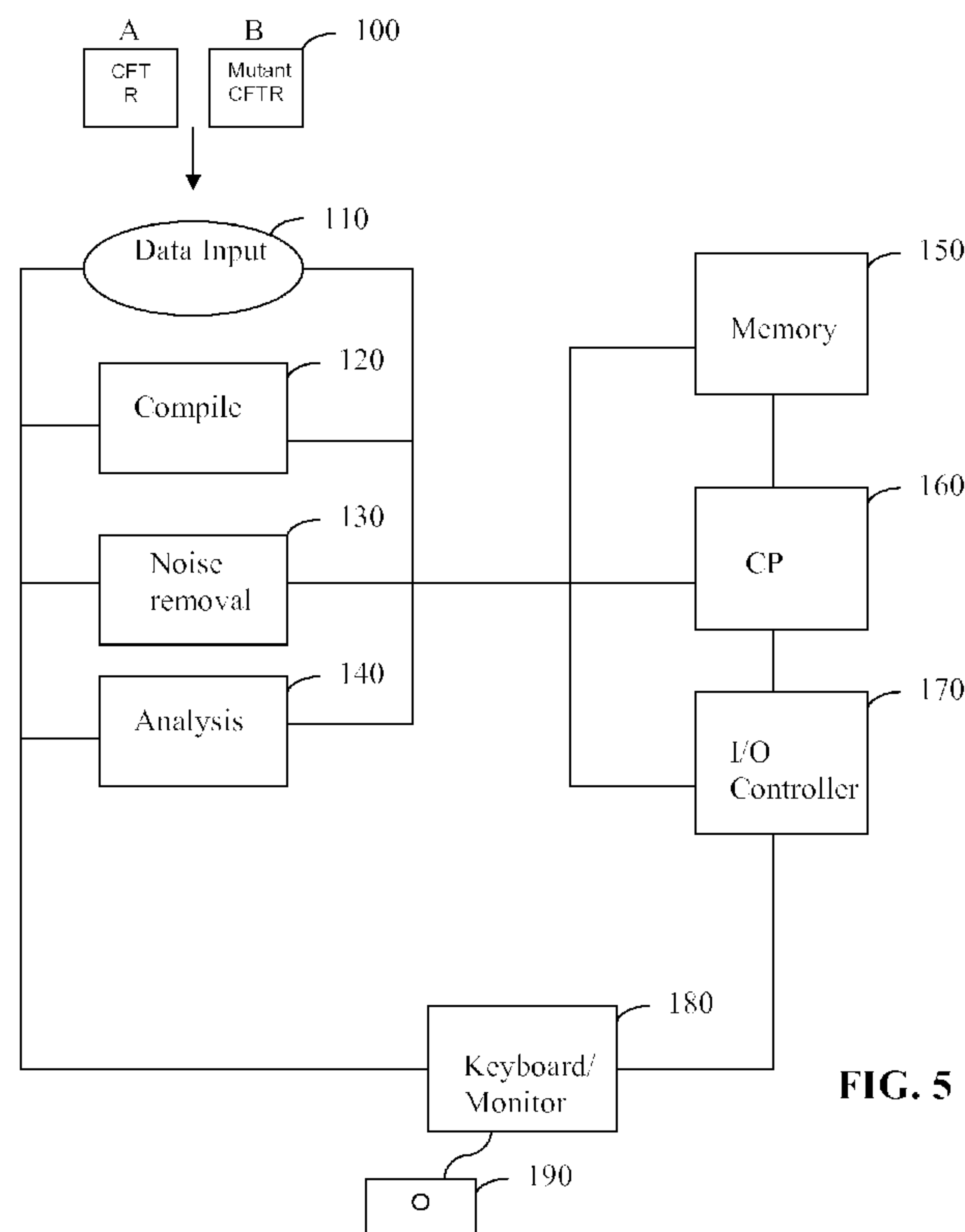




(86) Date de dépôt PCT/PCT Filing Date: 2011/03/22
(87) Date publication PCT/PCT Publication Date: 2011/09/29
(85) Entrée phase nationale/National Entry: 2012/09/20
(86) N° demande PCT/PCT Application No.: US 2011/029314
(87) N° publication PCT/PCT Publication No.: 2011/119531
(30) Priorités/Priorities: 2010/03/22 (US61/316,321);
2010/06/28 (US61/359,029)

(51) Cl.Int./Int.Cl. *C07K 16/28* (2006.01),
C07H 21/00 (2006.01), *C12Q 1/68* (2006.01),
C40B 30/00 (2006.01), *C40B 40/08* (2006.01),
G01N 33/48 (2006.01), *G01N 33/53* (2006.01)
(71) Demandeur/Applicant:
ESOTERIX GENETIC LABORATORIES, LLC, US
(72) Inventeurs/Inventors:
ROHLFS, ELIZABETH M., US;
SIRKO-OSADSA, DEBORAH ALEXA, US;
ROSENBLUM, LYNNE SARAH, US;
ZHOU, ZHAOQING, US;
HEIM, RUTH A., US;
NAGAN, NARASIMHAN, US
(74) Agent: MOFFAT & CO.

(54) Titre : MUTATIONS ASSOCIEES A LA MUCOVISCIDOSE
(54) Title: MUTATIONS ASSOCIATED WITH CYSTIC FIBROSIS



(57) **Abrégé/Abstract:**

The present invention provides novel mutations identified in the cystic fibrosis transmembrane conductance regulator (CFTR) gene that can be used for a more accurate diagnosis of cystic fibrosis (CF) and CF related disorders. Methods for testing a sample obtained from a subject to determine the presence of one or more mutations in the CFTR gene are provided wherein the presence of one or more mutations indicates that the subject has CF or a CF related disorder, or is a carrier of a CFTR mutation.



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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
29 September 2011 (29.09.2011)

(10) International Publication Number
WO 2011/119531 A1

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

(21) International Application Number:
PCT/US2011/029314

(22) International Filing Date:
22 March 2011 (22.03.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/316,321 22 March 2010 (22.03.2010) US
61/359,029 28 June 2010 (28.06.2010) US

(71) Applicant (for all designated States except US): **ESOTERIX GENETIC LABORATORIES, LLC** [US/US]; 358 South Main Street, Burlington, North Carolina 27215 (US).

(72) Inventor; and

(71) Applicant (for all designated States except US): **NAGAN, Narasimhan** [US/US]; 30 Daffodil Court, South Grafton, Massachusetts 01560 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ROHLFS, Elizabeth M.** [US/US]; 111 Winter Street, Hopkinton, Massachusetts 01748 (US). **SIRKO-OSADSA, Deborah Alexa** [US/US]; 149 Brigham Hill Road, North Grafton, Massachusetts 01536 (US). **ROSENBLUM, Lynne Sarah** [US/US]; 2 Teresa Road, Hopkinton, Massachusetts 01748 (US). **ZHOU, Zhaoqing** [US/US]; 10 Nouvelle Way Unit S707, Natick, Massachusetts 01760 (US). **HEIM, Ruth A.** [US/US]; 396 Lake Street, Shrewsbury, Massachusetts 01545 (US).

(74) Agents: **CALKINS, Charles W.** et al.; Kilpatrick Townsend & Stockton LLP, 1001 W. Fourth Street, Winston-Salem, North Carolina 27101-2400 (US).

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

[Continued on next page]

(54) Title: MUTATIONS ASSOCIATED WITH CYSTIC FIBROSIS

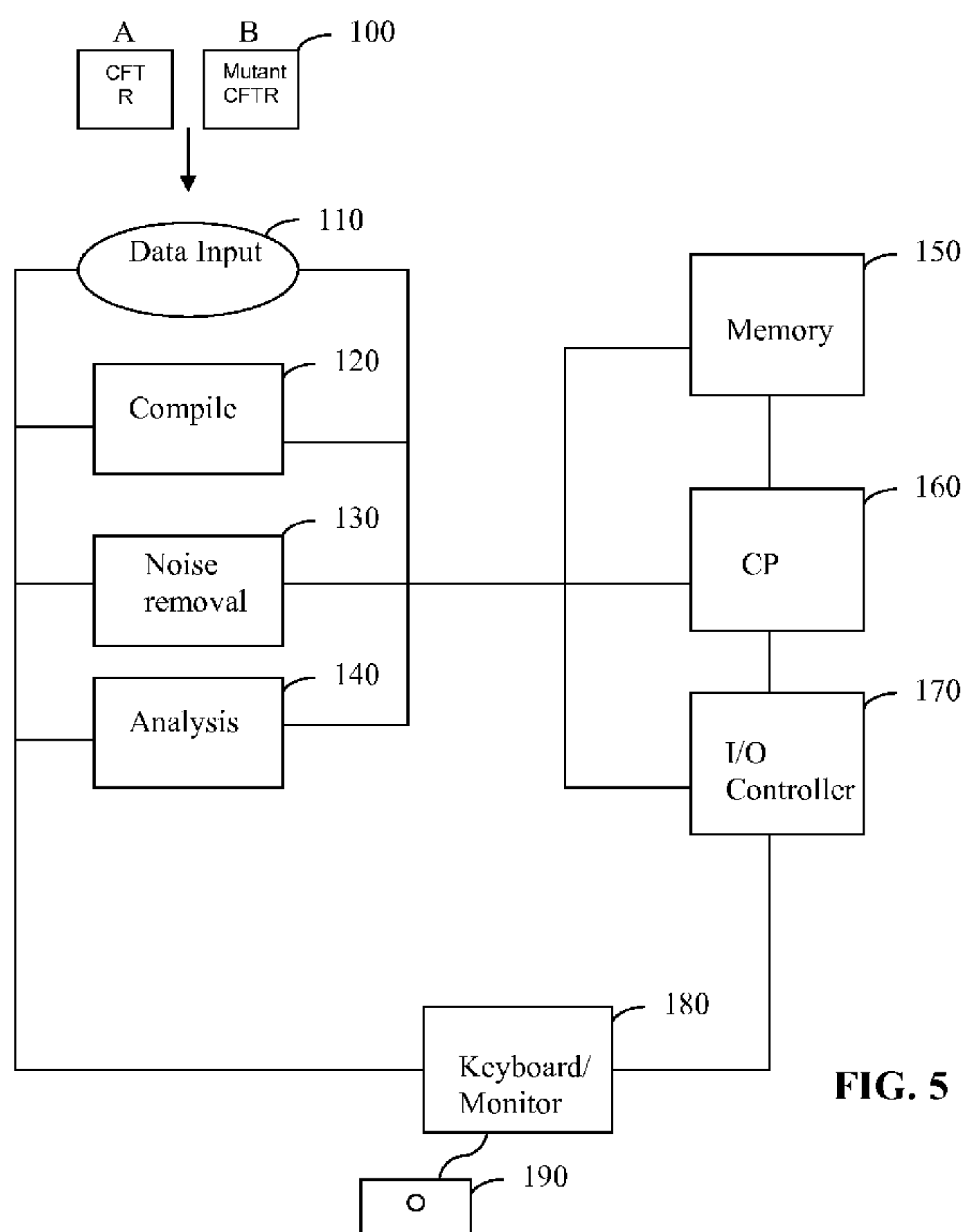


FIG. 5

(57) Abstract: The present invention provides novel mutations identified in the cystic fibrosis transmembrane conductance regulator (CFTR) gene that can be used for a more accurate diagnosis of cystic fibrosis (CF) and CF related disorders. Methods for testing a sample obtained from a subject to determine the presence of one or more mutations in the CFTR gene are provided wherein the presence of one or more mutations indicates that the subject has CF or a CF related disorder, or is a carrier of a CFTR mutation.

WO 2011/119531 A1



NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

MUTATIONS ASSOCIATED WITH CYSTIC FIBROSIS

This application claims priority to under 35 U.S.C. §119(e) to U.S. Patent Application Nos. 61/316,321, filed March 22, 2010 and 61/359,029, filed on June 28, 2010. The disclosures of each of U.S. Patent Application Nos. 61/316,321 and 61/359,029 are expressly
5 incorporated by reference in their entireties herein.

BACKGROUND OF THE INVENTION

Cystic fibrosis (CF) is the most common severe autosomal recessive genetic disorder in the Caucasian population. It affects approximately 1 in 2,500 Caucasian live births in North America (Boat *et al*, The Metabolic Basis of Inherited Disease, 6th ed, pp 2649-2680,
10 McGraw Hill, NY (1989)). The incidence of disease is lower in African American, Hispanic and Asian individuals. Approximately 1 in 25 Caucasian persons are carriers of the disease. The responsible gene has been localized to a 250,000 base pair genomic sequence present on the long arm of chromosome 7. This sequence encodes a membrane-associated protein called the “cystic fibrosis transmembrane regulator” (or “CFTR”). The CFTR gene contains 27
15 exons and encodes a protein of 1480 amino acids. Several regions are contemplated to have functional importance in the CFTR protein, including two areas for ATP binding, termed Nucleotide Binding Folds (NBF), a Regulatory (R) region that has multiple potential sites for phosphorylation by protein kinases A and C, and two hydrophobic regions believed to interact with cell membranes.

20 The major symptoms of classical cystic fibrosis include chronic pulmonary disease, pancreatic exocrine insufficiency, congenital absence of the vas deferens in males and elevated sweat electrolyte levels. The symptoms are consistent with CF being an exocrine disorder. Although recent advances have been made in the analysis of ion transport across the apical membrane of the epithelium of CF patient cells, it is not clear that the abnormal
25 regulation of chloride channels represents the only defect in the disease. Mutations in the CFTR gene are also associated with atypical CF and monosymptomatic diseases such as congenital absence of the vas deferens in males, idiopathic chronic pancreatitis and chronic sinusitis (Noone and Knowles, *Respir. Res.*, vol. 2, p. 328 (2001); Southern, *Respiration*, vol. 74, p. 241 (2007)). A variety of CFTR gene mutations are known. One of them leads to the
30 omission of phenylalanine residue 508 within the first putative NBF domain. This mutation, termed $\Delta F508$, accounts for about 70% of the CFTR chromosomes in Caucasian patients and was highly associated with the predominant haplotype found on chromosomes of Caucasian

CF patients (Kerem, et al., *Science*, vol. 245, p. 1073 (1989); Lemna, et al., *New Engl. J. Med.*, vol. 322, p. 291 (1990)). However, the haplotypes associated with Caucasian CF chromosomes without $\Delta F508$ also exist although less common, confirming that allelic heterogeneity is present in CF and CF related disorders.

5 Therefore, there is a need for more effective genetic screening for other CFTR mutant alleles which are present in the other 30% of Caucasian CF patients, as well as other alleles found in other racial and ethnic groups. Knowledge of such alleles can be used to design probes for screening and/or testing, as well as to devise other screening and/or testing methods. The more complete the set of probes available for CFTR mutant alleles, the more
10 accurate the diagnoses.

SUMMARY OF THE INVENTION

The present invention provides methods, products and systems relating to novel mutations identified in the CFTR gene that can be used for more accurate diagnosis of CF and CF related disorders.

15 In one aspect, the present invention provides a method for testing for mutations in the CFTR gene, which comprises testing a sample obtained from a subject to determine the presence of one or more mutations selected from Table 1, 2, 3, or 4 in the CFTR gene or protein, wherein the presence of the one or more mutations indicates that the subject has CF or a CFTR related disorder, is at risk of developing CF or a CF related disorder, or is a carrier
20 of a CFTR mutation. In some embodiments, the one or more mutations are selected from Table 1, 2 or 3. In some embodiments, the one or more mutations are selected from Table 1 or 2. In some embodiments, the one or more mutations are selected from Table 1. In some embodiments, the one or more mutations selected from Table 1, 2, 3, or 4 are part of a panel of CFTR mutations.

25 Yet other embodiments of the present invention comprise systems for performing the method. For example, the system may comprise a station or device for testing a sample obtained from a subject to determine the presence of one or more mutations selected from Table 1, 2, 3, or 4 in the CFTR gene or protein, wherein the presence of the one or more mutations indicates that the subject has CF or a CFTR related disorder, is at risk of
30 developing CF or a CF related disorder, or is a carrier of a CFTR mutation. In some embodiments, the one or more mutations are selected from Table 1, 2 or 3. In some embodiments, the one or more mutations are selected from Table 1 or 2. In some embodiments, the one or more mutations are selected from Table 1. In some embodiments,

the one or more mutations selected from Table 1, 2, 3, or 4 are part of a panel of CFTR mutations. Also, the system may comprise a device for analysis and/or interpretation of the data. For example, a computer having software to analyze the data for the presence of one of the mutations of the invention may be included in the system.

5 The following embodiments may be used in either the methods or the systems of the invention. In some embodiments, the sample contains an isolated nucleic acid. In some embodiments, the testing step comprises nucleic acid sequencing. In some embodiments, the testing step comprises hybridization. In some embodiments, the hybridization is performed using one or more oligonucleotide probes specific for a region in the CFTR gene (SEQ ID
10 NO:1) (FIG. 1) corresponding to the one or more mutations selected from Table 1, 2, 3 or 4, and under conditions sufficiently stringent to disallow a single nucleotide mismatch. In some embodiments, the hybridization is performed with a microarray. In some embodiments, the testing step comprises restriction enzyme digestion. In some embodiments, the testing step comprises PCR amplification. In some embodiments, the PCR amplification is digital PCR
15 amplification. In some embodiments, the testing step comprises primer extension. In some embodiments, the primer extension is single-base primer extension. In some embodiments, the testing step comprises performing a multiplex allele-specific primer extension (ASPE). In yet other embodiments, the testing step may comprise performing real-time PCR.

 In some embodiments, the sample contains purified or partially purified protein. In
20 some embodiments, the testing step comprises amino acid sequencing. For example, in certain embodiments, the system comprises a device for amino acid sequencing. In some embodiments, the testing step comprises performing an immuno assay using one or more antibodies that specifically recognize one or more epitopes corresponding to the one or more mutations selected from Table 1, 2, 3 or 4. In some embodiments, the testing step comprises
25 protease digestion (e.g., trypsin digestion). In some embodiments, the testing step further comprises performing 2D-gel electrophoresis.

 In some embodiments, the testing step comprises determining the presence of the one or more mutations using mass spectrometry. In some embodiments, the mass spectrometric format is selected from among Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight
30 (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform, and combinations thereof.

 In some embodiments, the sample is obtained from cells, tissue, whole blood, mouthwash, plasma, serum, urine, stool, saliva, cord blood, chorionic villus sample, chorionic villus sample culture, amniotic fluid, amniotic fluid culture, transcervical lavage

fluid, and combination thereof. In further embodiments, the sample is obtained from a pregnant woman, for testing the sample for the presence of one or more CFTR mutations in fetal nucleic acids contained therein. For example, in certain embodiments, the system comprises a station for processing of the samples.

5 In yet another aspect, the present invention provides a method for screening and/or testing for CFTR mutations, comprising steps of: (a) providing a sample obtained from a subject; (b) testing the sample for the presence of a mutation at a pre-determined position selected from Table 1, 2, 3 or 4, in the CFTR gene or protein; and wherein the presence of the mutation at the pre-determined position indicates that the subject has an increased risk of
10 having CF or a CF related disorder, or being a carrier of a CFTR mutation.

 Yet other embodiments of the present invention comprise systems for performing the method. For example, the system may comprise a station or device for testing a sample obtained from a subject to determine the presence of one or more mutations selected from Table 1, 2, 3, or 4 in the CFTR gene or protein, wherein the presence of the one or more
15 mutations indicates that the subject has CF or a CFTR related disorder, is at risk of developing CF or a CF related disorder, or is a carrier of a CFTR mutation. In some embodiments, the one or more mutations are selected from Table 1, 2 or 3. In some embodiments, the one or more mutations are selected from Table 1 or 2. In some embodiments, the one or more mutations are selected from Table 1. In some embodiments,
20 the one or more mutations selected from Table 1, 2, 3, or 4 are part of a panel of CFTR mutations. Also, the system may comprise a device for analysis and/or interpretation of the data. For example, a computer having software to analyze the data for the presence of one of the mutations of the invention may be included in the system.

 The following embodiments may be used in either the methods or the systems of the invention. In some embodiments, the testing step comprises determining the identity of the
25 nucleotide and/or amino acid at the pre-determined position selected from Table 1, 2, 3 or 4.

 In some embodiments, the presence of the mutation is determined by comparing the identity of the nucleotide and/or amino acid at the pre-determined position to a control.

 In some embodiments, the method further comprises a step of determining if the
30 mutation is listed in Table 1, 2, 3 or 4.

 In another aspect, the present invention provides products, e.g., reagents, for detecting novel CFTR mutations described herein. Such reagents may be used for detection of the mutations described herein in the protein sequence and/or the nucleic acid sequence.

In some embodiments, the invention provides a nucleic acid probe that specifically binds to a normal CFTR gene but not to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4. In some embodiments, the present invention provides a plurality of probes (e.g., as may be used for real-time PCR or sequencing), or an array containing one or more probes that specifically bind to a normal CFTR gene but not to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4. In some embodiments, the present invention provides a nucleic acid probe that specifically binds to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4 but not to a normal CFTR gene. In some embodiments, the array comprises one or more probes that specifically bind to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4 but not to a normal CFTR gene.

In some embodiments, the present invention provides an antibody that specifically binds to a normal CFTR protein but not to a mutant CFTR protein containing one or more mutations selected from Table 1, 2, 3, or 4. In some embodiments, the present invention provides an antibody that specifically binds to a mutant CFTR protein containing one or more mutations selected from Table 1, 2, 3, or 4 but not to a normal CFTR protein.

In some embodiments, the present invention provides a kit for comprising one or more reagents that differentiate a normal CFTR gene or protein from a mutant CFTR gene or protein containing one or more mutations selected from Table 1, 2, 3, or 4. Such kits may be useful, e.g., for screening and/or testing for CFTR mutations. In some embodiments, the one or more reagents comprises one or more nucleic acid probes. In some embodiments, the one or more reagents comprises one or more antibodies. In some embodiments, the one or more reagents are provided in a form of microarray. In some embodiments, the kit further comprises reagents for primer extension. Or, probes for the detection of mutations may be provided. In some embodiments, the kit further comprises a control indicative of a healthy individual. In some embodiments, the kit further comprises an instruction on how to determine if an individual has CF or a CF related disorder, is at risk of developing CF or a CF related disorder, or is a carrier of a CFTR mutation.

In still another aspect, the present invention provides a computer readable medium encoding information corresponding to one or more mutations shown in Tables 1, 2, 3 and 4. Such computer readable media may be part of the systems as described herein.

Other features, objects, and advantages of the present invention are apparent in the detailed description and claims that follow. It should be understood, however, that the detailed description, the drawings, and the claims, while indicating embodiments of the

present invention, are given by way of illustration only, not limitation. Various changes and modifications within the scope of the invention will become apparent to those skilled in the art.

FIGURES

5 FIG 1: Figure 1 is a genomic sequence of the CFTR gene according to an embodiment of the invention.

 FIG 2: Figure 2 is a cDNA sequence of CFTR according to an embodiment of the invention.

 FIG 3: Figure 3 is an amino acid sequence of CFTR according to an embodiment of
10 the invention.

 FIG 4: Figure 4 is a nucleotide sequence of the 5' end of the CFTR gene according to an embodiment of the invention.

 FIG. 5: Figure 5 is is a schematic of a system according to an embodiment of the invention.

DEFINITIONS

In order for the present invention to be more readily understood, certain terms are first defined. Additional definitions for the following terms and other terms are set forth throughout the specification.

5 In this application, the use of “or” means “and/or” unless stated otherwise. As used in this application, the term “comprise” and variations of the term, such as “comprising” and “comprises,” are not intended to exclude other additives, components, integers or steps. As used in this application, the terms “about” and “approximately” are used as equivalents. Any numerals used in this application with or without about/approximately are meant to cover any
10 normal fluctuations appreciated by one of ordinary skill in the relevant art. In certain embodiments, the term “approximately” or “about” refers to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such
15 number would exceed 100% of a possible value).

Antibody: As used herein, the term “antibody” refers to a polypeptide consisting of one or more polypeptides substantially encoded by immunoglobulin genes or fragments of immunoglobulin genes. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as myriad
20 immunoglobulin variable region genes. Light chains are typically classified as either kappa or lambda. Heavy chains are typically classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. A typical immunoglobulin (antibody) structural unit is known to comprise a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one
25 “light” (about 25 kD) and one “heavy” chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms “variable light chain” (VL) and “variable heavy chain” (VH) refer to these light and heavy chains respectively. An antibody can be specific for a particular antigen. The antibody or its antigen can be either an analyte or a binding partner.
30 Antibodies exist as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)₂, a dimer of Fab which itself is a light chain joined to VH-CH1 by a disulfide bond. The F(ab)₂ may be reduced under mild conditions to break the disulfide linkage in the hinge region thereby

converting the (Fab')₂ dimer into an Fab' monomer. The Fab' monomer is essentially an Fab with part of the hinge region (see, Fundamental Immunology, W. E. Paul, ed., Raven Press, N.Y. (1993), for a more detailed description of other antibody fragments). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of ordinary skill in the art will appreciate that such Fab' fragments may be synthesized *de novo* either chemically or by utilizing recombinant DNA methodology. Thus, the term "antibody," as used herein also includes antibody fragments either produced by the modification of whole antibodies or synthesized *de novo* using recombinant DNA methodologies. In some embodiments, antibodies are single chain antibodies, such as single chain Fv (scFv) antibodies in which a variable heavy and a variable light chain are joined together (directly or through a peptide linker) to form a continuous polypeptide. A single chain Fv ("scFv") polypeptide is a covalently linked VH::VL heterodimer which may be expressed from a nucleic acid including VH- and VL-encoding sequences either joined directly or joined by a peptide-encoding linker. (See, *e.g.*, Huston, *et al.* (1988) Proc. Nat. Acad. Sci. USA, 85:5879-5883, the entire contents of which are herein incorporated by reference.) A number of structures exist for converting the naturally aggregated, but chemically separated light and heavy polypeptide chains from an antibody V region into an scFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, *e.g.* U.S. Pat. Nos. 5,091,513 and 5,132,405 and 4,956,778.

Allele: As used herein, the term "allele" refers to different versions of a nucleotide sequence of a same genetic locus (*e.g.*, a gene).

Allele specific primer extension (ASPE): As used herein, the term "allele specific primer extension (ASPE)" refers to a mutation detection method utilizing primers which hybridize to a corresponding DNA sequence and which are extended depending on the successful hybridization of the 3' terminal nucleotide of such primer. Typically, extension primers that possess a 3' terminal nucleotide which form a perfect match with the target sequence are extended to form extension products. Modified nucleotides can be incorporated into the extension product, such nucleotides effectively labeling the extension products for detection purposes. Alternatively, an extension primer may instead comprise a 3' terminal nucleotide which forms a mismatch with the target sequence. In this instance, primer extension does not occur unless the polymerase used for extension inadvertently possesses exonuclease activity.

Amplification: As used herein, the term "amplification" refers to any methods known in the art for copying a target nucleic acid, thereby increasing the number of copies of a

selected nucleic acid sequence. Amplification may be exponential or linear. A target nucleic acid may be either DNA or RNA. Typically, the sequences amplified in this manner form an “amplicon.” Amplification may be accomplished with various methods including, but not limited to, the polymerase chain reaction (“PCR”), transcription-based amplification, isothermal amplification, rolling circle amplification, etc. Amplification may be performed with relatively similar amount of each primer of a primer pair to generate a double stranded amplicon. However, asymmetric PCR may be used to amplify predominantly or exclusively a single stranded product as is well known in the art (e.g., Poddar et al. *Molec. And Cell. Probes* 14:25-32 (2000)). This can be achieved using each pair of primers by reducing the concentration of one primer significantly relative to the other primer of the pair (e.g., 100 fold difference). Amplification by asymmetric PCR is generally linear. Additionally, methods such as real-time PCR may be utilized. A skilled artisan will understand that different amplification methods may be used together.

Animal: As used herein, the term “animal” refers to any member of the animal kingdom. In some embodiments, “animal” refers to humans, at any stage of development. In some embodiments, “animal” refers to non-human animals, at any stage of development. In certain embodiments, the non-human animal is a mammal (e.g., a rodent, a mouse, a rat, a rabbit, a monkey, a dog, a cat, a sheep, cattle, a primate, and/or a pig). In some embodiments, animals include, but are not limited to, mammals, birds, reptiles, amphibians, fish, insects, and/or worms. In some embodiments, an animal may be a transgenic animal, genetically-engineered animal, and/or a clone.

Approximately: As used herein, the term “approximately” or “about,” as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In certain embodiments, the term “approximately” or “about” refers to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100% of a possible value).

Biological sample: As used herein, the term “biological sample” encompasses any sample obtained from a biological source. A biological sample can, by way of non-limiting example, include blood, amniotic fluid, sera, urine, feces, epidermal sample, skin sample, cheek swab, sperm, amniotic fluid, cultured cells, bone marrow sample and/or chorionic villi. Convenient biological samples may be obtained by, for example, scraping cells from the surface of the buccal cavity. The term biological sample encompasses samples which have

been processed to release or otherwise make available a nucleic acid or protein for detection as described herein. For example, a biological sample may include a cDNA that has been obtained by reverse transcription of RNA from cells in a biological sample. The biological sample may be obtained from a stage of life such as a fetus, young adult, adult, and the like.

5 Fixed or frozen tissues also may be used.

Carrier: The term “carrier,” as used in the context of CF, refers to a person who is symptom-free but carries a CFTR mutation that can be passed to his/her children. Typically, a carrier has one CFTR allele that contains a disease causing mutation and a second allele that is normal or not disease-related. CF and CF related disorders are “autosomal recessive”
10 diseases, meaning that a mutation produces little or no phenotypic effect when present in a heterozygous configuration with a non-disease related allele, but produces a “disease state” when a person is homozygous, i.e., both CFTR alleles are mutant alleles that contain the same disease causing mutation or compound heterozygous, i.e., both CFTR alleles are mutant alleles that contain two different disease-causing mutations. A carrier status is whether or not
15 one is a carrier.

Coding sequence vs. non-coding sequence: As used herein, the term “coding sequence” refers to a sequence of a nucleic acid or its complement, or a part thereof, that can be transcribed and/or translated to produce the mRNA for and/or the polypeptide or a fragment thereof. Coding sequences include exons in a genomic DNA or immature primary
20 RNA transcripts, which are joined together by the cell’s biochemical machinery to provide a mature mRNA. The anti-sense strand is the complement of such a nucleic acid, and the encoding sequence can be deduced therefrom. As used herein, the term “non-coding sequence” refers to a sequence of a nucleic acid or its complement, or a part thereof, that is not transcribed into amino acid in vivo, or where tRNA does not interact to place or attempt
25 to place an amino acid. Non-coding sequences include both intron sequences in genomic DNA or immature primary RNA transcripts, and gene-associated sequences such as promoters, enhancers, silencers, etc.

Complement: As used herein, the terms “complement,” “complementary” and “complementarity,” refer to the pairing of nucleotide sequences according to Watson/Crick
30 pairing rules. For example, a sequence 5’-GCCGTCCCA-3’ has the complementary sequence of 5’-TGGGACCGC-3’. A complement sequence can also be a sequence of RNA complementary to the DNA sequence. Certain bases not commonly found in natural nucleic acids may be included in the complementary nucleic acids including, but not limited to, inosine, 7- deazaguanine, Locked Nucleic Acids (LNA), and Peptide Nucleic Acids (PNA).

Complementary need not be perfect; stable duplexes may contain mismatched base pairs, degenerative, or unmatched bases. Those skilled in the art of nucleic acid technology can determine duplex stability empirically considering a number of variables including, for example, the length of the oligonucleotide, base composition and sequence of the oligonucleotide, ionic strength and incidence of mismatched base pairs.

Control: As used herein, the term “control” has its art-understood meaning of being a standard against which results are compared. Typically, controls are used to augment integrity in experiments by isolating variables in order to make a conclusion about such variables. In some embodiments, a control is a reaction or assay that is performed simultaneously with a test reaction or assay to provide a comparator. In one experiment, the “test” (*i.e.*, the variable being tested) is applied. In the second experiment, the “control,” the variable being tested is not applied. In some embodiments, a control is a historical control (*i.e.*, of a test or assay performed previously, or an amount or result that is previously known). In some embodiments, a control is or comprises a printed or otherwise saved record. A control may be a positive control or a negative control.

Crude: As used herein, the term “crude,” when used in connection with a biological sample, refers to a sample which is in a substantially unrefined state. For example, a crude sample can be cell lysates or biopsy tissue sample. A crude sample may exist in solution or as a dry preparation.

Deletion: As used herein, the term “deletion” encompasses a mutation that removes one or more nucleotides from a naturally-occurring nucleic acid.

Epitope: As used herein, the term “epitope” refers to a fragment or portion of a molecule or a molecule compound (e.g., a polypeptide or a protein complex) that makes contact with a particular antibody or antibody like proteins.

Familial history: As used herein, the term “familial history” typically refers to occurrence of events (e.g., CF disease, CF related disorder or CFTR mutation carrier) relating to an individual’s immediate family members including parents and siblings. Sometimes, family history also may include grandparents.

Flanking: As used herein, the term “flanking” is meant that a primer hybridizes to a target nucleic acid adjoining a region of interest sought to be amplified on the target. The skilled artisan will understand that preferred primers are pairs of primers that hybridize 3’ from a region of interest, one on each strand of a target double stranded DNA molecule, such that nucleotides may be added to the 3’ end of the primer by a suitable DNA polymerase. For example, primers that flank mutant CFTR sequences do not actually anneal to the mutant

sequence but rather anneal to sequence that adjoins the mutant sequence. In some cases, primers that flank a CFTR exon are generally designed not to anneal to the exon sequence but rather to anneal to sequence that adjoins the exon (e.g. intron sequence). However, in some cases, amplification primer may be designed to anneal to the exon sequence.

5 *Genotype*: As used herein, the term “genotype” refers to the genetic constitution of an organism. More specifically, the term refers to the identity of alleles present in an individual. “Genotyping” of an individual or a DNA sample refers to identifying the nature, in terms of nucleotide base, of the two alleles possessed by an individual at a known polymorphic site.

10 *Heterozygous*: As used herein, the term “heterozygous” or “HET” refers to an individual possessing two different alleles of the same gene. As used herein, the term “heterozygous” encompasses “compound heterozygous” or “compound heterozygous mutant.” As used herein, the term “compound heterozygous” refers to an individual possessing two different alleles. As used herein, the term “compound heterozygous mutant”
15 refers to an individual possessing two different copies of an allele, such alleles are characterized as mutant forms of a gene. The term “mutant” as used herein refers to a mutated, or potentially non-functional form of a gene. (See “mutations of the CFTR gene.”)

Homozygous: As used herein, the term “homozygous” refers to an individual possessing two copies of the same allele. As used herein, the term “homozygous mutant”
20 refers to an individual possessing two copies of the same allele, such allele being characterized as the mutant form of a gene. The term “mutant” as used herein refers to a mutated, or potentially nonfunctional form of a gene.

Hybridize: As used herein, the term “hybridize” or “hybridization” refers to a process where two complementary nucleic acid strands anneal to each other under appropriately
25 stringent conditions. Oligonucleotides or probes suitable for hybridizations typically contain 10-100 nucleotides in length (e.g., 18- 50, 12-70, 10-30, 10-24, 18-36 nucleotides in length). Nucleic acid hybridization techniques are well known in the art. See, e.g., Sambrook, et al., 1989, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Press, Plainview, N.Y. Those skilled in the art understand how to estimate and adjust the stringency
30 of hybridization conditions such that sequences having at least a desired level of complementary will stably hybridize, while those having lower complementary will not. For examples of hybridization conditions and parameters, see, e.g., Sambrook, et al., 1989, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Press,

Plainview, N.Y.; Ausubel, F. M. et al. 1994, Current Protocols in Molecular Biology. John Wiley & Sons, Secaucus, N.J.

Insertion or addition: As used herein, the term “insertion” or “addition” refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, as compared to the naturally occurring molecule.

In vitro: As used herein, the term “*in vitro*” refers to events that occur in an artificial environment, *e.g.*, in a test tube or reaction vessel, in cell culture, *etc.*, rather than within a multi-cellular organism.

In vivo: As used herein, the term “*in vivo*” refers to events that occur within a multi-cellular organism such as a non-human animal.

Isolated: As used herein, the term “isolated” refers to a substance and/or entity that has been (1) separated from at least some of the components with which it was associated when initially produced (whether in nature and/or in an experimental setting), and/or (2) produced, prepared, and/or manufactured by the hand of man. Isolated substances and/or entities may be separated from at least about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, about 98%, about 99%, substantially 100%, or 100% of the other components with which they were initially associated. In some embodiments, isolated agents are more than about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, substantially 100%, or 100% pure. As used herein, a substance is “pure” if it is substantially free of other components. As used herein, the term “isolated cell” refers to a cell not contained in a multi-cellular organism.

Labeled: The terms “labeled” and “labeled with a detectable agent or moiety” are used herein interchangeably to specify that an entity (*e.g.*, a nucleic acid probe, antibody, *etc.*) can be visualized, for example following binding to another entity (*e.g.*, a nucleic acid, polypeptide, *etc.*). The detectable agent or moiety may be selected such that it generates a signal which can be measured and whose intensity is related to (*e.g.*, proportional to) the amount of bound entity. A wide variety of systems for labeling and/or detecting proteins and peptides are known in the art. Labeled proteins and peptides can be prepared by incorporation of, or conjugation to, a label that is detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical, chemical or other means. A label or labeling moiety may be directly detectable (*i.e.*, it does not require any further reaction or manipulation to be detectable, *e.g.*, a fluorophore is directly detectable) or it may be

indirectly detectable (*i.e.*, it is made detectable through reaction or binding with another entity that is detectable, e.g., a hapten is detectable by immunostaining after reaction with an appropriate antibody comprising a reporter such as a fluorophore). Suitable detectable agents include, but are not limited to, radionucleotides, fluorophores, chemiluminescent agents,
 5 microparticles, enzymes, colorimetric labels, magnetic labels, haptens, molecular beacons, aptamer beacons, and the like.

Multiplex PCR: As used herein, the term “multiplex PCR” refers to amplification of two or more regions which are each primed using a distinct primers pair.

Multiplex ASPE: As used herein, the term “multiplex ASPE” refers to an assay
 10 combining multiplex PCR and allele specific primer extension for detecting polymorphisms. Typically, multiplex PCR is used to first amplify regions of DNA that will serve as target sequences for ASPE primers. See the definition of allele specific primer extension.

Mutations of the CFTR gene: As used herein, the term “mutations of the CFTR gene” refers to one or more abnormal nucleic acid sequences as compared to a wild-type CFTR
 15 gene sequence. The “mutations of the CFTR gene” are also referred to as “mutant CF sequences.” Mutations of the CFTR gene encompass substitutions (e.g., single nucleotide polymorphisms (SNP)), deletions, insertions, additions, and/or duplications.

Primer: As used herein, the term “primer” refers to a short single-stranded oligonucleotide capable of hybridizing to a complementary sequence in a nucleic acid
 20 sample. Typically, a primer serves as an initiation point for template dependent DNA synthesis. Deoxyribonucleotides can be added to a primer by a DNA polymerase. In some embodiments, such deoxyribonucleotides addition to a primer is also known as primer extension. The term primer, as used herein, includes all forms of primers that may be synthesized including peptide nucleic acid primers, locked nucleic acid primers,
 25 phosphorothioate modified primers, labeled primers, and the like. A “primer pair” or “primer set” for a PCR reaction typically refers to a set of primers typically including a “forward primer” and a “reverse primer.” As used herein, a “forward primer” refers to a primer that anneals to the anti-sense strand of dsDNA. A “reverse primer” anneals to the sense-strand of dsDNA.

Polymorphism: As used herein, the term “polymorphism” refers to the coexistence of
 30 more than one form of a gene or portion thereof.

Pure or substantially pure: As used herein, the term “pure or substantially pure” refers to a compound, e.g., a protein or polypeptide that has been separated from components which naturally accompany it. Typically, a compound is substantially pure when at least

10%, more preferably at least 20%, more preferably at least 50%, more preferably at least 60%, more preferably at least 75%, more preferably at least 90%, and most preferably at least 99% of the total material (by volume, by wet or dry weight, or by mole percent or mole fraction) in a sample is the compound of interest. Purity can be measured by any appropriate method, e.g., in the case of polypeptides by column chromatography, gel electrophoresis or HPLC analysis. A compound, e.g., a protein, is also substantially purified when it is essentially free of naturally associated components or when it is separated from the native contaminants which accompany it in its natural state.

Real-time PCR: As used herein, the term “real-time PCR” refers to *quantitative real time polymerase chain reaction* (Q-PCR/qPCR/qrt-PCR) or *kinetic polymerase chain reaction* (KPCR), is a laboratory technique based on the PCR, which is used to amplify and simultaneously quantify a targeted DNA molecule. It enables both detection and quantification (as absolute number of copies or relative amount when normalized to DNA input or additional normalizing genes) of one or more specific sequences in a DNA sample.

Sense strand vs. anti-sense strand: As used herein, the term “sense strand” refers to the strand of double-stranded DNA (dsDNA) that includes at least a portion of a coding sequence of a functional protein. As used herein, the term “anti-sense strand” refers to the strand of dsDNA that is the reverse complement of the sense strand.

Specific: As used herein, the term “specific,” when used in connection with an oligonucleotide primer, refers to an oligonucleotide or primer, under appropriate hybridization or washing conditions, is capable of hybridizing to the target of interest and not substantially hybridizing to nucleic acids which are not of interest. Higher levels of sequence identity are preferred and include at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% sequence identity. In some embodiments, a specific oligonucleotide or primer contains at least 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 35, 40, 45, 50, 55, 60, 65, 70, or more bases of sequence identity with a portion of the nucleic acid to be hybridized or amplified when the oligonucleotide and the nucleic acid are aligned.

Subject: As used herein, the term “subject” refers to a human or any non-human animal. A subject can be a patient, which refers to a human presenting to a medical provider for diagnosis or treatment of a disease. A human includes pre and post natal forms. Particularly preferred subjects are humans being tested for the existence of a CFTR carrier state, CF disease or CF related disorder state.

Substantially: As used herein, the term “substantially” refers to the qualitative condition of exhibiting total or near-total extent or degree of a characteristic or property of

interest. One of ordinary skill in the biological arts will understand that biological and chemical phenomena rarely, if ever, go to completion and/or proceed to completeness or achieve or avoid an absolute result. The term “substantially” is therefore used herein to capture the potential lack of completeness inherent in many biological and chemical phenomena.

5 *Substantially complementary:* As used herein, the term “substantially complementary” refers to two sequences that can hybridize under stringent hybridization conditions. The skilled artisan will understand that substantially complementary sequences need not hybridize along their entire length. In some embodiments, “stringent hybridization
10 conditions” refer to hybridization conditions at least as stringent as the following: hybridization in 50% formamide, 5XSSC, 50 mM NaH₂PO₄, pH 6.8, 0.5% SDS, 0.1 mg/mL sonicated salmon sperm DNA, and 5XDenhart’s solution at 42 °C overnight; washing with 2XSSC, 0.1% SDS at 45 °C; and washing with 0.2XSSC, 0.1% SDS at 45 °C. In some
15 embodiments, stringent hybridization conditions should not allow for hybridization of two nucleic acids which differ over a stretch of 20 contiguous nucleotides by more than two bases.

Substitution: As used herein, the term “substitution” refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively, as compared to the naturally occurring molecule.

20 *Suffering from:* An individual who is “suffering from” a disease, disorder, and/or condition has been diagnosed with or displays one or more symptoms of the disease, disorder, and/or condition.

Susceptible to: An individual who is “susceptible to” a disease, disorder, and/or condition has not been diagnosed with the disease, disorder, and/or condition. In some
25 embodiments, an individual who is susceptible to a disease, disorder, and/or condition may not exhibit symptoms of the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition will develop the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, and/or condition will not develop the disease, disorder, and/or condition.

30 *Wild-type:* As used herein, the term “wild-type” refers to the typical or the most common form existed in nature. For example, a wild-type CFTR gene or protein refers to the typical or the most common form of CFTR gene or protein existed in a natural population. As used herein, “wild-type” is used interchangeably with “naturally-occurring.” In some embodiment, a wild-type CFTR gene or a locus thereof, refers to the CFTR gene sequence

which is found in NCBI GenBank locus ID M58478 (HUMCFTR) (SEQ ID NO:4) (FIG.4).
The CFTR gene is located on chromosome 7, which may be found in NCBI GenBank locus
AC000111 and AC000061, the contents of which are incorporated herein in their entirety by
reference. The cDNA for the CFTR gene is found in Audrezet et al., *Hum. Mutat.* (2004) 23
5 (4), 343-357.

DETAILED DESCRIPTION

The present invention provides, among other things, methods, products and systems
that use novel mutations in the cystic fibrosis transmembrane conductance regulator (CFTR)
gene in screening and/or testing for CF and CF related diseases, disorders or conditions. For
10 example, the novel mutations provided herein can be used to assist in clinical diagnosis of CF
disease, CF related disease, disorder or conditon, or carrier status and for genetic counseling
(e.g., for evaluation of an individual's risk for developing CF or being a carrier of a CFTR
mutation). The novel mutations provided herein can be used alone or in combination with
other known CFTR mutations as part of a panel of CFTR mutations.

15 Various aspects of the invention are described in detail in the following sections. The
use of sections is not meant to limit the invention. Each section can apply to any aspect of
the invention. In this application, the use of "or" means "and/or" unless stated otherwise.

Novel mutations in the CFTR gene

The CFTR gene was mapped to chromosome 7 and described in, for example, U.S.
20 Pat. No. 6,201,107 and U.S. Pat. No. 5,776,677, the disclosures of which are incorporated by
reference herein in their entirety. The CFTR genomic sequence is described in GenBank
Accession Number NC_000007 (range: 117120016..117308718; the entire contents of which
are herein incorporated by reference) (SEQ ID NO:1) (FIG. 1). The CFTR gene contains 27
exons. The exons are numbered 1, 2, 3, 4, 5, 6a, 6b, 7, 8, 9, 10, 11, 12, 13, 14a, 14b, 15, 16,
25 17a, 17b, 18, 19, 20, 21, 22, 23, and 24. The CFTR cDNA sequence is described in GenBank
Accession Number AR016032.1 (SEQ ID NO:2) (FIG. 2).

The CFTR protein is described in, for example, U.S. Pat. No. 5,543,399, the
disclosure of which is incorporated by reference herein in its entirety. The CFTR protein
sequence is also described in GenBank Accession Number AAC90840.1 (SEQ ID NO:3)
30 (FIG.3).

As described in Example I, the inventors of the present application identified various
novel mutations in the CFTR gene (Table 5). These mutations were identified by sequence
analysis of the CFTR gene in specimens submitted for clinical testing obtained from
individuals who were known to be affected with CF or likely to be a carrier because of

5 familial history, or suspected to be affected with CF based on other CF testing (see Clinical Indication listed in Table 5). The mutations were identified by comparing the CFTR gene sequence from patient samples to the wild-type CFTR gene or protein sequence (see SEQ ID NO:1-3). As shown in Table 5, patients carrying these mutations were from different ethnic groups including Caucasians, African Americans, Hispanics, and Asians. Thus, these mutations may be particularly useful for developing more effective genetic testing for patients from non-Caucasian racial groups.

10 Novel mutations described herein are located in introns (e.g., intron 3, intron 6a, intron 11, intron 14a, intron 19, intron 20, intron 21, and intron 23) and exons (e.g., exon 2, exon 3, exon 4, exon 5, exon 6a, exon 6b, exon 7, exon 9, exon 10, exon 11, exon 12, exon 13, exon 14a, exon 14b, exon 15, exon 16, exon 17a, exon 17b, exon 19, exon 20, exon 21, exon 22, and exon 24). Some of the novel mutations are nonsense mutations, i.e., mutations that result in a stop codon. Some of the novel mutations are missense mutations, i.e., mutations that result in amino acid substitutions. Some of the novel mutations cause in-frame insertions and/or deletions. Some of the novel mutations delete one or more nucleotides in such a manner as to lead to a shift in the reading frame. Some of the novel mutations alter the sequence at a splice junction, for example, consensus splice site ag/gt or other splice sites. Thus, most of the novel mutations described herein are likely to disrupt CFTR gene or protein expression or function.

20 The “ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007” (Richards SC et al. Genetics in Medicine, 10:294-300, which is incorporated herein by reference), provides interpretive categories and definitions of sequence variations which can be used, along with additional test results and clinical information to classify the novel mutations described herein into the following groups.

25 **Group I:** Patient has a novel sequence change that can be classified as category 2 according to the ACMG guidelines (i.e., nonsense, frame shift (FS), consensus splice site ag/gt). Patient has another well established CF disease causing mutation (i.e. F508, W1282X, etc). Patient indication is suspected of having CF, known to be affected with CF or identified through newborn screening. The Group I mutations, which are of particular interest, are shown in Table 1 and these mutations are expected to cause CF or CF related diseases, disorders or conditions.

30 **Group IIA:** Patient has novel sequence change that can be classified as category 3 according to the ACMG guidelines (i.e., missense, in-frame ins/del, other splice site mutations, etc). Patient has another well established CF disease causing mutation (i.e., F508,

W1282X, etc.). Patient is suspected of having CF, known to be affected with CF, or identified through positive newborn screening. The Group IIA mutations are shown in Table 2 (under subsection Group IIA).

5 **Group IIB:** Patent has a novel sequence change that can be classified as category 3 according to the ACMG guidelines (i.e., missense, in-frame ins/del, other splice site mutations, etc). Patient is suspected of having CF, known to be affected with CF, or identified through positive newborn screening. The Group IIB mutations are shown in Table 2 (under subsection Group IIB).

10 **Group III:** Patient has novel sequence change that can be classified as category 3 according to the ACMG guidelines (i.e., missense, in-frame insertions/deletions, other splice site mutations, etc). Patient has another well established CF disease causing mutation (i.e. F508, W1282X, etc). Patient indication is suspected of having CF, known to be affected with CF, or identified through newborn screening. Patient has an additional change(s) of unknown clinical significance. The Group III mutations are shown in Table 3.

15 **Group IV:** Mutations other than the Group I, II, and III mutations identified above. The Group IV mutations are shown in Table 4.

20 Novel CFTR mutations according to the invention however are not limited to the specific nucleotide or amino acid variations identified in Tables 1-4 and should encompass any abnormal nucleotides or amino acid residues, as compared to the wild-type CFTR gene or protein sequences, that may be present at any of the positions identified in Tables 1-4.

WO 2011/119531

PCT/US2011/029314

Table 1. Group I mutations

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
1824delA	Frameshift (FS)	n/a	Caucasian	F508del	Mutation was identified in a 22 year old patient with a known diagnosis of CF. This patient carried a second mutation known to cause CF (F508del).	e12
2957delT	FS	n/a	Caucasian	F508del	Mutation was identified in a 1 year old patient with a known diagnosis of CF. The patient carried a second mutation known to cause CF (F508del).	e15
4089ins4	FS	n/a	Caucasian	F508del	Mutation was identified in a 7 year old patient with a known diagnosis of CF. The patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	e21
4374+2T >C	Splice site mutation	n/a	1. Caucasian 2. Caucasian	1. F508del 2. F508del	Patient #1: Mutation was identified in a 45 year old patient with a suspected diagnosis of CF. The patient carried a second mutation known to cause CF (F508del). Patient #2: Mutation was identified in a 52 year old patient with a suspected diagnosis of CF. The patient carried a second mutation known to cause CF (F508del).	i23
3064A>T	Nonsense	K978X	African American	Q1042X	Mutation was identified in a 26 year old patient with a known diagnosis of CF. The patient carried a second mutation likely to cause CF Q1042X.	e16
246C>G	Nonsense	Y38X	Caucasian	F508del	Mutation was identified in a 1 month old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	e2
Group IIA						
269C>T	Missense (MS)	A46V	1) Caucasian 2) Black 3) African	1) 3849+1219 2G>A 2) F508del 3) none	Patient #1: Mutation was identified in a 32 year old patient who was tested due to abnormalities found on fetal ultrasound. The patient carried a second mutation of unknown clinical significance (3849+12192G>A). Patient #2: Mutation was identified in a 2 month old patient who was tested based on follow-up for a positive newborn screen. The patient carried a second mutation known to	e2

WO 2011/119531

PCT/US2011/029314

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
			American		cause cystic fibrosis (F508del). Patient #3: Mutation was identified in a 24 year old patient who was tested as a parental follow-up to a positive newborn screen.	
2902 G>T	MS	D924Y	Caucasian	F508del	Mutation was identified in a 1 month old patient with a suspected diagnosis of CF. The patient also had a positive sweat chloride test and carried a second mutation known to cause CF (F508del)	e15
3814G>A	MS	E1228K	Caucasian	F508del	Mutation was identified in a 1 month old patient with a suspected diagnosis of CF. The patient had a borderline sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	e19
502G>C	MS	G124R	Not Provided	F508del	Mutation was identified in a 2 month old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	e4
1520G>T	MS	G463V	Caucasian	F508del	Mutation was identified in a 17 year old patient with a known diagnosis of CF. Patient carried a second mutation known to cause CF (F508del).	e9
511_513dup TTA	In frame duplication	L1127dup	Caucasian, Asian	W1282X	Mutation was identified in a newborn with a suspected diagnosis of CF. The patient had clinical symptoms of CF including as a positive sweat chloride test, meconium ileus, echogenic bowel, and pancreatic insufficiency. The patient carried a second mutation known to cause CF (W1282X).	e4
978A>T	MS	F282D	1. not provided 2. not provided	1. 3120+1G>A 2) none	Patient #1: Mutation was identified in a 10 year old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (3120+1G>A). Patient #2: Mutation was identified in a 4 year old patient with a suspected diagnosis of CF and a family history of CF.	e6b
843G>C	MS	Q237H	Caucasian	F508del	Mutation was identified in a 2 month old patient with a known diagnosis of CF. The patient carried a second mutation known to cause CF (F508del).	e6a
829C>T	MS	L233F	Caucasian	D1152H	Mutation was identified in a 1 month old patient who was tested following a positive	e6a

WO 2011/119531

PCT/US2011/029314

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
					newborn screen. The patient carried a second mutation known to cause CF (D1152H).	
4096-6C>T	Splice site mutation	None	Caucasian	F508del	Mutation was identified in a 58 year old patient with a suspected diagnosis of CF. The patient carried a second mutation known to cause CF (F508del).	i21
4375-7delT	Splice site mutation	None	Caucasian	F508del	Mutation was identified in a 6 year old patient with a suspected diagnosis of CF. Patient has a family history, a borderline sweat chloride test and recurrent pneumonia. The patient carried a second mutation known to cause CF (F508del).	i23
1586G>C	MS	S485T	Caucasian	S1235R	Mutation was identified in a 2 year old patient with a suspected diagnosis of CF. The patient carried a second mutation S1235R (3837T>G) which has been reported in individuals with varying CF phenotypes.	e10
Group IIB						
875+4G>T	Splice site mutation	n/a	African American	none	Mutation was identified in a 1 month old patient who had a positive newborn screening test.	i6a
4005+3G>T	Splice site mutation	n/a	Caucasian	none	Mutation was identified in a 40 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	i20

Table 3. Group III Mutations

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
2711T>C	MS	1860T	Caucasian	F508del, E528E	Mutation was identified in a 58 year old woman with a suspected diagnosis of CF. The patient carried a second mutation known to cause CF (F508del) and an additional mutation of unknown clinical significance (E528E)..	e14a
3891G>C	MS	L1253F	Not provided	G85E, L15P	Mutation was identified in a 32 year old patient with a known diagnosis of CF. The patient carried a second mutation known to cause CF (G85E) and an additional mutation of unknown clinical significance (L15P).	e20
2524C>T	MS	P798S	African American	F508del, R74W, G921E, D1270N	Mutation was identified in a 5 year old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. This patient carried a second mutation known to cause CF (F508del) and three additional mutations of unknown clinical significance (R74W, G921E, D1270N).	e13
2894G>A	MS	G921E	African American	F508del, R74W, P798S, D1270N	Mutation was identified in a 5 year old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. This patient carried a second mutation known to cause CF (F508del) and three additional mutations of unknown clinical significance (R74W, P789S, D1270N).	e15

Table 4 Group IV Mutations

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
405+10247C >T	Possible splice site mutation	n/a	Caucasian	F508del	Mutation was identified in a 35 year old patient who was tested to determine if they were a carrier, there was no family history of CF. This patient carried a second mutation known to cause CF (F508del)	i3
405+10255d eIC	Possible splice site mutation	n/a	Not Provided	F508del, 124del23bp	Mutation was identified in a 10 year old patient. The patient carries two mutations know to cause CF (F508del and 124del23).	i3
1811+1643 G>T	Possible splice site mutation	n/a	1. Hispanic	1. F508del	Patient #1: Mutation was identified in a 1 year old patient with a known diagnosis of CF. Patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	i11

WO 2011/119531

PCT/US2011/029314

			2. Hispanic	2. F508del	Patient #2: Mutation was identified in a 6 year old patient with a known diagnosis of CF. The patient carried a second mutation know to cause CF (F508del).	
			3. Not provided	3. none	Patient #3: Mutation was identified in an 8 month old patient with a suspected diagnosis of CF.	
1812-13A>G	Splice site mutation	n/a	Caucasian	none	Mutation was identified in a 15 year old patient with a suspected diagnosis of CF. The patient has chronic sinusitis.	i11
2752-33insA	Possible splice site mutation	n/a	African American	F693L	Mutation was identified in a 6 year old patient with a known diagnosis of CF. The patient carries a second mutation of unknown clinical significance (F693L).	i14a
3849+12192 G>A	Possible splice site mutation	n/a	Caucasian	A46V	Mutation was identified in a 32 year old patient who was tested due to abnormalities found on fetal ultrasound. The patient carried an additional mutation of known clinical significance (A46V).	i19
724G>A	MS	A198T	Hispanic	none	Mutation was identified in a 4 month old patient with a suspected diagnosis of CF.	e6a
3899C>T	MS	A1256 V	Guyanese	none	Mutation was identified in a 45 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	e20
3986C>T	MS	A1285 V	Not Provided	none	Mutation was identified in a 23 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	e20
901G>A	MS	E257K	Hispanic	none	Mutation was identified in a 4 year old patient with a suspected diagnosis of CF. The patient has asthma and recurring pneumonia.	e6b
392 T>C	MS	F87S	Not Provided	none	The mutation was identified in a 1 month old patient with a suspected diagnosis of CF.	e3
3463T>C	MS	F1111L	Hispanic	none	Mutation was identified in a 6 year old patient with a suspected diagnosis of CF. The patient has asthma.	e17b
1757G>A	MS	G542E	Hispanic	none	Mutation was identified in a 25 year old patient who was tested to determine if they were a carrier, there was no family history of CF. The patient carried 2 copies of G542E..	e11
4025G>C	MS	G1298 A	Asian	G970D, Q1352H	Mutation was identified in a 34 year old patient with congenital absence of the vas deferens. The patient carried two other mutations of unknown clinical significance (G970D and Q1352H)	e21
4129G>T	MS	G1333 W	Not Provided	none	Mutation was identified in an 8 year old patient with a suspected diagnosis of CF. Patient had	e22

WO 2011/119531

PCT/US2011/029314

					recurrent respiratory infections and chronic cough.	
663T>G	MS	I177M	Caucasian	none	Mutation was identified in a 34 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	e5
3200I>C	MS	I1023T	Hispanic	none	Mutation was identified in a 34 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	e17a
4412 T>C	MS	I1427T	Asian	S1444S	Mutation was identified in a 34 year old patient who was tested to determine if they were a carrier, there was no family history of CF. The patient carried another mutation that is considered likely to be clinically benign (S144S).	e24
620A>C	MS	K163T	Caucasian	none	Mutation was identified in a 32 year old patient with a family history of CF.	e4
1738 A>G	MS	K536E	Hispanic	I488I	Mutation was identified in a 19 year old patient who's son had a positive newborn screening test. The patient carried another mutation that is considered likely to be clinically benign (I1488I)..	e11
3370A>C	MS	K1080 Q	Caucasian	none	Mutation was identified in a 9 year old patient with a suspected diagnosis of CF. The patient had asthma and failure to thrive.	e17b
1129 C>T	MS	L333F	Asian	none	Mutation was identified in a 37 year old patient who tested to determine if they were a carrier. there was no family history of CF.	e7
2383C>T	MS	R751C	Caucasian	2183delA A >G	Mutation was identified in a 36 year old patient who was being tested due to a partner being a CF carrier. The patient also carried a second mutation known to cause CF (2183delAA>G).	e13
2761delT CT	In frame deletion	S877del	Caucasian	F508del, D1152H	Mutation was identified in 1 month old patient who had a positive sweat chloride test. The patient carried two additional mutations known to cause CF (F508del and D1152H).	e14b
1106A>G	MS	Y325C	Caucasian	R334W	Mutation was identified in a 35 year old patient who was tested to determine if they were a carrier, there was no family history of CF. Patient carried a second mutation known to cause CF (R334W).	e7
622A>G	MS	T164A	Caucasian	none	Mutation was identified in a 3 month old patient with a suspected diagnosis of CF.	e5

Detection of CFTR mutations

A variety of methods known in the art can be used to detect CFTR gene mutations disclosed in the present invention. For example, methods that have been used to detect previously identified CFTR gene mutations have been described and are adaptable for use with the present invention. See e.g., Audrezet et al., "Genomic rearrangements in the CFTR gene: extensive allelic heterogeneity and diverse mutational mechanisms" *Hum Mutat.* 2004 April; 23(4):343-57; PCT WO 2004/040013 A1 and corresponding US application No. 20040110138; titled "Method for the detection of multiple genetic targets" by Spiegelman and Lem; US patent application No. 20030235834; titled "Approaches to identify cystic fibrosis" by Dunlop et al.; and US patent application No. 20040126760 titled "Novel compositions and methods for carrying out multiple PCR reactions on a single sample" by N. Broude, the entire contents of each of which are herein incorporated by reference.

