

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 March 2007 (22.03.2007)

PCT

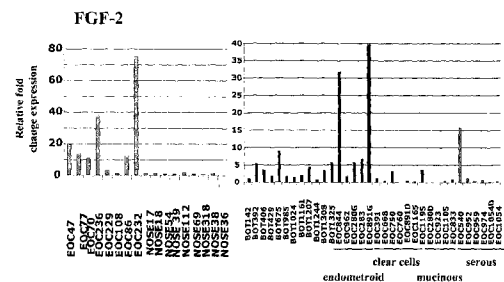
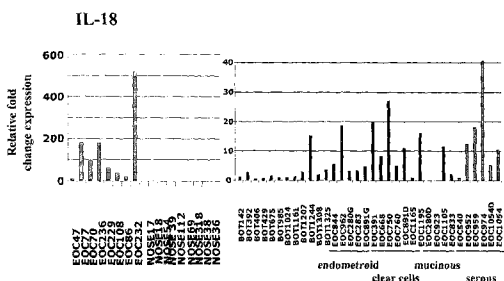
(10) International Publication Number
WO 2007/030949 A2

- (51) International Patent Classification:
C12Q 1/68 (2006.01) G01N 33/574 (2006.01)
C12Q 1/04 (2006.01) G01N 33/53 (2006.01)
- (21) International Application Number:
PCT/CA2006/001536
- (22) International Filing Date:
15 September 2006 (15.09.2006)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/716,941 15 September 2005 (15.09.2005) US
- (71) Applicants (for all designated States except US): VAL-
CHUM, S.E.C. [CA/CA]; 69, rue Sherbrooke Ouest, Mon-
treal, Quebec H2X 1X2 (CA). MCGILL UNIVERSITY
[CA/CA]; 845 Sherbrooke Street West, Montreal, Quebec
H3A 2T5 (CA).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): LE PAGE, Cécile
[FR/CA]; 75 Chemin de la Pointe Sud, Montréal (Québec),

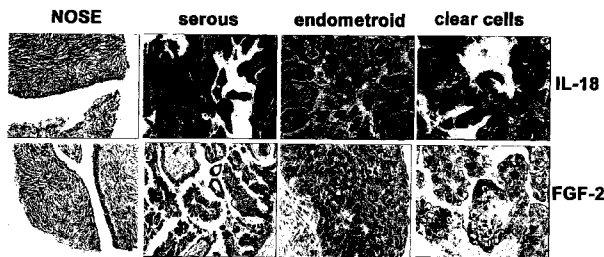
- H3E 1Z9 (CA). MES-MASSON, Anne-Marie [CA/CA];
260 Biscaye, Dollard-des-Ormeaux (Québec), H9H 1R8
(CA). PROVENCHER, Diane [CA/CA]; 75 Principale,
C.P.1160, St-Basile-le-Grand (Québec), J3N 1M6 (CA).
TONIN, Patricia [CA/CA]; 3544 Belmore Avenue,
Montréal (Québec), H4B 2B9 (CA). HUDSON, Thomas
[CA/CA]; 3561 Marlowe, Montréal (Québec), H4A 3L8
(CA).
- (74) Agents: DUBUC, J. et al.; GOUDREAU GAGE DUBUC,
2200-2000 Mc Gill College Avenue, Montréal, Québec
H3A 3H3 (CA).
- (81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP,
KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT,
LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU,
SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on next page]

(54) Title: METHODS OF DIAGNOSING OVARIAN CANCER AND KITS THEREFOR



(57) Abstract: A method comprising: providing a biological sample from a subject (subject sample); and detecting the expression level of each of the markers FGF-2 and CA125 in the subject sample. A kit comprising means for detection of an expression level of each of markers CA125 and FGF-2 in a biological sample from a subject (subject sample), and instructions to use said markers in a method of the present invention.



WO 2007/030949 A2



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

Published:

— *without international search report and to be republished upon receipt of that report*

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

TITLE OF THE INVENTION

[0001] METHODS OF DIAGNOSING OVARIAN CANCER AND KITS THEREFOR

CROSS REFERENCE TO RELATED APPLICATIONS

[0002] This application claims the benefit of United States Provisional Patent Application No. 60/716,941 filed September 15, 2005, which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

[0003] The present invention relates to methods of diagnosing ovarian cancer and kits therefor. More particularly, the present invention relates to the identification of markers associated with ovarian cancer and their use to detect ovarian cancer. The invention further relates to methods and reagents for the diagnosis of ovarian cancer.

BACKGROUND OF THE INVENTION

[0004] Epithelial ovarian carcinoma (EOC) is the most common malignant ovarian tumor, representing 80% of all ovarian malignancies (1). EOCs are thought to originate from either the normal ovarian surface epithelium (OSE) itself or from the crypts and inclusion cysts located in the stroma (1). EOCs are heterogeneous and are designated according to their histological subtype: serous, endometrioid, mucinous, clear cell, Brenner, undifferentiated or mixed (association of two or more sub-types) (2, 3). This cancer is often asymptomatic where over 70% of patients with ovarian cancer are diagnosed at an advanced stage of the disease. While up to 80% of the patients will initially respond to treatment, recurrence is generally observed within variable time intervals. Although 10-15% of the patients achieve and maintain a complete response to therapy, the remaining patients show persistent disease or eventually relapse thus requiring additional treatment. In contrast, borderline or low malignant potential (LMP) tumors, which represent 10-20% of all EOCs, have a more favorable prognosis compared to the invasive form of the disease, where the 5-year survival rate falls below 30% (1, 4).

[0005] Currently, there is no reliable method for screening early stage ovarian cancer. The clinically used CA125 serum marker (5) combined with trans-vaginal sonography, 3-dimensional ultrasound or power Doppler have yielded only minimal results (6). The reduced efficacy of CA125 for screening is largely related to its poor specificity. While elevated levels of CA125 are generally associated with the malignant disease, increased serum CA125 levels have also been observed with benign conditions (7), non-neoplastic conditions such as first trimester of pregnancy, menstruation, endometriosis, uterine fibrosis, acute salpingitis, hepatic diseases and inflammation of peritoneum, pericardium or pleura as well as with cancers of other sites. In addition, CA125 levels generally fail to rise in early stage disease, and lower levels are also associated with endometrioid and mucinous ovarian tumors (8). Thus, there is a need to develop reliable screening tools for EOC as these would be extremely valuable for improving cancer detection, clinical management and subsequently impact positively on survival.

[0006] Microarray technology is a powerful method for the analysis of cancer-specific gene expression by measuring tumor-specific expression of thousands of genes in hundreds of tumors (9), which can then be associated with specific clinical parameters. Candidate genes for diagnostic markers can further be characterized in combination with a large-scale quantitative polymerase chain reaction (Q-PCR) of RNA and immunohistochemical (IHC) analysis of protein expression using tissue arrays. However, such diagnostic techniques are difficult to implement since they require surgery to obtain the epithelial ovarian samples. Alternatively, if the differentially expressed gene encodes for a secreted protein circulating in peripheral blood, such a protein represents a potential serum based marker. The most common approach for testing such peripheral blood markers is through an enzyme-linked immunosorbent assay (ELISA). Although previous studies have investigated the potential of prostasin, osteopontin, mesothelin and HE4 (10-13) as diagnostic markers of EOC, no single marker has been shown to be sufficiently sensitive nor specific for proper diagnosis of ovarian cancer. Various combinations of different tumor markers have shown a higher specificity in differentiating benign from malignant disease (13, 14). However the efficacy and/or sensitivity of these markers were limited to advanced stage serous subtype tumors.

[0007] Therefore, ovarian cancer still remains a major source of morbidity and mortality and there is a clear need for the development of novel diagnosis method having the required sensitivity and specificity for early and reliable detection of ovarian cancer.

[0008] The present description refers to a number of documents, the content of which is herein incorporated by reference in their entirety.

[0009] Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of specific embodiments thereof, given by way of example only with reference to the accompanying drawings.

SUMMARY OF THE INVENTION

[0010] The invention relates to markers associated with ovarian cancer and corresponding methods, uses and products (e.g. probes, collections, kits, etc.) for the diagnosis of ovarian cancer.

[0011] Accordingly, in a first aspect, the invention provides a method comprising:

- (a) providing a biological sample from a subject (subject sample); and
- (b) detecting the expression level of each of the markers FGF-2 and CA125 in the subject sample.

[0012] In an embodiment, the above-mentioned subject is susceptible of having ovarian cancer.

[0013] In an embodiment, the above-mentioned subject is asymptomatic for ovarian cancer.

[0014] In an embodiment, the above-mentioned method further comprises:

- (c) comparing the expression level of each of the markers in the subject sample to corresponding pre-determined threshold expression levels for each of the markers, wherein an expression level of each of

4

the markers in the subject sample that is higher than the pre-determined threshold expression levels for each of the markers is an indication that the subject is affected by ovarian cancer.

[0015] In an embodiment, the above-mentioned method further comprises:

(c) comparing the expression level of each of the markers in the subject sample to the expression level of each of the markers in a control sample, wherein an expression level of each of the markers that is higher in the subject sample than in the control sample is an indication that the subject is affected by ovarian cancer

[0016] In an embodiment, the above-mentioned method further comprises:

(c) comparing the expression level of each of the markers in the subject sample to the expression level of each of the markers in a sample from the subject at an earlier time, wherein an expression level of each of the markers that is higher in the subject sample than in the sample from the subject at an earlier time is an indication that the subject is affected by ovarian cancer.

[0017] In an embodiment, the above-mentioned method further comprises:

(c) comparing the expression level of each of the markers in the subject sample to the expression level of each of the markers in a non-cancerous sample from the subject, wherein an expression level of each of the markers that is higher in the subject sample than in the non-cancerous sample from the subject is an indication that the subject is affected by ovarian cancer.

[0018] In an embodiment, the above-mentioned threshold expression level for each of the markers is determined by Receiver Operator Curves comparing the concentration of each of the markers in an ovarian cancer-free control population with that in a population affected by ovarian cancer.

[0019] In an embodiment, the above-mentioned expression is determined at the polypeptide level. In a further embodiment, the expression is determined using an immunoassay. In a further embodiment, the immunoassay is enzyme-linked immunosorbent assay (ELISA).

[0020] In an embodiment, the expression level of each of the above-mentioned markers is above the following pre-determined threshold expression levels: 50 U/ml for CA125 and 37 pg/ml for FGF-2.

[0021] In an embodiment, step (b) of the above-mentioned further comprises detecting the concentration of marker IL-18 in the sample.

[0022] In an embodiment, step (b) of the above-mentioned further comprises detecting the concentration of marker IL-18 in the sample, and the expression level of IL-18 in the sample is above the pre-determined threshold expression level of 215 pg/ml for this marker.

[0023] In an embodiment, the above-mentioned subject sample is a body fluid sample. In a further embodiment, the above-mentioned subject sample is selected from the group consisting of blood, plasma and serum. In a further embodiment, the above-mentioned subject sample is serum.

[0024] In an embodiment, the above-mentioned subject sample is primary culture cells derived from an ovarian tumor sample from the subject.

[0025] In an embodiment, the above-mentioned method is *in vitro*.

[0026] In an other aspect, the present invention provides a kit comprising means for detection of an expression level of each of markers CA125 and FGF-2 in a biological sample from a subject (subject sample), and instructions to use said markers in a method as recited above.

[0027] In an embodiment, the above-mentioned kit further comprises means for detection of an expression level of marker IL-18.

[0028] In an embodiment, the above-mentioned means for detection of expression level of each of the markers are antibodies.

[0029] In an other aspect, the invention provides a method of assessing the potential efficacy of a test compound for treating or inhibiting ovarian cancer in a subject, said method comprising determining the expression level of each of markers CA125 and FGF-2 in a biological sample from the subject (subject sample), before and after administration of the test compound to the subject, wherein a decrease in the expression level of the markers after administration of the test compound is indicative that said test compound is effective for treating or inhibiting ovarian cancer.

[0030] In an other aspect, the invention provides a method of assessing the potential efficacy of a therapy for treating or inhibiting ovarian cancer in a subject, said method comprising determining the expression level of each of markers CA125 and FGF-2 in a biological sample from the subject (subject sample), before and after administration of said therapy in said subject, wherein a decrease in the expression level of said markers after administration of said therapy is indicative that said therapy is effective for treating or inhibiting ovarian cancer.

[0031] In an embodiment of the above-mentioned methods, the methods further comprise detecting the concentration of marker IL-18 in the sample. In an embodiment, the above-mentioned expression is determined at the polypeptide level. In a further embodiment, the expression is determined using an immunoassay. In a further embodiment, the immunoassay is enzyme-linked immunosorbent assay (ELISA).

[0032] In an embodiment, the above-mentioned sample is a body fluid sample. In a further embodiment, the above-mentioned sample is selected from the group consisting of blood, plasma and serum. In a further embodiment, the above-mentioned sample is serum.

[0033] In an embodiment, the above-mentioned subject is a human.

[0034] In an embodiment, the above-mentioned ovarian cancer is epithelial ovarian carcinoma (EOC).

BRIEF DESCRIPTION OF THE DRAWINGS

[0035] In the appended drawings:

[0036] Figure 1 shows validation of gene expression profiles by Q-PCR on primary culture samples. (A) Two micrograms of RNA extracted from 9 NOSE and 8 primary culture cells of EOC were reverse-transcribed and the levels of IL-18 and FGF-2 quantified using specific primers (left hand panels). Each expression level was normalized to that of the control RNA. Relative fold change expression is the ratio of the NOSE 18 gene expression to that of other samples. Three micrograms of RNA extracted from 12 BOT (benign ovarian tumor) tissues and 22 EOC tissues were reverse-transcribed and the levels of IL-18 and FGF-2 quantified as in the left hand panels (right hand panels). Each expression level was normalized to that of the control RNA. Relative fold change expression is the ratio of the BOT142 gene expression to that of other samples; (B) shows the expression of IL-18 and FGF-2 in normal ovarian surface epithelial (NOSE) tissues and four histopathologies of EOC tissues. IHC was performed using antibodies against indicated proteins (left). Nuclei are counterstained with hematoxylin (blue). Brown color demonstrates specific peroxidase staining;

[0037] Figure 2 shows serum measurement of CA125 (A) IL-18 (B) and FGF-2 (C) by ELISA. Patients sera was tested for all CA125, IL-18 and FGF-2 and threshold levels (dashed lines) were determined for each serum marker. Solid lines show the median level of the serum marker for each group of patients. LMP: low malignant potential tumor patients (n=5). NOSE: normal ovarian surface epithelia patients (n=11). BOT: benign ovarian tumor patients (n=23). TOV: ovarian tumor patients (equivalent to EOC: invasive epithelial ovarian cancer patients) (n=42);

[0038] Figure 3 presents nucleic acid (SEQ ID NO: 1) and polypeptide (SEQ ID NO: 2) sequences for CA125;

[0039] Figure 4 presents nucleic acid (SEQ ID NO: 3) and polypeptide (SEQ ID NO: 4) sequences for IL-18; and

[0040] Figure 5 presents nucleic acid (SEQ ID NO: 5) and polypeptide (SEQ ID NO: 6) sequences for FGF-2.

DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0041] The present invention concerns markers which can be used to diagnose ovarian cancer in subjects.

[0042] "Selectivity" in the context of the present invention refers to the ability of a marker of the present invention to discriminate between a sample affected by ovarian cancer and one that is not i.e. a marker with high selectivity produces few false positives.

[0043] "Sensitivity" in the context of the present invention refers to the ability of a marker of the present invention to correctly identify a sample affected by ovarian cancer as such i.e. a marker with high sensitivity produces few false negatives.

[0044] "Marker" in the context of the present invention refers to, without being so limited, a nucleic acid or a polypeptide (or fragment thereof), which is differentially present in a sample taken from a subject having ovarian cancer as compared to a comparable sample taken from a control subject (e.g., a person with a negative diagnosis or undetectable cancer, normal or healthy subject).

[0045] "Subject" in the context of the present invention relates to any mammal including a mouse, rat, pig, monkey, horse. In a specific embodiment, it refers to a human.

[0046] As used herein the terms "sample from the subject at an earlier time" is meant to refer to a sample from a subject at a time where it was known that the subject was not affected by ovarian cancer.

[0047] The articles "a," "an" and "the" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article.

[0048] The term "including" and "comprising" are used herein to mean, and are used interchangeably with, the phrases "including but not limited to" and "comprising but not limited to".

[0049] The term "such as" is used herein to mean, and is used interchangeably with, the phrase "such as but not limited to".

[0050] Optionally, a marker can be modified before analysis to improve its resolution or to determine its identity. For example, the markers may be subject to proteolytic digestion before analysis. Any protease can be used. Proteases, such as trypsin, that are likely to cleave the markers into a discrete number of fragments are particularly useful. The fragments that result from digestion function as a fingerprint for the markers, thereby enabling their detection indirectly. This is particularly useful where there are markers with similar molecular masses that might be confused for the marker in question. Also, proteolytic fragmentation is useful for high molecular weight markers because smaller markers are more easily resolved by mass spectrometry. The markers can also be modified by the attachment of a tag of particular molecular weight that specifically bind to molecular markers, further distinguishing them. Optionally, after detecting such modified markers, the identity of the markers can be further determined by matching the physical and chemical characteristics of the modified markers in a protein database (e.g., SwissProt™).

[0051] Expression levels may in general be detected by either detecting mRNA from the cells and/or detecting expression products, such as polypeptides and proteins. Expression of the transcripts and/or polypeptides encoded by the nucleic acids described herein may be measured by any of a variety of known methods in the art. In general, the nucleic acid sequence of a nucleic acid molecule (e.g., DNA or RNA) in a subject sample can be detected by any suitable method or technique of measuring or detecting gene sequence or expression. Such methods include, but are not limited to, polymerase chain reaction (PCR), reverse transcriptase-PCR (RT-PCR), in situ PCR, quantitative PCR (q-PCR), in situ hybridization, Southern blot, Northern blot, sequence analysis, microarray analysis, detection of a reporter gene, or other DNA/RNA hybridization platforms. For RNA expression, preferred methods include, but are not limited to: extraction of cellular mRNA and Northern blotting using

labeled probes that hybridize to transcripts encoding all or part of one or more of the genes of this invention; amplification of mRNA expressed from one or more of the genes of this invention using gene-specific primers, polymerase chain reaction (PCR), quantitative PCR (q-PCR), and reverse transcriptase-polymerase chain reaction (RT-PCR), followed by quantitative detection of the product by any of a variety of means; extraction of total RNA from the cells, which is then labeled and used to probe cDNAs or oligonucleotides encoding all or part of the genes of this invention, arrayed on any of a variety of surfaces; in situ hybridization; and detection of a reporter gene. The term "quantifying" or "quantitating" when used in the context of quantifying transcription levels of a gene can refer to absolute or to relative quantification. Absolute quantification may be accomplished by inclusion of known concentration(s) of one or more target nucleic acids and referencing the hybridization intensity of unknowns with the known target nucleic acids (e.g., through generation of a standard curve). Alternatively, relative quantification can be accomplished by comparison of hybridization signals between two or more genes, or between two or more treatments to quantify the changes in hybridization intensity and, by implication, transcription level.

[0052] As used herein, "control sample" refers to a sample of the same type, that is, obtained from the same biological source (e.g. body fluid, tissue, etc.) as the tested sample but from a healthy subject, (i.e. who is not afflicted by ovarian cancer, and preferably who is not afflicted by any cancer). The control sample can also be a standard sample that contains the same concentration of the above-mentioned markers that are normally found in a corresponding biological sample obtained from a healthy subject. For example, there can be a standard control sample for the amounts of CA125, IL-18 and FGF-2 normally found in biological samples such as tissue, blood, plasma and serum.

[0053] The methods of the invention can also be practiced, for example, by selecting a combination of the above-mentioned markers and one or more additional markers for which increased or decreased expression correlates with ovarian cancer, such as CA72-4, hK6, hK10, HSCCE, kallikrein 4, kallikrein 5, kallikrein 6, kallikrein 8, kallikrein 9, kallikrein 11, CA15-3, CA19-9, OVX1, lysophosphatidic acid (LPA) or carcinoembryonic antigen (CEA), as well as other markers specific for other types of

cancer. Those skilled in the art will be able to select useful diagnostic markers for detection in combination with CA125, IL-18 and FGF-2. Similarly, four or more or five or more or a multitude of markers can be used together for determining a diagnosis of a patient.

[0054] In an embodiment, the expression level of the above-mentioned markers is determined at the polypeptide level.

[0055] Methods to measure polypeptide expression levels of the markers of this invention, include, but are not limited to: Western blot, immunoblot, enzyme-linked immunosorbant assay (ELISA), radioimmunoassay (RIA), immunoprecipitation, surface plasmon resonance, chemiluminescence, fluorescent polarization, phosphorescence, immunohistochemical analysis, matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry, microcytometry, microarray, microscopy, fluorescence activated cell sorting (FACS), flow cytometry, and assays based on a property of the protein including but not limited to DNA binding, ligand binding, or interaction with other protein partners.

[0056] The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymers.

[0057] In an embodiment, the expression level of the above-mentioned markers is determined using an immunoassay.

[0058] An immunoassay is an assay that uses an antibody to specifically bind an antigen (e.g., a marker). The immunoassay is characterized by the use of specific binding properties of a particular antibody to isolate, target, and/or quantify the antigen. The phrase "specifically (or selectively) binds" to an antibody or "specifically (or selectively) immunoreactive with," when referring to a protein or peptide, refers to

a binding reaction that is determinative of the presence of the protein in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and do not substantially bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to a marker from specific species such as rat, mouse, or human can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with that marker and not with other proteins, except for polymorphic variants and alleles of the marker. This selection may be achieved by subtracting out antibodies that cross-react with the marker molecules from other species.

[0059] Using the purified markers or their nucleic acid sequences, antibodies that specifically bind to a marker can be prepared using any suitable methods known in the art. See, e.g., Harlow & Lane, *Antibodies: A Laboratory Manual* (1988) and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986). Such techniques include, but are not limited to, antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice.

[0060] Generally, a sample obtained from a subject can be contacted with the antibody that specifically binds the marker. Optionally, the antibody can be fixed to a solid support to facilitate washing and subsequent isolation of the complex, prior to contacting the antibody with a sample. Examples of solid supports include glass or plastic in the form of, e.g., a microtiter plate, a stick, a bead, or a microbead. The sample is preferably a biological fluid sample taken from a subject. The sample can be diluted with a suitable eluant before contacting the sample to the antibody.

[0061] After incubating the sample with antibodies, the mixture is washed and the antibody-marker complex formed can be detected. This can be accomplished by incubating the washed mixture with a detection reagent. This detection reagent may be, e.g., a second antibody which is labeled with a detectable label. Exemplary

detectable labels include magnetic beads (e.g., DYNABEADS™), fluorescent dyes, radiolabels, enzymes (e.g., horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic beads. Alternatively, the marker in the sample can be detected using an indirect assay, wherein, for example, a second labeled antibody is used to detect bound marker-specific antibody, and/or in a competition or inhibition assay wherein, for example, a monoclonal antibody which binds to a distinct epitope of the marker is incubated simultaneously with the mixture.

[0062] Methods for measuring the amount of, or presence of, antibody-marker complex include, for example, detection of fluorescence, luminescence, chemiluminescence, absorbance, reflectance, transmittance, birefringence or refractive index (e.g., surface plasmon resonance, ellipsometry, a resonant mirror method, a grating coupler waveguide method or interferometry). Optical methods include microscopy (both confocal and non-confocal), imaging methods and non-imaging methods. Electrochemical methods include voltametry and amperometry methods. Radio frequency methods include multipolar resonance spectroscopy. Methods for performing these assays are readily known in the art. Useful assays include, for example, an enzyme immune assay (EIA) such as enzyme-linked immunosorbent assay (ELISA), a radioimmune assay (RIA), a Western blot assay, or a slot blot assay. These methods are also described in, e.g., *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993); *Basic and Clinical Immunology* (Stites & Terr, eds., 7th ed. 1991); and Harlow & Lane, *supra*.

[0063] In a further embodiment, the above-mentioned immunoassay is an enzyme-linked immunosorbant assay (ELISA).

[0064] The markers can be measured in different types of biological samples. The sample is preferably a biological fluid sample such as blood, plasma and serum. Other typical biological samples include, but are not limited to, tissue biopsy from ovarian tumor, sputum, lymphatic fluid, blood cells (e.g., peripheral blood mononuclear cells), tissue or fine needle biopsy samples, urine, peritoneal fluid, colostrums, breast milk, fetal fluid, tears, pleural fluid, or cells therefrom. Because all of the markers are found in blood serum, blood serum is a preferred sample source

for embodiments of the invention.

[0065] If desired, the sample can be prepared to enhance detectability of the markers. For example, to increase the detectability of markers, a blood serum sample from the subject can be preferably fractionated by, e.g., Cibacron™ blue agarose chromatography and single stranded DNA affinity chromatography, anion exchange chromatography, affinity chromatography (e.g., with antibodies) and the like. The method of fractionation depends on the type of detection method used. Any method that enriches for the protein of interest can be used. Sample preparations, such as pre-fractionation protocols, are optional and may not be necessary to enhance detectability of markers depending on the methods of detection used. For example, sample preparation may be unnecessary if antibodies that specifically bind markers are used to detect the presence of markers in a sample.

[0066] Typically, sample preparation involves fractionation of the sample and collection of fractions determined to contain the markers. Methods of pre-fractionation include, for example, size exclusion chromatography, ion exchange chromatography, heparin chromatography, affinity chromatography, sequential extraction, gel electrophoresis and liquid chromatography. The analytes also may be modified prior to detection. These methods are useful to simplify the sample for further analysis. For example, it can be useful to remove high abundance proteins, such as albumin, from blood before analysis. Examples of methods of fractionation are described in WO/2003/057014.

[0067] The methods for detecting these markers in a sample have many applications. For example, one or more markers can be measured to aid human cancer diagnosis. In another example, the methods for detection of the markers can be used to monitor responses in a subject to cancer treatment. In another example, the methods for detecting markers can be used to assay for and to identify compounds that modulate expression of these markers *in vivo* or *in vitro*.

[0068] In an embodiment, the subject is a human.

[0069] The present invention also relates to a kit for determining the likelihood of ovarian cancer in a subject, said kit comprising means for detection of expression of the markers CA125 and FGF-2 and, in more specific embodiments, the marker IL-18, in a biological sample from said subject together with instructions setting forth the above-mentioned method. Means for detection include probe, primer (or primer pair), or immunological reagent (e.g. antibody) in accordance with the present invention. For example, a compartmentalized kit in accordance with the present invention includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers may for example include a container which will accept the test sample (DNA, protein or cells), a container which contains the primers used in the assay, containers which contain enzymes, containers which contain wash reagents, and containers which contain the reagents used to detect the indicator products.