Nucleic acid analyses

In certain embodiments, CFTR gene mutations disclosed herein are detected at the nucleic acid level. For example, nucleic acid can be analyzed by sequencing, hybridization, PCR amplification, restriction enzyme digestion, primer extension such as single-base primer extension or multiplex allele-specific primer extension (ASPE).

Nucleic acid analyses can be performed on genomic DNA, messenger RNAs, and/or cDNA. In many embodiments, nucleic acids are extracted from a biological sample. In some embodiments, nucleic acids are analyzed without having been amplified. In some embodiments, nucleic acids are amplified using techniques known in the art (such as polymerase chain reaction (PCR)) and amplified nucleic acids are used in subsequent analyses. Multiplex PCR, in which several amplicons (e.g., from different genomic regions) are amplified at once using multiple sets of primer pairs, may be employed. Additionally, methods such as real-time PCR, as are known in the art, may be used to perform nucleic acid analysis.

In some embodiments, nucleic acids are amplified in a manner such that the amplification product for a wild-type allele differs in size from that of a mutant allele. Thus, presence or absence of a particular mutant allele can be determined by detecting size differences in the amplification products, e.g., on an electrophoretic gel. For example, deletions or insertions of CFTR gene regions may be particularly amenable to using size-based approaches.

Certain exemplary nucleic acid analysis methods are described in detail below.

Allele-specific amplification

In some embodiments, CFTR gene mutations are detected using an allele-specific amplification assay. This approach is variously referred to as PCR amplification of specific allele (PASA) (Sarkar, et al., 1990 *Anal. Biochem.* 186:64-68), allele-specific amplification
5 (ASA) (Okayama, et al., 1989 *J. Lab. Clin. Med.* 114:105-113), allele-specific PCR (ASPCR) (Wu, et al. 1989 *Proc. Natl. Acad. Sci. USA.* 86:2757-2760), and amplification-refractory mutation system (ARMS) (Newton, et al., 1989 *Nucleic Acids Res.* 17:2503-2516). The entire contents of each of these references is incorporated herein. This method is applicable for single base substitutions as well as micro deletions/insertions.

10 For example, for PCR-based amplification methods, amplification primers may be designed such that they can distinguish between different alleles (e.g., between a wild-type allele and a mutant allele). Thus, the presence or absence of amplification product can be used to determine whether a CFTR gene mutation is present in a given nucleic acid sample. In some
15 embodiments, allele specific primers can be designed such that the presence of amplification product is indicative of a CFTR gene mutation. In some embodiments, allele specific primers can be designed such that the absence of amplification product is indicative of a CFTR gene mutation.

In some embodiments, two complementary reactions are used. One reaction employs a primer specific for the wild type allele ("wild-type-specific reaction") and the other reaction
20 employs a primer for the mutant allele ("mutant-specific reaction"). The two reactions may employ a common second primer. PCR primers specific for a particular allele (e.g., the wild-type allele or mutant allele) generally perfectly match one allelic variant of the target, but are mismatched to other allelic variant (e.g., the mutant allele or wild-type allele). The mismatch may be located at/near the 3' end of the primer, leading to preferential amplification of the
25 perfectly matched allele. Whether an amplification product can be detected from one or in both reactions indicates the absence or presence of the mutant allele. Detection of an amplification product only from the wild-type-specific reaction indicates presence of the wild-type allele only (e.g., homozygosity of the wild-type allele). Detection of an amplification product in the mutant-specific reaction only indicates presence of the mutant allele only (e.g. homozygosity of the
30 mutant allele). Detection of amplification products from both reactions indicate (e.g., a

heterozygote). As used herein, this approach will be referred to as “allele specific amplification (ASA).”

Allele-specific amplification can also be used to detect duplications, insertions, or inversions by using a primer that hybridizes partially across the junction. The extent of junction overlap can be varied to allow specific amplification.

Amplification products can be examined by methods known in the art, including by visualizing (e.g., with one or more dyes) bands of nucleic acids that have been migrated (e.g., by electrophoresis) through a gel to separate nucleic acids by size.

Allele-specific primer extension

In some embodiments, an allele-specific primer extension (ASPE) approach is used to detect CFTR gene mutations. ASPE employs allele-specific primers that can distinguish between alleles (e.g., between a mutant allele and a wild-type allele) in an extension reaction such that an extension product is obtained only in the presence of a particular allele (e.g., mutant allele or wild-type allele). Extension products may be detectable or made detectable, e.g., by employing a labeled deoxynucleotide in the extension reaction. Any of a variety of labels are compatible for use in these methods, including, but not limited to, radioactive labels, fluorescent labels, chemiluminescent labels, enzymatic labels, etc. In some embodiments, a nucleotide is labeled with an entity that can then be bound (directly or indirectly) by a detectable label, e.g., a biotin molecule that can be bound by streptavidin-conjugated fluorescent dyes. In some embodiments, reactions are done in multiplex, e.g., using many allele-specific primers in the same extension reaction.

In some embodiments, extension products are hybridized to a solid or semi-solid support, such as beads, matrix, gel, among others. For example, the extension products may be tagged with a particular nucleic acid sequence (e.g., included as part of the allele-specific primer) and the solid support may be attached to an “anti-tag” (e.g., a nucleic acid sequence complementary to the tag in the extension product). Extension products can be captured and detected on the solid support. For example, beads may be sorted and detected. One such system that can be employed in this manner is the LUMINEX™ MAP system, which can be adapted for cystic fibrosis mutation detection by Luminex Corporation and is sold commercially as a universal bead array (TAG-IT™) (See, e.g., Example 2)

Additional ASPE methods and reagents are described in, e.g., U.S. patent publication number 2008/0138803 A1, the entire contents of which are herein incorporated by reference.

Single nucleotide primer extension

In some embodiments, a single nucleotide primer extension (SNUPE) assay is used, in which the primer is designed to be extended by only one nucleotide. In such methods, the identity of the nucleotide just downstream (e.g., 3') of the 3' end of the primer is known and differs in the mutant allele as compared to the wild-type allele. SNUPE can be performed using an extension reaction in which the only one particular kind of deoxynucleotide is labeled (e.g., labeled dATP, labeled dCTP, labeled dGTP, or labeled dTTP). Thus, the presence of a detectable extension product can be used as an indication of the identity of the nucleotide at the position of interest (e.g., the position just downstream of the 3' end of the primer), and thus as an indication of the presence or absence of a mutation at that position. SNUPE can be performed as described in U.S. Pat. No. 5,888,819; U.S. Pat. No. 5,846,710; U.S. Pat. No. 6,280,947; U.S. Pat. No. 6,482,595; U.S. Pat. No. 6,503,718; U.S. Pat. No. 6,919,174; Piggee, C. et al. *Journal of Chromatography A* 781 (1997), p. 367-375 ("Capillary Electrophoresis for the Detection of Known Point Mutations by Single-Nucleotide Primer Extension and Laser-Induced Fluorescence Detection"); Hoogendoorn, B. et al., *Human Genetics* (1999) 104:89-93, ("Genotyping Single Nucleotide Polymorphism by Primer Extension and High Performance Liquid Chromatography"), the entire contents of each of which are herein incorporated by reference.

In some embodiments, primer extension can be combined with mass spectrometry for accurate and fast detection of the presence or absence of a mutation. See, U.S. Pat. No. 5,885,775 to Haff et al. (analysis of single nucleotide polymorphism analysis by mass spectrometry); U.S. Pat. No. 7,501,251 to Koster (DNA diagnosis based on mass spectrometry); the teachings of both of which are incorporated herein by reference. Suitable mass spectrometric format includes, but is not limited to, Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform, and combinations thereof.

Oligonucleotide ligation assay

In some embodiments, an oligonucleotide ligation assay ("OLA" or "OL") is used. OLA employs two oligonucleotides that are designed to be capable of hybridizing to abutting sequences of a single strand of a target molecules. Typically, one of the oligonucleotides is

biotinylated, and the other is detectably labeled, e.g., with a streptavidin-conjugated fluorescent moiety. If the precise complementary sequence is found in a target molecule, the oligonucleotides will hybridize such that their termini abut, and create a ligation substrate that can be captured and detected. See e.g., Nickerson et al. (1990) Proc. Natl. Acad. Sci. U.S.A. 87:8923-8927, Landegren, U. et al. (1988) Science 241:1077-1080 and U.S. Pat. No. 4,998,617, the entire contents of which are herein incorporated by reference in their entirety.

Hybridization approach

In some embodiments, nucleic acids are analyzed by hybridization using one or more oligonucleotide probes specific for a region in the CFTR gene (SEQ ID NO:1) corresponding to the one or more mutations selected from Table 1, 2, 3 or 4, and under conditions sufficiently stringent to disallow a single nucleotide mismatch. In certain embodiments, suitable nucleic acid probes can distinguish between a normal CFTR gene and a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4. For example, suitable nucleic acid probes specifically bind to a normal CFTR gene but not to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4. Alternatively, nucleic acid probes specifically bind to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4 but not to a normal CFTR gene. Probes of the present invention include those that are capable of specifically hybridizing a mutant CFTR allele containing one or more mutations listed in Tables 1, 2, 3, or 4. Probes of the present invention also include those that are capable of specifically hybridizing a normal allele in a particular region of the CFTR gene and therefore capable of distinguishing a normal allele from a mutant CFTR allele containing one or more mutations listed in Tables 1, 2, 3, or 4. Thus, for example, one of ordinary skill in the art could use probes of the invention to determine whether an individual is homozygous or heterozygous for a particular allele.

Nucleic acid hybridization techniques are well known in the art. See, e.g., Sambrook, et al., 1989, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Press, Plainview, N.Y. Those skilled in the art understand how to estimate and adjust the stringency of hybridization conditions such that sequences having at least a desired level of complementary will stably hybridize, while those having lower complementary will not. For examples of hybridization conditions and parameters, see, e.g., Sambrook, et al., 1989, Molecular Cloning: A

Laboratory Manual, Second Edition, Cold Spring Harbor Press, Plainview, N.Y.; Ausubel, F. M. et al. 1994, Current Protocols in Molecular Biology. John Wiley & Sons, Secaucus, N.J.

5 In some embodiments, probe molecules that hybridize to the mutant or wildtype CFTR sequences can be used for detecting such sequences in the amplified product by solution phase or, more preferably, solid phase hybridization. Solid phase hybridization can be achieved, for example, by attaching the CFTR probes to a microchip.

10 Nucleic acid probes may comprise ribonucleic acids and/or deoxyribonucleic acids. In some embodiments, provided nucleic acid probes are oligonucleotides (i.e., "oligonucleotide probes"). Generally, oligonucleotide probes are long enough to bind specifically to a homologous region of the CFTR gene, but short enough such that a difference of one nucleotide between the probe and the nucleic acid sample being tested disrupts hybridization. Typically, the sizes of oligonucleotide probes vary from approximately 10 to 100 nucleotides. In some
15 embodiments, oligonucleotide probes vary from 15 to 90, 15 to 80, 15 to 70, 15 to 60, 15 to 50, 15 to 40, 15 to 35, 15 to 30, 18 to 30, or 18 to 26 nucleotides in length. As appreciated by those of ordinary skill in the art, the optimal length of an oligonucleotide probe may depend on the particular methods and/or conditions in which the oligonucleotide probe may be employed.

In some embodiments, nucleic acid probes are useful as primers, e.g., for nucleic acid amplification and/or extension reactions.

20 In some embodiments, nucleic acid probes are labeled with a detectable moiety as described herein.

Arrays

A variety of the methods mentioned herein may be adapted for use with arrays that allow sets of mutations to be analyzed and/or detected in a single experiment. For example, multiple novel CFTR mutations described herein (e.g., Tables 1, 2, 3 or 4) can be analyzed at the same
25 time. Additionally or alternatively, one or more novel CFTR mutations described herein (e.g., Tables 1, 2, 3 or 4) can be analyzed together with other CFTR mutations known in the art at the same time. In particular, methods that involve use of nucleic acid reagents (e.g., probes, primers, oligonucleotides, etc.) are particularly amenable for adaptation to an array-based platform (e.g., microarray). In some embodiments, an array containing one or more probes specific for
30 detecting CFTR mutations described herein (e.g., Tables 1, 2, 3 or 4) can be designed and adapted for various methods described herein. Additionally or alternatively, probes specific for

detecting CFTR mutations described herein (e.g., Tables 1, 2, 3 or 4) can be combined with probes specific for CFTR mutations known in the art. In some embodiments, an array containing multiple probes are known as a mutation panel. See, e.g., Wall et al. "A 31-mutation assay for cystic fibrosis testing in the clinical molecular diagnostics laboratory," *Human Mutation*, 1995;5(4):333-8, the entire contents of which are herein incorporated by reference. Other methods may include the use of real-time PCR with probes for detecting CFTR mutations as described herein.

Protein-based analyses

In certain embodiments, CFTR mutations are detected at the protein (or peptide or polypeptide level), that is, a gene product from a CFTR gene mutation is analyzed. For example, CFTR protein or fragment thereof can be analyzed by amino acid sequencing methods, or immuno assays using one or more antibodies that specifically recognize one or more epitopes corresponding to the one or more novel mutations described herein (e.g., Table 1, 2, 3 and 4). CFTR proteins can also be analyzed by protease digestion (e.g., trypsin digestion) and, in some embodiments, the digested protein products can be further analyzed by 2D-gel electrophoresis.

Antibody detection of mutant proteins

For example, specific antibodies that can differentiate between a normal CFTR protein and a mutant CFTR protein can be employed in any of a variety of methods known in the art to detect CFTR mutations. In certain embodiments, suitable antibodies can distinguish between a normal CFTR protein and a mutant CFTR protein containing one or mutations selected from Tables 1, 2, 3, or 4. For example, suitable antibodies specifically bind to a normal CFTR protein but not to a mutant CFTR protein containing one or more mutations selected from Table 1, 2, 3, or 4. Alternatively, suitable antibodies specifically bind to a mutant CFTR protein containing one or more mutations selected from Table 1, 2, 3, or 4 but not to a normal CFTR protein.

Antibodies against particular epitopes, polypeptides, and/or proteins (e.g., mutant or normal CFTR proteins) can be generated using any of a variety of known methods in the art. For example, the epitope, polypeptide, or protein against which an antibody is desired can be produced and injected into an animal, typically a mammal (such as a donkey, mouse, rabbit, horse, chicken, etc.), and antibodies produced by the animal can be collected from the animal. Monoclonal antibodies can also be produced by generating hybridomas that express an antibody of interest with an immortal cell line. For more details on methods of producing, and uses of,

antibodies to detect CFTR mutants, see, e.g., U.S. Patent No. 5,776,677, the entire contents of which are herein incorporated by reference.

In some embodiments, antibodies are labeled with a detectable moiety as described herein.

5 Antibody detection methods are well known in the art including, but are not limited to, enzyme-linked immunosorbent assays (ELISAs) and Western blots. Some such methods are amenable to being performed in an array format. For example, a variety of different antibodies, each of which is specific for different epitopes within the CFTR protein, could be immobilized in an array and used in an assay such as an ELISA.

10 ***Detectable moieties***

In certain embodiments, certain molecules (e.g., nucleic acid probes, antibodies, etc.) used in accordance with and/or provided by the invention comprise one or more detectable entities or moieties, i.e., such molecules are “labeled” with such entities or moieties.

15 Any of a wide variety of detectable agents can be used in the practice of the present invention. Suitable detectable agents include, but are not limited to: various ligands, radionuclides; fluorescent dyes; chemiluminescent agents (such as, for example, acridinium esters, stabilized dioxetanes, and the like); bioluminescent agents; spectrally resolvable inorganic fluorescent semiconductors nanocrystals (i.e., quantum dots); microparticles; metal nanoparticles (e.g., gold, silver, copper, platinum, etc.); nanoclusters; paramagnetic metal ions; enzymes; 20 colorimetric labels (such as, for example, dyes, colloidal gold, and the like); biotin; dioxigenin; haptens; and proteins for which antisera or monoclonal antibodies are available.

In some embodiments, the detectable moiety is biotin. Biotin can be bound to avidins (such as streptavidin), which are typically conjugated (directly or indirectly) to other moieties (e.g., fluorescent moieties) that are detectable themselves.

25 Below are described some non-limiting examples of other detectable moieties.

Fluorescent dyes

In certain embodiments, a detectable moiety is a fluorescent dye. Numerous known fluorescent dyes of a wide variety of chemical structures and physical characteristics are suitable for use in the practice of the present invention. A fluorescent detectable moiety can be 30 stimulated by a laser with the emitted light captured by a detector. The detector can be a charge-coupled device (CCD) or a confocal microscope, which records its intensity.

Suitable fluorescent dyes include, but are not limited to, fluorescein and fluorescein dyes (e.g., fluorescein isothiocyanine or FITC, naphthofluorescein, 4',5'-dichloro-2',7'-dimethoxyfluorescein, 6-carboxyfluorescein or FAM, etc.), carbocyanine, merocyanine, styryl dyes, oxonol dyes, phycoerythrin, erythrosin, eosin, rhodamine dyes (e.g., carboxytetramethyl-rhodamine or TAMRA, carboxyrhodamine 6G, carboxy-X-rhodamine (ROX), lissamine rhodamine B, rhodamine 6G, rhodamine Green, rhodamine Red, tetramethylrhodamine (TMR), etc.), coumarin and coumarin dyes (e.g., methoxycoumarin, dialkylaminocoumarin, hydroxycoumarin, aminomethylcoumarin (AMCA), etc.), Oregon Green Dyes (e.g., Oregon Green 488, Oregon Green 500, Oregon Green 514., etc.), Texas Red, Texas Red-X, SPECTRUM RED™, SPECTRUM GREEN™, cyanine dyes (e.g., CY-3™, CY-5™, CY-3.5™, CY5.5™, etc.), ALEXA FLUOR™ dyes (e.g., ALEXA FLUOR™ 350, ALEXA FLUOR™ 488, ALEXA FLUOR™ 532, ALEXA FLUOR™ 546, ALEXA FLUOR™ 568, ALEXA FLUOR™ 594, ALEXA FLUOR™ 633, ALEXA FLUOR™ 660, ALEXA FLUOR™ 680, etc.), BODIPY™ dyes (e.g., BODIPY™ FL, BODIPY™ R6G, BODIPY™ TMR, BODIPY™ TR, BODIPY™ 530/550, BODIPY™ 558/568, BODIPY™ 564/570, BODIPY™ 576/589, BODIPY™ 581/591, BODIPY™ 630/650, BODIPY™ 650/665, etc.), IRDyes (e.g., IRD40, IRD 700, IRD 800, etc.), and the like. For more examples of suitable fluorescent dyes and methods for coupling fluorescent dyes to other chemical entities such as proteins and peptides, see, for example, "The Handbook of Fluorescent Probes and Research Products", 9th Ed., Molecular Probes, Inc., Eugene, OR. Favorable properties of fluorescent labeling agents include high molar absorption coefficient, high fluorescence quantum yield, and photostability. In some embodiments, labeling fluorophores exhibit absorption and emission wavelengths in the visible (i.e., between 400 and 750 nm) rather than in the ultraviolet range of the spectrum (i.e., lower than 400 nm). For example, a suitable dye for use in real-time PCR procedures may include SYBR Green.

A detectable moiety may include more than one chemical entity such as in fluorescent resonance energy transfer (FRET). Resonance transfer results an overall enhancement of the emission intensity. For instance, see Ju et. al. (1995) Proc. Nat'l Acad. Sci. (USA) 92:4347, the entire contents of which are herein incorporated by reference. To achieve resonance energy transfer, the first fluorescent molecule (the "donor" fluor) absorbs light and transfers it through the resonance of excited electrons to the second fluorescent molecule (the "acceptor" fluor). In one approach, both the donor and acceptor dyes can be linked together and attached to the oligo

primer. Methods to link donor and acceptor dyes to a nucleic acid have been described previously, for example, in U.S. Pat. No. 5,945,526 to Lee et al., the entire contents of which are herein incorporated by reference. Donor/acceptor pairs of dyes that can be used include, for example, fluorescein/tetramethylrhodamine, IAEDANS/fluorescein, EDANS/DABCYL, fluorescein/fluorescein, BODIPY FL/BODIPY FL, and Fluorescein/ QSY 7 dye. See, e.g., U.S. Pat. No. 5,945,526 to Lee et al. Many of these dyes also are commercially available, for instance, from Molecular Probes Inc. (Eugene, Oreg.). Suitable donor fluorophores include 6-carboxyfluorescein (FAM), tetrachloro-6-carboxyfluorescein (TET), 2'-chloro-7'-phenyl-1,4-dichloro-6-carboxyfluorescein (VIC), and the like.

Enzymes

In certain embodiments, a detectable moiety is an enzyme. Examples of suitable enzymes include, but are not limited to, those used in an ELISA, e.g., horseradish peroxidase, beta-galactosidase, luciferase, alkaline phosphatase, etc.. Other examples include beta-glucuronidase, beta-D-glucosidase, urease, glucose oxidase, etc. An enzyme may be conjugated to a molecule using a linker group such as a carbodiimide, a diisocyanate, a glutaraldehyde, and the like.

Radioactive isotopes

In certain embodiments, a detectable moiety is a radioactive isotope. For example, a molecule may be isotopically-labeled (i.e., may contain one or more atoms that have been replaced by an atom having an atomic mass or mass number different from the atomic mass or mass number usually found in nature) or an isotope may be attached to the molecule. Non-limiting examples of isotopes that can be incorporated into molecules include isotopes of hydrogen, carbon, fluorine, phosphorous, copper, gallium, yttrium, technetium, indium, iodine, rhenium, thallium, bismuth, astatine, samarium, and lutetium (i.e., ^3H , ^{13}C , ^{14}C , ^{18}F , ^{19}F , ^{32}P , ^{35}S , ^{64}Cu , ^{67}Cu , ^{67}Ga , ^{90}Y , $^{99\text{m}}\text{Tc}$, ^{111}In , ^{125}I , ^{123}I , ^{129}I , ^{131}I , ^{135}I , ^{186}Re , ^{187}Re , ^{201}Tl , ^{212}Bi , ^{213}Bi , ^{211}At , ^{153}Sm , ^{177}Lu).

In some embodiments, signal amplification is achieved using labeled dendrimers as the detectable moiety (see, e.g., *Physiol Genomics* 3:93-99, 2000), the entire contents of which are herein incorporated by reference in their entirety. Fluorescently labeled dendrimers are available from Genisphere (Montvale, N.J.). These may be chemically conjugated to the oligonucleotide primers by methods known in the art.

Kits

In certain embodiments, the invention provides kits for use in accordance with the invention. Generally, inventive kits comprise one or more reagents that differentiate a normal CFTR gene or protein from a mutant CFTR gene or protein containing one or more mutations selected from Table 1, 2, 3, or 4. For example, kits may comprise one or more (e.g., any combination of) reagents as described herein, and optionally additional components. For example, a kit according to the present invention may also include reagents that can detect other CFTR mutations well known in the art.

Suitable reagents may include nucleic acid probes and/or antibodies or fragments thereof. In some embodiments, suitable reagents are provided in a form of an array such as a microarray or a CFTR mutation panel.

In some embodiments, provided kits further comprise reagents for carried out various detection methods described herein (e.g., sequencing, hybridization, primer extension, multiplex ASPE, immuno assays, etc.). For example, kits according to the invention may optionally contain buffers, enzymes, and/or reagents for use in methods described herein, e.g., for amplifying nucleic acids via primer-directed amplification, for performing ELISA experiments, etc.

In some embodiments, provided kits further comprise a control indicative of a healthy individual, e.g., a nucleic acid and/or protein sample from an individual who does not carry a CFTR mutation associated with CF or a CF related disorder. In some embodiments, provided kits further comprise a control indicative of known CFTR mutant alleles (such as $\Delta F508$). Kits may also contain instructions on how to determine if an individual has CF or a CF related disorder, is at risk of developing CF or a CF related disorder, or is a carrier of CFTR mutation.

In some embodiments, a computer readable medium encoding information corresponding to one or more mutations shown in Tables 1, 2, 3, and 4 is provided. Such computer readable medium may be included in a kit of the invention.

Systems

In an embodiment, the present invention provides systems for carrying out the analysis of the invention. Thus, in an embodiment, the present invention comprises a computer-readable medium on which is encoded programming code for the methods described herein. Also in an embodiment, present invention may comprise a system comprising a processor in

communication with a computer-readable medium, the processor configured to perform the methods described herein. Suitable processors and computer-readable media for various embodiments of the present invention are described in greater detail below and are illustrated in FIG. 5.

5 Thus, in certain embodiments, the invention comprises a system for predicting the activity of at least one gene comprising: a computer readable medium; and a processor in communication with the computer readable medium, the processor configured to estimate the effects of individual mutations in the at least one gene. The processor may, in certain
10 embodiments, be further in communication with a database comprising data for a plurality of sequences for the portion of the at least one gene, where the processor is configured to compare the nucleic acid and/or amino acid sequence of the portion of the at least one gene to the data of the plurality of sequences for the portion of the at least one gene to determine if there is a mutation in the portion of the at least one gene in the biological sample obtained from the subject.

15 In other embodiments, the invention comprises a computer readable medium on which is encoded program code for predicting the activity of at least one gene, the program code comprising code for applying a model to estimate the effects of individual mutations in the at least one gene. In certain embodiments, the programming code comprises code configured to compare the amino acid and/or nucleic acid sequence of the portion of the at least one gene to the
20 data for a plurality of sequences for the portion of the at least one gene stored in a database to determine if there is a mutation in the portion of the at least one gene in the biological sample obtained from the subject.

 Some embodiments of the systems and computer readable media of the invention may be applied to various genes. In certain embodiments, the at least one gene comprises the CFTR
25 gene.

 As noted herein, the sequence of the portion of the at least one gene and the biological activity of interest as assessed for a particular subject may be compared to a database of amino acid and/or nucleic acid sequences and biological activity as assessed for a plurality of subjects. Thus, in certain embodiments of the systems and computer readable media, the database
30 comprises data for the biological activity as measured in a plurality of samples from which the sequence of the portion of the at least one gene was determined.

Embodiments in accordance with aspects of the present subject matter can be implemented in digital electronic circuitry, in computer hardware, firmware, software, or in combinations of the preceding. In one embodiment, a computer may comprise a processor or processors. The processor may comprise, or have access to, a computer-readable medium, such as a random access memory coupled to the processor. The processor may execute computer-executable program instructions stored in memory, such as executing one or more computer programs including a sampling routine and suitable programming to produce output to generate the analysis described in detail herein.

Such processors may comprise a microprocessor, a digital signal processor (DSP), an application-specific integrated circuit (ASIC), field programmable gate arrays (FPGAs), and state machines. Such processors may further comprise programmable electronic devices such as PLCs, programmable interrupt controllers (PICs), programmable logic devices (PLDs), programmable read-only memories (PROMs), electronically programmable read-only memories (EPROMs or EEPROMs), or other similar devices.

Such processors may comprise, or may be in communication with, media, for example tangible computer-readable media that may store instructions that when executed by the processor, can cause the processor to perform the steps described herein as carried out, or assisted, by a processor. Embodiments of computer-readable media may comprise, but are not limited to, all electronic, optical, magnetic, or other storage devices capable of providing a processor, such as the processor in a web server, with computer-readable instructions. Other examples of media comprise, but are not limited to, a floppy disk, CD-ROM, magnetic disk, memory chip, ROM, RAM, ASIC, configured processor, all optical media, all magnetic tape or other magnetic media, or any other medium from which a computer processor can read. Also, various other devices may include computer-readable media, such as a router, private or public network, or other transmission device. The processor, and the processing may be in one or more structures, and may be dispersed through one or more structures. The processor may comprise code for carrying out one or more of the methods (or parts of methods) described herein.

The system may comprise a data compiling system as well as a means for the user to interact with the system as the analysis proceeds. Thus, in an embodiment, the present invention may comprise a system for collecting and/or compiling data from a plurality of assay measurements and/or sequencing data and transmitting the data to a computer, and a system for

transmitting the results of the analysis to a user. The systems of the present invention may be designed for high-throughput analysis of DNA and/or amino acid sequencing data. Thus, in an embodiment, the plurality of measured signals comprise a plurality of known DNA sequences isolated from at least one cell type.

5 Figure 5 shows an embodiment of the flow of information in a system comprising the software of the present invention. As discussed above, a computer processor or CPU may include, for example, digital logic processors capable of processing input, executing algorithms, and generating output as necessary in response to the inputs received from the touch-sensitive input device. As detailed herein, such processors may include a microprocessor, such as an
10 ASIC, and state machines, and/or other components. Such processors include, or may be in communication with, media, for example computer-readable media, which stores instructions that, when executed by the processor, cause the processor to perform the steps described herein.

 Thus, in an embodiment, the starting point may comprise data (100) that may comprise a normal CFTR gene (100A) and mutant CFTR gene (100B). Once the data has been collected
15 (110), it may be compiled (120) and/or transformed if necessary using any standard spreadsheet software such as Microsoft Excel, FoxPro, Lotus, or the like. In an embodiment, the data are entered into the system for each experiment. Alternatively, data from previous runs are stored in the computer memory (150) and used as required.

 At each point in the analysis, the user may input instructions via a keyboard (180), floppy
20 disk, remote access (e.g., via the internet) (190), or other access means. The user may enter instructions including options for the run, how reports should be printed out, and the like. Also, at each step in the analysis, the data may be stored in the computer using a storage device common in the art such as disks, drives or memory (150). As is understood in the art, the processor (160) and I/O controller (170) are required for multiple aspects of computer function.
25 Also, in a embodiment, there may be more than one processor.

 The data may also be processed to remove noise (130). In some cases, the user, via the keyboard (180), floppy disk, or remote access (190), may want to input variables or constraints for the analysis, as for example, the threshold for determining noise. The results of the analysis may then be compiled and provided in a form for review by a user (140).

30 The following examples serve to illustrate the present invention. These examples are in no way intended to limit the scope of the invention.

EXAMPLES

Example 1: Identification of novel mutations in the CFTR gene

Novel mutations in the CFTR gene were identified using the CF full sequencing assay. Typically, samples submitted for CF full sequencing assays were from individuals for whom testing in CF mutation panels has been uninformative, or partially uninformative. These individuals include 1) patients with idiopathic chronic pancreatitis; 2) patients with congenital bilateral absence of the vas deferens (CBAVD); 3) couples who test positive/negative by mutation analysis; 4) CF-affected or suspected patients in whom one or no mutations have been identified; 5) obligate carriers of a rare familial mutation; 6) patients with a family history of CF, for whom mutation analysis by other methodologies is negative; 7) patients with a CF related disease or condition. Sequence changes in the CFTR gene were identified by comparing the patient gene sequence to the wild-type gene sequence. Novel mutations in the CFTR gene that were unreported previously are summarized in Table 5.

As shown in Table 5, patients carrying these novel mutations were from different ethnic groups including Caucasians, African Americans, Hispanics, and Asians. Some of the mutations are located in introns (e.g., intron 3, intron 6a, intron 11, intron 14a, intron 19, intron 20, intron 21, and intron 23). Some of the mutations are located in exons (e.g., exon 2, exon 3, exon 4, exon 5, exon 6a, exon 6b, exon 7, exon 9, exon 10, exon 11, exon 12, exon 13, exon 14a, exon 14b, exon 15, exon 16, exon 17a, exon 17b, exon 19, exon 20, exon 21, exon 22, and exon 24).

As shown in Table 5, most of the novel mutations identified result in codon changes or altered gene splicing sites, which will likely affect the CFTR gene expression and/or protein function. In particular, some of the mutations are nonsense mutations (i.e., mutations predicted to result in the introduction of a stop codon). Some of the mutations affect consensus splice site ag/gt. Some of these mutations are insertion or deletion of at least one nucleotide. These mutations are category 2 mutations according to the ACMG guidelines, and are of the type expected to cause CF or CF related disease, disorder or condition.

Some mutations are missense mutations. Some are predicted to cause cause in-frame insertions and/or deletions. Some are likely to affect splice sites. These mutations are category 3 mutations according to the ACMG guidelines.

WO 2011/119531

PCT/US2011/029314

Thus, the novel mutations provided herein can be used, alone or in combination with other known CF mutations, to detect CF or a CF related disorder in CFTR testing assays including carrier testing.

Table 5. Novel Mutations in the CFTR gene

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
1824delA	Frameshift (FS)	n/a	Caucasian	F508del	Mutation was identified in a 22 year old patient with a known diagnosis of CF. This patient carried a second mutation known to cause CF (F508del).	e12
2957delT	FS	n/a	Caucasian	F508del	Mutation was identified in a 1 year old patient with a known diagnosis of CF. The patient carried a second mutation known to cause CF (F508del).	e15
4089ins4	FS	n/a	Caucasian	F508del	Mutation was identified in a 7 year old patient with a known diagnosis of CF. The patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	e21
4374+2T>C	Splice site mutation	n/a	1. Caucasian 2. Caucasian	1. F508del 2. F508del	Patient #1: Mutation was identified in a 45 year old patient with a suspected diagnosis of CF. The patient carried a second mutation known to cause CF (F508del). Patient #2: Mutation was identified in a 52 year old patient with a suspected diagnosis of CF. The patient carried a second mutation known to cause CF (F508del).	i23
3064A>T	Nonsense	K978X	African American	Q1042X	Mutation was identified in a 26 year old patient with a known diagnosis of CF. The patient carried a second mutation likely to cause CF Q1042X.	e16
246C>G	Nonsense	Y38X	Caucasian	F508del	Mutation was identified in a 1 month old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	e2
269C>T	Missense (MS)	A46V	1) Caucasian 2) Black 3) African American	1) 3849+1219 2G>A 2) F508del 3) none	Patient #1: Mutation was identified in a 32 year old patient who was tested due to abnormalities found on fetal ultrasound. The patient carried a second mutation of unknown clinical significance (3849+12192G>A). Patient #2: Mutation was identified in a 2 month old patient who was tested based on follow-up for a positive newborn screen. The patient carried a second mutation known to cause cystic fibrosis (F508del). Patient #3: Mutation was identified in a 24 year old patient who was tested as a parental follow-up to a positive newborn screen.	e2

WO 2011/119531

PCT/US2011/029314

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
2902 G>T	MS	D924Y	Caucasian	F508del	Mutation was identified in a 1 month old patient with a suspected diagnosis of CF. The patient also had a positive sweat chloride test and carried a second mutation known to cause CF (F508del)	e 15
3814G>A	MS	E1228K	Caucasian	F508del	Mutation was identified in a 1 month old patient with a suspected diagnosis of CF. The patient had a borderline sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	e19
502G>C	MS	G124R	Not Provided	F508del	Mutation was identified in a 2 month old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (F508del).	e4
1520G>T	MS	G463V	Caucasian	F508del	Mutation was identified in a 17 year old patient with a known diagnosis of CF. Patient carried a second mutation known to cause CF (F508del).	e9
511_513dup TTA	In frame duplication	L127dup	Caucasian, Asian	W1282X	Mutation was identified in a newborn with a suspected diagnosis of CF. The patient had clinical symptoms of CF including as a positive sweat chloride test, meconium ileus, echogenic bowel, and pancreatic insufficiency. The patient carried a second mutation known to cause CF (W1282X).	e4
978A>T	MS	E282D	1. not provided 2. not provided	1. 3120+1G>A 2) none	Patient #1: Mutation was identified in a 10 year old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (3120+1G>A). Patient #2: Mutation was identified in a 4 year old patient with a suspected diagnosis of CF and a family history of CF.	e6b
843G>C	MS	Q237H	Caucasian	F508del	Mutation was identified in a 2 month old patient with a known diagnosis of CF. The patient carried a second mutation known to cause CF (F508del).	e6a
829C>T	MS	L233F	Caucasian	D1152H	Mutation was identified in a 1 month old patient who was tested following a positive newborn screen. The patient carried a second mutation known to cause CF (D1152H).	e6a
4096-6C>T	Splice site mutation	None	Caucasian	F508del	Mutation was identified in a 58 year old patient with a suspected diagnosis of CF. The patient carried a second mutation known to cause CF (F508del).	i21

WO 2011/119531

PCT/US2011/029314

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
4375-7delT	Splice site mutation	None	Caucasian	F508del	Mutation was identified in a 6 year old patient with a suspected diagnosis of CF. Patient has a family history, a borderline sweat chloride test and recurrent pneumonia. The patient carried a second mutation known to cause CF (F508del).	i23
1586 G>C	MS	S485T	Caucasian	S1235R	Mutation was identified in a 2 year old patient with a suspected diagnosis of CF. The patient carried a second mutation S1235R (3837T>G) which has been reported in individuals with varying CF phenotypes.	e10
875+4G>T	Splice site mutation	n/a	African American	none	Mutation was identified in a 1 month old patient who had a positive newborn screening test.	i6a
4005+3G>T	Splice site mutation	n/a	Caucasian	none	Mutation was identified in a 40 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	i20
2711T>C	MS	I860T	Caucasian	F508del, E528E	Mutation was identified in a 58 year old woman with a suspected diagnosis of CF. The patient carried a second mutation known to cause CF (F508del) and an additional mutation of unknown clinical significance (E528E)..	e14a
3891G>C	MS	L1253F	Not provided	G85E, L15P	Mutation was identified in a 32 year old patient with a known diagnosis of CF. The patient carried a second mutation known to cause CF (G85E) and an additional mutation of unknown clinical significance (L15P).	e20
2524C>T	MS	P798S	African American	F508del, R74W, G921E, D1270N	Mutation was identified in a 5 year old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. This patient carried a second mutation known to cause CF (F508del) and three additional mutations of unknown clinical significance (R74W, G921E, D1270N).	e13
2894G>A	MS	G921E	African American	F508del, R74W, P798S, D1270N	Mutation was identified in a 5 year old patient with a suspected diagnosis of CF. The patient had a positive sweat chloride test. This patient carried a second mutation known to cause CF (F508del) and three additional mutations of unknown clinical significance (R74W, P789S, D1270N).	e15
405+1024 7C>T	Possible splice site mutation	n/a	Caucasian	F508del	Mutation was identified in a 35 year old patient who was tested to determine if they were a carrier, there was no family history of CF. This patient carried a second mutation known to cause CF (F508del).	i3
405+1025	Possible	n/a	Not	F508del,	Mutation was identified in a 10 year old patient.	i3

WO 2011/119531

PCT/US2011/029314

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
5d elC	splice site mutation		Provided	124del23bp	The patient carries two mutations know to cause CF (F508del and 124del23).	
1811+1643 G>T	Possible splice site mutation	n/a	1. Hispanic 2. Hispanic 3. Not provided	1. F508del 2. F508del 3. none	Patient #1: Mutation was identified in a 1 year old patient with a known diagnosis of CF. Patient had a positive sweat chloride test. The patient carried a second mutation known to cause CF (F508del). Patient #2: Mutation was identified in a 6 year old patient with a known diagnosis of CF. The patient carried a second mutation know to cause CF (F508del). Patient #3: Mutation was identified in an 8 month old patient with a suspected diagnosis of CF.	i11
1812-13A>G	Splice site mutation	n/a	Caucasian	none	Mutation was identified in a 15 year old patient with a suspected diagnosis of CF. The patient has chronic sinusitis.	i11
2752-33insA	Possible splice site mutation	n/a	African American	F693L	Mutation was identified in a 6 year old patient with a known diagnosis of CF. The patient carries a second mutation of unknown clinical significance (F693L).	i14a
3849+12192 G>A	Possible splice site mutation	n/a	Caucasian	A46V	Mutation was identified in a 32 year old patient who was tested due to abnormalities found on fetal ultrasound. The patient carried an additional mutation of known clinical significance (A46V).	i19
724G>A	MS	A198T	Hispanic	none	Mutation was identified in a 4 month old patient with a suspected diagnosis of CF.	e6a
3899C>T	MS	A1256V	Guyanese	none	Mutation was identified in a 45 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	e20
3986C>T	MS	A1285V	Not Provided	none	Mutation was identified in a 23 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	e20
901G>A	MS	E257K	Hispanic	none	Mutation was identified in a 4 year old patient with a suspected diagnosis of CF. The patient has asthma and recurring pneumonia.	e6b
392 T>C	MS	F87S	Not Provided	none	The mutation was identified in a 1 month old patient with a suspected diagnosis of CF.	e3
3463T>C	MS	F1111L	Hispanic	none	Mutation was identified in a 6 year old patient with a suspected diagnosis of CF. The patient has asthma.	e17b

WO 2011/119531

PCT/US2011/029314

Sequence Change	Type of mutation	AA Change	Ethnicity	Other Mutations	Clinical Information	Exon
I757G>A	MS	G542E	Hispanic	none	Mutation was identified in a 25 year old patient who was tested to determine if they were a carrier, there was no family history of CF. The patient carried 2 copies of G542E..	e11
4025G>C	MS	G1298 A	Asian	G970D, Q1352H	Mutation was identified in a 34 year old patient with congenital absence of the vas deferens. The patient carried two other mutations of unknown clinical significance (G970D and Q1352H)	e21
4129G>T	MS	G1333 W	Not Provided	none	Mutation was identified in an 8 year old patient with a suspected diagnosis of CF. Patient had recurrent respiratory infections and chronic cough.	e22
663T>G	MS	I177M	Caucasian	none	Mutation was identified in a 34 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	e5
3200T>C	MS	I1023T	Hispanic	none	Mutation was identified in a 34 year old patient who was tested to determine if they were a carrier, there was no family history of CF.	e17a
4412 T>C	MS	I1427T	Asian	S1444S	Mutation was identified in a 34 year old patient who was tested to determine if they were a carrier, there was no family history of CF. The patient carried another mutation that is considered likely to be clinically benign (S144S).	e24
620A>C	MS	K163T	Caucasian	none	Mutation was identified in a 32 year old patient with a family history of CF.	e4

Example 2. CFTR Mutation Detection Assay

The present example demonstrates that multiplex ASPE assay can be used to detect novel cystic fibrosis mutations described herein. Multiplex ASPE combines multiplex PCR and allele-specific primer extension. Multiplex PCR is performed to amplify target regions in the CFTR

gene containing novel sequence variations described herein from genomic DNA in a sample. Multiplex primer extension reactions are then performed using allele-specific primers, i.e., extension primers that possess a 3' terminal nucleotide, which form a perfect complement with the target sequence, are extended to form extension products and modified nucleotides (e.g., biotinylated dCTP) are incorporated into the extension product for detection purposes.

Alternatively, an extension primer may instead contain a 3' terminal nucleotide which forms a mismatch with the target sequence. In this instance, primer extension does not occur. Primer extension products are then hybridized to universal array beads with "anti-tag" sequence (sequences complementary to the tag sequence) for capture and detection purposes.

In some cases, the novel mutations described herein can be detected in combination of other known CF mutations, for example, mutations recommended by the American College of Genetics and American College of Obstetricians and Gynecologists, as well as other common and clinically relevant mutations, such as, for example, $\Delta F508$ (exon 10), G542X (exon 11), G551D (exon 11), R117H (exon 4), W1282X (exon 20), N1303K (exon 21), 3905insT (exon 20), 3849+10KbC>T (intron 19), G85E (exon 3), R334W (exon 7), A455E (exon 9), 1898+1G>A (exon 12), and/or 2184delA (exon 13).

Various ASPE kits can be used to carry out the detection methods described herein. For example, Luminex's TAG-ITTM kit and Data Analysis software can be modified to detect a panel of CF mutations including one or more novel mutations described herein. Mutation detection kit may use non-isotopic fluorescent technology, and a 96-well assay format that is compatible with automation such that result analyses and genotype calling are automated.

Allele Specific Primers

Allele specific primers can be designed based on the sequence variations shown in Table 5 and the CFTR genomic sequences (including exon and intron sequences) using various methods and software known in the art. A universal tag sequence can be added to allele specific primers.

Specimens and Assay Format

Specimens containing genomic DNA to be analyzed can be obtained from, but not limited to, the following sources: Whole blood (e.g., whole blood in EDTA, ACD-A, ACD-B), fresh or frozen tissue, amniotic fluid, CVS (chorionic villus sampling) tissue, cultured cells (e.g., CVS, amniotic fluid, fibroblasts, POC (product of conception)), blood spots, cord blood, mouthwash, genomic DNA extracted by an outside laboratory. Blood and bloodspot DNA samples are typically run undiluted at a 5 μ L input volume. An amount of 5 to 200 ng DNA is used as input. For testing prenatal and mouthwash samples, generally between 20 ng and 150 ng is used as input, for example, about 20 ng, 25 ng, 30 ng, 35 ng, 40 ng, 45 ng, 50 ng, 55 ng, 60 ng, 65 ng, 70 ng, 75 ng, 80 ng, 85 ng, 90 ng, 95 ng, 100 ng, 110 ng, 120 ng, 130 ng, 140 ng, or 150 ng.

A 96-well assay plate is used. Two genomic DNA controls are included with each assay plate. The specific controls are rotated sequentially through assay plates. Each assay plate also includes two cocktail blanks and ASPE (Allele-Specific Primer Extension) controls. A calibrating 96-well filter plate is also used during data acquisition.

Single-well multiplex PCR

Multiplex PCR are performed to amplify exons containing mutations described herein using consensus flanking intron sequences. Generally, amplicons range in size between about 150 bp and 600 bp (inclusive of endpoints).

Typically, 5 ng — 200 ng of DNA is amplified to produce a product containing multiple amplicons using PCR amplification conditions known in the art or optimized/modified using routine experimentation.

Enzymatic Post-PCR cleanup

PCR products are treated with Exonuclease I and Shrimp Alkaline Phosphatase to remove residual primers that will interfere with allele-specific primer extension reactions. PCR products are incubated with enzyme and then enzyme is heat-deactivated, according to standard protocols or modified protocols readily developed by one of ordinary skill in the art.

Single well Allele Specific Primer Extension (ASPE) reactions

Typically up to 100 sequence variations can be distinguished in a single-well reaction; using the Luminex bead set. For example, a set of allele-specific oligonucleotide (ASO) primers (including wildtype control ASOs) with tag sequences are used.

The Exo-SAP-treated PCR product is subjected to an allele-specific primer extension reaction containing tagged primers and biotinylated dCTP using PCR reaction conditions known in the art or modified readily by one of ordinary skill in the art..

Universal Array Sorting and Detection

5 Each bead is coupled with an anti-tag sequence complementary to the tag sequence ASPE primers. Therefore, any ASPE products, if present, can be captured for genotype analysis. Wild-type control for each amplicon is included. The signals from wildtype alleles serve as a control for each amplicon and provide information for allelic ratio calculation (typically obtained by calculating the ratio of signal for the mutant allele over signal for the wildtype allele), for the
10 detected mutations.

The ASPE product is added to the universal bead array containing anti-tags to the ASPE primers and incubated for hybridization. Hybridization reactions are then washed over a filter that captures the beads and removes any non-hybridized ASPE products containing biotin. Bead hybridization conditions are known in the art and can be adapted readily by one skilled in the art.

15 Streptavidin R-Phycoerythrin conjugate is added to the hybridized products on the filter plate and incubated at room temperature, followed by bead sorting and detection. For example, a modified LUMINEX™ 100 IS™ or 200 IS™ can be used. The LUMINEX™ 100 IS™ can upload sample sheets from text files or barcodes. Detection time averages 20-100 seconds per well.

20 *Results*

In the LUMINEX™ system, results are generated as a <.csv> file and exported in batches. The batch output file (.csv) is opened in TAG-IT™ Data Analysis Software (TDAS) version 6.0 where results are automatically generated based on pre-determined algorithms for allelic ratios on certain individually tested mutations and the presence or absence of signal on the
25 remaining mutations.

Mutation Confirmation

Samples positive for any of the mutations described herein can be confirmed by a second assay run. Positive samples can also be confirmed by direct DNA sequencing.

Example 3: Cystic Fibrosis Sequencing Assays

30 The Cystic Fibrosis full sequencing assay and single exon sequence assay can be used to detect mutations in the CFTR gene directly in a patient sample. The Cystic Fibrosis full

sequencing assay and single exon sequence assay can also be used to complement CF screening panels, and/or to serve as a confirmatory assay for samples that are positive for multiplex mutations or those without a normal counterpart in the CF mutation detection assay.

5 The CF full sequencing assay sequences the entire coding region of the CFTR gene plus 15 bp at the 3' end of each intron (30bp for e17b to cover a known mutation) and 6 bp at the 5' beginning of each intron.

In addition, the assay includes portions of introns 1, 3, 11, and 19 useful in identifying the exon 2, 3 deletion, the A>G mutation at 1811+1.6 kb, and the C>T mutation at 3849+10 kb.

Typically, the assay comprises analysis of 31 amplicons: e1, i1, e2, e3, i3, e4, e5, e6a, e6b, e7, e8, 10 e9, e10, e11, e12, e13a, e13b, e14a, e14b, e15, e16, e17a, e17b, e18, e19, i19, e20, e21, e22, e23, and e24. Each amplicon includes the complete coding region of the exon with the exception of 13.1 and 13.2, in which, due to the large size of the exon, the amplicon is divided into two fragments. The CF Single Exon Sequencing assay uses the same primers but on an individual basis as needed.

15 Samples tested in the CF single exon sequencing assay in this Example include those from individuals that 1) tested positive in a CF mutation detection assay (e.g., multiplex ASPE assay as described in Examples 2) but require confirmation; 2) tested positive in the CF full sequencing assay and require repeat testing; 3) are being tested for a known familial mutation(s); and/or 4) are being tested for a mutation that is not detectable in the CF mutation detection assay 20 of Example 2.

Specimens and Assay Format

Specimens to be analyzed can be extracted genomic DNA from any of, but not limited to, the following sources: Whole blood (e.g., whole blood in EDTA, ACD-A, ACD-B), blood spots, amniotic fluid, chorionic villus samples (CVS) (for single exon sequencing only), cultured cells 25 (e.g., CVS, amniotic fluid, fibroblasts, POC), mouthwash (for single exon sequencing only).

A 96-well format is used. Cocktail blanks are run for all amplicons on each assay.

PCR Amplification

Target regions containing mutations described herein are first amplified by PCR amplification. Typically, 5 ng - 200 ng of DNA is amplified in a 25 µL volume reaction. PCR 30 primers include 5' UPS tags - UPS1 for the Forward primers and UPS2 for the Reverse primers.

Table 6 presents sequences of exemplary primers used in amplification of certain exemplary target exon or intron regions.

Table 6: Primer sequences

Amplicon	Primer Name	Sequence (5'-3')	Amplicon Length	SEQ ID NO.
Primers for exonic sequences				
CF exon 1	UP1CFe1F	TTTAACCTGGGCAGTGAAG	373	5
	UP2CFe1R	AACCCAACCCATACACA		6
CF exon 2	UP1CFe2F	CAAATCAAGTGAATATCTGTTC	316	7
	UP2CFe2R	AGCCACCATACTTGGCTCCTA		8
CF exon 3	UP1CFe3F2	CTAAAATATTTGCACATGCAAC	333	9
	UP2CFe3R	TTTCTTAGTGTTTGGAGTTGG		10
CF exon 4	UP1CFe4F2	TCATTTTAAGTCTCCTCTAAAG	407	11
	UP2CFe4R	CGATACAGAATATATGTGCCA		12
CF exon 5	UP1CFe5F2	AACAACCTAGAAGCATGCCAG	394	13
	UP2CFe5R2	GTTGTATAATTTATAACAATAGTG		14
CF exon 6a	UP 1 CF6aF2	GGAAGATACAATGACACCTG	353	15
	UP2CFe6aR3	CTGAAGATCACTGTTCTATGC		16
CF exon 6b	UP1 CF6bF3	ATGACTTAAACCTTGAGCAGT	336	17
	UP2CFe6bR2	GGAAGTCTACCATGATAAACAT		18
CF exon 7	UP1CFe7F2	GAGACCATGCTCAGATCTTCC	507	19
	UP2CFe7R	ACTTTTATAACTTCCTAGTGAAG		20
CF exon 8	UP1CFe8F2	AAGATGTAGCACAATGAGAGTA	268	21
	UP2CFe8R	CAGTTAGGTGTTTAGAGCAA		22

WO 2011/119531

PCT/US2011/029314

Amplicon	Primer Name	Sequence (5'-3')	Amplicon Length	SEQ ID NO.
CF exon 9	UP1CFe9F	GTATACAGTGTAATGGATCATG	402	23
	UP2CFe9R4	CACCAAATTAAGTTCTTAATAG		24
CF exon 10	UP1CFe10F	TTCTGCTTAGGATGATAATTGG	479	25
	UP2CFe10R	GCATAGGTCATGTGTTTTATTA		26
CF exon 11	UP1CFe11F	CAGATTGAGCATACTAAAAGTG	240	27
	UTP2CFe11R	TACATGAATGACATTTACAGCA		28
CF exon 12	UP1CFe12F	GCTACTTCTGCACCACTTTTG	344	29
	UP2CFe12R	CAGTCTGTCTTTCTTTTATTTTA		30
CF exon 13a	UP1CFe13F3	CAAAATGCTAAAATACGAGAC	388	31
	UP2CFe13R5	TCCAGGAGACAGGAGCATC		32
CF exon 13b	UP1CFe13F4	CTCATGGGATGTGATTCTTT	714	33
	UP2CFe13R2	GATACACCTTATCCTAATCCTA		34
CF exon 14a	UP1CFe14aF3	ACCACAATGGTGGCATGA	299	35
	UP2CFe14aR	TGTATACATCCCCAAACTATC		36
CF exon 14b	UP1CFe14bF2	TGGGCATGGGAGGAATAGGTG	228	37
	UP2CFe14bR	TTACAATACATACAAACATAGTG G		38
CF exon 15	UP1CFe15F2	AAGTAACTTTGGCTGC	416	39
	UP2CFe15R2	CTGCCATTAGAAAACCA		40
CF exon 16	UP1CFe16F2	AAGTCTATCTGATTCTATTTGC	307	41

Amplicon	Primer Name	Sequence (5'-3')	Amplicon Length	SEQ ID NO.
	UP2CFe16R2	GTTTTTTTAATAATACAGACATAC T		42
CF exon 17a	UP1CFe17aF 3	TGTCCACTITGCAATGTGAA	317	43
	UP2CFe17aR 3	CAATAAAGAATCTCAAATAGCTC T		44
CF exon 17b	UP1CFe17bF 3	TAGTCTTTTTCAGGTACAAG	516	45
	UP2CFe 17bR6	CAATGGAAATTCAAAGAAATCAC T		46
CF exon 18	UP1CFe18F6	GAATACTTACTATATGCAGAGCA	416	47
	UP2CFe18R3	GTTCTTCCTCATGCTATTACTC		48
CF exon 19	UP1CFe19F	GCCCGACAAATAACCAAGTGA	494	49
	UP2CFe19R2	CTAACACATTGCTTCAGGCTA		50
CF exon 20	UP1CFe20F	AAGGTTGTTTGTCTCCATATAT	544	51
	UP2CFe20R	GCCTATGAGAAAACCTGCACT		52
CF exon 21	UP1 CFe21 F	ACATGGGTGTTTCTTATTTA	428	53
	UP2CFe21 R2	GTTAGGGGTAGGTCCAGT		54
CF exon 22	UP1CFe22F	GCTTGAGTGTTTTTAACTCTGTG	314	55
	UP2CFe22R	ATGATTCTGTTCCCACTGTGC		56
CF exon 23	UP1CFe23F	GTTCTGTGATATTATGTGTGG	226	57
	UP2CFe23R	CAAGGGCAATGAGATCTTAAG		58
CF exon 24	UP1CFe24F2	AGTTTCTGTCCCTGCTCT	356	59

WO 2011/119531

PCT/US2011/029314

Amplicon	Primer Name	Sequence (5'-3')	Amplicon Length	SEQ ID NO.
	UP2CFe24R	GAGCAAATGTCCCATGTCAAC		60
Primers for intronic sequences				
CF intron 1	UP1CFin1F2	AATGGTGTTTACCTACCTAGAGA A	250	61
	UP4CFin1R2	CCTCCTCTGATTCCACAAG		62
CF intron 3	UP3CFin3F3	CTGAGATTCTGTTCTAGGTGTG	366	63
	UP2CFin3R	CCTACACTCAGAACCCATCAT		64
CF intron 19	UP1CFin19F	TTCAGTTGACTTGTCATCTTG	223	65
	UP2CFin19R	AATATGTTGAAAGTTAAACAGTG		66
CF intron 11	UP1CFin11F	GTTACACTATAAAGGTTGTTTTAG AC	292	67
	UP2CFin11R	CACAGTTCCCATATTAATAGAAAT G		68
(Seq)	CFe9.SEQ.F	TTTTTAACAGGGATTTGGG	N/A	69
(Seq)	CFe6bF2	GATTGATTGATTGATTGATT	N/A	70
(Seq)	UPS1	GCGGTCGCATAAGGGTCAGT	N/A	71
(Seq)	UPS2	CGCCAGCGTATTCCCAGTCA	N/A	72

PCR conditions are as shown in Table 7.

Table 7:

PCR amplification conditions for CF full sequencing assay

Cycles	Temperature (°C)	Time	Function
1	95	5 min	Denaturation of enzyme
35	95	20 sec	Denaturation of dsDNA
	55	20 sec	Annealing
	72	40 sec	Extension
1	72	7 min	Final extension
1	8	Forever	End

5 *Enzymatic Post-PCR Clean Up*

PCR products are treated with Exonuclease I (Exo) and Shrimp Alkaline Phosphatase (SAP) to remove residual primers that may interfere with sequencing. The following incubation conditions are used:

- 37 °C for 30 minutes (enzyme digestion)
- 10 99 °C for 15 minutes (enzyme deactivation)
- Hold at 8 °C until storage
- Products can be stored, *e.g.*, at -80°C or -20°C.

Sequencing

- Exo-SAP treated products are diluted 1:2 in water, and 3 µL is added to 7 µL of each
- 15 forward and reverse sequence cocktail containing Big Dye v3.1 (ABI). In order to obtain bi-directional sequencing results, two sequencing reactions are performed for each amplicon, using both UPS1 and UPS2 primers. An additional forward sequencing reaction using gene specific primers is performed for Exons 6b and 9 to obtain readable sequence beyond the repeat regions. Cycle sequencing is performed in a thermocycler with the conditions shown in Table 8.

20

Table 8:
Thermocycler conditions for sequencing reactions for CF full sequencing assay

Cycles	Temperature (°C)	Time	Function
1	96	1 min	Denaturation of enzyme
25	96	10 sec	Denaturation of dsDNA
	53	5 sec	Annealing
	60	3 sec	Extension
1	8	Forever	End

5 Assay plates can be stored, *e.g.*, at -80°C for up to 2 weeks until analyzed or further manipulated.

Post-Sequencing Purification

Sequence products are purified using the Performa DTR Ultra 96 Well Plate (Edge Biosystems). Sequencing reactions are diluted 1:2 and 10 µL is purified through the Edge Plate.

10 *Sequencing run: ABI 3730 Genetic Analyzer and Data Analysis*

A 1 kV / 14 second injection is performed on the 3730x1 Genetic Analyzer. POP7 polymer and a 50 cm array are used for optimal resolution. Parameters for a typical sequencing run are shown in Table 9.

15

Table 9:
Parameters for typical sequencing runs

Feature	Parameter for CF full sequencing
Run Temp	60 °C
Pre Run Voltage	15.0 Kvolts
Pre Run Time	180 sec
Injection Voltage	1.0 Kvolts
Injection Time	14 sec
Voltage number of steps	30
Voltage Step Interval	15 sec
Data Delay Time	240 sec
Run Voltage	13.4 Kvolts
Run Time	2400 sec

Sequence data Analysis is performed using SEQSCAPE™ software (ABI).

5

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the appended claims. The articles “a”, “an”, and “the” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to include the plural referents. Claims or descriptions that include “or” between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention also includes embodiments in which more

10

15

than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process. Furthermore, it is to be understood that the invention encompasses variations, combinations, and permutations in which one or more limitations, elements, clauses, descriptive terms, etc., from one or more of the claims is introduced into another claim
5 dependent on the same base claim (or, as relevant, any other claim) unless otherwise indicated or unless it would be evident to one of ordinary skill in the art that a contradiction or inconsistency would arise. Where elements are presented as lists, e.g., in Markush group or similar format, it is to be understood that each subgroup of the elements is also disclosed, and any element(s) can be removed from the group. It should be understood that, in general, where the invention, or
10 aspects of the invention, is/are referred to as comprising particular elements, features, etc., certain embodiments of the invention or aspects of the invention consist, or consist essentially of, such elements, features, etc. For purposes of simplicity those embodiments have not in every case been specifically set forth herein. It should also be understood that any embodiment of the invention, e.g., any embodiment found within the prior art, can be explicitly excluded from the
15 claims, regardless of whether the specific exclusion is recited in the specification.

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one act, the order of the acts of the method is not necessarily limited to the order in which the acts of the method are recited, but the invention includes embodiments in which the order is so limited. Furthermore, where the claims recite a
20 composition, the invention encompasses methods of using the composition and methods of making the composition. Where the claims recite a composition, it should be understood that the invention encompasses methods of using the composition and methods of making the composition.