[0070] Kits for evaluating expression of nucleic acids can include, for example, probes or primers that specifically bind a nucleic acid of interest (e.g., a nucleic acid, the expression of which correlates with increased likelihood of ovarian cancer). The kits for evaluating nucleic acid or polypeptide expression can provide substances useful as standard (e.g., a sample containing a known quantity of a nucleic acid or polypeptide to which test results can be compared, with which one can assess factors that may alter the readout of a diagnostic test, such as variations in an enzyme activity or binding conditions). Kits for assessing nucleic acid or polypeptide expression can further include other reagents useful in assessing levels of expression (e. g. buffers and other reagents for performing amplification reactions, or for detecting binding of a probe to a nucleic acid or binding of an antibody to a polypeptide). The kits can provide instructions for performing the assay used to evaluate gene/polypeptide expression for determining likelihood of ovarian cancer based on the results of the assay. For example, the instructions can indicate that levels of expression of a gene of interest (e.g., relative to a standard or a control), correlate with increased likelihood for ovarian cancer.

[0071] The invention further provides a method of assessing the potential efficacy of a test compound for treating or inhibiting ovarian cancer in a subject, said method comprising determining, in a biological sample from said subject, the expression of the markers CA125 and FGF-2 and, in more specific embodiments, the marker IL-18, before and after administration of said test compound in said subject, wherein a decrease in the expression of said markers after administration of said test compound is indicative that said test compound is effective for treating or inhibiting ovarian cancer.

[0072] In an other aspect, the invention provides a method of assessing the efficacy of a therapy for treating or inhibiting ovarian cancer in a subject, said method comprising determining, in a biological sample from said subject, the expression of the markers CA125 and FGF-2 and, in more specific embodiments, the marker IL-18, before and after administration of said therapy in said subject, wherein a decrease in the expression of said markers after administration of said therapy is indicative that said therapy is effective for treating or inhibiting ovarian cancer.

[0073] In an embodiment, the above-mentioned ovarian cancer is epithelial ovarian carcinoma.

[0074] The present invention is illustrated in further details by the following non-limiting examples.

EXAMPLE 1

Clinical samples

[0075] Tissue samples and sera were obtained with informed consent from participants. Tumor samples were collected from surgeries performed at the Centre Hospitalier de l'Université de Montréal (CHUM). Histopathology, grade and stage of tumors were assigned according to the International Federation of Gynecology and Obstetrics (FIGO) criteria. Normal controls were defined as tumor-free patients. Primary cell cultures from normal ovarian surface epithelia (NOSE) and EOC samples were established as described (15, 16) and used for microarray analysis. Cells in primary culture were maintained in OSE media consisting of 50:50 medium 199:105

(Sigma) supplemented with 10% fetal bovine serum (FBS), 2.5µg/mL amphotericin B and 50 µg/mL gentamicin (15). Independent cohorts for microarray, ELISA and tissue array IHC studies were used and are presented in Table 1 below.

Quantitative PCR

[0076] Linear amplification of RNA from primary culture cells was performed as described previously (17). The cDNA synthesis was done according to the protocol of the SuperScript™ First-Strand Synthesis System for Q-PCR (Invitrogen Life Technologies) with a starting amount of 2 mg RNA and reverse transcription performed with random hexamers. The PCR reaction (temperature, specificity) was performed using conventional PCR conditions with a Rotor-gene™ 3000 Real-Time Centrifugal DNA Amplification System (Corbett tumor tissues Research, NSW, Australia). The Quantitect™ SYBR Green PCR (Qiagen Inc., ON, Canada) reaction mixture was used according to the manufacturer instruction. Serial dilutions were performed to generate a standard curve for each gene tested in order to define the efficiency of the Q-PCR reaction and a melt curve was done to confirm the specificity of the reaction. Based on the stability of its expression in microarray experiments, primers for the ERK1 gene were used as an internal control. Experiments were done in duplicate. Positive and negative controls were introduced in all experiments. The sequences for IL-18 primers are: Fwd 5'-CGCTTCCTCTCGCAACAACTAT-3' (SEQ ID NO: 7) and Rev 5'-CCGGGGTGCATTATCTCTACAGT-3' (SEQ ID NO: 8); FGF-2: Fwd 5'-CGCGCAGGAGGGAGGAGA-3' (SEQ ID NO: 9) and Rev 5'-ACGCCGCCTGGGGAGAG-3' (SEQ ID NO: 10) and finally ERK1: Fwd 5'-GCGCTGGCTCACCCCTACCT-3' (SEQ ID NO: 11) and Rev 5'-GCCCCAGGGTGCAGAGATGTC-3' (SEQ ID NO: 12). The Pfaffl analysis was used method to measure the relative quantity of gene expression (18).

RNA preparation and microarray

[0077] Total RNA was extracted with TRIzol™ reagent (Gibco / BRL, Life Technologies Inc., Grand Island, NY, USA). RNA was extracted directly from cells grown to 80% confluency. The quality of the RNA was monitored by gel electrophoresis and a 2100 Bioanalyzer using the RNA 6000 Nano LabChip™ kit (Agilent Technologies, Germany). Biotinylated hybridization target was prepared from

total RNA as described (19). HuGeneFI™ 6800 GeneChip™ microarray experiments were performed at the McGill University and Genome Québec Innovation Centre and raw data was processed using the Affymetrix™ MAS4 software. Detailed protocols are known in the art and are available at www.genomequebec.mcgill.ca/center.php. The raw data of each experiment was normalized according to the mean of the global intensity adjusted to 100 units. Arrays with global intensity below 100 were eliminated. After normalization, all values below 20 were considered as technical noise and expression values below this threshold were transformed to this value. All the EST's were next filtered, which had "A" call (ambiguous signal) across all samples. To detect differentially expressed genes in ovarian tumor samples versus normal ovarian cells, two statistical tests were used to identify classifiers. A parametric and a non-parametric (Mann-Whitney (U)) test were performed using GeneSpring™ software (Silicon Genetics). Candidate genes identified in common in the two analyses were selected for further analysis.

Tissue array and IHC

[0078] The following monoclonal antibodies were used in immunohistochemistry (IHC): anti-IL-18 (R&D system), anti-FGF-2 (Santa Cruz Biotechnology). A tissue array containing 94 cores of ovarian epithelial tissues (see Table 1 below) was built and used for IHC studies. Briefly, the tissue array was heated at 60°C for 30 min, deparaffinized in toluene and rehydrated in a gradient of ethanol. To unmask antigen the slides were submerged in 90°C citrate buffer (0.01M citric acid + 500ul tween-20/L adjusted to pH 6.0) (J.T. Baker Philipsburg, NJ) for 15 min. The tissue was blocked with a protein-blocking serum-free reagent (DakoCytomation Inc., Mississauga, ON) and incubated with the different antibodies overnight at 4°C in a humid chamber. The optimal concentration for each primary antibody was determined by serial dilutions. Subsequently, endogenous peroxidase activity was quenched by treatment with 3% H₂O₂. The array was then incubated with a secondary biotinylated antibody (DakoCytomation Inc., Mississauga, ON) for 10 min followed by incubation with a streptavidin-peroxidase complex (Dako Diagnostics Canada Inc.) for 10 min at room temperature. Reaction products were developed using diaminobenzidine (brown stain) containing 0.3% H₂O₂ as a substrate for peroxidase and nuclei were counterstained with diluted hematoxylin (blue stain). Epithelial zones were scored according to the intensity of staining (value of 0 for

absence, 1 for weak, 2 for moderate, 3 for high intensity). Each array was independently analyzed in a blind study by two independent observers. Statistical analyses were performed using the T-test.

ELISA

[0079] Patient's blood was centrifuged for 30 min at 2500 rpm and the separated serum was immediately frozen at -20°C until further use. Before measurement, all sera were re-centrifuged for 10 min at 8000 rpm. The sera were further tested by ELISA for CA125 (Panomics BC1013), FGF-2 (R&D System, item DFB50) and IL-18 (R&D System, item 7620) concentration according to the manufacturer's instructions. The limit of detection for IL-18 was 20 pg/ml, 10 U/ml for CA125 and 20 pg/ml for FGF-2. Independent experiments were calibrated with at least two samples. Statistical analyses were performed using SPSS software. For small sample set sizes (<10) the Mann-Whitney U test was applied, otherwise statistical analysis relied on the T-test.

EXAMPLE 2

Identification of two genes up-regulated in ovarian cancer and encoding for cytokines

[0080] Comparative analysis of gene expression profiles of ovarian epithelial cells was performed using 11 primary cultures of normal ovarian epithelial surface (NOSE) samples and 39 primary cultures of EOC samples. The 39 EOC represented different grades, stages and pathologies of ovarian cancer (see Table 1 below). To gain insight into genes exhibiting dominant expression levels in ovarian tumors, the expression profiles were analyzed using two different supervised classification algorithms. Among a total of 177 candidate genes that were common to both supervised analyses, several genes encoding for secreted proteins were identified but only two genes encoding for cytokines, IL-18 and FGF-2, were present. In order to maximize the chance of sampling differential gene expression in serum, these latter two genes were selected for further study.

[0081] **TABLE 1: SAMPLE SETS USED IN EACH EXPERIMENT**

Histopathology	Sample size	Tumor grade					Tumor stage	
		B	I	2	3	Mixed	Low	high
Microarray set (n=50)								
Normal	11							
Serous	29	6	1	7	15		4	25
Endometrioid	7			3	4			7
Mixed	1				1			1
Clear cell	2				2			2
Total tumors	39	6	1	10	22		4	35
Tissue array (n=114)								
Normal	20						NA	NA
Serous	21	4	5	5	7		NA	NA
Endometrioid	27		13	7	5	2	NA	NA
Clear cell	17			5	9	3	NA	NA
Mixed	5				3	2	NA	NA
Mucinous	24	21	3				NA	NA
Total tumors	94	25	18	17	24	6	NA	NA
ELISA set (n=70)								
Normal and benign	25							
Serous	29	3	2	3	20	1	3	26
Endometrioid	3		3				2	1
Clear cell	5				4	1	3	2
Mixed	3		1	1	1		0	3
Brenner	2					2	1	1
Mucinous	3	2	1				2	1
Total tumors	45	5	7	4	25	4	11	34
PCR tissues (n=34)								
Normal and benign	12							
Serous	6	1		1	2		2	4
Endometrioid	5		1		4			5
Clear cell	7				5	2	1	6
Mucinous	4	2	1		1		1	2
Total tumors	22	3	2	3	12	2	4	18

Grade B are low malignant potential tumors. Low stage: stage I and II tumors; high stage: stage III and IV tumors.

EXAMPLE 3

Validation of the differential gene expression of IL-18 and FGF-2

[0082] Q-PCR was used to validate the differential expression of the IL-18 and FGF-2 RNA as observed in the microarray analysis. For this purpose, 9 NOSEs and 8 EOCs randomly chosen from the previous set of primary cultures, as well as 12 benign tumors (BOT) and 22 EOCs from fresh tissues, were compared and their expression levels correlated with the results obtained by microarray analysis (Figure 1A, left hand panels). IL-18 and FGF-2 RNA were weakly detectable in NOSE samples while they were readily detectable in the majority of malignant samples serving as an independent confirmation of their differential expression in EOC. To determine IL-18 and FGF-2 expression in tissues, RNAs isolated from 12 benign and 22 malignant ovarian tumor tissues were also tested (Table I). Most malignant tissues, with the exception of two mucinous and one serous tumor, showed an overexpression of IL-18 (Figure 1A, right hand panels). Highest FGF-2 RNA

expression was seen in endometrioid tissues, although the difference between benign and malignant tissues was less striking (Figure 1A, right hand panels).

EXAMPLE 4

Protein expression of IL-18 and FGF-2 in ovarian tissue specimens

[0083] To address the expression of FGF-2 and IL-18 in EOC, IHC was performed with IL-18 and FGF-2 specific antibodies on ovarian tissues using a tissue array containing 20 NOSE and 94 EOC tissue cores from 114 independent patients. The 94 EOC cores represented the different grades and pathologies of ovarian cancer with the exception of Brenner tumors (see Table 1 above). Scoring results from the IHC analyses are summarized in Table 2 below. IL-18 and FGF-2 were expressed in NOSE as well as EOC tissues. In NOSE tissues, heterogeneity of staining intensity was observed among the different cores (see Table 2 below). In addition, IL-18 and FGF-2 staining was also present in the stroma of NOSE tissues, which may be due to their direct expression by stromal cells or to the secretion of these cytokines by adjacent epithelial cells. EOC tissues showed a slightly more marked staining of IL-18 and FGF-2. The staining was a significantly stronger for IL-18 in serous, endometrioid and clear cells tumors ($p < 0.05$) and for endometrioid and clear cell tumors with FGF-2 (Figure 1B and Table 2).

[0084] TABLE 2: INTENSITY OF IMMUNOSTAINING OF TISSUE ARRAY WITH ANTI-IL-18 AND ANTI-FGF-2 ANTIBODIES

Histopathology	<i>p</i>	Staining intensity			
		0	1+	2+	3+
Antibody anti-IL-18					
Normal		3	13	4	0
Clear cells	<0.001	0	1	14	2
Endometrioid	0.001	1	9	15	2
Serous	0.03	0	13	8	0
Mixed	0.05	0	2	3	0
Mucinous	0.20	8	11	5	0
Total tumors	0.005	12	36	45	4
Antibody anti-FGF-2					
Normal		5	7	5	3
Clear cells	<0.001	0	2	12	3
Endometrioid	0.01	1	6	15	5
Serous	0.45	7	1	13	0
Mixed	0.33	2	0	3	0
Mucinous	0.08	11	8	6	0
Total tumors	0.14	21	17	49	8

*0, absence; 1, weak; 2, moderate; 3, for high intensity.

EXAMPLE 5**Serum IL-18 and FGF-2 proteins as markers of EOC**

[0085] IL-18 and FGF-2 were studied as individual markers in comparison to CA125. For this purpose a total of 72 patients was selected: 25 patients were free of cancer and 47 patients had ovarian cancer (see Table 1 above). Among the cancer-free patients, six presented with benign ovarian (BOV) or (benign) tumors. Among the 47 ovarian cancers, five were low malignant potential (LMP) tumors, eight grade 1, four grade 2, and 26 grade 3 tumors (see Table 1 above). Six different pathologies were represented in the set of selected patients with EOC (serous, endometrioid, clear cells, Brenner, mucinous and mixed).

[0086] CA125 was significantly elevated in patients with EOC ($p < 0.001$) (see Table 3 below). No significant difference was observed in patients with or without benign tumors ($p = 0.31$). Patients with LMP tumors showed a lower level of CA125 (median 75 U/ml) than malignant EOC (median level 350 U/ml) (see Figure 2 and Table 3 below). This observation was consistent with the increased levels of CA125 which correlated with increased tumor grade ($r = 0.33$, $p = 0.004$) and stage in independent studies (8). The increased level of CA125 also correlated with the histopathology (Spearman's Rho test $p = 0.002$) where CA125 was more elevated in serous tumors than in endometrioid and clear cell tumors (see Table 3 below).

[0087] While IL-18 was also significantly more elevated in EOC patients (median level pg/ml $p = 0.003$) it was correlated with tumor grade (Spearman's Rho test, $p = 0.172$). Serous tumors showed the highest level of IL-18 expression (median level 305 pg/ml) but no significant correlation was observed between IL-18 and pathology disease (Spearman's Rho test, $p = 0.173$). As observed with CA125, there was no significant difference between patients with or without benign tumors ($p = 0.99$) (see Table 3 below).

[0088] FGF-2 levels were higher in EOC patients compared to cancer free patients although with a weaker significance compared to CA125 or IL-18 ($p = 0.04$). In accordance with the results obtained in tissue arrays, serum FGF-2 levels were

highest in association with clear cell tumors. A correlation between increased FGF-2 serum levels and tumor grade was also detected (Spearman's Rho test $p=0.02$) (Table 3).

[0089] TABLE 3: EXPRESSION LEVEL OF MARKERS CA125, IL-18 AND FGF-2 IN SERUM

	CA125 (U/ml) [Median/average (p^*)]	IL-18 (pg/ml) [Median/average (p^*)]	FGF-2 (pg/ml) [Median/average (p^*)]
NOSE + benign	37/92	204/215	29/35
All EOC	306/474 (<0.001)	264/315 (0.001)	39/50 (0.037)
Normal	44/114	203/212	34/43
Benign	32/63 (0.31)	207/219 (0.99)	27/25 (0.31)
LMP	75/100 (0.30)	236/257 (0.31)	68/21 (0.175)
Invasive EOC	350/545 (<0.001)	258/327 (0.003)	49/56 (0.006)
Grade 1	339/336 (0.03)	282/267 (0.04)	31/31 (0.70)
Grade 2	260/285 (0.04)	251/283 (0.008)	43/36 (0.011)
Grade 3	484/683 (<0.001)	307/370 (0.003)	66/68 (0.002)
Low stage	75/419 (0.18)	233/239 (0.62)	39/47 (0.65)
High stage	350/450 (<0.001)	281/282 (<0.001)	44/42 (0.016)
Serous	419/544 (<0.001)	305/358 (<0.001)	44/54 (0.11)
Endometrioid	339/380 (0.16)	281/252 (0.37)	23/23 (0.63)
Mucinous	38/46 (0.79)	263/247 (0.29)	21/32 (0.68)
Clear cells	34/492 (0.55)	242/330 (0.30)	69/49 (0.10)

NOSE, normal ovarian surface epithelia; EOC, epithelial ovarian cancer; LMP, low malignant potential tumor; low stage, stage I and II; high stage, stage III and IV. p^* Mann-Whitney test.

EXAMPLE 6

Diagnostic potential of serum CA125, IL-18 and FGF-2 as markers

[0090] Receiver Operator Curves were used to determine threshold values for the three serum markers to compare the diagnostic potential of the individual cytokine markers with CA125. The greatest accuracy in differential diagnosis of malignant tumors was achieved with a threshold of 50 U/ml for CA125, 215 pg/ml for IL-18 and 37 pg/ml for FGF-2. Sensitivity, namely the fraction of patients correctly diagnosed with ovarian cancer, was more accurate when considering CA125 or IL-18, as individual markers. Sensitivity as determined by CA125 and IL-18 was 82% and 78% respectively, compared to 58% with FGF-2 (Table 4). To ensure that there was no difference in sensitivity between CA125 and IL-18 the number of samples was increased to 97 (data not shown). In this larger set, CA125 and IL-18 sensitivity levels remained similar (75% and 74%, respectively).

[0091] Specificity was defined as the fraction of samples correctly diagnosed

as non-malignant, including serum from patients with normal ovaries or benign disease. Individual analysis of patients with either normal ovaries or benign disease gave similar results (data not shown). Specificity was best provided by FGF-2 (72%). CA125 and IL-18 showed relative low similar specificities of 60% and 64% respectively (Table 4). In the larger set, CA125 and IL-18 specificity levels remained similar (61% and 64% respectively, data not shown).

[0092] TABLE 4: SPECIFICITY AND SENSITIVITY OF CA125, IL-18 AND FGF-2 IN UNIVARIATE OR MULTIVARIATE ANALYSIS

Patient type	CA125 (U/ml)		IL-18 (pg/ml)		FGF-2 (pg/ml)		CA125+IL-18 + FGF2	
	n >50 U/ml	%+	n >215 pg/ml	%+	n >37 pg/ml	%+	n	%+
Specificity								
NOSE + benign	10/25	60	9/25	64	7/25	72	5/25	80
Sensitivity								
All EOC	37/45	82	35/45	78	26/45	58	35/45	78
LMP	3/5	60	3/5	60	1/5	20	3/5	60
Invasive EOC	34/42	81	34/42	81	25/42	60	32/42	76
Low stage	6/11	55	6/11	55	7/11	64	7/11	64
High stage	31/34	91	29/34	85	20/34	59	28/34	73
Serous	28/29	97	26/29	90	19/29	66	27/29	93
Endometroide	2/3	67	2/3	67	1/3	33	1/3	33
Clear cell	1/5	20	3/5	60	4/5	80	3/5	60
Mucinous	1/3	33	2/3	67	1/3	33	1/3	33
Brenner	2/2	100	0/2	0	0/2	0	1/2	50
Mixed	3/3	100	2/3	67	1/3	33	2/3	67

EOC, Epithelial ovarian cancer; LMP, low malignant potential tumor; low stage, stage I-II; high stage, stage III-IV; % +, corresponds to percentage of NOSE + benign which do not score above the threshold (specificity).

EXAMPLE 7

Diagnostic potential of serum IL-18 and FGF-2 as combined markers with CA125

[0093] The estimated correlation among the three serum markers (IL-18, FGF-2 and CA125) were low suggesting that they were complementary to each other and that a multivariate approach might outperform the CA125 assay alone. To validate this hypothesis a multivariate analysis was performed using a logistic binary regression algorithm. As shown in Table 5, FGF-2, but not IL-18, increased the diagnosis potential of CA125 (Odd Ratio from 5.24 to 6). However addition of both FGF-2 and IL-18 achieved a superior diagnostic potential (Odd Ratio= 6.94, 0.95 (1.99-24,39), p=0.002) suggesting that the combination of both IL-18 and FGF-2 with CA125 allows a better sensitivity and specificity.

[0094] Scoring samples as malignant was also tested based on whether ELISA values were above the threshold for at least two of the three markers. In this analysis (Table 4), a sensitivity of 78% was achieved which was similar to that obtained with CA125 or IL-18 alone, but the specificity of diagnosis was dramatically increased from CA125 (60%), IL-18 (64%) or FGF-2 (72%) alone to 80% the combination of these serum markers (Table 4). Similar result was obtained in a larger set of samples (77%, data not shown).

[0095] TABLE 5: LOGISTIC BINARY REGRESSION (LBR) ANALYSIS OF MULTIVARIATE ANALYSIS OF CA125, IL-18 AND FGF-2

		p	OR	CI
LBR	CA125	<0.001	5.24	2.07-13.33
	CA125+IL-18	0.002	4.78	1.81-12.66
	CA125+FGF-2	0.002	6	1.93-18.61
	IL-18+FGF-2	0.014	2.25	0.726-6.96
	CA125+IL-18+FGF	0.002	6.94	1.99-24.39

OR+: odd ratio. CI: confidence interval 95%

[0096] Although the present invention has been described hereinabove by way of specific embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention as defined in the appended claims. Throughout this application, various references are referred to describe more fully the state of the art to which this invention pertains. The disclosures of these references are hereby incorporated by reference into the present disclosure.

REFERENCES

1. Auersperg, N., Wong, A. S., Choi, K. C., Kang, S. K., and Leung, P. C. Ovarian surface epithelium: biology, endocrinology, and pathology. *Endocr Rev*, 22: 255-288., 2001. Agarwal, P., Bagga, R., Jain, V., Kalra, J. & Gopalan, S. Familial recurrent molar pregnancy: a case report. *Acta Obstet Gynecol Scand* 83, 213-4 (2004).
2. Serov, S. F., Scully, R., and Sobin, L. H. *Histological typing of ovarian tumours.*, Vol. 9. Geneva: World Health Organization, 1973.
3. Chuaqui, R. F., Cole, K. A., Emmert-Buck, M. R., and Merino, M. J. Histopathology and molecular biology of ovarian epithelial tumors. *Ann Diagn Pathol*, 2: 195-207, 1998.
4. Crispens, M. A. Borderline ovarian tumours: a review of the recent literature. *Curr Opin Obstet Gynecol*, 15: 39-43, 2003.
5. Bast, R. C., Jr., Klug, T. L., Schaeztl, E., Lavin, P., Niloff, J. M., Greber, T. F., Zurawski, V. R., Jr., and Knapp, R. C. Monitoring human ovarian carcinoma with a combination of CA 125, CA 19-9, and carcinoembryonic antigen. *Am J Obstet Gynecol*, 149: 553-559, 1984.
6. Modugno, F. Ovarian cancer and high-risk women-implications for prevention, screening, and early detection. *Gynecol Oncol*, 91: 15-31, 2003.
7. Woolas, R. P., Xu, F. J., Jacobs, I. J., Yu, Y. H., Daly, L., Berchuck, A., Soper, J. T., Clarke-Pearson, D. L., Oram, D. H., and Bast, R. C., Jr. Elevation of multiple serum markers in patients with stage I ovarian cancer. *J Natl Cancer Inst*, 85: 1748-1751, 1993.
8. Meyer, T. and Rustin, G. J. Role of tumour markers in monitoring epithelial ovarian cancer. *Br J Cancer*, 82: 1535-1538, 2000.
9. Le Page, C., Provencher, D., Maugard, C. M., Ouellet, V., and Mes-Masson, A. M. Signature of a silent killer: expression profiling in epithelial ovarian cancer. *Expert Rev Mol Diagn*, 4: 157-167, 2004.
10. Mok, S. C., Chao, J., Skates, S., Wong, K., Yiu, G. K., Muto, M. G., Berkowitz, R. S., and Cramer, D. W. Prostatein, a potential serum marker for ovarian cancer: identification through microarray technology. *J Natl Cancer Inst*, 93: 1458-1464, 2001.
11. Kim, J. H., Skates, S. J., Uede, T., Wong, K. K., Schorge, J. O., Feltmate, C. M., Berkowitz, R. S., Cramer, D. W., and Mok, S. C. Osteopontin as a potential diagnostic biomarker for ovarian cancer. *Jama*, 287: 1671-1679, 2002.
12. McIntosh, M. W., Drescher, C., Karlan, B., Scholler, N., Urban, N., Hellstrom, K. E., and Hellstrom, I. Combining CA 125 and SMR serum markers for diagnosis and early detection of ovarian carcinoma. *Gynecol Oncol*, 95: 9-15, 2004.
13. Hellstrom, I., Raycraft, J., Hayden-Ledbetter, M., Ledbetter, J. A., Schummer, M., McIntosh, M., Drescher, C., Urban, N., and Hellstrom, K. E. The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. *Cancer Res*, 63: 3695-3700, 2003.
14. Woolas, R. P., Conaway, M. R., Xu, F., Jacobs, I. J., Yu, Y., Daly, L., Davies, A. P., O'Briant, K., Berchuck, A., Soper, J. T., and et al. Combinations of multiple serum markers are superior to individual assays for discriminating malignant from benign pelvic masses. *Gynecol Oncol*, 59: 111-116, 1995.

15. Kruk, P. A., Maines-Bandiera, S. L., and Auersperg, N. A simplified method to culture human ovarian surface epithelium. *Lab. Invest.*, 63: 132-136, 1990.
16. Lounis, H., Provencher, D., Godbout, C., Fink, D., Milot, M. J., and Mes-Masson, A. M. Primary cultures of normal and tumoral human ovarian epithelium: a powerful tool for basic molecular studies. *Exp. Cell Res.*, 215: 303-309, 1994.
17. Ouellet, V., Provencher, D., Maugard, C. M., Le Page, C., Ren, F., Lussier, C., Novak, J., Ge, B., Hudson, T. J., Tonin, P. N., and Mes-Masson, A. M. Discrimination between serous low malignant potential and invasive epithelial ovarian tumors using molecular profiling. *Oncogene*, In Press, 2005.
18. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*, 29: e45, 2001.
19. Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E. S., and Golub, T. R. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc Natl Acad Sci U S A*, 96: 2907-2912., 1999.

WHAT IS CLAIMED IS:

1. A method comprising:
 - (a) providing a biological sample from a subject (subject sample);and
 - (b) detecting the expression level of each of the markers FGF-2 and CA125 in the subject sample.
2. The method as recited in claim 1 wherein said subject is susceptible of having ovarian cancer.
3. The method as recited in any one of claims 1 and 2 wherein said subject is asymptomatic for ovarian cancer.
4. The method as recited in any one of claims 1 to 3, further comprising (c) comparing the expression level of each of the markers in the subject sample to corresponding pre-determined threshold expression levels for each of the markers, wherein an expression level of each of the markers in the subject sample that is higher than the pre-determined threshold expression levels for each of the markers is an indication that the subject is affected by ovarian cancer.
5. The method as recited in any one of claims 1 to 3, further comprising (c) comparing the expression level of each of the markers in the subject sample to the expression level of each of the markers in a control sample, wherein an expression level of each of the markers that is higher in the subject sample than in the control sample is an indication that the subject is affected by ovarian cancer.
6. The method as recited in any one of claims 1 to 3, further comprising (c) comparing the expression level of each of the markers in the subject sample to the expression level of each of the markers in a sample from the subject at an earlier time, wherein an expression level of each of the markers that is higher in the subject sample than in the sample from the subject at an earlier time is an indication that the subject is affected by ovarian cancer.

7. The method as recited in any one of claims 1 to 3, further comprising (c) comparing the expression level of each of the markers in the subject sample to the expression level of each of the markers in a non-cancerous sample from the subject, wherein an expression level of each of the markers that is higher in the subject sample than in the non-cancerous sample from the subject is an indication that the subject is affected by ovarian cancer.

8. The method as recited in claim 4, wherein the threshold expression level for each of the markers is determined by Receiver Operator Curves comparing the concentration of each of the markers in an ovarian cancer-free control population with that in a population affected by ovarian cancer.

9. The method as recited in any one of claims 1 to 8, wherein the expression is determined at the polypeptide level.

10. The method as recited in claim 9, wherein the expression is determined using an immunoassay.

11. The method as recited in claim 10, wherein said immunoassay is enzyme-linked immunosorbent assay (ELISA).

12. The method as recited in claim 10, wherein the expression level of each of the markers is above the following pre-determined threshold expression levels: 50 U/ml for CA125 and 37 pg/ml for FGF-2.

13. The method as recited in any one of claims 1 to 12, wherein step (b) further comprises detecting the concentration of marker IL-18 in the sample.

14. The method as recited in claim 12, wherein step (b) further comprises detecting the concentration of marker IL-18 in the sample, and wherein the expression level of IL-18 in the sample is above the pre-determined threshold expression level of 215 pg/ml for this marker.

15. The method as recited in any one of claims 1-14, wherein the

subject sample is a body fluid sample.

16. The method as recited claim 15, wherein the subject sample is selected from the group consisting of blood, plasma and serum.

17. The method as recited in claim 15, wherein said subject sample is serum.

18. The method as recited in any one of claims 1-17, wherein the subject sample is primary culture cells derived from an ovarian tumor sample from the subject.

19. The method as recited in any one of claims 1-18, wherein said subject is a human.

20. The method as recited in any one of claims 1-19, wherein said ovarian cancer is epithelial ovarian carcinoma (EOC).

21. The method as recited in any one of claims 1-20, wherein said method is *in vitro*.

22. A kit comprising means for detection of an expression level of each of markers CA125 and FGF-2 in a biological sample from a subject (subject sample), and instructions to use said markers in a method as recited in any one of claims 1 to 21.

23. The kit as recited in claim 22, further comprising means for detection of an expression level of marker IL-18.

24. The kit as recited in any one of claims 22 and 23, wherein said biological sample is selected from the group consisting of blood, plasma and serum.

25. The kit as recited in claim 24, wherein said biological sample is serum.

26. The kit as recited in any one of claims 22-25, wherein said subject is a human.

27. The kit as recited in any one of claims 22-26, wherein said ovarian cancer is epithelial ovarian carcinoma (EOC).

28. The kit as recited in any one of claims 22-27, wherein the means for detection of expression level of each of the markers are antibodies.

29. A method of assessing the potential efficacy of a test compound for treating or inhibiting ovarian cancer in a subject, said method comprising determining the expression level of each of markers CA125 and FGF-2 in a biological sample from the subject (subject sample), before and after administration of the test compound to the subject, wherein a decrease in the expression level of the markers after administration of the test compound is indicative that said test compound is effective for treating or inhibiting ovarian cancer.

30. A method of assessing the potential efficacy of a therapy for treating or inhibiting ovarian cancer in a subject, said method comprising determining the expression level of each of markers CA125 and FGF-2 in a biological sample from the subject (subject sample), before and after administration of said therapy in said subject, wherein a decrease in the expression level of said markers after administration of said therapy is indicative that said therapy is effective for treating or inhibiting ovarian cancer.

31. The method as recited in any one of claims 29 to 30, further comprising detecting the concentration of marker IL-18 in the sample.

32. The method as recited in any one of claims 29 to 31, wherein said expression level is determined at the polypeptide level.

33. The method as recited in claim 32, wherein said expression level is

determined using an immunoassay.

34. The method as recited in claim 33, wherein said expression level is determined using enzyme-linked immunosorbent assay (ELISA).

35. The method as recited in any one of claims 29-34, wherein said subject sample is selected from the group consisting of blood, plasma and serum.

36. The method as recited in claim 35, wherein said biological sample is serum.

37. The method as recited in any one of claims 29-36, wherein said subject is a human.

38. The method as recited in any one of claims 29-37, wherein said ovarian cancer is epithelial ovarian carcinoma.

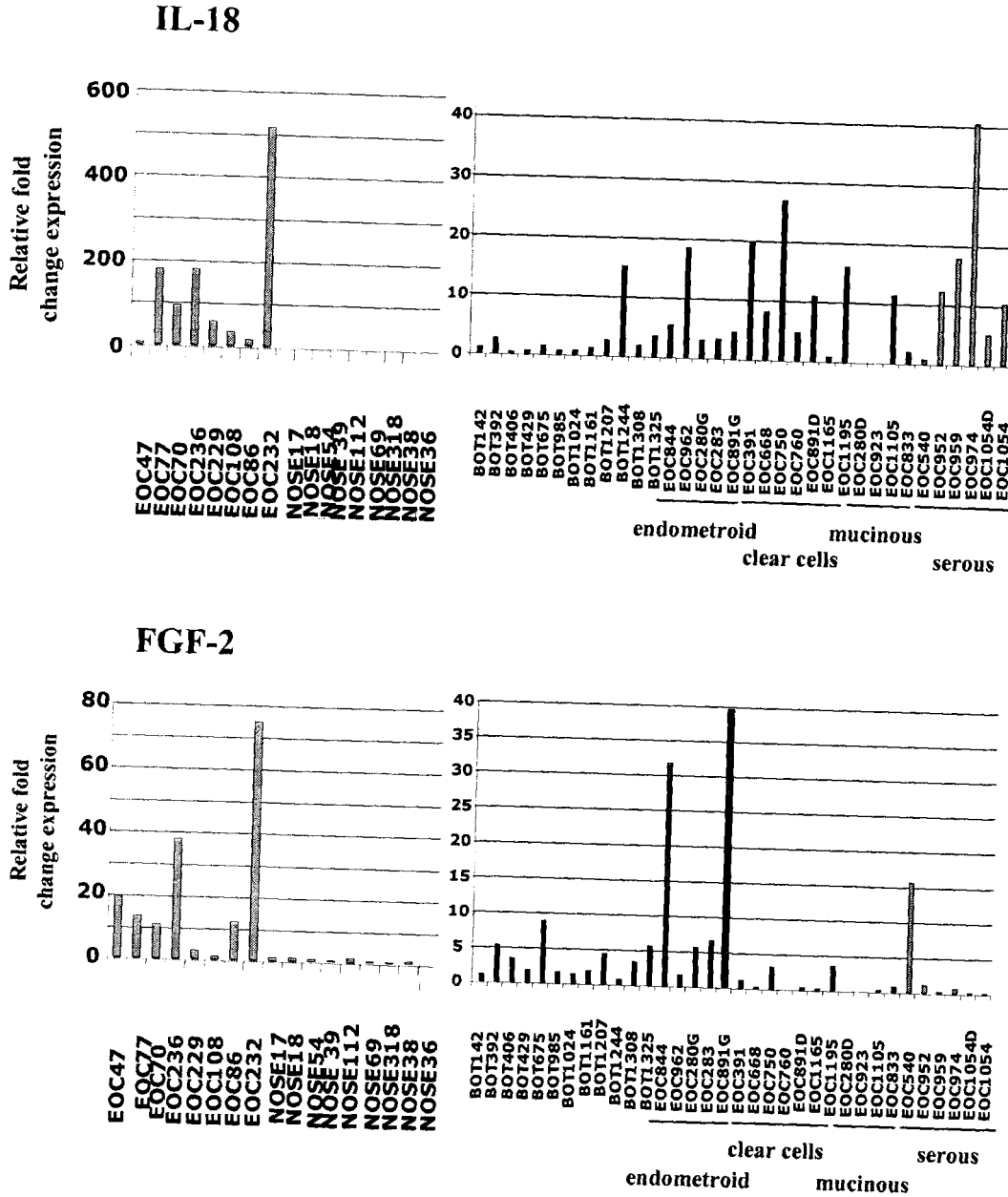


Figure 1a

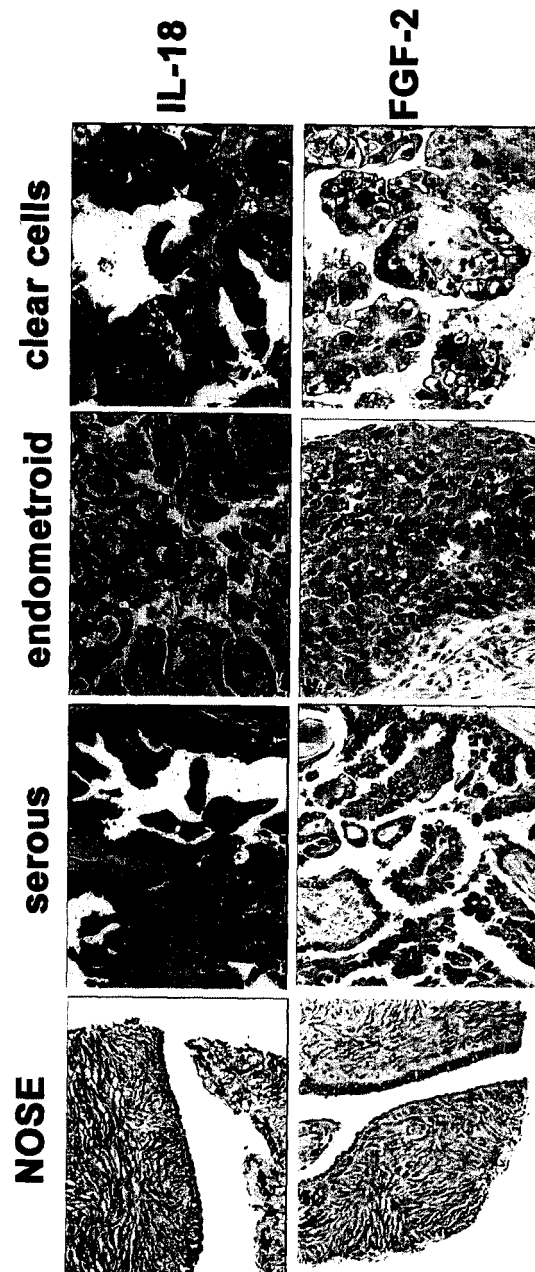


Figure 1b

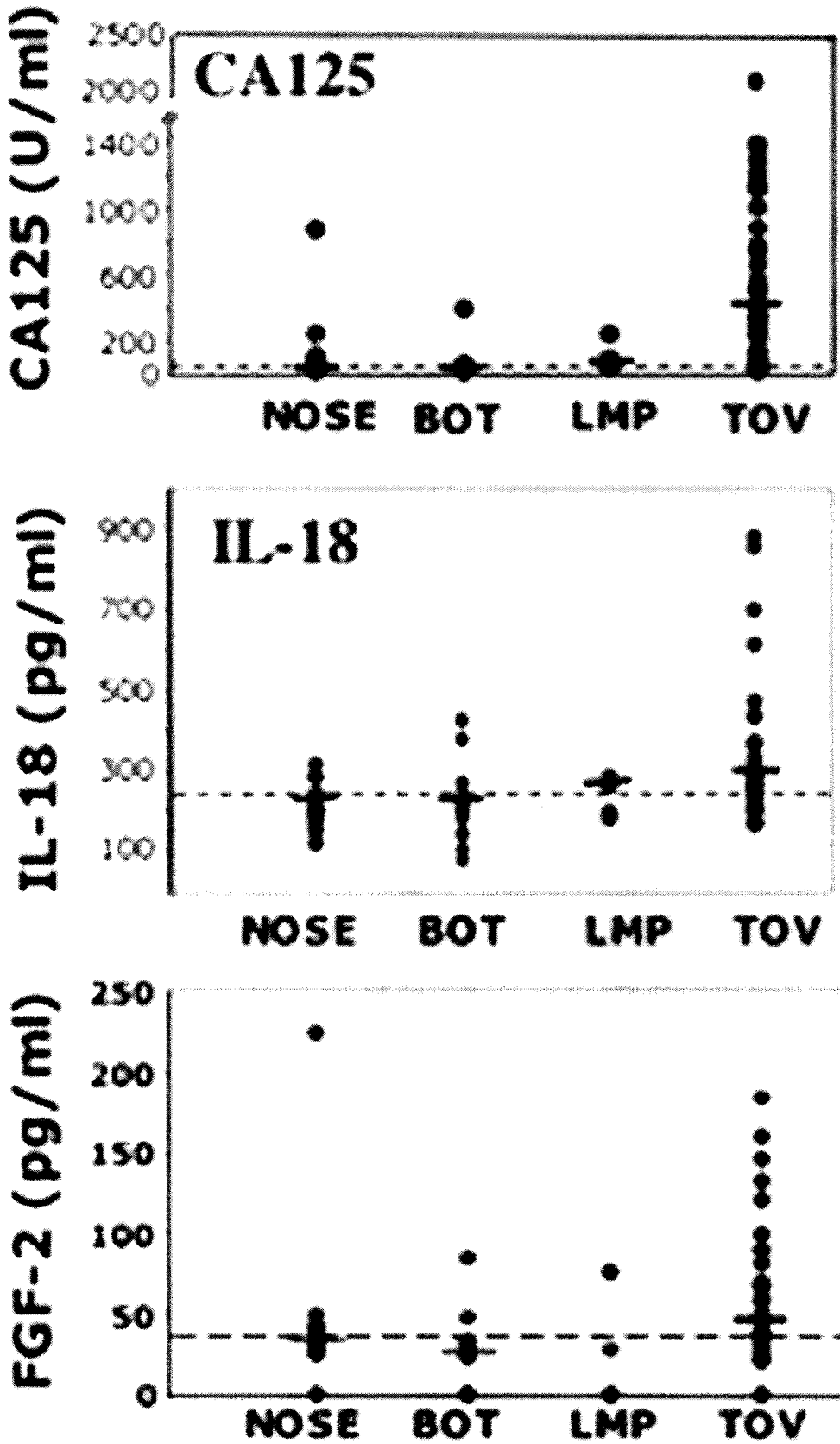


Figure 2

Nucleic acid and polypeptide sequences of CA125

Nucleic acid sequence (SEQ ID No:1)

```

1 aagcgttgca caattcccc aacctccata catacggcag ctcttctaga cacaggtttt
61 cccagggtcaa atgcggggac cccagccata tctcccaccc tgagaaattt tggagtttca
121 gggagctcag aagctctgca gaggccaccc tctctgaggg gattcttctt agacctccat
181 ccagaggcaa atgttgacct gtccatgctg aaaccctcag gccttctctg gtcattctct
241 cccacccgct ccttgatgac agggagcagg agcactaaag ccacaccaga aatggattca
301 ggactgacag gagccacctt gtcacctaag acatctacag gtgcaatcgt ggtgacagaa
361 catactctgc cctttacttc cccagataag accttggcca gtccctacac ttcggttgtg
421 ggaagaacca cccagtcttt gggggtgatg tcctctgctc tccctgagtc aacctctaga
481 ggaatgacac actccgagca aagaaccagc ccatcgctga gtccccaggt caatggaact
541 cctcttagga actaccctgc tacaagcatg gtttcaggat tgagttcccc aaggaccagg
601 accagttcca cagaaggaaa tttaccaa gaagcatcta catacacact cactgtagag
661 accacaagtg gcccagtcac tgagaagtac acagtcccca ctgagacctc aacaactgaa
721 ggtgacagca cagagacccc ctgggacaca agatatattc ctgtaaaaat cacatctcca
781 atgaaaacat ttgcagattc aactgcatcc aaggaaaatg ccccagtgtc tatgactcca
841 cctgagacca cagtactgca ctcacatact ccaggaagga caaacccatc atttgggaca
901 ctttattctt ccttccttga cctatcacct aaagggaccc caaattccag aggtgaaaca
961 agcctggaac tgattctatc aacctgga tatcccttct cctctctgca acctggctct
1021 gcaggacaca gcagaataag taccagtgcg cctttgtcat catctgcttc agttctcgat
1081 aataaaatat cagagaccag catattctca gccagagatc tcacctcccc tctgtctcct
1141 ggggtgcccg aggccagagc cagcacaatg cccaactcag ctatcccttt tccatgaca
1201 ctaagcaatg cagaacaag tgccgaaagg gtcagaagca caatttctct tctggggact
1261 ccatcaatat ccacaagca gacagcagag actatcctta ccttccatgc cctcgctgag
1321 accatggata taccagcac ccacatagcc aagactttgg cttcagaatg gttgggaagt
1381 ccaggtaccc ttggtggcac cagcacttca gcgctgacaa ccacatctcc atctaccact
1441 ttagtctcag aggagaccaa caccatcac tcacgagtg gaaaggaaac agaaggaaact
1501 ttgaatacat ctatgactcc acttgagacc tctgtctcct gagaagagtc cgaaatgact
1561 gccaccttgg tccccactct aggttttaca actcttgaca gcaagatcag aagtccatct
1621 caggtctctt catcccacc aacaagagag ctcagaacca caggcagcac ctctgggagg
1681 cagagttcca gcacagctgc ccacgggagc tctgacatcc tgagggcaac cacttccagc
1741 acctcaaaag catcatcatg gaccagtgaa agcacagctc agcaatttag tgaacccag
1801 cacacacagt ggggtggagc aagtcctagc atgaaaacag agagaccccc agcatcaacc
1861 agtgtggcag cccctatcac cacttctggt cctcagtggt tctctggctt caccaccctg
1921 aagaccagct ccacaaaagg gatttggctt gaagaaacat ctgcagacac actcatcgga
1981 gaatccacag ctggccaac caccatcag tttgtgttcc cactgggat ttcaatgaca
2041 ggaggcagca gcaccagggg aagccagggc acaaccacc tactcaccag agccacagca
2101 tcacttgaga catccgcaga tttgactctg gccacgaacg gtgtcccagt ctccgtgtct
2161 ccagcagtga gcaagacggc tegtgtctca agtcctccag gagggacaaa gccatcatat
2221 acaatggttt cttctgtcat cctgagaca tcactcttac agtctcagc tttcagggaa
2281 ggaaccagcc tgggactgac tccattaaac actagacatc ccttctcttc cctgaaacca
2341 gactctgcag gacacaccaa gataagcacc agcatttctc tgttgtcatc tgcttcagtt
2401 cttgaggata aagtgtcagc gaccagcaca ttctcacacc acaagccac ctcatctatt
2461 accacagggg ctctgaaat ctcaacaaag acaagccca gctcagccgt tctttctctc
2521 atgaccctaa gcaatgcagc aacaagtctt gaaagagtca gaaatgcaac tccccctctg
2581 actcatccat ctccatcagg ggaagagaca gcaggagtg tcctcactct cagcacctct
2641 gctgagacta cagactcacc taacatccac ccaactggga cactgacttc agaatcgtca
2701 gagagtecta gcaactctcag cctcccaagt gtctctggag tcaaaaccac attttcttca
2761 tctactcctt cactcatct atttactagt ggagaagaaa cagaggaaac ttcgaatcca
2821 tctgtgtctc aacctgagac ttctgtttcc agagtaagga ccaccttggc cagcacctct
2881 gtccctaacc cagtattccc caccatggac acctggccta cacgttcagc tcagttctct
2941 tcattcccacc tagtgagtga gctcagagct acgagcagta cctcagttac aaactcaact
3001 ggtcagctc ttctaaaaat atctcacctc actgggacgg caacaatgtc acagaccaat
3061 agagacacgt ttaatgactc tgtgcaccc caaagcaciaa ctgggccaga gactagtccc
3121 agattcaaga cagggttacc ttcagcaaca accactgttt caacctctgc cacttctctc
3181 tctgtactgt taatgtctc taaattcact tctccagcaa ctagtctccat ggaagcaact
3241 tctatcaggg aacctcaac aacctcctc acaacagaga ccacgaatgg ccaggtctct
3301 atggctgtgg cttctaccaa catcccaatt ggaagggct acattactga aggaagattg
3361 gacacaagcc atctgcccac tggaaaccaca gcttctctctg agacatctat ggattttacc

```

Figure 3A

3421 atggcceaag aaagtgtctc aatgtcagta tctccatctc agtccatgga tgtctgtggc
 3481 tcaagcactc caggaaggac aagccaattc gttgacacat tttctgatga tgtctatcat
 3541 ttaacatcca gagaaattac aatacctaga gatggaacaa gctcagctct gactccacaa
 3601 atgactgcaa ctccacctcc atctcctgat cctggctctg ctagaagcac ctggcttggc
 3661 atcttgtcct catctccttc ttctcctact cccaaagtca caatgagctc cacattttca
 3721 actcagagag tcaccacaag catgataatg gacacagttg aaactagtctg gtggaacatg
 3781 cccaacttac ctccacgac ttccctgaca ccaagtaata ttccaacaag tggtgccata
 3841 ggaaaaagca cctgggttcc cttggacact ccatctccag ccacatcatt ggaggcatca
 3901 gaagggggac ttccaacctt cagcacctac cctgaatcaa caaacacacc cagcatccac
 3961 ctcgagcac acgctagtcc agaaagtcca agcaccatca aacttaccat ggcttcagta
 4021 gtaaacctg gctcttacac acctctcacc tccccctcaa tagagaccca cattcatgta
 4081 tcaacagcca gaatggctta ctcttctggg tcttcacctg agatgacagc tcttgagag
 4141 actaacactg gtagtacctg ggaccccacc acctacatca ccactacgga tcctaaggat
 4201 acaagttcag ctcaggtctc tacaccccac tcagtgagga cactcagaac cacagaaac
 4261 catccaaaga cagagtcogc caccocagct gcttactctg gaagtcctaa aatctcaagt
 4321 tcaccaatc tcaccagtcc ggccacaaaa gcatggacca tcacagacac aactgaacac
 4381 tccactcaat tacattacac aaaattggca gaaaaatcat ctggatttga gacacagtca
 4441 gctccaggac ctgtctctgt agtaatccct acctccccta ccattggaag cagcacattg
 4501 gaactaactt ctgatgtccc aggggaaccc ctggctcttg ctcccagtga gcagaccaca
 4561 atcactctcc ccattggcaac atggctgagt accagtttga cagaggaaat ggcttcaaca
 4621 gaccttgata tttcaagtc aagttcacc atgagtacat ttgctatttt tccacctatg
 4681 tccacacctt ctcatgaact ttcaaagtca gaggcagata ccagtgccat tagaaataca
 4741 gattcaacaa cgttggatca gcacctagga atcaggagtt tgggcagaac tggggactta
 4801 acaactgttc ctatcaccac actgacaacc acgtggacca gtgtgattga aactcaaca
 4861 caagcacagg acaccctttc tgcaacgatg agtctactc acgtgacaca gtcactcaaa
 4921 gatcaaacat ctataccagc ctccagcatcc ccttcccac ttactgaagt ctaccctgag
 4981 ctcgggacac aagggagaag ctctctgag gcaaccactt tttggaacc atctacagac
 5041 acactgtoca gagagattga gactggccca acaaacattc aatccactcc acctatggac
 5101 aacacaacaa cagggagcag tagtagtggg gtcaccctgg gcatagccca ccttcccata
 5161 ggaacatcct cccagctga gacatccaca aacatggcac tggaaagaag aagttctaca
 5221 gccactgtct ctatggctgg gacaatggga ctccctgtta ctatgtctcc aggaagaagc
 5281 atcagccagt cattaaggaag agtttctct gtcctttctg agtcaactac tgaaggagtc
 5341 acagattcta gtaaggaag cagcccaagg ctgaacacac agggaaatac agctctctcc
 5401 tcctctcttg aaccagcta tgcggaagga agccagatga gcacaagcat cctctaacc
 5461 tcatctccta caactcctga tgtggaattc atagggggca gcacattttg gaccaaggag
 5521 gtcaccacag ttatgacctc agacatctcc aagtcttcag caaggacaga gtcacagtc
 5581 gctaccctta tgtccacagc tttgggaagc actgaaaata caggaaaaga aaaactcaga
 5641 actgcctcta tggatcttcc atctccaact ccatcaatgg agtgacacc atggatttct
 5701 ctcaactctca gtaatgcccc caataccaca gattcacttg acctcagcca tggggtgac
 5761 accagctctg cagggacttt ggccactgac aggtcattga atactggtg actagagcc
 5821 tccagattgg aaaacggctc tgatactct tctaagtccc tgtctatggg aaacagcact
 5881 cacacttcca tgactgacac agagaagagt gaagtgtctt cttcaatcca tccccgact
 5941 gagacctcag ctctggagc agagaccact ttgacttcca ctctggaaa cagggccata
 6001 agcttaacat tgcctttttc atccattcca gtggaagaag tcatttctac aggcataacc
 6061 tcaggaccag acatcaactc agcaccatg acacattctc ccatcaccac accaacaatt
 6121 gtagggacca gtacaggcac aattgaacag tccactcaac cactacatgc agtttcttca
 6181 gaaaaagttt ctgtgcagac acagtcaact ccatatgtca actctgtggc agtgtctgct
 6241 tccccaccac atgagaattc agtctcttct ggaagcagca catcctctcc atattctca
 6301 gcctcacttg aatccttga ttccacaatc agtaggagga atgcaatcac ttcttggtca
 6361 tgggacctca ctacatctct cccactaca acttggccaa gtactagttt atctgaggca
 6421 ctgtcctcag gccattctgg ggtttcaaac ccaagttcaa ctacgactga atttcactc
 6481 ttttcagctg catccacatc tgctgctaag caaagaaatc cagaacaga gacctatggt
 6541 cccagaata cagccgagcag tactttgac actgatgcat cctcgtcac aggtctttct
 6601 gagactcctg tgggggcaag tatcagctct gaagtccctc ttccaatggc cataacttct
 6661 agatcagatg tttctggcct tacatctgag agtactgcta acccgagttt aggcacagcc
 6721 tcttcagcag ggaccaaatt aactaggaca atatccctgc ccactcaga gcttttgggt
 6781 tcctttagaa tgaacaagga tccatggaca gtgtcaatcc ctttggggtc ccatccaact
 6841 actaatcag aaacaagcat cccagtaaac agcgcaggtc cacctggctt gtccacagta
 6901 gcatcagatg taattgacac accttcagat ggggctgaga gtattcccac tgtctcctt
 6961 tccccctccc ctgatactga agtgacaact atctcacatt tcccagaaaa gacaactcat
 7021 tcatttagaa ccatttcac tctcactcat gagttgactt caagagtgac acctattct
 7081 ggggatttga tgagttcagc tatgtctaca aagcccacag gagccagtc ctccattaca