INCORPORATION OF REFERENCES

25 All publications and patent documents cited in this application are incorporated by reference in their entirety to the same extent as if the contents of each individual publication or patent document were incorporated herein.

What is claimed is:

1. A method comprising:
 testing a sample obtained from a subject to determine the presence of one or more
 mutations selected from Table 1, 2, 3, or 4 in the cystic fibrosis transmembrane conductance
 5 regulator (CFTR) gene or protein,
 wherein the presence of the one or more mutations indicates that the subject has cystic
 fibrosis (CF) or a CF related disorder, is at risk of developing CF or a CF related disorder, or is a
 carrier of a CFTR mutation.
2. The method of claim 1, wherein the one or more mutations are selected from Table 1, 2
 10 or 3.
3. The method of claim 1, wherein the one or more mutations are selected from Table 1 or
 2.
4. The method of claim 1, wherein the one or more mutations are selected from Table 1.
5. The method of claim 1, wherein the one or more mutations are part of a CFTR mutation
 15 panel.
6. The method of any one of the preceding claims, wherein the sample comprises nucleic
 acid.
7. The method of claim 5, wherein the testing step comprises nucleic acid sequencing.
8. The method of claim 5, wherein the testing step comprises hybridization.
- 20 9. The method of claim 7, wherein the hybridization is performed using one or more
 oligonucleotide probes specific for a region in the CFTR gene (SEQ ID NO:1) corresponding to
 the one or more mutations selected from Table 1, 2, 3 or 4, and under conditions sufficiently
 stringent to disallow a single nucleotide mismatch.
10. The method of claim 7 or 8, wherein the hybridization is performed with a microarray.
- 25 11. The method of claim 5, wherein the testing step comprises restriction enzyme digestion.
12. The method of claim 5, wherein the testing step comprises PCR amplification.
13. The method of claim 11, wherein the PCR amplification is digital PCR amplification or
 real-time PCR.
14. The method of claim 5, wherein the testing step comprises primer extension.
- 30 15. The method of claim 13, wherein the primer extension is single-base primer extension.

16. The method of claim 5, wherein the testing step comprises performing a multiplex allele-specific primer extension (ASPE).
17. The method of claim 1, wherein the sample comprises protein.
18. The method of claim 16, wherein the testing step comprises amino acid sequencing.
- 5 19. The method of claim 16, wherein the testing step comprises performing an immuno assay using one or more antibodies that specifically recognize one or more epitopes corresponding to the one or more mutations selected from Table 1, 2, 3 or 4.
20. The method of claim 16, wherein the testing step comprises protease digestion.
21. The method of claim 19, wherein the protease digestion is trypsin digestion.
- 10 22. The method of claim 19, wherein the testing step further comprises performing 2D-gel electrophoresis.
23. The method of claim 1, wherein the testing step comprises determining the presence of the one or more mutations using mass spectrometry.
24. The method of claim 22, wherein the mass spectrometric format is selected from among
15 Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform, and combinations thereof.
25. The method of claim 1, wherein the sample is obtained from cells, tissue, whole blood, mouthwash, plasma, serum, urine, stool, saliva, cord blood, chorionic villus sample, chorionic villus sample culture, amniotic fluid, amniotic fluid culture, transcervical lavage fluid, and
20 combination thereof.
26. The method of claim 1, wherein the method is used in conjunction with an additional test for CFTR mutations.
27. A nucleic acid probe that specifically binds to a normal CFTR gene but not to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4.
- 25 28. An array comprising one or more probes that specifically bind to a normal CFTR gene but not to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4.
29. A nucleic acid probe that specifically binds to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4 but not to a normal CFTR gene.

30. An array comprising one or more probes that specifically bind to a mutant CFTR gene containing one or more mutations selected from Table 1, 2, 3, or 4 but not to a normal CFTR gene.
31. An antibody that specifically binds to a normal CFTR protein but not to a mutant CFTR protein containing one or more mutations selected from Table 1, 2, 3, or 4.
32. An antibody that specifically binds to a mutant CFTR protein containing one or more mutations selected from Table 1, 2, 3, or 4 but not to a normal CFTR protein.
33. A kit comprising
one or more reagents that differentiate a normal CFTR gene or protein from a mutant CFTR gene or protein containing one or more mutations selected from Table 1, 2, 3, or 4.
34. The kit of claim 32, wherein the one or more reagents comprises one or more nucleic acid probes.
35. The kit of claim 32, wherein the one or more reagents comprises one or more antibodies.
36. The kit of claim 32, wherein the one or more reagents are in a form of microarray.
37. The kit of claim 32, wherein the kit further comprises reagents for primer extension.
38. The kit of claim 32, wherein the kit further comprises a control indicative of a healthy individual.
39. The kit of claim 32, wherein the kit further comprises an instruction on how to determine if an individual has CF or a CF related disorder, is at risk of developing CF or a CF related disorder, or is a carrier of a CFTR mutation..
40. A computer readable medium encoding information corresponding to one or more mutations shown in Tables 1, 2, 3 and 4.
41. A method comprising steps of:
(a) providing a sample obtained from a subject;
(b) testing the sample for the presence of a mutation at a pre-determined position selected from Table 1, 2, 3 or 4, in the CFTR gene or protein; and
wherein the presence of the mutation at the pre-determined position indicates that the subject has an increased risk for developing CF or a CF related disorder or is a carrier of a CFTR mutation..
42. The method of 40, wherein the testing step comprises determining the identity of the nucleotide and/or amino acid at the pre-determined position selected from Table 1, 2, 3 or 4.

WO 2011/119531

PCT/US2011/029314

4

43. The method of claim 41, wherein the presence of the mutation is determined by comparing the identity of the nucleotide and/or amino acid at the pre-determined position to a control.

44. The method of claim 40, wherein the method further comprises a step of determining if the mutation is listed in Table 1, 2, 3 or 4.

5

1 aattggaagc aaatgacatc acagcaggtc agagaaaaag ggttgagcgg caggcaccca
61 gagtagtagg tctttggcat taggagcttg agcccagacg gccctagcag ggaccccagc
121 gcccagagaga ccatgcagag gtgcctcttg gaaaaggcca gcgttgctct caaacttttt
181 ttcaggtgag aaggtggcca accgagcttc ggaaagacac gtgccacga aagaggaggg
241 cgtgtgtatg ggttgggttt ggggtaaagg aataagcagt ttttaaaaag atgcgctatc
301 attcattggt ttgaaagaaa atgtgggtat tgtagaataa aacagaaagc attaagaaga
361 gatggaagaa tgaactgaag ctgattgaat agagagccac atctacttgc aactgaaaag
421 ttagaatctc aagactcaag tacgctacta tgcacttggt ttatttcatt tttctaagaa
481 actaaaaata cttgttaata agtacctaag tatggtttat tggttttccc cttcatgcc
541 ttggacactt gattgtcttc ttggcacata caggtgccat gcctgcata agtaagtgtc
601 cagaaaacat ttcttgactg aattcagcca acaaaaattt tgggtaggt agaaaatata
661 tgcttaaagt atttattggt atgagactgg atatatctag tatttgtcac aggtaaatga
721 ttcttcaaaa attgaaagca aatttggtga aatattttatt ttgaaaaag ttacttcaca
781 agctataaat tttaaaagcc ataggaatag ataccgaagt tatatccaac tgacatttaa
841 taaattgtat tcatagccta atgtgatgag ccacagaagc ttgcaaactt taatgagatt
901 ttttaaaata gcatctaagt tcggaatctt aggcaaagtg ttgtagatg tagcaottca
961 tatttgaagt gttctttgga tattgcatct actttgttcc tgttattata ctgggtgtgaa
1021 tgaatgaata ggtactgctc tctcttggga cattacttga cacataatta cccaatgaat
1081 aagcatactg aggtatcaaa aaagtcaaat atgttataaa tagctcatat atgtgtgtag
1141 gggggaagga atttagcttt cacatctctc ttatgtttag ttctctgcat gtgcagttaa
1201 tcctggaact ccggtgctaa ggagagactg ttggcccttg aaggagagct cctccctgtg
1261 gatgagagag aaggacttta ctctttggaa ttatcttttt gtgttgatgt tatccacctt
1321 ttgttactcc acctataaaa tcggcttctc tattgatctg ttttctagt ccttataaag
1381 tcaaaatggt aattggcata aattatagac tttttttagc agagaacttt gaggaaccta
1441 aatgccaacc agtctaaaaa tgcagttttc agaagaatga atatttcatg gatagttcta
1501 aatactaattg aactttaaaa tagcttacta ttgatctgtc aaagtgggtt tttatataat
1561 tttcttttta caaatcacct gacacattta atataggtta aaaaatgcta tcaggctggg
1621 ttgcaaagaa aatgtattac aaaggctgct aagtgtgtta agagcatact catttctggt
1681 ctccaaaata tttcataagg tgctttaaga ataggtatgt ttttaaaagt taagttccta
1741 ctattttatag gaactgacaa tcacctaaaa taccaatgat tacaaaactc cttctggcct
1801 tctggactgc aattctaaaa gtgtaaaaaa catattttct gcattaagtt aggcagtatt
1861 gcttagtttt caaagtggta ggctttggag tcagattatt ttgattcaga tcctacatct
1921 actgttttagt agctctgttg cctgaggcag gtcccttaac atctctgtgt gtgacttgac
1981 ctttaaaatt tggagactgt cataggggtt aatcccttga gaaaatgaat gtgaaaagtt
2041 agcctaattg taactgctat tattatggat taccatattt tcacattcat cacagtacat
2101 gcaccttggt aatataagat gctcaattca tctttgagta taattttgtg actctcaatc
2161 tggatatgca atgagtgggc ctgtatgaga atttaattta tgaaaaattg tgtttcacat
2221 ggccttacca gatatacagg aaacacgtca catgtttcta ttgtatgttg ttaaagcct
2281 tagaatttaa ctttctgaat aggatccctt cagtttgaga gtcataaaag agtaaaatta
2341 ttatggtatg agttatagat tgtattgaat atctctttat atgtctaggt tttgtcattg
2401 gaaaaccaa aagtttgga aaaaaatcta agttatttct tactttctta attttgtgtg
2461 gatttcacat caagtataaa atttgaagaa catctgaact atcataatcc atatatatat
2521 ataaaataaa cataatctaa gagagaattt caccatgaaa aattcaggta gttcatgact
2581 atcagagcaa acaagtacat taaattgaaa cttttatgaa aataacattt atgaaatagg
2641 aagctatttt taaaactagaa gtgatataat agcatataat ttataattca tatacaagtg
2701 ggattgattt ataaatggtc accaacagag attgtgctat ttaatttggg aaaatttttt
2761 aaatttacat tttctcacia cttttaaggt agttattcag tttgttctc tctgtctctt
2821 ctctcatgcc ctgaattttt catatttctg ttagttgtaa gagtgtatat caaacctgt
2881 gtcacatgac ataacttgaa ttttctgtgt gatatctgtg ctatgtctag gtctatactg
2941 aggaactgtg ggaacccccc agaatccaag tatacagtgc cactgatttc ttacaaggga
3001 tgtggggtct cctgtaaaact ctgcagttag tctcaagtaa gaccaaagag taaaatattg
3061 ttaggatcta aggtggaaat tcagcaaaga atcacatagt ctaagtctcg agtttaacag
3121 taagataatt tgagatactt ttgtaattat taaacacaaa gtaatgagag attttaaaac
3181 aaacaaatac acctgaattt atatatcaga ataggtatgg tggttcaaaa tagctatcta
3241 ataaaaacca cactcctatt ctaaacattt gcctttgato aaaataattt tgggtctctt

FIG. 1

3301 attatgaaat tgcctttcta aataatacat aaattttotc tcataagtat atattagcca
3361 cattatthta ttgttattgt tttatattca tagcttgott tagattaaaa attatattac
3421 ccagactggg ctcttggaact tgcttccaag tgacttttga ctgtatcaca aaatcaaatt
3481 cactctgaaa atataaagat ttttcatcat aatttccttt gttaacagcc aagtgtacc
3541 taatttttagg tgttttcatt aaaaaaaaaat gcattgcaaa ctttaaagac aattcttttg
3601 tttgtttgtt tttaaaagac agagtctcac tctgttgccc aggctagagt gcagtgacac
3661 aatcataact cactgcaacc tccacctcct gggctcaagt gagccttcca tottgccctca
3721 cgagtagctg ggtcttcagg tgtacagggtg tgtaccacca tgccctggcta actttttttt
3781 tttttaagtt atatagagac agtatctcac tatgttgccc aggctgctct tggagctcct
3841 ggccctcaagt tatcctccca ctcagtctcc caaagtgtctg ggattacagg cgtaagccac
3901 ctcaacctgt cagcctaaag acagtgttta atgaagagaa atataagtgc tttgagcaat
3961 ggaagtataa ttaaaattat actatgaaag atttataaag atgaccattt tgaatgggac
4021 cacacttatt tggttatata aattatgata cactattaaa aattcatcat gatgattttg
4081 tattttacatt ttattttacat gtttgcaatt tgtgaggaaa gctaaaatta tggctaagcc
4141 ataaatattt ttgcagtttg ttgaggggtg ttgtaaaagt gttgccaagg aagaccagtt
4201 ggctacccaa acaaggggtt agtctagggtc tgatcaatac atacacatta totcagggtt
4261 gtctatcaga aaaaccttag gttatccaaa tcaaaataaa atagatgcat aaaacaaagg
4321 ccaatatgtg ttgaacaatt atattgtgat atacaactgc caagcattcc cgattaccat
4381 gactccattt agtcagtcca tgggcaaatg ccatcaatga ggacagccca gggtttccat
4441 attctctctt ggctttacat cctataggaa ttggaggggc ccacctctgg gataggagcc
4501 cttctgtctt gaacaatgtt gtctgaacac taacaaatgt tgactttcta caccagtccc
4561 tcaatagtct tttctattta tcttttgct gacctgttt tgttattaca cagttgagat
4621 ttttcagctg ggaatctgtg ttaattttgt attaattttg attagcttaa ctctcagagt
4681 tctaaaagta cctcctgtac ctgatatatg acaaaaatta taattacatt tatttatata
4741 taaaatatct ttgtatatgt aaaatatctt tgtatatata attatataat tgtttctttt
4801 aattttgcaa attttaaaaa gttctccttt gtttggaagt ttattcctat agttttttat
4861 atgctagtta aattattaat cacttgattc aagtaatatt cttatatact tataaggaat
4921 agtgtagttt taatatttaa ttccttgcta aagagagaag tggaatctat tttcttagc
4981 tacttcatca atattttatg tttgatgtga cagtcaaaat atccctcaga gctaactgtt
5041 aactagggga aatcacgggt ttcagtttt ccatttatgt gttatgggag ggagtggac
5101 ttagtgtaat aatattcaat acataaatgt taacacttgt ttaaagggtc ttgagtgagt
5161 actgctataa aatgcattat tattgctagt gtcatttcac aagagcctat aatttcagtg
5221 tgatagagct acaatataag tatagtattg caaaaccatc aggaaggggtg ttaactattt
5281 agcatgcagt tatgtgttggt ttgtcaaaac gttaaaaaca tctctgaact agcagcaatt
5341 ttggcaattt tgatcctgag gcatctgtgt agggcatctt cctggagaaa aacctctgag
5401 atgcaatgag gtcaaaagggt gaaaacagac tatgataaag atcaagttgt ttggagatct
5461 tgtagaaaga ttaatttaca aatatgtcaa gtgcattatc atggaggaaa acattgctat
5521 ttctgttggt tctcttcaga gctctagaat caatttacca catagtgtgt tcagtgtgaa
5581 attagcatta cagagtggct ttacggcctt actgtagggc attgtgtcag caaagagctt
5641 aggcctcttt tagcaagaag cttgtaaaaa tttaatthac tottagattg ctgatgtag
5701 agaattacat tctacagag ctctgaaaaa tcttttttca gagtttttca cagctgtatt
5761 caagttgcaa ggcttggtcaa ctttgctatt tttctgtgca gctctgttaa cttattatta
5821 tcttttgaca taaattatga ttccaaattg taaagctctg gatgtcaggg ccttttctaa
5881 tttgtttagt atgatattca gaccatttca agactcttcc gtggaacaat ttaataaaga
5941 tttttttgtg atgttaatga gttcatgggt atcaacctta gagacctgtg tctattgtag
6001 atcgatgaca ttcaacagtc ctgcagtgtc ggcatcattt tgataaaaag gggctaaaagc
6061 aagtgggact gtgggcagat ttttaatgct tagaacaatt attccatcga agttttcttg
6121 tgtcccttct gccttagcct ttgtaggata gcatgcttgc taatttcttg ctcatggggg
6181 aaggaaatga agatttttgc taggtccgta ggattattag gactactcag gcctgaagct
6241 atgcctggat atagccagaa aactctccca tagcttgctc caaggagctg agatacagca
6301 gtacttcctt tgtaggctcat gattctgggt aacctggaag atgacctcat tcatattctg
6361 tattctatgt gagacgttaa gaaggtagag gtggccaaga aggaaattgt tgctgccttt
6421 atggaacaaa ttatctgaaa ccagctttc tcgagggctt cattgaagta ctcaactggg
6481 gcacttaacc cagtctaagg ctggtcaagg aaggcttgct gggggaagtg tcttttgtat
6541 tcacacctaa aggaggttat tcaattagaa ttatccaaag agggtaggga tgggctagga
6601 aaaattttaa caggtagtgt ggaggactga caggataagt aagcatggca ccttcaaaat

FIG. 1 (cont.)

6661 atcctgagaa gttccctatg acgggaacat aaaatatgtg acagagattt gtgggagatg
6721 ggtctggaaa ctctagcagg ggccagatcg taagggggct ttgtaggctt tgtaggcttt
6781 gtttgggctt tatcatactg gaagtgaaaa gccatggctt ttaaacagga gagggacata
6841 atcagttcat atactgttgc agttttgtaa aagaaaagat gagctgaaag agtggccatg
6901 gtggaggtgg gtggggtggg ggggaggggg cggggagaga gagagagaga gagagatttg
6961 aaagacattt aggaggtaaa atcaactggg ttggtaatca attagtagtt gaaggagaag
7021 gaaagagaag agttaaggat aacatctata tttgttgatt tggataatag aggggacagt
7081 ggtgctgctt attgaatgag aaaatttaat cggagaagaa ggcatggagc aggagtgcag
7141 acctatgtga ctctacttct ctcaaaacca gaaacggaaa tgatgtatat ggctcagggg
7201 taggtaatat ggttatttga aaatgtatta aagtgattta gagcttagtc ttaggtaaga
7261 gatataagat gtctgaggtg acagttttat aaatatgtag agtgcccact tgtttggcct
7321 tattgtggca tagtgtgacc tgagagtgtt aggaagaagc agctgagttc tagggacagt
7381 actgggtaaa ttctacttag aaattatact tagaactctc ctatataacc tgctaactga
7441 tgtctgaacc tcctgataac ttcactcctt taggcagtgc tttcacatc acgggacaca
7501 acatatgaga gatcatagaa attcaatgtg gtatgaaaat ctgcttggga cttcagatat
7561 tgtctccagt gattgaataa aaataggagc tcacctacta tgatgaggtt tctgtgtgtg
7621 ttaaaagaag gttttcatta cttttgaaaa ggttatgtat ccttgtttta tgttaaaact
7681 ttgagctttg ttaaatatgc agagtctctc ttcttagcat ggactacaga ggtgcaacta
7741 cctcctacct gacttcacat ctactcccaa atgcctagtg aaggcttaat aatttcaaaa
7801 agggactcta gaatttcatt tgataccagt cagacaaatg tgtgaaaatt aagcataata
7861 ggcagaatcc caggggtact gacagctgta ttaagaggtg attcaagggc taaaccttag
7921 agtccagcat tggttatggg tgtgacaaga aaatgaagcc tatgltggct gggattagca
7981 accacagttc tagaggaagc aagggtggaga aactatatag ggggctccct ttgtacgttt
8041 tatttatttt aaacatctct ataaactcta gaaattaaaa caacaatacc aacacaaaag
8101 catcactttt tcgaccaaag accattgcta tacttttttg tgtaaagggc tagatagtaa
8161 atatttttcag ctttgtgggc cacataagtc tctgcaatag acaatatgca aacaaataag
8221 catggctgtg tttcaattaa actttattat gaacattaaa atttgaattt catataactt
8281 ttacatgttg caaaatatte tttattttaa ttctattgca atatgcttta aaagatacag
8341 ttttttagtct ttcttagttt aaaataaaat ctagaaaaaa ttttaagtct tctataactt
8401 tttttcggta actgaataat tttaaaagta agtgaaacat ttagacatgc aaaatggact
8461 tttcagaaga agaaaatggg agcttaacag ttattagatt attgtccaga ataatttttg
8521 acttataagt ctctgttgac catttcattg cctctttttt tggaatatgc atcttttaat
8581 gtgtccttca aggcaaaggc tctatcttat ctatcttgtg tcttgcattt tcccagggca
8641 atgtttttca caattttttt aaaaaacaat actgtaatca attttcaaat aaaattttcc
8701 atgggaccgc agtgtataca aatagcagtg acaataaaaag ataataactc tcccataaat
8761 acaaagaaac agttaaccta gtgctctaaa gtaaaggcta cagtgatttt gtataacatt
8821 tatatgtaat tttcttgatc ctacatgggt gtgtttttca cagtgttatg tttctgaaat
8881 cgagatgcct tttataattg atgtcaaaaag aaacttgtca gccacaaggc ccaggaataa
8941 gttgtaatat gggaacttag caatacataa aggtatatat actcctgtga cctcagctga
9001 attatttgca ttggttgcat cccacaaggg tgactcttaa ataaatttag tttgttgctt
9061 gaaattttctt gggataaatt actttgtgat gtagttttga aaaaaaaca ggtaatatth
9121 agtctgaagt ttgtctgaca tactaagcaa tgtaattaaa gtagaagtcg cctaagctca
9181 gcactttatt atgccttgaa attatactgc ctgtccaca ggtgaagggt ttatgaatgc
9241 agtttgtcac tgtaactcta ttcatagctc tgaaaggctg agagtgactc agaagaatat
9301 ttttgctctg aatatgaaga acgcttagac taaaacttta attacgatgc tgaagaagaa
9361 agtggtaggt gattgcatga ataagtatgt aatattgtta atttctaaaa actgtgtata
9421 gttaatgtag tgcttctttt tggaaaggct attgttaaat tgatggtaaa ttctataacc
9481 aatatcacct taaagcaagt acgcatgata aagtattata aaaccatgat aatatcatat
9541 gtggcttatt attgttccct gagtgttgta caactctgtt atgctgtgat gaaacctcat
9601 gcaaacagggt atgtcaaaga tatgatgggc tgttaactga gcttggcca catatggtgt
9661 agtgacatgc tcactaatgc agtgagaga taaccaataa cagatcataa caggtttaaa
9721 tatgtgcaag gagatgtcag cagaagcttt cctacatagt gaatactaaa caagcctgac
9781 agcccaggat catgttcgga tcaatctagt gtgctaaaaat taacatatag tctacattt
9841 gagaatgtgt gattttcttg gttcctgtct ataaaaaat attttaaaat acatacattt
9901 caaatcagaa gttggtgaat tcaactgaaat atttctagag aacactaggt attggggctc
9961 atagtgtgaa aaccactgac ttaattcttc ccccatcttg gttgttctg atcttccctt

FIG. 1 (cont.)

10021 gtgtcccccatt tccagccatt tgtatccotta gaaaatgato tcatattcta cttcatcttt
10081 atcttcattg tcaactgtca ggtagcaata tatgatggaa gaagcatgta ctttggaatc
10141 agacagacct ggctggaato ctaactctgt cacttattaa caatgtgato ttaggcaatt
10201 tacttaatct ctctgaacct cagctactct cgtcagtaca atgagttatc cttatcttta
10261 catggcacag tattattatg atatcaaaaa ttcattgagt atttactctg catattagtc
10321 aagggttctcc agagaagtag aaccaatgat acacacacac acacacacac acacacacac
10381 acacacacac acaattttatt ataaggaatt gacttacatg attatgatgg ctaacaagtc
10441 caaaatctgc agtatgggtc agctggcagg aaacccagga gagtcaatgt tccagtttga
10501 gtctgaaggc agtctgttgg ggaatttcgt ccttctcttg gaggccagcc tttttgttct
10561 atacaggcct tcaaccgatt ggatgaagtt cacctttatt agtgagggca atctgcttta
10621 accaaagttt actgatttaa atgttaatct catccaaaaa caccaccca gttgacacat
10681 aaaattaacc atcactctct gtaagcactt tctatgcatt aagtgatagc aaataatgcc
10741 agacataggg cgtctttaat aaatggtaag cactgttato agcaacaaca ggattattat
10801 aattagcacc ttttcatctt tctgtctggg ctctgagaaa gtacctctct tctctaaatt
10861 tatccctcct ttctatgaa ttagacccag tgctttctct gaattatgaa ggtcacactc
10921 ctacaaatgc cccttcccaa ttgcacatct gtgggtttc tttgccattg acttttatct
10981 ctagctttta aatttacagg catatgtcag ttaacaatgg gaatgcgttc tgggtaatat
11041 gtcccttaggc aattttatcg ttgtgagaat actatagagt atacctacac aagcctagat
11101 gtcgtatagc ctactacaca cctaggcaat atgacatagt cttttgcttc taggctacaa
11161 acctgtacgg cttgttacta tactgaatac tgcaggcagt tgtgacacag tggattttgc
11221 atatoggaac atgtctaaac acagaaaagg tgcactaaaa atactatgta gtgatctcat
11281 gggaccacca ttgtatatgc agtctgctgt agactgaaat gtcatgcagt gcataactgt
11341 atcttaaata ctcaaagtat cacctttgtt tgtttgtccc cttgtgtgca tcatcctaac
11401 gtggaatttc tctgttgatt agggccagcg tattagtttg ctagggctac cataacaaaa
11461 taccacaaat ttggtggctt aaataacagg aatttattat cttatggttt tgaagactag
11521 aagtacaaga tcaaggtgtt ggcaggtttt tcttctaagg gccatgagga agagtctatt
11581 ccatgccttt ccctacctt ctgggtggtt gctagaaatc cttggcattc cttgacttac
11641 agaggcatca ccctgatctc tgttttctac ttcacatggc attctccctg tgagcctgtc
11701 tctgtgtcca aacttcttta ctattaatat aaggacacca gtcataattg attagggctc
11761 acttttagtga cctcattgga atgttattac ctctgtaaaag atcctatctc taaataaggt
11821 cacatcctta ggtaccgggg gttaggactc aaacatacct ttttttgggg aaacacaatt
11881 caacctataa caattgataa cactctttag gagcagaatg cgatatggaa gtaatttgag
11941 accataaagt atatacatgt agggagttaa tctatgaaac ctattgaaag ccatatatac
12001 ctcatgtata gtgggtccata aatagcatgg agacattgca gaggatgtta agtgatatga
12061 tacaggaaca atccaagaag gtcataagaa aaaggacctt ttgctottga gaggactgaa
12121 gaatgacttt ccattttatga aattttggta catgtccact aaaaatagga tgaaggccaa
12181 acttaggaag aatatttttga taatggagaa ggttgcatat aaaaacattt tattgaggac
12241 aattaaataa tgttggctgg aagttttagg atgatcatct ttaggactca gaaaaagaga
12301 agaaacatta tttaaagaatt gtccctgaac aagtataggc acctcacat ttgcattgca
12361 ttactatag aattgaaaaa tgttttgacc tttttttttt ggcttttaat atatttgacc
12421 aagagtaaca gctaagcaat acctatttgc aatcagtgtc atcatgtggg ctccaaacat
12481 atcatgtttg tgtaattaat tgattgaccc attaatattgt tcaatttctg ctctgttcca
12541 ggcactgaac aacatgatgg agataaaaga taaatattac acctgccttg tcccaagaa
12601 gttagtcttc tgagggaag aaattagcaa acaaattgta atctcagtta tgtgccatgt
12661 tccatgctgg gcacagggga tacagtagtt taaaaaaaac acaagatcta taagggtgtt
12721 cttcttgtgg accttacagt ctagggtgct tggaacatg gggcgttggc agacaagtaa
12781 atacacattt tgtggtaaag gctcaggtag aagaagtaca ggatagaata gagcacacca
12841 tggggaatta atctagactt cagagaggct cacacataca taatttatgt gtgactattt
12901 caatgcattt gaggtttctt ggaaatagag gttaggtttt attttaagga agttaccatt
12961 ttttttttca gtgtgatgtg gttgaaccaa agaatgccat gccagtgat ggtaatagga
13021 taatcttttt aaaaattaag agccacctaa taaatcaata gtttcattca gcgggagctc
13081 ctgcagagtt caaaaagaag agaactctggc acagcgtttc ctttaaagtt cattttccta
13141 gagtgtgaat ggaagcaaga gattataaca ttttgaggtc aaaaaaattc tgaaatgcct
13201 ataaaaatta ttttctccaa attatcatca tttgtgcttt taatgacctg attgcaaaga
13261 tgaacatttt gaattcttaa attgcttatt aggattggtt aatgaatcaa ttatctatta
13321 ctgtatgttt tgctattgga aaaaatagca acttaagtgt tttgcagacc tttacttagg
13381 tatatgttgc ttttatgaaa aaaaagatgt aatatattaag taaaagggat ttaaagcaag

FIG. 1 (cont.)

13441 gctttttgagg tagagtctta ttaattcctt ggtaaaccctt gagccaattg ttgtctatgt
13501 tctctgcctc tgtcttgctc cttccttctg ggattcactg tgggaatgcg ggattgttaa
13561 tctggggatg ctgtccaatc ctgcctctct caagctttgc tattgatctc cctcccagtg
13621 ataataaagc ttgaagaaaa tgaaagtagc gttagtattg gtccotcaaac tcaagaacag
13681 gatgaaactt aaatcttgag tcatacaatt gtgtctacat actgctcccc aaaaagagaa
13741 gtaaagaaga tgctaacttt cctttttaat ttgcagtact tagcaatttg ttttcttgag
13801 ggttaagtaa taacagtgga agaaaaaagg gttaaaatgc caccaagaac ccaattccat
13861 gtttagtttg aaagtgggaa atcagctgcc actgggaagt ctgaatccaa tgccatgatg
13921 ttctttgaat cttcttgaga aataatcatg tgtagccata acatacctgt ataacagagc
13981 agagaacata aacaaatgaa ggtgaaggga agattaagac agaagagaaa aattccagaa
14041 tcgactgatc atttttatct gtttagatga ttccaggcag aatcctagag accaacttta
14101 tcacaactga attttaaaaa tcaccagctt tgtcattgtg atgcagcatc agtttcagta
14161 ttatccttgg agtattaatt cttaatcatc ttcactcttag aacatttttg aggtcacttc
14221 tagtctctat ttcaccagtg aagaaacaaa aatccccaaa ctatatcagg tgggaattaca
14281 cagtattttt tttttaattt tggggaaagt cgattcaagg cagtaacttg caagctagtg
14341 ttagaaaagga ttttaataaat agtggttttt ctgtacacat agtgagaggt cattacatca
14401 tttggttgtt gaaagtcata aggatgtcta gcatgcgctt tgccctgtagt ggttcatgcc
14461 aggcagattc ctgactccta taaccagag cttatcagag catttatgtc cccaaagaga
14521 aatgtcacct ccatctttca ataaacactt tagcaaagaa aaatcaagta ctttaattcc
14581 aaatcttgag ttaattccag aataacaatg atggctcgga aaaatatggg tatttctgtc
14641 aaaggacaga gaaacctagt agagagtatt tactttgggt cctagtgatg gtatctgaac
14701 aagctaggtg aacaaagagc ctcaataagg gattttgagg tctagaaaaa gagaggaaat
14761 accaaataaa tggaataatt ataaaataaa taaccagcaa gttaaataca tatatcatgt
14821 gggagatata cttatatcac tcatgtgatt tctattttgt tcctatatta ggccaaggag
14881 aggtggaact tgttttcctt tttccctctc agctacgaat ggacatactt aaaactgttt
14941 ctctgcttct gttctctaaa atgtgattgt ctaacagtaa ccgtgatgac gttttgacag
15001 ttgcacaagt ttctttcttt aagctttaaa aatgccagcc agtaaccag tggcatttct
15061 actataaaat cttaaggcca atccatttcc ccttttcctt attttcttgg tttcaaatat
15121 atttttattg ccaatggaaa taaaaatcct aaattagaga gcaatggcat ccttgtctt
15181 gtgaataaag agctcctaaa tgtgaactta tacaggatgc agcaatttat agggtagtta
15241 atcattcttc tttctagcca gttgttccag ctacagtttt gtggctcttg ttagtggctt
15301 cattcccaga tagaataaaa atcaaaccaa aatcctggaa aggcactctg aggatgcttc
15361 tctaaagtag atgggcatca actataaaat acaatgcctt gtttctctg ttatgtttca
15421 agatgggtgg gatttttttt gtagcattac ttattattgc ctctcaagtg cttgagcttt
15481 tgaaatccaa gtcatgtgag tgaattagat acagctgtta gaagtggcct ttcaatgcca
15541 atggtacaca ttcccttggtt tctttacgat actattgctc ttacaacttt tatctgaagt
15601 cataaattca tagttgtccc agaagttaag ttccctgctt ctagaggaca gaaaacaaac
15661 aatttacaca actcatggtg catgtcacca gtccttagat ctcatgaaat atgcatgaaa
15721 tcttaaatca cttgctgtag ccaccagacc attgacatat ttgaaagact ttagtgtatc
15781 aaagtcacta taatgaaaat tttgatttca ccagttctag gagtgaaaaa tcaaatgttt
15841 agtaaaactt tctaaaatta aactgacag ttgatttctg tatactgttg ttcttaataa
15901 tagctttatt gagatataat tcataatcaa aacaacttac ccatttaaaag catacaatcc
15961 aatgattttt tagtatcttc aaagagttgc ctatcaccat aaccaatttt agaacacttt
16021 catcactgta aaaagaaaact ccattcctat tagcagtcac tccttattcc aaatccccct
16081 gctcgcccta gacaactaca aatgtacttt ccatctctat agatttgcct gttctggaaa
16141 ttttatgtaa atagaacaaa gtgttctttt gtgactggct tatttcaactt agcatttttt
16201 ttcaaagatt catccctgtt gtagcgtgta tcagtgcac attctttttt attttttttag
16261 agacagggcc ttgctctgtt gccaggttg gaatgtgcag tggcatgatc atgggtcaact
16321 atagctttga agtcataggc gaaagcggtc ctcccacctc agtctcccga gtagctgaga
16381 ctacaggctt gcaccacatg actgtctaatt ttataatttt ctttagagac agggctcttgt
16441 tatgttgtct aggtgtctct caaactccag ggctcaagtg gtcctcctcc cacagcatcc
16501 taaagtgttg ggattatagg tgtgagccac agcacctggc ttgcatcatt ctttttattg
16561 ttgaataata tcccacttgt aagaatatgt attttattta tcttttcccc agttaataga
16621 tatttcgatt gttcctaatt cttgtctatt ataaataatg gtgctatgaa catttgtgta
16681 caagtttttg tgcagacatc cattttcctt tcttttgggc atatacctac gagtgtaatg
16741 gatgggcat atagtaactt tatgtttaat attttgagga tttttcaaac tgttttccaa

FIG. 1 (cont.)

16801 agtggctgca tcatttttaa ttccttccac catttgtgtga gtgtttcaat ttctccacat
16861 atttgcaaca cttactatta totactctta aaaattacag ccatcctact gggcatgaag
16921 tggatatttca ttgtgagttt tttttttctt tttctttttt tctttttttg ctaatgtttg
16981 tggatttttct tttcattttc ttgatgggtgt cctttgaagc acaaaagtat ttaattttga
17041 taattttcaa tttatttttt gttattgctg tttgtgcttc tgggtgttga tctaagtgtg
17101 tgctacttta aaaaattagt tgtaatatgg caaattggat acatgtgtag gctttgggtg
17161 cacaatccta atttttaa attgactctg cccttgacaa attaactaat taagcttcct
17221 tagcctcagt ttctcaactg taagtgggag atattacca gacctacct ttgaattggt
17281 gtggggatca gatgaaataa tgtatgtgaa atatttagaa ttatgcaagt ctgtggtaat
17341 gaatacta atgttagctatc attattgtta taatcccaat aataaattct ggtgctttga
17401 aaattaaacc aaagccaagc agttgatatg aagaagcatg taataatgta cagacataat
17461 gctttataga caacattgaa tttggctctc atgaacatca ggaatagtgg tcatggtagt
17521 tattatctcc agcaggaact gtagctgaga gatcttcaga gctttttcca aggcgatctc
17581 actgggaaat aatagagaca aggttacaag ctagggtgtg gttttcttct taaaatcttt
17641 agttcagttt ttttcaataa cagatttgta gtaggcatca ggtgactggg gattcgtatt
17701 cttcaagttg aaatattacc ttgttgagaa agaaaccatg tgtgagacaa ccatgttgag
17761 aaagaaaaag tgattttata gaaaattaat attgatagtg agcattatat gaaaatcatg
17821 aagttagaac atatttggcc agaaaattta cattaatagt taccatagc aattaatgca
17881 ttataattac acataccttt tctttaatga aaaagaattc tttccttcca agttatgca
17941 tgctattggt aaacattaga gaatatagag aagcaaaaaa gaaaatatct tttttgatat
18001 tttcttaaca tacgtctgtt cctaataatg tttatagttt agaagcattg catgaaatgg
18061 gtagatcaat tttctattta atgtttggat tcattaggta cgaagttagc aaattaattt
18121 ccattagggt gcctgtatgg ttgtaaatcc tggacctgca gaagattttt cagtattggt
18181 ttgtagtctt ttgttttagc gcaaataatt agttctccag agcttctgaa attaattgac
18241 cactttaatg gtgtttacct acctagagaa agaaaaagaa cttctccaag tcccttggtg
18301 aaattaagcc tcatgaacaa ttaactcaa tatacacaag gcttgtcttt agcgagcata
18361 tactccctaa agttgattaa gctgaccaag tgattactgc ttataaattc accattttat
18421 ggagaagaag caaacactgc taaatacctt gtggaatcag aggaggggaa attagtaact
18481 tgaccccaat actgcgattt taaattgaat tcttgaagcc tacaagtttt acacaggact
18541 ttagagagct ggatagtatc accttgctaa gtccactctt tactatgatt ctttgagaaa
18601 aatacatctg actaaataac tctgaatcta aattggataa aataaatgtg acattcaaaa
18661 tgttattttat gatttttagaa aaatatcctt atagacacta gatgagtttt agtctcaaat
18721 caatcctccc tatcatagtc acctatcaaa ataactaaag caaagtggta gagctgtgct
18781 ctagaagttt gggattttatg atcacaatct tttccaatga gtccctctct tctctgct
18841 gtcttcaaca tttgtttttt tttttttttg gttaggacta tccagattgt gtggcctatt
18901 tcaaaactcat ggcaaataca ttggatgata agaaattttc taatgtattt gaatttgtct
18961 acacaaacta gagtaattgc tattaattcc tcaagtgtta attatttcat gcaaaaagga
19021 aaaaggctat tagtctttta gtgtattagt atgtcaatat ttgggagaag tgtcatgcaa
19081 ttagtggttt gaatttccta ttttatttta ttgcatttta ttttatttgc ctagtcaaat
19141 aaaaagtaat gttaaataca tgggaagcatg attgttttct acactaaaaa tcattttgac
19201 ttgaaaagat ctgatatcca tgaccttc atgaagtttt ggcagatgaa aatgtcagat
19261 gcgtcttttg gattaataaa aggcaaaagt cagatcgaaa aatgagtata agctttaatt
19321 atatgacttt aggaggatat gttatgaaaa tcaaagcttt aatagtgatt ataattggca
19381 agttcttttt ttataaggaa ttacaagtca ctctatacaa aaattggaat ttttgccta
19441 agaaatgaaa tttactatag tttcatctgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt
19501 ttaaaaaatc aagtgatagg gcttttctct aataaaatct gaaatctctt atagttaagt
19561 gaacagaaca gtgtatctag gatgctagac ttttttttca aagttagttt aaaacttata
19621 catagtaaaa tctgtatgcc ttagggtatc ctgtttgcta tcccatagtg aatgattaat
19681 tagtttctgt tagaaatagt cagaactagg ctgggtgtgg tgggtggctc tgccgtgaat
19741 tccaggactt tgggaggcca aggcaggagg atctcttaag cccaggaatt tgcaaccagc
19801 ttgggcaggc tggtgagatc ctatctctac aaaaacaaac aaacaaaca aggacaataa
19861 gaaagaaaga aatagccaga gctttgaaca aaatttctaa gtagaccaat gtaaaagtct
19921 gtcgtcaata tgtagtggct atgaatggag gttatgaatg aaagagaagg ataagatgaa
19981 ctagagggtg gaggggaaga cagcaggccc aagtgaaagg cagagccgag tttattgctt
20041 tttggttatt ccagggtgtg ctgctttgtc tcatgaaaca cctggatgat cactgatttc
20101 tagtggaaga aatgctgaaa agtccttact gtgcatttaa acattctagg tttaataata

FIG. 1 (cont.)

20161 tcaggggtttt tcaaaagaaa ggggtggctgg agttttgcac taactaatat ttcataaagt
20221 gtctaagtat agatgtcttg tttttttttg tattttotaag actggcctga ggtaggcatg
20281 gagaattcct tgatgggaca taattttctt cctttctttt tttttttttt tttttttttt
20341 tgagacggag ttttgctctt gttgcccagg ctggagtgca atggcacaa ctcggctcac
20401 tgcaacctcc gcctcccagg ttcaagcaat tctcccacct cagcctcccg cgtagctggg
20461 attacaggca tgtgccccca tgctgggcta attttttttg tatttttagt agagatgggg
20521 tttctccatg ttggtcaggc tggctctgaa ctcttaacct caggtgatcc acccaacctg
20581 gcctcccaaa gtgctgggat tacaggcgtg agccaccgag cctggcctga tgggacatat
20641 ttttcattca attttattga tttaacctca caaaataaaa tatttcctta agatgactct
20701 gtggtcattg ttgggcagca taagcttaat ggatttttagt tatcataatt taccttaaac
20761 ccaatttgta tttcaggata taaatagagg tttattgtag tgaatcttcc aggaaatact
20821 aagtgatact aataattata gatggtgaac ttaagtcttt atattactga atttgtttgg
20881 tttgatgatg ctaggctatg gcattcttgc taatcaaaac gatgtgtcat ggtgtaacat
20941 aacttattaa aatgggcaca gataacacag gaagcttttt ataaaagcag ctcacaaatt
21001 gtgttacttt gaactgaact ggccatttat gggaaaggtc actgggttgt aaataaggac
21061 caaaagagtt acgtttatat tttttaaaag agattgagga gatttatttt tacatttctt
21121 gaaaatgcct tattttggta tggattgac agatagtga attctgctca tttgtaaata
21181 tagtgtcata ttttaataat ttcaaacata ttgaaaatgc agaatttatt aatagtggga
21241 gcacattttc ctttttacta aatgttctac aggttctttt ctttccatcc acacacagtg
21301 ccattacctt cattctaagc ctttcaaaac tctggcagta agtgatctgc tgcacttagc
21361 tctttccagc tgagctgatt tttaaatttt cagaaaattt gtgagctaatt tgttaaacad
21421 ggccattatt aaaaattaaa ttatttcaac ttataattaa ataaattata ttaaaacaaa
21481 agtattaaaa actcaaaagt tggctgggag cactggctca cgtctgtaat ccagcactt
21541 tgggagaccg aggcagggtg attgcctgaa gtcaggggtt cgagaccaac ctgaccaaca
21601 tggagaaaacc ctgtctctac taaaaatata aaaaaatagc cgggcatggt ggtgcatgcc
21661 tgtaatccca gctactcagg aggctgaggc aggagaattg cttgaaccca ggagggtggag
21721 gttgtggtga gctgagattg cgcatttgcg ctccagcctg ggcaacaaga gtgaaactct
21781 gtctcaaaaa aaaaaaaaaa aaaaaaaaaa aaacaaaaaa aaaaaaaaaa caaaaagcaa
21841 acaaacaaaa aaacaaaaat tatcacttcc taattatttt gcattttact attatctatg
21901 ctattaacgt tatttgcctt cattgtattt gaaagggtga ctatattcta ttgcacttcc
21961 attgtactat attctaatat gcaactgtgt atcccttccc aactctgtgt tcaatgactt
22021 tatatttggg tgccttaaaa tgatgacgat gagagtattt atatcataga aattggcaaa
22081 tgccgtaagt cagtttttgt ttttgttttt gttttccgga gaggggattg ttaaataatt
22141 gcctgcatgc aacaccacta catgcagtct gctatctttt gttcttccctg ctttcaggct
22201 cctctcccag ctgtctgtct agcacaaccc agcataccaa atttttotaa atagggaaag
22261 ttgaacatgg taaaagaatg aatgaagtca aaagaatgtg gaaagacctt ggctttgcc
22321 tttagtaaag tttagcatct ctaagcctcc atctctttat caataaaatt gagcaatgat
22381 cccttttagt tctacccatt taagaagatt ttcaaataaa aaccacaacc tgctcatggt
22441 tatgaaggca ctttggaagc cgctaaatac acgggttttt attagtagta aacacttact
22501 tcaccttttt cacttcttga ctttagttta caagggtcca taatctaaat tatatcataa
22561 attgctgtcc cagatttttt tacagcctaa ttgccacctg tatgttcgac tttccttctg
22621 ttctttatgt tagatactgg gatagtatgc accagggtgg tgtgccatca ctttctcaga
22681 tgatgtccac tgaagacctt gcctgatcat ggcattcatt ttctgtgtgt attcagactg
22741 gcctcaacta ttttctttat tgctctccag gaaaaattac aaatgaatca gactgggcaa
22801 tgaagggtaa acctaatatt cgctctttgt taaagacagc tcttggttaa atgcggatat
22861 tgcaaattaa tggaaaaaat atgacatagt aaaccatact cacttattaa tatcttagta
22921 aggaataatt gatgaagtta cttaacctta gagccctaat tcagttaagt tttaatgaag
22981 gacaagtgtg agagatatcg agaaccaggc gcagggtgct actgaagaag ttccagacca
23041 aggaagtata aagaaggacc tgggtgggag cagtgaattt ggatatgagg gccactggca
23101 aagttttgcc ccagaacagt gtcaaaatgt ttgcatttgg catagccctt tctctttttg
23161 ttctgaatgg ctttgctaga atatcttttc tataatgaat ttatcctgct tctcagatat
23221 tgctaaagca ctcccttttg aattttgggtg ctttaacatg cattttgata cattaccaaa
23281 taaggctctga atgacacaaa ttttagaact ctccagagaa aagaaagatg ctgagggaaa
23341 aagcataggg ttgggactca ctaaatccca gttcaattcc tttctttaat aaatatatcc
23401 aattttacct gagaaagctc tctgtctctc gaattttatt tagaaatttc tctttgtaca
23461 tgattgattt cacaatcctt cttctgcctc ctcttctact ttcttcttcc tagattttcc
23521 tatctttatg aagattatcc tgccttatcc tcaacagtta gaaacaatat ttttgaaaat

FIG. 1 (cont.)

23581 cactacggta tcctgcatag tgattttccca tgccaaacttt actaattttcc attataaatt
23641 attattttatt gatgcctaga gggcagatga gtgtagctgc tatggagtga ggagacaaaa
23701 cataagaaag ttatgatcct accctcaggt aatgattcag acatgataat taagtcaaca
23761 aattgataga aactaatcac taactctctg gctatagtca ttctttcaat gaatagctca
23821 ttactgagta tgcattgctac agtaacaaaa ttatataagg ctgttgatta aatgttgatt
23881 aagtgcattgt cttattcaga gttttttttat atttgaaatg gaagaggctg gacttcagta
23941 atttgctata aactgctagt atatgattat ttgggggagc ttatttttta aagaataatt
24001 taaatatgga atgttttagca gtttgttttt tccctgggaa aaaccatact attattccct
24061 cccaatccct ttgacaaagt gacagtcaca ttagttcaga gatattgatg ttttatacag
24121 gtgtagcctg taagagatga agcctgggat ttatagaaat tgacttattt tattctcata
24181 tttacatgtg cataattttc catatgccag aaaagttgaa tagtatcaga ttccaaatct
24241 gtatggagac caaatcaagt gaatatctgt tctctctctc tttatttttag ctggaccaga
24301 ccaattttga ggaaaggata cagacagcgc ctggaattgt cagacatata ccaaatccct
24361 tctgttgatt ctgctgacaa tctatctgaa aaattggaaa ggtatgttca tgtacattgt
24421 ttagttgaag agagaaattc atattattaa ttatttagag aagagaaagc aaacatatta
24481 taagtttaat tcttatattt aaaaatagga gccagtatg gtggctaatt cctgtaatcc
24541 caactatttg ggaggccaag atgagaggat tgcttgagac caggagtttg ataccagcct
24601 gggcaacata gcaagatggt atctctacac aaaataaaaa agttagctgg gaatggtagt
24661 gcatgcttgt attcccagct actcaggagg ctgaagcagg agggttactt gagcccagga
24721 gtttgagggt gcagtgagct atgattgtgc cactgcactc cagcttgggt gacacagcaa
24781 aaccctctct ctctaaaaaa aaaaaaaaaa aggaacatct cattttcaca ctgaaatggt
24841 gactgaaatc attaaacaat aaaatcataa aagaaaaata atcagtttcc taagaaatga
24901 ttttttttcc tgaaaaatac acatttggtt tcagagaatt tgtcttatta gagaccatga
24961 gatggatttt gtgaaaacta aagtaacacc attatgaagt aaatcgtgta tttttgcttt
25021 caaaaccttt atatttgaat acaaatgtac tccctgggaa gtcttaaggt aatggctact
25081 ggttatcaaa caaatgtaaa aattgtatat ttttgagtac ctgttacatg ccaggtagaa
25141 tatctctctc cagccactct gaggggaaag catcattatc tctattttac agaaaagcaa
25201 actgaggctc agagagataa tatactttgc cagttaatga atgatggagc catgattcca
25261 gctgagggtc gtattgcctt gctctctagg aatggtagtc cccccataa agaatctctc
25321 agtttccttt ccaatcaaaa ggtaggagtc cttttgattg ccagtgacag aaacccaatt
25381 tactagctta agtaaaataa aggaacgaat ttattggctc atgaagcctg aactatgtga
25441 agacctaggt ggagaactgg ccttaggaac tcaatgggac caaggactca aatgccacct
25501 ggtggcattt gccttatgct ggttttattt tctcagaccg gaccagcttt ctacataaag
25561 tgggtccctg gttagaactc tttgctccta tctttaagga ccacgaaaga aggagccctt
25621 tgtccttggc taaatgtgaa aaatcccaga gactcttgag tcatagtgtc tacccttggg
25681 gccactcata gtctagaatg aactaggctg agtctcgtgc caacagcaca ggcctgatgc
25741 cagataaaaag ggtgagtga gggggataaa aaataagaca tagctactaa attattgcac
25801 caaagtaaaa acattgagtt gacttgcaat ttgtttcttt taattaaatt catttccttt
25861 ttttggcatt ttgaaggcaa agtaagatat taaactttat ttttattgat tttattcaaa
25921 gaattaagct agtgggagta gcagattcac acttctaaga tcaagggcca gcttctatta
25981 ttgaacactt ggtgtgtgca aatgccatga ggtagggata ctttgttttg ttttttattt
26041 tttattgggt tcgatctctt ttgtttatga tgtatcccca agtgccctaga atagggcctg
26101 gcatatggta tatactcaat aaatatttgt tgaatgaatc catgatggaa tgtgaaatgg
26161 ctagcattac atagaaacct gtagcattgc tggagagata aaatatataa acataatcca
26221 ttgcaggat attgacaagt tcaaaataat ataatgggta ttgaatatct aaatgtttgt
26281 tgttggtgtt gctgttggtt ttgagacaga gtcttgctct gttgccaggt ctggagtgtg
26341 atggtgcaat tttggctcac tgcaaaactc gtctcctggg ttcaagtgat tctcctgcct
26401 cagcctctcg agtagctggg ttacaggcca ctgcacaaa tgccctggct atttttgtat
26461 tttagtagat gtggagtttc gccatgttgg ccaggctggg cttgaactcc tgacctcaag
26521 tgatctgccc accttggcct cccaaaatgc tgggattata ggtgtgagcc actatgccc
26581 gctttgaata tctaagtttt aattggatgc tgagggaatg attaatacaga gtagggctgg
26641 gttaattgaa aaatgtgata catttgattt tatggccaga tagagaacat gaatctgaat
26701 ttgcagaatt atctggctta acattttttt ctttccagtt ttcactgtat ccccatgtt
26761 gattcaattt aaaaaatata cctattttac ttcaattcaa caatgctatg ccagtacaaa
26821 cccatacgtt ctattatttt tgttttggtt tgtttttgta tctccacctt gttacttctt
26881 ttcttataaa attgggtattt gaaatttatt gaaatatttt ggaagagtga cataccattt
26941 ttggtacttt gtacctctgc acccttggga agtgacctg gcttcacatt tcataactgc

FIG. 1 (cont.)

9/63

27001 cttgtgacca tggccctcaa gtggttgcca gatggttgaa gaacattaac ctatctggct
27061 caattttgtg accatggatt gaatcctcta cataactgca gtgtgcaaac cacacatccg
27121 ttccaagatt gtagtcagga tatgaacttt ttaagaataa aacttcttcc cttctgatct
27181 gggcctggta tgtggtccta ctagaaccac atcacctact cttggtgcta acaatttgtg
27241 gcaccaagtt gttcaagttt caccatttaa agaaattccc cgacctggcc ttctcctcag
27301 gtaactaccc cattctattt tttctttcat agctaacatt ctctgctctc ctggctctc
27361 tacttcaactt tcattttacat ctcagctcct gaagtatggg ttccaccatg ttccataaac
27421 tacattgccc agggtcacta gagacctctt atgaaatata acaacacctt tctacattac
27481 ttccgtgtgg accacttttt cacattgaac ccattttgtt ggtttatgta cacacccctt
27541 ccttggtttt cccatctgat ccattttctcc tttgatggag aaggtagctc tgctccatat
27601 ttagcttctt actctgagta accaaatgtt atggatggga ggtagctct gtgtgtgaga
27661 gaaaggtgga gaagcatgtg gggagggaaa tagatgggaa aaggtaatta ggctttatag
27721 aagggtctct attagcaagc ttctagggga tgccaagatc catgcttaga gattgccagg
27781 cttgtcttca aatctcagct gtgtattact cctttatgtt tttgtttgt ttgtgttgtt
27841 tgtttttgag acagagtctc gctgtgtcac ccaggctgga gtgtagtggg gtgatctcag
27901 ctcactgcaa actctgcctc ctgggttcaa gcgaatctca gtctcctgag tagctgggac
27961 tacaggcatg caccaccagg cctggctaata tttgttagag acggggtttt gctatgctgg
28021 ccaggctggg cttgaaactc tgacctcaag tgatctgccc gccttggcct cccaaagtgt
28081 tgggattagt ggcgtgagcc actgccccgg cctattactc ctttagagtg atttagagcc
28141 atgtttactt atggtaactt gacagtaatg ggaataacca ctgatgaaac gttaaagcctt
28201 tgtctaattg tttacctagt tcttcttgtt ggttcatgaa atttttctac tctgtacagt
28261 ttgaaaatta agatgataat atttagagat attttattcc tttgtgaaga gaaaaaggc
28321 tttcattaac agaaatcagt ggcaataact taataaatac aatcagctgg tgttctata
28381 gtattttaaa gaaaacagaa agtttactag atttcagcca gttttcagac tatttaatgt
28441 ctattcttac tataatagaa aatatataat ttgatcttgt tctcattttt caaagacctt
28501 taatacatga ttttagtagt tgaaaatgaa gttaaataat agtttatgcc tctactttta
28561 aaaacaaagt ctaacagatt tttctcatgt taaatcacag aaaaagccac ctgacatttt
28621 aacttgtttt tgatttgaca gtgaaatctt ataaatctgc cacagtctta aaccaataaa
28681 gatcaaggta taagggaaaa atgtagaatg tttgtgtgtt tattttttcc accttgttct
28741 aagcacagca atgagcattc gtaaaagcct tactttattt gtccacctt ttcatgttt
28801 tttagaagcc caacactttt cttaacaca tacaatgtgg ccttttcatg aaatcaattc
28861 cctgcacagt gatatatggc agagcattga attctgccaa atatctggct gagtgtttgg
28921 tgttgtatgg tctccatgag attttgtctc tataatactt gggttaatct ccttgगतat
28981 acttgtgtga atcaaaactat gttaagggaa ataggacaac taaaatattt gcacatgcaa
29041 cttattgggc ccacttttta ttcttttgca gagaatggga tagagagctg gcttcaaaga
29101 aaaatcctaa actcattaat gcccttcggc gatgtttttt ctggagattt atgttctatg
29161 gaatcttttt atatttaggg gtaaggatct catttgtaca ttcattatgt atcacataac
29221 tatattcatt tttgtgatta tgaaaagact acgaaatctg gtgaatagggt gtaaaaaatat
29281 aaaggatgaa tccaactcca aacactaaga aaccacctaa aactctagta aggataagta
29341 aaaatccttt ggaactaaaa tgtcctggaa cacgggtggc aatttacaat ctcaatgggc
29401 tcagcaaaat aaattgcttg cttaaaaaat tattttctgt tatgattcca atcacatta
29461 tcttactagt acatgagatt actgggtgcct ttattttgct gtattcaaca ggagagtgtc
29521 aggagacaat gtcagcagaa ttagggtcaaa tgcagctaata tacatatatg aatgtttgta
29581 atattttgaa atcatatctg catgggtgaat tgtttcaaag aaaaacacta aaaatttaaa
29641 gtatagcagc tttaaatact aaataaataa tactaaaaat ttaaagttct cttgcaatat
29701 attttcttaa tatcttacat ctcatcagtg tgaaaagtgt cacatctgaa aatccaggct
29761 ttgtgggtgt taagtgcctt gtatgttccc cagttgctgt ccaatgtgac tctgatttat
29821 tattttctac atcatgaaag cattatttga atccttgggt gtaacctata aaaggagaca
29881 gattcaagac ttgtttaatc ttcttggtta agctgtgcac aataatttgc ttggggcggt
29941 tacttatcat atggattgac ttgtgtttat attggtcttt atgcctcagg gagttaaaca
30001 gtgtctccca gagaaatgcc atttgtgtta cattgcttga aaaatttccag ttcatacacc
30061 cccatgaaaa atacatttaa aacttatctt aacaaagatg agtacactta ggcccagaat
30121 gttctctaata gctcttgata atttctaga agaaattttt ctgacttttg aaataataga
30181 tccataatat atattcttat ggaaatctga aaccatttgg gcatttgggg gtaaaaagta
30241 ttttattagt aaatttaaat gaggtagctg gataattaaa ttacttttaa gttaccttgc
30301 agatgatttt tctcaatcag agcaccaccc agagctttga gaaacaattt tattcacagc
30361 ttctgattct atttgatgta attttttagaa aataagtttt gctgggtgct ttgaatcagg

FIG. 1 (cont.)