Figure 3A (continued)

7141 ctgggagaga gaaggacaat cacctctgct gctccaacca cttccccat agttctcact
7201 gctagtttca cagagaccag cacagtttca ctggataatg aaactacagt aaaaacctca
7261 gatatccttg acgcacggaa acaaatgag ctcccctcag atagcagttc ttcttctgat
7321 ctgatcaaca cctccatagc ttcttcaact atggatgtca ctaaacacagc ctccatcagt
7381 cccactagca tctcaggaat gacagcaagt tcctccccat ctctcttctc ttcagataga
7441 ccccaggttc ccacatctac aacagagaca aatacagcca cctctccatc tgtttccagt
7501 aacacctatt ctcttgatgg gggctccaat gtgggtggca ctccatccac ttaccacc
7561 tttacaatca cccacctgt cgagacaagc tcggccctat tagcctggtc tagaccagta
7621 agaactttca gcaccatggt cagcactgac actgcctccg gagaaaatcc tacctctagc
7681 aattctgtgg tgacttctgt tccagacca ggtacatggg ccagtgtagg cagtactact
7741 gacttacctg ccattgggctt tctcaagaca agtccctgag gagaggcaca ctcacttcta
7801 gcatcaacta ttgaaccagc cactgccttc actccccatc tctcagcagc agtggtcact
7861 ggtaccagt ctacatcaga agccagtctt ctactacga gtgaaagcaa agccattcat
7921 tcttaccac agacccaac tacaccacc tctggagcaa actgggaaac ttcagctact
7981 cctgagagcc ttttgtagt cactgagact tcagacacaa cacttacctc aaagattttg
8041 gtcacagata ccattctgtt ttcaactgtg tccacgccac cttctaaatt tccaagtacg
8101 gggactctgt ctggacttc cttccctact ttactcccgg aactccagc catccctctc
8161 actgccactg agccaacaag ttcattagct acatcctttg attccacccc actggtgact
8221 atagcttcgg atagctttgg cacagtccca gagactacc tgaccatgct agagacctca
8281 aatggtgatg cactggttct taagacagta agtaaccag ataggagcat ccctggaatc
8341 actatccaag gagtaacaga aagtccactc catccttctt ccacttcccc ctctaagatt
8401 gttgctccac ggaatacaac ctatgaaggt tcgatcacag tggcactttc tactttgctt
8461 gcgggaacta ctggttccct tgtattcagt cagagttctg aaaactcaga gacaacggct
8521 ttggtagact catcagctgg gcttgagagg gcatctgtga tgccactaac cacaggaagc
8581 cagggatag ctagctctgg agaatcaga agtgggtcca ctcactcaac tggaaacaaa
8641 acattttctt ctctccctct gaccatgaac ccaggtgagg ttacagccat gctgaaatc
8701 accacgaaca gactgacagc tactcaatca acagaccca aagggatacc tgtgaagccc
8761 accagtgtct agtcaggcct cctaacacct gtctctgctt cctcaagccc atcaaaggcc
8821 tttgcctcac tgactacagc tccccatca acttggggga tcccacagtc taccttgaca
8881 tttgagttt ctgaggtccc aagtttgat actaagtccg cttctttacc aactcctgga
8941 cagtccctga acaccattcc agactcagat gcaagcacag catcttctc actgtccaag
9001 tctccagaaa aaaaccaag ggcaaggatg atgacttcca caaaggccat aagtgaagc
9061 tcatttcaat caacagttt tactgaaacc cctgagggat ctgcctccc ttctatggca
9121 gggcatgaac ccagagtccc cacttcagga acaggggacc ctgatatgc ctgagagagc
9181 atgtcttatc cagacccaag caaggcatca tcagctatga catcgacctc tcttgcatca
9241 aaactcacia ctctctcag cacaggtcaa gcagcaagg ctggttctag ttctctccc
9301 ataagcctat cactgagaa agaaacaagc ttctttccc cactgcatc cacctccaga
9361 aagacttcac tatttcttg gccttccatg gcaaggcagc ccaacatatt ggtgcatctt
9421 gactctcag ctctgacact ttctccaaca tccactctaa atatgtcca ggaggagcct
9481 cctgagttaa cctcaagcca gaccattgca gaagaaggagg gaacaacagc tgaacacag
9541 acgttaacct tcacaccatc tgagaccca acatccttgt tacctgtctc ttctcccaca
9601 gaaccacag ccagaagaaa gatttctcca gaaacatggg caagctctat ttcagttctt
9661 gccaaagcct ccttggttga aacaactgat ggaacgctag tgaccaccat aaagatgtca
9721 agccaggcag cacaaggaaa ttccactggg cctgccccag cagaggagac ggggaccagt
9781 ccagcaggca catcccagc aagccagaa gtgtctacca ctctcaaaa catgagctcc
9841 aaggaacca gcatcagccc agagatcagg tccactgtgc tccacttctc ttggaagact
9901 ccagaaacaa ctggttccat ggagaccaca gtggaaccag tcacccttca gtcacagcc
9961 ctaggaagtg gcagcaccag catctctcac ctgcccacag gaaccacatc accaaccaag
10021 tcaccaacag aaaatatgtt ggctacagaa aggtctccc tctccccatc cccactgag
10081 gcttgacca accttattc tggactcca ggaggacca ggcagtcact gccacaatg
10141 tctctgtct ccctagagtc accaactgct agaagcatca cagggactgg tcagcaaagc
10201 agtccagaac tggtttcaa gacaactgga atggaattct ctatgtggca tggctctact
10261 ggagggacca caggggacac acatgtctct ctgagccat cttccaatat cctgaagac
10321 cctgtaacca gcccaactc tgtgagctca ttgacagata aatccaaaca taaaaccgag
10381 acatgggtaa gcaccacagc cattccctcc actgtcctga ataataagat aatggcagct
10441 gaacaacaga caagtcgac tgtgatgag gcttattcat caactagttc ttggtcagat
10501 cagacatctg ggagtgcac cacccttggg gcatctctg atgtcacaaa cacattatac
10561 atcacctcca cagcacaac caccctacta gtgtctctgc cctctggaga ccaaggcatt
10621 acaagcctca ccaatcctc aggagaaaa acaagctctg cgtcatctgt catctctct
10681 tcaatagggc ttgagactct gagggcaat gtaagtgcag tgaaggtga cattgcccct
10741 actgctgggc atctatctca gacttcatct cctgcggaag tgagcatcct ggacgtaacc
10801 acagctccta ctccaggtat ctccaccacc atcaccacca tgggaaccaa ctcaatctca

Figure 3A (continued)

10861 actaccacac ccaaccaga agtgggtatg agtaccatgg acagcacccc gggccacagag
 10921 aggcgcacaa cttctacaga acacccttcc acctgggtctt ccacagctgc atcagattcc
 10981 tggactgtca cagacatgac ttcaaacttg aaagttgcaa gatctcctgg aacaatttcc
 11041 acaatgcata caacttcatt cttagcctca agcactgaat tagactccat gtctactccc
 11101 catggccgta taactgtcat tggaccagc ctggtcactc catcctctga tgcctcagct
 11161 gtaaagacag agaccagtac aagtgaaga acattgagtc cttcagacac aactgcatct
 11221 actcccattc caacttttcc tctgtccag aggatgagca tctcagttcc tgacatttta
 11281 agtacaagtt ggactccag tagtacagaa gcagaagatg tgcctgtttc aatggtttct
 11341 acagatcatg ctagtacaaa gactgaccca aatacgcccc tgtccacttt tctgtttgat
 11401 tctctgtcca ctcttgactg ggacactggg agatctctgt catcagccac agccactacc
 11461 tcagctcctc agggggccac aactccccag gaactcactt tggaaacat gatcagccca
 11521 gctacctcac agttgccctt ctctataggg cacattacaa gtgcagtcac accagctgca
 11581 atggcaagga gctctggagt tactttttca agaccagatc ccacaagcaa aaaggcagag
 11641 cagacttcca ctcagcttcc caccaccact tctgcacatc cagggcaggt gccacatca
 11701 gcagcaacaa ctctggatgt gatccacac acagcaaaaa ctccagatgc aacttttcag
 11761 agacaagggc agacagctct tacaacagag gcaagagcta catctgactc ctggaatgag
 11821 aaagaaaaat caaccocaag tgcaccttgg atcactgaga tgatgaattc tgtctcagaa
 11881 gataccatca aggaggttac cagctcctcc agtgtattaa aggacctga atacgctgga
 11941 cataaacttg gaatctggga cgacttcac cccaagtttg gaaaagcagc ccatatgaga
 12001 gagttgcccc ttctgagtcc accacaggac aaagaggcaa tccacccttc tacaacaca
 12061 gttagagacca caggctgggt cacaagtccc gaacatgctt ctattccac tatcccagcc
 12121 cactcagcgt catccaaact cacatctcca gtggttacia cctccaccag ggaacaagca
 12181 atagtttcta tgtcaacaac cacatggcca gagtctacia gggctagaac agagccta
 12241 tccttcttga ctatgaaact gagggacgtc agccttaca tggacaccag ctcaaccaca
 12301 caaacaagta ttatctcttc cccaggttcc actgcgatca ccaaggggccc tagaacagaa
 12361 attacctcct ctaagagaat atccagctca ttccctgccc agtctatgag gtcgtcagac
 12421 agccctcag aagccatcac caggctgtct aactttctctg ccatgacaga atctggagga
 12481 atgatccttg ctatgcaaac aagtccacct ggcgctacat cactaagtgc acctacttg
 12541 gatacatcag ccacagcctc ctggacaggg actccactgg ctacgactca gagatttaca
 12601 tactcagaga agaccactct ctttagcaaa ggtcctgagg atacatcaca gccaaagcct
 12661 ccctctgtgg aagaaaccag ctcttctctt tccttggtac ctatccatgc tacaacctcg
 12721 cctccaataa tttgttgac atcacaaggg cacagtccct cctctactcc acctgtgacc
 12781 ttcagtttct tctctgagac ctctggcctg ggaagacca cagacatgtc gaggataagc
 12841 ttggaacctg gcacaagttt acctcccaat ttgagcagta cagcaggtga ggcgttatcc
 12901 acttatgaag cctccagaga tacaagga attcatcatt ctgcagacac agcagtgacg
 12961 aatatggagg caaccagttc tgaatattct cctatcccag gccatacaaa gccatccaaa
 13021 gccacatctc cattggttac ctcccacatc atgggggaca tcacttcttc cacatcagta
 13081 tttggctcct ccgagaccac agagattgag acagtgtcct ctgtgaacca gggacttcag
 13141 gagagaagca catcccaggt ggccagctct gctacagaga caagcactgt cattaccat
 13201 gttctagtgt gtgatgtctac tactcatgtc accaagacac aagccacttt ctctagcggg
 13261 acatccatct caagccctca tcagtttata acttctacca acacatttac agatgtgagc
 13321 accaaccctc ccacctctct gataatgaca gaatcttcag gagtgaccat caccaccaa
 13381 acaggtccta ctggagctgc aacacagggg ccatatctct tggacacatc aacctgctc
 13441 tacttgacag agactccatt agctgtgact ccagatttta tgcaatcaga gaagaccact
 13501 ctcataagca aaggtcccaa ggatgtgacc tggacaagcc ctccctctgt ggcagaacc
 13561 agctatccct ctccctgac acctttcttg gtcacaacca tacctctgc cacttccacg
 13621 ttacaagggc aacatacatc ctctcctggt tctgagactt cagttcttac ctctggactg
 13681 gtgaagacca cagatatggt gaacacaagc atggaacctg tgaccaattc acctcaaat
 13741 ttgaacaatc catcaaatga gatactggcc actttggcag ccaccacaga tatagagact
 13801 attcatcctt ccataaacia agcagtgacc aatatgggga ctgccagttc agcagatgta
 13861 ctgcattcca ctctcccagt cagctcagaa ccatctacag ccacatctcc aatggttctc
 13921 gcctccagca tgggggacgc tcttgcttct atatacaatc atggttctga gaccacagac
 13981 attgaggag agccaacatc ctcccgtact gctggagcaa aagagaacag caccctccag
 14041 gatgtaact caactacaga gtcaaacatc atcctctcca atgtgtctgt gggggctatt
 14101 actgaagcca caaaaatgga agtcccctct tttgatgcaa cattcatacc aactcctgct
 14161 cagtcaacia agttcccaga ttttttctca gtagccagca gtagactttc aaactctcct
 14221 cccatgacaa tatctacca catgaccacc acccagacag ggtcttctgg agctacatca
 14281 aagattccac ttgccttaga cacatcaacc ttggaaacct cagcagggac tccatcagtg
 14341 gtgactgagg ggtttgccca ctcaaaaata accactgcaa tgaacaatga tctcaaggac
 14401 gtgtcacaga caaacctcc ctcttcagat gaagccagct ctccctcttc tcaagcact
 14461 gtccttgtca caaccttacc ttcttctgtt gctttcacac cgcaatggca cagtacctcc
 14521 tctctgttt ctatgtctc agttcttact tcttctactgg taaagaccgc aggcaaggtg
 14581 gatacaagct tagaaacagt gaccagttca cctcaaagta tgagcaaac tttggatgac
 14641 atatcggctca cttcagcagc caccacagat atagagacaa cgcactcttc cataaacaca

Figure 3A (continued)

14701 gtagttacca atgtggggac caccggttca gcatttgaat cacattctac tgtctcagct
 14761 taccagagc catctaaagt cacatctcca aatgttacca cctccaccat ggaagacacc
 14821 acaatttccc gatcaatacc taaatcctct aagactacaa gaactgagac tgagacaact
 14881 tcctccctga ctccctaaact gagggagacc agcatctccc aggagatcac ctcgtccaca
 14941 gagacaagca ctgttcctta caaagagctc actggtgcca ctaccgaggt atccaggaca
 15001 gatgtcactt cctctagcag tacatccttc cctggccctg atcagtccac agtgtcacta
 15061 gacatctcca cagaaaccaa caccaggctg tctacctccc caataatgac agaatctgca
 15121 gaaataacca tcaccaccca aacaggctct catggggcta catcacagga tacttttacc
 15181 atggaccat caaatacaac cccccaggca gggatccact cagctatgac tcatggattt
 15241 tcacaattgg atgtgaccac tcttatgagc agaattccac aggatgtatc atggacaagt
 15301 cctccctctg tggataaaac cagctcccc tcttctcttc tgtcctcacc tgcaatgacc
 15361 acaccttccc tgatttcttc taccttacca gaggataagc tctcctctcc tatgacttca
 15421 cttctcacct ctggcctagt gaagattaca gacatattac gtacacgctt ggaactgtg
 15481 accagctcac ttccaaatth caccagcacc tcagataaga tactggccac ttctaaagac
 15541 agtaaagaca caaaggaaat ttttctctct ataaacacag aagagaccaa tgtgaaagcc
 15601 aacaactctg gacatgaatc ccattcccct gcactggctg actcagagac acccaaagcc
 15661 acaactcaaa tggttatcac caccactgtg ggagatccag ctccctccac atcaatgcca
 15721 gtgcatggtt cctctgagac tacaacatt aagagagagc caacatattt cttgactcct
 15781 agactgagag agaccagtac ctctcaggag tccagcttcc ccacggacac aagttttcta
 15841 ctttccaaag tcccactgg tactattact gaggtctcca gtacaggggt caactcttct
 15901 agcaaaatth ccaccccaga ccatgataag tccacagtgc cacctgacac cttcacagga
 15961 gagatcccca ggtctctcac ctctcttatt aagacaaaat ctgcagaaat gacgatcacc
 16021 acccaagcaa gtccctctga gtctgcatcg cacagtacc tcccttggga cacatcaacc
 16081 acactttccc agggagggac tcattcaact gtgactcagg gattcccata ctcagaggtg
 16141 accactctca tgggcatggg tcttggaat gtgtcatgga tgacaactcc ccctgtggaa
 16201 gaaaccagct ctgtgtcttc cctgatgtct tcacctgcca tgacatcccc ttctctgtt
 16261 tcctccacat caccacagag catcccctcc tctcctcttc ctgtgactgc acttctact
 16321 tctgttctg tgacaaccac agatgtgttg ggcacaacaa gccagagtc tghtaaccagt
 16381 tcacctcaa atttgagcag catcactcat gagagaccgg ccacttcaa agacactgca
 16441 cacacagaag ccgccatgca tcattccaca aacaccgag tgaccaatgt agggacttcc
 16501 ggtcttgga ataaatcaca atcctctgtc ctgactgact cagagacatc gaaagccaca
 16561 cctctgatga gtaccacctc caccctgggg gacacaagt tttccacat aactccta
 16621 atctctcaga ctaaccaaat tcaaacagag ccaacagcat cctgagccc tagactgag
 16681 gagagcagca cgtctgagaa gaccagctca acaacagaga caaatactgc ctttcttat
 16741 gtgcccacag gtgctattac tcaggcctcc agaacagaaa tctcctctag cagaacatcc
 16801 atctcagacc ttgatcggcc cacaatagca cccgacatct ccacaggaat gatcaccagg
 16861 ctcttcacct ccccatcat gacaaaatct gcagaaatga ccgtcaccac tcaaacact
 16921 actcctgggg ctacatcaca ggtatcctt ccttgggaca catcaaccac acttttccag
 16981 ggagggactc attcaaccgt gtctcagga tttccacact cagagataac cactcttcgg
 17041 agcagaacct ctggagatgt gtcattgag acaactcccc ctgtggaaga aaccagctct
 17101 ggttttccc tgatgtcacc ttccatgaca tccccttctc ctgtttctct cacatcaca
 17161 gagagcatcc cctcctctcc tctcctctgt actgcacttc ttacttctgt tctggtgaca
 17221 accaccaatg tattgggac aacaagccca gagaccgtaa cgagttcacc tccaaattta
 17281 agcagcccca cacaggagag actgaccact tacaagaca ctgagcacac agaagccatg
 17341 catgcttcca tgcatacaaa cactgcagtg gccaacgtcg ggacctccat ttctggacat
 17401 gaatcacaat cttctgtccc agctgattca cacacatcca aagccacatc tccaatgggt
 17461 atcaccttgc ccatggggga tacaagtgt tctacatcaa ctctgcctt ctttgagact
 17521 agaattcaga ctgaatcaac atcctctttg attcctggat taagggacac caggacgtct
 17581 gaggagatca acactgtgac agagaccagc actgtccttt cagaagtgcc cactactact
 17641 actactgagg tctccaggac agaagtatc acttccagca gaacaaccat ctcagggcct
 17701 gatcattcca aatgtcacc ctacatctcc acagaaacca tcaccaggct ctccactttt
 17761 ccttttghaa caggatccac agaaatggcc atcaccaacc aaacaggctc tatagggact
 17821 atctcacag ctacccttac cctggacaca tcaagcacag cttcctggga agggactcac
 17881 tcacctgtga ctcagagatt tccacactca gaggagacca ctactatgag cagaagtact
 17941 aagggcgtgt catggcaaa cctcctctct gtggaagaaa ccagttctcc ttcttccca
 18001 gtgcctttac ctgcaataac ctcacattca tctctttatt ccgagatc aggaagtagc
 18061 cccacttctg ctctcctctg gacttccctt ctcacctctg gcaggaggaa gaccatagac
 18121 atgttgagca cacactcaga acttgtgacc agctccttac caagtgaag tagcttctca
 18181 ggtgagatac tcacttctga agcctccaca aatacagaga caattcactt ttcagagaa
 18241 acagcagaaa ccaataggg gaccaccaat tctatgcata aactacattc ctctgtctca

Figure 3A (continued)

18301 atccactccc agccatccgg acacacacct ccaaaggtta ctggatctat gatggaggac
 18361 gctattgttt ccacatcaac acctggttct cctgagacta aaaatggtga cagagactca
 18421 acatcccctc tgactcctga actgaaagag gacagaccg ccctggtgat gaactcaact
 18481 acagagtcaa aactgtttt ctccagtgtg tcctggatg ctgctactga ggtctccagg
 18541 gcagaagtca cctactatga tcctacattc atgccagctt ctgctcagtc acaaaagtcc
 18601 ccagacattt ccactgaagc cagcagcagt cattctaact ctctccctt gacaatatct
 18661 acacacaaga ccactgccac acaaacaggc ccttctgggg tgacatctct tggccaactg
 18721 accctggaca catcaacct agccacctca gcaggaactc catcagccag aactcaggat
 18781 tttgtagatt cagaacaac cagtgtcatg aacaatgatc tcaatgatgt gttgaagaca
 18841 agccctttct ctgcagaaga agccaactct ctctcttctc aggcacctct ccttgtgaca
 18901 acctcacctt ctctgtaac ttccacattg caagagcaca gtacctctc tcttgtttct
 18961 gtgacctcag taccacccc tacactggcg aagatcacag acatggacac aaacttagaa
 19021 cctgtgactc gttcacctca aaatttaagg aacaccttgg ccacttcaga agccaccaca
 19081 gatacacaca caatgcatcc ttctataaac acagcaatgg ccaatgtggg gaccaccagt
 19141 tcaccaaagt aattctattt tactgtctca cctgactcag acccatataa agccacatcc
 19201 gcagtagtta tcacttccac ctcgggggac tcaatagttt ccacatcaat gcctagatcc
 19261 tctgcatgta aaaagattga gtctgagaca actttctccc tgatatttag actgagggag
 19321 actagcacct ccagaaaaat tggctcatcc tcagacacaa gcacggtctt tgacaaagca
 19381 tctactgtct ctactactga ggtctccaga acagaactca cctcctctag cagaacatcc
 19441 atccaaggca ctgaaaagcc cacaatgtca cgggacacct ccacaagatc tgtcaccatg
 19501 ctttctactt ttgctggcct gacaaaatcc gaagaaagga ccattgccac ccaaacaggc
 19561 cctcataggg cgacatcaca ggtaccctt acctgggaca catcaatcac aacctcacag
 19621 gcagggacc cactcagctat gactcatgga ttttcacaat tagatattgtc cactcttacg
 19681 agtagagttc ctgagtacat atcagggaca agcccacct ctgtggaaaa aaccagctct
 19741 tcctcttccc ttctgtcttt accagcaata acctcaccgt cccctgtacc tactacatta
 19801 ccagaaagta ggcgtcttc tcctgttcat ctgacttcac tcccacctc tggcctagtg
 19861 aagaccacag atatgctggc atctgtggcc agtttacctc caaacttggg cagcacctca
 19921 cataagatac cgactacttc agaagacatt aaagatacag agaaaatgta tccttccaca
 19981 aacatagcag taaccaatgt ggggaccacc acttctgaaa aggaatctta ttgctctgtc
 20041 ccagcctact cagaaccacc caaagtcacc tctccaatgg ttacctctt caacataagg
 20101 gacaccattg ttccacatc catgcctggc tcctctgaga ttacaaggat tgagatggag
 20161 tcaacattct ccgtggctca tgggtgaaag ggaaccagca cctcccagga ccccctgta
 20221 tcacacagaa aaagtgtgtt ccttcacaag ctgactcagtg gtgactatga gacctctagg
 20281 acagaagtgt cctctctag aagaacatcc attccaggcc ctgatcattc cacagagtca
 20341 ccagacatct ccactgaagt gatcccagc ctgctatct cccttggcat tacagaatct
 20401 tcaaatatga ccactcatc tcgaacaggt cctcctcttg gctctacatc acagggcaca
 20461 ttacacttgg acacaccaac tacatcctcc agggcaggaa cacactcgat ggcgactcag
 20521 gaatttccac actcagaaat gaccactgtc atgaacaagg accctgagat tctatcatgg
 20581 acaatocctc cttctataga gaaaaccagc ttctcctctt cctgatgcc ttcaccagcc
 20641 atgacttcac ctctgtttc ctcaacatta ccaaagacca ttcacaccac tcttctctc
 20701 atgacctcac tgctacccc tagcctagt atgaccacag acacattggg cacaagccca
 20761 gaacctaaa ccagttcacc tccaaatttg agcagtacct cacatgtgat actgacaaca
 20821 gatgaagaca ccacagctat agaagccatg catccttcca caagcacagc agcgactaat
 20881 gtggaaacca cctgttcttg acatgggtca caatcctctg tcctaactga ctcagaaaaa
 20941 accaaggcca cagctccaat ggataccacc tccacatgg ggcatacaac tgtttccaca
 21001 tcaatgtctg ttctctcga gactacaaaa attaagagag agtcaacata tctcttgact
 21061 cctggactga gagagaccag catttcccaa aatgccagct tttccactga cacaagtatt
 21121 gttctttcag aagtccccac tggactact gctgaggtct ccaggacaga agtcacctcc
 21181 tctggtagaa catccatccc tggccttctc cagtccacag ttttgccaga aatatccaca
 21241 agaacaatga caaggtctt tgcctgccc accatgacag aatcagcaga aatgaccatc
 21301 cccactcaaa caggtccttc tgggtctacc tcacaggata cccttacctc ggacacatcc
 21361 accacaaagt cccaggcaaa gactcattca actttgactc agagatttcc acactcagag
 21421 atgaccactc tcatgagcag aggtcctgga gatatgtcat ggcaaagctc tccctctctg
 21481 gaaaatccca gctctctccc ttccctgctg tctttacctg ccacaacctc acctcctccc
 21541 atttctcca cattaccagt gactatctcc tctctctctc ttctgtgac ttcacttctc
 21601 acctctagcc cgtaacgac cacagacatg ttacacacaa gcccagaact tgtaaccagt
 21661 tcacctcaa agctgagcca cacttcagat gagagactga ccaactggca ggacaccaca
 21721 aatacagaag ctgtgcattc ttccacaaac acagcagcgt ccaatgtgga gattcccagc
 21781 tttggacatg aatccccctc ctctgcectta gctgctcag agacatccaa agccacatca
 21841 ccaatgttta ttacctccac ccaggaggat acaactgttg ccatatcaac ccctcacttc
 21901 ttggagacta gcagaattca gaaagagtca atttctccc tgagccctaa attgagggag
 21961 acaggcagtt ctgtggagac aagctcagcc atagagacaa gtgctgtcct tctgaagtg
 22021 tccattgggt ctactactga gatctccagg acagaagtca cctcctctag cagaacatcc
 22081 atctctggtt ctgctgagtc cacaatgttg ccagaaatat ccaccacaag aaaaatcatt

Figure 3A (continued)

22141 aagttcccta cttccccat cctggcagaa tcatcagaaa tgaccatcaa gacccaaaca
 22201 agtccctcctg ggtctacatc agagagtacc tttacattag acacatcaac cactccctcc
 22261 ttggtaataa cccattcgac tatgactcag agattgccac actcagagat aaccactctt
 22321 gtgagtagag gtgctgggga tgtgccacgg cccagctctc tccctgtgga agaaacaagc
 22381 cctccatctt cccagctgtc tttatctgcc atgatctcac cttctcctgt ttcttcaca
 22441 ttaccagcaa gtagccactc ctcttctgct tctgtgactt cacctctcac accaggcaa
 22501 gtgaagacta ctgaggtgtt ggacgcaagt gcagaacctg aaaccagttc acctccaagt
 22561 ttgagcagca cctcagttga aatactggcc acctctgaag tcaccacaga tacggagaaa
 22621 attcctcctt tcccaaacac ggcagtaacc aaagttggaa cttccagttc tggacatgaa
 22681 tcccttctct ctgtcctacc tgactcagag acaaccaaag ccacatcggc aatgggtacc
 22741 atctccatta tgggggatac aagtgtttct acattaactc ctgccttatac taacactagg
 22801 aaaattcagt cagagccagc ttcctcactg accaccagat tgagggagac cagcacctct
 22861 gaagagacca gcttagccac agaagcaaac actggtcttt ctaaagtgtc cactgggtct
 22921 actactgagg tctccaggac agaagcctac tcttttagca gaacatccat gtcaggccct
 22981 gagcagtcca caatgtcaca agacatctcc ataggaacca tccccaggat ttctgcctcc
 23041 tctgtcctga cagaatctgc aaaaatgacc atcacaaccc aaacaggtcc ttcggagtct
 23101 aactagaaa gtacccttaa tttgaacaca gcaaccacac cctcttgggt ggaaaccac
 23161 tctatagtaa ttcagggatt tccacaccca gagatgacca cttccatggg cagaggtcct
 23221 ggaggtgtgt catggcctag ccctcccttt gtgaaagaaa ccagccctcc atcctccccg
 23281 ctgtctttac ctgccgtgac ctcacctcat cctgtttcca ccacattcct agcacaatc
 23341 cccccctctc cccttctgt gacttcaact ctcacctctg gcccgggcag aaccacagat
 23401 atcttgggta caagcacaga acctggaacc agttcatctt caagtttgag caccacctcc
 23461 catgagagac tgaccactta caaagacact gcacatacag aagccgtgca tcttccaca
 23521 aacacaggag ggaccaatgt ggcaaccacc agctctggat ataaatcaca gtcctctgtc
 23581 ctgactgact catctccaat gtgtaccacc tccaccatgg gggatacaag tgttctcaca
 23641 tcaactcctg ccttccctga gactaggagg attcagacag agctagcttc ctccctgacc
 23701 cctggattga gggagtcag tggctctgaa gggaccagct caggcaccaa gatgagact
 23761 gtcctctcta aagtgccac tgggtgctact actgagatct ccaaggaaga cgtcacctcc
 23821 atcccaggtc ccgctcaatc cacaatatca ccagacatct ccacaagaac cgtcagctgg
 23881 ttctctacat ccctgtcat gacagaatca gcagaaataa ccatgaacac ccatacaagt
 23941 cctttagggg ccacaacaca aggcaccagt actttggcca cgtcaagcac aacctctttg
 24001 acaatgacac actcaactat atctcaagga ttttcaact cacagatgag cactcttatg
 24061 aggaggggtc ctgaggatgt atcatggatg agccctcccc tctggaaaa aactagacct
 24121 tctttttctc tgatgtcttc accagccaca acttcaactt ctctgtttc ctccacatta
 24181 ccagagagca tctcttctc tctcttctc gtgacttca tctcactgc ttgcttggca
 24241 aaaactacag atatgttgca caaagctca gaacctgtaa ccaactcacc tgcaaatttg
 24301 agcagcacct cagttgaaat actggccacc tctgaagtca ccacagatac agagaaaact
 24361 catccttctt caaacagaac agtgaccgat gtggggacct ccagttctgg acatgaatcc
 24421 acttcccttg tctagctga tccacagaca tccaaagtca catctccaat ggttattacc
 24481 tccaccatgg aggatacag tgtctcaca tcaactcctg gcttttttga gactagcaga
 24541 attcagacag aaccaacatc ctccctgacc cttggactga gaaagaccag cagctctgag
 24601 gggaccagct tagccacaga gatgagcact gtcctttctg gagtgccac ttggtgccact
 24661 gctgaagtct ccaggacaga agtcacctcc tctagcagaa catccatctc aggctttgct
 24721 cagctcacag tgtcaccaga gacttccaca gaaaccatca ccagactccc tacctccagc
 24781 ataatgacag aatcagcaga aatgatgatc aagacacaaa cagatcctcc tgggtctaca
 24841 ccagagagta ctcaactgt ggacatatca acaacaccca actgggtaga aaccactcg
 24901 actgtgactc agagattttc aactcagag atgaccactc ttgtgagcag aagccctggt
 24961 gatatgttat ggcctagtca atcctctgtg gaagaaacca gctctgcctc ttccctgctg
 25021 tctctgctg ccacgacctc acctctctct gtttctctc cattagtaga ggatttccct
 25081 tccgcttctc ttctgtgac ttcttctctc accctggcc ttggtgataac cacagacagg
 25141 atgggcataa gcagagaacc tggaaaccagt tccacttcaa atttgagcag cacctccat
 25201 gagagactga ccactttgga agacactgta gatacagaag acatgcagcc ttccacacac
 25261 acagcagtga ccaacgtgag gacctccatt tctggacatg aatcacaatc ttctgtccta
 25321 tctgactcag agacacccaa agccacatct ccaatgggta ccacctacac catgggggaa
 25381 acgagtgttt ccatatccac ttctgacttc tttgagacca gcagaattca gatagaacca
 25441 acatcctccc tgacttctgg attgagggag accagcagct ctgagaggat cagctcagcc
 25501 acagagggaa gcaactgtct ttctgaagtg cccagtggtg ctaccactga ggtctccagg
 25561 acagaagtga tatcctctag ggaacatcc atgtcagggc ctgatcagtt caccatatca
 25621 ccagacatct ctactgaagc gatcaccagg ctttctactt cccccattat gcagaaatca
 25681 gcagaaagtg ccatcactat tgagacaggt tctctggggg ctacatcaga ggtaccctc
 25741 accttggaca cctcaacaac aaccttttgg tcagggacc actcaactgc atctccagga
 25801 tttcacact cagagatgac cactcttatg agtagaactc ctggagatgt gccatggccg
 25861 agccttccct ctgtggaaga agccagctct gtctcttctc cactgtcttc acctgccatg
 25921 acctcaactt cttttttctc cgcattacca gagagcatct cctcctctcc tcatcctgtg
 25981 actgcacttc tcacccttgg cccagtgaag accacagaca tgttggccac aagctcagaa

Figure 3A (continued)

26041 cctgaaacca gttcacctcc aaatttgagc agcacctcag ctgaaatatt agccacgtct
 26101 gaagtcacca aagatagaga gaaaattcat ccctcctcaa acacacctgt agtcaatgta
 26161 gggactgtga tttataaaca tctatcccct tcctctgttt tggctgactt agtgacaaca
 26221 aaaccacat ctccaatggc taccacctcc actctgggga atacaagtgt ttccacatca
 26281 actcctgcct tcccagaaac tatgatgaca cagccaactt cctccctgac ttctggatta
 26341 agggagatca gtacctctca agagaccagc tcagcaacag agagaagtgc ttctctttct
 26401 ggaatgccca ctgggtgctac tactaaggct tccagaacag aagccctctc cttaggcaga
 26461 acatccaccc caggctcctgc tcaatccaca atatcaccag aaatctccac ggaaaccatc
 26521 actagaatth ctactcccct caccacgaca ggatcagcag aaatgaccat ccccccaaa
 26581 acaggtcatt ctggggcctc ctcaaacagt acctttacct tggacacatc aagcagagcc
 26641 tcctggccag gaactcactc agctgcaact cacagatctc cacactcagg gatgaccact
 26701 cctatgagca gaggtcctga ggatgtgtca tggccaagcc gcccatcagt ggaaaaaact
 26761 agccctccat cttccctggg gtctttatct gcagtaacct caccttcgcc actttattcc
 26821 acaccatctg agagtagcca ctcatctcct ctcgggtga ctcctctttt cttctctttc
 26881 atgatgaaga ccacagacat gttggacaca agcttggaac ctgtgaccac ttcacctccc
 26941 agtatgaata tcacctcaga tgagagtctg gccacttcta aagccaccat ggagacagag
 27001 gcaattcagc tttcagaaaa cacagctgtg actcagatgg gcaccatcag cgctagacaa
 27061 gaattctatt cctcttatcc aggcctccca gagccatcca aagtgcacatc tccagtggtc
 27121 acctcttcca ccataaaaga cattgtttct acaaccatac ctgcttctc tgagataaca
 27181 agaattgaga tggagtcac atcccctctg acccccacac caagggagac cactcactcc
 27241 caggagatcc actcagccac aaagccaagc actgttcctt acaaggcact cactagtgcc
 27301 acgattgagg actccatgac acaagtcatg tcctctagca gaggacctag cctgatcag
 27361 tccacaatgt cacaagacat atccagtcaa gtgatcacca ggctctctac ctccccatc
 27421 aaggcagaat ctacagaaat gaccattacc acccaaacag gttctcctgg ggctacatca
 27481 aggggtacct ttaccttggg cacttcaaca acttttatgt cagggacca ctcaactgca
 27541 tctcaaggtat tttcacactc acagatgacc gctcttatga gtagaactcc tggagatgtg
 27601 ccatggctaa gccatcccctc tgtggaagaa gccagctctg cctctttctc actgttctca
 27661 cctgtcatga cctcatcttc tcccgtttct tccacattac cagacagcat ccactcttct
 27721 tcgcttctctg tgacatcact tctcacctca gggctgggtga agaccacaga gctgttgggc
 27781 acaagctcag aacctgaaac cagttcaccc ccaaatttga gcagcacctc agctgaaata
 27841 ctggccacca ctgaagtcac tacagataca gagaaactgg agatgaccaa tgtggtaac
 27901 tcaggttata cacatgaatc tccttctctg gtcctagctg actcagtgc aacaaagcc
 27961 acatcttcaa tgggtatcac ctacccccaca ggagatacaa atgttctc atcaaccct
 28021 gccttctctg acaccagtag gattcaaaca aagtcaaagc tctcactgac tctgggtg
 28081 atggagacca gcatctctga agagaccagc tctgccacag aaaaaagcac tgtcctttct
 28141 agtggtgcca ctgggtgctac tactgaggctc tccaggacag aagccatctc ttctagcaga
 28201 acatccatcc caggccctgc tcaatccaca atgtcatcag acacctccat ggaaaccatc
 28261 actagaatth ctacccccct cacaaggaaa gaatcaacag acatggccat ccccccaaa
 28321 acaggtcctt ctggggctac ctgcgagggt acctttacct tggactcatc aagcacagcc
 28381 tctggccag gaactcactc agctacaact cagagatttc cacagtcagt ggtgacaact
 28441 cctatgagca gaggtcctga ggatgtgtca tggccaagcc cgctgtctgt ggaaaaaac
 28501 agccctccat cttccctggg atcttcatct tcagtaacct caccttcgcc actttattcc
 28561 acaccatctg ggagtagcca ctctctcct gtcctgtca cttctctttt cacctctatc
 28621 atgatgaagg ccacagacat gttggatgca agtttggaac ctgagaccac ttcagctccc
 28681 aatatgaata tcacctcaga tgagagtctg gccacttcta aagccaccac ggagacagag
 28741 gcaattcagc tttttgaaaa tacagcagcg tcccattgtg aaaccaccag tctacagag
 28801 gaactctatt cctcttcccc aggettctca gagccaacaa aagtgatac tccagtggtc
 28861 acctcttctc ctataagaga caacatgggt tccacaacaa tgctggctc ctctggcatt
 28921 acaaggattg agatagagtc aatgtcatct ctgaccctg gactgaggga gaccagaacc
 28981 tcccaggaca tcacctcatc cacagagaca agcactgtcc tttacaagat gtcctctggg
 29041 gccactcctg aggtctccag gacagaagtt atgcctcta gcagaacatc cattcctggc
 29101 cctgtctcagt ccacaatgtc actagacatc tccgatgaag ttgtcaccag gctgtctacc
 29161 tctcccatca tgacagaatc tgcagaaata accatcacca cccaacagc ttattctctg
 29221 gctacatccc aggttaccct tcccctgggc acctcaatga cctttttgtc agggaccac
 29281 tcaactatgt ctcaaggact ttcacactca gagatgacca atcttatgag caggggtcct
 29341 gaaagtctgt catggacgag ccctcgcttt gtggaaacaa ctagatcttc ctcttctctg
 29401 acatcattac ctctcagcag ctcaactttct cctgtgtcct ccacattact agacagtgc
 29461 ccctcctctc ctcttctctg gacttcaact atcctcccag gcctgggtgaa gactacagaa
 29521 gtgttgata caagctcaga gcctaaaacc agttcatctc caaatttgag cagcacctca
 29581 gttgaaatac cggccacctc tgaatcatg acagatacag agaaaattca ccttctctca
 29641 aacacagcgg tggccaaagt gaggacctcc agttctgttc atgaatctca ttctctgtc
 29701 ctagctgact cagaaacaac cataaccata cttcaatgg gtatcacctc cgctgtggac
 29761 gataccactg ttttcacatc aaatcctgcc ttctctgaga ctaggaggat tccgacagag
 29821 ccaacattct cattgactcc tggattcagg gagactagca cctctgaaga gaccacctca
 29881 atcacagaaa caagtgcagt cttttatgga gtgccacta gtgctactac tgaagtctcc

Figure 3A (continued)

29941 atgacagaaa tcatgtcctc taatagaaca cacatccctg actctgatca gtccacgatg
 30001 tctccagaca tcatcactga agtgatcacc aggctctctt cctcatccat gatgtcagaa
 30061 tcaacacaaa tgaccatcac cacccaaaaa agttctcctg gggctacagc acagagtact
 30121 cttaccttgg ccacaacaac agcccccttg gcaaggaccc actcaactgt tctcctaga
 30181 tttttacact cagagatgac aactcttatg agtaggagtc ctgaaaatcc atcatggaag
 30241 agctctccct ttgtggaaaa aactagctct tcatcttctc tgttgcctt acctgtcacg
 30301 acctcacctt ctgtttcttc cacattaccg cagagtatcc cttcctctc tttttctgtg
 30361 acttcaactcc tcaccccagg catgggtgaag actacagaca caagcacaga acctggaacc
 30421 agttttatctc caaatctgag tggcacctca gttgaaatac tggctgcctc tgaagtcacc
 30481 acagatacag agaaaattca tccttcttca agcatggcag tgaccaatgt gggaaccacc
 30541 agttctggac atgaaactata ttctctgttt tcaatccact cggagccatc caaggetaca
 30601 taccacagtgg gtactcctc ttccatggct gaaacctcta tttccacatc aatgctctgt
 30661 aattttgaga ccacagatt tgaggctgag ccattttctc atttgacttc tggatttagg
 30721 aagacaaaaca tgtccctgga caccagctca gtcacaccaa caaatacacc ttcttctcct
 30781 ggggtccaact accttttaca gagttccaag actgatttca cctcttctgc aaaaacatca
 30841 tccccagact ggctccagc ctcacagtat actgaaatc cagtggacat aatcaccccc
 30901 tttaatgctt ctccatctat tacggagtcc actgggataa cctccttccc agaatccagg
 30961 tttactatgt ctgtaacaga aagtaactat catctgagta cagatttgc gccttcagct
 31021 gagactatgt ccactggcac agtgatgctt tctctatcag aggccatgac ttccatttgc
 31081 accactggag ttcccagagc catctcaggt tcaggtagtc cattctctag gacagagtca
 31141 ggccctgggg atgctactct gtccaccatt gcagagagcc tgccttcac cactcctgtg
 31201 ccatttctct cttcaacct cactaccact gattcttcaa ccatcccagc cctccatgag
 31261 ataacttctt cttcagctac cccatataga gtggacacca gtcttgggac agagagcagc
 31321 actactgaag gacgcttggg tatggtcagt actttggaca cttcaagcca accaggcagg
 31381 acatcttcaa caccatttt ggataccaga atgacagaga gcgttgagct gggaacagtg
 31441 acaagtcttc atcaagttcc ttactctca acacggttga caagaactga tggcattatg
 31501 gaacacatca caaaaatacc caatgaagca gcacacagag gtaccataag accagtcaaa
 31561 ggccctcaga catccacttc gcctgccagt cctaaaggac tacacacagg agggacaaaa
 31621 agaatggaga ccaccaccac agctttgaa accaccacca cagctttgaa gaccacttcc
 31681 agagccacct tgaccaccag tgtctatact cccactttgg gaacactgac tcccctcaat
 31741 gcatcaaggc aaatggccag cacaatctc acagaaatga tgatcacac cccatagtgt
 31801 ttccctgatg ttccagaaac gacatctca ttggctacca gcctgggagc agaaacagc
 31861 acagctcttc ccaggacaac cccatctggt ctcaatagag aatcagagac cacagcctca
 31921 ctggctctctc gttctggggc agagagaagt ccggttatc aaactctaga tgtttcttct
 31981 agtgagccag atacaacagc ttcatgggtt atccatctg cagagaccat cccaactggt
 32041 tccaagacaa cccccaattt tttccacagt gaattagaca ctgtatcttc cacagccacc
 32101 agtcatgggg cagacgtcag ctcagccatt ccaacaaata tctcacctag tgaactagat
 32161 gcaactgacc cactggctac tatttcgggg acagatacta gtacaacatt cccaactg
 32221 tctcaagtccc cacatgaaac agagacaaga accacatggc tcaactatcc tgcagagacc
 32281 agctcaacta ttcccagaac aatccccaat ttttctctc atgaatcaga tggccacact
 32341 tcaatagcca ccagtcctgg ggcagaaacc agttcageta ttccaattat gactgtctca
 32401 cctgggtgcag aagatctggt gacctcacag gtcactagtt ctgggacaga cagaaatag
 32461 actattccaa ctttgactct ttctcctggt gaaccaaga cgatagcctc attagtcacc
 32521 catcctgaag cacagacaag ttgggccatt ccaacttcaa ctatctgcc tgctgtatca
 32581 cggttgggtg cctcaatggt caccagtttg gcggcaaga caagtacaac taactgagct
 32641 ctgacaaact cccctgggtg accagctaca acagtttcat tggctcacga tctctgacag
 32701 accagcccaa cagttccctg gacaacttcc atttttttcc atagtaaact agacaccaca
 32761 ccttcaatga ccaccagtca tggggcagaa tccagttcag ctggtccaac tccaactggt
 32821 tcaactgagg taccaggagt agtgaccctc ttggtcacca gttctagggc agtgatcagt
 32881 acaactattc caattctgac tctttctcct ggtgaaccag agaccacacc ttcaatggcc
 32941 accagtcag ggaagaagc cagttctgct attccaactc caactgttcc acctggggta
 33001 ccaggagtgg tgacctctct ggtcactagt tctagggcag tgactagtac aactattcca
 33061 attctgactt tttctcttgg tgaaccagag accacacctt caatggccac cagtcatggg
 33121 acagaagctg gctcagctgt tccaactggt ttacctgagg taccaggaat ggtgacctct
 33181 ctggttgcta gttctagggc agtaaccagt acaactcttc caactctgac tctttctcct
 33241 ggtgaaccag agaccacacc ttcaatggcc accagtcag gggcagaagc cagctcaact
 33301 gttccaactg tttcaactga ggtaccagga gtggtgacct ctctggctac tagttctagt
 33361 ggagtaaaca gtacaagtat tccaactctg attctttctc ctggtgact ctgaaaccaca
 33421 ccttcaatgg ccaccagtca tggggcagaa gccagctcag ctggtccaac tccaactggt
 33481 tcacctgggg tatcaggagt ggtgaccctc ctggtcacta gttccagggc agtgaccagt
 33541 acaactattc caattctaac tctttcttct agtgagccag agaccacacc ttcaatggcc
 33601 accagtcag ggtagaagc cagctcagct gttctaactg tttcaactga ggtaccagga
 33661 atgggtgacct ctctggctac tagttctaga gcagtaacca gtacaactat tccaactg
 33721 actatttctt ctgatgaacc agagaccaca acttctattg tcaaccattc tgaggcaagc
 33781 atgatttctcag ccattccaac ttagctgtc tcccactg tacaagggtc ggtgacttca
 33841 ctggctcacta gttctgggtc agagaccagt gcgttttcaa atctaactgt tgccctcaagt

Figure 3A (continued)

33901 caaccagaga ccatagactc atgggtcgc tcatcctgga cagaagcaag ttctgttgtt
 33961 ccaactttga ctgtctccac tgggtgagccg tttacaaata tctcattggt caccatcct
 34021 gcagagagta gctcaactct tcccaggaca acctcaaggt tttcccacag tgaattagac
 34081 actatgcctt ctacagtcac cagtcctgag gcagaatcca gctcagccat ttcaactact
 34141 atttcacctg gtataccagg tgtgctgaca tcaactggtca ctactctctg gagagacatc
 34201 agtgaacctt ttccaacagt gcctgagctc ccacatgaat cagaggcaac agcctcatgg
 34261 gttactcatc ctgcagtcac cagcacaaca gttcccagga caaccctaa ttatttctat
 34321 agtgaaccag acaccacacc atcaatagcc accagtctctg gggcagaagc cacttcagat
 34381 tttccaacaa taactgtctc acctgatgta ccagatatgg taacctcaca ggtcactagt
 34441 tctgggacag acaccagtat aactattcca actctgactc tttcttctgg tgagccagag
 34501 accacaacct catttatcac ctattctgag acacacacaa gttcagccat tccaactctc
 34561 cctgtctccc ctgggtgcatc aaagatgctg acctcactgg tcatcagttc tgggacagac
 34621 agcactagaa ctttcccac actgacggag accccatag aaccagagac aacagccata
 34681 cagctcattc atcctgcaga gaccaacaca atggttccca agacaactcc caagttttcc
 34741 catagtaagt cagacaccac actcccagta gccatcacca gtccctgggccc agaagccagt
 34801 tcagctgttt caacgacaac tatctcacct gatatgtcag atctgggtgac ctactggtc
 34861 cctagttctg ggacagacac cagtacaacc ttcccacat tgagtgagac cccatagtaa
 34921 ccagagacta cagtcacgtg gctcactcat cctgcagaaa ccagcacaac ggtttctggg
 34981 acaattccca acttttccca taggggtca gacactgcac cctcaatggt caccagtcct
 35041 ggagtagaca cgaggtcagg tgttccaact acaaccatcc caccagtat accaggggta
 35101 gtgacctcac aggtcactag ttctgcaaca gacactagta cagctattcc aactttgact
 35161 ccttctctct gtgaaccaga gaccacagcc tcatcagcta cccatcctgg gacacagact
 35221 ggcttctactg ttccaattcg gactgttccc tctagtgagc cagatacaat ggcttctctg
 35281 gtcactcatc ctcccacagac cagcacacct gtttccagaa caacctccag tttttcccat
 35341 agtagtccag atgccacacc tgtaatggcc accagtccca ggacagaagc cagttcagct
 35401 gtactgacaa caatctcacc tgggtcacca gagatgggta cttcacagat cactagttct
 35461 ggggcagcaa ccagtacaac tgttccaact ttgactcatt ctctgggtat gccagagacc
 35521 acagccttat tgagcaccca tcccagaaca gggacaagta aaacatttcc tgcttcaact
 35581 gtgtttcctc aagtatcaga gaccacagcc tcaactacca ttagacctgg tgcagagact
 35641 agcacagctc tcccactca gacaacatcc tctctcttca cctacttgt aactggaacc
 35701 agcagagttg atctaagtcc aactgcttca cctgggtgtt ctgcaaaaac agccccactt
 35761 tccaccatca cagggacaga gaccagcaca atgattccaa ctccaactc tctccttggg
 35821 ttactagaga ctacaggctt actggccacc agctcttcag cagagaccag cacagtagt
 35881 ctaactctga ctgtttcccc tgctgtctct gggctttcca gtgcctctat aacaactgat
 35941 aagccccaaa ctgtgacctc ctggaacaca gaaacctcac catctgtaac ttcagttgga
 36001 ccccagaat tttccaggac tgtcacaggc accactatga ccttgatacc atcagagatg
 36061 ccaacaccac ctaaaaccag tcatggagaa ggagtgagtc caaccactat cttgagaact
 36121 acaatgggtg aagccactaa tttagctacc acaggttcca gtcccactgt ggccaagaca
 36181 acaaccacct tcaatacact ggctggaagc ctctttactc ctctgaccac acctgggatg
 36241 tccacttgg cctctgagag tgtgacctca agaacaagtt ataaccatcg gtccctggatc
 36301 tccaccacca gcagttataa ccgtcggtag tggacctctg ccaccagcac tccagtgact
 36361 tctacattct ccccagggat ttccacatcc tccatcccca gctccacagc agccacagtc
 36421 ccaatcattg tgccattcac cctcaacttc accatcacca acctgcagta cgaggaggac
 36481 atgcggcacc ctggttccag gaagtccaac gccacagaga gagaactgca ggtctgctc
 36541 aaaccttgt tcaggaatag cagtctgaa tacctctatt caggtctgac actagctca
 36601 ctcagggcag agaaggatag ctcagccatg gcagtggatg ccatctgcac caatcgccct
 36661 gaccctgaag acctcggact ggacagagag cgactgtact gggagctgag caatctgaca
 36721 aatggcatcc aggagctggg cccctacacc ctggaccgga acagtctcta tgtcaatggt
 36781 ttcaccatc gaagctctat gccaccacc agcactcctg ggacctccac agtggatgtg
 36841 ggaacctcag ggactccatc ctccagcccc agccccacgg ctgctggccc tctcctgatg
 36901 ccgttcaacc tcaacttcac catcaccaac ctgcagtagc agggaggacat gcgtcgcact
 36961 ggctccagga agttcaacac catggagagt gtctgctcaa gtctgctcaa gccctgttcc
 37021 aagaacacca gtgttggccc tctgactct ggtgcagat tgaccttgt caggcccag
 37081 aaagatgggg cagccactgg agtggatgcc atctgcacc accgccttga ccccaaaagc
 37141 cctggactca acagggagca gctgtactgg gagctaagca aactgaccaa tgacattgaa
 37201 gagctgggccc cctacaccct ggacaggaac agtctctatg tcaatgggtt caccatcag
 37261 agctctgtgt ccaccaccag cactcctggg acctccacag tggatctcag aaacctcagg
 37321 actccatcct cctctccag ccccacaatt atggctgctg gccctctct ggtaccattc
 37381 accctcaact tcaccatcac caacctgcag tatggggagg acatgggtca cctggctcc
 37441 aggaagtcca acaccacaga gagggctctg cagggctctgc ttgggtccat attcaagaac
 37501 accagtgttg gccctctgta ctctggtgac agactgacct ctctcaggtc tgagaaggat
 37561 ggagcagcca ctggagtgga tgccatctgc atccatcctc ttgaccccaa aagccctgga
 37621 ctcaacagag agcggctgta ctgggagctg agccaactga ccaatggcat caaagagctg
 37681 ggcccctaca ccttgacag gaacagtctc tatgtcaatg gtttcacca tcggacctct
 37741 gtgcccacca ccagcactcc tgggacctcc acagtggacc ttggaacctc agggactcca
 37801 ttctccctcc caagccccgc aactgctggc cctctcctgg tgctgttcc cctcaacttc