10/63

30421 gtatggagta cagttcactc tgatcctatc atataaatca tgtaagtata taacattttc
30481 aataagtgat tgttggattg aagtgaatga tatttcaagt aattgttatg tcatggccaa
30541 gatttcagtg aaactcaaaa tttctcctgg ttgtggtctc cattgcatgc tgcttctatt
30601 gattaaccta agcactactg agtagaagct ggaagagggg tctaattaga aggccctttt
30661 ctatgctctg cttggcttgt aaaataattt atttctctag atcccaccaa catagtagtt
30721 tcatgtatgc aaaaacaccc acctaaatgt caaagtttgt atgatacatg gacatatcta
30781 tagaattttt ttggtctgg tgcatgccaa aaaataaaca tgatatagaa gaatttaata
30841 tttattgagt acctaatctg ttccagttca atatgaaggt ctttatgcag attattttac
30901 ttaatttttc tagtaactcc atggagcaaa aattatctct aatttatata acaggaagtt
30961 gagcgtgagg caaattaagt aactttccca aagttacaca tatggtaagt ttgagagata
31021 tcccagtcct tttagctcca aagcctttga ccttttcacc ataccagatt atgattgcta
31081 ttaatatata attataatta taatgattgt atttaggtac tcaacagaat ggtgactcta
31141 gtaaccagcc ttggttctgc tgagcttctc tgcgtcttct caggagacac aggctacaga
31201 gcttgaaggc tgaggattct tccagggcca cttcaggggc aaatctgaaa ctttcttcag
31261 gacaggaatc aacgagatct tctcacttac ttatacctgg gggaggaact gtatgaaatc
31321 caccaagaa ccagtcatgc taagggccaa acctatagac aaaaaagggt ataggagaat
31381 ggagtatgta tggagaaaga ctaaattgtt cttaaaacttc tcaagcttaa aaatatccca
31441 gcaaaaagaga tcgtaaaagc cttcatgggc gtattaatta tccatgcag ggggtgagtg
31501 gaaaggtact cctgagcccg aggctacagc tttggaaacta gcagcacctt tgaaggggaa
31561 agcgtgtttc catcatctca actcctactg ataaccaatg gaatattggt gagtaaagga
31621 tcctggggga agaagcagct gaaatgtgta ggtgagaagg cagagagaag aatatttata
31681 ttgggaatgg cacaagtgtg atgaggctgc aggtttttca ccttgtcat agagaaaaaa
31741 ccacgctgac accatgcagt tttaaatagt gagaaatttg caaattgtta gatcttaaat
31801 aatttagata aacatagtgg ccatttagat tattgcagtt ttttcaggat atctgatctc
31861 ttgatttcat tctttttgtc tcttataaga ataaaagggg gggagaaaat ttagccatta
31921 tagtatttct ctacattttc tctgtccttt tacataactt acaccagtgc cttcctattt
31981 atggtattat ttatgggtat ttcttctttt ctttcactga gcaaggataa atgagccagg
32041 gattcttgaa actactgtaa cacttctctt agaaatagat ggtcatactt tcagaatctc
32101 tacacattct tagtccctct aaacaatgat agttgtggca taaaaatatt tgcttggttt
32161 caggactgat agagaaaagt actataaaat ttgctgttaa ctgtgaaagg ttaaaaaaaa
32221 ggaggtgcc a tcatgaagga gctaactctt ctgaagtact gctgtagttt taaatattat
32281 tagctatgac ttctcaccat taactatgca cttgcttttt cttcatctga ctcagcagcc
32341 agatagatgc aacattgtct ttaacattta agactcctag caagtccggg cacgggggct
32401 cacacctgta atcccagcac tttgggaggc cgaggtgggc aaatcacaag gtcaggagtt
32461 tgagaccagc ctggccaata tggtgaaacc ctgtctctac taaaagtaca aaaatcagcc
32521 aggtgtggtg gcgtggtggc gggcacctgt ggtcccagct acttgggagg ctgaggcagg
32581 agaatagctt gaaactggga ggcagaggtt gcagtgagct gagatcgcac cactgcactc
32641 cagcctgggt gacagagcga gactccatct caaaaaaaaa aaaaaaaaaa aagactccta
32701 gcatggaaga gaaactggct gttgaaaacc tgaatgtgag agtcagtcaa ggatagtttg
32761 agggaagcca agtagaggaa gctctcacia gcagattggt gagagaatat gattatacaa
32821 tgcatttatt atgataagaa attcacaagc attcattcaa aatactcttg attcctaggc
32881 agctctgggc atatttccac caacaaattg aggcataatg cagtgcagcc taggtcagac
32941 tacctttttt cattaaacct cacaaaatta aaggacatac aggagaagtc ctggtactca
33001 tggtgcagac tacagtctat atggcaaagg aggatctctg tcccttatgt ttggatgaaa
33061 acattgggta ggcatttgaa tacaagccta ctgctaatat ggggctaagg tctttggccc
33121 cctaaagggt tgcgtgaaata ttactgacag gaggcagatt gataagagga aaagcacata
33181 aatgtatttg acatgtatac atgggagcct tcaggatgaa gacctaccct ctcagtgcag
33241 tatggaagct tgtataccat cttgaggtta cagaaagaat gggggttttg atctttgtaa
33301 aacaggtttc agtggcaaga caggttatga gaaggagaaa ggaagagact tgggtagcaa
33361 agggggctct gttttgtagg taaatcgttg gcagcccaca gagaaaatag atggagaatg
33421 tttcttttca gaccttggca ggtgtcagat tctcagttaa tctctcctag atttgaaaaa
33481 aaaaaaaaaa gtctagaaaag ggagagcctg gctgcactaa cacattttct acagatgcaa
33541 atttctccca caaaatacag ctttgcaggt ccacttctat ctgctgggcc tgtggcaacc
33601 atttcaaaat atgtgaatga aatatatgtg ggggtaaact atttttatth acttccttaa
33661 agaagggatg gtgttctctc ggggaattctg tgcatagaga gcctgtggct taggcacttt
33721 gatttatgta tatctcttcc tgtgattggc tatctaggga ctgctatctc cagcaaatct
33781 tctaaatgct tgccatgtag aattcctttc tcatctttct gtctcacccc cttatctago

FIG. 1 (cont.)

33841 tgctttctcta accctagagt gacactgcac tccccacaat ctccctatgtc ctgaatatatt
33901 taccocatcc taaactccat ctctaacaca gatgcacttt cttgtgctgc ctactgcatt
33961 gtacatcttc cccttagttc ccatgatgca actctgccct accccagaaa atgtaattta
34021 attgggtctgg gataaaacct gggacactat cattcttgaa atattcccca agcgattota
34081 attatatagc caaagttgag aactatttgt agacaggcat cagcatgac acttaatgat
34141 ttgacttttg ctagatctaa ggtgaggaaa ttggagagtg gtatccatag gaagaactgt
34201 ttagtttaat ttttttttta tttttcttc taaaaaaaaa tccaacaacg agatacatgt
34261 gcggaacatg caggtttggt acatagggtat aatgtgccat ggtagttgt tgcacctatt
34321 gaccocatcc ctaagttccc tcccctactc cttacttccc aacaggccct ggtgtatggt
34381 gttccctctc ctgggtccac ctgttctcaa tgttcaactc ctttttacga gtgagaacac
34441 atggtgtttg attttctggt cctgtgttaa tttgctgagg atgatagttt ccagcttcat
34501 ccaogtccct gcaaaggaca tgatctcatt cttttttatg gctgcatagt attccatgat
34561 gtatatgtac cacattttct ttatccagtc tgtcattgat gggcatttgg gttggttcca
34621 tgtctttgct attgtaaata gttctgcagt aaacatatat gtccatgtgt ctttatagta
34681 gaatgattta tattactttg ggtatatacc cagtaatgag attgctgggt caaatggcat
34741 ttctggttct agatacttga ggaatcgcca cactgtcttc cacaatgggt gaactaattt
34801 acactcccac taacagtgtg aaagcgttcc tatttctcca cagcctcacc agcatctatt
34861 gtttcctaac attttaataa ctgctattct gactggcatg agatggtagc tcatttgtgt
34921 ttgatttgc atttatctga tgatcagtga tgcctgagatt tttaaaatat gtttgttggc
34981 catgtaaatg tcttttctga agtgtctgtt catatccctt gccacctta atagggtttt
35041 tttttcttg tgaatttggt taagtgcctt gtaaattctg gaaattagat ctttgtcaga
35101 tggatagatt gcaaaaattt tctcccattt tgtaggttgc ctgttctact tgatgatagg
35161 ttcttttgct gtgcagaagc tcttttagttt aattagatcc aatttgtcaa ttttggcttt
35221 ttttgcaatt gcttttggca ttttctctgt gaagtctttg cccgtgccta tgtcctgaat
35281 ggtattgctg aggttttctt ctagggtttt tatagttttg ggttttacct ttaagtcttt
35341 aatacatctt gagttaattt ttgtataagg tataaggaag gggccagtt tcagttttat
35401 gcataatggc taggcagttt tcccaccacc atttactgaa taggagatct tttcctcatt
35461 gcttggtttt gtcagatttg tcgaagatca gatggttgta gatgtgtggt gttatttctg
35521 aggtctctgt tctgcacct tggctatata gtctgttata gtaccagtc catgctgttt
35581 tggttaccgt agccttgtag tatattttga agtctggtag cgtgatgcct ccagctttgt
35641 tctttttgct taggattgtc ttggctatat ggagtcttct ttgattccat atgaaattta
35701 aaataatttt tttttattct gtgaagaatg tcaatggtag ttgatggga atagcattga
35761 aattataaat tactttgggc agtatagcca tgttcacaat attgattctt tctatccgta
35821 aggacgacac tttttccatt tgtttgtgtt ctctcttatt tccctgagca gtggtttgta
35881 gttctcctta aagaggctct tcacatcctt tgttagctgt gttcctaggt attttgttct
35941 ctttgtagtg attgtgaatg ggaattcatt cttgatttgc ctctctgctg cctgttgttg
36001 gtgtaaacaa aattcatttc ttgttcttat ttgtgaaatt ttggaaccaa atctattttc
36061 aaattagaaa ttgcttgtga taatggtttt gcaacttaga ctggatatga gacgatgaga
36121 tattagttct ttcattcctt tgtaggaata tgggtgcact tgcattattt tagctaacta
36181 gtgtccttta atgactaatg aatatgacat ggtgaaacaa agtaaaatat atatgatgca
36241 ctaagtatgc attgtttcca aaggttcagc attttttttt tgttaactct gctgggatct
36301 gctttatgca ctgataacat aacttatttt atgatcttaa gcaaataaaa acacttatct
36361 ggacctcagt ttcccttaact gtacaactga gggaaaactg atagtatagc tatagtacag
36421 tataccatct ttaccgtcac ttccatcttt taaattatgt gtatataaga tagggcctag
36481 ataaatggta tttatcttaa attacagtga tactagctta taacttaatt tgctaggtoa
36541 tgttgaaactg ataacaatgt gtgaactgat gagcaactga gaagtaacca ggttgtgtta
36601 taacagtttg tttttgattt agggttatca gtgaggggtg cgggtggggag gggactttgg
36661 agtctaactg tctagttcaa atattagttt ttgtttattt ttatttttaa tttttgtggg
36721 tacatagtag atgtatatat ttatggggta catgtgatgt ttccatatag gcatgcaatg
36781 tgaaataagc acatcataga gaatggggta tccatccctc caaacactta tcttttgagt
36841 taccaacaat ccaatgacac tctttaagtt atcaaatcac agttttgcca gctactagcc
36901 atgtgatttt gggtagggtta cttaaattct cttcatctca atttcattat tgtaaagtgg
36961 agataatgat agcacatttt ttctttttct tttttctttt attttttatt attatacttt
37021 aagttgtgtg atacatgtgc agaatgtgca ggtttgttac ataggatatca acaactctat
37081 aaaacatggt ctatccagga aaagaaacta tcatcagagt gaacaggcaa cttacggaat
37141 gggagaaaat gtttgcaatc tagatggcga ttgcaatggc ggttcgtctg atccatcagc
37201 ccatcatcta cattagggtat ttctccta at gctatccctc ccttgctcc ccacccctc

FIG. 1 (cont.)

37261 acaggccccc gtgtgtgatg ttcccctccc tgtgtccatg tgttctcatt gttcaactcc
37321 cacttatgag tgagaacatg tgggtgttgg ttttctgttc ttgtgttagt ttgctgagaa
37381 tgatggtttc cagcttcac ccatgttcctg caaggacatg aactcatcct tttttatggc
37441 tgtatagtat tccatgggat atatgtgcca cattttcttt atccagtcta tcattgggtg
37501 acatttgggt tggttccaag tctttgctat tgtgaacgct gcagcaatga acatacataa
37561 gcatatgtct ttctagtcaa ataagttata atcctttggg tatgtacca gtaatgggat
37621 tgctgggtca aatgggtatt ctgggtctag attcttgagg aatcgccaca ctgtcttcca
37681 caatggttga attaatatt acctccacca acagtgtaga agcattccta tttctccaca
37741 tccgctccag catctgttgt ttctgactt tttaatgat accattctaa ctggtgtgag
37801 atgggtatct attgtgggtt tgatttgcac ttctctaatt actagtgat atgagcttct
37861 tttcatgttt gttggctgca taaatgtctt cttttgagaa gtgtctgttc atatccttcc
37921 cccacttttt gatgggggtt ttttttccct gtaaatttgt ttaagttcct tgtagatttt
37981 ggatattagc cctttgtcag gtggatagat tgcaaacatt ttctcccatt ctgtaagttg
38041 cctgttcact ctgatgatag tttcttttgc tggatagaac atgttttata gaggttgtgt
38101 gagaattaaa tgcattaagc acatagaata gattctggta catagcaagt gctctctcta
38161 tatatggaac tctatatgta gttggtgcaa aagtaattgt ggttttcacc attgaaagta
38221 atggcaaaga ccatcattac cttttcacca atttaaatat atggaaggaa tatatatata
38281 aaacctatat atatatgtca catatatgtc tctaaccocat tattataata tataatacaa
38341 tatatattat aattataaatt gtatataaca tatgttatat aataatatag taatatttat
38401 tctaaataaa tatataatac tataaataat ataataattt atatatatga ttataatata
38461 taataggcta tattatatat tattaacata tacatatgtg tatatatatg tctttcatag
38521 acttaaatat atagagcaat aataggttag aaaatagcaa acatgtatat ataaacatat
38581 atacatatag aaaacatata taaaaacata tatatatata tatatatgtg tgttttctgc
38641 ctttcatttt tagagacagg gtctcatcat gttgcccagg ctggtctcaa actcctgggc
38701 tcaagtgatc ctactgcttt ggactcccga agtgctggga ttccagacat gagacactgc
38761 acccagtcga gtccctgtct ttttaaatag actctctacc taagtgcaca aatactcatt
38821 atttacattt agttatttct gtatatatgc tataagcaaa tcttgtagca ccagtttgat
38881 ttttataagg cacaagaata tattttacta atgctttaaa atggcagcta gattctagta
38941 ttactttaga aattaaaatt aatattttaa cacatctttc attattgtgt tatctgaacc
39001 aaacctatta ttgctgctat ttcagcaaat ccaggggctt tttcttataa aatatgaaga
39061 atatagctta gatttctagt gaagatgtta ccagtaataa ttaataaaat cagtaagcac
39121 taaaaggaaa atacccaaaac taaagcattt tgaattagtc attgaatcta aaagaaaggt
39181 agattttttt ctgagattct gttctagggtg tggatatatg gtatttttgc aaaaactata
39241 aacaattgtg gcaaaatgaa ggaaatattt aaaaacaaac ctcttaattc ttcagtggat
39301 taagcgtgaa tatgttttta ttttctatga tgaatatgga aaaattcatt tccttagcaa
39361 tttgtatgag cccaaaaact attgtcagac tctgctgtat caaaatagac aaaaaattga
39421 cactcacttt taacctgcca aaagcaaaat cttaaacttt tgcttttagta tataagccag
39481 cattcattgt atcctatgat gggttctgag tgtagggtga tttgctttct tccatttttt
39541 gtatgcatgt tttcttttta tttattattg taagttgtat gaaattttta tccaaatttt
39601 tattttcttc tgattaataa tcagaataat cagataatta ctggtaaatt tgatgttaat
39661 ccttccagct ttttcccatg ggaatttata cttaataaag gggagaagtc atcattacat
39721 aatgtgcata ttaatotgct tctcccttta atgtgttgtg aatgcctttc catgtcatta
39781 gatgtttttc taacctagtta ctttcatgaa tcatatggct gtacctgat ttatttaatc
39841 agttcctcat cattgagtat gttaaattgc tccatttttt tattactata aaaggctcct
39901 cagtacacac ccctttaaaa gctgactctt agaagggtgt cttgactctc tacctaagtg
39961 taaaaataca aataaattgc tttccagaaa aggtgcacta ctattttact ttctgatac
40021 taaactatga aaattcagtc ctaacaatag atatttaaat aaagttttta aaatgccag
40081 tgaaaaagag catattatta ttttcatttg cattactttt ggttcctggg gaggtttaatc
40141 tgtttttgta tattaattat gcatttatat ttctttttgt gtgtgtgaat tgccttccat
40201 gttctttgtg tgtttttatt ttgttgtatt tgtctcttcc ttgatatatg agagaatatt
40261 ttccctagcc tgtcaattgc cttgtaattt tgtttctagt gaggtttttt tttttttttt
40321 acaattaaaa gctttaattt ttgaaaattt tgctggcaaa tctatatatc tttttctttg
40381 tttctgctt tgacattatt cttttataaa ggcccatgcc acccaaatat tatgtaagca
40441 tgcactatg tttttattac ttcacttttt acatttaaat atctactcta tttagaattc
40501 attgtgatgc atgtatgagg tagaaatota atttcaaaaa gatgagtatc cagtttgtcc
40561 atcatttatt gcatgatctc tttctccact gaattaaaaa gcogtatttt ataatatatt

FIG. 1 (cont.)

40621 aaagtattac atgtgcttgg acatgttcoet ggactttttga gataaatcag totattttcctt
40681 tgtcatgtca catattatta tggettttatg atttaatatc cagtaatgta aaccctctga
40741 cacattattc ttattcctca aatgtttttg atgagttttc ttccaaatga aattttataat
40801 catttttattc attgattcaa caaatatttg ttgaatggat attctgtgct tgggtattgtg
40861 catgggtatta ggattgttgc aaaaattgag actgacagtc cctactctta cgggtgctaaa
40921 aattcacttc caaaaaaatc tttaaatggt gatgaagatt gcactaatct tataaaaataa
40981 cttggagggg aatgtaatct ttgcaacatt aagttcttca ttttagaaag ttttaagact
41041 ctccatttat ttgagacttt taaaatatgt cccaataatg ttttgtgaga tgtatatattt
41101 aagatatata tcttattgct attacattgt atcttttggt atattgttac tatgaatggg
41161 ataactattt aattagatgt catttttggt atatagaaat ctattttcctt agcatagtca
41221 ttttttaaac ctcgatctat taaattcttg attcatttac atttgttaca caatcatatt
41281 ctatgctgat aatacttctt gottctttcc aatatttgta cctcgatcat ttttcttggt
41341 gagttgtatt agctagaagt tctagaaaaa tgttaaatgg tagtaatagc tagtattctg
41401 ttttttctg actctaaatg taatgcactc agacttttat aattatggca ttgattgtaa
41461 cattttgagg aagaaatcct ttttcagggt aataatgtat ctttatattc aagtttatta
41521 agaacattta ttggaaacat attgaaattt tatcagattc cttttcagtt gttactgaga
41581 taatcatagg ttcttctgta ttcttttaat taatttctca aaattaaact gtcctattat
41641 tcttggaata acgacatata aagtactgta tatttaaaag aagttaaaat gataatgggtg
41701 attttattaa gtgacctcac acaatagaaa acagtgtagc cttagaagtt ttccaaagtga
41761 ccattctact tagaaacaac cctgctttgg gatcagaact gtaattttta aagtaaagtt
41821 ttctgggttt aattcattta gtgtaattac aagcatgagt tcaggtttct atttttttca
41881 cctgaacttt ccttcatggg ttgaatatct agaaaaagca gactttccta tctctagact
41941 aaacatttga tcctatctta ggtatgcatt acaatttttt aaccataaat ggttaaagaa
42001 tttagactca tctacaataa ctttgaagct ctggctctga agaacatgtg agaaatgaga
42061 tataactcct agaagatata ggagacattt ttagtcttcc aaattttccc tgggaggctg
42121 atctaaattg agtcacaaaa ttgttcccac caggaatgca atcacttgag ctgttttcta
42181 atctgagccc ctctaccag atgatcttct gaactcatac tgttcagact ttcactcttc
42241 tgagtagaaa acagccatag tcatggcagg atgagggcta ggacaattac ccaaggaatt
42301 cttggcctct gccatgggac tctgcagact cagatcatat aatcagagat gtttagcactg
42361 gaggggacat cacaattagc tttctccacc tcttagttta tcagtgagga aaactgtcca
42421 gagcgcgga gagactaaaa taacacagcc aatgtaggta atgtgctgga taagaatttg
42481 gaattcacga ttttgaattc agtggtttatt tcaccatcac gctggccttac acgttggtat
42541 caggcttctt ctattattga agtgagccat taagtgaatt ccacttgat ttgtgtctga
42601 tacagagtaa taaactattt tattaaatat ccaaataatt atacattcct ccttcttaca
42661 tgcaagccta agtttgottg tactatttca tgtggtagca aatcaggaog cttcttgtgt
42721 ctctgaaaaat actctgagta atggagtaca gtcagctttc ttgtaccaag aatatagggga
42781 ctatgtttct cccagtcatt ctggggataa tttttgtgaa ggattgcact tcataggtta
42841 agctaggtat cagttaccag tgttttttcc aaataaaaaa aaaatcaggt gatatctgta
42901 aatggttcca ttgtaaatat taaagaacat gatgcttaaa acagattagg gaaaactata
42961 gaaggggtgg ggtttcggag tgctaatttt gtoccttgaat ggtaacagct ccatgtgggtg
43021 gtgaggttta tgttggtttg ctgtttgag atgatottat tattagaatt tttcataccg
43081 aaaataaact gcatttttagt ttgtaaacat gcccttccag agtaatgcta ccagttcttt
43141 gtgaaatagc tactgttggt caaaggatga ctatgtcctc ttcggttgag gaaagatgac
43201 aacaaaactca gtaatgacat gtaaaatagg tattacaaac caggtatggg ggcatgagcc
43261 tgtaatccca gctacttgag aggctaaagc aggaggatct gttgatctat ggatttgagg
43321 ctgtagtgtg ttgtgatggc acctatgaat agcccttgca ctccagccca agcaacaaag
43381 caagactgtc tctgaatttt tgttttggtt tgttttttgt tttttttttt ttgagacaga
43441 gtcttgctct gtcacccagg ctgaagtgca gtggcgcgat ctccactcac tgcaagctcc
43501 gcctcctggg ttcaogccat tctcctgcct cagcctcccg agtagctagg actacaggcg
43561 cccgcctcca cgcccagcta aattttttgt attttttagta gagacgaggt ttcactgtgt
43621 tagccaggac ggtcttgatc tctgacctt gtgatcctcc tgccctgggc tctgaatttt ttaaaaaggc
43681 ctgggattac aggcgtgagc caccgcgcc ggccctgtc tctgaatttt gttaaattaca actgaataaa
43741 attccactca aattaatata cattttaatt gtgttttggt gtaaattaca actgaataaa
43801 aattcagcaa ataagtotgt tgtggtaggg aaaagtctat tgtgatctgg aaaatataat
43861 ggagaaatcc agtggaagag attttatttc acattactca aaataaaaaa atcttataca
43921 agtctttaca cttgtaactt gaaaaattct gtgctaaaat ttagcttggt tgctaaaata
43981 tttctctttt tttctcagaa gcttcttttt agcatcctat agacacaagt tactttttta

FIG. 1 (cont.)

44041 aatattttgca taacttgcttt gcaatgtatt gtttatcagt agttotatat tctttgagat
44101 agtctatcca gtctttctgt atttatcgta tgtctgtata gatatatatt agcagataaa
44161 tgagttctga aaggggagaa atgtgattat gctaatacatg atataaagaa ttgactttat
44221 aagcagtgtt cacaggtcac acctttcccg ttactgtctt acagtgaaca agaaatgatg
44281 ctttgtctgg tatgcatggc aaataatgcc ccttgctctc tgcttcatga tcacatgtga
44341 tacttctaac atagatagca catgtaaatc cagtggcctt gactgcaact caagagagca
44401 ttttggccaa gtacaaaacc actagtcacg aaaaaaaaaa aaaaaccaa tcaaagtaaa
44461 ttgatgggat tgacatttgt ctatgaaaaa caacataata tagaacaatt ctggggtaaa
44521 atattgatct aaaataattt taaggattaa atattgccat tgtaagcata ctatgagcaa
44581 ttatgtttgt aatgcagata tatttataat tttaaatcca agatttacct taattgtaca
44641 ttttcctaata ttaaaaaagt tttttgaaa aaaaaatcct cgaatctaga gaaaggttgg
44701 caaacataa tggaactttg taaaaaacat ccagggcagc actttcactg attgcagtag
44761 cttaggagtg aaaaacaaca caactgctcc aatgtatggc aatgggcaaa tatcccgatt
44821 tattcacagg gtggcatggt aggcagtgtc tagaataaat gagttgggta tacaagtatc
44881 aatagggata aatgtgaaaa acacagtgtt aagtttttaa aaagttgtaa aaagcacagt
44941 aggatgttat ttatataaaa tttaaaaacc tcaaaaacca ttcttctttg atatatattc
45001 taaagatgaa catatatgta atagaagtac aaaacataca taaaataata tacactatgc
45061 agtcatttgt gtaacttactt ttcaaaaata tttcagtaga tatagcaaac agttaacatg
45121 taatatttgg ataggagggt ggcaattttc tttttagcac ctgcctgtct gctatcattc
45181 aaactcacat ttaaaatgtg gctatgtgag atgagagAAC tataatattc caggtttgtg
45241 attagtttgg aaacttttta aaagtttgaa tgtggctctga gagatagttt gttataattt
45301 ctgttctttt acatttgctg aggagagctt tacttccaac tatgtgggtc attttggaat
45361 aggtgtgggtg tgggtgctgaa aaaaatgtat attctgttga tttgggggtg agagttctgt
45421 agatgtctat taggtctgct tgggtgcagag ctgagttcaa ttctgggta tcttgttga
45481 ctttctgtct cgttgatctg tctaattgtt acagtggggt gttaaagtct cccattatta
45541 atgtgtggga gtctaagtct ctttgtaggc cactcaggac ttgctttatg aatctgggtg
45601 ctctgtattt ggggtgcataa atatttagga tagttagctc ctctgttga attgatccct
45661 ttaccattat gtaatggcct tctttgtctc ttttgatctt tgttggttta aagtctgttt
45721 tatcagagac taggattgca acccctgctt tttttgttt tccattggct tggtagatct
45781 tctcccatcc ttttattttg agcctatgtg tgtctctgca cgtgagatgg gtttctgaa
45841 tacagcacac tgatgggtct tgactcttta tccaatttgc cagtctgtgt cttttaattg
45901 gagcatttag tccatttata tttaaagtta atattgttat gtgtgaattt gatcctgtca
45961 ttatgatgtt agctgggtgat tttgctcatt agttgatgca gtttcttctt agtctcgatg
46021 gtctttacat tttggcatga ttttgcagtg gctgggtactg gttgttctct tccaggttta
46081 gcgcttcctt caggagctct tttagggcag gcctgggtgg gacaaaatct ctcagcattt
46141 gcttgtctat aaagtatttt atttctcctt cacttatgaa gcttagtttg gctggatata
46201 tctcagacca cagtgcacac aaactagaac tcaggattaa gaatctcact caaagccgct
46261 caactacatg gaaactgaac aacctgctcc tgaatgacta ctgggtacat aacgaaatga
46321 agacagaaat aaagatgttc tttgaaacca acgagaacaa agacaccaca taccagaatc
46381 tctgggatgc attcaaagca gtgtgttagag ggaaatttat agcactaaat gcctacaaga
46441 gaaagcagga aagatccaaa attgacaccc taacatcaca attaaaagaa ctagaaaagc
46501 aagagcaaac acattcaaaa gctagcagaa ggcaagaaat aactaaaatc agagcagaac
46561 tgaaggaaat agagacacaa aaaacccttc aaaaaatcaa tgaatccagg agctgggttt
46621 ttgaaaggat caacaaaatt gatagacogc tagcaagact aataaagaaa aaaagagaga
46681 agaatacaat agacacaata aaaaatgata aaggggatat caccaccaat cccacagaaa
46741 tacaaactac catcagagaa tactacaaac acctctacgc aaataaacta gaaaatctag
46801 aagaaatgga tacattctct gacacataca ctctcccaag actaaaccag gaagaagttg
46861 aatctctgaa tagaccaata acaggctctg aaattgtggc aataatcaat agtttaccac
46921 ccaaaaagag tccaggacca gatggattca cagccgaatt ctaccagagg tacaaggagg
46981 aactggtaac attccttctg aaactattcc aatcaataga aaaagaggga atcctcccta
47041 actcatttta tgaggccagc atcattctga taccaaagcc gggcagagac acaacccaaa
47101 aagagaattt tagaccaata tctttagatg acattgatgc aaaaatcctc aataaaatac
47161 tggcaaacog aatccagcag cacatcaaaa agcttatcca ccatgatcaa gtgggcttca
47221 tccctgggat gcaaggctgg tccaatatac gcaaatcaat aaatgtaatc cagcatataa
47281 acagagccaa agacaaaaac cacatgatta tctcaataga tgcagaaaaa gcctttgaca
47341 aaattcaaca acccttcacg ctaaaaaact tcaataaatt aggtattgat gggacgtatt
47401 tcaaaataat aagagctatc tatgacaaac ccacagccaa tatcatactg aatgggcaaa

FIG. 1 (cont.)

15/63

47461 aactggaagc attccctttg aaaactggca caagacaggg atgccctctc tcaccgctcc
47521 tattcaacat agtgttggaa gttctggcca gggcaatcag gcaggagaag gaaataaagg
47581 gtattcaatt aggaaaagag gaagtcaaat tgtccctgtt tgcagacgac atgattgttt
47641 atctagaaaa ccccatcgtc tcagcccaaa atctccttaa gctgataagc aacttcagca
47701 aagtctcagg atacaaaatc aatgtacaaa aatcacaagc attcttatac accaacaaca
47761 gacaaacaga gagccaaatc atgagtgaac tcccattcac aattgottca aagagaataa
47821 aatacctagg aatccaactt acaagggatg tgaaggacct cttcaaggag aactacaaac
47881 cactgctcaa ggaaataaaa gaggacacaa acaaatggaa gaacattcca tgctcatggg
47941 taggaagaat caatatcgtyg aaaatggcca tactgcccac ggtaattttac agattcaatg
48001 ccatcccat caagctacca atgactttct tcatagaatt ggaaaaaact actttaaagt
48061 tcatatggaa ccaaaaaaga gcccgcatcg ccaagtcaat cgtaagccaa aagaacaaag
48121 ctggaggcat cagctacct gacttcaaac tatactacaa ggctacagta accaaaacag
48181 catggtactg gtacccaaac agagatatag atcaatggaa cagaacagag ccctcagaaa
48241 taacgccgca tatctacaac tatctgatct ttgacaaacc tgagaaaaac aagcaatggg
48301 gaaaggattc cctattttaat aaatggtgct gggaaaactg gctagccata tgtagaaagc
48361 tgaaactgga tcccttcctt acaccttata caaaaatcaa ttcaagatgg attaaagatt
48421 taaacgttag acctaaaacc ataaaaaccc tagaagaaaa cctaggtatt accattcagg
48481 acataggcgt gggcaaggac ttcattgtcca aaacaccaa agcaatggca acaaaagcca
48541 aaattgacaa atgggatcta attaaactaa agagcttctg caaagcaaaa gaaactacca
48601 tcagagtga caggcaacct acaacatggg agaaaathtt cgcaacctac tcatctgaca
48661 aagggtctaat atccagaatc tacaatgaac tcaaacaaat ttacaagaaa aaaacaaaca
48721 accccatcaa aaagtgggcyg aaggacatga acagacacta ctcaaaagaa gacatttatg
48781 cagccaaaaa acacatgaag aaatgctcat catcactggc catcagagaa atgcaaatca
48841 aaaccactat gagatatcat ctacacaccag ttagaatggc aatcattaaa aagtcaggaa
48901 acaacaggtg ctggagagga tgtggagaaa taggaacact tttaactgtg tgggtgggact
48961 gtaaaactagt tcaaaccattg tggaaagtca tgtggcgatt cctcagggat ctagaactag
49021 aaataccatt tgacccagcc atcccattac tgggtatata cccaaaggac tataaatcat
49081 gctgctataa agacacatgc acacgtatgt ttattgcggc actattcaca ataggaaaga
49141 cttggaacca acccaaatgt ccaacaatga tagactggat taagaaaatg tggcacatat
49201 acaccatgga atactataca gccataaaaa atgatgagtt catgtccttt gtagagacat
49261 ggatgaaatt ggaaaccatc attctcagta aactatcgca agaacaaaaa accaaacacc
49321 gcatattctc actcataggt ggggaattgaa caatgagatc acatggacac aggaagggga
49381 atatcacact ctggggactg tggtggggct gggggagggg ggagggatag cattgggaga
49441 tataccta at gctagatgac acgttagtggt gtgcagcgca ccagcatggc acatgtatac
49501 atatgtaact aacctgcaca atgtgcacat gtaccctaaa acttagagta taaaaaaaaa
49561 aaaaaaaaaa gtttgaatgt tttcttgcat tcagagcctt gggtgacata gttaattaaa
49621 aataaaacat tgtatataaa gcacagaatg agcagctaca caaagctgct caatcaatga
49681 cagctctata tgggttaggg tttcttggtg ggatgacatt gatgtagaaa gcatggctcat
49741 ctattgagaa tgatggggct ggaggtattg gatacttgag gtttagaaaa tacattgtag
49801 aaaatggaca aaaacccctc aaattaaggg atgaggcaga ataatgcttg gcaataccag
49861 gggtaggctg cagtctttct tggaaatata tattttaaat ggaaccaatt atcatagcat
49921 catttcctct cagggttacc ctctgatccc tattttacta aatcgttata aaacaaaatg
49981 aggaattatg tgtccttccc ttttgaagcc aatgtaacaa gatgggtaag aattagacct
50041 cctgagttca aaatccctgg attcagatct attcctgtat attcaggaga agtggttaata
50101 aattcgatgg acaatttggc ttagttagtc attgaggacc ctgatgaggt atatttggga
50161 aaacataact tccgctctct ctcatagact cacgggcctt tgaggagtcc aggagtcat
50221 ggaatctggc ctgagggttg ggctgctggc aaaactcctt ccccaaagtc cattcctatt
50281 gctgactgag aagggactag cattggaagt ggctgatttt aaataccgct agtgctgggtg
50341 tgctcctccc tcccattccc agctctgctt tgtgtagttg ccttgagaag ctaagttcat
50401 tctgaaaata atgccattgc acaaaacact tttgaaagtt ctagtittgaa attacatcag
50461 gtcacttggc ctgtgtggcc tcagtttctt catctgccat gtgaaaataa taatgcctac
50521 tctgtagcaa agaaagtctc tatagtaaac aaaaaaaaaa cctactctga tactgaaagt
50581 tgttatgaaa aataaaaaag ggaaatgctt tagaaactgt taagtgctat gtagatgtta
50641 ctaattaaca aaccatttca gaaactatac tttttathtt atggccacta ttcactgttt
50701 aacttaaaat acctcatatg taaacttgct tcccactgtt gctataacaa atcccaagtc
50761 ttatttcaaa gtaccaagat attgaaaata gtgctaagag ttccacatat ggtatgacco
50821 tctatataaa ctcattttaa gtctcctcta aagatgaaaa gtcttggtgtt gaaattctca

FIG. 1 (cont.)

50881 ggggtattttta tgagaaataa atgaaattta atttctctgt ttttccccctt ttgtaggaag
50941 tcaccaaagc agtacagcct ctcttactgg gaagaatcat agcttcctat gacccggata
51001 acaaggagga acgctctatc gcgattttatc taggcatagg cttatgcctt ctctttattg
51061 tgaggacact gctcctacac ccagccattt ttggccttca tcacattgga atgcagatga
51121 gaatagctat gtttagtttg atttataaga aggtataact tccttgcaca ggccccatgg
51181 cacatatatt ctgtatcgta catgttttaa tgcataaat taggtagtga gctggtacaa
51241 gtaagggata aatgctgaaa ttaatttaat atgcctatta aataaatggc aggaataatt
51301 aatgctctta attatccttg ataatttaat tgacttaaac tgataattat tgagtatctt
51361 ctgtaaactg cctctgttgt agtttttttt ttctcctaact catgttatca tttttttgga
51421 atccatgggt tcctgttaag atgactcaca cagcctacat aaaagtaatt gacaaaatat
51481 catcttatag taaaatgcc aatatcttta tgttcagcaa gaagagtata atatatgatt
51541 gttaatgata acccaaacaa caaaagattt caccttaact ggttgtcata agtagtagta
51601 tccaccgcct tattttgagt tggattttta tcatcctatg agcctacaa atttaaagtt
51661 tttggaacag cacgtgcatt gaaccataa gaacctactc tgcttttctg catgtattgt
51721 ccagacaaga gaccaaattg ccgaggcatc atttaggtga attctaatta acatttagct
51781 accctacaac cacaattcaa ggttgtttca aaggcatgtg cttgcatcat cctgattcac
51841 taccatgtgt tactaacttg gatctgcaaa gtcattataa aaagctgttt tgatggactt
51901 atttgatat tgctttaccc ttctctctct tttcttttta tcaatgtaa aacattatat
51961 gttaaatact tggcttttaa gagcatagat ctgaaatctg cctctagcaa ataaccata
52021 acacttctaa gatatacctg caaggccaat tgtgttgtaa aaccttgata accatacttt
52081 attgttcaaa aaagcctttt atgaaggcag aagttaaaaa aaaaaaaca aaaaaacaga
52141 gtccacagtt atcacctcag ctacaatctc atcagttcac aagtaccagc aaacatgtg
52201 ataagtcaac aaatgtttta tttcaatctg aacattttac gtaagtgaag actttgttag
52261 atatcatttg gaatgtggaa tctacacagt tggcatatca gagaaggttg aattcagttt
52321 aataaatgtt tatagaaagt gcttgttatc ataataata tagctcagga tgtgcatgac
52381 aagcttttaa gcgattgggt acactatctc atttgatctt ctgcacaact attaatggta
52441 ggtactatta tccctatctt atggataagt aaactaagat ttaaaaagta cagaacatgg
52501 tgtgaacact gcttcaaaaat ttctaaaata ggtaaatcac gatctctaaa ctggaggggt
52561 gtccaaccac tagggacaat agagtaactga tatttagtgg tcagactgta atgcgggaag
52621 agacaggcat gggctaaaacg ggtgtagaga tcaataaagg ggcagggttag tttgtaaaaca
52681 tgtccatatg taacatttag cacaaataca ggatataagg gctttcagac ccagctgcat
52741 tgataaaaag ttaggtggta ttgtatctgt ctctcttctt caatgttgca tatctgtgtt
52801 cttgccaggt ttgcttcac cctctagcca cacttattgg cctacaatgg catcatcacc
52861 aaagaaggca atcccatctc cgtgtggcct tggtttgctc cctaaagtaa acctgtgtgt
52921 tacttttccc aggtctcatg ctttcccata tctgacctgt tttgtcctca tggccaggat
52981 atgtgggacc tttcctacaa tgttccaaag tttgtaatag agctcttctc tgctttgttc
53041 caaattctgc aacattttac tttaaataat gaatttaaat acaaacaaac ttgagctttg
53101 cctatacttt tcaagaatgc agagataact aaattaataa aaatattcat tgagtcctta
53161 ctgtgcacac agctctatgt taagccttgt gcagaactca aagtcactcg agattaagcc
53221 tgttactaag ttatgtgcaa tttagctcag tggatttccc ccacttcata ttgctctgat
53281 aatgtttttg aattaactgc cttgattcct tcttttctct gcttgtctat acactattta
53341 ttattctaca ccactcctaaa ttctaactcc tcaagaaaat ccttcocagat gatttttcta
53401 accaggagtt ttaacttctt ttttaactacc ctattacttt ctacttcctt aactcatcta
53461 tcatattata tttagttatt tatatactag gtcgccttga agaagggatt gtgttttcat
53521 aaatcttaat aatccctgag gcatcaagta cagtgatttg catttactaa atgctcaaca
53581 aatatgtgag ggattcactt gaaactaata ttagataatt ccagtcaaa gtgatctaata
53641 agcaaataca ttcttcagtt ttataggcaa agtatgactc tggttttcca taatcataat
53701 taatttgtca actttataat ttttaattaag taaatttaat tggtagataa ataagtagat
53761 aaaaaataat ttacctgctt aactacgttt catatagcat tgcatttttc tttgtaaaat
53821 ttaagaattt tgtattaata aactttttta caaaagtatt aattattcag ttattcatca
53881 tatactttta ttgacttaaa agtaatttta ttcaaaagag ttagtatagg actacatgaa
53941 aaattcaagg ccaaggctta atttcaaaat tcaactgcct tggctctatc ttttaaaaca
54001 aaacaaaaaa ctcccgcaca atatcaatgg gtatttaagt ataatatcat tctcattgtg
54061 aggagaaaaa ataattattt ctgcctagat gctgggaaat aaaacaacta gaagcatgcc
54121 agtataatat tgactgttga aagaaacatt tatgaacctg agaagatagt aagctagatg
54181 aatagaatat aattttcatt acctttactt aataatgaat gcataataac tgaattagtc
54241 atattataat tttaactata atatatttgt attttgtttg ttgaaattat ctaactttcc

FIG. 1 (cont.)

54301 attttttcttt tagacttttaa agctgtcaag ccgtgttcta gataaaataa gtattggaca
54361 aottgttagt ctcctttcca acaacctgaa caaatttgat gaagtatgta cctattgatt
54421 taatctttta ggcactattg ttataaatta tacaactgga aaggcggagt tttcctgggt
54481 cagataatag taattagtgg ttaagtcttg ctcagctcta gcttccctat tctggaaact
54541 aagaaaggtc aattgtatag cagagcacca ttctggggtc tggtagaacc acccaactca
54601 aaggcacctt agcctgttgt taataagatt tttcaaaact taattcttat cagaccttgc
54661 ttcttttttaa aactttaaat ctgttatgta ctttggccag atatgatacc tgagcaattc
54721 ttgttctggg ttgtcttatg tgaaaaataa attcaaggtc cttgggacag ataatgtgtt
54781 ttatttatct ttgcatatcc attacttaaa acagcattgg acccacagct ggtacaaaat
54841 taattactgt tgaattgagc aaatatattat tctaaatgtc tctgtcaaat gacagagtgt
54901 ggttgtgtgg attaagtccc tggagagagt tctttgttct ctcagtgtct atgctgtggt
54961 tcttgcttta tgcaaaaaga agtaagttac ttaaaacctg gacatgatac ttaagatgtc
55021 caatcttgat tccactgaat aaaaatatgc ttaaaaatgc actgacttga aatttgtttt
55081 ttgggaaaac cgattctatg tgtagaatgt ttaagcacat tgctatgtgc tccatgtaat
55141 gattacctag attttagtgt gctcagaacc acgaagtgtt tgatcatata agctcctttt
55201 aottgctttc tttcatatat gattgttagt ttctaggggt ggaagataca atgacacctg
55261 tttttgctgt gcttttattt tccagggact tgcattggca catttcgtgt ggatcgctcc
55321 tttgcaagtg gcaactcctca tggggctaatt ctgggagttg ttacaggcgt ctgccttctg
55381 tggacttggg ttcttgatag tcttgccct ttttcaggct gggctagggg gaatgatgat
55441 gaagtacagg tagcaaccta ttttcataac ttgaaagttt taaaaattat gttttcaaaa
55501 agcccacttt agtaaaaacca ggactgctct atgcatagaa cagtgatctt cagtgtcatt
55561 aaattttttt tttttttttt tttttgagac agagtctaga tctgtcacco aggctggagt
55621 gcagtggcac gatcttggct cactgcactg caacttctgc ctcccaggct caagcaattc
55681 tcctgcctca gcctccggag tagctgggat tagaggcgca tgccaccaca ccagctaatt
55741 ttttgtattt tagtagagac agggtttcac cagggtgccc aggctgggtct cgaatgcctg
55801 acctcaggtg atccgccac ctccgcctcc caaagtaact atattacagg catgagctac
55861 cgcgccgggc ctaaaaaata ctttttaaga tgggtgtaaat attactttct gtatcaatgg
55921 tacatttttt acttgtcagt ctctagaatt tctttataaa tatgttgatt cagttcattt
55981 ttgtagatta taaaacaggt aaaaaggat aaacatttta tgtgaattaa agggaatacc
56041 taatttttgt gtagagttta ttagctttta ctactctggt ttatggatca tcacaccaga
56101 gccttagtta ctttgtgtta cagaataact aatatgagt aatgaatgac ttacacaagt
56161 cactgcttag gataaagggc ttgagtttgt cagctagagt atgacagaaa gtatctaagt
56221 tttggagtca aatagcactt tgtttgaatc ccagattgca tgcttactag ttatgtgacc
56281 ttagtcaagc cacttcacct cactgagtct ttgctttttt catctctaaa atagagatac
56341 ccaccgctca taggctgtca taagggatag agatagcata tggaatgagt ctgtacagcg
56401 tctggcacat aggaggcatt taccaaacag tagttattat ttttgttacc atctatttga
56461 taataaaaata atgcccactt gttgaataaa agaaatatga cttaaaacct tgagcagttc
56521 ttaatagata atttgacttg tttttactat tagattgatt gattgattga ttgattgatt
56581 tacagagatc agagagctgg gaagatcagt gaaagacttg tgattacctc agaaatgatt
56641 gaaaatatcc aatctgttaa ggcatactgc tgggaagaag caatggaaaa aatgattgaa
56701 aacttaagac agtaagttgt tccaataatt tcaatattgt tagtaattct gtccttaatt
56761 ttttaaaaat atgtttatca tggtagactt ccacctcata tttgatgttt gtgacaatca
56821 aatgattgca ttttaagttct gtcaatatcc atgcattagt tgcacaaaatt cactttcatg
56881 ggctgtagtt ttatgtagtt ggtccagggt gttatttttat gctgcaagta tattatactg
56941 atacgttatt aaagaatttc ctacatatgt tcaactgctgc tcaatacatt tatttcgtta
57001 aaacaattat caagatactg aaggctgatt ggtaactcac atgggaactgg gagagtatac
57061 aattctgaac caaatagatg attctctatt attatatctt aatttatgtg ttatggtata
57121 ttaaacadga aaaaaattgt atttggttag aatatgtttg ctcttcootta actcgggaat
57181 gacatagggg aatattcaca gattgggttc ctataaatcc tccacttgaa gtgaagtcag
57241 ttcaagtaat gaaagctacc tcctgagata gaatcagtac ttggcaccta tctctagtgt
57301 tctttcacct catataacct ttcactgatt agtaaagatt atatccaaca aagaaagtac
57361 agcacagact gagatatgat tactgagata aatttgggca aaatataaac tacagcattt
57421 ctgtagcaat gagaccattt ttcttcagtt gagctccatg ttctacaaac ttcaatcaaa
57481 aaaggttcta ggagactcag tgaaagttga tacactgttc aaggaacaaa taatttcagc
57541 acatgggaat ttcacaggga aaaatatact aaaaagagag gtaccatttt ggatggtgtc
57601 aatatgggtt atgaggaatt caggctgctg agtccagtgt acaatggaaa ctgagctgca
57661 ggtgtgtgat tgtaacaaca aaagaaatgc tgaaatatta agtcctttgc catgtaaata

FIG. 1 (cont.)

57721 gaaaaagagt atttattttcc caaacattat tgctcacctg tttttgttat gcctttcaag
57781 ataaatccag gaaaggaatt gcattttctt tccagaaaac aagttcttgg gggaattggt
57841 caattggtag atggtgtttt tctcattaac aagtgagtgc tccatcacac ttgctgagtg
57901 ctccatcaca cttgctctct gcattactcc tctgcctgca aacacatata tagcaagggg
57961 gatgacaagg atatcagagg gtctggtttt ctcaaactca tgataaactc atggctgggt
58021 cattcttggg gctgatttta ctttgttttt tgttgttatt gttccctctt cctcaaaaga
58081 tgaaatctat cctcttact tggaatttct ctttgatata tagcgaatgt ttggttgtaa
58141 cctgtataat ctggcatgaa attgtcactc gaaaaggcta gaagtgttga cataaatatg
58201 ggacagcaag agttgctcct actcaagaga gcaaataata tgttctggaa gagattggca
58261 gaattcacat caaaggagtg attacttcag cctgggccac tgttgtagtg gtcaaaaggc
58321 tgtgcaaagc tctctgaaaa tccactcttt tattgctctt tagtaataaa gtcactttca
58381 attttaaaaa taacaaactg atatattttt atgactcata aaatgttagc aattatatta
58441 tggagaatct actttctggg tgattcttac aaatgttctt ggatctattt tttttctta
58501 tagtacctat tcttccatt tttctcagct ctagttaata tatttcaaca acagttcaac
58561 aaatttaaca tttttataaa aagtgtttcc tatcatttta taaataccag cctagtccat
58621 gttattcctt ttcttggtga ggagaaagga cacacattgt aaattcaaat atagacctct
58681 actgtgctat ttaatcttgg taacaactcc acaaaggaga tgacatgttt tcttctata
58741 gaggtagatt ctgtaaagtt agaggaaga gtgacttgct taagatggca taagctgtaa
58801 ctggcagaac caggattcaa agccagggtg gatgccaaa tcataatctg tcttcagtgt
58861 caagttactg aaattggtaa acattagacc taaatagacg gaattgcaat ccgggttggg
58921 cacattaaac tccattttct tcatcaatgt gctcagatta cattttactt ttcaggctaa
58981 aaatggaaaa aaagagtccc tcttagttct gcacttgaga atgagaatag cttttctgaa
59041 ttatacaagg aagaagaact aatgccccaa tgccaggtag ccacatgcac tatgccatgg
59101 cacagctggt gccccctttc accagagccc tctctctgta tcttggttga cctttccttg
59161 ggcaagagct ggggtggggag gatcacaagt gactccaatt tggatggctt cgggaagact
59221 gggaccgagc tgaaggcagt gttgtcctct gcactccctg ttttctgtct gctggagcac
59281 tgaagcctca catatgtatt aaaaaataa tttccatttg catttcagac tagaagattg
59341 aacgtatagt gtaatgtgat tgcaaataat tatattgaaa tgagacagag aggatgtagt
59401 atctactgtc ataatttttc aaaacccacc tgcaacttga attaaaagaa ccacttgggt
59461 tttttttttt gtttcaaacg caaatcctgg aaacctactg agactcattc agtcagtatc
59521 tctaagaggc aagcttgaga ctgtatatth aaaaagcatc tcagggtgatt tttacacatg
59581 ctaaggctta agaaccactt ctctgtagct tatatgttat tttcaatgtt cctcaaagcc
59641 aagttagaat ttccaaagtg ttaagaatcc attagacaat cacagaattg tcttttctct
59701 ttataaatct tgcaatgttg ttctcatttc catacttaat tacttaaaac accaaccaac
59761 caacaagcaa aaaatgatta gtctaactaa tattacaagt taataatgaa gtaaagggtt
59821 aaaaataatg tcataataat gttaataaca aattattaat tataatthaa aaataatatt
59881 tataatthaa aaataatatt tacaagtagt acaagcaaaa cactggtagt ttcattgtta
59941 tcttttcata taaggtaact gagggccaga gagattaaat aacatgocaa aggtcacaca
60001 ggtcatatga tgtggagcca ggttaaaaat ataggcagaa agactctaga gaccatgctc
60061 agatcttcca ttccaagatc cctgatattt gaaaaataaa ataacatcct gaattttatt
60121 gttattgttt tttatagaac agaactgaaa ctgactcgga aggcagccta tgtgagatac
60181 ttcaatagct cagccttctt cttctcaggg ttctttgtgg tgtttttatc tgtgcttccc
60241 tatgcactaa tcaaagggaat catcctccgg aaaatattca ccaccatctc attctgcatt
60301 gttctgcgca tggcgggtcac tcggcaattt ccttgggctg taaaaacatg gtatgactct
60361 cttggagcaa taaacaaaat acaggtaatg taccataatg ctgcattata tactatgatt
60421 taaataatca gtcaatagat cagttctaht gaactttgca aaaatgtgcg aaaagataga
60481 aaaagaaatt tcttccacta ggaagttata aaagttgcca gctaatacta ggaatgttca
60541 ccttaaacct tctctagcat ttctctggac agtatgatgg atgagagtgg cattttatgc
60601 caaattacct taaaatcccc ataatactga tgtagctagc agctttgaga aattctaaag
60661 ttttcaagtg ataagactca atttatacaa agctaattgg ataaacttgt atatgattaa
60721 gaagcaaata aatacttatt atgctttttt gctgtttatt taaatattta acccagaaaa
60781 taagtcactg tgacagaaat aaaaatgaga gagaaggggt agccactctt aggtagttct
60841 ggcattatth aatctaggcc agaggttgca aatgggtgtc catagaacta attttggctc
60901 ctagacctgt cttattthaac ctttcattta aaaaatttgt attgggtgcc agcaattaaa
60961 aattgggaga tgtctcacac acacacacac ataaacacac acactcatgt gtgcagcctc
61021 ttttgaagaa ttggaataac tagtcaactg cgtcctcctt ttccacaagc tgtgacagct
61081 cctgctcac agagcacctg cctctcctg ttcatcatgc tctcttctca gtccatttcc

FIG. 1 (cont.)

61141 ttcattatat cacctatttg gtcctgagac taagtgagtt tgagatctgt gatttagaca
61201 aagtggtgaa tctagctctg aatcatagta agtagctctg ggaatcatct tgtcttctgt
61261 tagccattg agagagaaat agagagagag agagagagaa agaaagaaga agaaacagat
61321 ctggggagag tcaactgaatg ggagcataga gacagagaaa cagatctaga aaaccaaact
61381 gggagaaaat gagagaaacc aaaagagagg tagagaggag cagagaagaa aatgaagaag
61441 caaggcaagg accaggcttt ttcattatth cttatggcca agacttcagt atgcgtggac
61501 ttaattcttc cttatgctcc taccttccct agggaaactg atttggagtc tctaatagag
61561 cccttctttt agaatcacag tttgatgcct taaaactagt tatatacctt cacatgcttc
61621 cttaaccacac agaagtgatg ctaatgaggc cottaataag gagcgtgcta ttaagatgaa
61681 gacattcatt ttttttctcc gtccaatgtt ggattaaggc acattagtgg gtaattcagg
61741 gttgctttgt aaattcatca ctaaggttag catgtaatag tacaaggaag aatcagttgt
61801 atgttaaato taatgtataa aaagtthtat aaaatatcat atgttttagag agtatatttc
61861 aaatatgatg aatcctagtg cttggcaaat taactthaga acactaataa aattattthta
61921 ttaagaaata attactatth cattattaaa attcatatat aagatgtagc acaatgagag
61981 tataaagtag atgtaataat gcattaatgc tattctgatt ctataatatg tttttgctct
62041 cttttataaaa taggattthct tacaaaagca agaataaag acattggaat ataacttaac
62101 gactacagaa gtagtgatgg agaatgtaac agccttctgg gaggaggta gaattthta
62161 aaaattgttt gctctaaaca cctaactgtt ttcttctttg tgaatatgga tttcatccta
62221 atggcgaata aaattagaat gatgatataa ctggtagaac tgggaaggagg atcactcact
62281 tattttctag attaagaagt agaggaatgg ccagggtgct atggttgtaa tcccagcact
62341 ttgggagacc aaggcgggtg gatcacctga ggtcaggagt tcaagaccag cctggccaac
62401 atggtaaaac ccggtctcta ctaaaaatac aaaaaattaa ctgggcatgg tggcagatgc
62461 ttagtccca gctgctcggg aggctgaggc aggagaatca cttgaacctg ggaggcggag
62521 gttgcagtga gctaagatca cggcactgca ctccagcctg ggcaacaagg cgagactctg
62581 tctgaaaaag aaaaaaaaaat aaaaaataaaa ataaaaagaa gtggaggaat attaaatgca
62641 atataaaaagc tttttttatt tttaaagtcac acaatttggt tcacataaca gatcaggaaa
62701 taatacacag atcataagtt ttggagctgg gtttgaatcc tggctctgcc atttactthc
62761 tgtgtaatct aagtcaagtt actgaactth gtgggcctc tggctctcca tgtgtaaaat
62821 ggagaatatt aatatthacc ttgcaagttt gttgtgaaga ctgaaggaga gaatttaggt
62881 aaaacattca tcagagtacc atgcacacag ttgttctca ataaacatta gcttctctga
62941 ttgcaagttc cagtctaaag tgctthtat ataccagcca ataaaaggat gcgagagaga
63001 tataccagtg tattgtthtc taccattthta aacctattht catccactgt tacaaattct
63061 atcatactgc tccacataaa aaatattatc aatgatttht agtctctgaa gtgcaatatt
63121 tgattattga gcacacctgt tgaagttthta gtttcttctc acttacatgg tttgtgtaaa
63181 ggtaggagggt ataaaaccag tgtcctaggt ctaaatctth cttaatgtca tactttggat
63241 tcattgatat aagtaacttg agcaccagcg cttcattthta cttcatttht taaagatata
63301 gtaagagtaa ttcccatctg cctagcaaaa ttgttttgta gaaaagtttg tggatcagat
63361 ttattthtact ttgattthtag gaattthcaag tgtcttctgc ggcataaagg aaaaatatgc
63421 agtttgacat tttctactac tttcagggtca ttattthtct actctggtgc aaaaaccctc
63481 aattcctgtc tcactccatc taatcaataa ggtagcatgc ttgagccctt actatgtgoc
63541 aggcactagg ataagcactt tataatgttht gtccaatta attctcacag catttctatg
63601 acctaaataa aattaatatt ttcattthcac caataataaa atggaggctt caaaaagttt
63661 agggacttgg ctgagctcac acaactggca aggactgaaa atggattthta gtcccaaatg
63721 tcataggcta gagccctthc actaaactgt tgtcttccat ctgggtggcat cctcttctc
63781 cagtctttgt cacctaaact ctgggcaccc cttgatggca tttacttatg atgggtgatgc
63841 ttgttaaact tctgtttgac gactthcaacg tccatataaa tgagtcttcc aatactgtac
63901 ttagaactta tatthttag tagcttcttht aaaagctthc tctcttagtc atactctgag
63961 ttttgttagc acctggactt accttacttht ggaaatgttg cactctgaaa tctctthtct
64021 agcttggaat ttcttaatct tccaactgtt tgagtcttht aattctacat ttactgcctt
64081 tccattthcat caggattthct agtctctthta attcttctct ttgaactcct cctgattthaa
64141 cctctgctta ttggaagaac aataattthta ttctctcagc tgcactctca attccttht
64201 cctthtgggtg attthtcttht ttctacagaa acacttactt tatcagttth ggagaaggaa
64261 gtgctatctg ggtaacagta gtgctatctg ttgactctag tcaactgtaa gttthtataca
64321 tttattgttht aaaccttata tgggtctata atcttcttg ggaaatcctt tcatttgtct
64381 ttaattthcct ttaccattthc cctaaaggct attccagatt tttatcacat tcacaaaatt
64441 cccgtcttht ctgaggatct gttcaccccc agtagatagc cttgtctccc acaatacatg
64501 gagaaaatag aggccaccgt catattthgaa tgtthccaac ttctctcttc accttggaa

FIG. 1 (cont.)

64561 ttatcttttt cttcttttgt gtctaagaga aagatgtata cttctttctta cccttgtctg
64621 aactactcta ttttgcttca tcttctcaga acaggggacc agcaattatt cttoctccag
64681 aagcttcaac atcttttgtc aactgactcc ttctcatgtt taaatatttt caagttaaac
64741 aatttctttc ctgactttcg ctcacgcaac ctcatgccca aaaccttato actcttcttc
64801 cctttgctgt caaggctgtt ctcacttctt cactttttgt ggacttctcc ccactacaac
64861 atagattctg ctatcaccaa tctattaaaa ctgttatact cttgtggaat ttatcattta
64921 atttagcttc agtgaaccgt tctttccaga ttattttggc ctcagaccat gacttctaag
64981 tctgccgtgc ttgccactta agtgatgatg ggccagtggg tccccacctt ggctctgtg
65041 ttagtctgtt ttcatgttgc tgataaagac ataccaaga atgggcaatt tacagaagaa
65101 aggggtttga gggactcaca gttccatgtg actggggagg ctcacaatc atggtggatg
65161 atgaaaggca tgtctcacat ggaggcagat aagagcatag aacttgtgca gggaaacttc
65221 cctttattaa accaccaggt cttgtgagac ttcttacta tcacgagaat aggatgggca
65281 agacctccc ccattgttca attatctccc actgggtccc tcccacaaca catgggaatt
65341 atgggagcta taattcaaga tgagatttgg gtgaggacat agccaaacca tatcagctc
65401 cttctggctt tttatgttct ccgtgggtga cctctctcag gctcaagtga taaccaatgt
65461 gctgatgact ctcaaatgcg catctctggc ttcagtttct tcttgaact tcatacatat
65521 gtttccaaat ttctgctg tacctcaagg ttcttgttca tcacttccca agcttcataa
65581 acgcactcat tttagtgtat tctctgtctc ctttgatagc atccctgaga ggcaagtccc
65641 tgggtgagtta tatacaactc ctcccttgct ccaaacctga gagtaagtaa cattcctatt
65701 aacatattag gaagctgagg cttagacagt ttaagtaact caagcatggg tacacaacta
65761 gctagggcag agctaaaatg tcaggctagg cttctgtgac tccaaagccc tttctcactt
65821 agcatatcat cacttatttt tttttttaat cacatatatg attttttttt ctttaagaga
65881 tagaatcttg ctctatcacg tgggctggag tgcagtggca caatcatagc tcaactgtaac
65941 cttgaacttg ggctcaagtg atcctcctgc cttagcctac tgagtagcta gggctacaga
66001 cacacaccac catgcctagc taattttatt ttattttatt ttattttttg agacagagtc
66061 tcaactctgc acccaggctg gagtgcagtg gtgcgatctt ggctcactgg aacctctgct
66121 gcccggttc aagcgattct cctgcctcag cctcctgagt agctgggatt acaggtgcct
66181 gccactgtgc ccagctaatt tttgtatttt tagtagagac ggggtttcac catcttggcc
66241 aggcttgtct tgaactcctg acctcgtgat ccactcgct cggcctccca aagtgtggg
66301 attacaggtg tgagccacca cgcttgcca cctacctaat ttttaatttt tttgtagaga
66361 cagggtctca ctacgttgcc caggctggtc ttgaactcct gttctcaaac aatcctcctg
66421 cctcggacac cccaagtgca gggattacag gcatgagtca ttgcagctga cctgtatata
66481 tgatttttag tatatgtaaa tatacatatt tattaatgt aaatataaat ataatgtgt
66541 ggagtgatat ccattgaaat gttaaacata gttctcagt gtacaactac aggtgatttc
66601 tcttttctta tttctggttt tctgtgtttt ccaaatttct tgaaatgtgt cttctgtaat
66661 cagaaataaa agttattagt aacaacagtc ttccactggg acaagtgcct attggataaa
66721 agtcccactt ctaagcatga tactcacaa ctttaggtta atagcctttg tcaccttgcc
66781 atatacatct gatccagcca ctcacaccat tcttgagata tattttgttc ctttgtgcct
66841 aaatcattgt gcatgcagat ccattctcct ggaacacctt taacctttc ttagtcctgt
66901 gaaatcctac ttacatcctt catagcctag catgtatgtc atttattttg tcaaggggta
66961 gttggttggt ctcttgaatg tactgccata tgacgtgggt tgatttcaat tgtagacca
67021 agctcattgc aatattaatt cgtttgtcat tctcccatgt aggatgtttg aagtagtttc
67081 taacacagag attatactca ataaatattt attagataaa taaatgaata agggaataac
67141 aaatgccttt gtctcatttt aaaatacttt cattgttagc taccatata ataaaaaact
67201 aaaagcagta gttttcaagc atgattgttt atgtatgcct taaaagaatt ttgaaaacct
67261 atgtaccctt gacacacttt taagttaact tataaatttt tcaacatagt ttaagtggg
67321 ggcaaatgat gtagtttctt gtgtatttta aactgcttaa gtatgctata catggatttc
67381 ttcaaaaccc tgaagctgca gtttcagtgc attcaattta tggaaaagaa attaatttat
67441 aaaattgggt cttattgtca agtcaatcag ctaaatataa cttgctttct gtcaggaaaa
67501 gtctgacttt aaaatacaga taagtaataa ctattattaa ttaattaaat tattaaaatt
67561 aaaataatta aataatttgt taattaaaat gccttattcc cctacttatt tctgcaattt
67621 gactctaaga atagatagga catgtagatt gccttaggtt tgaaatctgg gtgaaataag
67681 atactgcctc cttcagtatt tctgcctttg cttttatggg agcctctttc aagaaaaagt
67741 cattctctca tgggtccctt gtttgagtcc cagaggtttt cctactccag aaagtgcaac
67801 gtagtgagac tagtactata ctcccttgca tggtaagtga gaaggctgtc tgtataaaat
67861 gagggaagga ctcatgagag ggaagtaggt caggagaaat gataggttct caggcaggtt

FIG. 1 (cont.)

67921 aatTTTTagga aagagtgaat agagTccctt aaaacaaggt gcatctgctt cctcctgato
67981 aatcTTTTagg actgtttact ttgatttgaa gaccactatg ctaaagcttc ccacgggggc
68041 aatagtgagg caaggaatTT ttaaaaggga attacttctt cgtagctact tttgtgaaat
68101 gaattcattt gaattatctg gcaatctctt catatttata ttcaacaata attacttaaa
68161 gaaatgcttt gagcttctca gaggagggTg ctaccagtgt gatggagtag aattcagatt
68221 tgggtagtga ctttaaagct gtgtgacttt agtcatttaa ctgctgagtc acagtctaca
68281 gctttgaaag aggaggatta taaaatctat ctcatgttaa tgctgaagat taaataatag
68341 tgtttatgta ccccgcttat aggagaagag ggtgtgtgtg tgtgtgtgtg tgtgtgtgtg
68401 tgtatgtgta tgtatacatg tatgtattca gtctttactg aaattaaaaa atctttaact
68461 tgataatggg caaatatctt agtttttagat catgtcctct agaaaccgta tgctatataa
68521 ttatgtacta taaagtaata atgtatacag tgtaatggat catgggccat gtgcttttca
68581 aactaattgt acataaaaaca agcatctatt gaaaatatct gacaaactca tcttttattt
68641 ttgatgtgtg tgtgtgtgtg tgtgtgtttt tttaacaggg atttggggaa ttatttgaga
68701 aagcaaaaca aaacaataac aatagaaaaa cttctaattg tgatgacagc ctcttcttca
68761 gtaattttct acttcttggT actcctgtcc tgaaagatat taatttcaag atagaaagag
68821 gacagtTgtt ggcggTtgct ggatccactg gagcaggcaa ggtagtTctt ttgttcttca
68881 ctattaagaa cttaatTtgT tgtccatgtc tctttttttt tctagtTtgt agtgcTggaa
68941 ggtattttTg gagaaattct tacatgagca ttaggagaat gtatgggtgt agtgtcttgt
69001 ataatagaaa ttgttccact gataatttac tctagtTttt tatttctca tattattttc
69061 agtggctttt tcttccacat ctttatattt tgcaccacat tcaacactgt atcttgcaca
69121 tggcgagcat tcaataactt tattgaataa acaaatcatc cattttatcc attcttaacc
69181 agaacagaca ttttttcaga gctggTccag gaaaatcatg acttacattt tgccttagta
69241 accacataaa caaaaggTct ccattttTgt taacattaca attttcagaa tagatttaga
69301 tttgcttatg atatatata aggaaaaatt atttagTggg atagtTtttt gaggaatac
69361 ataggaatgt taatttatTc agtggTcatc ctcttctcca tatccaccc taagaacaac
69421 ttaacctggc atattTggag atacatctga aaaaatagta gattagaaa gaaaaacagc
69481 aaaaggacca aaactttatt gtcaggagaa gactTtgtag tgatcttcaa gaatataacc
69541 cattgtgtag ataatggtaa aaactTgtc tcttttaact attgaggaaa taaatttaaa
69601 gacatgaaag aatcaaatta gagatgagaa agagctTtct agtattagaa tgggcTaaag
69661 ggcaataggT attTgcttca gaagtctata aaatggTtcc ttgttcccat ttgattgtca
69721 ttttagctgt ggtactTtgt agaaatgtga gaaaaagTtt agtggTctct tgaagctttt
69781 caaaataact tctagaatta taccgaataa tctaagacaa acagaaaaag aaagagagga
69841 aggaagaaa gaggaaatga ggaagaaagg aagtaggagg aaggaaggaa ggaaagaagg
69901 aaggaagtaa gagggaagca gtgctgctgc tgtaggtaaa aatgttaatg aaaatagaaa
69961 ttaagaaaga ctctgaaag gcaattattt atcaatatct aagatgagga gaacctatt
70021 ttgaagaatt gaatatgaga cttgggaaac aaaatgccac aaaaaatttc cactcaataa
70081 atttggTgtc aggctgggtg cagtggctca cacttgtaat cctagcactt ttggaggcag
70141 aggcaggTga attgcttgag tccaggagtT tgagaccagc gtgggcaaca tggcaaaccc
70201 cacctctaca aaaaacacaa acaaaagaaa atagctgggt gtggTggTgt gtgcctgtag
70261 tcccagctac ttgggaggct gaggtgggag gatcacctga gcctgagaag tggaggctgc
70321 agtgagccat gattgcacca ctgtacccta gcctaggTga taggctcaaa aaaaaaaaaa
70381 attggTgttt gcaatgctaa taatacaatt tggTtgtTt tctctccagt tgttttctta
70441 catacgaaac agcttttaaa acaaaatagc tggaaTtgtg cattttttct taaaaaaaca
70501 ttttctttct taaaatgtta ttatttttct tttatatctt gtatattatt actagcagtg
70561 ttcactatta aaaaattata ctataggagg ggctgatact aaataagtta gcaatggTct
70621 aaacaaggat gtttattttat gaaaaggtag taattgtgtt tcatagaatt tttaaaatta
70681 attctgcgta tgtcttcaag atcaattcta tgatagatgt gcaaaaaatag ctttgggaatt
70741 acaaattcca agacttactg gcaattaaat ttcaggcagt tttattaaaa ttgatgagca
70801 gataattact ggctgacagt gcagttatag cttatgaaaa gcagctatga aggcagagtT
70861 agaggaaggc agtggTccct tgggaatatt taaacacttc tgagaaacgg agtttactaa
70921 ctcaatctag gaggtgcct tttagtagta ttaggaatgg aacactttat agtttttttt
70981 ggacaaaaga tctagctaaa atataagatt gaataattga aaatattaac attttaagtt
71041 aaatcttacc cactcaatac aatttggtaa tttgtatcag aagcttaaaa gataacctaa
71101 tagttcttct acttctataa cttacccaaa tatgtttgca gagatcttat gtaaagctct
71161 tcattataac actgctttca ggagccaaaa attgggtggg ggagcccat aaatgttgaa
71221 taataggggt ttgattagat aaattttggT gtagtTctat aatggcgtgt tattcagcca
71281 ataaaaggTt tgttaaagaa tgaactgtgac ggatgtatat gatatactct taagtgaata

FIG. 1 (cont.)

71341 aagagttaca aaatggttatg tacaagttac aaaatgtatg tacattatga tccatttttc
71401 ataaaaatcat atgtatgtat atatgtgtgt ctggaaggat aaatttatca agttgttato
71461 tctgaaattt tgggtatatatt ttatatttct agattttctg ttactttgtt actttactga
71521 taaagtaata acgttggttga cttttgtcac tctccctat taataatcat ctaggctgca
71581 aaaggatcat gtcttcttta tttttatatt ccaaggactg tcaacaagtg cctagcactt
71641 gacaggtata ttatagaaat ttaactgaat atcttttagga aatagatttt tgtttgtagt
71701 tgttctagtc tacattaaat gtcttgctgt tatgaaactt ccttgaatta ttttagtgaa
71761 gcaatattag tatagaattt tgcactactg gatgcccttg actgaaagct ggcttatggc
71821 atctcaccag tgtgtgggga gtttcagtc ttctgttgtc tgcactcacag ctgaagcagt
71881 gctgttgctg acaattcctg acaccacctt gtctctatta ttgatcattg cctcactatg
71941 gtactgagtt ttagcttatt cttgtaataa ctgggactca tatgtataga ataagctatt
72001 agctcacgtt tttgcttgct ttttatacag aatacatgtc tgcaaatagt tttatcaata
72061 ttttggaatt ttgggagata tgaagttaaa aacatcattg aatatatata tatacacaca
72121 cacatatata tatgacacta tacatgattt attttattta atttttaaaa ttttattctt
72181 tttagagatt aggtcttact ctgtcaccca ggctgaactt cagtgggtgtg atcatagctc
72241 actgtaacct tgaactcctg ggctcaattg acctttccgc ttcagcctcc caaagtgctg
72301 ggtttatagg catgagccac tgtgtctggt ccaatatgca tatatatatt tttaacctgg
72361 attatcagag ctatatgtgt tttaggttta taaagctgta ctatgtgaaa atatcacttc
72421 taggtttaat tttgtacaaa ggaattttat atagaaatga ggtaattcag attttttccc
72481 atgtaataag aattgtaaaa tttactgaaa caaacatcaa aaagatatct gttacatgac
72541 cttcctttct tttgaatata tttcagggtga tattatttat taaaatttaa aaatgaaaat
72601 taaaatatat aaaaagttga aaattattcc tttctttact gtctctcato tgtccatttt
72661 ccattctcct gcattccctc atccaacca ggtagccaat ccaggtaact ttttttagta
72721 tcttcccaga gatgtttctc tctatatata taatcaatat acatttttta ttattcccca
72781 cctctctttt tatgtaacaa tatgcagagt tttgcttctt gcttttccca ctatcttgga
72841 caactttcca tattcaaagc acagaggact tgcacatatg ttcagactgc tgaatatttc
72901 tgtctctccc ctgccattca tatgttgaaa tctaattcc caaggtgatg gtattgcagg
72961 gtggggcctt tgggaggtga ttagtccatg aggggtgaagt ctttagtaaa tgagattagt
73021 gtctttataa aagaaaacct agagagaccc tcacacctta gagagacctt cacccttttc
73081 tgccatgtga gaacacagca ggaagacagc tggctatcca ggattcagga gtctcttagc
73141 agacccaaat ctgctggcac cttgatcttg gacttcccag cctccagaac tgtgagaaat
73201 aaattcctgt tgtttataag ccacacagtt catgggtattt tgttatagca gcctgaacaa
73261 ggacacacac acacacacac acacatgcac acacatttaa atagatgcat agtattctat
73321 catatggatg gatattctat gatataatga atcactattg attgacattt gggttgtttc
73381 caatattttg ttaacacaaa gaacaacact acaaataact ttatatacat atcatttagc
73441 acatctgcaa ttgtatcagt aggtctccta taagtggcca agcattttgtg tacttgtgat
73501 tttggtagat gttgtcaaat gtccttccct gaaatttgta ccaattcgta ctcatgccat
73561 acactctaaa tagagtgctg atttccccac agcattacta acagatgata ttatctaatt
73621 taaaaagttt ctcatcttat agggaaaata gtatgtcaat gtattcttaa cttgcatttc
73681 ttttattata agtagtgtaa aatatcattt caacttatac acaggaggaa tttctctcta
73741 tataaagtga tctagaatc ataataaaaa atatcaccaa ctcattagga aaatgtacaa
73801 aggattgaat agatatctca tcaaaaataa aaatataagt ggccttttaa cattgaaagg
73861 taacatttga acaaagactt gcaggaggtg agggattagg gaatgcagac tctgggaaga
73921 gtcttccaag tagcaggtga agcaagtgc aagctttcag atgggactga ctatacctgt
73981 ctggtttgaa gaacagtaag gaggtcactg aggtggcat agagtaagac agggagggtg
74041 gaatactgtc agagaagtaa tggcggtgg aggtaggggg taaaccataa agtgctcgta
74101 aagactaagg cttattttctc tgggtgagat tagaggccac tggagagttt taaacagaag
74161 taacagggcc actttggcta atgttttttag gctattctgt agggagacaa gggaggaagc
74221 aaggagatga gttaggagtc tattgtgcca gttcaggcaa gtgatgatgg tggcttgatc
74281 caggtagtag tgggaagtag atagtaggaa gtgatcagat tcaggacatg ctttgaagga
74341 agatccaata ggattaatgg ataagttgaa caatggcata tgagaaaagt cacagaggag
74401 tcaaagatga ttccaagctt tctggactga gtaactggaa ggataaatgt gccgtttact
74461 agaaagataa tgggagaaac aggttttgga tggagcttgg tttgggaata ttaagtttga
74521 aatgcctatt tgacatccaa atagagatgt tagttggatg tacaagtcta gtttcaagga
74581 agagggggct ggtagtgtga agatggggct ggataagatt ctaaaggaaa gagggttgat
74641 aagaagagaa aggggtgtag gggttagcct aagggcattc taagtattag aggttaagga
74701 ggtgggtgaa gaaaacccaa taaaataaaa gtctgagaag acaaagctag tgaatgaatg

FIG. 1 (cont.)

23/63

74761 tgggtatcccc gaacccaact gatgtcaagc agaaggggtgt tatcaactag gtcaaatgct
74821 cattcatcaa gtaagatgaa actgttataa ttaaccgggtg tcttctgaaa tacggagata
74881 actcgtgact taatgaaagc aatagtagag aagggtcaaac ttgaccagaa tgaaattaga
74941 aagaataaga ggaaagaaaa gaccaaatac agacaaccat tgatgcctta ttcttttgat
75001 atactcctgg agtccacttg ctaatacaat tgacccttaa acaatacagg cttgaactgc
75061 atgggtccac ttatttgtga attttttttc agttaatata ttggaaaatt tttggggttt
75121 tttgacaatt tgaaaaaact caciaaactgt ctagcctaga aataccgaga aaattaagaa
75181 aaagtaagat atgccatgaa tgcataaaat atatgtagac actagcctat tttatcattt
75241 gctactataa aatatacaca atctattata aaaagttaaa atttatcaaa acttaacaca
75301 cactaacacc taccctacct ggcaccattc acagtaaaga gaaatgtaaa taaacataaa
75361 aatgtagtat taaaccataa tggcataaaa ctaattgtag tacatatggt actactgtaa
75421 taatttggaa gccacttctt gttgctatta cggtaagctc aagcattgtg gatagccatt
75481 taaaacacca cgtgatgcta atcatctccg tgtgagcagt tctctctcca gttaaattgca
75541 tattgcagta aaaagtgatc tctagtgggt ctgcataatt tttcatcatg tttagtgc aa
75601 tgccataaac cttgaataac atcaagcaat ccatacaaa tgccactagt gatgcacgga
75661 aaagttgtaa cagtacaaga aaaaagttga gttgcttgggt atttaccata tattgaggct
75721 tgcagctaca gttgcctgca atttcgagat aaatgaaccc agtataaaga ctggtgtaac
75781 aaaagaaaaag aaaatgtgaa accatcagtg cagctatgcc agcagggtgtg aagtcttgca
75841 ctttttgcaa aatacaaaaat atgaaatatg tgttaattga ctgtttatgt tatctgtaag
75901 gtttccactc aacaataggc tattagtagt taagtttttg tggagtcaaa aattatacgt
75961 ggatttttga ctatacagtg ggttggcacc cctaaccctc atgttgataa agggccaatg
76021 gtatattatt taattttttt gtatttatat tcataaataa gattaaatct atatttccaa
76081 gtaatctcta taagattttg ttattaatat tactattatt tttgagacag agtcttactg
76141 tcaccaggct ggagcacagt ggtgcgatct cggctcactg caacctctgc ctcccgggct
76201 caagcaattc tctgctcoa cctcccaag tagctgggac tacaggcacg cacaaccaca
76261 ctgagcta at ttttgtat ttagtagaga cggggtttca ccatgttggc caggatggta
76321 ttgatctctt gacctcatga tctgcctgcc tcggcctccc aaagtgttgg gattacaggc
76381 atgagccact gtgcacagcc attaatatta ttgttacc caataaaaaa tttggaaact
76441 tgtcttcttt tccctgatt ctgtttaaat agcactggag ttacctgtt tgaatttttt
76501 ttccaagcgg tcccttatga gttttctcta tgttttattt gtttcatttc tttttttttt
76561 tttttttttt tttttgagac ggagtctcgc tctgtcgccc aggctggagt gcagtggcgg
76621 gatctcggct cactgcaagc tcgcctccc gggttcacgc cattctctg cctcagcctc
76681 ccaagtagct gggactacag gcgcgcgcca ctacgcgcgg ctaatttttt gtatttttag
76741 tagagacggg gtttcaccgt tttagccggg atgggtctga tctcctgacc tcgtgatccg
76801 ccgcctcgg cctcccaaag tgctgggatt acaggcgtga gccacgcgc cggcctgtt
76861 tcattttctta tategtat tttgcaactcc tttattgata cttttcttcc tgattaggtt
76921 totactaaaa ccaaacaagc tttccatgaa ttagctttta gatttactta ttagtttaac
76981 tgttctgttg tattgtaact cattaattta taattttatc tttattaatt attctatttt
77041 tottcgcttt tttgttgttt ttctagtttt tgagttagat gtttgacgct tttttaaaaa
77101 gctgtgcatt ttctctggg taatacttta gctgtatatt atgtattctg atatatagtg
77161 tttccattac attgttttct agaaaatctg tagctttgat ttatatgtt ttctctttg
77221 acctagata tctaaggga aaatttaaca tttccagaa agaaaacaaa ttttctttgt
77281 tttccaagaa tgttgttcaa attatttcta ctgcttgga tttttatcat ttttgtgtat
77341 ccagtaaata gtcaatattt gtacttgctc tctgaccaca taaaagaata tattcgtgta
77401 gtttctatta atagattaga gttcaattca gatattaaat gtacatcatt attcatgata
77461 tttaggtctt ctacatcttc acttatcttt tttctacttg ctttgccatt aacagataaa
77521 gttgaattaa aggtttctac tacatacatt tctcctgtt attccttata gggtctgtaa
77581 tttttgcttc aagaatattg ctttttaaat ttaatatata gatacttata attacactct
77641 agcattataa agagcctttt ctttttcatt gaatgtattt gggcctgcat atgtctaaca
77701 tgaaaattat agtccttttt ttgtttcttt gtttgtattt acagttttaa gttccatttt
77761 caacctttat gcactctttg ctttaggtgt gtctctttta gttagcataa agttagggtt
77821 gtctttaatt tcacctgaag tcttttcttc ttaatagatg ggtaagcca actgaaaaat
77881 aaaactgact tatatacttt tatttcaagt atgtcctcca caaatatttt ttgaatagat
77941 tagcttatat actttggaat ttgttaaaaa aagattttta taaaaataa ttgtggtgaa
78001 atgtacataa cataaaattt atcattttga ccatttttaa gggcatagct ctgtggcata
78061 aagtatactc acatagttgt gcaactatca cctccttttg attttttttt actaattttg
78121 taaatttggt tcatctgagc tgtcttatta tgttttggtt tatgtttttc tttcctttat

FIG. 1 (cont.)

24/63

78181 tatgaagtca ctgtattgtc tgtaggctat atgtatctgt gagtgtgtgt gtatatgtgt
78241 gtattatggt ttttaaaaaa gtctatatatt gttttccagt ggctatactt aatactaata
78301 actttatggt aaatttttca ttctatgtga ctctagttca ctaatatgag ctctgataaa
78361 atcagtgctt ttctogaggtt aggagatcaa gaccatcctg gctaacacag tgaaactccg
78421 tctctactaa aaatacaaaa aattagccag acgtgatggc gggtgcccg agtcccagct
78481 actcgggagg ctgaggcagg agaatggcgt gaaccagga ggcagaactt gcagtgagcc
78541 gagatcgcg cactgcactc tagcctgggt gacagagtga gactctgtct ctaaataaat
78601 aaataaataa ataaataaat aaataaaatc agtgcttttt ctccctctgc tacctccttt
78661 ccttctactc agtttttagtc agtagtatta tcttttttca gatttatctt tgtattgtta
78721 aatctgctta tgcttctatt actttattta ttagctttta atgatacctt ttgactttca
78781 gcttttctta ataaagcaat cagcaaattt cctttacact ccacacttat accccatttc
78841 ctttgtttgt ttatttggtt ttactttcta acttttctta ttgtcaggac atataacata
78901 tttaaacttt gtttttcaac tcgaattctg ccattagttt taatttttgt tcacagttat
78961 ataaatcttt gttcactgat agtccttttg tactatcatc tcttaaataga ctttatactc
79021 caagaaaggc tcatgggaac aatattacct gaatatgtct ctattactta atctgtacct
79081 aataatatga aggtaatcta ctttgtagga tttctgtgaa gattaaataa attaatatag
79141 ttaaagcaca tagaacagca ctcgacacag agtgagcact tggcaactgt tagctgttac
79201 taacctttcc cattcttctc ccaaacctat tccaactatc tgaatcatgt gcccttctc
79261 tgtgaacctc tatcataata cttgtcacac tgtattgtaa ttgtctcttt tactttccct
79321 tgtatctttt gtgcatagca gagtacctga aacaggaagt attttaaata ttttgaatca
79381 aatgagttaa tagaatcttt acaaataaga atatacactt ctgcttagga tgataattgg
79441 aggcaagtga atcctgagcg tgatttgata atgacctaat aatgatgggt tttatttcca
79501 gacttcactt ctaatggtga ttatgggaga actggagcct tcagagggta aaattaagca
79561 cagtggaaga atttcattct gttctcagtt ttcttgatt atgctggca ccattaaaga
79621 aaatatcatc tttgggtgtt cctatgatga atatagatac agaagcgtca tcaaagcatg
79681 ccaactagaa gaggtaagaa actatgtgaa aactttttga ttatgcatat gaaccttca
79741 cactacccaa atttatatatt tggctccata ttcaatcggg tagtctacat atatttatgt
79801 ttctctatg ggtaagctac tgtgaatgga tcaattaata aaacacatga cctatgcttt
79861 aagaagcttg caaacacatg aaataaatgc aatttatatt ttaaataatg ggttcatttg
79921 atcacataa atgcatttta tgaaatgggt agaattttgt tcactcatta gtgagacaaa
79981 cgtcctcaat ggttatattat atggcatgca tataagtgat atgtgggatc ttttaaaag
80041 ataccacaaa atatgcatct ttaaaaaatat actccaaaaa ttattaagat tattttaata
80101 attttaataa tactatagcc taatggaatg agcattgatc tgccagcaga gaattagagg
80161 ggtaaaattg tgaagatatt gtatccctgg ctttgaacaa ataccatata acttctagtg
80221 actgcaattc tttgatgcag aggcaaaatg aagatgatgt cattactcat ttcacaacaa
80281 tattggagaa tgagctaatt atctgaaaat tacatgaagt attccaagag aaaccagtat
80341 atggatcttg tgctgttcac tatgtaaatt gtgtgatggg gggttcagta gttattgctg
80401 taaatgttag ggcagggaat atgttactat gaagtttatt gacagtatac tccaaatagt
80461 gtttgtgatt caaaagcaat atctttgata gttggcattt gcaattcctt tatataatct
80521 tttatgaaaa aaattgcaga gaaagtaaaa tgtagcttaa aatacagtat ccaaaaaaat
80581 ggaaaagggc aaaccgtgga ttagatagaa atggcaattc ttataaaaag ggttgcatgc
80641 ttacatgaat ggctttccat gtatatactc agtcattcaa cagttttttt tttagagccc
80701 cattcttatt ttttatacac tttgagagca taatgaaaag aaaagctacc tgcaaaagtt
80761 ttggacttac ctcaaagagg atatacttca ttctcaaaa ggccttcttc caggaatagt
80821 atttcataac ctggagggtg gaaaaatctg gatttgttac aaaaaaatct gagtgtttct
80881 agcggacaca gatatttgtc taggagggga ctaggttgta gcagtggtag tgccttacia
80941 gataaatcat gggctttatt tacttaogag tggaaaagtt gcggaagggt ccttacagac
81001 tttttttttg cgttaagtat gtgttttccc ataggaatta atttataaat ggtggtttga
81061 tttcctcaag tcaaccttta aaagtatatt tagccaaaat atagcttaaa tatattacta
81121 gtaataaatt tagtactgtg ggtctctcat tctcaaatg agcatttact aatttctgaa
81181 cactgtgcta ggtcctggga ataccaaatt gaataagaca tagtctattt ttctgaaggg
81241 tttatagcag agtccctgt gttaataatg aaggagtgtg tggtatgtga atcatatct
81301 aatagggttg ttaaaaaata tgaaaaaagg agaagaggaa gaacatcttt tttttttctg
81361 attgcacggg cagccttaaa attatttttg aagtgtacaa ttcagtgttt ttttagcata
81421 ttcacagggg tgtattatca tcaccatatt tttggcctct tgaaaagaaa tctgtgctt
81481 attagcatcc aattaccgtt cctttgtagc taagtctccc ccattccagc tttaaacaat
81541 caccatcta ctttctgtct ctataaattt gtctcttttg gacatttcac ataaatgaaa

FIG. 1 (cont.)

25/63

81601 taatataata ggggttttttg tgcctaaata agcttctaaa gaagaataag gtaaggaatc
81661 atcattcagc aaatattttat taagacttgc tttatttttat acagtgtact aggagctgga
81721 gatgaaaata tgtgtagaac atgaatcata tacttcggga atttgtggac tagtgggaaa
81781 gattgacata tcaataacaa atcgaattag tgatgtaata gaggcatttt tacaggagta
81841 aaatgaggta gcatggactc tatctgggtc tgaataatgt gaggagtaac ctcccttacac
81901 aaagaggcac aaggctaatag tcctctgatg gaatgattca ccatgcaatt ctaagggatga
81961 caagaatgaa agttagggcc ttgaagaaat attttgatta agagctgcca ataaagtaga
82021 gtaaagatta gattgatgtg aagaagtggg agattaatga gtaaagtgtc actggcttgt
82081 tgagaagatt aaatgagatg tacatgtaat gtacctaaaca caacgtcttg tacaaagtag
82141 ccattcagta gagactagct tgtattatct ccctttgagg taaagaaaac tgttagaaat
82201 agtattttcta ctactgatag tatttcttct acttatgcct ccctttgagg tgaagaatac
82261 tgttagaaaa catgacatag gagaaatacc cctgagagac agttcttatt agtgactact
82321 gtgcagaaaa gatggagggtt ggtgtaatta aggagaagga aagccatgaa gccaaagtat
82381 tatgaaaaag catcaatatg aattttcatg ttgacaaagt ggtataaaaag ataattataa
82441 agatggtcac ttataaatac ggtagttctg tgtgacacaa tttacagaag ttggtatatac
82501 gtgtggaaga aaacagcata agatcctgaa ggtttgaaact gtgggcacat tggctccatg
82561 ctcaggaaat ggcaatgggg ttgggaagtg attccacttt atgtcccttt cagacacata
82621 aaaattactt gtgtgagtat cttatgccag acactattca ctgtgtagtg agcatggtgg
82681 gtatgaaatg acaactttat tgtctttcct gtcaaagaac ttgtaggctg gttgggggaa
82741 agagaccatt tcaatatgaa gtgctgagct agaggtaccc ttagggcact acagaagcct
82801 agctgatggc ttttagcctg gctagacagt tcaggatctc taaaagcagg tgccctgaag
82861 gctgagtcaa atacaaaaat gtattttgga cagaggaaat tgtatgaaca gaaacacaga
82921 acatgaaact acttggttgg tgcaggggat catcagcata gaaccagaca gaaccagagt
82981 gtaaataagc cagaaggcca tgtcatggag gccttgtata ccagtctcag gaatttggtt
83041 gtggagagct ttcatcaggg gaatgatgta atcagcttgg aaatgtagat atatcactga
83101 ctgtgatagt gaggagcaga attaaggtgg acgtgattag aagctttgtg aatagcagaa
83161 agaacataga ttttgaaaagc tggcagacgt aggttactga agaaagttac ttaaccttgc
83221 tatgtcttta gttttatcct ctgcaatatg gggataatac tgccatattt gtagagtctt
83281 gtggattctt ctggcatata taatagaaaa taaaacagct attattatta ttgttgatgg
83341 tactatttgc tatatctgac tacaaggaga aagactaata ggaaaccatt tcaggaatcc
83401 agatatggtc atgatggaca ggaagagaca agagttacat agaggaatcc tgggaagata
83461 agaaatgtca tttttatgta ctgtttgcac ccatacagaca aggcatacagg aaaaatgatac
83521 cttcaggaaa gagtgatttt ttttcttcaa gaaattagaa gaggggagaa atttggtttaa
83581 gattaaggac tccatgcata agagaaaactg ggaggggaaga caggtagaaa tgctatgggg
83641 ttaggaagga agaatgcaga ggtggattac ttagaattga gacatctgat caagacagag
83701 ggatcacagc ttttgctaac aaagtactag tggaggatgc cactaggtga ggtttaataa
83761 ataattgttg acaataagtt ccatttaaaa aataaacaat ttatgcttct tctttgccta
83821 agtgtcaaat aaaacattca gatttttatt tcaaagtatc cctgagtccc tgttcccttt
83881 tttgtcctgc tgaacttttg aactgattta ggcttcctta gtcctctcat aatagaaaaa
83941 atcagccagg tatttcctac atttcttgta ttttaaaaaa atgtaatgga tgtaatgaat
84001 ttttaagcaa tgtaatgaat acaataagta acttagtata tgctgttttc ttctctatgc
84061 tgaatgtttc atacatgtta ttttctatac aactacatgg tcaattcctt gaaaatatca
84121 actccaaaat ctttattttg gtatactcca cgtagcacat tgagagagtt ttaaactctt
84181 gttggatgac tgtttcaaaa gtgttttgaa gtaggcattg cagttgcaaa aagtttgctc
84241 agcaaagtgt gttctgtctc acagtctcag acattgagca gatgattaca tgacagcacg
84301 tgattgctgg gagtaacaga caaaagtaac tgaaaagtgt cggttatctt gacagtcaaa
84361 atcaaaagtg tcccctattt tcagtgcctt aagagtttct ttttgtgttt ttggtattgt
84421 tgttaaataa gtgttctcac ctttgaaaag gtcaataaga attcaataca gtataatgtc
84481 tgtgtgccaa atgaagggtc cccttatttt taagtgtgga ggagttttga tcataagaac
84541 ttgaaatacc tacagaatcc ttgatggtta agcagctggg gccagcacia gaatccctca
84601 atatgttctc tatgaagccc cgatcaccaa atgcaaacat tcatgattca gtatattttc
84661 atcttgactg ccaaagttga tctgtttctt aatatattac atctagactt ggaactggag
84721 atgagaacag aatattatct tctcattttt tgtgtttttg ttcaactcta atgtctgcaa
84781 agcacttgcg tatgtaatga tgctcagtgat cataggagca ggcaggtaag tgtaaatttg
84841 tctggatagg agaaagcatg cacaacatat ttcacatagt tttctgattt cagtttgttt
84901 ttgcaaatta ttcactcagt gagatagctt aaagacgtta tcacagggaa aggcattggag
84961 atagttctgt gttgatagaa aacttgtaat gtacagccat gagtgagaag tcaggttcag

FIG. 1 (cont.)

85021 attctttcacc ttcagtcctc ctcttttcata aacagctcca tgtcctatatt tacatatcct
85081 acttttaaaac gagattatag aagaatgaat ttctaggcaa agtgacactt atttttaaaat
85141 actattacgt atccctgtgc ccattaaactt atccctaccat ttttcttccc ctgtgtccaa
85201 accacottta gaatctccta aatatttgta gctattgtaa acagcactgg agactttgct
85261 agttttaaag gagaaatcaa cgcaattaag ccctagttaa tttacttato ccttatgaga
85321 ttataattgt attttgttat taaaaggggg acagagtaca ctgttctctt gcotttttta
85381 tttccagact accacttctc ctgcacttga caataccgca gtctaccacg tagtcccatg
85441 gctgacagga ggagaattct aggcaggcca gtgtttgagt agtgagtaat tggactgtct
85501 ttaccagca actcactgtt ttgtaaatgt acctgagttt ggagaagtaa ttggctttta
85561 taaggggtgc ggggtggagg gttggggtgg ggagagtgag aaggaggtca gagctttagg
85621 atatataatt ggtctccaca aagttgttgt gatacttttg gaaccacgta atgggtctca
85681 ttaactaagt gtctgtcatg acagccatta catatgcatt ataataaaaa tttatttaca
85741 gtgtaagttg aagaaggtaa aatctggatg tagtttctaa actctgcttg gcagttttca
85801 tatttaagcc actagaagaa aaaaattggg agggaagctg agaagaattt actgaaagaa
85861 aaaaatactt gggagggaaa ttggcaagaa gtatgaaaaa gcttgggagg gaagtaagca
85921 aataaatgag ttaatgactg ttctggaaaa taaactctat catgcagata tcacatgact
85981 gattaaattt gaatttgacc tctgtcttc caggtctggt aaaaactaac ctgtaagaac
86041 ttgaaactta gcotttgaat ggtcaatcca ccactgtagg agaatttatg aatgttcagt
86101 tgagagaact gaaaataaag aagtaaccata ggaattaaca tttgcattca gtagccaaga
86161 tataatggac atctgaaaca ggtatttgag gccaggcgtg gtgtctcatg cctgtaataa
86221 tagcactttg ggaggccgag gtgggtggat cacaggaggc caggagtcca agaccagcct
86281 actaaaacac acacacacac acacacacac acacacacac acactagcca ggcgtggtgg
86341 tgcacgtttg tagtccaagc tacttggggg gctgaggcat gagaatagct tgaaccaga
86401 aggcggaggt tgctgtgagc tgagattgag ccactgcact ctagcctggg tgacagagtg
86461 agactctgtc tcaaaaataa aataaaacat atatttgaaa cacattgaat tatgtccctt
86521 aaacaagaat aaacatcact aaatgactgt accttgaact acctgtaatt ttctcctgat
86581 aggtaattaa gcttcaaagt actgacactt atttactgta atatgaagca ataacttaaa
86641 aaaaaaaaaa aactattgaa ccagaaccaa acaggaatgc catagcattt tgtaaactaa
86701 actgctattt catttcattt gagccctgga acttgaaaat aaatgctagc taacatctgt
86761 gaacagaaca taccatcag tactgtgcta agcacctttc atgaactggg cattaaatcc
86821 tcactttcca tttatttagt gacaacttca cccagagttt gcagtcaaag tgaaaatgtg
86881 ctgaattcca aaagtgtgag ctaggtttta gaagttaatc acaattctgg aacaaattac
86941 tagcttaaca aatgagagtt cttatgtctc taaaaccaa atagccctaa gtctgtccct
87001 cccagtaaga tttgggccag tcaatggaac agtaatatat aaatataatt acagctgtct
87061 aggagcaaac tatcctatga atagataata aaattaagac acttaagcca tgttttcata
87121 ttaaaacaca aagtaaaaaa tcattgtttt ccaaagataa aagccatact gtatcatgac
87181 atatatatgc ccgatgtttc gacctcttg aagaattgag attctcgact ctacactctt
87241 agcgtttttc atattgaaca gatgtttaat ttaaggaggt caagagaaat cttacactta
87301 ttttttaatg gtacottaga catagaagga acctcagaaa tctctggctg aatatcca
87361 tctgcagatg atcatgtcat taggcttctg actctatagc catagaaaaa tattcatgaa
87421 gacctttcag gaagggaatg ttggtatttc taaaaattga gtacaagtat tctctagaca
87481 aaacagctct tgaaatggca gattgtattc ccattattat atttcagaat caagacatta
87541 atacctactt tttatttaac aggttttagt atccttgaat tagattttat aaattaaaga
87601 aatagatttc aataaatatt tgttgagttc ctagtatgga aacatcgtgt ttggcaccag
87661 ggatgttgcc tgcaagtata acaggagttc gtatttgtaa tgagtttatg atttacagat
87721 atttgggggg caaagatatc attcggtaaa tacttatgag tgcaaaacttt gaactaggga
87781 ctggggccaa ctctaggaac atatttgatg acagagacac aatccctgtc ctcaaggagc
87841 tttcattcta gtagagaaga tgaaaaccag tacagttttg taagttagat gatattggtt
87901 aatgtagggg tcttatgtaa gtctagagaa gtagcattta atctgttctt agaaggctag
87961 gaaagatttc cctggaggaa gtgacattta agctgagaga ggatggataa acaggagtca
88021 tctgagttaa caacagggag aacattccag aaagagaaca aaatgtacga ggctgatgc
88081 caagagagaa cattcattgc attggggaac tatagtcact tctgtgtggc tgggatgtag
88141 aatgaaatga gcctggaccc aagagagcac tttgcccttt ggggaagctg taggtattac
88201 agtaagggtg gagtctggaa agaaaggggt atattgtgag atctgaattg ggagaggaca
88261 gttatatcca gacctttata tgctccagta agaagactga actttacact gggggccatg
88321 ggactcactg aatggcatta aatttgagag tggcatatg accagatttg cattttacaa
88381 agattgtcat tgactgcaac atgaagtatg gagtattgga ggagcggtaa ggctgggtggc

FIG. 1 (cont.)

88441 agggagataa tttaggagggc tttagggtgag ggatgataat gacttgccag gtaggaagga
88501 gtaaatttct tctcagtggg taattagaag attgaatgga tggacttggt cactatttgg
88561 tatagaaggg gaaaaaagat gtcaaagatg atgccaatTT ttaaaaataa tttaacattt
88621 atttttaaat attttttcag ccttattaag gtataatgga caacaattgt aggtatatgt
88681 catttacaac atgatgtttt gatttatgta tacattgtga aatgactgcc atagtcaagc
88741 tcattaacat atccatcact cacataatta acattttgtg tgtatgcagt gagaacatca
88801 ggctctactc tcttagcaat tttcaagtat agattacatt tgttaccaac tatagtggcc
88861 acactataca atagagctcc aggacttatt catcctgcct aactaaaact ttgtactctt
88921 tgaccaacat cttcccatTC gtctctcctc cccatgccaa gtttccatct tggtcagttg
88981 ggtggatagt agtactatct gccgaggcag gttggtaggg tgaaaacaat gtgttcctt
89041 ttggaaatgc tgaggtgacc agggaacttc caagggaatc tgtctggatc tagagcttag
89101 aagagatgtt tgggctggaa acagacatca ggtattcttc agtatatggg ttgtaaatga
89161 agtcacagga gtgggtgata tcaccaatgg tgagtgtagt ataagaagac tggactgagg
89221 acagatttcc aaggaatttc aatacttaag aggtacgcag agaaaagagg ggctgtgaag
89281 gacaccaagg aggagactaa gagccaggag ggaaaacttt caagagagta ttgcattatg
89341 gaagggaaga agagagaaca ttttaaatga tacgcaatgc tcaataatgg tatccgcttt
89401 ggagaggcca agtaagattc ctaagtacc attggatcaa ggtccttaat cttacaaaaa
89461 cttatgcaaa tcaataataa agagatgata acccgataat caaaaataga caaggcatat
89521 aagaagaaaa tgaattaaaa atattcaaag cattcaacat atacaaatgc gctcaatctg
89581 atatataatg aaagaaaagt aaattaaaa acaatgggc atgactaaat aacagtatga
89641 gggagcctga ggagaaggag catttgaaat ttcagtacag aagagaaaag gggtgactta
89701 tagaaaaagg agacagaaac catagaacat gtttggagga taagactcaa acaggtagtg
89761 gggacccttt tctagagtag gatgaaaaca ggtaatgtgt gtggatgcaa atatgaggt
89821 ggatgtaatg ggaagttgag cgaattcata tttagtcatt cattcaaaaa tacttaattg
89881 agttactgct gtgtggcaag catcattcta caaacagagg gcacagtgat aagcaagcca
89941 gtttgtactc tegtgtact tacattctac tttgagaaga cagattataa ataggttaaa
90001 aagtcaataa tatgatgttt cagcatcaac aataaaaaat tagggtgata tatagagtgc
90061 cagggaaagt gctttcatgg acctcttcac tctctcctct cctgggtgtc taagctactc
90121 cttcatccat gctgccattt ctcttggttt acggttccag tatagtactc atcacattat
90181 tactatagag ccatccacct tatgaagggtg aagggtgtcca tctccttact taaaaaaaaa
90241 aaaaacaaac aaaaaaacia aaaccccgaa aaacaaaaaa agaggcagaa agacagaagg
90301 tctccacta actttcacgt gccatgtaac cagcgaaatc caattatttt acagcattct
90361 agctatagaa gagtttgga agcgtagtgc ttagtgttct agcctttgta gcacaggaaa
90421 gggcctggaa ggaaaggaat tgtgtcttc gcagttgctt ttctttatgg ggaagtgcta
90481 tagcccaaac aatatttttag gaattttcat ctattgtcaa tatgcaaact ggaaggggat
90541 aatgaaaatg ttgtggttag aagtttatga aatattgtta ttcacatttt aaagtaaaaa
90601 gagggaatgt ttaagagact tgtttaagat cacatgtctc ataattgggtg ggaccagcaa
90661 tacaatccaa atctaactac ttatcttttt gctatgccct attagtgttc atattagaaa
90721 agaaattcta tctcagacac taatgatttg ttctttggac accaatgact ttaagttaaa
90781 acttcatact agttaattta attatggtgt agcagtatta ttaaaactatc aagactataa
90841 attttctatt tgtaaaggag attatgatac caaagattag tgaactaatg atattgagaa
90901 ttctatgaca taattttgaa aaatatttgc aggatattta tttttgtgta aatgatgctt
90961 tcaagctacc ataatcctaa gtaagtgtat atttgggaaa accacctatt ctaacacact
91021 tgaaatttaa ataagtcagg aaattttttt ccagatcttc tcccaaatTA tcttcatctt
91081 tttcctctcc ccttgggaaa gaatctcttc atgctcata atatcaaatt taaactatgg
91141 aagtccagggt ggtggacagt cagcaaaggg gaagatgaga agcttgtgtt ataaagccag
91201 ctcttgctcag aataaggatc tggtaggaac ttcagaagtg atgggtaggt aagtatgaag
91261 gccaggctcc aagatctaaa ttacaaagca gaagacttac ttaccaggga gctggaaaac
91321 atgttaggaa atccagagca ggaacagatt tcaagatago acaataatat agcagtgaag
91381 tactgagaaa agagtttttt tcacgggttg gatttattct agcatttttag gcagcatttg
91441 ggcatttcta agtggtcaga cttagaggag atagttaaagg aattagcagc tgctaaatgc
91501 caattcttag accagttgaa tcaaatcat ctaaaaagct ttcagaaacc agacttttta
91561 agggccattt gagagactct caaatctgga atccagaaat ctatagctag atgagtttaa
91621 ggtagagcca gaataagaaa aataaaatag tttgtttgtt tcaggatatct tttccaatat
91681 tatttccgaa cctaccccaa acaccttaaa tcaactgcatt ctatagccat tcttttaaaa
91741 atgcttgagt tattagtttt caaaaacaaa tacaaatctg cacacataca gaaataaaca
91801 ttaaagagac ataaagatat taaacagagt tacatatact tacaacttca tacatatata

FIG. 1 (cont.)

91861 ttatatataa aactgaatat taagtgtttg atattagtga caaaatctgt aacatccatt
91921 atattagtgc tttttgtact ttttgttggg tgtagtaaaa attgcattcg aatttgagtt
91981 ttctgctata tttttggtca gtccctatca gtgaaggaaa aacctttttt tattatttta
92041 ttgttttttt attttttgag acggagtcct gctctgttgt ccaggetgga gtgcagtggc
92101 atgatcttgg ctcaactcaa cctctgcctc ccgggttcaa gogattctcc tgcctcagcc
92161 tcctgagtag ctgggactac aggcacctgc caccaggtcc agctaatttt tgtattttta
92221 gtagaaatgg ggttttgcca tgttggccaa gttggtctgg aactcctgac ctocaggtgat
92281 ctgcctggct tggcctccca aagtgtctgga attacaggtg taagtcacca cgcctggccc
92341 ctttttattt tttaagctga ttgaagattc ttagttctca tgctttctag tggtgattaa
92401 tcttttagcca atatttctat atacagttat tagtaatcat gtttgactta ggtcaacaaa
92461 caatctttcc taaaaaaaca gaaccccaat tttaatttct gaattattta gtatctattt
92521 totgctgtgg aagttgaatt atgttgatag atatcataca gggccatgta acactctcag
92581 atacacgttc acatgtatag tagctgtata caaaaatggt acttcattct ctctctcttt
92641 ataatactct tggctctctt acgttctctc acacactcta ctcttccctt cctctgttct
92701 ttctacttgt tccctctgct cctaccacac ttattccccc ctgttccatt ttcttgtgct
92761 ataaagcaca agtgcttagt aattatcaaa tattaataac aatgacacta accacccaat
92821 gatttagtgt taatgacatg ctttattgaa tggcattacc tctaaagttc atgtttcctt
92881 taccoacca agcttcttac cctcctccct taccacaagc atctatattg tcaaggttgt
92941 tataaagagt aataagccag ccattaaaaa agggtttatg gtattttcct atctacaaag
93001 tcacaggaag ctcaaatgta ctcagtaaat attgcaaaat tacacaggac cattaaatgt
93061 aacactccac cctttctctc tctctctctc tctcttgctc tctctctctc tttctgtcaa
93121 tatagcaaca ccctatatca ttgccctttg tatgtgcaaa tcagagttaa taagctttat
93181 attagcaatt actccttaac aacttctggt ttgtttggtc cagttgaata atgtaagcac
93241 ttaaaaaaat gaaattataa acatttatgt gaaaagtgc tatatcacat tggatatgtt
93301 gttatgcact ccttaataat aaagtaagtt aatctttatt gcacacttat tataatatta
93361 ctttgaccct ctctagtact ctttatctaa gtattctcaa gtgctttaca atctcaaaca
93421 gacccaatgt gttgtataca cagaatcctt tgaagctgac atttgccctt ctgaccagct
93481 tgttgtaaag gaaatcagcc aaaaaacaag tatctagatg agtagctcaa acattagtac
93541 acatagtaat cacaggtcaa aatgcagata gattaacctg tccaaattct cctgagtaag
93601 agtaggtgaa acatttttaa ataagctccc caggtgattc tgaaattggg ccaaggacca
93661 catattaaga actaatgatc caaacaattt gaacttttat tgtagattaa accatgctga
93721 gaaaattatt aaaaattgaa atggcagtggt aggatggttt gaaagaaagg tttttcaggg
93781 ccctttcaac aataaaatta attgaacaca atattaaaac tctatatttg atttaagact
93841 aaggttttca ttgtttttta atctcagtaa tttttatgta acaggtcaat tcatacccag
93901 catcttaatt ccaatgaatg atttcccaca acaatttttg tggataactc caagggaaact
93961 cgaaggaagt tgtagtatga acaaagagaa gtagaatttg tccctgtgtg taaggcttct
94021 ctgataagca gcacaggctc tcatactgct ttttaaaaaa attatgatag catcaagtgg
94081 aattaatttt ttttagatta tactttcatg gaagggaaga tctactgtga aggctggaaa
94141 accaacaccc ttaagataaa tatattacca gatttgagcg ctcttagtaa tcagcaaaga
94201 taaatgttta acagtgcata caaatgaag tgttttatgt taaatcaaat agagaaagcc
94261 aaacactaat aatgtggtta caaatgaaca ataaattagg taatcagaac aggtacagac
94321 attaatagca ggatattggt attattaatg tattttgttt taaaataatg aacttaatta
94381 caattctcct catcctaccc cactatttta ttttattcca gattcagcag cttcatatta
94441 tgtctctgaa acacttatta tttaaagttat ccaaagtac acatttctct ttatataaat
94501 gtttcagtcc agaaaaggag gccaaataca ttagctcaga acatcaaate ttctcagatg
94561 tgggaatctt ttattttcac acttttaaag gtaatctgta tttctagcgt ctattataga
94621 cagaaaactt tcatatgaca acattcctat tttcttaact gccttgatag gggcgaagac
94681 aaattctaag taggactttt taccccattc ttcttaccat cattctttca caaaaccccc
94741 agcttttagac aatcgctatt atgaatttga catgtactat tccaatccat tcccataaat
94801 ttacacccat atatacatat agttatctat gaacaatatt tagtagcttt tttgtgtgtg
94861 gctttaaaaa ttacataaat tgtataattt gtgcacattc ttctttaatt tgccttcttg
94921 gctacggtta tctttttgag atctagctat gctgctggta tgtagaattc tatttcattc
94981 ttttttcatt gttgttttgt acccataacg tgtcacattt tatttatacc ttctgttctt
95041 gatggacatt tagattcttc caggatttta ctcaatactg caatgaaaaa ctttgaattt
95101 ttctcttttg cacatattca agagactttt ctgacatata tatctatagg tgaattgtgt
95161 agtcatatga tacatacaca cattttaaat ttcactagat actgccaaat tgccttttga
95221 aatagccata caatttatag taccaccagc cacttatgaa agttcccat tctcaaatc

FIG. 1 (cont.)

95281 tttgaaagtt cttattataa acagacatat taattottgc cattctgatt tgtaaatacag
95341 aatctctatt gttctacctc tagttctaata ttggaattcc ccaattaact gtaagatgct
95401 atatattttc atgtttgtta gtcattctga ttccatatcc tttaaccaatt atctttttgg
95461 taagttattg tgggtggccat gagatgtgcc ttacagaggc cttgctagag ggaatgtgat
95521 tgaatgagag ccccagatgc tgtgtattaa aatcctgcac tgagtttgtc tcaagatttc
95581 ttgcacgtga atgaatgagt acagctggga tactaaagca gatgtgtatt tgggagatat
95641 gagacttctt tagtggctga tttttggctc ataaatgact ttgccaaaacc ttccttagac
95701 tgctcagtgt tctaacatct tccatccagc cttctaccct tctttccttt actaggggat
95761 tgaatttaca ttgaggtctc atagccttct ctgctctctc ccttatttcc ttttatacaa
95821 atatttcccc taataaatcc atgcacattt aataccattt tgctatttgc aacctgcagg
95881 tcctggacta acacagttct atacattgca ttaccattct ctagagtggg atcttttggt
95941 gtagagagtt ttaaaatttt tatgtagtca cttttatcca tatttttctt tatggtttat
96001 attttttgtg cttctcttta acacatcttt tctagcagaa ttcataaata tattattcta
96061 tattgccaaa agtttgaaaag ttgcaatcat tagaattaat ttttgtatat tgtgtaagtt
96121 aagaatctaa ttttattggt tttcattgga aagccatttg tcccaagata attttttagt
96181 agtccctcct tcccctattg tcattctgac atatttttcc taggttccga tctatgcatg
96241 tgtttcttta tggaagagtt ggccctttgt atctttgagt ttcaaatacca tggattcaat
96301 caaccacaga tagaaaatat ttagaaaagc gtcagaattg aacatgtaca tacattttgc
96361 ttgtcattat tccctaaaca atatagtata acaactattt atgtaggatt tacattgtat
96421 taggtattgt aagtaatcta gagatgattt aaagtataca ggaagatgtg catatgttac
96481 atgcaaatac taccccattt atataagggt cttgagcatt catggatttt ggtatccaca
96541 gagagtctg gaaccaattc cccacagatg ccaaggcaca actgtattta ttctatcatc
96601 tacttgttta atctcacatc agtatctact tttgaaataa caataacttt attatttaac
96661 tttttttatt acttaggatt agagaatttc ctctggtgag gcatcatagt gtctcaagct
96721 ggccataaag acaagtgagg gctaggatcg gtaagactgg gcagaggaag atacaacaga
96781 tctcctatgc atgaagcaaa agtgcagctc agaagccagc tctttcatta agttgtcctc
96841 tataccctca ctagattgta agctcttgaa atgagaggct ataccttaat tgtctctggt
96901 atctaaaata cttccactca ctgcttgga catattgctt gcaataatta agcttgccct
96961 ggctcccaa gcatagagca aatcacactc ctccccttgc ctttgagaag ctacagctct
97021 tcgaaggtag agatatgtga acagataaga aaatggatga caggagaaca gaaacgcatg
97081 actgtcagag aagtcattgg agactttaca gaggaatta aatttttatt gatcttgaaa
97141 gagtttgcca gatgaagtag aggacaggca ttttagacaa agggaacagg aatgtgaaa
97201 acacaaagtg atggaagtca tgggtgagttt ggagaactat aaaacttcaa tgtggctgaa
97261 gggtaagggt gatatagagg agtgctggga ggtgaggctg aagaaataag ctaggaaatg
97321 tctttttatg ccatttttta aagtttgga cttattctga agttcacatg gatccaatat
97381 tttttgtttt gtgttgtttt aagcagaagc gtgacatgat cagcttgaat gatgaacaac
97441 ttgaattggt taaagtggat cacacagctc actgttttac agttattctt tgaccaagat
97501 attctttatt aactgaggaa aaaaagggtt ttctgaatt ttgcagtcac gggatatatg
97561 ataagcattc ttgatttatc atcttcaatc ctgttacata acataataac cattgttatt
97621 acctttagca atgctttcct cagtattatc taatggccta taaaatgtga ctttcatttg
97681 caaatacagt acatctaaca agaacttacc acagctgcta tgcaaaatac caatacaatt
97741 gacccttgga caatgtgggg gttaggggtg ctgattcccc atgcagttga acatgttaca
97801 taacataata cataaccatt gttattatgt aacaggattg aaaatgataa atctttggaa
97861 agtggggcaa atgaattctt atgaattcca tatcttccac atgtgtttta cttttttgat
97921 aagaagtagt aacctagttc agaaagaaaa taatcatccc cttttactta tgcaggatac
97981 caagtctatc ttagcaccat aatagtgaat gataggaatc aagctctatg aatacattca
98041 catgtacata tatatggcta tataggacac atgcatgcac atatacatat atacacttgc
98101 atatatgtgt atatacatgt acatatatgc atgtatattc aattgtatat gtgtatatag
98161 ccaagttatt gtacagttga cctttgaaca acacgggttt gaactatgca ggtccactta
98221 cacgtatttt ttttttccgt ttctgacacc cctaaggcaa caaggccaac tctcccctt
98281 gctcttctc ctcagctgac tcaacatgaa aactatgagg acgaagacct ttatgaagat
98341 tcacctccac ttaatgaata gtacatacat ttctttttcc ccatggtttt ctttaataaca
98401 ttttcttttc tctagcttgc tttattgtaa taatatagta tataatacat ataacatacc
98461 aagtatgtgt taattgactg cttatgttat cagtaaggct tctggccaac agtagactat
98521 tgctagttaa gtttctggta gttacaagtt atatgtgggt gttcgactgc atggggagtc
98581 agcaccocaa ccctcatggt gtccaagggt gttgtccaag ggtcagttgt aattgggtatt
98641 ttggatagca gctgtggtaa attctgggtta gatgtactat atttataaat gaaactcaca

FIG. 1 (cont.)

30/63

98701 ttttataggc cattaaatat tattgaggag agcattttcta agggtaaaat cttgtctaata
98761 gotttgaaaca ttttcatttt cctgtcagtt tagatctttt tgaagtaatt ctgaaaatct
98821 ctctttttaag ctaaatttaa cacaaccaa tagccaaata ttaagttcc actaatgaag
98881 atatctaaat ttctgttaaa aatttaagat atatgttaaa ccttctaat ataactcttc
98941 tctcagtcaa actttttttt ttaacagttg ctttgcttct tctttcaaag tcatacttca
99001 acaaagttgc tattgaatat gtctgactaa acatgttagc tatatgataa gatggctgga
99061 taagagataa atatagaaaa tgtagctttt tttctacttg caataaccct ttaggaatta
99121 aaatggaaaa ctaataacta tttgattcat aatagtagca aaccgtaaaa tatttagaca
99181 taaatctact aagaaattta taagacatat atggagaaaa ttcaattgaa taaaccgtta
99241 ttgaagtata taaaataaga tctggatgaa tagaaagatc ataattttta ataaaatttt
99301 gcatcttaaa aagtgaacct tctccaaata tatgcacatt taataaaatt ataaatacat
99361 cccaatgagg ttggttttga aattttgtta attggaactt aaatttcacc taagaagaaa
99421 aaataaagaa tagttaagag tgcctgcttt gtagacaaat tgccttagtt agaactcttg
99481 ctctatcatc tattagctat gttatctttg ggataacatt catcttttct tatagatatg
99541 cttaaaaacag tgcctgacat atagtaagca caaatatcca ttagctatto ttcttattat
99601 ttatgttatt agtattgtta atatttgtaa ttatatggaa gactaaatga ccaaagagag
99661 tcaagaaatt tatgaataag atttatgcgt tgttagatat tagagccatt aaaaaaaaaa
99721 aaaccaaagt gccaaaaaac ctgacacagt gttaatacag gaataaaaaa atggatcaga
99781 ggaaccaaac agaaaagcca gaaatggatc ttaggaaaca tgagaatatg atatatgata
99841 gatgctaaat gaattcagta taaaaatatt aatgtaataa atcatgcttg ctattcaagt
99901 aaaagaaaat gaggttagat tcatgtctca taccaaatat aaccataaat tataccttga
99961 ttaaattttt taattaaaaa gcaataatat ttgaaaagaa atataggata ctcaatgtat
100021 aacctgaagg ttgggtagta cttttcaaca aatataggaa tttttcactt gaaatactag
100081 aagaaaaaaaa gatagcaaac aaatacagga attccaattt caagcagata taatgatttc
100141 atgaaatgtt aactgtgcac atgatagatg gtctatggat agtgcaaaag aaaaagagaa
100201 aagaaaaaat gttttttaac atatgcagca aaaaaggttt ttaacatcta ttacatacaa
100261 ataaaaatga atgtataaca cagacttcaa taaaaatagg catttcacag gagaacaatt
100321 cagatggcca gtattttaca tttcataggt attaaggaaa atacaaatta aaatggcaaa
100381 ttagcaaaaa ttgaggtgtg attatattaa tatctgttgg tgggtggtgat tatggggaaa
100441 agggtacttt caaaaacttg taatataaat ataattcttt tggttgtttt gtaaaggaac
100501 ctgacaatat cttttaaaaa taaagaaaa gcatactttt gacctagcca tccattcat
100561 gagggtatgt cttagaaaaa taagatcaca aaatcataga gatttatgtg caatgatatt
100621 attggtaggt catttttatg aggaggggtg tggatagtaa atgccagggt aaatcacata
100681 gcatctaata aacgtattta tgaactacaa aagcttacac tttcagtcta gtctagtcca
100741 gactgcaaat aaatgtgagc aagtgaattc aagcacagaa gtgcttgaag gcaggtttca
100801 taaatctact ttcttacagt atcctgatat tgacttatcg agacagttac tgtggggttg
100861 attattaaaa tatttatgta tctaggtatt tttcattcag tagtatgtta ttcaattagc
100921 aacaagtgtg gggattttaa gatattcttg tttgttttta ctgctgaaac atattctagt
100981 ggaaatttcg aataaacgat tagtcatcct aaaagcaaga tacattttct cagaaaagac
101041 aaggtaaaga acttgatat cctccctcaa ttcgtttata aggtaataag atgaataaaa
101101 atatcatagt acaatttagc attgtaaaaa aaaattaatt ggtcatctct agtgtggtcg
101161 tgcttggaag gtgaaagaag ccaagatctt gtctgggaat atcatgtcta ccttgacctc
101221 acccttaaga atcctagcct ttagttttaa atcacatggc tacatacata ccaacttcaa
101281 caatagtaca tctggcaagg tcatgcaaac ctgggacttg agottctgat tctaagtcca
101341 gtgctttttg tgtacatcat ctcttgta caacctatga tgatatgcta ataaaagcta
101401 cgtgatcagg octtaaaaaat ctgctttttt tttgtaatgg tagaatgggg catattatca
101461 catcaggtaa acaactctatt caaggataaa tggaaatgaa tgtcatatat agatcattga
101521 taaatatctc attacaaaat tatgagagtt accaatgttt gagtgtatat tatgggccag
101581 cccttttatat taaattactt caaattttta caactgttaa aggaagatat tattataccc
101641 attttataga tggacaagtt agggccagaa aagacttcct caaagctgtt agtccagtaa
101701 tggagacagg gctagaaaac aggtcatttt gctctttgac taatgttact actcatgttt
101761 tgtattttgt ttaaagtttt attttatttt gctttattta ttttttgaga caagatctta
101821 ctctgtcacc caggctggag tgcaatggag tgatcacggg tcattgcagc cttgacctcc
101881 tgggctcaag cgatcctccc acctctcaat ctccagagta gctaggacta ctacaggtgt
101941 gtgccaccat acctggctaa attttgcatt ttttgtgggg acagggtttc actatgttgc
102001 ccaggctggg cttgaactcc tgggctccag cgattcacct gccttgacct ccaaagtgc
102061 cagtatcaca ggcttgagcc accatgtcca gccagtttt attttagaat taaaaaaaaa

FIG. 1 (cont.)