Figure 3A (continued)

37861 accatcacca acctgaagta tgaggaggac atgcatcgcc ctggctccag gaagttcaac
 37921 accactgaga gggtcctgca gactctgctt ggtcctatgt tcaagaacac cagtgttggc
 37981 cttctgtact ctggctgcag actgaccttg ctcaggtcog agaaggatgg agcagccact
 38041 ggagtgatg ccatctgcac ccaccgtctt gacccccaaa gccctggact ggacagagag
 38101 cagctatact gggagctgag ccagctgacc aatggcatca aagagctggg cccctacacc
 38161 ctggacagga acagtctcta tgtcaatggt ttcaccattt ggatccctgt gccaccagc
 38221 agcactcctg ggacctccac agtggacctt gggtcaggga ctccatcctc cctccccagc
 38281 cccacagctg ctggccctct cctggtgcca ttcaccctca acttcaccat caccaacctg
 38341 cagtacgagg aggacatgca taccaccagg tccaggaagt tcaacaccac ggagcgggtc
 38401 ctgcaggggtc tgcttgggtcc catgttcaag aacaccagtg tcggccttct gtactctggc
 38461 tgcagactga ccttgctcag gtccgagaag gatggagcag ccactggagt ggatgccatc
 38521 tgcacccacc gtcttgacc caaaagccct ggagtgagca gggagcagct atactgggag
 38581 ctgagccagc tgaccaatgg catcaaagag ctgggtccct acaccctgga cagaacagc
 38641 ctctatgtca atggtttcc ccatgagacc tctgcgcca acaccagca tcctgggacc
 38701 tccacagtgg accttgggac ctcagggact ccactctccc tccccagccc tacatcngct
 38761 ggccctctcc tggtnccntt caccctcaac ttcaccatca ccaacctgca gtacgaggag
 38821 gacatgcggc acccnggntc caggaagtcc aacaccacng agagggtnct gcagggctctg
 38881 ctnaagcccc tnttcaagag caccagtgtt ggccctctgt actctggctg cagactgacc
 38941 ttgtccaggt ccgagaagga tggagcagcc actggagtgg atgccatctg caccaccgt
 39001 cttgacccca aaagccctgg agtggacagg gagcagctat actgggagct gagccagctg
 39061 accaatggca tcaaagagct gggtccttac accctggaca gaaacagtct ctatgtcaat
 39121 ggtttcacc atcagacctc tgcgccaac accagcactc ctgggacctc cacagtggac
 39181 ctggggacct cagggactcc atcctccctc ccagcccta catctgctgg ccctctcctg
 39241 gtgccattca cctcaactt caccatcacc aacctgcagt acgaggagga catgcatcac
 39301 ccaggctcca ggaagttcaa caccacggag cgggtcctgc agggctctgt tgggtccatg
 39361 ttcaagaaca ccagtgtcgg ccttctgtac tctggctgca gactgacctt gctcaggcct
 39421 gagaagaatg gggcagccac tggaatggat gccatctgca gccaccgtct tgaccccaaa
 39481 agccctggac tcaacagaga gcagctgtac tgggagctga gccagctgac ccatggcatc
 39541 aaagagctgg gccctacac cctggacagg aacagtctct atgtcaatgg ttccaccat
 39601 cggagctctg tggccccac cagcactcct gggacctcca cagtggacct tgggacctca
 39661 gggactccat cctccctccc cagccccaca acagctgttc ctctcctggt gccgttacc
 39721 ctcaacttta ccatcaccaa tctgcagat ggggaggaca tgcgtcacc tgggtccagg
 39781 aagttcaaca ccacagagag ggtcctgcag ggtctgcttg gtccctggt caagaccctc
 39841 agtgtcggcc ctctgtactc tggctgcaga ctgatctctc tcaggctctga gaaggatggg
 39901 gcagccactg gagtggatgc catctgcacc caccacctta accctcaaag ccctggactg
 39961 gacagggagc agctgtactg gcagctgagc cagatgacca atggcatcaa agagctgggc
 40021 ccctacacc tggaccggaa cagtctctac gtcaatgggt tcaaccatcg gagctctggg
 40081 ctaccacca gactccttg gacttcaca gttgacctg gaacctcagg gactccatcc
 40141 cccgtcccca gccccacaac tgctggcctt ctctgggtgc catcaccct caacttcacc
 40201 atcaccacc tgcagtatga ggagcactg catcgccctg gatctaggaa gttcaacacc
 40261 acagagaggg tctgcaggg tctgcttagt cccattttca agaactccag tgttggccct
 40321 ctgtactctg getgcagact gacctctctc aggcccgaga aggatggggc agcaactgga
 40381 atggatgctg tctgcctcta ccaccctaat cccaaaagac ctggactgga cagagagcag
 40441 ctgtactggg agctaagcca gctgaccac aacatcactg agctgggccc ctacagcctg
 40501 gacagggaca gtctctatgt caatggttcc acccatcaga actctgtgcc caccaccagt
 40561 actcctggga cctccacagt ctactgggca accactggga ctccatctc ttccccggc
 40621 cacacagagc ctggccctct cctgatacca ttcactttca actttaccat caccaacctg
 40681 cattatgagg aaaacatgca acaccctggt tccaggaagt tcaacaccac ggagagggtt
 40741 ctgcaggggtc tgctcaagcc cttgttcaag aacaccagtg ttggccctct gtactctggc
 40801 tgcagactga cctctctcag gccgagaag gatggggcag caactggaat ggatgctgtc
 40861 tgcctctacc accctaatac caaaagacct gggctggaca gagagcagct gtactgggag
 40921 ctaagccagc tgaccacaa catcactgag ctgggcccct acagcctgga caggacagt
 40981 ctctatgtca atggtttcc ccatcagaac tctgtgcccc ccaccagtac tcctgggacc
 41041 tccacagtgt actgggcaac cactgggact ccactcctct tccccggcca cacagagcct
 41101 ggccctctcc tgataaccatt cactttcaac tttaccatca ccaacctgca ttatgaggaa
 41161 aacatgcaac accctggttc caggaagtcc aacaccacgg agagggttct gcagggctctg
 41221 ctcaagccct tgttcaagaa caccagtgtt ggccctctgt actctggctg cagactgacc
 41281 ttgctcagac ctgagaagca tgaggcagcc actggagtgg acaccatctg taccaccgc
 41341 gttgatccca tcggacctgg actggacagg gagcgctat actgggagct gagccagctg
 41401 accaacagca ttaccgaact gggaccctac accctggaca gggacagtct ctatgtcaat
 41461 ggcttcaacc ctgggagctc tgtgccaacc accagcactc ctgggacctc cacagtgcac
 41521 ctggcaacct ctgggactcc atcctccctg cctggccaca cagcccctgt ccctctctg
 41581 ataccattca ccctcaactt taccatcacc aacctgcatt atgagaaaaa catgcaacac
 41641 cctggttcca ggaagttcaa ccccaggag agggttctgc agggctctgt caagccctg
 41701 ttcaagaaca ccagtgttgg ccctctgtac tctggctgca gactgacctt gctcagacct
 41761 gagaagcatg aggcagccac tggagtggac accatctgta cccaccgcgt tgatcccatc

Figure 3A (continued)

41821 ggacctggac tgnacagnga gcnqctntac tgggagctna gccanctgac caannncatc
 41881 nnnagactgg gnccctacac cctggacagg nacagtctct atgtcaatgg tttcacccat
 41941 cnganctctg ngcccaccac cagcactcct gggacctcca cagtgnacnt nngnacctcn
 42001 gggactccat cctccntccc cngccncaca tctgctggcc ctctcctggt gccattcacc
 42061 ctcaacttca ccatcaccaa cctgcagtac gaggaggaca tgcatacccc aggctccagg
 42121 aagttcaaca ccacggagcg ggtcctgcag ggtctgcttg gtcccatggt caagaacacc
 42181 agtgatcgcc ttctgtactc tggctgcaga ctgaccttgc tcaggcctga gaagaatggg
 42241 gcagccactg gaatggatgc catctgcagc caccgtcttg accccaaaag ccctggactc
 42301 gacagagagc agctgtactg ggagctgagc cagctgacct atggcatcaa agagctgggc
 42361 ccctacaccc tggacaggaa cagtctctat gtcaatggtt tcaccatcg gagctctgtg
 42421 gccccacca gcactcctgg gacctccaca gtggaccttg ggacctcagg gactccatcc
 42481 tccctcccca gcccacaac agctgttccct ctctggtgce cgttcaccct caacttacc
 42541 atcaccaatg tgcagtatgg ggaggacatg cgtcaccctg gctccaggaa gttcaacacc
 42601 acagatgagc tctgtcaggg tctgctatgt cccttgttca agaactccag tgtcggcct
 42661 ctgtactctg gctgcagact gatctctctc aggtctgaga aggatggggc agccactgga
 42721 gtggatgcca tctgcacca ccaccttaac cctcaaagcc ctggactgga cagggagcag
 42781 ctgtactggc agctgagcca gatgaccaat ggcatcaaag agctgggccc ctacacctg
 42841 gaccggaaca gtctctactg caatggtttc acccatcgga gctctgggct caccaccagc
 42901 actccttggg cttccacagt tgaccttggg acctcaggga ctccatcccc cgtccccagc
 42961 cccacaatgc ctggccctct cctggctcca ttcaccctaa acttcaccat caccaactg
 43021 cagtatgagg agggactgca tgcacctgga tctaggaagt tcaacgccac agagagggtc
 43081 ctgcagggtc tgcttagtcc catattcaag aactccagtg ttggccctct gtactctggc
 43141 tgcagactga cctctctcag gcccgagaag gatggggcag caactggaat ggatgctgtc
 43201 tgcctctacc accctaatcc caaaagacct ggactggaca gagagcagct gtactgggag
 43261 ctaagccagc tgaccacaa catcactgag ctgggcccct acagcctgga cagggacagt
 43321 ctctatgtca atggtttcac ccatcagagc tctatgacga ccaccagaac tctgatacc
 43381 tccacaatgc acctggcaac ctcgagaact ccagcctccc tgtctggacc tacagccgc
 43441 agccctctcc tgggtctatt cacaatcaac tgcaccatca ccaacctgca gtacgaggag
 43501 gacatgctgc gcaactggctc caggaagtcc aacaccatgg agagtgtcct gcagggtctg
 43561 ctcaagccct tgttcaagaa caccagtgtt ggccctctgt actctggctg cagattgacc
 43621 ttgctcagcc ccaagaaaga tggggcagcc actggagtgg atgccatctg caccaccgc
 43681 cttgacccca aaagccctgg actcaacagg gaggcagctg actgggagct aagcaactg
 43741 accaatgaca ttgaagagct gggccctac accctggaca ggaacagtct ctatgcaat
 43801 ggtttcacc atcagagctc tgtgtccacc accagcactc ctgggacctc cacagtggat
 43861 ctcagaacct cagggactcc atcctccctc tccagcccca caattatgnc nctgncct
 43921 ctctgntnc cnttcacnt caactnacc atcaccacc tgcantangn ggannacatg
 43981 cnnccccng gntccaggaa gttcaacacc acngagaggg tctacaggg tctgctcagg
 44041 cccttgttca agaacaccag tgtcagctct ctgtactctg gttgcagact gaccttctc
 44101 aggcctgaga aggatggggc agccaccaga gtggatgctg cctgcacctg ccgcccctgat
 44161 cccaaaagcc ctggactgga cagagagcaa ctatactggg agctgagcca gtaaccacc
 44221 agcatcactg agctgggacc ctacacctg gacagggtca gtctctatgt caatggcttc
 44281 aacctctgga gctctgtgcc aaccaccagc actcctggga cctccacagt gcacctggca
 44341 acctctggga ctccatcctc cctgctggc cacacancn ctgnccctct cctgntnccn
 44401 ttcacntca acttnaccat caccaacctg cantangngg annacatgcn nccccnggn
 44461 tccaggaagt tcaacaccac ngagagggtt ctgcagggtc tgcctaaacc cttgttcagg
 44521 aatagcagtc tgaatacct ctattcagc tgcagactag cctcactcag gccagagaag
 44581 gatagctcag ccattggcagc ggatgccatc tgcacacatc gccctgacc tgaagacctc
 44641 ggactggaca gagagcagct gtactgggag ctgagcaatc tgacaaatgg catccaggag
 44701 ctgggcccct acacctgga ccggaacagt ctctacgtca atggtttcac ccatcggagc
 44761 tctgggctca ccaccagcac tcttggact tccacagttg accttggaa ctcagggact
 44821 ccatcccccg tccccagccc cacaactgct ggccctctcc tggtgccatt caccctcaac
 44881 ttcaccatca ccaacctgca gtatgaggag gacatgcatc gccctggttc caggagggtc
 44941 aacaccacgg agagggttct gcagggctg ctacgccct tgttcaagaa caccagtgtt
 45001 ggccctctgt actctggctg cagactgacc ttgctcagac ctgagaagca agaggcagcc
 45061 actggagtgg acaccatctg taccaccgc gttgatccca tggacctgg actggacaga
 45121 gagcggctat actgggagct gagccagctg accaacagca tcacagagct gggaccctac
 45181 acctggata gggacagtct ctatgtcaat ggcttcaacc ctggagctc tgtgccaacc
 45241 accagcactc ctgggacctc cacagtgcac ctggcaacct ctgggactcc atcctcctg
 45301 cctggccaca cagcccctgt ccctctcttg ataccattca cctcaactt taccatcacc
 45361 gacctgcatt atgaagaaaa catgcaacac cctggttcca ggaagttcaa caccacggag
 45421 agggttctgc aggtctgtct caagcccttg ttcaagagca ccagcgttgg ccctctgtac
 45481 tctggctgca gactgacctt gctcagacct gagaacatg gggcagccac tggagtggac
 45541 gccatctgca ccctccgctc tgatcccact ggtcctggac tggacagaga gcggtatac
 45601 tgggagctga gccagctgac caacagcgtt acagagctgg gccctacac cctggacagg

Figure 3A (continued)

45661 gacagtctct atgtcaatgg cttcaccat cggagctctg tgccaaccac cagtattcct
 45721 gggacctctg cagtgcacct ggaacctct gggactccag cctccctccc tggccacaca
 45781 gccctggcc ctctcctggt gccattcacc ctcaacttca ctatcaccaa cctgcagtat
 45841 gaggaggaca tgcgtcacc tggttocagg aagtccagca ccacggagag agtcctgcag
 45901 ggtctgctca agcccttgtt caagaacacc agtgtcagct ctctgtactc tggttgcaga
 45961 ctgaccttgc tcaggcctga gaaggatggg gcagccacca gagtggatgc tgtctgcacc
 46021 catcgtctctg accccaaaag ccctggactg gacagagagc ggctgtactg gaagctgagc
 46081 cagctgacct acggcatcac tgagctgggc ccctacacc tggacaggca cagtctctat
 46141 gtcaatgggt tcacccatca gagctctatg acgaccacca gaactcctga tacctccaca
 46201 atgcacctgg caacctcgag aactccagcc tcctgtctg gacctacgac cgccagccct
 46261 ctctggtgac tattcacaat taacttcacc atcactaacc tgcggtatga ggagaacatg
 46321 catcaccttg gctctagaaa gttaaacacc acggagagag tccttcaggg tctgctcagg
 46381 cctagtgtca agaaccacc tgttggccct ctgtactctg gctgcagact gaccacgctc
 46441 aggcccaaga aggatggggc agccacaaa gtggatgcca tctgcaccta ccgcccgat
 46501 cccaaaagcc ctggactgga cagagagcag ctatactggg agctgagcca gctaaccac
 46561 agcatcactg agctgggccc ctacaccag gacagggaca gtctctatgt caatggcttc
 46621 acccatcgga gctctgtgcc aaccaaccagtt atcctggga cctctgcagt gcacctggaa
 46681 acctctggga ctccagcctc cctccctggc cacacagccc ctggccctct cctggtgcca
 46741 ttcacctca acttcaact caaccaactg cagtatgagg aggacatgc tcacctgggt
 46801 tccaggaagt tcaaccacc ggagagagtc ctgcagggtc tgctcaagcc cttgttcaag
 46861 agcaccagtg ttggccctct gtactctggc tgcagactga ccttgctcag gcctgaaaaa
 46921 cgtggggcag ccaccggcgt ggacaccatc tgcactcacc gcctgacctc tctaaacca
 46981 ggactggaca gagagcagct atactgggag ctgagcaaac tgaccctgg catcatcgag
 47041 ctgggcccct acctcctgga cagaggcagtt ctctatgtca atggtttcc ccatcgacc
 47101 tctgtgcccc ccaccagcac tcctgggacc tccacagtgg accttggaa ctcagggaact
 47161 ccattctccc tccaagccc gcancnntc gnccctctcc tgnntcncac caccntcaac
 47221 ttnaccatca ccaacctgca ntangnggan nacatgcnnc ncccngntc caggaagttc
 47281 aacaccacng agagggtcct gcagactctg cttggtccta tgttcaagaa caccagtgtt
 47341 ggcttctgt actctggctg cagactgacc ttgctcaggt ccgagaagga tggagcagcc
 47401 actggagtg atgccatctg caccaccctg ctgacccca aaagccctgg agtggacagg
 47461 gagcaactat actgggagct gagccagctg accaatggca ttaaagaact gggcccctac
 47521 acctggaca ggaacagtct ctatgtcaat gggttcacc attggatccc tgtgcccacc
 47581 agcagcactc ctgggacctc cacagtggac ctggggtcag ggactccatc ctcccctccc
 47641 agccccacaa ctgctggccc tctcctgggt ccgttcacc tcaacttcac catcaccaac
 47701 ctgaagtacg agggagcat gcattgccct ggctccagga agttcaacac cacagagaga
 47761 gtctgcagag gtctgcttgg tccatgttc aagaacacca gtgttggccc tctgtactct
 47821 ggctgcagac tgaccttgc caggctccag aaggatggag cagccactgg agtggatgcc
 47881 atctgcacc acctcttga ccccaaaagc cctggagtgg acagggagca gctatactgg
 47941 gagctgagcc agctgaccaa tggcatcaa gagctgggtc cctacacct ggacagaac
 48001 agtctctatg tcaatggttt caccatcag acctctgccc ccaaccagg cactcctggg
 48061 acctccacag tggaccttgg gacctcaggg actccatcct cctcccag ccctacancn
 48121 nctgnccctc tctgntncc nttaccntc aacttnacca tcaaacct gcantangng
 48181 gannacatgc nncnccng ntccaggaag tcaaacacca cngagngngt nctgcagggt
 48241 ctgctnnnc cctnttcaa gaacnccagt gtnggcnctc tgtactctgg ctgcagactg
 48301 acctnctca ggnngagaa gnatggngca gccactggan tggatgcat ctgcanccac
 48361 cncnctnanc caaaagnc tggactgnac agngagcngc tntactggga gctnagccan
 48421 ctgaccaann ncatcnnga gctgggnccc tacaccctgg acaggnacag tctctatgtc
 48481 aatggtttca cccattggat cctgtgccc accagcagca ctctgggac ctccacagtg
 48541 gaccttgggt cagggactcc atcctcctc cccagccca caactgctgg cctctcctg
 48601 gtgccgttca cctcaactt caccatcacc aacctgaagt acgaggagga catgcattgc
 48661 cctggctcca ggaagttcaa caccacagag agagtctctg agagtctgct tggteccatg
 48721 ttcaagaaca ccagtgttgg ccctctgtac tctggctgca gactgacctc gctcaggtcc
 48781 gagaaggatg gagcagccac tggagtgat gccatctgca cccaccgtgt tgaccctaaa
 48841 agccctggag tggacaggga gcagctatac tgggagctga gccagctgac caatggcatc
 48901 aaagagctgg gtccctacac cctggacaga aacagtctct atgtcaatgg tttcaccat
 48961 cagacctctg cgcccaacac cagcactcct gggacctcca cagtgnactn nggnacctcn
 49021 gggactccat cctccntccc cngccncaca tctgctggcc ctctcctggt gccattcacc
 49081 ctcaacttca ccatcaccaa cctgcagtag gaggaggaca tgcataccc aggtccagg
 49141 aagttcaaca ccacggagcg ggtcctgag ggtctgctg tccccatgtt caagaacacc
 49201 agtgtcgccc ttctgtactc tggctgcaga ctgaccttgc tcaggcctga gaagaatggg
 49261 gcaaccactg gaatggatgc catctgcacc caccgtcttg accccaaaag cctggactg
 49321 nacangagc ngctntactg ggagctnagc canctgacca annnctcnn ngagctgggn
 49381 cctacacccc tggacaggna cagtctctat gtcaatgggt tcaaccatcn ganctctgng
 49441 cccaccacca gcaactcctg gacctccaca gtgnactnng gnacctcng gactccatcc
 49501 tcntccccn nctgnccct cctcctnctc gntccagnt cnttaccnt caactnacc
 49561 atcaccaacc tgcantangn ggannacatg cncnccng gntccaggaa gttcaacc

Figure 3A (continued)

49621 acngagaggg ttctgcaggg tctgctcaaa cccttgttca ggaatagcag tctggaatac
 49681 ctctattcag gctgcagact agcctcactc aggccagaga aggatagctc agccatggca
 49741 gtggatgcc tctgcacaca tgcacctgac cctgaagacc tggactgga cagagagcga
 49801 ctgtactggg agctgagcaa tctgacaaat ggcattccagg agctgggccc ctacacctg
 49861 gaccggaaca gtctctatgt caatggtttc acctcagaa gctctatgcc caccaccagc
 49921 actcctggga cctccacagt ggatgtggga acctcagga ctccatcctc cagcccagc
 49981 cccacgactg ctggccctct cctgatacca ttcaccctca acttcaccat caccaacctg
 50041 cagtatgggg aggacatggg tcaccctggc tccaggaagt tcaacaccac agagagggtc
 50101 ctgcagggtc tgcttggtcc catattcaag aacaccagtg ttggccctct gtactctggc
 50161 tgcagactga cctctctcag gtctgagaag gatggagcag ccactggagt ggatgccatc
 50221 tgcattccatc atcttgacc caaaagccct ggactcaaca gagagcggct gtactgggag
 50281 ctgagccaac tgaccaatgg catcaaagag ctgggcccct acaccctgga caggaacagt
 50341 ctctaatnca atggtttcac ccatcggacc tctgtgcca ccaccagcac tctgggacc
 50401 tccacagtgg ccttggaac ctcagggact ccattctccc tcccaagccc cgcaactgct
 50461 ggccctctcc tgggtgctgt caccctcaac ttcaccatca ccaacctgaa gtatgaggag
 50521 gacatgcatc gccctggctc caggaagttc aacaccactg agagggctct gcagactctg
 50581 cttggctccta tgttcaagaa caccagtgtt ggccttctgt actctggctg cagactgacc
 50641 ttgctcaggt ccgagaagga tggagcagcc actggagtgg atgccatctg caccaccctg
 50701 cttgaccca aaagccctgg actgnacagn gagcngctnt actgggagt nagcctctg
 50761 accaannc a tcnngagct gggncctac accctggaca ggnacagtct ctatgtcaat
 50821 ggtttcacc atcnngactc tngngcccacc accagcactc ctgggacctc cacagtgnac
 50881 ntnggnacct cngggactcc atcctcctc cccngccnca cancnctgn cctctctctg
 50941 ntnccttca contcaact naccatcacc aacctgcant angngganna catgcnncnc
 51001 ccnggntcca ggaagttcaa caccacngag agagtccttc agggctctgct caggcctgtg
 51061 ttcaagaaca ccagtgttgg cctctgtac tctggctgca gactgacct gctcagccc
 51121 aagaagatg gggcagccac caaagtggat gccatctgca cctaccgccc tgatcccaaa
 51181 agccctggac tggacagaga gcagctatac tgggagctga gccagctaac ccacagcatc
 51241 actgagctgg gccctacac ccaggacagg gacagtctct atgtcaatgg cttcaccat
 51301 cggagctctg tgccaaccac cagtattcct gggacctctg cagtgcacct ggaaaccact
 51361 gggactccat cctccttccc cggccacaca gagcctggcc ctctcctgat accattcact
 51421 ttcaacttta ccatcaccaa cctgcttcat gaggaaaaca tgcaacacc tggttccagg
 51481 aagttcaaca ccacggagag ggttctgcag ggtctgctca cgccttgtt caagaaccac
 51541 agtgttggcc ctctgtactc tggctgcaga ctgacctgca tcagacctg gaagcagag
 51601 gcagccactg gagtggacac catctgtacc caccgcttg atcccatcg acctggactg
 51661 gacagagagc ggctatactg ggagctgagc cagctgacca acagcatcac agagctggga
 51721 ccctacacc tggatagga cagtctctat gtcgatggct tcaaccctg gagctctgtg
 51781 ccaaccacca gcaactcctg gacctccaca gtgcacctgg caacctctg gactccatcc
 51841 cccctgctg gccacacagc cctgtcctc ctcttgatac cttaccct caactttacc
 51901 atcaccgacc tgcattatga agaaaacatg caacaccctg gttccaggaa gttcaaccac
 51961 acggagaggg ttctgcaggg tctgctcaag cccttgttca agagcaccag cgttggccct
 52021 ctgtactctg gctgcagact gacctgtctc agacctgaga aacatggggc agccactgga
 52081 gtggacgcca tctgcacctc cgccttgat cccactggct ctggactgga cagagagcgg
 52141 ctatactggg agctgagcca gctgaccaac agcatcacag agctgggacc ctacacctg
 52201 gatagggaca gtctctatgt caatggcttc aacccttggg gctctgtgcc aaccaccagc
 52261 actcctggga cctccacagt gcaactggca acctctggga ctccatcctc cctctctggc
 52321 cacacaactg ctggccctct cctggtgcg ttcacctca acttcaccat caccaacctg
 52381 aagtacgagg aggacatgca ttgccctggc tccaggaagt tcaacaccac agagagagtc
 52441 ctgcagagtc tgcattggtc catgttcaag aacaccagtg ttggccctct gtactctggc
 52501 tgcagactga ccttgctcag gtccgagaag gatggagcag ccactggagt ggatgccatc
 52561 tgcacccacc gtcttgacc caaaagccct ggactgnaca gngagcngct ntactgggag
 52621 ctnagccanc tgaccaannn catcnngag ctgggnccct acaccctgga caggnacagt
 52681 ctctatgtca atggtttcac ccatcnganc tctgngcca ccaccagcac tctgggacc
 52741 tccacagtgn acntngnac ctcngggact ccatectcn tcccngccn cacancnct
 52801 gnccctctcc tgnntcctt caccntcaac ttnaccatca ccaacctgca ntangnggan
 52861 nacatgcnnc nccnggntc caggaagttc aacaccacng agngngtnt gcagggctctg
 52921 ctnnnncccn tnttcaagaa cncagtgtn ggcctctgt actctggctg cagactgacc
 52981 tnnctcaggn cngagaagna tggngcagcc actggantgg atgccatctg canccaccn
 53041 cntnaccca aaagnccctg actgnacagn gagcngctnt actgggagt nagcctctg
 53101 accaacagca tcacagactc gggaccctac acctggata gggacagtct ctatgtcaat
 53161 ggtttcacc atcgaagctc tatgccacc accagtattc ctgggacctc tgcagtgcac
 53221 ctggaaacct ctgggactcc agcctcctc cctggccaca cagcccctgg cctctctctg
 53281 gtgceattca cctcaactt cactatcacc aacctgcagt atgaggagga catgctcacc
 53341 cctggttcca ggaagttcaa caccacggag agagtcctgc agggctctgct caagccctg
 53401 ttcaagagca ccagtgttgg ccctctgtac tctggctgca gactgacct gctcaggcct

Figure 3A (continued)

53461 gaaaaacgtg gggcagccac cggcgtggac accatctgca ctaccgcct tgaccctcta
53521 aacctggac tgnacagnga gcnctntac tgggagctna gccanctgac caannncatc
53581 nnnagctgg gncctacac cctggacagg nacagtctct atgtcaatgg ttaccaccat
53641 cnganctctg ngcccaccac cagcactcct gggacctcca cagtgnacnt nggnacctcn
53701 gggactccat cctccntccc cngccncaca ncnctgncc ctctcctgnt nccttcacc
53761 ntcaacttna ccatcaccaa cctgcantan gnggannaca tgcnncccc nggntccagg
53821 aagtccaaca ccacngagn ngtnctgcag ggtctgctnn nccccnttt caagaacncc
53881 agtgnngcc nctgtactc tggctgcaga ctgacctnnc tcaggncnga gaagnatggn
53941 gcagccactg gantggatgc catctgcanc caccnctn ancccaaaag ncctggactg
54001 nacagngagc ngctnactg ggagctnagc canctgacca annncatcnn ngagctgggn
54061 ccctacaccc tggacaggna cagtctctat gtcaatgggt ttaccctcg gagctctgtg
54121 ccaaccacca gactcctgg gacctccaca gtgacctgg caacctctgg gactccatcc
54181 tcctcgctg gccacacagc cctgtccct ctcttgatac cattaccct caactttacc
54241 atcacaacc tgcattatga agaaaacatg caacacctg gttccaggaa gttcaacacc
54301 acggagcggg tctgcaggg tctgcttggg cccatgttca agaacacaag tgtcggcctt
54361 ctgtactctg gctgcagact gacctgtctc aggctgaga agaatggggc agccactgga
54421 atggatgcca tctgcagcca cctcttgac cccaaaagcc ctggactgna cagnagcng
54481 ctntactggg agctnagcca nctgaccaan nncatcnnng agctgggnc ctacacctg
54541 gacaggnaca gtctctatgt caatggtttc acccatcnga nctctgngcc caccaccagc
54601 actcctggga cctccacagt gnaentngn acctcnggga ctccatctc cntccccngc
54661 cncacancnn ctgncctct cctgntncn ttaccntca acttnaccat caccaacctg
54721 cantangngg annacatgcn ncnccnggn tccaggaagt tcaacaccac ngagngngtn
54781 ctgcagggtc tgctnnncc cntntcaag aacnccagt tngccntct gtactctggc
54841 tgcaactgga cctnctcag gncngagaag natgngcag ccactggant gtagccatc
54901 tgcancacc ncnctnanc caaaagncct ggactgnaca gngagcngct ntactgggag
54961 ctngaccanc tgaccaann catcnnngag ctgggncct acacctgga cagnacagt
55021 ctctatgtca atggtttcac ccatcagaac tctgtgcccc ccaccagtac tctgggacc
55081 tccacagtgt actgggcaac cactgggact ccatcctct tccccggcca cacagagcct
55141 ggccctctcc tgataccatt cactttcaac ttaccatca ccaacctgca ttatgaggaa
55201 aacatgcaac accctggttc caggaagttc aacaccacgg agagggttct gcagggtctg
55261 ctacgcccct tgttcaagaa caccagtgtt ggccctctgt actctggctg cagactgacc
55321 ttgtcagac ctgagaagca ggaggcagc actggagtgg acaccatctg taccaccgc
55381 gttgatccca tcggacctgg actgnacagn gagcngctnt actgggagct nccanctg
55441 accaannca tcnngagct ggnccctac accctggaca ggnacagtct ctatgtcaat
55501 ggtttcacc atcnganctc tngcccacc accagcactc ctgggacctc cacagtgnac
55561 ntngnacct cngggactcc atcctcctc cccngccna cancnctgn cctctctctg
55621 ntncnttca cntcaactt naccatcacc aacctgcant annggganna catgcnnnc
55681 cnggntcca ggaagttcaa caccacngag nngntctgc aggtctgct nnnccccntn
55741 ttcaagaacn ccagtgtngg cntctgtac tctggctgca gactgacctn nctcagncn
55801 gtagagnatg gngcagccac tggantggat gccatctgca nccaccnncn tnanccaaa
55861 agnctggac tgnacagnga gcnctntac tgggagctna gccanctgac caannncatc
55921 nnnagctgg gncctacac cctggacagg nacagtctct atgtcaatgg ttaccaccat
55981 cggagctctg tgccaaccac cagcagtoct gggacctcca cagtgcacct ggaacctct
56041 gggactccat cctcctgccc tggccacaca gccctgtcc ctctcttgat accattcacc
56101 ctcaacttta ccatcaccaa cctgcattat gaagaaaaca tgcaacacc ttggtccagg
56161 aagtccaaca ccacggagag gttctgcag ggtctgctca agccctgtt caagacacc
56221 agtgttggcc ctctgtactc tggctgcaga ctgacctgac tcagacctga gaaactggg
56281 gcagccactg gactggagc catctgcacc ctccgcttg atcccactgg tctggactg
56341 nacagngagc ngctnactg ggagctnagc canctgacca annncatcnn ngagctgggn
56401 ccctacaccc tggacaggna cagtctctat gtcaatgggt tcaccatcnn ganctctgng
56461 cccaccacca gactcctgg gacctccaca gtgnaentng gnacctcngg gactccatcc
56521 tcntccccn gccncacanc nctgncct ctctgntnc cnttaccnt caactnacc
56581 atcacaacc tgcantagn gannacatg cncnccccng gntccaggaa gttcaacacc
56641 acngagngng tntgcaggg tctgctnnnn cctnttca agaacnccag tgtnggcent
56701 ctgtactctg gctgcagact gacctnctc agnncngaga agnatggngc agccactgga
56761 ntggatgcca tctgancca cncnctnan cccaaaagnc ctggactgna cagnagcng
56821 ctntactggg agctnagcca nctgaccaan nncatcnnng agctgggnc ctacacctg
56881 gacaggnaca gtctctatgt caatggtttc acccatcgga cctctgtgccc caccaccagc
56941 actcctggga cctccacagt gcaactggga acctctggga ctccatctc ctgctggc
57001 cacacagccc ctgtccctct ctgtatacca ttaccctca actttaccat caccaacctg
57061 cagtatgagg aggacatgca tcgccctgga tctaggaagt tcaacaccac agagagggtc
57121 ctgcagggtc tgcttagtcc catttcaag aactccagt tggccctct gtactctggc
57181 tgcaactgga cctctctcag gccgagaag gatggggcag caactggaat gtagctgtc
57241 tgctctacc accetaatcc caaaagacct gggctggaca gagagcagct gtactgcgag
57301 ctaagccagc tgaccacaa catcactgag ctgggcccct acagcctgga caggacagt
57361 ctctatgtca atggtttcac ccatcagaac tctgtgcccc ccaccagtac tctgggacc

Figure 3A (continued)

57421 tccacagtgt actgggcaac cactgggact ccatacctcct tccccggcca cacancnct
57481 gncctctcc tgntncntt cacntcaac ttnaccatca ccaacctgca ntangnggan
57541 nacatgcnnc nccnngntc caggaagttc aacaccacng agngngtntc gcagggctg
57601 ctnnnncccn tnttcaagaa cnccagtgtg ggcntctgt actctggctg cagactgacc
57661 tnnctcaggn cngagaagna tggngcagcc actggantgg atgccatctg canccaccnn
57721 cntnanccca aaagnctgg actgnacagn gagcngctnt actgggagct nagccantg
57781 accaannca tcnnngagct ggnccctac accctggaca ggnacagtct ctatgtcaat
57841 ggtttcacc attggagctc tgggctcacc accagcactc cttggacttc cacagttgac
57901 cttggaacct cagggactcc atccccgtc cccagccca caactgctgg cctctcctg
57961 gtgccattca ccctaaactt caccatcacc aacctgcagt atgaggagga catgcatcgc
58021 cctggatcta ggaagttaa cgccacagag agggctcctg agggctctgt tagtccata
58081 ttcaagaaca ccagtgttg ccctctgtac tctggctgca gactgacctt gctcagacct
58141 gagaagcagg aggcagccac tggagtggac accatctgta cccaccgct tgatccatc
58201 ggcactggac tgnacagnca tggctgtac tgggagctna gccantgac caannncaatc
58261 nngagctgg gnccctacac cctggacaggg nacagtctct atgtcaatgg tttcaccat
58321 cngantctg ngcccaccac cagcactcct gggacctcca cagtgnacnt nggnacctcn
58381 gggactccat cctcctccc cngccnaca ncnctgncc ctctcctgnt nccntcacc
58441 ntcaactna ccatacacia cctgcantan gnggannaca tgnncnccc nggntccagg
58501 aagttcaaca ccacngagng ngtntgcag ggtctgctnn nccntntt caagaacnc
58561 agtgtagggc ntctgtactc tggctgcaga ctgacctnnc tcaggnca gaagnatgg
58621 gcagccactg gantggatgc catctgcanc caccnctn anccaaaag nccctggactg
58681 nacagngagc ngctntactg ggagctnagc canctgacca annnctcnn ngagctgggn
58741 ccctacacc tggacagga cagtctctat gtcaatggtt tcacctatog gagcttggg
58801 ctcaccacca gcaactcctg gacttcaca gttgacctg gaacctcagg gactccatcc
58861 cccgtccca gcccacaac tgctggcct ctctgggtg cattcaccat aaactcacc
58921 atcacaacc tgcagatga ggagacatg catgcctctg gctccaggaa gttcaacc
58981 acggtagggc tcctcaggg tctgtctacg ccctgttca ggaaccacag tgtcagctct
59041 ctgtactctg gttgcagact gacctgtctc aggcctgaga aggatggggc agccaccaga
59101 gtggatgctg tctgcacca tctgctgac cccaaaagcc ctggactgna cagngagcng
59161 ctntactggg agctnagcca nctgaccaan nncatcnng agctgggncc ctacacctg
59221 gacaggnaca gtctctatgt caatggttcc accatcnga nctctngcc caccaccagc
59281 actcctggga cctccacagt gnacntngn acctcnggga ctccatctc cntccnngc
59341 cncacnncn ctgnccctc cctgntnccn ttcacctca acttnaccat caccaacctg
59401 cantangngg annacatgcn ncnccnngn tccaggaagt tcaaccac ngagngngtn
59461 ctgcagggtc tgctnnncc cntnttcaag aacnccagt tngcctct gtactctggc
59521 tgcagactga cctnctcag gncngagaag natgngcag ccaactggant ggatgccatc
59581 tgcancacc ncnctnanc caaaagnct ggactgnaca gngagcngct ntactgggag
59641 ctngaccanc tgaccaannn catcnngag ctgggnccct acacctgga caggnacagt
59701 ctctatgtca atggttccac ccattggatc cctgtgcca ccagcagcac tctgggacc
59761 tctcagctgg acctgggtc agggactcca tcctcctcc ccagccccc aactgtggc
59821 cctctcctgg taccattcac cctcaacttc accatcacca acctgcagta tggggaggac
59881 atgggtcacc ctggctccag gaagttcaac accacagaga gggctctgca gggctctgtt
59941 ggtcccatat tcaagaacac cagtgttggc cctctgtact ctggctgcag actgacctt
60001 ctcaggtccg agaaggatgg agcagccact ggagtggatg ccactctgat ccactctt
60061 gacccccaaa gccctggact gnacagngag cngctnact gggagctnag ccantgacc
60121 aannnctcn nngagctgg nccctacacc ctggacaggn acagtctcta tgtcaatggt
60181 ttcaccatc nngantctgn gccaccacc agcactcctg ggacctccac agtgnacntn
60241 ggnacctcng ggaactccatc ctccntccc ngcncacan cncctgnccc tctcctgntn
60301 ccttccaccn tcaacttnac catcaccac ctgcantang nggannacat gennncccn
60361 ggntccagga agttcaacac cacngagngn gtntctcagg gtctgctnnn nccntnttc
60421 aagaacncca gtgtngccn tctgtactct ggctgcagac tgacctnct cagngcngag
60481 aagnatggng cagccactgg antggatgct atctgcanc accnncntna nccaaaagn
60541 cctggactgn acagnagcn gctnactgg gagctnagcc anctgaccaa nncatcnnn
60601 gagctgggnc cctacacctt ggacaggnac agtctctatg tcaatggtt caccatcag
60661 accttgcgc ccaacaccag cactcctggg acctccacag tggacctgg gacctcagg
60721 actccatcct ccctccccag ccctacatct getggccctc tctgggtgcc attcaccctc
60781 aacttcacca tcaccaacct gcagtagcag gaggacatgc atcaccagc ctccaggaag
60841 ttcaacacca cggagcgggt cctgcaggtg ctgcttggtc ccatgttcaa gaacaccagt
60901 gtcggccttc tgtactctgg ctgcagactg acctgtctca ggctgagaa ggcctgagca
60961 gccaccagag tggatgctgt ctgcacctat cgtcctgacc ccaaaagccc tggactgnac
61021 agngagcngc tntactggga gctnagccan ctgaccaann ncatcnnga gctgggnccc
61081 tacacctgag acaggnacag tctctatgct aatggtttca cccatcngan ctctngccc
61141 accaccagca ctctgggac ctccacagtg nacntnggna cctcngggac tccatcctcc
61201 ntcccngcc ncaagcccc tgctcctctc ttgataccat tcacctcaa ctttaccatc
61261 accaacctgc attatgaaga aaacatgcaa caccctggtt ccaggaagtt caaccaccg

Figure 3A (continued)

61321 gagaggggttc tgcaggggtct gctcaagccc ttgttcaaga gcaccagcgt tggccctctg
61381 tactctggct gcagactgac cttgctcaga cctgagaaac atggggcagc cactggagtg
61441 gacgccatct gcaccctccg ccttgatccc actggctctg gactggacag agagcggcta
61501 tactgggagc tgagccagct gaccaacagc gttacagagc tggggcccta caccctggac
61561 agggacagtc tctatgtcaa tggcttcacc cagcggagct ctgtgccaac caccagtatt
61621 cctgggacct ctgcagtgca cctggaaacc tctgggactc cagcctccct ccctggccac
61681 acagcccctg gccctctcct ggtgccattc accctcaact tcactatcac caacctgcag
61741 tatgaggtgg acatgctca ccttggttcc aggaagtca acaccacgga gagagtcctg
61801 cagggctctgc tcaagccctt gttcaagagc accagtgttg gccctctgta ctctggctgc
61861 agactgacct tgctcaggcc tgaaaaacgt ggggcagcca ccggcgtgga caccatctgc
61921 actcaccgcc ttgaccctct aaaccctgga ctggacagag agcagctata ctgggagctg
61981 agcaaaactga cccgtggcat catcgagctg ggcccctacc tcctggacag aggcagtctc
62041 tatgtcaatg gtttcaccca tcggaacttt gtgccatca ccagcactcc tgggacctcc
62101 acagtacacc taggaacctc tgaaacctca tctccctac ctagacccta agtgcctggc
62161 cctctcctgg tgccattcac cctcaacttc accatcacca acttgcaqta tgaggaggcc
62221 atgcgacacc ctggctccag gaagttcaat accacggaga gggctctaca gggctctgctc
62281 aggcccttgt tcaagaatac cagtatcggc cctctgtact ccagctgcag actgaccttg
62341 ctcaggccag agaaggacaa ggcagccacc agagtggatg ccatctgtac ccaccacct
62401 gaccctcaaa gccctggact gaacagagag cagctgtact gggagctgag ccagctgacc
62461 cacggcatca ctgagctggg cccctacacc ctggacaggg acagtctcta tgtcgatggt
62521 ttcactcatt ggagcccct accgaccacc agcactcctg ggacctccat agtgaacctg
62581 ggaacctctg ggatcccacc ttcctccct gaaactacan cnnctgnccc tctcctgntn
62641 ccnttcaccn tcaacttnac catcaccaac ctgcantang nggannacat gcnnccccn
62701 ggntccagga agttcaacac cacngagagg gttctgcagg gtctgctcaa gccctgttc
62761 aagagcacca gtgttggccc tctgtattct ggctgcagac tgaccttgct caggcctgag
62821 aaggacggag tagccaccag agtggacgcc atctgcacc accgcccctga ccccaaatc
62881 cctgggctag acagacagca gctatactgg gagctgagcc agctgaccca cagcatcact
62941 gagctgggac cctacaccct ggataggac agtctctatg tcaatggttt caccagcgg
63001 agctctgtgc ccaccaccag cactcctggg actttcacag tacagccgga aacctctgag
63061 actccatcat ccctccctgg cccacagcc actggccctg tcctgctgcc attcacctc
63121 aattttacca tcaactaacct gcagtatgag gaggacatgc atcgccctgg ctccaggaag
63181 ttcaacacca cggagagggt ccttcagggt ctgcttatgc ccttgttcaa gaacaccagt
63241 gtcagctctc tgtactctgg ttgcagactg acctgtctca ggcctgagaa ggatggggca
63301 gccaccagag tggatgctgt ctgcaccat cgtcctgacc ccaaaagccc tggactggac
63361 agagctgggc tgtactgga gctgagccag ctgaccacag gcatcactga cctgggccc
63421 tacaccctgg acaggcacag tctctatgtc aatggtttca cccatcagag ctctatgacg
63481 accaccagaa ctctgatac ctccacaatg cacctggcaa cctcgagaac tccagcctcc
63541 ctgtctggac ctacgaccgc cagccctctc ctgggtgctat tcacaattaa cttcaccatc
63601 actaacctgc ggtatgagga gaacatgcat caccctggct ctagaaagt taacaccacg
63661 gagagagtcc ttcagggctc gctcaggcct gtgttcaaga acaccagtgt tggccctctg
63721 tactctggct gcagactgac cttgctcagg cccaagaagg atggggcagc caccaaagtg
63781 gatgccatct gcacctaccg ccctgatccc aaaagccctg gactggacag agagcagcta
63841 tactgggagc tgagccagct aaccacagc atcactgagc tgggccccta caccctggac
63901 agggacagtc tctatgtcaa tggtttcaca cagcggagct ctgtgccac cactagcatt
63961 cctgggacct ccacagtgga cctgggaaca tctgggactc cagtttctaa acctggtccc
64021 tcggctgcca gccctctcct ggtgctattc actctcaact tcaccatcac caacctgagg
64081 tatgaggaga acatgcagca cctggctcc aggaagtca acaccacgga gagggtcctt
64141 cagggcctgc tcaggctcct gttcaagagc accagtgttg gccctctgta ctctggctgc
64201 agactgactt tgctcaggcc tgaaaaggat gggacagcca ctggagtgga tgccatctgc
64261 acccaccacc ctgaccccaa aagccctagg ctggacagag agcagctgta ttgggagctg
64321 agccagctga cccacaatat cactgagctg ggccactatg cctggacaa cgacagcctc
64381 tttgtcaatg gtttcaactc tcggagctct gtgtccacca ccagcactcc tgggacccc
64441 acagtgtatc tgggagcatt taagactcca gcctcgatat ttggcccttc agctgccagc
64501 catctcctga tactattcac cctcaacttc accatcacta acctgaggta tgaggagaac
64561 atgtggcctg gctccaggaa gttcaacact acagagagg tccctcagg cctgctaagg
64621 cccttgttca agaaccaccag tgttggccct ctgtactctg gctccaggct gacctgtctc
64681 agccagagaa aagatgggga agccaccgga gtggatgcca tctgcacca ccgcccctgac
64741 cccacaggcc ctgggctgga cagagagcag ctgtatttgg agctgagcca gctgaccac
64801 agcatcactg agctgggccc ctacacactg gacagggaca gtctctatgt caatggtttc
64861 acccatcgga gctctgtacc caccaccagc accggggtgg tcagcgagga gccattcaca
64921 ctgaacttca ccatcaacaa cctgcgtac atggcggaca tgggccaacc cggctccctc
64981 aagttcaaca tcacagacaa cgtcatgaa cactgtctca gtccttgtt ccagagagc
65041 agcctgggtg cacggctacac aggetgcagg gtcatgcac taaggctctgt gaagaacggt
65101 gctgagacac ggtggacct cctctgcacc tactctgcagc ccctcagcgg cccaggtctg

Figure 3A (continued)

65161 cctatcaagc aggtgttcca tgagctgagc cagcagacc atggcatcac cgggctgggc
65221 ccctactctc tggacaaaga cagcctctac cttaacggtt acaatgaacc tggcttagat
65281 gagcctccta caactcccaa gccagccacc acattcctgc ctectctgtc agaagccaca
65341 acagccatgg ggtaccacct gaagaccctc aactcaact tcaccatctc caatctccag
65401 tattcaccag atatgggcaa gggctcagct acattcaact ccaccgaggg ggtccttcag
65461 cacctgctca gacccttggt ccagaagagc agcatgggcc ctttctactt gggttgcaa
65521 ctgatctccc tcaggcctga gaaggatggg gcagccactg gtgtggacac cacctgcacc
65581 taccaccctg accctgtggg ccccgggctg gacatacagc agctttactg ggagctgagt
65641 cagctgacc atggtgtcac ccaactgggc ttctatgtcc tggacagga tagcctcttc
65701 atcaatggct atgcaccca gaatttatca atccggggcg agtaccagat aaatttccac
65761 attgtcaact ggaacctcag taatccagac cccacatcct cagagtacat caccctgctg
65821 agggacatcc aggacaaggt caccacactc tacaaggca gtcaactaca tgacacattc
65881 cgcttctgcc tggtaaccaa ctgacgatg gactccgtgt tggctactgt caaggcattg
65941 ttctcctcca atttggacc cagcctgggt gagcaagtct ttctagataa gaccctgaat
66001 gcctcattcc attggctggg ctccacctac cagttgggtg acatccatgt gacagaaatg
66061 gagtcatcag ttatcaacc aacaagcagc tccagcacc agcacttcta cctgaatttc
66121 accatcacca acctaccata ttcccaggac aaagcccagc caggcaccac caattaccag
66181 aggaacaaaa ggaatattga ggatgcgctc aaccaactct tccgaaacag cagcatcaag
66241 agttatTTTT ctgactgtca agtttcaaca ttcaggtctg tccccaacag gcaccacacc
66301 ggggtggact ccctgtgtaa ctctcgcga ctggctcgga gagtagacag agttgccatc
66361 tatgaggaat ttctgaggat gaccgggaat ggtaccagc tgcagaactt caccctggac
66421 aggagcagtg tccttgtgga tgggtattct cccaacagaa atgagccctt aactgggaat
66481 tctgaccttc cttctgggc tgcatectc atcggcttgg caggactcct gggactcatc
66541 acatgcctga tctgcggtgt cctggtgacc acccgccggc ggaagaagga aggagaatac
66601 aacgtccagc aacagtgcc aggtactac cagtcacacc tagacctgga ggatctgcaa
66661 tgactggaac ttgccggtgc ctggggtgcc tttccccag ccagggtcca aagaagcttg
66721 gctggggcag aaataaacca tattggtcgg aaaaaaaaa aaaaa

Figure 3A (continued)

Polypeptide sequence (SEQ ID No:2)