102121 tccacttgga ttgttacatt ttatctcatt gctttatatt tatagaatta ctttataaat
102181 gccactttct taattttcat agtttagcact ctttatgaaa cataaactat tatttgacco
102241 aggtttttgt tagaggaatt gagtcagaga gctgttaagt aactgagatt tcacaataag
102301 ccagacagac cagggttcaa attctgggtc tcacattatc caattcaata ttccagcttt
102361 gttacttatt gagcaaccac tacaagcaca gtttacatga catctgatag ctctcaaaat
102421 gaattttaca aacataattc agatttcaac tcagcagtga ctcaggagaa aggacacttg
102481 gatgcatttc tttatggcat ttttcccagg gtacacgcaa cctggaagat ctcccagta
102541 tgggggaagg tttcacctg aggaatccca ttccctctaa tctgggacaa gggggaggag
102601 agtactgtct cttatcagcc atctccccag ggaggcctgg gccctcctgg aatgcatacc
102661 atggcttact gactcaaagt gttgaaaaga ccaggcattg ggacacacaa cactactctt
102721 aaaataaaaa aagaatcaga gtagcttggt gttataattg aaatggacag agtaacatgg
102781 taccaagaaa ctattagcaa ttccctccct aaatccctca ttttcttaaa gcattttctc
102841 cttttcctca acaagcttta agttggattt gaagaatgat aagactaaaa ggagggtgt
102901 ttctggtctt tggaggaatt tgatattcca ttcgatctga gtgtgcaaag cctgagttca
102961 catgaactct tctgatctct ttctctaata ttttttcacc ttattcatat gggaaagaag
103021 gaggggaata ctttagttcc attctccctc ctccatattc cttgacttgt ttaaaatata
103081 aatgttatag acacctaaga tagaaatttg actgaaacag cctcttaatt attgtcttaa
103141 aaaattggta taatgaaatt gcatttgtag tctttggaca tttaaatcca gaagggatat
103201 tttctttttc ttttttaaaa atttaattca atagtttttg ggctacagggt ggtttttggg
103261 tacatggata agtgcttttag tgggtgattc tgagattttg atatacccat cacctgagca
103321 gtgtgcactg tacccaatat gtagcttttt atcccccccc cgctccaccc ttcccttctc
103381 gtccccaag cacattatat aattattatg cctttgcagc ctcatgggtt agctccact
103441 tgtaagtgag aacatgcgat atttggtttt ccattcctga gttacttcat ttagaataaa
103501 ttgtctctag ctccattcaa gttgctgcaa aggcattat ttcatccgt tttttggctg
103561 aatagtattc catagtgtat atatgccaca tttctttat ccacttggtg attgataggc
103621 atttaggttg gacccatatt ttgcgaatta tgaattgtac tgctgtaaac atgagtgtgc
103681 tttttttttt tccatataat gacttctttt cctttgggta gataccagc agtgggactg
103741 ctggatcgaa tggtagttct ccttttagtt ctttaaggaa tctccatact gttttccaca
103801 gtggttgtag tagtttacaa ccccaccagc agtgtaaaac tgttccattt tcagcacatc
103861 catgccaaca tctattattt tttagctttt taattgtggc tattcttgca ggagtaagat
103921 ggtatctcat tgtggtttta atttgcattt ccctgataat cagtgatgtt gagcattttt
103981 tcctgtgttt gttatttggt tgtatatctt gagaattatc tattctgtcc tttgcccact
104041 ttttgatgga attatttggt tttttttctt gctgatttgt ttgagttcct tgtagatcct
104101 ggatactagt cctttatcgg atgcatagtt tatgaatatt ctttcccact ctgtaggttg
104161 tctgtttacc atgctaatta tttattttgc tgtgcaaaag cttttcagtt taattatttc
104221 ccatctattt atttttgttt ctgttttatt tgcttttggt atcttagtca tgaacttttt
104281 acctaaacca atgactataa gagtttttcc aatgttatct tctagaatgc ttatgttttc
104341 tggctctaga ttttaagtct tgattcatct tgagttaatt tttgtataag gtgagcattg
104401 aggatccagt ttcatcttc tacgtgtggc ttgccagttt tcccagcacc atttattaga
104461 tagggtatcc tgtcccccact ttatgttttt gtatgctttg tcaaagatca gttgacttta
104521 agtatttggt tttatttctg ggttctctat tctgttccat tgtctacttg cctatttggt
104581 taccagtacc aggtgttttt agtaactata gcctgtagt ataatttgaa gtcgggtaat
104641 atgatgcctc cagatttggt ctttttgctt agtattcctt tagctatgtg ggctcttttt
104701 tagttcccta tgaatttttag gatttttttc tagttctgtg aagaattatg atgatatttt
104761 gatgggaatt gtattgaatt tgtagattgc ttttggcagt atggtcattt tcatagtatt
104821 gattctaccc atccatgagc atgggatgtg tttccatttg tttgtgtcac ctgtgatttc
104881 tttgagcagc attttgtagt tttccttgta gagatcttta acctccttg ttaagtatat
104941 tttcatgtat tttagttttt tttttttggt tgttttggtt tgttttggtt tgtttttgca
105001 gctgttgtaa aagggttgga gttcttgatt tgattctcag cttgggttgt gtcagcaggg
105061 acattttcta aagtatagac tgtagttcct tatcttctat ctgtttctta ctgtcccctt
105121 cagtattctt gtcccttttt cccgctatta tctttttgac cttttaatat atagatatct
105181 acttctactt ctgacaattt ttgcttctcc aattttcttt ctttttctcc tctgcacaca
105241 tttattttatt ttcttctatg taactcttta tttttaactt aatatttgat taacttccct
105301 tccctgtctc ttttccctct ttccataaat cttcattaat tgccctgcact gagctaggat
105361 tctatactct ctaaataaat aatctatttt ctatagtcaa ctgtgttata atcgtactgt
105421 caagataact acttattttt aatacttaaa aatattttga aattttaacc aatttaatta
105481 atacaatggt gagttcaaat ttgaaaaaaa caatggaaaa ctgtaataat tctagcaacc

FIG. 1 (cont.)

105541 tcoctgctttt taataatgta ttagaaaaatt tgcoctctttt tcaaaagcct acagtgaatc
105601 tattcataca aggcaaaagc aaaccattct ctccattctc tttttttctc caaaagattt
105661 aagtgttttt tgtttgtttg ttttgttttg ttttttagat attgagtcct gctctgtcat
105721 ccaggctgca gtgcagtggt gtgatcatag ctogctatag cctcgaatcc ctgggttcaa
105781 gcaatcctcc tccctcacco tcoctgagtag ctggggctac aggtgcatgc taccatgccc
105841 agctaattta aaaggaaaaa aattgtgtag agatgggtct tgctatgttg ccagaggctgg
105901 tctcaaactt ccaatctcaa gcatttctcc caccagcat cctgaagtgc tgagattata
105961 agtgagccac tatgcccac cagatttagt ttttaaaaag agaatacgat ttgaaaaagg
106021 aaaaatgtga ggcaggagag aagaaataca cacacgagct gttttgtaat tgctgtaaaa
106081 ctgaaatctt cagcctcact aaaggagcac ttgcatgaac acctctaaat taccttatta
106141 ccttctaaat taggtgtgaa gtctaacttc taaattatga gtgaaatcca ctgcaattct
106201 tgttatttgg atggaatcct aggtatgtgg tccagttcat gagttgaaca aaagcatgct
106261 catttaggcc aggtagaaag aaataaagac ctatgtttta catgtctcat aaccactgaa
106321 ggtccttctc ataagcagtg cttatgggta ttaacgacct ctctatatat tacttctcca
106381 gtgcctaagt agccgagtc actgagtcct gotacatctc ctccaacatg tcagcatttt
106441 tttcacaggc cttttgttac tctagatcag aaatgttgat agcaacagtt ccttgagggc
106501 agcagctagc atgatgccag ccaacaggaa ccaccaaagt gttcttaata taaattacta
106561 cttattaatc tatttacttt gtgcatttgg agttttgcat gtaaagtcct atttatgtcc
106621 atatggtaga taaatggaac aaatgaataa cagaagtaac cattttgata ctttagatat
106681 agataatatt ggattatttc tggattgtga aagaagaagg aagaagcata tggagagaaa
106741 gtttttagtag aggggaggaa ggaggaggtg gaaacgaatg tacaaggatg ggaggagaaa
106801 agggagagag actttttttt ttttaaggcg agagtttact acctatctaa ctcttcgcat
106861 tottgaaagt tcagacccaa tcccatcggt ttgaaagcct ctagggtatt ctatctattg
106921 tatacttctg ttatgtacaa aattaatttg ccaattaatt gtgaactggt ttataaacta
106981 tottaaaatg gttagttaaa tctttgggat agtattttagc tttctccagg attatgactt
107041 accttctaaa ttagacatac aatgcctagg agtcaaggac tattttgcat aaattccagt
107101 cttcttttac aatgcctaga atgattgtta ccacagaaat attcattacc tgggagaaaag
107161 gatgacagga ggggcagaat gaatggagag aggtcgtgag aatgaggtgc tgaggatgga
107221 cgaggaagaa agctgtttta gttgggagga taggtgacag aagcatggaa aggaattgcc
107281 ttggacccat ggaagcccag tgaagatact tagatcctgc aggggtgtga ataagtctct
107341 tttagtttct cttcttagga ggtttgttca ttttgggaga tttcttttga aaagagtga
107401 cttaaattgg agaaaagtac atttttagtat gttgataaca tttgaatttg taaaatggac
107461 ctatggatga tctacacata tttatatacc cataaatata cacatatatt aatttttggg
107521 attttataat tattatttaa tgatcattca tgacatttta aaaattacag aaaaatttac
107581 atctaaaatt tcagcaatgt tgtttttgac caactaaata aattgcattt gaaataatgg
107641 agatgcaatg ttcaaaattt caactgtggt taaagcaata gtgtgatata tgattacatt
107701 agaaggaaga tgtgcctttc aaattcagat tgagcatact aaaagtgact ctctaatttt
107761 ctatttttgg taataggaca tctccaagtt tgacagagaaa gacaatatag ttcttggaga
107821 aggtggaatc aactgagtg gaggtcaacg agcaagaatt tcttttagcaa ggtgaataac
107881 taattattgg tctagcaagc atttgcctga aatgtcattc atgtaaaaaa attacagaca
107941 tttctctatt gcttttatatt ctgtttctgg aattgaaaaa atcctggggg tttatggcta
108001 gtgggttaag aatcacattt aagaactata aataatggta tagtatccag atttggtaga
108061 gattatgggt actcagaatc tgtgcccgta tcttgggtgc agtgtatttg tttgctcat
108121 agtatagttt actacaaatg gaaaactcta ggattctgca taatactgga cagagaagat
108181 gtaaatatct gttagttcca tcatagacco tgccactcca atgtacacac cagctttagg
108241 cttcttggta tagataaaca tacatttttca aaatttttca tcataatttt cataacaaaa
108301 taggaaggca aatgatgtca cttggcttaa aatctataat atttaaaata aacaggacaa
108361 atgcattaac attgttgggg gaggaggtcc cttagtagaa acactcttgg tccaagcatt
108421 ttaaagctgt caaagagatg taaatataga taatgtatgt caaggagaga gctttgtggt
108481 taaactgtaa ctttcagttt aaacaattat tggtgactct gatgtcaaat gtttctcaag
108541 ctttatctga acaaaattct tctcactttg ttgccaaagt cgttaacaag aaatcacatt
108601 gactcattga tgttttggct cctttccctt actttctgtt gctttccaaa agctgagaca
108661 ggaaactaac cctaactgag cacctgcaat tgccctggtg tattctagtc atgtgtgtac
108721 ttttgtgtgt atgtaatccc cttacagctc tgcaaagtaa gaattgttct cctgcttta
108781 cagaagagat cataagataa ttgaggtgtg tagatgttaa cttgccaaaa gccatacagg
108841 aaaatggtag agtcacagtt tgaaccaggt ccttttgatt ctttacatta aaccatgctt
108901 tgatcttgga aatacactgt aaggcaataa atcaatagat acggataatt cacaggcttc

FIG. 1 (cont.)

108961 taaataaatg gaagttgatt gtttttatct gtgagccaaa gtaagactta ttctaagaat
109021 tccacaaatt tagataagat agagtatatg gcttctagac atccaacata gaactgagtt
109081 tgtgttatca gtttaagatt tgggttttgct gtaaggtgca cacactttga ggaactaaaa
109141 ataattgtct gttcttattc tgatcagaat gtgtaatgtg ttgtccagtt ttggatgatg
109201 aatttcttat ttctaattctc ataagaaact tgtcatagat gtgagggaga gaattaagaa
109261 cagagtgtgg ggaagaaact gtgtacattt tgatgggac cattatgtag ctcttgcata
109321 ctgtcttcaa aaataagtta cactataaag gttgttttag acttttaaag ttttgccatt
109381 ggtttttaaa aaaattttta aattggcttt aaaaatttct taattgtgtg ctgaatacaa
109441 ttttctttat tacagaagta ccaacaatta catgtataaa cagagaatcc tatgtacttg
109501 agatataagt aaggttacta tcaatcacac ctgaaaaatt taaatggtat gaagaaatta
109561 tctcatttct attaatatgg gaactgtgtc ttcactttta ttactgttct aaggccaact
109621 caatgtagat tttacttgct tatggtttca tatttttagct aaatagtaaa ataatatgga
109681 tatacatttt gttgtgaact actcatactt tccttatttg gaacttttat gaatatgata
109741 tagagactga aactacaagg aacaaaatgc aatatcaatt atacagttgt ggcagcactg
109801 ctatcaattt gttgatagtg gttaacactt agaaaaacat tttaaaaata atttcacata
109861 agtaatgtaa tttattagct gtctctgaca ttttacagtt tggaatagtt tattttcttt
109921 ttggtgtcct caccaaaacc caacatcttc aagggcagga actgtataat ttttgccatt
109981 gtattttgag cacatagcat ggtacttgcc tctaaataga tactattggt aaaatatattt
110041 ttaaggtaat attttaaagt gtatgctatg gtacagttca gtttgtgact tttgctagtt
110101 tatgccactt acagtttagc aaatcacttc agcagttctt ggaatgttgt gaaaagtgat
110161 aaaaatcttc tgcaacttat tcttttatcc ctcattttaa ataatctacc atagtaaaaa
110221 catgtataaa agtgctactt ctgcaccact tttgagaata gtgttatttc agtgaatcga
110281 tgtggtgacc atattgtaat gcatgtagtg aactgtttta ggcaaatcat ctacactaga
110341 tgaccaggaa atagagagga aatgtaattt aatttccatt ttcttttttag agcagtatac
110401 aaagatgctg atttgtattt attagactct ccttttggt acctagatgt tttaacagaa
110461 aaagaaatat ttgaaaggta tgttctttga ataccttact tataatgctc atgctaaaat
110521 aaaagaaaga cagactgtcc catcatagat tgcattttac ctcttgagaa atatgttcac
110581 cattgttggc atggcagaat gtagcatggc attaaactca atctgatctg cctactggg
110641 ccaggattca agattaactt cattaaaacc ttttctcacc gctcatgct aaaccagttt
110701 ctctcattgc tatactgtta tagcaattgc tatctatgta gtttttgcag tatcattgcc
110761 ttgtgatata tattacttta attattatta taacttaacat ttttatttac tttttgtgtt
110821 agtattttat tctgtcttct ccttagatag taaccttctt aagaaaatat atatgctaag
110881 tgttttactg gtttaatatg cttagactac tcatctacct caatacttcc ttggagatct
110941 cctcctcagt cacacagagc tcaggactta tatttccttg gaactcctgt tagggctcaa
111001 tgtacatgaa attccctaga cagacagaca gtcagttata tggcttgatt tcaaagtttc
111061 aaaatgattt aatggactat caagtagttt attaggagaa cagttattat actcttctaa
111121 aaataaagac ttttaagcaat aaagatgtat atgtatataa aatggctggg ttattcctag
111181 aagtaccttt cttagaattt agttaaattt aatatccaag atactatctt ttcaacctg
111241 agattgtgaa aagtaacttc tatcaatata aactttacta catttgtatt gtgttagtgt
111301 gttacagtat aatctagaac aatgtgtctt tctatatgat atatgacatt ttaatgccta
111361 aaaaaactga tatgtcttag atgattctag tcaggattta cttctagaat agattaaaat
111421 tctatttgag gagagtcaaa ttaattatcg aattctcagt tgttattatt gctgttttat
111481 ttttagtgaa acagattagt cttaatgtaa acacttgaga aataaattga tggccaacct
111541 aaaatgtaaa aaagaaatta atagaaaatt taaagagcaa caaagctctg acatttaaaa
111601 gaaatgaagt acaaatctct agggacctta aagatcatct aataatttcc tcattttcta
111661 gataaataaa ctgagagacc ccgaggataa atgatttgct caaagtcaaa tatctactta
111721 atataggaaa ttttaatttc ttctcagctc gttaacatgc aacttttcaa tatagcatgt
111781 tatttcatgc tatcagaatt cacaaggtac caatttaatt actacagagt acttatagaa
111841 tcatttaaaa tataataaaa ttgtatgata gagattatat gcaataaaac attaacaaaa
111901 tgctaaaata cgagacatat tgcaataaag tatttataaa attgatattt atatgttttt
111961 atatcttaaa gctgtgtctg taaactgatg gctaacaaaa ctaggatttt ggtcacttct
112021 aaaatggaac atttaagaa agctgacaaa atattaattt tgcataagg tagcagctat
112081 ttttatggga cattttcaga actccaaaat ctacagccag actttagctc aaaactcatg
112141 ggatgtgatt ctttcogacca atttagtgca gaaagaagaa attcaatcct aactgagacc
112201 ttacaccgtt tctcattaga aggagatgct cctgtctcct ggacagaaac aaaaaacaa
112261 tcttttaaac agactggaga gtttggggaa aaaaggaaga attctattct caatccaatc
112321 aactctatac gaaaattttc cattgtgcaa aagactcctt taaaaatgaa tggcatcgaa

FIG. 1 (cont.)

112381 gaggattctg atgagccttt agagagaagg ctgtccttag taccagattc tgagcagggg
112441 gaggcgatac tgcctcgcat cagcgtgata agcactggcc ccacgcttca ggcacgaagg
112501 aggcagtcctg tcctgaacct gatgacacac tcagttaacc aaggtcagaa cattcaccca
112561 aagacaacag catccacacg aaaagtgtca ctggcccccctc aggcaaactt gactgaactg
112621 gatatatatt caagaagggt atctcaagaa actggccttg aaataagtga agaaattaac
112681 gaagaagact taaaggtagg tatacatcgc ttgggggtat ttcaccccac agaattgcaat
112741 tgagtagaat gcaatatgta gcatgtaaca aaatttacta aaatcatagg attaggataa
112801 ggtgtatctt aaaactcaga aagtatgaag ttcattaatt atacaagcaa cgttaaaatg
112861 taaaataaca aatgatttct ttttgcaatg gacatatctc ttcccataaa atgggaaagg
112921 atttagtttt tggctcctca ctaagccagt gataactgtg actataagtt agaaagcatt
112981 tgctttatta ccatcttgaa cctctgtggt gaagagggtg agtataaata actgtataaa
113041 taaatagtag ctttcattat ttatagctcg caaaataatc tgtatggaag tagcatatat
113101 aaggatatata aacatttagc ctcttgatag gactaactca cattctgggt tgtatatcag
113161 tcttgectga atttagctag tgtgggcttt tttttatctt gtgagtttgc tttatacatt
113221 gggtttctga aaagatttct tttagagaat gtatataagc ttaacatgta ctagtgccaa
113281 tottcagaca gaaattttgt tctattaggt tttaagaata aaagcatttt atttttaaaa
113341 caggaaataa tataaaaagg agagtttttg ttgttttagt agaaaactta atgccttgga
113401 tgaaatgagc catgggcagg gttgtaatga attgatatgt ttaatagtat agatcatttg
113461 tgaataatat gacctttgac aagacacaag ccattaacat ctgtaggcag aagtttcctt
113521 ctttgtaaaa tgaggggaata aaatagatcc ctaaagtgtg taatttttagt atttctaaac
113581 tttatgaagg tttcctaaat gataattcat ctatatagtg tttttttgtg tgtttgtttg
113641 tttgtttgtt tgagatggag tctcgctctg tcacctaggc tggagtgcaa tggtgcaacc
113701 tcggctcact gcaacctctg cctcctgggt tcaagctaat ctcctgcctc agcctcctga
113761 gtagctgaga ttacaggcat gcaccaccat gccagagctaa tttttgtatt tttagtagag
113821 aaggggtttc atcatgttga ccaggctggt ctggaactcc tgacctgtg atcccccac
113881 ctgagcctcc caaagtgtcg gtattacagg cgtgtgccac cacgtccagc ctgagccact
113941 gcgcccagcc catctatata gtttaatatc aatctaaatg aatttctcag tctgagcct
114001 aaaaatttag ttgtaaagaa tgatatcctt gactaataat agtttctatt aatggattgc
114061 atctagtgtc aggtggcata tatttagtcc ccacaactac cctggaagggt atttaaaatt
114121 tttcacattt gcagataagg aaactaaagt tcagagttcg gcaacatgct tgaattcaag
114181 cagctcctag gatgttaatg gtggagggtg ggttcaaate cagatctgtc tgactcaaaa
114241 aatgcatact cctaaccagt gcactatata ccaattccat aggagccctt ctttgtgatt
114301 catagcactt tcccatgagt tttgttgatt ttgtgagaaa caaaactcct tttcctttgg
114361 actgtctgga atctctcttt ttc aaatttt tgaaatgtat ttctatgcca aaagacaaag
114421 atttctagag gaatatgcct aggatgagaa ttatgtaatt taaatcacag ctggaaagag
114481 agaaagtcct aagttactaa gaaatgttca aacacaaatg agctttcagt ctattggaag
114541 acctttatag ctagaagtat actgaactgt acttgtccat ggacccctga agaaacagggt
114601 taaatcaaag agagttctgg gaaacttcat ttagatggta tcattcattt gataaaagggt
114661 atgccactgt taagccttta atggtaaaat tgtccaataa taatacagtt atataatcag
114721 tgatacattt ttagaatttt gaaaaattac gatgtttctc atttttaata agctgtgttt
114781 gctccagtag acattattct ggctatagaa tgacatcata catggcattt ataattgattt
114841 atatttgttt aaatacactt agattcaagt aatactatct ttttattttc atatattaaa
114901 aataaaacca caatggtggc atgaaactgt actgtcttat tgtaatagcc ataattcttt
114961 tattcaggag tgcttttttg atgatatgga gagcatacca gcagtgacta catggaacac
115021 ataccttcga tatattactg tccacaagag ctttaatttt gtgctaattt ggtgcttagt
115081 aatttttctg gcagaggtaa gaatgttcta ttgtaaagta ttaactggatt taaagttaaa
115141 ttaagatagt ttggggatgt atacatatat atgcacacac ataaatatgt atatatacac
115201 atgtatacat gtataagtat gcatatatac acacatatat cactatatgt atatatgtat
115261 atattacata tatttgtgat tttacagtat ataattggat agattcatat agttcttagc
115321 ttctgaaaaa tcaacaagta gaaccactac tgatatttta ttatttcata ttacatataa
115381 aatataattt aatacaaata taagaagagt ttttaataga tttttaataa taaaggttaa
115441 gagattcgaa agctcaaagt agaaggcttt tatttggtt gaaattaaac aattagaatc
115501 actgttgata ttttattatt tcatattaca tataaaatat atttaaatat aaagataaga
115561 gtttttaata gattttataa taaatgttaa gagattaaaa aactgaaaat agaaggcttt
115621 tatttggatt gaaattaaag gccaggcatg gtggttcatg cctgtaatcc cagaatttta
115681 ggagactgag tggggaggat tgcttgagcc caggggtcaa gaccagcctg ggcaacacag
115741 tgagacaccg tatctacaaa ataattaaaa aattagctgg gcatgggtgt gtgtgcctgt

FIG. 1 (cont.)

115801 atgctacccat taactaagga ggctgaggtg ggagaatcgc ttgagcctgg gaggtcaagg
115861 ctgccctgaa ctgtgattgt gccattgcat tccagcctgg gtgccagaga gagaccctat
115921 ctctaaataa ataaataagt aaataaataa acagcaacaa caaaaacact caaagcaa
115981 ctgtactaaa ttttgaattc attctgagag gtgacagcat gctggcagtc ctggcagccc
116041 tcgctcactc tcagggcctc cttgaccttg acgcccactc tggctgtgcg tgaggagccc
116101 ttcagccctc ccctgcactg tgggagcccc tttctgggct ggccaaggcc agagccggct
116161 ccctcagctt gcggggaggt gtggaggag aggcgctggg ggaactgggg ctgcgggtgc
116221 cttgtgggcc agcgcgagtt ctgggtgggt gtgggctggg caggccccgc actcggagca
116281 gccggccggc ccgcgcagcc ccaggcagtg aggggcttag cacctgggcc agcagctgct
116341 gtactcgatt tctcactggg ccttagctgc ctccctgcgg ggcagggctc gggacctgca
116401 gcctgccatg cctgagcctc cccccaacct gcgctgcag tgggctcctg cgtggcccaa
116461 gcctcctgac gagcaccgcc cctgctcca cggcaccag tcccatagac cggccaaggg
116521 ctgaggagtg tgggtgcagg gcgcagggtc ggcaggcagc tccacctgca gcccagtg
116581 gggatccact ggggtgaagcc agctgggctt ctgagtctgg tggggacttg gaggatcttt
116641 atgtctagct aagggtattgt aaatacacca atcagcactc tgtatctagc tcaaggtttg
116701 taaacacacc aatcagcacc ctgtgtctag ctcagggttt gtgaatgcac caatcagcac
116761 tctgtatcta gttaatctgg tggagacttg gagaaccttt atgtctagct aagggtattgt
116821 aaatatacca atgtgcactc tgtatctagc tcaaggtttg taaatacacc aatcagcact
116881 ctctgtctag ctcagggttt gtaaatacac caatggacac tttgtatcta gctaacttag
116941 tgaggagggt gagaactttt gtgtctagct cagggtattgt aaacgcacca atcagcccc
117001 tgtcaaaacg gaccaatcag ctctctgtaa aaccaatctg ctgtctgtaa aatggacc
117061 tcagcaggat gtgggtgggg ccagataaga gaataaaagc aggcctgctg agccagaagt
117121 ggcaacctgc tggggctctgt agaagctttg ttcttttgtt ctttgcaata aattttgcta
117181 ctgctcactt tttgggtccg cattgcgttt atgagctgtg acactcactg ggaaggctct
117241 cagcttcact cctgaagcca gcgagatcac gaaccacca gaagaaagaa actcctaaca
117301 catccgaaca tcagaaggaa caaactcagg acacgcggcc tttagaactc ataactca
117361 ctgcaaggtt ccttggcttc attctcgaag tcagtgcagc caagaacca ccaattccgg
117421 acacaatttg actgcagaaa atggatgtcc aacctgtgg tttccctggg ccacattgga
117481 agaagaaagg agttgtcttg ggccacacat aaaatacact tactatagca gatgagctaa
117541 agaaaagaaa aaagtccatg cgtaatcttt gtgatatgtg ccaccacca taagcaaat
117601 tgttctctta ttcaaaagg tggacacagc tgctctagat attttattat taaatagca
117661 ggcaattact gtttaaataa agatttcctc acagaatgag attaaaagta tatattagtg
117721 gcttagcatt cattttagac aaccttttta gagattcaaa tcacacactt gcttacagaa
117781 attttgttgt cttcaatgtc cccttgtgg tttctttacc aagcctctac tgttcttcac
117841 atcaccaagt taaaaaaaa aaaggggcgg gggggcagaa tgaaaattgc atggtaggcc
117901 acaagttcag atcctcatcg acacaagagg tgccctgaagc agtggatgag gcttttctat
117961 ggatcatgag cagccacata aatgcttaa agggcctggc agggagcatc agtgggtgat
118021 gtggctggga ggctgaatgg agagcatttg ttcttcagtt atctatagaa ggcagctgtc
118081 actcagcacc agctaagggc tcccatgag ggaactgggg atcagggttt ccagatcttt
118141 ttatgtaaca ggataagaca gagatccagc tttttttggg taattatttc ctattttaaa
118201 atacgggtag ttgattaaat aaaaacaaac gaatgaacac catatgggca caacaaaaca
118261 catctgtggc ttggattcag cttgtgaatg attactgcag atatttattc tagaggacac
118321 ccctgggtat gtccataatat aaaacctaaa tctaaactca agtcccatgc taccttcaga
118381 gaataaatga ccagaaaaa gaaccacctc tcctaaggaa gtataaatth gtataaact
118441 gagacccaaa cttacaactg tacatttttc ttattgttgg gctgttgcta acctcaatta
118501 agaaggcttg atgatatttg taaagtgtca tcactccacc atggtccagt aacatctgat
118561 cactccacca tgggtccagta acatctgaat ggtcaagaaa tatctaaacg tatgtacca
118621 aaatttgtgt atactactgt accaataaac catttgtttc catttgatct ctgagtgtgg
118681 taatacatgt tatttgccct gctgttgtaa ataaacaaac caaatggagg cttgatgcaa
118741 gatgcagtgt agcatagtgc caactctgga ctccgactac tcagggtgta aattctaact
118801 ctgttctatt aacaccatga aactgagcaa gttagttaaa actcgtggg cccattttct
118861 catttataca atggagattt taatagtaca gctacatagg ccattttgtg gtttaaaata
118921 catcatgatt atgaaacact taatgtaggg cttgctacat aatgagcaag gtttgttgct
118981 gttatcatta atatccttaa ttctcattat tataaaactt gagatagtat gaggtgaaca
119041 agttcataac agcaatataa tgaaaatttt aataattcct tttatacttt aacaaaaata
119101 cgagattggg taatttatta tttttacatg agtaataaat attgcattaa aatatattta
119161 aaatttacca cattaatgtc tgccagtcac gccaaatgac caacatgaat gtgaataaaa

FIG. 1 (cont.)

119221 ctcagttotgt gccattttaa tottaaccaa ccotttataa ttgttaatga tttgaacctc
119281 tgccttgaaa gatcacatta cttgattgtc ttcaacttat ctgaatgtgg tagtgatttc
119341 tgtaaaattta taggaccttt gtctcatgca gctccatgga gttgaactta tgcaccttta
119401 aaatggtata tacttaatta attaagtggt gatctgcttc acatgtgtat aatattatta
119461 gctcactaaa ccaagaaaac agtggtcctt tagggaaaga aactaaatta caacagagaa
119521 tataaatacc atataaatat ctattattta ttgaactgtc acaattattg caaaaaatta
119581 ccttttagtg gacaaaacaa ttgatattgc ccttttctgg aaaagaaata atgtaatata
119641 tgatgaatag ttttggccag tatcctctag accttgccag ttaactggct ctcaaaattt
119701 tgaataataa aaacttggtg atagtagaaa aatagtaatt ttttaaaagt atgtgcacaa
119761 ttatacaact aaacaattca ttcaccagtg ttcacaattc tattgccttc tttgaatcaa
119821 aatttacata gtttttcttt tagactaagc tcctttatga taccagtgtg cccattttct
119881 attaccattg aaatgtotca tgagcatgtc acattctggt acaactgcta atccaggatg
119941 acagtttagt tcttttaaat ccaattgaga gccttctact catgaccaga gaacctaaag
120001 aaaggttaag atacatttat tccttggtgt aagtgatttg tctattttta gttttcctaa
120061 gggtcataatt tcaatttaga ttttttttta taggttaggt aaaataggct tcccttttgc
120121 aatatgaaat atgtagtctt ttaaaaaatt tcttcaaagc tattaaactg aaaaaaatt
120181 aatttggtct attcagtttg ttgacctta ccattttgga aagagagtga ctctactttt
120241 gtatttggtta acattttccc tactacaggg cagtatcttt tgtaagttct tagatattag
120301 caccaaataa ataggcaaaa aaaatctatt atgttaattc ttagaacccc tgcttggcag
120361 tgcatcattg actagatgga gaagaaatga aaataataca ttaggaagca gtttctggt
120421 tcttttgaaa acaactagag agtcttggtg ttgactggaa tatctgaaga tcoctgtttaa
120481 tgctttcatt ctatgattgt taagaatatg tcatagaact gctgtatcct gtttctttat
120541 gtcttccctt ctgtttggtg attagaaatc cctgagtggc tttacattat tagtacagta
120601 gatatgtagt atattcccat aataccactg ctgctattga ctaatagtaa taattttagg
120661 gcagctttat gacagttggt ttatgtttta ggggtgcatt tgacttgtga agcattgaaa
120721 tctgggtatt aagcacactg ttttctatgt ggtatggaat gattcttaaa gccctgagaa
120781 aatggaaaat aaaaatattt ttcttttta ccataatcac ctatgactgt cactctatca
120841 taaactgcat aaactttata acctcaaac attttggaat tgaaatgaca gaacttgctt
120901 actcaattgc ttctatatac accaaatatt tttttaaaagt attatgttaa gtccttgaaa
120961 atattttggt ctactcaata gaagcagttt aggttggtag ttctatgtgg aaaccgtgag
121021 gaaataattt tatattatga tgactagacc agtctttgaa catcactttg gttattgttc
121081 cattagtaaa tattataatt atttctgaga tttactcacc ttcaaagaat gttggcaatg
121141 ccagcattat taacactcct ctagttagaa caaagaggaa atgtaataac aaaacataat
121201 aatagccaaa taaagagtga cttagaatgt acacccttat ctaggatcct gagtaattcg
121261 attattctta ggaaatacac ttttgtgcta gaacaaagac ttttgaaata gctaatttct
121321 gggtttcttt tcattttgaa ttaacttgaa tttcaaggaa acaagggtag tttttacaga
121381 tacagtgcac agaagctctg tgtacaatga agaaaagtag gaaagtgaga aaaaatgccat
121441 tagatttttc atcgttatac tatctgatat gtgaatttaa ctaaaactta tatacctcat
121501 tatagtactt cctaattgaa tttcttaatt taagtgttcc ccataagggt tttttttata
121561 taaacttaag tactgttaaa tatttaaggc aaattcaggt ataaaaataag acttgttgat
121621 atcttattcc aagcatattt gtttctctcc tatttatttt tattctgtgt tcatttccaa
121681 aattgtttta ctcacaactg tttgtttttt ctgttttcatt ctgtggtaaa ggtatcattt
121741 ggctaattgt ataatttcag tgtcatttct aatattccaa ttgtgatagt atcaacacaa
121801 gattaaattt ctctacatgg tttatgagaa tggaatgcca aattgaaata gaacagagca
121861 cagatgatct aaatataaaa agaactacaa aaatcacagt tgtttaaaaa ggttttttgt
121921 ttgtttatat atggtgcaga acatttggtc cttagccaaa tgtttccacc ttgagaaagc
121981 tatagagatt ctatgtagtc ctagtaccaa taatatgttt taacctgaat gtacctatc
122041 tttattcata aactgtgact ttttacctg ctgaaacttt tttttttaag acaatctcac
122101 tctgtcgtcc agtctggagt gcagcagtgg tgtgatcttg gctcactgca acctctacct
122161 tctgtgttca agcaattctg gtgcctcggc cacctgagta gttgggatca caggtgtaca
122221 ccaccaggcc tggctaatag tttttgatat ttctagtaga gatgagtttt gccacattgg
122281 ccaggctggc ctgaaaactc tggcctcaag tgatctgcct gccttggcct cccaaagtgt
122341 tggatttaca agtgtgagcc actgtgcctg gcctgaaact cataattcat ttccattaat
122401 attaattcca ccttttccaa taattaattg atttcacaag tattagtccc ctataatcat
122461 tgaatggcta ataaaattat ttatagcaaa cagattaatt atctgccagc agtctgagat
122521 tagtttcttt aaaaaatgtt tattatttaa aacattcagc tgtgatcttg gctttcttgt
122581 gaggttcaat agtttctatt gagtaaagga gagaaatggc agagaattta cttcagtga

FIG. 1 (cont.)

37/63

122641 atttgaattc cattaactta atgtggtctc atcacaaata atagtaactta gaacacctag
122701 tacagctgct ggacccagga acacaaagca aaggaagatg aaattgtgtg taccttgata
122761 ttggtacaca catcaaattg tgtgatgtga atttagatgt gggcatggga ggaataggtg
122821 aagatgttag aaaaaaaatc aactgtgtct tgttccattc caggtggctg cttctttggt
122881 tgtgctgtgg ctcccttgaa agtgagtatt ccattgtcta ttgtgtagat tgtgttttat
122941 ttctgttgat taaatattgt aatccactat gtttgtatgt attgtaatcc actttgtttc
123001 atttctccca agcattatgg tagtggaag ataaggtttt ttgtttaaat gatgaccatt
123061 agttgggtga ggtgacacat tcctgtagtc ctagctcttc cacaggctga cgcaggagga
123121 tcacttgagc ccaggagttc agggctgtag tgttgatca ttgtgagtag ccaccgact
123181 ccagcctgga caatatagtg agatcctata tctaaaataa aataaaataa aatgaataaa
123241 ttgtgagcat gtgcagctcc tgcagtttct aaagaatata gttctgttca gtttctgtga
123301 aacacaataa aaatatttga aataacatta catatttagg gttttcttca aattttttaa
123361 tttataaaag aacaactcaa tctctatcaa tagtgagaaa acatatctat tttcttgcaa
123421 taatagtatg attttgaggt taagggtgca tgccttcta atgcaaaata ttgtatttat
123481 ttagactcaa gtttagttcc atttacctgt attggaaatt cagtaagtaa ctttggtgc
123541 caaataacga tttcctattt gctttacagc actcctcttc aagacaaagg gaatagtact
123601 catagtagaa ataacagcta tgcagtgatt atcaccagca ccagttcgtt ttatgtgttt
123661 tacatttacg tgggagtagc cgacactttg cttgctatgg gattcttcag aggtctacca
123721 ctggtgcata ctctaatac agtgtcgaaa attttacacc acaaaatggt acattctgtt
123781 cttcaagcac ctatgtcaac cctcaacacg ttgaaagcag gtactttact aggtctaaga
123841 aatgaaactg ctgatccacc atcaataggg cctgtggttt tgttggtttt ctaatggcag
123901 tgcctggcttt tgcacagagg catgtgcctt ttgttgaacc tccatttgac tggcatgcac
123961 atgtctcaga tattataggt tatcatatat tgttgctcct aatatttctg tgttagataa
124021 ttagagtagc ttggtttgta agaattgtat gttggtggga ctgtagcaga acaagaaggc
124081 ccttatgggt cagtcatacc tctcttttca aatatttggg ctagctctct tctgggcctc
124141 ttgttgccaa tatatagtat tgcctcaaaag ggcaggagat ttgaagtgat caaggaaaat
124201 atattttttc tattgattaa gtcttttgat ggggtagaat aatctaattt catgtaactg
124261 ctcaaagtta tatggtaggg ggatcccaaa tgtattttaa aactattttt atatcatcat
124321 atttgaagta atagaaagtc agagtagcag aataaaggta ctaaaaattt taaaaactaa
124381 taaggtaact tgaaagaaat caattatggt gattcctcat taaacaaatt tgcacttaaa
124441 gactgagggt aataaggatt tccccaggt ttttcatagc aacctgtgag cactttctct
124501 gttgaggcat ttatggtatg aaaagatgag taaggcacag ttcttgccct ggagaaggct
124561 acaggtgaga ggaggagtgt acacagaaac atttgatata aagcaaggaa taaattccaa
124621 gactaaaatt ttcagaaatc taaaaaactc aagataagaa aaaccatta tattttctgg
124681 gtaacaaaat ttcagtgtta ttaacatgta ggaagatctt gatattttatt ctgaagccca
124741 tgtgtgttgc tgaaatattg ccgcatttgc atatactcat caccatcctc tgttttggag
124801 ctaagaattt tagactcaag atgtctaatt aagttgatcc attgatttta ttttttatgg
124861 aaatctgaga cccacagaag gcaggggatt tgccacatt tctagaagag tcagacatga
124921 gcgatgaggc acagtggaaa gaacatgagc attgcctgag ctctgagttg gcgctataag
124981 agcagtgate atgggcaagt gactcttctg agccttggcc tctcaccctg ttaagtgaag
125041 aaaagaatat ttcagaagat ctttgtgaga atgaaacaag gcaatttact tgccctgctac
125101 atagccaatg ggaaatcaat ataagttccc cgtggttccc ttctgtgggg ttttgttccc
125161 acagaggggt cactggccat tccacttctt cttttccaag ctctcattc cttttaacgc
125221 tgttcatagt tggttccaaa ccatttgaaa tataataagc accaggatgg ttttttcttt
125281 ccaccaaagc aaatttcatt ttctaaacac tgtttataaa tatcaatggc tattttttca
125341 atttttgatt atcatgaaaa tatacaataa tgtttaatta aatatgctaa agaattgatt
125401 aataaatatg tattaataaa ttctacata taaggccttt ttgcttgggg tatgggtgat
125461 acaaaataaa tgtggcatga acccactgac ctctagcaat ttataaccta gaaaaagagt
125521 tatgatatgt ttataagttc ctgtgatata agacatgcat atagtcatta taacagaggt
125581 gcaacaaga tgtatcaagt atgtccagag gaggaagaga ttaatcccag ctggaggaaa
125641 cactgatgct ttcttgagc aggggcattt gagttgagaa agggaggaaa catagatttt
125701 gacaatgaga gctgagggga aaggggtttc aggtggaggg aaccgcatgt ggaaagcagg
125761 gaggtaggaa agtgtagagt gtgtttaaag aatagaccag tttggctgaa acaggatatt
125821 tgagcagagg aagcttgtac taggtagggt ggttgaggcc aaattatgca aggcattaaa
125881 tattaaacta ggaattttgg actttatcct gcagtttatg ggggttaa atgataagattc
125941 aatatcactt tatttgtaca gtattatggt acattttatc taattgtttg ttttaattcct
126001 gtctagacaa tgaattcctc aagggcaagg agcatggctt attcacctca gtaatttcag

FIG. 1 (cont.)

WO 2011/119531

PCT/US2011/029314

38/63

126061 tgcctagcat tgtgcctggg acaaagtggg caottgtata taaccttttt taattgaagc
126121 aacaagttgt caaccttaca aatgtgaatc cgtgattcag atgacagggt gaaatgtaga
126181 ttgtctgcga agagggcgaga aagagagtat gacaaaggag gacaagacag tggggcaggc
126241 agggagagag agcagccagg gtctcggtag aggtatgtca aaaaggatat gaagtcagag
126301 gagaaggaga cccctatggt atagaataca aatggaaggg aaatgatgac aacagtaagt
126361 tgtcattaaa tgcaaggttg caaaagtaag attgtaaaagc aggatgagta cccacctatt
126421 cctgacataa tttatagtaa aagctatttc agagaaattg gtcgttactt gaatcttaca
126481 agaactctgaa acttttataaa aggtttataaa gtaaaagaca ataacttgaa cacataatta
126541 tttagaatgt ttggaaagaa aaaaaatttt ctaagtctat ctgattctat ttgctaattc
126601 ttattttgggt tctgaatgcg tctactgtga tccaaactta gtattgaata tattgatata
126661 tctttataaaa attagtgttt tttgaggaat ttgtcatctt gtatattata ggtgggattc
126721 ttaatagatt ctccaaagat atagcaattt tggatgacct tctgcctctt accatatttg
126781 acttcatcca ggtatgtaaa aataagtacc gttaagtatg tctgtattat taaaaaaca
126841 ataacaaaag caaatgtgat tttgttttca ttttttattt gattgagggt tgaagtcctg
126901 tctattgcat taattttgta attatccaaa gccttcaaaa tagacataag tttagttaat
126961 tcaataataa gtcagaactg cttacctggc ccaaacctga ggcaatccca catttagatg
127021 taatagctgt ctacttggga gtgatttgag aggcacaaaag gaccatcttt cccaaaatca
127081 ctggccacaa agtgtgacat tttggcattg gcactactat ttgatggaag ccaacctccc
127141 cccaaaaggc ctgtattaga atgaagatgg attccctggg tgggttacac ttgaaactag
127201 cctcacccat gaacactttg gcacagatta gctagcccat tccccacag taaggacct
127261 aaggaagggg cagaagcaaa gataagtttt agaacaaaag agaggggaaa gaaaaaatct
127321 agggttttat gagggctgtc cctgagtgat agatgtgaat aggcctccag ggcaggctgg
127381 ctgagaggct gactctttgg gttggggtga ctgattggtg gtgaggatgg agaagaaaag
127441 gggagtggag gaggtgaaag tgacctggg acattagggt ccataagtg acaggattta
127501 aggagtgttg taagctgtgg ttgttgacc aggtttaagc acagcttctt gagcttctg
127561 actggtttag gtcaagctcc agagagcaaa tgccacagtc tcagtgatct cctggagaa
127621 acagttggaa taggatgttg cccatgttgg gatgagtcat tgtccgctct tgctctttcc
127681 ctacctctgc aaaataataa tactgtattt gattgaacat ataaaacaaa agaaggatta
127741 tcacataagt atgtatatat aaccaacatt ggcagggtgc gaaaaaccag actgtcagtt
127801 tgcctcatct gaaatgattg acacaaacaa atatatattac tgtcccaagt gaactttggc
127861 attttggata tcttctagtt gttctgttta aagatataac ttagaagcag ctgatggaat
127921 atttaaattc atgcgttgaa ttcatgcatt caaagaaaca tgtcctgagt cactaaatgc
127981 tgacatttgt ttttcatgtt aagagtgtaa ataactggtc ccaaataata tattattaca
128041 tcagataaaa actggaatgt gaacctctta acttgattgt gaaagtattt gccaatggtg
128101 cctcttgata attatttgag gctcacttca gaactctctt ggaagggtta atttttaaat
128161 agtcattttta taaattaaca tttttgacat atgtgatggc tctcaaattt tttcttttat
128221 gccagtttga atcattttctg ctcaattttt ttttttaatt gggatggagt ctactctgt
128281 tgcccaggct ggagtgcagt gatgcaatct tggctgactg caacctccac ctctcggtt
128341 caagcgattc tctcgcatca gcctccagag tagctgggat tacaggcgcg caccaccatg
128401 cctggataat ttttgtatta ttactagaga tggggtttca ccacgttggc caggctggtc
128461 ttgaactcct gaactcctga cctcaagtga tccacctgcc tcagcctctt aaagagctgg
128521 aattataggt gtgagccact gcaccaggcc ctgttcaact tttaatgcta agattcattt
128581 gttgttgttt cacaagtgat taggcagagg tcttttataat taattttacc attttatttg
128641 taagagagtc tcatattaag gaagcataat atatgacaat ccaaatacag taaaaatttg
128701 gttaattttg attttgttaa ataattaatc acaggggtcc ttcaaattgt gagctcctct
128761 ggttatactt atgtttttacc tctgggtata ctaattttca aacaaatgaa atttcattct
128821 attcatgata tttcagaagc agatctgttg cacaaaataa agcataccta taaattttct
128881 ttttttataaa aaaagtctct gtctactcta ttttctatta ttttctctt tttaaaattt
128941 gaattttatt gtggcaagtc cacttaacat gagatttacc ctcttaacag atttttatgt
129001 gtaaaataca atattgttca ccatgggtaa atgttgcaca gcagatctct ggaacttatt
129061 cattttgcac tactgaaatt ttatacctgt tgattagtat ctccccattt cctctctccc
129121 cctgtcctgt tacctatggt tctgttcttt gcttctttga gtttgagtat tttgatacct
129181 catgtaatct tcattctatt ttctaacttt gacaatgttc tgacaaaatt gctttccgga
129241 ttggagcact gtatagttaa aattgaaaat cttgggttatt ttctacagat tccactatt
129301 ttaccttgag cagacactta tcttgaaagg tctcagattt gtcacttgta gaatggggaa
129361 tataaacctg ataatgggtc ctttcagttc taaagttata tcagttgaaa atacatgtgt
129421 cacttatggg aacgggtaga gaactggctc actgaacagc atatggatat tataaagtgg

FIG. 1 (cont.)

39/63

129481 ttttttttaa tcctttctgc agacagttac tttatacttt attcaaatgg attattgtga
129541 agtacatggt agcggacttt gtacctttta aaaatgtatg tatttggtgt aatgtagaaa
129601 tatagaaatt tattaagtat gatttatttc aatgttaagc atgagaaaat atgctccgaa
129661 aggttagata gcttgccata atgacaagct tgtattttcaa gcagaacttt ctgaatcaaa
129721 agactccaag acgaatgcc agctttcaaa aactgtctaa ccaaaataaa tcctaagatt
129781 caccttcata ctaaaattat ttaaaaatag tttattttta attaatattc acttaaaatg
129841 tatttatcat gcaatacttt aaagtgtctg ggaaatgaaa atatccaaag atcaaagaac
129901 accatgtttt caaacttcaa aaatgttatc agtgacctaa acaattttta aaattttcat
129961 agagcctatg aaaaatgtac ttgcaaatgg ctactttctg actaggaata gaatggggag
130021 agtatttagt ccaacaatga tagactggat taagaaaatg tggcacatat acaccatgga
130081 acactatgca gccataaaaa atgatgagtt catgtccttt gtagggacat ggatgaaatt
130141 ggaaaacatc attotcagta aactatcgca agaacaaaaa accaaacacc gcatattctc
130201 actcataggt gggaattgaa caatgagatc acatggacac aggaagggga atatcacact
130261 ctggggactg ttgtggggtg gggggagggg ggagggatag cactgggaga tataccta
130321 gctagatgac gagttagtgg gtgcagtgc ccagcatggc acatgtatac atatgtaact
130381 aacctgcaca atgtgcacat gtaccctaaa acttaaagta taataaaaaa aataaaaaa
130441 agtttgaggt gtttaaagta tgcaaaaaa aaaaaagaaa taaatcactg acacactttg
130501 tccactttgc aatgtgaaa tgtttactca ccaacatggt ttctttgatc ttacagtgtg
130561 tattaattgt gattggagct atagcagttg tcgcagtttt acaacctac atctttgttg
130621 caacagtgcc agtgatagtg gcttttatta tgttgagagc atatttcctc caaacctcac
130681 agcaactcaa acaactggaa tctgaaggta tgacagtga tgtgcgatac tcactttgta
130741 aaaaagctat aagagctatt tgagattctt tattgttaat ctacttaaaa aaaattctgc
130801 ttttaaactt ttacatcata taacaataat ttttttctac atgcatgtgt atataaaagg
130861 aaactatatt acaaagtaca catggatttt ttttcttaat taatgaccat gtgacttcat
130921 tttggtttta aaataggtat atagaatctt accacagttg gtgtacagga cattcattta
130981 taataaactt atatcagtca aattaaacaa ggatagtgtc gctattacta aaggtttctc
131041 tgggttccca aatgatactt gaccaaattt gtccctttgg ctgtttgtct tcagacaccc
131101 tttcttcatg tgttgagct gccatttctg gtgcccccaa actctacttg agctgttagg
131161 gaatcacatt ttgcagtgc agccttagtg tgggtgcatt ttcaggcaat actttttcag
131221 tatatttctg ctttgtagat tattagctaa atcaagtcac ataaacttcc ttaatttaga
131281 tacttgaaaa aattgtctta aaagaaaatt ttttagtaa gaattaattt agaattagcc
131341 agaaaactcc cagtggtagc caagaaagag gaataaatat tgggtggtaat tttttaagtt
131401 cccatctctg gtagccaagt aaaaaaagag ggtaactcat taataaaata acaaatcata
131461 tctattcaaa gaatggcacc agtgtgaaaa aaagcttttt aaccaatgac atttgtgata
131521 tgattattct aatttagtct ttttcaggta caagatatta tgaaattaca ttttgtgttt
131581 atgttatttg caatgttttc tatggaaata tttcacaggc aggagtccaa ttttactca
131641 tcttgttaca agcttaaaag gactatggac acttcgtgcc ttoggacggc agccttactt
131701 tgaaactctg ttccacaaag ctctgaattt acatactgcc aactggttct tgtacctgtc
131761 aacactgcgc tggttccaaa tgagaataga aatgattttt gtcacttctc tcattgctgt
131821 taccttcatt tccattttta caacaggtac tatgaactca ttaactttag ctaagcattt
131881 aagtaaaaaa ttttcaatga ataaaatgct gcattctata gggttatcaat ttttgatata
131941 tttagagttt agtaattaac aaatttggtg gtttattatt gaacaagtga tttctttgaa
132001 tttccattgt tttattgtta aacaaataat ttcttgaaa tcggatatat atatatatat
132061 gtatatatat atatatatat atatatatat acatatatat atatagtatt atccctgttt
132121 tcacagtttt aaaaaccgat gcacacagat tgtcagatag caattctgtg attgaagggg
132181 aaatatgtca cctcttcata ctcatattgg tgaagggctc tagcttcaaa attaatagat
132241 tcctaaagag gggaaatgaa acatccgcat ttacacacac acacacacac acacacacag
132301 agttcctctt gtcggttaagt tttgtttttt ttaaactctc actagataaa atttgttatc
132361 taattgtgag ttttacacaa agaaaaactg tcacagaaaa gaaagacagt gtcacatttt
132421 tcaaaagaaa aagaagaaaa gaaagtgcc tgtttttcaa atacaaatgt tctggattga
132481 ttttaggata tttagtgaaa aacaaagtat ttcataataa gtaaaataaa aatctatgta
132541 ggtaaatttg tttctcta ataaagaatt gaatttctga gtatttatga taagtgttga
132601 aataacttct tatatgtgac agtgaatact ggcagagcaa atgccaaatc aatgocaaat
132661 ctgtaggata atttgattgt aggaacagaa ttctactcaa accgaaagca ggcatttgct
132721 ggagttacag aaaggcctca tggaacaccg agaaggtggg gccattcgac tcttaaagaa
132781 gctgcaacag gcacaagaga gtcagctgca gctcttcttc ttgagtctat atctgtcctg
132841 ggtccattcc tttttgtggt tgcttcattc ctttctctct ctgaagactg gtttttctgg

FIG. 1 (cont.)

40/63

132901 tctaccaggg ctatgccaca ttgactttat gtagtgtctc cattctggcc tcctgaattt
132961 acaggagagt tcctctgtac aaactcaaag tcctggagag aacagaaaao agcttccttt
133021 tggctcaggg gtccaactgc agtctactct gctgctatga ggatagtggg ttcaccacct
133081 ttgttggtct ctcagctagg gcagtgggaa atgactctat gaaaggaata tacatgggca
133141 ggcaaatgta ctaatcctca tcagtactgt aattttaagc aactttaaaa aattctttta
133201 agttatttga aaataagatc aaagaaggct gaattacata aatgaagatt tgttaacaat
133261 taattcaaac caatataaca catgctataa catgggtgag tgtgattgag tcttgattta
133321 ttaggggcaa taatcaaaac atttaacaat cattatagta cagaacttac caatcaaac
133381 agatgctcag ccggagtgga tgttggccac ccagctatta ttatccctgg ctcaattggg
133441 cttcagctgt gttaacttgc aaacattaat taactatcta agccctcat tttcctcaag
133501 tgtaaataga cacaataata ttacctatc cataggtgtg gggggaatag taaatgtaat
133561 aatttgtcca aaacacttag tatagtgcct ggtccatggg aaatactaaa taaatgttat
133621 ctgacttatt attaaaattt tatcttctca gottaacctt cagaacagta atatatggg
133681 gtctagataa atcttgctta tatgaaaata atttaatact acatgcagat atatgctgtg
133741 tatattatgc cttctgttag aggaattgca gaaacaaaaa tttcaattaa taataagatg
133801 aattatttct cccaattgta gaatcttttg acaattttat catgcattac agatgtaaga
133861 actcttgatt gggacttgat agtctaactt tataataatt taagaacatt cctcttagag
133921 aatttctatg gccataatac tgaacacatg aattttaatt agctgtcctc tttagcccta
133981 aaaaaaaaaa tactgtaatt taacacttaa gtgttggtct tcccaggtag agtaatcttt
134041 tttttttttt tttttttttt ttgcatagag ggtaatcttt tctctttcca aatggcgaaa
134101 ctgttagttt tctgactgtc cggtgaaatt ctaagtccac ttacttccca atagcatgca
134161 attagcaaag gtctctcttg caaaggcaca gaacacacct aaacatcttg cagatgctgt
134221 ttggacactc ttccctgct tttggctctc ttgtaaagca gctcatctgg atacaggatc
134281 tcttttcccc attgcccatt ctaatatatg ttaccgttat tacttataga ataatagtag
134341 aagagacaaa tatggtacct acccattacc aacaacacct ccaataccag taacattttt
134401 taaaaagggc aacactttcc taatattcaa tgcctctttg atttaaaatc ctgggtgaat
134461 acttactata tgcagagcat tattctatta gtagatgctg tgatgaactg agatttaaaa
134521 attgttaaaa ttagcataaa attgaaatgt aaatttaatg tgatatgtgc cctaggagaa
134581 gtgtgaataa agtgcgtcac agaagagaga aataacatga ggttcattta cgtctttgt
134641 gcatctatag gagaaggaga aggaagagtt ggtattatcc tgacttttag catgaatac
134701 atgagtacat tgcagtgggc tgtaaactcc agcatagatg tggatagctt ggtaagtctt
134761 atcatctttt taacttttat gaaaaaatt cagacaagta acaaagtatg agtaatagca
134821 tgaggaagaa ctatataaccg tatattgagc ttaagaaata aaacattaca gataaattga
134881 ggggtcactgt gtatctgtca ttaaatcctt atctcttctt tccttctcat agatagccac
134941 tatgaagatc taataactgca gtgagcatcc tttcacctgt ttccttatcc aggattttct
135001 aggagaaata cctagggggt gtattgctgg gtcataggat tcacccatgc ttaactgagt
135061 ggtgccaaat tgtcctcaag tctgtgtgac tgatatatat ccccatcaag agagtacaag
135121 aattctcata gctatgtatc ttcaacaaca cttgggtgtc ggtagatgtg aagtgattac
135181 taaaaatata ggggaagctgc atacataatt attggctttt gctgttctct tacattaatt
135241 tcttattcat gttgattact catttgtcac ctagtttttt cttccttaat taaattgtag
135301 gaatttatga attatggatt gatcatcagc tctatacatt tcaaacataa tccctcagtc
135361 agtggcttgg cttatagagt cttttgatga aaagaagctt ttaagttaa taaagttcaa
135421 tttattgtct tttcctttat gttttgtgct tttgggtatc tgattaagaa ctccctcctt
135481 atattggggt ctcaaattta gcagcataac attttcatac tattatttaa atttttttca
135541 cattatttag tgatagcacc tttcttatcc ctaaaagtgt tatcattgcc ttctgtcttt
135601 ctgcttgata aatattgcca cacatttgta tactttatta gtgtgtacaa agaccacatt
135661 ttagttgtgt tatttctctt gttttgggtt tctagaatgc agagccatta atattatagt
135721 aatgcttatg tgctaatacc atatcagggg cacaaatccc attgcagcgg gactgagaaa
135781 ttaaaggaaa tgatgcacat ttactcattt ttgtttaaaa aatcaaagtc atatttttca
135841 atcagactat atggttggtc tggatagctt catcattgaa tttttaaagt atttttgtac
135901 tactgtatth aaaattatcc attcaccact gottttgtag atggtttaga aaccaagtt
135961 aggaatgact gtgcaacact attattatac tcttttttaa attatacttt ttgcttaagt
136021 ttctttcctt gttctctgag acagtgttca tgttcccaaa ccacacacat ttattcagct
136081 ataaaatttg tatgatcaac tctgtcaga acaaacatca ttataaaaaa tatctccagg
136141 aaaaagaaaa cctttttaat gctctcttct ggttcattgt tcttttatt ttctttaagc
136201 attttcataa cccattgagc tgtaatttaa ttggaacatg atttatacta aagttgggtt
136261 ctttaccttt aaottttttt tttagtttga tcagctctct ttagcttctg tagttcggtc

FIG. 1 (cont.)

136321 tttaattcca ttccagtatg cttttggagt tgggtotcat aaatgtatag aaatgtttct
136381 gttgggaaac agcaggagaa tattaataa atattgtgct tacatctatt taattctttg
136441 cccaactttc tacaactttg actttacatt taagctctc atgcacttac atgtttcttt
136501 acctaataat atcttttcac catgggtgtg tacaattcct ttgtccttgc tgtattaatt
136561 ttcttggttt acatagtagc ctctacacat tgatgtcaaa acctctgttt ggtgcatttc
136621 tactctgcgt gttcaatctc catgaaagtt totgtaaggt attttcattc ctctagtttt
136681 tcacatgtgc atcctggctt tgtgacctgt gctttgatat cgtgcctttc atcttgtggc
136741 attgaaggat ctttgcaagg acctattgtg ttataataca gtctatgaaa aatatcaata
136801 tttgcatttg atcacattta aaaaaatcac attcttttgt ttgaatatca aagctaatat
136861 gtgagtgatt tccctgccaa atagcacaa tagcctttcc tgggtgttta tgggcattta
136921 totggttaat gattcccatc atagtgtgtg caccatgcc attgctaaac ttatacagta
136981 acttttttgt tttcacctca gcatatgttg agagtaggaa atagatagga ctatgccctc
137041 aaattttacg tttatatgat gttaatccta aaggctcctg tgacttctga agtaaaaact
137101 cagtgttgtc attttactta ctgaattgtt agctgagttt agagttgagt ttacaatgga
137161 gtaacaagg tgtttagttt gatgtatgct tttagtcttt cagaaaaaaa tgtttatact
137221 tggaaagaat agtttattta cccatctggc ctagttttaga caaaaacaca gagtcaaatg
137281 tcaacagaat tctgaagtta taaaaatgac agtgtggctt tttttttttt taaccttcca
137341 cctggtgctt atgoccaaagt gctagcttt ctttagctct caactaataa aggtaatggt
137401 tagataacat ttaacgttaa gttgcattgt gtttatgac acatatctca aatattggta
137461 cacgaaactg tacaacaacc ttttttatta gattttccta cgaaattcct tattatattc
137521 cctaagatag ctttttccca ccttctctct ccttctccct tctcaggtgc tccaataatt
137581 ccaacccctg cagccagtga ctttattata tottttttta aaaatctaaa aaaaaaatt
137641 gatgcaacca ggaagaattt tctcatttct ctccaccagt tgtaccagcc tactgcacct
137701 ctctcatgc accaccttct gctgtgttc ttgtctctat attcaggagc aagtaatatg
137761 caatacctcc ctctttgtgg gatctttctc attagcataa aaatactttc ccttgatctc
137821 cagctactac cccatttctt tgacctacat atagcaaaat atttgagaaa ggaccacttt
137881 ccacttttct ctcaatctac ttccattttt ttotcaatcc actttcattt cattgttctc
137941 ctcaacccat totttccaca acctacttca ttttatttcc atcagcccca taactcagga
138001 tcaacatctt gccagagcca atttccctgt ctcccttaac agctccagca gtatttatgc
138061 catggacaaa ttattottct tgtgatactt totctcttgc ttccatgaca ctactccac
138121 ttcattttct ttctacctct ctggctcttc cttggtccct tttcctggcc ccttctctct
138181 ttcagatctc taaacatcag ctatatctca gccctgttct actgacactc tctagctggt
138241 attttctaaa cccatgtttc agaaaccata tottgatgaa tcttggaagg ccgaggcagg
138301 cgaattactt gaggtcggga gtttgagacc agcctggcca acgtggtgaa accccatctc
138361 tcctaaaaat acaaaaatta cctggccgtg gtggcatgca cccagctact tgagaggctg
138421 aggcacaaga atcgcttgaa cctgggaggt ggaggtttca gtgagccgag atcctgccac
138481 tgcactccag cctgagcaat agaggagact ccgtctcaca cacacacaca cacacacaca
138541 cacaaagaaa ataaaccatc tcttgatgaa tcataaattt gtgtctctag tttagacctc
138601 tatcctgctc tctaaatgat gtatccaaat atcatcttga caccatcata tgttcataaa
138661 acataattat agaatatctt tcagtaggct tgacatttta aggcattgagt ttccgttcag
138721 tatctcctta aaatataccc aggttctcag gagactattc aaacaggaca aagcttctat
138781 tctacttact aatgtgtctg gccctatttg gcaggttgga taaaaagtca tctgaacatt
138841 gtcactttat gaataatata gtttaatagt ttgtgaatca cccctgcaat ttaaaaaata
138901 gtaaaattat cagaatctaa ttttaataatt cctattggaa caccocatgt taggggattt
138961 ccagttattt caattgatat ctcaatgttt taaagattgt ttatttctat tactaattca
139021 ctctttattt taacataaat tgtggctatc tatctctatt catttcaatt atatttctca
139081 taccattcta tagatggggg gaaaagaaaa gtgttaattt tttaaaactc catacctcaa
139141 atactatatg aatttatagt tgttattgct aaagcaatta tottacatct tttcctccaa
139201 aacaaagtta tgtgctgggt tattttcttt gtactcataa gatgccttcc atttttagta
139261 acataagtct tgtctttctc ctattottag ctacttaagc attatgtagc ttaataaagc
139321 actaaagatt cctatctgta tgaaaaaata aagattaaat aaataagatc tagaaagggg
139381 gacaagggtg tgcttcaaaa tgaaccatac caagccatct agcgattgat aaattactca
139441 cactcataat cacattgttg gaaagaagcc attgacaatt cagtttggtt cacaactgtc
139501 tatcacatag tgagcacaac taaaagacta ctttttgtct tttactgctt gttttgttga
139561 tcaagtgact gattgtacaa tgaccaacaa gaagtctgat gtgtagagaa aaggggaacc
139621 tggcttttct gccttactcc tgatgcctaa ttctgagcat gtgaatatta ttctgtttct
139681 ttaattctcc aagtgaagca gcagataaac catccttgtt tccattagct gtctacctgc

FIG. 1 (cont.)

42/63

139741 ttcaactgtg tgttttctaataacataagaa taagaaagcc accaggggtga gcaggggaagg
139801 caatgagtct gcaagggttg tggatagatt tctgttagtg aggcctctaga aagttcttcc
139861 aagattgatg caatctgaga agagttttct gtcaatacaa actccctggg tttctccttt
139921 gtcctttttac tgcctgtgtt tgttttgggt tccagtaaag atcaagtgac tgattgtacc
139981 atgaccaaca agaagcctga tgtgtggaga aaaggggaac ctggcttttc tgcattactc
140041 ctaatgccta attttcttgt actgaaagta gtttttgctg taagaatctg aggggaggag
140101 tcatttcttc aatttttttt tttggctctc ttttaatggt ttcttgatca tgtctatcct
140161 tatttttctg ttttcacaaa tttttgtggt atattttcct ctcatgacct ctgtctcaag
140221 acttctttcc atccatctct tctcatttca tctgttagag tgtctgtggt aagagccctg
140281 cattctactc tggccttgcc atgtgtggcc ttgggcaagt cctagcctcc ttgaggggtct
140341 tatttttctc atttgtaaaa tgaaacagtt tgatgagaag ttttctaagg ttcttcaag
140401 ctttgacaat ctctctcttc tggatctttt tcccatgaaa aatttcaact cttgattagc
140461 atgtaggcag ggattattcc acatccttat aggaatcaca tttctgctac tgtccctgaa
140521 tgctagagtc cattgattaa gttattcact gctgcaattg tcagagctga tcaaagaact
140581 ctgaaccagt gtgttaactag aactaacaaa gaaaatgcc ttagatggt ctagagtctt
140641 gaattagtag aagaggttta ataagaacct taagggattg ctagaatggt aaaaacaaac
140701 aaacaaaaaa aaagggtgaa aagtttagaa aattcactgg tctttgtgcc catcatttta
140761 cttccagggt ttagataatc tcatttttgc aatgaaggaa tggattagat cacaagtctt
140821 catcctagta gcacatgcag aatctttata aaaacacaga gtagccagggt gcggtggctc
140881 atgcctgtaa tcccagcact ttgagagcct ggggcagggt gatcacttga gaataggagt
140941 tgaagacca gctggccaac atggcaaaac cctgtatcta ctaaaaatc aaaaattagc
141001 caggcatgat ggcacatgcc tcccagctac tggggagggt gaggcaggag aatcgattga
141061 acccgggaga tggaggttgc agggagctga gatagctcca ctgcactcca gcctggtgac
141121 agggtgagac tccatcacaa acaaaacaaa acaaaagaaa gcaaaaacac agattactca
141181 gggtcacta agaccagtga agtcagttct cttggtaggg ggcagggtga ctgagcatga
141241 tgtttgtaat tttaaaagtg ctccagggtga ttctagcgtg tatcaagcaa gacttgtgaa
141301 ccactgaact acatgctaag actcatttta gctctgattt tctgtgagtc atagcagagg
141361 gctcagcaaa ctttttctat aaatgctaag atagtaaata ttttcagctt tgtgggctgt
141421 atcgtcttta tgacaactca actcagctct tgtagagaaa agcagctgta cataatatgt
141481 aaactaatgg gagtagctag atgtgtcctg tgggccatag ttttgcctga tcctggtcta
141541 tgtcatagaa tttccttttg aattgatgga ccaccagcaa atgatttttg tcctgtatca
141601 atcaatgata catacataaa tctctacaag acatgtaaag gatgaggctt aatgacagag
141661 tactttgggg aagacataat attgcaaaat taagatgctt agagaaaaat catattaaaa
141721 tagtgaaaac tgtgagaagg tatttttgatt tgttgttttg gattcctctt tttgcaaatt
141781 cttttgaaat attttcagtg gaagctacat agatccaatt gtattcacca agctagattg
141841 taattaagct ccagagtaag taatagattt gatgagtgat gtccaacctt ttacatggaa
141901 gagtaagttt gagtcttctt ttgccattg acacacttag taccatgttt accaaagttc
141961 ttagttattg aaatgggcac cagcatattt tgaaacgttg gtgttaactt gggatatgcc
142021 ttttgtcatg ttgcaaatag attttgtttc tgttttgtga agatcaccat ctctgtcact
142081 tctgatagaa aaagtgcac tgacttctca agtgatttga cacaggttaa aatatgtaaa
142141 ccatttctgt agagagcaag ctgtaataat atactaaagg gctaggttta tagtataata
142201 taaataactc atttatgctg ttaataattt atagcaacat ggcatttgac tgacttttta
142261 tgtgctctag tcatgtaagt aatagatgtg gaaacataga ccagagtttc aagaacatgt
142321 tttgggcaga gtctgttttc ttgctattat ctcttaagtt tatgttcctg gcctaaagat
142381 tatgctaatt gatctgcctt ggtcttgggt gtcagggtct tgttagcgag tattgaaaag
142441 catagttttt gcctactggg aaggatttat gatttaaaaag ccctaaatct ccccttttat
142501 gtacttcata cttagaaaat ttttctgtta aactgtgtga cttttttaca ttgtgccagt
142561 tttctagatg actctcgtca tatttatttc ttgcaatcct tctataacta tcagttatga
142621 agtctcttta tagtgttgcc agccaggctc cagggtgtgtg aaatgtattt tctattatgg
142681 attttgggggt atgatggcac atagtttggg tgttaatgcc taatcttgat gtactggctt
142741 ctgaacaacc aaaaggatga aaggaaatag aacaaatatt tttgtgaggg agaggagtct
142801 ggcttcttga cttactctag aaaaagcctg taagcctcct cttccctcct tgtcacacaa
142861 agtgacaaag aaaatcaaga attgttttct tcttggctta aatgcacccc ttataaagta
142921 aggctgagat caggctgtga agctatcttt ttgtcaagac tgtcataatt ccaaaacact
142981 ttgttcttct aatgcttagg ttagtaactt taaacatttt tataaagata gtgagggtcca
143041 gttttaagga ttgaacctt ctcaaggggc tcagaagagg ttttggagaa taataaaatt
143101 aaataatgaa accaataatt taaaccagat catgatcctt aagaaaaaat cccatcaaat

FIG. 1 (cont.)

43/63

143161 ttggggctaaa ctctaataata cagaggtctg cacaacttat gtcaagtatt cttccccaca
143221 aatgaagaat ggggttcatt gtgtcattgg ttgggtctca ttttggttc atcttctatt
143281 tctcaaagtc taagaaaagt gtcctacgg aagtgggtgt tggctatcat gagactttgc
143341 tgctggcagg ccagcttgct gctctagaca gagatatccc tcgacctcc ttggacaact
143401 gttttctgtg cacaggaagc agcaggctgg ggttaaggag tttgccaatc cagtcattct
143461 gataattgct gaatatgaat ttctatccag cacaatctag gtagctacaa tggcacagta
143521 gtttttatgt atcaggtgaa aatgtttaat aggcactcta aatgagagaa aaggttaagt
143581 gaggttaaaa gctcaatgaa aacaaataga tgagactaaa aatagttcaa taggttgtaa
143641 cttccatctc atccaaacag caatgaatat tttgaggctg aggcgctgag gggtaaaatt
143701 gcagcctgga ctacttgcta atgtagacct acagcactgt cattcttact gcacagacac
143761 tgctttctgc ataggaggta gaataatgaa ttcatttatt attaacaaag atttattaag
143821 tgactgcatg gtgctaacca ctagatgggg agggatgttt tgaactgtcc attgtttgac
143881 tataacaagg aacgctttga acgaggttac tatcataggc agaatttggt taacatgaag
143941 cctatgagac ataagccaca ggtcctctca cgtgcaggaa ctcccttgaa ggccctatac
144001 ttaattttat atgcatagtt tggatttgga ttcttttttt ttaagagtt ccccaaatta
144061 cttaaagcttc aggcctccaca aaacctggat ctacccctgg tagcagctat gaatctttga
144121 ctatgaaatt aagtgtacaa gaaatatgac tttacttttt ctgtgattga gtttattttc
144181 tatttgagca cgcattccac tgagtgaag aaataatctc attgaattca gagattttgc
144241 tgggttctaa gtggagttta cagaatgcca tgatattagg aattaaggag tgtgttgccc
144301 tacatcatct tttgtcctg ctcactgtct ctgaggcact gatgttcta tgtgacctag
144361 aggggcatgg tccaggtaga tggagtctgt ccttggtctc actgtgagct ctgcttgct
144421 gacccttctt cagtttcttc catgcccctg aggggtaaaa agattcaaat ctgaagctat
144481 atcaagccat ctgtgcatag acattccaag caacctggt cactctactg ctcccatgct
144541 atgcaaggca caggaagctt cactatggca tgagtatttc ctgggctttg ccttgggaatt
144601 gaggcacggg cctcctttgt tctaaaattc cccaaatcta cttgaggata gaaccaggat
144661 ttggttgcaa ggcagaactt ttcttagagg acctggtatc taaacctct tgttaccccc
144721 atttatggac cccatttatg ggtgaggag agtgactgct tctaatecat cataattttt
144781 gtctatggct actgtttttg catagacact atgttttgag tccttaggct ttggcttttg
144841 gcgcttaatg gccaatattc acatggctca aaattttcaa atgatccata tctgacttga
144901 gtttcaaaag tcagtttttg aaacttaaat gatcagaatt gatttgttct gctctggttc
144961 tgatgtggcc tctccttcca gaggtactgg aggtagaata tccaagggtg aaagcccacg
145021 actacaagga attggttagt aattcataat gttagctgtc cacatctatt cagtaatggc
145081 atttcagtgg ctgcacaact gacctgggtg aaagtgtctg cacaagccac tttttcttcc
145141 tgtcagaaaa tgttctcacc cactgaattg aatgactgtc tgctcatatg ctgtgaatga
145201 gtgcccagtc ttaagattaa atcacacgtt cttggctatg catatttggg catgctgtgg
145261 ggagttataa taggctgtct tagagtcaca ttaagcagct agacagacaa tgagttggaa
145321 agttacattt tctaaatttg attggtacat tccatttgtc acatttgaca ttagaagttc
145381 tggattcacc ctctatggtg agcttcacta atggagaatg taatttgcaa tgcctaaaca
145441 caagtcctaa acagaaaaca ttgtatgtta cattccagtg ctacccaaat agtggttttg
145501 aaagtcctta ttttctaata ctactatgtg taattttgag tcatttagat agcaacagtt
145561 aaatgtttta tagattgttt ggaagtatta aaatgtgaag gatttttggt atatagtgtc
145621 tttcctatct tgcttaataa aatataagtt tagaattgtg tatagaatta acatgcaaaa
145681 atatcaagtc tcaactttat acagttaatc tacatttgtg tatacccttc aattatttca
145741 agagagggat actattctta tgcaggataa atacaataag atattttaaa tgaattttaa
145801 ctacatctct ggcagtttca tctcaatagt agttgtaatt ttatctccca gacctatta
145861 tagactagca gctctctatg aaaattagtg acagtgtgag tgtattttta ttcaaagtta
145921 atcaagaatg actgagtcaa gagttagcta cccctgaaag taactcataa ttcagaattt
145981 aaaatattac atgtggaaca atcatgacta tatgcctttt actttctcta tcattattta
146041 gggtgtgggc tttgggtcct tttcacatcc gttaacagtg ggcttgactt caaaggatta
146101 ttttcttgaa tcttgaataa ttgctgaaga caatttgaag atattttcaa gatgaaggaa
146161 actgaagcac agaatcacta gagtgaaaaa agaacttcac aaacagtgca ggcttgatca
146221 atggcatggg aaaacaggca atacagttag aattgctaag atggaatttt aacgttcaat
146281 taaggatcta tctctaaaact cctctgcttt atccaccaat cattccatat taaagatgaa
146341 gaattgttcc catttcacct tttgataagg aaaaatagaa ataacagaag caaatacact
146401 tttgccaca tttttttcca aaaagaataa tttttgaagt ctaaacgttt ggtgtaaata
146461 agatgatgtg ttaatatgtt aaaggaaagc tagttaagtt tttgactgaa taaagccagc
146521 atcaataatt actagtaaga ctaaaaataa gagcagtaaa attgtgtcta atcagctact

FIG. 1 (cont.)

146581 aatatctggg aaggattgag ccacaggatc aaagatggta tcttttaaaa atagaagttg
146641 agtgaattcg gtcttcaaat tctttctttt tattcattta tttttattta ctcatlagta
146701 tattcattcc tttattcatg tattgttcaa atatatattg ggtacttatt atatgccaa
146761 ttgtttttta aatcacattc caaattcccg taagtcataa ttattcagag atgtatgttt
146821 tttttaaaaa aaattgaaca cctttaaaaa ttatcaagtc cttttatttc tgtatgcatt
146881 aaagataaac tttactaaat gttacatgaa tagatttata aagcagataa atatttaatt
146941 tcaaatataa ccottatatg caattatatt ttcttagca ctaaaaatga atatttaagt
147001 aatttatatt aaaagtgtaa ttatttaact gcagatgtat gccaatgact taaattgttt
147061 aaagattata gcaaagttgt ttaaaattgt ctaatcatga agagttcact taaccacctg
147121 gttgacacat aaaattatag ttagttacta aggtagttcg agagaaagag aagaatcttc
147181 agtagtggtt ttgaggtgtg gtacatttta ttataatata ccggttatac agcattgtgc
147241 agtgctgctc atagtagaaa taaattttct ctttgatgtc atctattccc ttgtgtggct
147301 tacataactg agaattaggt gatcacaaaa ataaacaggc ctatacagag cccatttata
147361 taagtcctgg ttatttctct tcagttaaac ttttaattat atccaattat ttctgttag
147421 ttcattgaaa agcccgacaa ataaccaagt gacaaatagc aagtgttgca tttacaagt
147481 ttttttttag gaagcatcaa actaattgtg aaattgtctg ccattcttaa aaacaaaaat
147541 gttgttattt ttatttcaga tgcgatctgt gagccgagtc ttttaagttca ttgacatgcc
147601 aacagaaggt aaacctacca agtcaaccaa accatacaag aatggccaac tctcgaaagt
147661 tatgattatt gagaattcac acgtgaagaa agatgacatc tggccctcag ggggccaaat
147721 gactgtcaaa gatctcacag caaaatacac agaaggtgga aatgccatat tagagaacat
147781 ttcttctca ataagtcctg gccagagggt gagatttgaa cactgcttgc tttgttagac
147841 tgtgttcagt aagtgaatcc cagtagcctg aagcaatgtg ttagcagaat ctatttgtaa
147901 cattattatt gtacagtaga atcaatatta aacacacatg ttttattata tggagtcatt
147961 atttttaata tgaaatttaa tttgcagagt cctgaaccta tataatgggt ttattttaaa
148021 tgtgattgta cttgcagaat atctaattaa ttgctaggtt aataactaaa gaagccatta
148081 aataaatcaa aattgtaaca tgttttagat ttcccatctt gaaaatgtct tccaaaaata
148141 tcttattgct gactccatct attgtcttaa attttatcta agttccattc tgccaaacaa
148201 gtgatacttt ttttctagct tttttcagtt tgtttgtttt gttttctttt gaagttttaa
148261 ttcagacata gattattttt tcccagttat ttactatatt tattaagcat gagtaattga
148321 cattattttg aaatccttct tatggatccc agcactgggc tgaacacata gaaggaactt
148381 aatatatact gatctctgga attgattctt ggagacaggg atggtcatta tccatatact
148441 tcaggctcca taaacatatt tcttaattgc cttcaaatcc ctattctgga ctgctctata
148501 aatctagaca agagtattat atattttgat tgatattttt tagataaaat aaaagggagc
148561 tgaaaactga attgcaaact gaattttaaa actttatctc tctgtgggta attgcaaaca
148621 cagatacaaa aatatagaga gagatacagt tagtaaagat gttagggtcac cgttactaac
148681 actgacatag aaacagtttt gctcatgagt ttcagaatat atgagtttga ttttgcccat
148741 ggattttaga atatttgata aacatttaat gcattgtaca aattctgtga aaacatatat
148801 ataggatgtg cgaaaagtcc ctgtgtatca tgtgaaatgg cttaaaacag aacaccatag
148861 gtattcatat cagtgaatac cataggtagc tgaaagtgtt ttttctggg gtcgccaaga
148921 tgaatgcaa aagtgatatc attattataa acaatagcca gaatagggtg gtataaacct
148981 ggtagaaagc cttgataaat tgactttctc tctcctgac atcctgocac ccctttgctt
149041 tgctgatgct catttgtcca cttaaattaa ctcaagcaag ccctagtaaa gtaatagaat
149101 ttgtggagtc ctcatlagta taggaagttt ccctgatgtg agattagtaa ttagagatgt
149161 agcaaaatga gaaagaagta atatgcttag atatttcatt ttctctgaac ctgtatatac
149221 aaaataggcc atgcgtgttc agtaactatt cactgcaagg cactctctag gtactttggg
149281 ggaattggaa attactcaca taaggctatg gattgtgcca tttgtcaaaa gacaaaatga
149341 caacaaattt agtttaaaga cctcagtcag ctttattttc tattctagat ttggacagtc
149401 cttcatttca caaattggag taagtgttcc aataagttga gcaaaggagc ttggctttat
149461 agacccaaaa aaagggccaa aggaagcaga aacaaagaac aataagagaa ttggctcattt
149521 caaagttact tttcttgaaa ggtggggaca aggagacaga ataatagaaa agtcactgat
149581 tggttaacat tggattaaga attaaaacag aggaaacttt aagattgaag tttgaaactg
149641 acttgttttg gaaatcaggc tgtcttcttt cttgatttct tagaaggccg gataacaact
149701 gagttttgct ttggtgaaca tgggtgaact catttttact tttagtctgg tctgttgagg
149761 cctcgtgaga gagcttaatc taaaacaatg acttctata atttttgttt gacacatcca
149821 aagagggact ctaatattta ttgagagctt atcatatctt aagtaactgtt taaacacttt
149881 ttttgctat tacatttgat cttattataa ctctaaaggc agaaatgatt gcttttattt
149941 tccacaatgg aggaaactga ggttcaatta agtgagtaag gaagcaggga tcttaaaccc

FIG. 1 (cont.)

150001 agataccatt gctcctcttt aaaggtggaa gaacagaaaa catggggcag gggaagagag
150061 aaagtttctg tcccaggaca tgataatcta aaagggaaaa cgtaagatcc actgaaacct
150121 gaggcagatt tattgtggca ataacaaagc ttaagtttca cagaccttca tttgcctgag
150181 ccaactttga aggccatgta tctaattttg tttttataat tctataatct ttattcttga
150241 aaagagccct ccctccaaat ttacaagctt tgggccccca aaatccttga aatgcccttg
150301 aataagagat atccaggtaa atgctatggg aattcagagg aggaagcagt tagtatcagt
150361 tggcggagag ttaggctatt aagagaaggt tttatatagg aagtggcatt tagaatgaag
150421 ctttgagaac tgagctgtgt atttgaacaa gtaaaggtgg tgttgcagaa ttttgctcct
150481 tagttctatt aaaaacccgg gttcttgtca catgatccgg aaaatttagg cacacagata
150541 cattgaagca tgagtagagc aggattttat tgggcaaaaa ggaaaaaaag aaaactcagc
150601 aaatcgagat ggagtcttgc tcacagattg aatcccaggc caccacaaag gaactgaaga
150661 gatcgggctt ctcctctgca taaggtgcaa attcccatg gctccacca cttccctta
150721 gtgtgcatgt ggggctccag tccacgggtg gcattgccag acaagccttg ggcaggttcc
150781 ctcatctgtg caaaagcatc tgatgtaaac acttgagggg tggttcggag attctctggg
150841 acccttttat tttcttatct gcctaggcat ttggctgtct cagtgggttg gaaaggtgct
150901 tccaggcaaa gggcataaca tgaggcaaa ggcattgcaca gaaaacagtg actggttcag
150961 tcaggttggg ggatgccaaa ggaagtaatg ggagacaaga ttggagcaag atagataaga
151021 gattgtggat tttttttctt ttttatctat ataaatacag agacaggggtc tcactatggt
151081 gccaggctg gtctcaaact cctggcctca agtgatcctc ccacctcacc ctcccaaagt
151141 gctaggatta caggcatgag gcactgtgcc caacctccaa ttttggattt tgagagctaa
151201 agcaatatag tcgaaaactc agataatcca ggtagatttt gctattaggt gctatttggg
151261 tcctggtaca gagctaaaac ccttgggaatt tcctaagtga taagagctac aggagcatct
151321 tttgttatat gtttccccc ctagtctctg aaatagctct agagaaatac aggtgaataa
151381 catcctttgt tattcatatc aagccctat caaccatacc ccagtctcta tttatgaagt
151441 ggcttttggg aagtccctaa agacaggagt ggggaaaggc tggttgtcag ggggatgggt
151501 tgaaactttc atcttccccc cttgacctcc agggagggat gagtggctga aaattgtgta
151561 aaatcaacaa tggccagtga tttaatcaac catgcctatg taatgaagcc acccgataag
151621 cettaactgg aacttttttg agagcctcca ggctggtgaa gacattgagg tgctcagaag
151681 gtggtattcc agagagagca cagaatctct gttcccttc ccacattcat tttgctatgc
151741 atctctccca tctggctggt cttgagaggt atccgtttat aataaaactg taacctagta
151801 agtaaactgt taccctgagt tctgtgagcc attctagcaa attatcaaac ctaaagagtt
151861 catggatacg tgcaatttac agatgcacag tcagaagcac agatgacaat ctgggcttgc
151921 cattggcatt tgaagtgtgt tgggaggcag tottacagga atgagccctt atcctgtggg
151981 gtctatgcta ataacagaca gttgtcagca ttgcttgggt tcgaaaaccc acattggttg
152041 tgtcagaagt attgtcagta ggatagggaa aacagtttgt tttctttttt tagtggctct
152101 tggtcactct taagagcagg gcttctcaaa gtgtggtcct tgaaccagca tcacctgtac
152161 cacgtaagaa cttatgagaa atgttcattc ttgggcccc acaaagaatt aaaaattctg
152221 aggggtgtgaa cggggtctga gtttcagcac aacttccga ccattgctgat gcattcttgc
152281 ccaagcatga aagccctccc ttgtttaaga aggcattag ggccgggtgt ggtggctcat
152341 gcttgtaatc gagcactttg agaggacata gtgggaggat cacttgagcc ctggagttct
152401 agacaagcct gggcaacatg gcaaaatgct gtctccacaa aaatcacaaa aattaggtgg
152461 gcgtgtgttg tgtgcctata ggcccagcta cttaggagac tgaggcagga ggatcgcttg
152521 agcccaggag attaaggctg cagcgagctg tgatggcacc actacagcct ggatgacaga
152581 gtgagacact gtctcaaaaa aaaaaaagaa aaagaaaaag aaaaaagaaa ggaaaatgaa
152641 aaagaacgcc attaggtata aaggagcaat ggtaaaagac cagttgcaaa aggttaggga
152701 atgggtgggt actgaaataa gaagctatgt agaacactag tgttgggtggc aggaagtaga
152761 aagcaagagc actgctctgt gggggatggg catagcaaat gcaatatgga ggcatttgcc
152821 tctgcaactga ggagaaaact atcttttcca agataggagg aaaggagata agtggaaatta
152881 aagagaacct ttgagcacag agttgggaaa ctgaagggtat ttgtgttgtg ctccctcaat
152941 cttttaattc aactataagc taaacccatg aaacttgagt agtttcagtt atctgacttt
153001 tttcttctct tttgatacag tgttggctat tctgggtctt ttgcctctct ttatgtactt
153061 aagaatcagt ttgccaatgt atgcaaaata actggctggg attttgattg tgattggctt
153121 gaatctatag atggagttag gaaggactga catcttgaca atgttgaagc ttccatttca
153181 tcattatgaa atatttctcc atttgtttga ttctttgatt tcttttatca gaatttagtt
153241 ttccctcatat agtcttttaa aatattttgt tatattttgt tcaagtattt tgtttttgag
153301 gaatgccaat gtaaattggta ttgtgatttt aatttcaaat tccaattttt cattgctggt
153361 atataggaaa atgatttttt ttgcatgtta gccttatatc tttcaacttt gctataatca

FIG. 1 (cont.)

153421 attattgata gtttcaagga ttttttgggc aattattttg aatctttctac atagattatc
153481 atcatctgaa cttagtttta tttcttccct cccaatctgt atacctttat ctccctttct
153541 tatttcatta gctaggactt ccagtatgat gttgaaagta gtggtgagag gggatatctt
153601 ggtcttggtc ttgatcttag tgggaaaact tcaagtttct tatcattaag tatgatttta
153661 gctggagggc ttttgtagaa gttttttttt ttttaagttga agaagctcc ttctattttt
153721 agtttgctga tttttaaaaa gaatcaggaa tgggtgttaa attttgtaa atgctttct
153781 gcaactattg atttgagcac tttatttttc ttctttggct tgttgatgtg aagtacatta
153841 attgattttt gaatgctgaa tcaacctttt gtacctgaga ttaatcccg tgggtgtgg
153901 tatataatta tttgtataca tgttgagttc gatttgctaa tactttttga gaatttttgc
153961 attggtgttc atgaaaaaat attggtgtgt agttttttgt gacatcttta tctgcttatg
154021 gttttaaggc aatgctggcc tcatagcatg agttaggag tatttctct acttttacat
154081 ttgagaagag attgcagaga attagtaaaa ttctacttt aaatatattg tggaaatcac
154141 cagtgaacct atctggacct ggtgctttct gttttggaag gtcattaatt attttaaaat
154201 agatataggc ctattcagat tactatttt ttctcatgcg agtttttagc gattgtcttt
154261 caaggaattg gtctattttca tttaggttat caaatatgtc aacgtagagt tattcatagt
154321 attcttttat tatcctttta atgtgcaagg gatctgtagt gatgtccct ttttgtttt
154381 attgatatta gcaatttggt tcacatcttt tattttgctt tgttagccag gctagagata
154441 tctctatttt tgatgttttt gatgaaccaa ctttttgtt tattgatttt ctctgttgat
154501 ttctgtattt caatttcag atttttaaat tatgcttaca tttgatttaa ttgatcttc
154561 ttttgctagt tatccaagg ggaagcttat attgttaaga tccctttgca ttcttatgca
154621 ttcaatgatg taaatttccc tctaagcact gctttttctg catctcaca atattcatga
154681 gttgtatttt catgttcatt tagtttgaaa tatttttaaa ttctcttga tatttctctt
154741 ttgacctatg tgttacttag aagtgtgttg tttaatcac atttttaaaa attttctagc
154801 tatctttctg ttattgatt ctagtttaat tccattgtgg tctgagagca tatattgtat
154861 aattttaatt ttataaaaat ttgttaagg gtgatttatg gcccagaat tggctatct
154921 tgggtgaatg tccatgtaag ctttggaaga ctgtgtattc tgotatat ttgaaggta
154981 gtctatagac atcaattatg tccagttgat tgatggtgct gttgaattca actatgtct
155041 tactgatttt ccacctgcta gatctgtcca ttctttgcag agggacactg aagtctccaa
155101 ctctagtagt gaatattcta tttcttggt cagttttatc aacttctgct tcatgtcttt
155161 tgatgctttg ttgctagaaa catacacatg aagaattgg atgtcttttg gagcatgacc
155221 catttatcct catataatgc cctcattat ttctcgccc tgatgtctgt tctctctgaa
155281 agaaatatag cctctccagg tctcttttgg ttggtgttaa aatgaactaa cttcttttat
155341 ccccttact tttagtttat atgtggtttt aaatttaaa tgggtttctt gtagacagca
155401 aatagttcag agttgttttt cgatccact tgacaatctt tgtcttttaa ttggtatatt
155461 tggactattg atattttaag tgattattga tatagttaga taaacatcta ctatatttat
155521 tactgttttc tgtctgttac actactgtt ctttgtttat atttttattg tctactcttt
155581 ttctttccat tgtggtttta atcgagcatt ttatatgtt ccattttctt ttcttagcat
155641 agtaattctt ctttaaaaaa acatttttta gtggttgccc ctagagtttg caatatacat
155701 ttacaactaa tctaagtcca ttttcaata atactaaata atttcatgt tagtgcaagt
155761 accttttaat aataaaacac tcccagttcc accttcagt ctcttgatt atagctataa
155821 tttagttcac ttacatatat ggggtatacct aagtatatac attatcatat ttatgattga
155881 atatattgat gaaattatt tgaaaaaact gttatcgta aatcaattaa gagtaagaaa
155941 aatagttcta attttattat aaaatgaaat accttcatt attcattct taatacactt
156001 tctttcttta tgtagatcca agtttctgac ctgtataatt ttcttttct ctcttcagct
156061 tctttgaaca tttcttacca gccagacct ctgacaacaa ttttcccaa ttttgtttg
156121 tctgatagag actttatttc ttcttgact ttgaagaata attccacagg gcacagaact
156181 ctgattgggt gatttcttcc cctcaaaccc ttaaataatt cattccactg ccttcttgc
156241 tgcattgttt ctgagaagtt agatataatt ctatctttg cctttctata ggtaagatgt
156301 ttttctctct ggcttctatc aagatttttt ctatatgaac atgatatgcc tttctttttg
156361 aacatgatat gcctttcttt ttgaacatga tatgccttg tgtcggatt ttttggcat
156421 tattctgctt ggttttctct gagtttcttg gatatgtgg atggtatctg acactaattt
156481 ggaaaaatc tcagtcatta ttgcttcaaa tatttctct gttctttttt ttctttatt
156541 ctcttctggt tattcccatt acatgtatgt tacagtttt gtagtcatcc cgtgttttg
156601 gatattctgt ttttttcagt ttttttttcc ttgcatttc agtgttgga gtttctattg
156661 acatattctc aacctcagag attctttctt cagctgtgt cagtctacca atgagtccat
156721 caaaggcatt ttacattttt attacagaat ttttgacct tagaatttct tttgattcca
156781 tctttgaatc tccatttctc ttctgcttt catctgtct tgcattgtgc ctacttttct

FIG. 1 (cont.)

47/63

156841 catgaaaacc tttagctttt tttttttttc tttttgaggt ggagtctcac tgttgcccag
156901 gctggagtgc agtgggtgtga tcttggtcca ctgcaacctc tgcctcctgg gttcaagtga
156961 ttctcctcct cagcctccca agtagctggg attacaggtg cctgccacca tgcctgagta
157021 atttttgtat ttttagtaga gatgggggtt tatcatgttg gccaggcggg tcttgaactc
157081 ctaacctcaa gtgatctgcc caccttagcc tcccaaattg ctgggattat aggtgtgagc
157141 caccatgccc tgccttttagc atgttaatca tagttgtttt aaattcctga tctgttaatt
157201 ccaacatccc tgtcatactc gactgtgggt ctgatgcttg ctctgtgttt tcaaaggtg
157261 tttttttttt tttgcctttt agtaagcctt gtaatttttt attgaaaggt ggacatgatg
157321 tgctgggtaa aaggaactgt agtaaatagg ccttttagtaa tgtactggta ggtgtagcag
157381 agggtgaggg aagtattctg tagtcctatg attagggttt agtcttttag tgagcctgtg
157441 cgcctgcagc ttggaagcac ttgtgaagtg ttttttcacc ctttttggtg ggacatagtg
157501 actagtgtga gcgggagttg agtatttccc ttcccctagg tcagttaggc tctgaaaaaa
157561 ccctgatagg ttaggcattg taaaatagtc tcttttgagg gcaggcattg ttataagaat
157621 agaatgctct ggggccaggt gcggtggctc acgcctgtaa tccccgcact ttgggaggct
157681 aaggcaggtg gatcacctga ggtcaggagt togagaccag cctggccaac atggtgaaac
157741 cccgtctcta ctaaaaatac aaaaatcagc cagggtgtggg ggcacacacc tataatccca
157801 gctactcagg aggctgaggc aggagaactg cttgaaccca gtaagtggag gttacagtga
157861 cccaagattg tgccactgca gtctagtctg ggtgacagag caagactcog tctcaaaaaa
157921 aaaagaatgc tctggcatat ttgaaaatgg ttacttttcc ctttttttct ctgatcttca
157981 ctgtgagaac ctggtaagca tcctataggc aaaattcata aaagtataga agtcggccag
158041 tgacttggac ccacttggaa ttttcttgct ctacatcat gcacactgaa tctccagcaa
158101 tttttcactt acagtttagg ttttctacc ctactactgg ttctctcaga ggtttctgct
158161 tattggtttc tgttttgtaa gttgtgattc tctgtaccta actgcctgtc tcccattttg
158221 gggggcagtg gtttgccctg tgacctcact tctctgacag atctaagaaa agttgtttat
158281 ttttcagtgt gctctgcttt ttacttgtta cgatgaagcc aaccactttc agaatttcta
158341 caaaccagat cagaatctgg aagtcctgtt tttttatttt ttttatccct ttgtttagca
158401 tgttacctat cttaacacat tttaaataag tgaatgcata gcttatatct acttctaggt
158461 tatatgcttc cttagaatag gaattgattc ttaaaatgtc gttctgctca cgctgtaat
158521 tccagcactt tgggaggcca aggcaggcgg atcaactggg gtcaggagtt caagaccagc
158581 ctggtcaaca tggtaaaacc ctgtgcctgc aaaaaataca aaaattagct gggcatggtg
158641 gtggccatct gtaatcccag ctactaggga agctaaggca tgagaatcac ttgaacctgg
158701 gaggtggagg ttgcagtgag ctgagatcgc gccactgcac tccagcctgg gtgacaagag
158761 caaaactcca tctcataaat aaataaataa ataaataaat aaataataaa aataaaaaaa
158821 taaaataaaa caaaaatttt attctgagca gtctctgaag aatataaatt ctactgcctt
158881 gcctttagaa cttataacag catctcgcaa actatcacaa gatgctccaa acatacttct
158941 tatgtgctga attaagaagt caactcaaat ttagtatact agtaatatatt ttggatatcc
159001 caaaacactg ccagctcagc tttaggctgc ctttcttggg ggggaaaaaa gcagttgaaa
159061 tttaggactt aagtgggcat ctcgtttaat ttttaatgga tttctatgtt gttggttatg
159121 gtgaagaggt gaaaagaata aatattctgt gcagaaaaat tattcagttc tcatgtgaaa
159181 acactttgtc catagcaatt actttatgaa aaagatgtgg tattactttc tttgctctta
159241 actgagacct ttaatttaaa gaacctatac tttacaagtt tttattttca atgcatgaaa
159301 aatgtagcag ctatttcaca acctttactt ttaaaatcca ttttcttttt taatctcaaa
159361 tagttttttc ttaaaacctt ttgacttttt atctaaattg taatagccag agcaccttcc
159421 cacaactaga atatctcatc ctttttgtct tttctttttc ctctcaaaat gcctactggg
159481 aacttaattt ggagtcagat tcttcatgat aaatctggac ttaatcaaaa ttctcatat
159541 ggtatattgt atatatcaca gtactggata gtcctctgat taaatagata tttgatagta
159601 ctttaaggtc tatacttttg gatgaactta actgctttct ccatttgtag tctcttgaaa
159661 atacagaaat ttcagaaata atttataaga atatcaagga ttcaaatcat atcagcacia
159721 acacctaaat acttgtttgc tttgttaaac acatatccca tttctatctc tgataaacat
159781 tgggtgtaaag tagttgaatc attcagtggg tataagcagc atattctcaa tactatgttt
159841 cattaataat taatagagat atatgaacac ataaaagatt caattataat caccttgtgg
159901 atctaaattt cagttgactt gtcactttga tttctggaga ccacaaggta atgaaaaata
159961 attacaagag tcttccatct gttgcagtat taaaatggcg agtaagacac cctgaaagga
160021 aatgttctat tcatggtaca atgcaattac agctagcacc aaattcaaca ctgtttaact
160081 ttcaacatat tattttgatt tatcttgatc caacattctc agggaggagg tgcattgaag
160141 ttattagaaa acactgactt agatttaggg tatgtcttaa aagcttattt gcgggaagta
160201 ctctagcctt attcaacaga tcactgagaa gcctggaaaa acaaatcccg gaaactaatt

FIG. 1 (cont.)

160261 attatgtgcc agttatataa acaagaagac tttgttgggt acaaaccagt gattccttgc
160321 ctttgaaaaa tgtgtcagat atcatgcatt accagcagtt caatgatata aggaaaccag
160381 agtaatagct aaaaccttta aagctaaacc aaagatttac aaattgcctc ttcattccagt
160441 ctttcccaac ctaaaaactg agttctctaa aaatttttagt atttttttct gaagaaaagg
160501 gaacatggac atttatctaa tcctcattag aaatctgact aatgataaca aggatttaga
160561 cctcaagcac ttcttaccaa aattcttgat atgaccttat agcaaattac tttcacctgt
160621 tgaactttcc tttcttttat tccctgtac ctacctgca ctgggcata tcaagttgct
160681 tatacaacac tttactattg tgttagaaaa atcatgacac atgatgaatg tgtttgtgca
160741 acatgagctg attcataaat gaaaatgtgc attgaaattc cacaatattt taaaattagg
160801 agtttatcta gcaattgaac aaaattgatt aaatccatta tttgttagat cagctaaatt
160861 acataagttc attcatctgc tcataaatcc atccattctt ccatctggct atcccttagt
160921 caattcaaat aaatatttat ggggcacttt gggtaagcca ggtgctaaga attcaatgca
160981 aaacaagata gactccctgt tccttgttga acttataatt ttggtacaaa caaaagcaat
161041 aatcaagaaa aaataaaaaa agtactgatt gtgattaata atatgaagaa attcaacaga
161101 gtattgtact taacatttga ttgatctgat tttctcagtt gtctgagaac aaacatttgt
161161 gaaaatctca ttgtagagtt cttacgatgg ataggggggtc aactgtgtca ttattgctta
161221 tcagcttatc ccaaagacct agtttattac cagattgcaa atagtgttca ataaattatt
161281 cttattaagg gttgttatgt actctaaaac atttattgtg gtcccttcac tggttctggg
161341 ttacaaactt acttttctat gatgacatag tatagaaatt gagagtgaat atttagaagt
161401 tcatttttat tatatatattt tgaagtattg atatgtagtg aattagaaat ttaaaaagaa
161461 aacaaaactg tccttcacta cagattgaaa agcattatac taaaagacca tttgctcagt
161521 tatagtatat aaaggccaaa tgacttaaaa acaaattatg taaggagaag gaaacaacca
161581 tttattcagt gccactaact gtcagccagt ttttctcagt gtcagttaat gactgcagta
161641 gtgttctacc ttgctcaaaag caccctcctc aagttctggc atctaagctg acatcagaac
161701 acagagttgg ggctctctgt gggtcacctc tagcacttga tctcctcatg cagtgcattg
161761 tgctctcacg tctatgctat gttcttatgg tcttttaggt acaagaataa ttttctttct
161821 tttccttact atacattttg ctttctgaaa ttcccttctc gccaatccag gtgaatgtca
161881 gaatgtgatt tgacaactgt ccaaagtact cattcactga ggagtggtaa ggcttcgcgc
161941 caacctgcct tctctgggaa tatactgctg cctgaacata tcattgttta ttgccaggct
162001 tgaacttcac caaatataatt tattaggggtc aacatctaaa tattagaact atttcagatt
162061 aattttttaag tcgtatccac tttgggtact agatcaaatt gcaggctctct gcttctggct
162121 tgagcctatg tttagagatg atgtgcatga agacactctt tgcttttctt ttatgcaaaa
162181 tgggcatttt caatcttttt gtcattagta aaggctcagt ataaaggaag tctgcatcag
162241 gggccaatt ccttatggcc agtttctcta ttctgttcca aggttggttg tctccatata
162301 tcaacattgg tcaggattga aagtgtgcaa caaggtttga atgaataagt gaaaatcttc
162361 cactggtgac aggataaaat attccaatgg tttttattga agtacaatac tgaattatgt
162421 ttatggcatg gtacctatat gtcacagaag tgatcccatc acttttacct tataggtggg
162481 cctcttggga agaactggat caggggaagag tactttgtta tcagcttttt tgagactact
162541 gaacactgaa ggagaaatcc agatcgatgg tgtgtcttgg gattcaataa ctttgcaaca
162601 gtggaggaaa gcctttggag tgataccaca ggtgagcaaa aggacttagc cagaaaaaag
162661 gcaactaaat tatatttttt actgctattt gatacttgta ctcaagaaat tcatattact
162721 ctgcaaaaata tatttgttat gcattgctgt cttttttctc cagtgcagtt ttctcatagg
162781 cagaaaagat gtctctaaaa gtttggaatt ctcaaattct ggttattgaa atgttcatag
162841 ctttgatagt gtttttcaga agaccaaatt tacagtggga gccttgggct tttgtttttt
162901 aacagctctt ttttgttcct gcttcagtgg cctgacctcc aagttagcaa tcgccagggt
162961 gagaaatgct ttgogagaca taacagatgc tcctgaaata acaaacactt ggaatcatga
163021 ggtagtggaa ttgaaaatag aaagtgtagt gattgttttt tgttatttgg atgggatgaa
163081 caatgtcaga ttagtctgta actatttttt tttaatgtca ctctgatttg gtcacaaagg
163141 atctctagtc tcattgcctt agtatcattc tacgaattag aatgtgttac tgtgtaagag
163201 cacttcttgt atatgagaga aatagcaaca gttccagttt aaagtgatat aaatggaaac
163261 caagaaatgt ctttaactggg accaaatctg gacagcattt actgtatttt tgctgggtatt
163321 ttctctagtc tttccgggta tattcacatt taatgatcac ttttctccct ttgtgctaatt
163381 ggacactgaa tccattccac taccatagtt cttgctaata ctactctact ttttacacaa
163441 aattaaaatg ccaggagcac ctccaggtag actgactata aatctagact gaaaaaaaag
163501 cttgtatttc ttaacagatt acctgtgga acatttgctc ctttcaacta atgaggcact
163561 aaatattgta actgctcaac tgggtctttt aattttattg tctagacttt gtcagtgtgc
163621 cagaagcttt atcctgggtg gagttttgaa aacagtattg tttcttcaga aagaaaaaag

FIG. 1 (cont.)

163681 ggattgtcag atgatctaaa aataaagaaa cactggaaat acaagtatcc caaggtgata
163741 gcattaggca agataaaaaat gttgaaaagc gaaaaagaac tggttgatag agaagtgttg
163801 ttattcagta gaacctaaagt cttgtggtcc cattttttaat gaaaaatggg gaattttttg
163861 gtttttattg ttcttggtca cacaaatctg cccattagaa taagccaagc cctaaaaatt
163921 aatttcagtt tccactgggaa tccttttagtt tatctactat gtagtagaga ggttttgttt
163981 tattgcatgt ttgacgtagg aacgtatata tgcaagacat ggaggaaaac caagtgggcc
164041 agagttttga aaattcttta tcttttcttt ctgccaaagt gagtctccca agtttgtctt
164101 ttttttttca tttccactct tctatggttt ctagcattat ataaacccaa caaaaaaat
164161 acgttcagag attccttcag aaatgctgga tgatcttgat atcgatgctt ttcatatatg
164221 tgtttatgat gctggtttct ggggctggct ctcagtatca caaagatgtc tgtaaacaga
164281 atatgctatt tcttctttgt gacaaatttt gaacattatg tgaatgtcca agaaagagca
164341 aaagagggca aacttotcat acatttttga tgtcgaaacc aagagacgct tttattttcc
164401 taacttttct ttgaaagttc aaattaagta attttatcct gtcctaaagt ttaaaaagaa
164461 aaaaaaagg aagaaggaat taaaaatcca aagaaaatta tgtttgtttg cttttctggt
164521 tttttcttcc ttccaactcc gagactttgc aagggcatag ttctgaagat ctctgacact
164581 gagacattag agatctctgt atcaatggat catttgtttt cagacatatg aaacaggaac
164641 tttgaacaag aaatttcccc tctttttctc atagtgatcc tgagacatca gctgtggaat
164701 cacaacacgt cattagtttt ggcaggtcct tgcagggtgt ttgttttggt ttattaatgt
164761 tcttccctcc tgtagctaga cagcaatctt ggagaatctg ccagcttgga agactattgt
164821 gtaaatttca aggtggagcc tcctttaatt tgttctgtgt tacctgtgag ctgtgaggtc
164881 atgaagagga gacaatgagg ctaatcatga gagccccatt ggtttaggca attagaacaa
164941 caagatctaa aatggtttat tagccttgaa ttgtgttaag cacataattc ataaaaaaca
165001 gaaaaaatat ttttaaagt atgtctaaat cttcagttac aagtttgaaa ggtgacaaac
165061 tattctgagg aaatgattag gcctattctt gcaacgagtc tttatgatct gaaaagaatc
165121 tatgtccaca cataactccc acctcaaaga tggggcatct tttgctctgg gagatatcaa
165181 atgcgaccaa aacaagtgtt tgtagatttg aatgatgatt cagcagtgtg gcagttctca
165241 ctcattttat aataattaac aacttaataa ttaattatta aactcctaca tgcttaacat
165301 tataagtatg ataaactctg tggttacata aaagatatac atagcaactg tccttgatct
165361 gtcacagtga ggtcccaatc caacctatga gcttcaaatg aaaagttcaa aattacactc
165421 attgtcataa gtcagagatc aaaggaagaa aggatttaac caaatgata aattaaatat
165481 aggtgattaa atatagtcat ggttcaaggc atgggccagt tagggagtgt gatgtgggta
165541 attatgaaag gccagctccc aagccctgtt gttgctactc ccccatca gtcatccttc
165601 ctttttttct acttctactg cagtgccttc ctcatctttt cccttgcatc cctccattat
165661 atgagtcata caaattagac ttttcaaagc aacattaaca ttgtgtgaat ttggggtttt
165721 tgactaatcc caacattcca ccccatcatt ccagtccac atgggatttg gagccttggt
165781 tataaacctg gcacttctaa tatatcttat cttagagtaa tcttltgtatt tgtttaattt
165841 ccacttagca ttgtaaatac ttgcagggtat cctagttaag aaagcaagggt ttaaacacaa
165901 aatcatcac aattaaagca ggctagataa agaatgtaat agaaatgcta gataaaacag
165961 attttttctt actaagtttt ctgtccctta tagagtgcac aacacaataa cttgcttgat
166021 aagaattcaa tgtacattgt tttgtgctga atcactaaat gcttgatttc tgtaacaaga
166081 gattgtgggt ccacagtat ctggatttta gtctgtgtaa tcttaggcaa gttatttgat
166141 ttctctgtgc ctctgttttc ttgtctgtaa aatgagtata atggtagtaa ctaattcatt
166201 gtgtttttgt gaggattaaa tgagttaata actagtactc ctccctggca catagtaagt
166261 acaatatgct gtgctgtggg ggttggttatt attttttata gttccttgag caaaagaaat
166321 aatgtcccca tcttagtata atattggagg tatataccat agaagtgaac aaaagaatat
166381 agtttcacaa agaaagtgat aattaaggcg gttcataaag ggtcataaag cttgtagatt
166441 ttagaaatgt gggggcatga ggatgtggag agggatttcc aggatgccag acaggagat
166501 tatggatgag tactaagatg agaactagaa aaagctgagg ggcaaaagggt cagaggaggc
166561 cacaagttag ggagtattag gaaaaagaag ttaatacttg acaagtgcc aatggcttc
166621 acgaggaatg ggttgggcct ttttgagtga ggaagaggct ggtgaaaggg tggaggagga
166681 cactgctgct gctgatggca tgggggtgtg gtggcaggag aggcagggac atgagctagg
166741 aaactctcca gctatgaagt gatgagtctg gagtaatata aggacagtag gggaggagt
166801 ctgaacttaa gggaggagag aaaaataatt ggtatggaag taggtacaat gcaattttat
166861 tatttctgag cctaaaaatg tgaaattttt gattatttgg tcagaccagg gaagtatttt
166921 cttttatgct atctctgaaa atgtatacac taaaagttg tagtataaaa aggttgtaaa
166981 gcattaagta attttagagg aaacaataat ttggatattt tacatgcaat catttatatg
167041 caaatatatg taaatattac aaaattatto tctatttggt acaaacctta aatatttttg

FIG. 1 (cont.)

167101 actgaggaat atttttattca tctaattata gctactttgt tctaactaat agatattcct
167161 gaaaacaaag caacactttt ttggagacag agtcttgca tgcacactag acttgagtgt
167221 gttaccttga actccagggc tccagtgatc ctcccacctc agtctcttgg gtaggtggat
167281 tacaggccca cactaccatg cccagctgta ttagtccatc ctttcattgc tataaagaaa
167341 taccggaaac tgggtaattt ataaagaaaa taaatgtaac tggctcacgg ttcttcagge
167401 tgtacgggaa gcatagcagc atctgcttct gaggaggcct caggaagttt tcaatcatgg
167461 tggaaggcaa ataagaagca ggcattgttac acgacgaatc aggagcaaga caaagtgagg
167521 gaggaggtgc cacacacttt gaaatgagca gatctcatga gaacagcgcc aagaggatgg
167581 tgctataccg ttcattgagaa atccaccccc atgatccagt taacctccac caggccccgc
167641 ctccaacact gggaattaca attcaacatg agatttgggc agagacacag atccaaacca
167701 taccaccagc taataccaaa aaaaaaaaaa aatttttttt ttaagacatg gtcttactat
167761 gttctacagg ctggtcttaa actcctggcc tcaagtgatc ctcccacctt ggctcccaa
167821 agcactggga attcagacat gagtaacagt gcctggccaa taactatttt taaacattct
167881 ctaccataaa cttaggatct tgatttgttc acattgaaca gattttttatt atacagattg
167941 aattttataag aaaatgttgc agacattgtc aaaaagggac gtocaaacca ctgtgatatt
168001 tataagcatt tgggccacat ttgatagaa ctatacacgg agtgtgtgtg tgtgtgtgtg
168061 tgtatatata tatacacaca cacattattt atatatatgt atatatgtat atatatatat
168121 gtatttatat atatatgtgt atatgtatgt acacattatt taacctacct ctgtgtgagt
168181 gtgtgcata atacacgcac acacacacac acaaataat atatttccct tctgagacaa
168241 agccaaacag cactgtatgc ttaaagaaaa acagtccac ttcccactta tgtaatttat
168301 attacatcca gtcaccacac cagccaaact gctttattgt tttttgtttg acatccaatg
168361 ctaaagcata atgcctgttg cagtgaata tacatgagca acctgagaa ctcaatatag
168421 cctcacgtgt tgccactgag ttgagttgag gagtcaagct gtagcaaaaa ggtttgtcac
168481 cgggtgagta atggtgctct tatttttctc tgggtctcaa gaagtgcctt ttatgacata
168541 tatggcatta aataaatatc agatatttgc acatcctaac tttcctattg gtgaagtttc
168601 ttaaaagaga gataaagggc cattgtgtga ttgatagttt caggatatatt tttgctgcac
168661 agtcagtcct agtgtaccac gtaggggcaa ccacgtaact tctcagggcc ttgactgttt
168721 catttgtaaa ccagagaaaa ggacttgggt gacctccaaa gacctttcaa atttggagat
168781 gagtttgtgg aaagttcaaa cagtttagaa aacagaacta agacacccac tggcaccctt
168841 ggaagcaaga gagtgccagg tactatttgt aatacaggaa tgaaatacct aattgtatga
168901 aattgaattc taactgaacc agtttgttca gttaaatttt ttttttcaat tagagtgcct
168961 acttcagtat ctaacactag acagtaaaact gtagacaaaa gacctacaga atttctgaat
169021 ggtatcaaat tcaccacact taaaactttg ggatgtctaa tttcaaccaa cagctttcct
169081 tcttcataat gttgaatata tgtgtatcta ttttagctaa atttaatata tatcaatata
169141 ctttgataga tatttttatat aaactattag actatagtat tatgagtaaa agaccacca
169201 tttcccaagc aattataaag aacgatcaaa attttaatgg gttgttagta ttatttcttt
169261 aaagattgtg atactgataa atatttggcc acattttaat agaattatac atgggatgtg
169321 tgtgtgtgtg tgtgtgtgtg tatatgtgtg tgtgtatata tatatggcag tagagatata
169381 tatatctaca cacatctaga tatatatata catgtatata tatatataca cacatatatc
169441 tgtgtgtata tatacatatg tatatatacc tacatacata tgtacatata catacatgca
169501 tatatctgta catatatata tagtgtgtgt gtgtgtatat atatatatat atatatattt
169561 ttttttctct gagccaaaac aaaatactag gttgtaatag ctgttctttc agaaggaaga
169621 aaaacaacat gtgctgaact ctgagtttga tgtttttgta ttttacttcc tattttcata
169681 tcagtcattt tattttattc ggaagaattt attgagcata tattatgaac acagcttttg
169741 ctaaggacag ggtatgcagc agttatggcc tagtaggaga tatggatgtt aaaaacaaaa
169801 tgctcacaaa tgcacatata atcttaatac tcattgtaag ctatgaaagc agagtgtgag
169861 tattatgaga ccatatgttg ggagatttta tttggtattg aggatcagga aagatacccc
169921 tgaggaagtg atattttaatt tgaaacctaa agaaagcagt tggccatggg aagaaggtag
169981 ggaatgagat tcccaagcaa taggaatcca atgtgtgaag aagctgaggg agtgaaagaa
170041 agctagtgtg gtggcaggaa gaaagagaag agaatggaga agggcactaa atgagtcaga
170101 gaagtaggag gggctaaacc atgtagggtc gtgtaggcca tcttaaaggc ctgagtgtag
170161 tggaaaacct ttgaaggttt gttaaaagg caatgaaatg ttctaatttc tgttgtagtg
170221 aattgctttg attgctgaat gcgaatggat gggtagagat gcaagagtga aaggaagaa
170281 atcaattagg aggtctcttg cctgctccag ataggactga taattaattt tatttgggaa
170341 gatcagggag aaagataagt catgaatgac tcccaagttt ctggattgaa gaaatgaagg
170401 taccatacac tgagatggga aagcctaggg gtagagtgc tttgagaaga aaggtagcat
170461 ttccccattt cataaaacat ggaagaacaa agaggctgga ttctgtttg tagacatacc

FIG. 1 (cont.)

51/63

170521 ttccaggcca gaactgcatt actacaacat ctttgcaagc cacattgcct ttcataacto
170581 tgtgtcagtg ttgatgccgt aacatctttg gccttccccc taccatcctc ccgcagtcct
170641 ccatgataat gccattattc cgtttcaaat tgtgtgcttc cattggatgt gtgagtctcc
170701 ttgaaagtta taatgaggct gtagcccata tgaaatgctt caactcaggt cctgcatagg
170761 aagaggaagc taatctctcc aggaactgag cctgtggcta gagggatgga taattgttta
170821 aataaagaat atgctgctga gtactgatgg gctctttatg taccattttg gctgctgctg
170881 cccaaccttt aatctttcct gagctttaa taggaaggaa aaaatggctc acaaaggatt
170941 tgagccattt tgotgtggtg atgaggagca cgggtttaga gacaaacact cctgtgtttg
171001 aattccagct cctactatct cctagctaag tgaccttga caagtcactt accttctcca
171061 acctgctgtt tcttcatgta cgtaatagga tttacctcat gaggttgaca tgaagattga
171121 aagaggtaac atatagaatg agcctgtccc aggacatggg tcatgataag tctgccataa
171181 atgggagcta tgtgtcccac ccttttgagg gagataactg ttctgtagca ggtaatatat
171241 tgtttgatac ttggttaacc cttacaatta tcatttcctg ttcttctcaa taatgctaga
171301 aaccttttat ttaaagaacc acaatataaa atgaaaaata tataaaaaaa gcaaatggaa
171361 aaattctatt ggcaaggctt tttaaacttta tatactaaat aaatccaatt gcttaaataa
171421 tgaactgact caagttotca gcactgcttc ttgtttaatt ctctttagtt ttcagaatt
171481 ctccaataat gacctttgtc tactctcttc agtttattca gaaattactt ttatttacat
171541 agaagtttgg aagtggatac acaaacatat ccctcacata tcttatgata ctatgagtca
171601 tatactcatc tcttatattc cctctgtaaa gcaatgtagg tacctttcag gaaggtgatt
171661 tttatgtagg ttgagaaata tcagcatgga ggtcctagct gacctctcta gagagtttct
171721 gagacatttg acaacaactt tttctttaag tcatcagtta tgccccgggg tatgaaattt
171781 ctaacatgat cctcagtaaa cttggctgcc ttgctgagga tactctccat ctgcctgaga
171841 gacacagaca ccattaattg ggaattgact tgacttgtgt ggttccttgt ggaccagatg
171901 gccactaaat attctcattt caaggcaatt ggtaaaaact acacttcaag aaatttcatt
171961 cttaattccc cttagtggat gttattaacc aaaggcaaaa gaaaaaaagg gtaaaaaaaa
172021 tattctaaat gttaatatca aaaatattat tttcaattca cccaggcac agagaactaa
172081 gtattattat tgotattgca ccggcattcc ccaatgagac agtgattttc ttttaagaca
172141 tttttaaata atataggcag aattaagtag acggtgatct ggtaagtaga tgtttcaggg
172201 taacagctgt gcaatgctcc atgcaggga ttagattgtc attttattcc ttaccaggaa
172261 catacattca gttaaacaat tatttgactt ctgctcttcc actgatttct aagttgagggc
172321 tctctcttgt gctgtctga tcagataagt agagttgtgc cttggtttat agatgagata
172381 aatgtgtatt tgaataagca taagttaaag aaattttaaa atcccttagg aagctaggct
172441 tatcagagaa atccaaggaa atacattaac aaactaggaa tttgttctaa caggttaatt
172501 ataactcata aacttattgg gtttttttac cttttaattt tatattacat ttgcttataa
172561 taaggaatat tgotaggaat aaaatttttt aatattctac aattaacaat tatctcaatt
172621 tctttattct aaagacattg ggattagaaa aatgttcaca agggactcca aatattgctg
172681 tagtatttgt ttcttaaaag aatgatacaa agcagacatg ataaaatatt aaaatttgag
172741 agaacttgat ggtaagtaca tgggtgtttc ttattttaaa ataatttttc tacttgaaat
172801 attttacaat acaataaggg aaaaataaaa agttatttaa gttattcata ctttcttctt
172861 cttttctttt ttgotataga aagtatttat ttttctgga acatttagaa aaaacttgga
172921 tccctatgaa cagtggagtg atcaagaaat atggaaagt gcagatgagg taaggctgct
172981 aactgaaatg attttgaaag gggtaactca taccaacaca aatggctgat atagctgaca
173041 tcattctaca cacttttgtt gcattgtatgt gtgtgcacaa ctttaaaatg gagtacccta
173101 acatacctgg agcaacaggt acttttgact ggacctacc ctaactgaaa tgattttgaa
173161 agaggtaact cataccaaca caaatgggtg atatggctaa gatcattcta cacactttgt
173221 gtgcatgtat ttctgtgcac aacttcaaaa tggagtacc taaaatacct ggcgcgacaa
173281 gtacttttga ctgagcctac ttctctctc actggtatgg ctccaaccat caggccctat
173341 cttggtccat ttaggctgct aaaataaaat accaaagact gagctgctta taagcaatct
173401 ttggaggctg agaagtcaaa gatcaagggt ccagcagggt tgotgtctcg tgagagcata
173461 cttcctgggt cattgatggg gctttcttgc tgtgtcctca cataatggaa agggcaagac
173521 ctctctgggt tctcttttac aatggcacta atcccatcat gagggctttg ttctcatgac
173581 ctaatcacct cccacatgtc ctacattcta atactatcac cttgggggtt aggattttaa
173641 catatgaatt tgaggagggt ggggggggga cacaaatatt tagaccatag catttcactc
173701 ctgacctcca aagttcatgt cttcttcaca tgcaaaatac attcattcca tccaatagc
173761 ccccaaagtc ttaacttgtt ccagcatcaa cttacaaggo taaagtccaa ggtttcatct
173821 aaatatcagc taaatcagca caaacagcta aatcaggtag agtgggactt aaggtgtgat
173881 tectctttag gcagattgct ctccaactat gaaattgtga aatcaaacct attatgtact

FIG. 1 (cont.)

52/63

173941 ttcaaaaataa aatgggtgaaa caggcacagg ctagacagtc ccattttcaaa aaagagaaat
174001 agaaaagaaa aaaggagtga cagggtctcta taagtctaaa actttaaggc ttgagaataa
174061 tttgctttgc tttgcctcca ggctcactgg ggtggtgtct tacctctgga cacactgggg
174121 tggaggctct atcctcatgg atttgagtgt ctcatctctt gtggcaggtc tgtgctccaa
174181 tcccacacct atggctccct gagtgtgcaa ttgcatgcct ggtggttcta ctggtctggg
174241 attgcatagg tggcccagcc ttcatagctc cactgggcat tgccctaata tgggctctat
174301 gtggtgacct caccctggg cctctacctg ggccctgtga ctccctgggt tcttgaaatc
174361 taggtggagg cagccatccc cctacagttg tgctgagtgt agtgcatag tgctggggtc
174421 tgctagagct atacctaggg tgggtggagat gtatggcaat ggagtatggg gagctgatat
174481 ggtttgggtg tgtccccacc caaatcttgt cttgaattat aatttccata atctccatgt
174541 gttgagggag ggacctggtg agagggtgact ggatcatggg catgggttttc ccatgctggt
174601 catgtgatag tgagtgagtt ctcacgagat ccaatggttt cataaggcag ttttccctgc
174661 tcttgacccc tctttcttgc ctgtcaccat gtaagacata actctttccc ttcggccatg
174721 attgtaagtt tcctgaggcc ttcccagcca tgtggaactg tgagtcaatt aaacctcttt
174781 tctttataaa ttaccagtc tctttacagc aatgtgaaaa tgtgctaata caggagcaaa
174841 gactgcagtg tgagggtggca atgtgaagtc tgcaatgtga ggtggcacgg ggcagttgta
174901 gcccctcctt tgaaatcttt ctccctacc ccaggcctct gcactctgaa ctatgatggg
174961 aaaggcagct tggaagatct ccaaattggc ttggagtcat tcttccattg tcttgacta
175021 taaattctgg ctctgttta ggtggctgac taatatcccc actgtctgaa tgcatagcac
175081 ctagtctctg ttgagatggc tagtccatag taatttactt atcaaatttg gccacacct
175141 ttgtattctc tcctgagcag gctttctcat ctctcacaat atggataggc tgagaatttt
175201 ccaaattttg aagttctgct tccttttga tcaataatc cattttaaag tcatttctca
175261 tcttgaattt tactatgagc agtcaagagt aactaagctg ctcttcaac tttgcttgga
175321 tatttctctc gtcaaacatt caatttcatt gctttcaagt tctgccttc acaaaacact
175381 aggacacaaa cagctcagcc aagttctttg acattttata agaaggatag ctttctctcc
175441 attgtccaat aacatgttcc tcatttccat ctgaaaaccc atcagattgg cctttaccgt
175501 ccatattttc tgggaacattc tgctcatgac cacttaggta ttcggtgaaga agatagtagc
175561 tttctctata gctctctctc tctctggagc cctcaccaga atggccttta attgtccatt
175621 cacagcaatg taggcttttt ctagcatgta cctgaaaact ctccagcct ctactcatta
175681 ccttgttcca aagctgcttc cacattgagt atttgttaca gcagtaccca gatcccagta
175741 ccaatattct gtcttagtcc attggggcta ctacacgatg tcttataaac aacagtaaaa
175801 tttatttttc acagttgtgg aggctgggaa gttcaaaatc tgggtgccagc agattttgtg
175861 tctggtgaag gccttcttcc tcacagatgg ctgtgttctc actgtgttgt tacatggcag
175921 aagagtgggc aggttagctc tctgggatgt cttttataag ggcagtaatc caaatcatgg
175981 gtttagggta gagccctcat gacctaaatc acctcccaa gggccacact cctaatacca
176041 gcatctttga agttaggatt tcaacatatg actttggcag ggggacagaa gctttcagtt
176101 tatagcaaac cctataggta gcactacttt gtcttttct aatcaatttg cgtcaatgaa
176161 acatgaatta gaagagacct aggcgactcc actatactgg gattattccc agtataaatt
176221 atcatctctc cacaccttct catctactcc ctatctgagt tctgaagctc tccactacaa
176281 gaaggaggct ttggtttgac ttgatatact tctctgggaa acaggtttag cataaaacag
176341 tgatgctcat totagaacac ctgcaaatga caatagtttt ctttcgaagt cgccaggaat
176401 cgtctgcctt tgggtatgtg gctgtgagca ctgccgggca aaatgccata tgacctagat
176461 gaggcataat ccatcctttg aagccattag gacattatat aggaaatata ttaactaaaa
176521 tgggaataaaa ttttctaaat aacaccttat gtttatccaa cagggtggttc attatacttg
176581 agagcattat acagaggaat ttgatgggga ggagagctgg agaaattctc gaaattctgg
176641 gtttcttttaa cagaatactc tagctataaa cttataattt taaaaaataa gcattatatt
176701 aaagaaaagg gaacataaat tattttgttt tattaaactt agtccaaag gtctggattg
176761 tggcagaata ggatcagggg acctaaaatg ttgagcctca aaggctctct tagagaacaa
176821 ctgtattcca ctattagcgc ttttggctct tttagcccaa tttctgttta tcccaaatgt
176881 tcttcccttt tctgccttcc ttcacagtgg acctgccag gagctttgaa atgcctgtga
176941 gtgttaaaca cttaccctatt gagtgcccaa ccttaacatg cccctaataa aatgtactta
177001 gattaaccgt tttcattatc aaagtttctt tattacccaa caaacacagg cgctttaaag
177061 aaaacattaa ctaaattgca agtgacacat ttaagatct ttgatatgac ttcagagaat
177121 gcactatagg aacacaatgc aatgggaggg aaacttggga ggggaagacat tagcctttat
177181 aaaatctgca agtattgcc aatcaaaata aaatttacag gaaagcagga tcataaatat
177241 aatctaaaat cttagaacct gtggttatga ttttaaatac taatacaatg caaaattttt
177301 acctgtttag gtttttattt catcagttca tatttaggta tatactttta ctgttctct

FIG. 1 (cont.)

177361 tttttataat ttaccattca caaagatgat gatgttagtc taactttaat gtcatgagtg
177421 ctttgagtag tagtgctaag tttttgttga gtagtagtgt gcttttttga ttagtagtga
177481 taggtttttg atgagtaagc ctgctagcag catacaaaca aacaagcaag tatcagccta
177541 gagaagcaga aaaggcattt gggtttcaaa gtcacaaggc ctaggcctta gtctaataca
177601 gctgataata caatttgtcc aaacaggaca tttttgggtg tgtcaaacac taaactggac
177661 aggacattat gacaaaagtg caaagcagga ctttccgggg caaaccagga tgtatgtcat
177721 ctccactgagt cctctctttg tccttgccat gactagtatc tctagaggta aatgaacaga
177781 gtaatgacaa atagccagac acctgaatct tatcccaaca gcacctccta cataattccc
177841 cattatccca aatggaaatt aaaaatatat acagtgataa ttccaggcca agaaatgctt
177901 tattttctagc ttggacttgg cttccatgtc cagtgtagaa tcttatcctt gctgatctgg
177961 actgtatctc atgaagccat gacttgtacc tagttactag ctggaaggct tagaacaaaa
178021 gctgggtccag agagcctcct ttttccttat ttccctgggtc cacaccttta ccatggcagt
178081 ctgcctatca tttgatggag gaatttaaag caagtccaag ggaagggaag agagtttcta
178141 aaatctagaa cttggatagt ttaatttacc tatcccaaaa cagcttaggc ccagacagct
178201 tctctccaag attgggtgcca aactgaaatt accagctgtg tagaccaaag agaatttcaa
178261 aagaaaactga atcccaagag aaaaaaaaaa gacttctggc attgtggccc aataaattgg
178321 taggattggt gtgacttttc aagtttacat gtaaaatggg ccagcgcag tgccctggcaa
178381 atatgggtac taagtaaaag taactataat catgtttttt taatctggac ttcacttggg
178441 catcctttta atggtgtctg acagaatcct agttcttgtc tcactttact tagtttccct
178501 gggaaaatttc atgtgtcctt ttggctttaa ttaatatctc tattttgatg acctccatta
178561 totgcctatt ccagagcctt tccacctgat atctcagcac atgaaaagca ccttatgtca
178621 ataagtgagt tccttccctg ccccaccaca tacctgtcct gtgttcccta ttccactgaa
178681 tggcatccca tcctccagtt tcccaaggcc aagacctggg actcatcttt cactctcaag
178741 ttcctccacg ggtaccacaca tgtcacatcc tgtcaatgct gtccctgggg agtatctgaa
178801 atatatccac ttttcttcat ttccacctga caccactatt aacacttgca caaatttctg
178861 aggttccctg ctcatttccc tcattgacct ccaatagttc attctgctct ttgcagctct
178921 ggtgatcttt ccaaacccca catctgatca cttgtttctt cccttcatat ggctccttaa
178981 tgccctcttg actaagtcca cactgcttaa ggtggcttac caggtccttc atgattttgt
179041 ctttgttttg ctttctacac tcactgccc acttcccctt acttcccatg attcagttat
179101 actgaatttc tttgggtctc taaagcacat gtgctttctg ttctgcagag gcttttttgt
179161 tcacttgcta ttctctacct gggaaactcc ccagccctt cactgcctcc ttctaccatc
179221 tttcaggcct ctccttacac atcacttctt tccaaaaatc tgcccttgaca ctccaggctc
179281 cggtttccta ggtgtaccct ataactccac ccctttcata gcattttctc ctctggctgg
179341 agattttacc ttttaacttgt ccatgtcccc cactggagtg gaagttcctg gaggtcaggg
179401 attatatcct attaattggt gtatttccag tgccctagagt agtcttgcac acatggatgg
179461 tattcaataa atattggttg aatgaataag gagttctttc atttcatatg taatagatca
179521 tggaaatagc cttgtgattg atacacagca ggtattacca tcctcacttt agaatgagga
179581 ctgagagcct tgagatgtct gagggccttg actgggacag ctggcagatg caggagcaga
179641 gctgcatcac ccctgtgggc tatctcaggg ttgtctgtaa tctaagtaca atgtctgttg
179701 attttggaact gaaggctttt tgggtaattg tttgcttttt caatacttat aaaatagtth
179761 ccatccttac tcattgatag taaggttagt tatttttagaa aacaagctaa atagcagaaa
179821 tagtggcctt ttaagttgaa aatttaccct gaaaaatcta cagagtagca aacagagtat
179881 caaaaggagt tgactgtatc tatttttata actgccactt atggattatt cagtaaaacc
179941 acaattcact tttatgatth tttttcatgt ttctctgtca caagagcaaaa ctcttgctcc
180001 ataataacat tccagaatac agcaatagca aaagtcaaca ttttgaatcc tttaaaaact
180061 cttagacatt tttttttttt tagtttaaca tgttacaaaa caaaatttct tcttttttca
180121 cagcagtttg ggaagtacat actatthatt agctcatcag catgaagctg gaaaattctt
180181 tttcctaaag ttctttatat ctacaaaactg ttgatgtttt catttattta tttttaatgc
180241 tacgttgtaa tgaaaatcat tggaaaactt tagattctag taattttgaa gtcttcttag
180301 tttggacagg actgagctaa agtttgtact ttttttaatt tattgaaaaa tgggtttctaa
180361 tgatagtatt aacaagatta tattgggggc aggacgcagt ggctcacact tgtaatccta
180421 gcactttggg aggccgaggc ggttggatca cctgaggtca ggagttcaag accagcctgg
180481 ccaacatgta gaaatccctt ctccactaaa atacaaaaat tagctgggca tgggtggcagg
180541 cactgtaatc ccagctactt gggaggctga ggcaggagaa ttgtttgaac ctgggagtcg
180601 gaggttgcag tgagcccaga tgcaccact gcactccagc ctgggcaata gagcaagatt
180661 ctgtctcaaa aaggaagaaa gaaagattat attggggata tatatgtgtg tgtgtgtgtg
180721 tgtgtgtata tacacacaca tatatatata catatataca tatatataca tatttaaagg

FIG. 1 (cont.)

180781 ataaaggatt ctgctgccac agatcactaa atcagatgat ctctagcaat ttcoctgtttg
180841 tttgtttttt gcccatagtg cttatctctt tgaacagtaa tttccactt actatttttc
180901 tccccttttg gaccataatt tcctttaagg cagagcctcc tgttactcat ctttgaatct
180961 ggggtctgtc agagtaccta gaatttaata aactctcatt aagagccagt tgaaagaata
181021 tatgactaag cagtcattta catccaaaag atccgtagga gaattcttat cagcacatgt
181081 gattggtaac aataactttg tacttttcaa aaacaattac taatctatct tgctttccat
181141 tatctcacca aaacctatta gcatgtctgg cagaaaatag atacttaata aatttcttaa
181201 atgtttactg acttcaatth taagttttat taactatgth gacttttctc taatgaagat
181261 gattctaaaa agcttttttac tatacttcac agtgaataaa acagtgagat aggaatattg
181321 caaaatgtcc cctgtgttgg tcagtcttag tgtcattcat tttaaaaatt ctgttctcta
181381 aatattgaca gttatatata aatttatgta attgtttact tctaataaag aatttcatct
181441 ggggaaaaaac atactttgct cagctctttg ccacaagtgc aaagtctaag acagtcaaht
181501 agcttttcta gtacggcctt aggaacttag tatatgactg gtgtgaatct agagggagca
181561 tactgcattc tgaccaaaat ctccaccctg ttaactatggc catcactaac ttcgcagtat
181621 tgcagtactt cctgctagct tagttcccaa ggcaacttgt gaaggaaaat ttttaciaag
181681 ctgttgtcac acaaaggtag tgtttcagth cctgagccca tgtccttgga gttgcccagg
181741 ctccaataat actaataatt actgtacatt aggtacttac catgtgccat attctgtggg
181801 agcgcctttc cacaaattat ctctggtaat ccttgaaca accctttgac atcaatatta
181861 ttattttctc cattttttta catatgagat aaatgagact taaaataatg tgcctgatat
181921 catcagcaaa tgagctgagg agggcagatt caaagctgat tgtgtttgac tctagagctg
181981 cagtcttaag ccagaccttt tcttgctggg taattttact gaaaaaaaaa aaaaaaaaaa
182041 aaaaccctca aatactgctg attgatctaa agtactaaca tttctatcag tgttagggaa
182101 attttaatth tataatthga ttttgtgaga aatttatagc atcttgaata ctccatgca
182161 aagtgatatg tcttagataa cattttacaa tggcagagct taagccagtg ctccagtcatt
182221 cattcatcct caagtthtga ttcatttatc attcatcaaa actctgttht gtttggccac
182281 ccacattcta ggagctcagt acatatttga taaatgaatg aattgttgag gttgacagth
182341 acccaggact ggcattagga acacagagct gaagagcacg tttttaccct caagaagctt
182401 acagtctaac gagggaaact gcacaaatac tactatcact aggtgcctgg ttgaatggct
182461 taagagatga tcagggatat tcagaaggat atgtcaggct cagcaatggc atcacttgag
182521 agcatcaagg tgtttaggga actacaagat gtttggttct gctgggaata agagtgaagg
182581 gggctccatt tggatgcctc atacaccagg tgagagatct tagattttat tccaccagga
182641 ggagaactac cataggatth aaaacagaaa tgatatggct aaacctacat cttaggaaga
182701 tccctggggg gtttgtatgg tggacttgca atthgactaa ttgagatthg taggatgatt
182761 cttaaagagat gatgatgacc cagactggga tccactataat agagttggta aggaggagaa
182821 tgatttaaaa agtagttgga agaattctag ggatggagat aaacatttga aaattattaa
182881 cttataggth gtcatcaata cctgaaaat gactgggact tcagaggaga gtctggagag
182941 ttggaaatga caaagactaa tattcaaggg ggcaggaaga gggagagthg ttcacacatg
183001 acaataggaa gaaatggcca tagagtgtgt ggtttctctc aagccaagga atagatgtht
183061 taagaaagga aaattcttgt ggtgggaagc agtagagatg acagatacac attaatthct
183121 tgagatthct agatgactaa atgggcagat gttgaatgat agctaaagga gaaccagaa
183181 acaaggggagg gattthgtth ttgtthttta aaaaagatag accatagcag ctatcatagac
183241 tgaaacaata aaaaagthga aggcacaaag aaagacacag gtctctaac tccctgccca
183301 gtgcccttht ttcataatth cagcacttgt atthctaatg tttatgtthg agtcttcggg
183361 gatacatcag agtagtcccc cttgtctaat aaatgtgtht acatttcttg ccataccaga
183421 aacccttctc aaactthaat gaatttctac aagggtgagat tactthaatg agaaaccaac
183481 caaggaaagg agtatcatct gcaatatact ttcaaatgth ttttgcttgt ttgtthcttg
183541 tccagctaaa aaaaaaaaaa aaaaacaagc cattggctct aacacaactt tcatattcta
183601 cccaatatc aaagaggctt aaaatctct ggtcgtgtga tgggcacaca gthaattht
183661 tgtgaacaaa cacagtgtta tgggccatth ctgaattht ctctgaaatc ataagattht
183721 ttctgagcca ttatctcatt ctatattaca gtcagggtga gcccatctta cctctcata
183781 ctaaattcta gacttctcaa gggcaggaga caatcatctg tatatctctt tggccttcat
183841 aacttcagga gtacttgcca aaaataaaca tthaatgcac atthattthga ataattgata
183901 agatccaata ctcaataaac tttgtcatat ttttatagaa tgggtthcta tatctcatt
183961 gcattthcaa actthactth tactgtctag cthtaaaaaa aaagcctthg actctaatac
184021 agcctcata tctacccca atatctaaga ggcctthatat ctctagtgt tgtaccacta
184081 tthtaactcc agtattthtt acttcatagt tttacctatt tgttacagth agthtttatg
184141 aattcaagag atgaatagca atthtccata tgtaatthaa aaaaccacc agthgactat

FIG. 1 (cont.)

184201 tttatgctat cttttgtcct cagtcatgac agagtagaag atgggaggta gcaccaagga
184261 tgatgtcata cctccatcct ttatgctaca ttctatcttc tgtctacata agatgtcata
184321 ctagagggca tatctgcaat gtatacatat tatcttttcc agcatgcatt cagttgtggt
184381 ggaataattt atgtacacct ttataaacgc tgagcctcac aagagccatg tgccacgtat
184441 tgttttctta ctactttttg ggatacctgg cactaataag acactcattg aaagtctcct
184501 aatgaatgaa gtacaaagat aaaacaagtt atagactgat tcttttgagc tgtcaagggt
184561 gtaaatagac ttttgctcaa tcaattcaaa tggcggcagg tagtgggggt agagggattg
184621 gtatgaaaaa cataagcttt cagaactcct gtgtttattt ttagaatgtc aactgcttga
184681 gtgtttttaa ctctgtggta tctgaactat cttctctaac tgcaggttgg gctcagatct
184741 gtgatagaac agtttcctgg gaagcttgac tttgtccttg tggatggggg ctgtgtccta
184801 agccatggcc acaagcagtt gatgtgcttg gctagatctg ttctcagtaa ggcgaagatc
184861 ttgctgcttg atgaaccag tgctcatttg gatccagtgt gagtttcaga tgttctgtta
184921 cttaatagca cagtgggaac agaatcatta tgctgcttc atggtgacac atatttctat
184981 taggctgtca tgtctgcgtg tgggggtctc cccaagata tgaaataatt gcccagtgga
185041 aatgagcata aatgcataatt tccttgctaa gagtcttggt ttttcttcct aagatagttt
185101 ttagtttcat acaaactcct ccccttgctc aacacatgat gaagctttta aatacatggg
185161 cctaactctga tccttatgat ttgcctttgt atccattta taccataagc atgtttatag
185221 ccccaaataa agaagtactg gtgattctac ataatgaaaa atgtactcat ttattaaagt
185281 ttctttgaaa tatttgcctt gtttatttat ggatacttag agtctacccc atggttgaaa
185341 agctgattgt ggctaacgct atatcaacat tatgtgaaaa gaacttaag aaataagtaa
185401 tttaaagaga taatagaaca atagacatat tatcaaggta aatacagatc attactgttc
185461 tgtgatatta tgtgtggtat tttctttctt ttctagaaca taccaaataa ttagaagaac
185521 tctaaaacaa gcatttgctg attgcacagt aattctctgt gaacacagga tagaagcaat
185581 gctggaatgc caacaatttt tgggtgagtct ttataacttt acttaagatc tcattgcctt
185641 tgtaattctt gataacaatc tcacatgtga tagttcctgc aaattgcaac aatgtacaag
185701 ttcttttcaa aaatatgtat catacagcca tccagcttta ctcaaatag ctgcacaagt
185761 ttttcacttt gatctgagcc atgtggtgag gttgaaatat agtaaatcta aaatggcagc
185821 atattactaa gttatgttta taaataggat atataactt tttgagccct ttatttgggg
185881 accaagtcac acaaaatact ctactgttta agattttaaa aaaggctcct gtgattcttt
185941 caataactaa atgtcccatg gatgtggtct gggacaggcc tagttgtctt acagtctgat
186001 ttatggtatt aatgacaaag ttgagaggca catttcattt ttctagccat gatttgggtt
186061 caggtagtac cttttctcaac caccttctca ctgttcttaa aaaaactgtc acatggccag
186121 gcacagtggc ttacatctgt aatcccaata ctttgggagg ctgaggtggg gggattactt
186181 gaggccagga attcaagacc agcccaggca acatagttag gcccacatctg tctttattaa
186241 aacaaaacaa aactgtcaca gcttctttca agtgatgttt acaaatccct tatggtttag
186301 tcacaaggaa gttctgagga tgatgtatca cgtcatttct gttcaggctt ttgagcctcc
186361 tggaggtaaa tgggttcctt actgaaggct tgttattacc atgattatca ctaagcttga
186421 agtaacaaat taggggggca gactcacaac ctcttgccct gccatggaca agttcaagaa
186481 tctaagtaaa gtectctatt gtctgatctt ggatttgctc aacctgaaca agccaaggag
186541 gtgtattaaa ctcaggcaca tcctgaccaa tttggaattc ttaagcttca gatcactgtg
186601 gaagaggctc aactctttat ggtgctgtag acttacgctc attttctagg taatttataa
186661 gggacctaat attttgtttt caaagcaact tcagttctac taaacctccc tgaagaatct
186721 tccagctgct gagtagaaaa tcacaactaa tttcacagat ggtagaacct ccttagagca
186781 aaaggacaca gcagttaaat gtgacatacc tgattgttca aaatgcaagg ctctggacat
186841 tgcattcttt gacttttatt ttcccttgag cctgtgccag tttctgtccc tgctctggct
186901 tgacctgctt tctgtcccag atctcaacta cagccatttc cctaggtcat agaagagaac
186961 aaagtgcggc agtacgattc catccagaaa ctgctgaacg agaggagcct cttccggcaa
187021 gccatcagcc cctccgacag ggtgaagctc tttcccccac ggaactcaag caagtgcaag
187081 tctaagcccc agattgctgc tctgaaagag gagacagaag aagaggtgca agatacaagg
187141 ctttagagag cagcataaat gttgacatgg gacatttgct catggaattg gagctcgtgg
187201 gacagtcacc tcatggaatt ggagctcgtg gaacagttac ctctgcctca gaaaacaagg
187261 atgaattaag ttttttttta aaaaagaaac atttggtgag ggggaattgag gacactgata
187321 tgggtcttga taaatggctt cctggcaata gtcaaattgt gtgaaaggta cttcaaatcc
187381 ttgaagattt accactgtg ttttgcaagc cagattttcc tgaaaacct tgccatgtgc
187441 tagtaattgg aaaggcagct ctaaatgtca atcagcctag ttgatcagct tattgtctag
187501 tgaaactcgt taatttgtag tgttgagaa gaactgaaat catacttctt agggttatga
187561 ttaagtaatg ataactggaa acttcagcgg tttatataag cttgtattcc tttttctctc

FIG. 1 (cont.)

WO 2011/119531

PCT/US2011/029314

56/63

187621 ctctcccat gatgtttaga aacacaacta tattgtttgc taagcattcc aactatctca
187681 tttccaagca agtattagaa taccacagga accacaagac tgcacatcaa aatatgcccc
187741 attcaacatc tagtgagcag tcaggaaaga gaacttccag atcctggaaa tcagggttag
187801 tattgtccag gtctaccaa aatctcaata tttcagataa tcacaatata tcccttacct
187861 gggaaagggc tgttataatc tttcacaggg gacaggatgg ttcccttgat gaagaagttg
187921 atatgccttt tcccaactcc agaaagtgac aagctcacag acctttgaac tagagtttag
187981 ctggaaaagt atgttagtgc aaattgtcac aggacagccc ttctttccac agaagctcca
188041 ggtagagggg gtgtaagtag ataggccatg ggcactgtgg gtagacacac atgaagtcca
188101 agcattttaga tgtatagggt gatggtggtg tgttttcagg ctagatgtat gtacttcatg
188161 ctgtctacac taagagagaa tgagagacac actgaagaag caccaatcat gaattagttt
188221 tatatgcttc tgttttataa ttttgtgaag caaaattttt tctctaggaa atatttattt
188281 taataatggt tcaaacatat ataacaatgc tgtattttta aagaatgatt atgaattaca
188341 tttgtataaa ataattttta tatttgaaat attgactttt tatggcacta gtatttctat
188401 gaaatattat gttaaaactg ggacagggga gaacctaggg tgatattaac cagggggccat
188461 gaatcacctt ttggtctgga ggggaagcctt ggggctgatg cagttgttgc ccacagctgt
188521 atgattccca gccagcacag cctcttagat gcagttctga agaagatggg accaccagtc
188581 tgactgtttc catcaagggg aactgcctt ctcaactcca aactgactct taagaagact
188641 gcattatatt tattactgta agaaaatata acttgtcaat aaaatccata catttgtgtg
188701 aaa

FIG. 1 (cont.)

WO 2011/119531

PCT/US2011/029314

57/63

1	aattggaagc	aaatgacatc	acagcaggtc	agagaaaaag	ggttgagcgg	caggcaccca
61	gagtagtagg	tctttggcat	taggagcttg	agcccagacg	gccctagcag	ggaccccagc
121	gcccgagaga	ccatgcagag	gtgcctctg	gaaaaggcca	gcgttgtctc	caaacttttt
181	ttcagctgga	ccagaccaat	tttgaggaaa	ggatacagac	agcgcttggg	attgtcagac
241	atataccaaa	tcccttctgt	tgattctgct	gacaatctat	ctgaaaaatt	ggaaagagaa
301	tgggatagag	agctggcttc	aaagaaaaat	cctaaactca	ttaatgccct	tcggcgatgt
361	ttttctgga	gatttatgtt	ctatggaatc	tttttatatt	taggggaagt	caccaaagca
421	gtacagcctc	tcttactggg	aagaatcata	gcttcctatg	acccggataa	caaggaggaa
481	cgctctatcg	cgatttatct	aggcataggc	ttatgccttc	tctttattgt	gaggacactg
541	ctcctacacc	cagccatttt	tggccttcac	cacattggaa	tgcagatgag	aatagctatg
601	tttagtttga	tttataagaa	gactttaaag	ctgtcaagcc	gtgttctaga	taaaataagt
661	attggacaac	ttgttagtct	cctttccaac	aacctgaaca	aatttgatga	aggacttgca
721	ttggcacatt	tcgtgtggat	cgctcctttg	caagtggcac	tctcatggg	gctaactctg
781	gagttgttac	aggcgtctgc	ctctgtgga	cttggtttcc	tgatagtect	tgcccttttt
841	caggctgggc	tagggagaat	gatgatgaag	tacagagatc	agagagctgg	gaagatcagt
901	gaaagacttg	tgattacctc	agaaatgatt	gaaaatatcc	aatctgttaa	ggcatactgc
961	tgggaagaag	caatggaaaa	aatgattgaa	aacttaagac	aaacagaact	gaaactgact
1021	cggaaaggcag	cctatgtgag	atacttcaat	agctcagcct	tcttctctc	agggttcttt
1081	gtggtgtttt	tatctgtgct	tccttatgca	ctaatacaag	gaatcatcct	ccggaaaata
1141	ttcaccacca	tctcattctg	cattgttctg	cgcatggcgg	tcactcggca	atttccttgg
1201	gctgtacaaa	catggtatga	ctctcttgga	gcaataaaca	aaatacagga	tttcttacia
1261	aagcaagaat	ataagacatt	ggaatataac	ttaacgacta	cagaagtagt	gatggagaat
1321	gtaacagcct	tctgggagga	gggatttggg	gaattatttg	agaaagcaaa	acaaaacaat
1381	aacaatagaa	aaacttctaa	tggtgatgac	agcctcttct	tcagtaattt	ctcacttctt
1441	ggtactcctg	tctgaaaga	tattaatttc	aagatagaaa	gaggacagtt	gttggcggtt
1501	gctggatcca	ctggagcagg	caagacttca	cttctaata	tgattatggg	agaactggag
1561	ccttcagagg	gtaaaattaa	gcacagtgga	agaatttcac	tctgttctca	gttttcttgg
1621	attatgcctg	gcaccattaa	agaaaatac	atctttggtg	tttctatga	tgaatataga
1681	tacagaagcg	tcataaaagc	atgccaaacta	gaagaggaca	tctccaagtt	tcagagaaaa
1741	gacaatatag	ttcttgga	aggtggaatc	acactgagt	gaggtcaacg	agcaagaatt
1801	tctttagcaa	gagcagtata	caaagatgct	gatttgtatt	tattagactc	tcttttggga
1861	tacctagatg	ttttaacaga	aaaagaaata	tttgaaagct	gtgtctgtaa	actgatggct
1921	aacaaaacta	ggattttggt	cacttctaaa	atggaacatt	taaagaaagc	tgacaaaata
1981	ttaattttgc	atgaaggtag	cagctatttt	tatgggacat	tttcagaact	ccaaaatcta
2041	cagccagact	ttagctcaaa	actcatggga	tgtgattctt	tcgaccaatt	tagtgcagaa
2101	agaagaaatt	caatcctaac	tgagacctta	caccgtttct	cattagaagg	agatgctcct
2161	gtctcctgga	cagaaacaaa	aaaacaatct	tttaaacaga	ctggagagtt	tggggaaaaa
2221	aggaagaatt	ctattctcaa	tccaatcaac	tctatacgaa	aattttccat	tgtgcaaaag
2281	actcccttac	aatgaatgg	catcgaagag	gattctgatg	agcctttaga	gagaaggctg
2341	tccttagtac	cagattctga	gcagggagag	gcgatactgc	ctcgcacag	cgtgatcagc
2401	actggcccca	cgttcaggc	acgaaggagg	cagtctgtcc	tgaacctgat	gacacactca
2461	gttaaccaag	gtcagaacat	tcaccgaaag	acaacagcat	ccacacgaaa	agtgtcactg
2521	gcccctcagg	caaacttgac	tgaactggat	atatattcaa	gaaggttatc	tcaagaaact
2581	ggcttgga	taagtgaaga	aattaacgaa	gaagacttaa	aggagtgcct	ttttgatgat

FIG. 2

58/63

2641	atggagagca	taccagcagt	gactacatgg	aacacatacc	ttcgatatat	tactgtccac
2701	aagagcttaa	ttttgtgct	aatttgggtgc	ttagtaattt	ttctggcaga	ggtggctgct
2761	tctttggttg	tgctgtggct	ccttggaaac	actcctcttc	aagacaaagg	gaatagtact
2821	catagtagaa	ataacagcta	tgcagtgatt	atcaccagca	ccagttcgta	ttatgtgttt
2881	tacatttacg	tgggagtagc	cgacactttg	cttgctatgg	gattcttcag	aggtctacca
2941	ctggtgcata	ctctaatac	agtgtcgaaa	attttacacc	acaaaatgtt	acattctgtt
3001	cttcaagcac	ctatgtcaac	cctcaacacg	ttgaaagcag	gtgggattct	taatagattc
3061	tccaaagata	tagcaatttt	ggatgacctt	ctgcctctta	ccatatttga	cttcateccag
3121	ttgttattaa	ttgtgattgg	agctatagca	gttgctgcag	ttttacaacc	ctacatcttt
3181	gttgcaacag	tgccagtgat	agtggctttt	attatgttga	gagcataattt	cctccaaacc
3241	tcacagcaac	tcaaacaact	ggaatctgaa	ggcaggagtc	caattttcac	tcattctgtt
3301	acaagcttaa	aaggactatg	gacacttcgt	gccttcggac	ggcagcctta	ctttgaaact
3361	ctgttcacac	aagctctgaa	tttatact	gccaaactgt	tctgtacct	gtcaacactg
3421	cgctggttcc	aaatgagaat	agaaatgatt	tttgtcatct	tcttcattgc	tgttaccttc
3481	atttccattt	taacaacagg	agaaggagaa	ggaagagttg	gtattatcct	gacttagcc
3541	atgaatatca	tgagtacatt	gcagtgggct	gtaaaactcca	gcatagatgt	ggatagcttg
3601	atgcatctg	tgagccgagt	ctttaagtgc	attgacatgc	caacagaagg	taaacctacc
3661	aagtcaacca	aaccatacaa	gaatggccaa	ctctcgaaag	ttatgattat	tgagaattca
3721	cacgtgaaga	aagatgacat	ctggccctca	ggggggccaaa	tgactgtcaa	agatctcaca
3781	gcaaaataca	cagaaggtgg	aaatgccata	ttagagaaca	tttcttctc	aataagtcct
3841	ggccagaggg	tgggcctctt	gggaagaact	ggatcaggga	agagtacttt	gttatcagct
3901	tttttgagac	tactgaacac	tgaaggagaa	atccagatcg	atggtgtgtc	ttgggattca
3961	ataactttgc	aacagtggag	gaaagccttt	ggagtgatac	cacagaaagt	atttattttt
4021	tctggaacat	ttagaaaaaa	cttggatccc	tatgaacagt	ggagtgatca	agaaatatgg
4081	aaagttgcag	atgaggttgg	gctcagatct	gtgatagaac	agtttctctg	gaagcttgac
4141	tttgtccttg	tggatggggg	ctgtgtccta	agccatggcc	acaagcagtt	gatgtgcttg
4201	gctagatctg	ttctcagtaa	ggcgaaagtc	ttgtgcttg	atgaaccag	tgctcatttg
4261	gatccagtaa	cataccaaat	aattagaaga	actctaaaac	aagcatttgc	tgattgcaca
4321	gtaattctct	gtgaacacag	gatagaagca	atgctggaat	gccacaactt	tttggtcata
4381	gaagagaaca	aagtgcggca	gtacgattec	atccagaaac	tgctgaacga	gaggagcctc
4441	ttccggcaag	ccatcagccc	ctccgacagg	gtgaagctct	ttccccaccg	gaactcaagc
4501	aagtgcaagt	ctaagcccca	gattgtctgt	ctgaaagagg	agacagaaga	agaggtgcaa
4561	gatacaaggc	tttagagagc	agcataaatg	ttgacatggg	acatttgctc	atggaattgg
4621	agctcgtggg	acagtcacct	catggaattg	gagctcgtgg	aacagttacc	tctgcctcag
4681	aaaacaagga	tgaattaagt	tttttttaa	aaaagaaaca	tttggttaagg	ggaattgagg
4741	acactgatat	gggtcttgat	aaatggcttc	ctggcaatag	tcaaattgtg	tgaaagggtac
4801	ttcaaatect	tgaagattta	ccacttggt	tttgcaagcc	agattttcct	gaaaaccctt
4861	gcatgtgct	agtaattgga	aaggcagctc	taaatgtcaa	tcagcctagt	tgatcagctt
4921	attgtctagt	gaaactcggt	aattttagt	gttggaagaag	aactgaaatc	atacttctta
4981	gggttatgat	taagtaatga	taactggaaa	cttcagcggg	ttatataagc	ttgtattcct
5041	ttttctctcc	tctccccatg	atgtttagaa	acacaactat	attgtttgct	aagcattcca
5101	actatctcat	ttccaagcaa	gtattagaat	accacaggaa	ccacaagact	gcacatcaaa
5161	atatgccccca	ttcaacatct	agtgagcagt	caggaaaagag	aacttccaga	tccttgaaat
5221	cagggttagt	attgtccagg	tctacaaaaa	atctcaatat	ttcagataat	cacaatacat
5281	cccttacctg	ggaaagggct	gttataatct	ttcacagggg	acaggatggt	tccttgatg

FIG. 2 (cont.)

WO 2011/119531

PCT/US2011/029314

59/63

5341	aagaagtga	tatgcctttt	cccaactcca	gaaagtgaca	agctcacaga	cctttgaact
5401	agagtttagc	tggaagaaagta	tgtagtgca	aattgtcaca	ggacagccct	tctttccaca
5461	gaagctccag	gtagagggtg	tgtaagtaga	taggccatgg	gcactgtggg	tagacacaca
5521	tgaagtccaa	gcatttagat	gtataggttg	atggtggtat	gttttcaggc	tagatgtatg
5581	tacttcatgc	tgtctacact	aagagagaat	gagagacaca	ctgaagaagc	accaatcatg
5641	aattagtttt	atatgcttct	gttttataat	tttgtgaagc	aaaatttttt	ctctaggaaa
5701	tatttatttt	aataatgttt	caaacatata	ttacaatgct	gtattttaa	agaatgatta
5761	tgaattacat	ttgtataaaa	taatttttat	atttgaaata	ttgacttttt	atggcactag
5821	tatttttatg	aaatattatg	ttaaaactgg	gacaggggag	aacctagggt	gatattaacc
5881	agggggccatg	aatcaccttt	tggtctggag	ggaagccttg	gggctgatcg	agttgttgcc
5941	cacagctgta	tgattcccag	ccagacacag	cctcttagat	gcagttctga	agaagatggt
6001	accaccagtc	tgactgtttc	catcaagggt	acactgcctt	ctcaactcca	aactgactct
6061	taagaagact	gcattatatt	tattactgta	agaaaatatc	acttgtcaat	aaaatccata
6121	catttgtgta					

FIG. 2 (cont.)

1	mqrsplekas	vvsklffswt	rpilrkgyrq	rlslsdiyqi	psvdsadnls	eklerewdre
61	laskknpkli	nalrrcffwr	fmfygiflyl	gevtkavqpl	llgriiasyd	pdnkeersia
121	iylgiglc11	fivrtlllhp	aifglhhigm	qmriamfsli	ykktklssr	vldkisigql
181	vsllsnnlnk	fdeglalahf	vwiaplqval	lmgliwellq	asafcglgfl	ivlalfgagl
241	grmmmkyrdq	ragkiserlv	itsemieniq	svkaycweea	mekmienlrq	telkltrkaa
301	yvryfnssaf	ffsgffvvl	svlpyalikg	iilrkiftti	sfcivlrnav	trqfpwavqt
361	wydslgaink	iqdflqkqey	ktleynlttt	evvmenvtaf	weegfgelfe	kakqnnnnrk
421	tsngddslff	snfsltgtpv	lkdinfkier	gqllavagst	gagktsllmm	imgelepseg
481	kikhsgrisf	csqfswimpq	tikenifgv	sydecyrysv	ikacqlcedi	skfaekdniv
541	lgeggitlsg	gqrarislar	avykdadlyl	ldspfgyldv	ltekeifesc	vcklmanktr
601	ilvtskmehl	kkadkililh	egssyfygtf	selqnlqpdf	ssklmgcdsf	dqfsaerrns
661	iltetlhrfs	legdapvswt	etkkqsfkqt	gefgekrkns	ilnpinsirk	fsivqktp1q
721	mngieedsde	plerrls1vp	dseqgeailp	risvistgpt	lqarrqsvl	nlmthsvnqg
781	qnihrkttas	trkvslapqa	nlte1diysr	rlsqtglei	scceinedlk	ccffddmesi
841	pavttwntyl	ryitvhksli	fvliwclvif	laevaaslvv	lwl1gntplq	dkgnsthsrn
901	nsyaviitst	ssyyvfyyv	gvadtllamg	ffrglplvht	litvskilhh	kmlhsvlqap
961	mstlntl1kag	gilnrfskdi	ailddl1plt	ifdfiqllli	vigaiavvav	lppyifvatv
1021	pvivafimlr	ayflqtsqq1	kqlesegrsp	ifthlvtslk	glwtlrafr	qpyfetlfhk
1081	aln1htanwf	lylstlrwfq	mriemifvif	fiavtfisil	ttgegegrvg	iiltlamnim
1141	st1qwavnss	idvds1mrsv	srvfkfidmp	tegkptkstk	pykngqlskv	miienshvkk
1201	ddiwpsggqm	tvkdltakyt	eggnaileni	sfsispqgrv	gllgrtgsgk	stllsaflrl
1261	Integeigid	gvswdsit1q	qwrkafgvip	qkvfifsgtf	rknldpyeqw	sdqciwkvad
1321	evglrsvieq	fpgkl1dfvlv	dggcvlshgh	kqlmclarsv	lskakillld	epsahldpvt
1381	yqi1rrtlkq	afadctvile	ehrieamlec	qqflvieenk	vrqydsiqkl	lnerslfrqa
1441	ispsdrvk1f	phmsskcks	kpqiaalkee	teeevqdt1rl		

FIG. 3

61/63

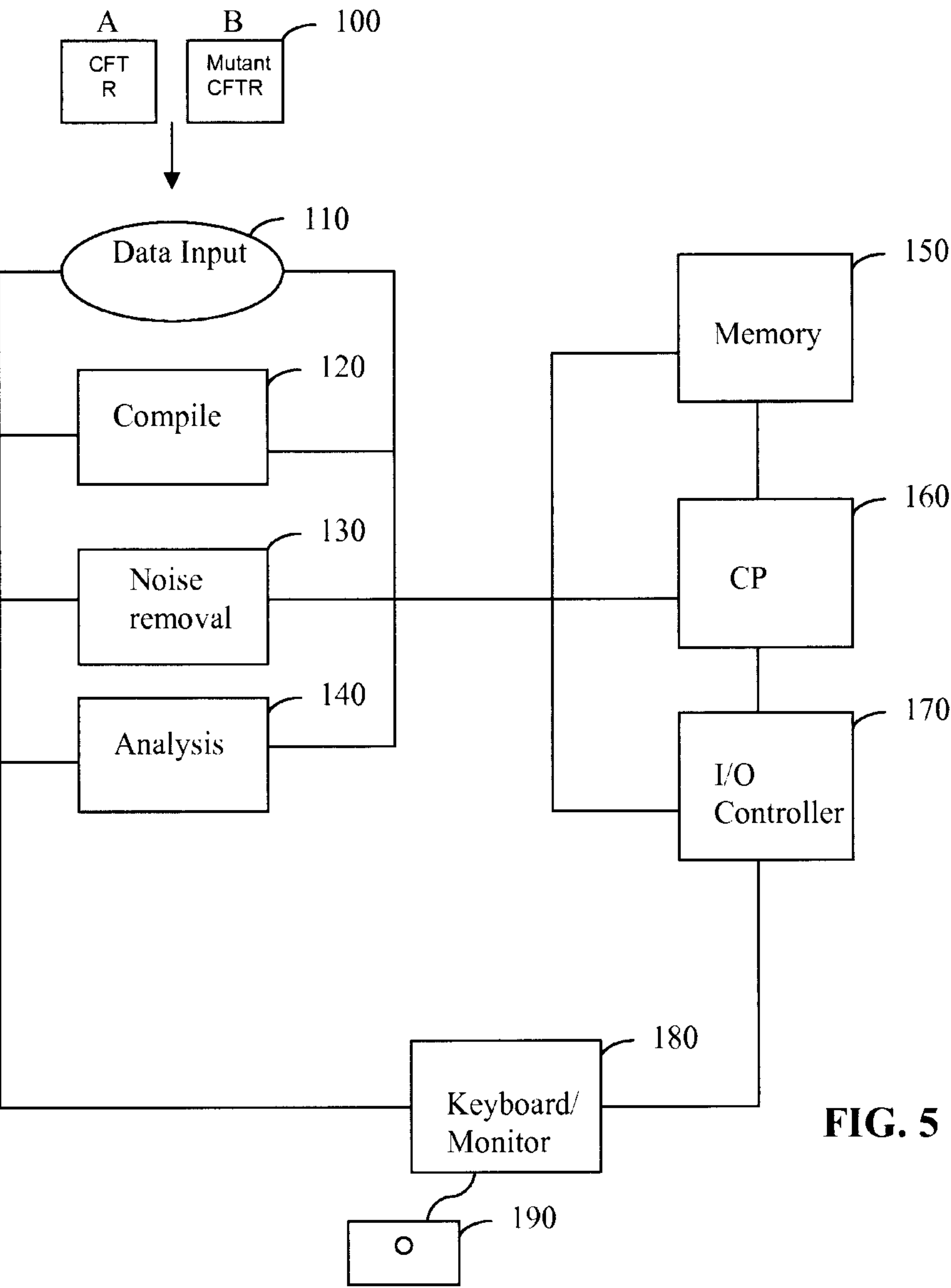
1 gaattcaaag gaaaacataa gatgcaatc gtgcctcaa ggaggttgta gggaagaggg
 61 gttatgaatg tatgtaaata gaagttgggtg tgcgtgtgtg ttataaaca gaattgtcag
 121 accaaacatt attttggaag cagtaaaagt aaactagaat ctggcctagt catgtcccag
 181 gacacctctt tcaagtcctg aaacatcttt gtaagactgt aatgtgtgtt tacatcctag
 241 gtaatcactg tggccactg ttgaagagct gtggctgttc ttaccctct agttagataa
 301 acttataagc acaaccagac tacatatatg aagctgaaga gacctgtct tttttaacg
 361 agcttttctt cccgatagga gtgactattt ctttctct tccacattt caggtttag
 421 tgtactgtg attgctacc acttatcact attaaagtct actcaggaga gaatctgaga
 481 aacactctca aattaagttg aacatgatgg ataagtaaag tattgtgaaa gtccactctc
 541 atgatttcta atggtgaaac ctggcagggt gactaatctt tgacgagaag gttatcactt
 601 ataatcttc atatattgag atcatttgta agaagcacc agcacattgc tgaacacaaa
 661 gtaggtatta aataaatgtt ggcttctt tctctactc atctctgctc ttcttttaa
 721 tataccttta aaatgatgcc acagaaatgg ccaccaatc ttctatatt aaggtcagtt
 781 cttgcattag gaaattctat aggggaagta tgtgaagtat gtgtagtcag tcattaaatg
 841 cttgggctct ggccacagat tgttaggtt taaatcccag ttctcttt tattattaat
 901 tgtgcaactt gcttgggaaa acatgaaact tgttttct caggttcatt atctgtaata
 961 tatagtgaat gaagaagttt cctgtccat gaaggtgttg taaagattaa aaaaggcaaa
 1021 ttaggtgtg tattgtcat aataattggc atatatgga agtgaccaac aaccataagg
 1081 tattataaaa ttgtataaa atgatatgag ctatcattga gcagcatgaa agaagagctt
 1141 cactgttca ctactatca ccctggccca ttaatctct tctgttct gacattcag
 1201 agatacgtt aggatttcaa tcatgacct aagccacatt tgaacaatt tctggtggat
 1261 aagtcctcat tcccacatta tgtatgtacc tagatgcaa tctgaatat catgtcgcaa
 1321 ttagtgcatc tggacatgct tgctaactgt gtaaaagctc tgaataatgg taaagtttta
 1381 ttttaccaa aacaaatttg ggctgtaatg tttatgata aaaatctgtg gtcttctat
 1441 gtacatgtgt gtgtacatgc taaaatgca atgttatagt taaatgta tcatataaag
 1501 tatgtaactc cagtggctac ttagtgtgc tacttggtt gtagattct gcttctgt
 1561 ttcattgtta aacaggtcta gaagttatta ttcatgaaa ctaatgtgag gaaaaagact
 1621 atgtgatata ataagtgaca ttatataaat acatgaggga tgattgatt agaagcagta
 1681 ttacacagt ataggagtaa tggtttagaa ctgactcag gttgaatct tagctctatc
 1741 attataggca ttacttaac tttctgtt tgettaactg aaaactgaag ataataacac
 1801 ctattfacat ggttggtata aggttatat gaataatgc tggcaatag taagaactca
 1861 agtaactgt tctctctc cagaaggaga ttggctgaaa aatattgga gtctcctcca
 1921 gccatattcc ttggtcagct tctatgacc ttttgagc ttaattct atcccttat
 1981 ttctactgc ttgtgataa caaagaagaa ctaattatta attatttca aaatgcatgt
 2041 attatattg atgggccaca ctaacagtta taaacaaac aacagattgg gaatggggaa
 2101 gtgatgtgg tgagttcaat cacatgtctg ggaaaagtca atagtgaaga cagagtctca
 2161 caatttttg tcataatgga gagatgaaa cacaggtaga ggattcaaa caacagagt
 2221 gatggtgagt taaaatgct gaaattctt cctggtgtct aactaatgc aatgtgttt
 2281 atcttttgc ttttctct actattcaaa ttaggataa taaagattaa atgttctaa
 2341 atcttactt acaatatcaa gaaaaaagg tatgttttg cccacggaag ggcaaagcag
 2401 agctatgaaa acctgtgaa cacattctt atttcaaca caggtcttg tcttccatc
 2461 atgaaatgca cttttattt gtactgtatt tgggtgacca caagtcaaca acaagataat
 2521 tcacaagacc ctgccttag atgtgtcggc aataaagtaa tcaggccaaa attttactt
 2581 tctttgaat tttcaattc aaacacaatg tatgttctt ttacacagt aggggtcagg
 2641 gattagaggg ttggtcctt taaaaccgtc agagacacag gcaatctac aaaaattct
 2701 cagaaggaag gcgcctacgc ctgggaatgc ccagatgcc ctcagagagt tgaagatggc

FIG. 4

62/63

2761 gtttctctga gtcagggtcaa agttaacaca ttaccttcgc tcaaagact gcttggttc
2821 ctttcggtgg attagtcaag atgttttgc tactgagact aggaaatcta taggagggcg
2881 ggtagttta cattgttct tgcattatc gctaaaacac tcaaagcct tccttaaaa
2941 tgcgcactgg gctaaaaagg atagacaagg aacacatcct gggccggtaa ttacgcaaag
3001 cattatctcc tctacctcc ttcagattt tttttctct ttacgtacgt gtcctaagat
3061 ttctgtcca ccttggagt tctctacct aaacctgaaa ctaataaagc ttggttctt
3121 tctccgacac gcaaaggaag cgctaaggta aatgcatcag acccactg ccgcggaact
3181 ttteggctct ctaaggctgt atttgatat acgaaaggca catttctct ccctttcaa
3241 aatgcacctt gcaaacgtaa caggaaaccg actaggatca tcgggaaaag gaggaggagg
3301 aggaaggcag gctccgggga agctggtggc agcgggtcct ggtcttggc gacctgacg
3361 cgaaggaggg tctaggaagc tctccgggga gccggtctc ccgccgttg cttctctgt
3421 cctccagcgt tgccaactgg acctaaagag aggcgcgcac tgcgccac ctgcgggatg
3481 ggcctggtgc tgggcggtca ggacctgac ctggaaggag cgcgcgcgag ggaggaggc
3541 tgggagtcag aatcgggaaa gggaggtgcg ggcggcgag ggagcgaagg aggagaggag
3601 gaaggagcgg gaggggtgct ggcgggggtg cgtagtgggt ggagaaagcc gctagagcaa
3661 atttggggcc ggaccaggca gactcggct ttaacctgg gcagtgaagg cgggggaaag
3721 agcaaaagga aggggtggtg tgcggagtag ggtgggtgg ggggaattgg aagcaaatga
3781 catcacagca ggtcagagaa aaagggttga gcggcaggca ccagagtag taggtcttg
3841 gcattaggag cttgagccca gacggcccta gcagggaccc cagcggcga gagaccatgc
3901 agaggtcgcc tctggaaaag gccagcgtt tctccaaact ttttctagg tgagaagggtg
3961 gccaacgag cttcggaag acacgtgcc acgaaagagg agggcgtgtg tatgggttg
4021 gtttgggta aaggaataag cagttttta aaagatgcgc tatcattcat tgtttgaaa
4081 gaaaatgtgg gtattgtaga ataaacaga aagcattaag aagagatgga agaataact
4141 gaagctgatt gaatagagag ccacatctac ttgcaactga aaagtagaa tctcaagact
4201 caagtacgt actatgcact tgtttattt cattttcta agaaactaaa aatactgtt
4261 aataagtacc taagtatggt ttattggtt tccccctca tgccttgac acttgattgt
4321 cttcttgca catcacggtg ccatgcctgc atatagtaag tgctcagaaa acattcttg
4381 actgaattc

FIG. 4 (cont.)



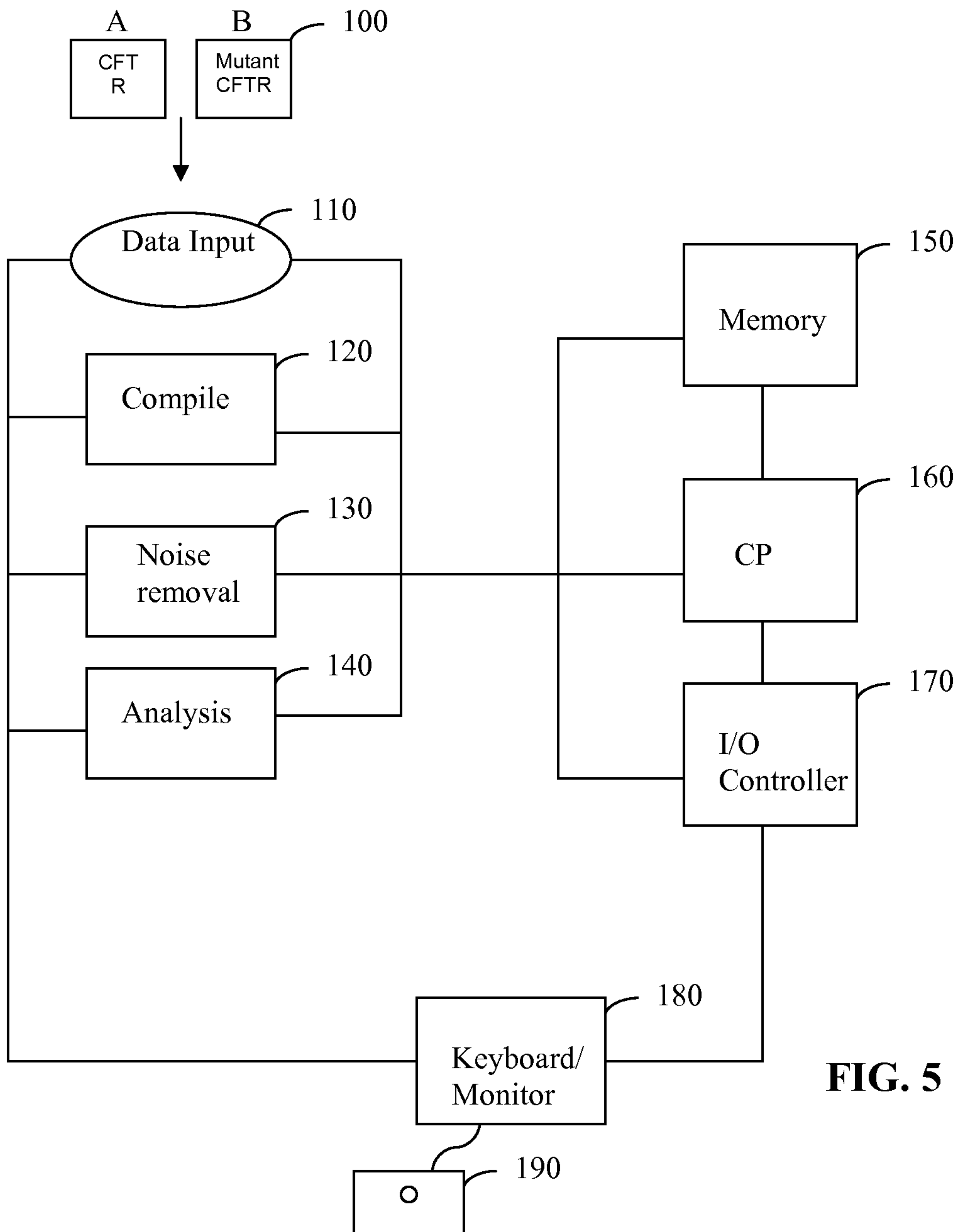


FIG. 5