MLKPSGLPGSSSPTRSLMTGSRSTKATPEMDSGLTGATLSPKTS
 TGAIIVVTEHTLPFTSPDKTLASPTSSVVGRTTQSLGVMSSALPESTSRGMTHSEQRTS
 PSLSPQVNGTPSRNYPATSMVSGLSSPRTRTSSTEGNFTKEASTYTLTVETTSGPVTE
 KYTVPTETSTTEGDSTETPWDTRYIPVKITSPMKTFADSTASKENAPVSMTPAETT
 DSHTPGRTNPSFGTLYSSFLDLSPKGTPNRGETSLELILSTTGYPFSSPEPGSAGHS
 RISTSAPLSSASVLDNKISETSI FSGQSLTSPLSPGVPEARASTMPNSAIPFSMTLS
 NAETSAERVRSTISSLGTPSISTKQTAETILTFHAFNAETMDIPSTHIAKTLASEWLGS
 PGTLLGGTSTALTTTSPSTTLVSEETNTHHSTSGKETEGTLNTSMTPLETSAPGEESE
 MTATLVPTLGFTHLDSKIRSPSQVSSHPTRRELRTTGSTSGRQSSSTAAGSSDILRA
 TTSSTSKASSWTSESTAQQFSEPQHTQWVETSPSMKTERPPASTSVAAPITTSVPSV
 SGFTTLKTSSTKGIWLEETSADTLIGESTAGPTTHQFAVPTGISMTGGSSTRGSQGT
 HLLTRATASSETSADLTLATNGVPSVSPAVSKTAAGSSPPGGTKPSYTMVSSVIPET
 SSLQSSAFREGTSLGLTPLNTRHPFSSPEPDSAGHTKISTSIPLSSASVLEDKVSAT
 STFSHHKATSSITGTPEISTKTKPSSAVLSSMTLSNAATSPERVRNATSPLTHPSPS
 GEETAGSVLTLTSAETDTPNIHPTGTLTSESSESPSTLSLPSVSGVKTTFSSSTPS
 THLFTSGEETEETSNPVSQFETS SVSRVRTLASTSVPTPVFPTMDTWPTSAQFSS
 HLVSELRATSSSTSVTNSTGSALPKISHLTGTATMSQTNRDTFNDSAAQSTTWPETSP
 RFKFTGLPSATTTVSTSATSLSATVMVSKFTSPATSSMEATSIREPSTTILTTETNGP
 GSMAVASTNIPIGKGYITEGRDLTSHLPIGTTASSETSMDFTMAKESVSMVSPSQSM
 DAAGSSTPGRTSQFVDTFSDVYHLTSREITIPRDTGSSALTPQMTATHPPSPDPGSA
 RSTWLGILSSSPSSPTPKVTMSSTFSTQRVTTSMIMDTVETSRWNMPNLPSTTSLTPS
 NIPTSGAIGKSTLVPLDTPSPATSLEASEGGLPTLSTYPESTNTPSIHLGAHASSESP
 STIKLTMASVVKPGSYTPLTFPSIETHIHVSTARMAYSSGSSPEMTAPGETNTGSTWD
 PTTYITTTDPKDTSSAQVSTPHSVRTLRTTENHPKTESATPAAYSGSPKISSPNLTS
 PATKAWTITDTEHSTQLHYTKLAEKSSGFETQSAPGFVSVVIPTSPITIGSSTLELTS
 DVPGEPLVLAPSEQTTITLPMATWLSTSLTEEMASTDLDISSPSSPMSTFAIFPPMST
 PSHELKSEADTSAIRNTDSTTLDQHLGIRSLGRTGDLTTVPITPLTTTWTSVIEHST
 QAQDTLSATMSPHVTQSLKDQTSIPASASPHLTEVYPELGTQGRSSSEATTFWKPS
 TDTLSREIETGPTNIQSTPPMDNTTTGSSSSGVTLGIAHLPIGTSSPAETSTNMALE
 RSSTATVSMAGTMGLLVTSAPGRSISQSLGRVSSVLSESTTEGVTDSKSSPRLNTQ
 GNTALSSSLEPSYAEQSOMSTSIPLTSSPTTPDVEFIGGSTFWTKEVTTVMTSDISKS
 SARTESSSATLMSTALGSTENTGKEKLRASMDLPSPTPSMEVTPWISLTLNAPNTT
 DSLDLSHGVHTSSAGTLATDRSLNTGVTRASRENGSDTSSKSLSMGNSTHTSMTDTE
 KSEVSSSIHPRPETSAPGAETTLTSTPGNRAISLTLPFSSIPVEEVISTGITSGPDIN
 SAPMTHSPITPPTIVWTSTGTIEQSTQPLHAVSSEKVSQVQSTPYVNSVAVSASPTH
 ENSVSSGSSTSSPYSSASLESLDSTISRRAITSWLWDLTSLPPTTWPTSLSEALS
 SGHSGVSNPSTTTEFPFSAASTSAAKQRNPETETHGPQNTAASTLNTDASSVTGLS
 ETPVGASISSEVPLPMAITRSRDVSGLTSESTANPSLGTASSAGTKLRTTISLPTSES
 LVSFRMKNKDPWTVSIPLGSHPTTNETSIPVNSAGPPGLSTVASDVIDTPSDGAESIP
 TVSFSPSPDTEVTTISHFPEKTTTHSFRTISSLTHELTSRVTPIPGDWSSAMSTKPTG
 ASPSITLGERRTITSAAPTTSPIVLTASFTESTVSLDNETTVKTSDILDARKTNELP
 SDSSSSDLINTSIASSTMDVTKTASISPTSISGMTASSPFLFSSDRPQVPTSTTET
 NTATSPSVSSNTYSLDGGSNVGGTPSTLPPFTITHPVETSSALLAWSRPVRTFTMVS
 TDTASGENPTSSNSVVTVPAPGTWASVSGSTTDLFAMGFLKTSPAGEAHSLLASTIEP
 ATAFTPHLSAAVVTGSSATSEASLLTTSSEKAIHSSPQTPPTTPTSGANWETSATPESL
 LVVTETSDTTLSKILVTDITLFTVSTPPSKFPSTGTLGASFPPTLLPDTPAIPLTA
 TEPTSSLATSFDSTPLVTIASDSLGTVPETTLTMSETSNGDALVLKTVSNPDRSIPGI
 TIQGVTESPLHPSSTSPSKI VAPRNTTYEGSITVALSTLPAGTTGSLVFSQSSENSET
 TALVDSSAGLERASVMPLTTGSQGMASGGIRSGSTHSTGKTFSSLPMTMNPGEVTA
 MSEITTNRLTATQSTAPKGI PVKPTSAESGLLTPVSASSSPSKAFASLTAPPSTWGI
 PQSTLTFEFSEVPSLDTKSASLPTPGQSLNTIPDSASTASSLSKSPKPNPRARMMT

Figure 3B

STKAI SASSFQSTGFTET PEGSASPSMAGHEPRVPTSGTGDPYASESMSYPDPKAS
SAMTSTSLASKLTLFSTGQAARSGSSSSPISLSTEKETSFLSPTASTSRKTSLEFLGP
SMARQPNILVHLQTSALTLSPTSTLNMSQEEPELTSSQTIAEEEGTTAETQTLTFP
SETPTSLLPVSSPTEPTARRKSSPETWASSISVPAKTSLVETTDGTLVTTIKMSSQAA
QGNSTWPAPAEETGTSPAGTSPGSPEVSTTLKIMSSKEPSISPEIRSTVRNSPWKTP
TTVPMETTVEPVTLQSTALGSGSTSI SHLPTGTTSPKSPENMLATERVLSLSPSPE
AWTNLYSGTPGGTRQSLATMSSVSLESPTARSITGTGQOSSPELVSKTTGMEFSMWHG
STGGTTGDTHVSLSTSSNILEDPVTS PNSVSSSLTDKSKHKTETWVSTTAIPSTVLNKN
IMAAEQQTSRSVDEAYSSTSSWSQDTS GSDITLGASPDVNTLYITSTAQTSLVSLP
SGDQGITSLTNPSGGKTSSASSVTS PSIGLETLRANVSAVKSDIAPTAGHLSQTS SPA
EVSILDVTTAPTPISTTITMTGNSISTTTPNPEVGMSTMDSTPATERRTSTHEPS
TWSSTAASDSWTVTDMTSNLKVARSPTI STMHTTSFLASSTELDSMSTPHGRITVIG
TSLVTPSSDASAVKTETSTSERTLSPSDTTASTPISTFSRVQRMSISVPDILSTSWTP
SSTEAEADVPMVSTDHASTKTDPNTP LSTFLFDSLSTLDWDTGRSLSSATATTSAPQ
GATTPQELTLETMISPATSQLPFSIGHITSAVTPAAMARSSGVTF SRPDPTSKKAEQT
STQLPTTSAHPGQVPRSAATLDVIPH TAKTPDATFORQQTALTEARATSDSWNE
KEKSTPSAPWITEMMNSVSEDTIKEVTSSSSVLKDEYAGHKLGIWDDFIPKFGKAAH
MRELPLLSPPQDKEAHPSTNTVETTGWVTSSEHASHSTIPAHSASSKLTSPVVTST
REQAIVSMSTTWPESTRARTEPN SFLTIELRDVSPYMDTSSSTQTS LISSPGSTAIT
KGPRTETSSKRISSSFLAQSMRSDSPSEAITRLSNFPAMTESGGMILAMQTSPPGA
TSLSAPTLDTSATASWTGTPLATTQRFTYSEKTTLFSKGPEDTSQPSPPSVEETSSSS
SLVPIHATTSPSNILLTSQGHSPSSTPPVTSVFLSETSGLGKTTDMSRISLEPGTSLP
PNLSSTAGEALSTYEASRDTKAIHHSADTAVTNMEATSSEYSPIPGHTKPSKATSPLV
TSHIMGDITSSTSVFGSSETTEIETVSSVNOGLQERSTSQVASSATETSTVITHVSSG
DATTHVTKTQATFSSGTSISSPHQFITSTNTFTDVSTNPSTSLIMTESSGVTITTQTG
PTGAATQGPYLLDSTMPYLTETPLAVTPDFMQSEKTTLISKGPKDVTWTSPPSVAET
SYPSSLTPFLVTTIPPATSTLQGHSTSSPV SATSVLTSGLVKTTDMLNTSMEPVTNSP
QNLNPSNEILATLAATTDIETIHPSINKAVTNMGTASSAHLHSTLPVSSEPSTATS
PMVPASSMGDALASISIPGSETTDIEGEPTSSSLTAGRKENSTLQEMNSTTESNII LSN
VSVGAI TEATKMEVPSFDATFIPTPAQSTKFPDIFSVASSRLSNSPPMTISTHMTTQ
TGSSGATSKIPLALDSTLETSAGTSPSVVTEGFAHSKIITAMNNDVKDVSQTNPPFQD
EASSPSSQAPVLVTTLPSSVAFTPW HSTSSPVSMSSVLTSSLVKTAGKVDTSLETVT
SSPQMSMNTLDDISVTSAAATTDIETHPSINTVVTVNGTTGSAFESHSTVSAYPEPSK
VTSPNVTTSTMEDTISRIPKSSKTRTETETSSSLTPKLRETSISQEITSSSTETST
VYKELTGATTEVSRDVTSSSSSTSFPGPDQSTVSLDISTETNRLSTSPIMTESAEI
TITTTQTPHGATTSQDFTFTDPSNTPQAGIHSAMTHGFSQLDVTTLMSRI PQDVSWTS
PPSVDKTSSPSSFLSSPAMTTPSLISS TLPEDKLSSEPMTSLTSGLVKIIDILRTRLE
PVTSSLPNFSSSTSDKILATSKDSDKTK EIFPSINTEETNVKANNSGHESHSPALADSE
TPKATTQMVITTTVGDPA PSTSMVHGSSETTNIKREPTYFLTPRLRETSTSQESSFP
TDTSFLLSKVPTGTITEVSS TGVNSSSKI STPDHDKSTVPPDFTTGEIPRVFTSSIKT
KSAEMTITTTQASPPESASHSTLPLDST TSLSQGGTHSTVTQGFYSEVTTLMGMGPGN
VSWMTTPPVEETSSVSSLMSSPAMTSPSPVSSTSPQSI PSSPLPVTALPTSVLVTTTD
VLGTTSPESVTSSPNNLSSITHERPATYKDTAHTAAMHSTNTAVTNVGTSGSGHKS
QSSVLADSETSKATPLMSTTSTLGDTSVSTSTPNISQTNQIQTEPTASLSPRLRESST
SEKTSSTTETNTAFSYVPTGAITQAS RTEISSSRSTSISDLDRPTIAPDISTGMITRFL
TSPIMTKSAEMTVTTQTTTPGATSQGILP WDTSTTLFQGGTHSTVSQGFPHSEITTLR
SRTPGDVSWMTTPPVEETSSGFSLMSPSMTSPSPVSSTSPESIPSSPLPVTALLTSVL
VTTTNVLGTTSPETVTSSPNNLSSPTQERL TTYKDTAHTAAMHSTNTAVANVGT
ISGHESQSSVPADSHTSKATSPMGITFAMGDTSVSTSTPAFFETRIQTESTSSLIPLG
RDTRTSEEINTVTETSTVLSEVPTTTTTEVSRTEVITSSRTTISGPDHSKMSPYISTE
TITRLSTFPFVTGSTEMAITNQ TGPIGTISQATLTLDTSSSTASWEGTHSPVTQRFPHS
EETTTMSRKTGKVSQWQSPSVEETSSPSSPVPLPAITSHSSLYSAVSGSSPTSALPVT
SLLTSGRRKTI DMLDTHSELVTSLSASSFSGEILTSEASTNETIHFSENTAETNM
GTTNSMHHKLHSSVSIHSQPSGHTPPKV TGSMMEDAI VSTSTPGSPETKNVDRDTSPL
TPELKEDSTALVMNSTTESNTVFSSVSLDAATEVSRAEVTTYDPTFMPASAQSTKSPD
ISPEASSSHSNSPPLTISTHKT IATQTGPSGVTS LGQLTLDTSTIATSAGTSPSARTQD

Figure 3B (continued)

FVDSETTSVMNNDLNDVLKTSPPFSAEEANSLSSQAPLLVTTSPSPVTSTLQEHSTSSL
VSVTSVPTPLAKITDMDTNLEPVTRSPQNLNRNTLATSEATTDHTMHPSINTAMANV
GTTSSPNEFYFTVSPDSDPYKATSAVVITSTSGDSIVSTSMRSPSSAMKKIESETTFSL
IFRLRETSTSQKIGSSSDTSTVFDKAFTAATTEVSRTELTSSSRSTSIQGTEKPTMSPD
TSTRSVTMLSTFAGLTKSEERTIATQTGPHRATSQGTLTWDTSITTSQAGTHSAMTHG
FSQLDLSTLTSRVPEYISGTSPPSVEKTSSSSSLLSLPAITSPSPVPTLLPESRPSSP
VHLTSLPTSGLVKTTDMLASVASLPPNLGSTSHKIPTTSEDIKDTEKMYPSTNIAVTN
VGTTSSEKESYSSVPAYSEPPKVTS PMVTSFNIRDTIVSTSMPGSSEITRIEMESTFS
VAHGLKGTSTSQDPIVSTEKSAVLHKLTTGATETSRTTEVASSRRTSIPGPDHSTESP
ISTEVIPLSLPISLGITESSNMTIITRTGPPPLGSTSOGTFTLDTPTTSSRAGTHSMATQ
EFPHSEMTVMNKDPEILSWTIPPSIEKTSFSSSLMPPSPAMTSPPPVSSSTLPKTIHTTP
SPMSTLLTPSLVMTTDTLGTSPPEPTSSPPNLSSTSHVILTTDEDTTAIEAMHPSTST
AATNVETTCSGHGSQSSVLTDESEKTKATAPMDTTSTMGHSTTVSTSMSVSSETTKIKRE
STYSLTPGLRETSIQNASFSTDTISIVLSEVPTGTTAEVSRTEVTSSGRTSIPGPSQS
TVLPEISTRMTRLFASPTMTEAEMTIPTQTGPGSSTSQDTLLTDTSTTKSQAKTHS
TLTQRFPHSEMTTLMRGPDMWQSSPSLENPSSLPSSLPLPATTSPPPISSTLPVT
ISSSPLPVTSLTSSPVTTTDMMLHTSPELVTSSPPKLSHTSDERLTTGKDTTNTTEAVH
PSTNTAASNVEIPFSGHESPSSALADSETSKATSPMFITSTQEDTTVAISTPHFLETS
RIQKESISSLSPKLRETGSSVETSSAIETSAVLSEVSI GATTEISRTEVTSSSRSTSIS
GSAESTMLPEISTTRKIKFPTSPILAESSEMTIKTQTSPPGSTSESTFTLDTSTTPS
LVITHSTMTQRLPHSEITTLVSRGAGDVPRPSSLPVEETSPSSQLSLSAMISPPVVS
STLPASSHSSASVTSPLTPGQVKTTEVLDAEAPETSSPPSLSSTSV EILATSEVTT
DTEKIHFPNTAVTKVGTSSSGHESPSSVLPDSETTKATSAMGTISIMGDTSVSTLTP
ALSNTRKIQSEPASSLTRLRETSTSEETSLATEANTVLSKVSTGATTEVSRTEAIF
SRTSMGPEQSTMSQDISIGTIPRISASSVLTESAKMTITTQTGPSESTLESTLNLNT
ATTPSWVETHSIVIQQFPHPEMTTSMGRGPGGVSWSPPPFKETSPPSSPLSLPAVTS
PHPVSTTFLAHI PPSPLPVTSLTSGPATTTDILGTSTEPGTSSSSSLSTTSHERLTT
YKDTAHEAVHPSTNTGGTNVATTSSGYKSQSSVLADSSPMCTTSTMGDTSVLTSTPA
FLETRRIQTELA SSLTPGLRESSGSEGTSSGKMSVLSKVPTGATTEISKEDVTSIP
GPAQSTISPDISTRVSWFSTSPVMTESAEITMNTHTSPLGATTQGTSTLATSSTTSL
TMHTSTISQGFHSQMSTLMRRGPEDVSWMSPPLEKTRPSFSLMSSPATTSPSPVSS
TLPESISSSPLPVTSLTSGLAKTTDMLHKSSEPVTNSPANLSTSV EILATSEVTTD
TEKTHPSSNRTVTDVGTSSSGHESTSFVLADSQTSKVTS PMVTSIMEDTSVSTSTPG
FFETSRIQTEPTSSSLTLGLRKTSSSEGTSLATEMSTVLSGVPTGATAEVSRTEVTSS
RTSISGFAQLTVSPETSTETITRLPTSSIMTESAEMMIKTQTDPPGSTPESTHTVDIS
TTPNWVETHSTVTRQRFHSEM TTVLSRSPGDMLWPSQSSVEETSSASSLLSLPATTSP
SPVSSSTLVEDFPSASLPVTSLTLPGLVITTD RMGISREPGTSSSTNLSSTSHERLTTL
EDTVDTEDMQPSTHTAVTNVRTSISGHESQSSVLS DSETPKATSPMGTTYTMGETSVS
ISTSDFFETSRIQIEPTSSSLTSGLRETSSSERISSATEGSTVLEVP SGATTEVSRTE
VISSRGTSMGPDQFTISPDISTEAITRLSTSPIMTESAESAITIETGSPGATSEGLT
TLDTSTTTTFWSGTHSTASPGFHSSEM TLMRTPGDVPWPSLPVVEEASSVSSSLSP
AMTSTSFSSALPESISSSHPVTAALLTLGPVKTTDMLRTSSEPETSSPPNLSSTSAEI
LATSEVTKDREKIHPSNTPVNVGTVIYKHLSPSSVLADLVTKPTSPMATTSTLGN
TSVSTSTPAFPETMMTQPTSSSLTSGLREISTSQETSSATERSASLSGMPTGATTKVSR
TEALSLGRTSTPGPAQSTISPEISTETITRISTPLTTTGAEMTITPKTGHSGASSQG
TFTLDTSSRASWPGTHSAATHRSPHSGMTTPMSRGPEDVSWPSRPSVEKTSPPSSLV
LSAVTSPSPLYSTPSESSHSSPLRVTSLFTPVMMKTTDMLDTSLEPVTSPSPMNI
DESLATSKATMETEAIQISENTAVTQMGTISARQEFYSSYPGLPEPSKVTS PVVTSST
IKDIVSTTI PASSEITRIEMESTSTLTPTPRETSTSQEIH SATKPSTVPYKALTSATI
EDSMTQVMSSSRGSPDQSTMSQDISSEVITRLSTSPIKAESTEMITTQTGSPGATS
RGTLLTLDTSTTFMSGTHSTASQGFHSQM TALMSRTPGDVPWLSHPSVVEEASSASFSL
SSPVMTSSSPVSSSTLPDSIHSSSLPVTSLTSGLVKTT ELLGTSSEPETSSPPNLSST
SAEILATTEVTTDTEKLEMTNVVTSGYTHESPSSVLADSVTTKATSSMGITYPTGDTN
VLTSTPAFSDTSRIQTKSKLSLTPGLMETSISEETSSATEKSTVLSVPTGATTEVSR
TEAISSSRSTSIPGPAQSTMSSDTSMETITRISTPLTRKESTDMAITPKTGPSGATSQG
TFTLDSSTASWPGTHSATTQRFQSVVTT PMSRGPEDVSWPSPLSVEKNSPPSSLV
SSSVTSPSPLYSTPSGSSHSSPVVTSLFTS IMMKATDMLDASLEPETTSAPNMNITS
DESLATSKATTETEAIHV FENTAASHVETTSATEELYSSSPGFSEPTKVISPVVTSST
IRDNMVSTTMPGSSGITRIEIESMSSLTPLGRETRTSQDITSSSTETSTVLYKMSSGAT

Figure 3B (continued)

PEVSRTEVMPSSRTSIPGPAQSTMSLSDISDEVVTRLSTSPIMTESAEITITTQTGYSL
ATSQVTLPLGTSMTFLSGTHSTMSQGLSHSEMTNLMRGPESLSWTSRPFVETTRSSS
SLTSLPLTTSLSVSSSTLLDSSPSSPLPVTSLILPGLVKTTTEVLDTSSSEPKTSSSPNL
SSTSVEIPATSEIMTDTEKIHPSNNTAVAKVRTSSSVHESHSSVLADSETTITIPSMG
ITSAVDDTTVFTSNPAFSETRRIPTPTFSLTPGFRETSTSEETTSITETSAVLYGVP
TSATTEVSMTEIMSSNRTHIPDSQSTMSPDIIITEVITRLSSSSMMSESTQMTITTQK
SSPGATAQSTLTLATTTAPLARTHSTVPPRFLHSEMTTLMRSPENPSWKSSPFVEKT
SSSSLLSLPVTSPSVSSTLPQSI PSSSFSVTSLLT PGMVKT TDTSTEPGTSLSPNL
SGTSVEILAAASEVTTDEKIHPSSSMAVTNVGTTSSGHELYSSVSIHSEPSKATYPVG
TPSSMAETSISTSM PANFETTGF EAEPF SHLTSGFRKTNMSLDTSSVTPTNTPSSPGS
THLLQSSKTDFTSSAKTSSPDWPPASQYTEIPVDIITPFNASPSITESTGITSFPESR
FTMSVTESTHHLSTDLLPSAETISTGTVMPSLSEAMTSFATTGVPRAISGSGSPFSRT
ESGPGDATLSTIAESLPSSTPVFSSSTFTTTDSSTIPALHEITSSSATPYRVDTSLG
TESSTTEGRLVMVSTLDTSSQPGRSSTPILDTRMTESVELGTVTSAYQVPSLSTRLT
RTDGIMEHITKI PNEAAHRGTIRPVKGPQTSTSPASPKGLHTGGTKRMETTTTALKTT
TTALKTTSRATLTSVYPTTLGLTLP LNASRQMASTILTEMMITTPYVFPDVPETTSS
LATSLGAETSTALPRTTPSVLNRESETTASLVSRGAERSPVIQTL DVSSSEPDTTAS
WVIHPAETIPTVSKTTPNFHSELDTVSSSTATSHGADVSSAIPTNISPEL DALTPLV
TISGTDSTTTFPTLTKSPHETETRTTWLTHPAETSSTIPRTIPNFHSHESDATPSIAT
SPGAETSSAIPIMTVSPGAEDLVT SQVTSSGTDNRMTIPTLTLSPGEPKTIASLVTHP
EAQTSSAIPSTISPAVSRVLVSMVTS LAAKTSTTNRALTNSPGEPATTVSLVTHPAQ
TSPTVPWTTSIFFHKSDDTTPSMTSHGAESSAVPTPTVSTEVPGVVTPLVTSRAV
ISTTIPILTLSPGEPETTPSMATSHGEEASSAIPPTVSPGVPGVVTSLVTSRAVTS
TTIPILTFLSGLGEPETTPSMATSHGTEAGSAVPTVLPEVPGMVTSLVASSRAVSTTLP
TLTLSPGEPETTPSMATSHGAEASSTVPTVSPVPGVVTSLVTSSSGVNSTSIPTLIL
SPGELETTSPMATSHGAEASSAVPTPTVSPGVSGVVTPLVTSRAVSTTIPILTSS
SEPETTPSMATSHGVEASSAVLTVSPEVPGMVTSLVTSRAVSTTIPTLTSSDEPE
TTTSLVTHSEAKMISAIPTLAVSPTVQGLVTSLVTSSSGSETSAFNLTVASSQPETID
SWVAHPGTEASSVPTLVSTGEPFTNISLVTHPAESSSTLPRTTSRFSHSELDTMPS
TVTSPEAESSAISTTISPGIPGVLTSLVTSSSGRDISATFPTVPESPHESEATASWVT
HPAVTSTTVPRTPPNYSHSEPDTTPSIATSPGAEATSDFPITVSPDVPDMVTSQVTS
SGTDTSTITPTLTLSSGEPETTTSFITYSETHTSSAIPTLVSPGASKMLTSLVSSG
TDTTTFPTLTPETPYEPETIAIQLIHPAETNTMVKPTPKFHSKSDTTLPVAITSPG
PEASSAVSTTTISPDMSDLVTSLVPSGGTDTSTTTFPTLSETPYEPETTVTWLTHPAET
STTVSGTIPNFHSHRSDTAPSMVTS PGVDTRSGVPTTIPPSIPGVVTSQVTSSATDT
STAIPTLTPSPGEPETTASSATHPGTQTGFTVPIRTVPSSEPDTMASWVTHPPQTSTP
VSRTTSSFSHSSPDATPVMATSPRTEASSAVLTTISPGAPEMVTSQITSSGAATSTTV
PTLTHSPGMPETTALLSTHPRGTGSKTFPASTVFPQVSETTASLTIRPGAETSTALPT
QTTSSLFTLLVTGTSRVDLSPASPGVSAKTAPLSTHPGTETSTMIPTSLSLGLET
TGLLATSSSAETSTSTLTLTVSPAVSGLSSASITDKPQTVTSWNTETSPSVTSVGGP
EFSRTVTGTTMTLIPSEMPTPKTSHGEGVSPTTILRTTMVEATNLATTGSSPTVAKT
TTTFNTLAGSLFTPLTTPGMSTLASESVTSRSTSYNHRSWISTTSSYNRRYWT PATSTP
VTSTFSPGISTSSIPSSAATVPFMVPTLNFTITNLQYEEDMRHPGSRKFNATEREL
QGLLKPLFRNSSLEYLYSGCRLASLRPEKDSSAMA VDAICTHRPDPEDLGLDRERLYW
ELSNLTNGIQELGPYTLDRNSLYVNGFTHRSMPTTSTPGTSTVDVGTSGTPSSSPSP
TAAGPLLMPFTLNFTITNLQYEEDMRRTGSRKFNTEMESVLQGLLKPLFKNTSVGPLY
GCRLTLRPEKDGAATGVDAICTHRLDPKSPGLNREQLYWELSKLTNDIEELGPYTL
RNSLYVNGFTHQSSVSTTSTPGTSTVDLRTSGTPSSLSPTIMAAGPLLVPFTLNFTI
TNLQYGEDMHPGSRKFNTERVLQGLLGP I FNKNTSVGPLYSGCRLTSLRSEKDGAAT
GVDAICIHHLDPKSPGLNRERLYWELSQLTNGIKELGPYTLDRNSLYVNGFTHRTSVP
TTSTPGTSTVDLGTSGTPFSLPSPATAGPLLVLFTLNFTITNLQYEEDMRHPGSRKFN
TTERVLQTLGPMFKNTSVGLLYSGCRLTLRSEKDGAATGVDAICTHRLDPKSPGLD
REQLYWELSQLTNGIKELGPYTLDRNSLYVNGFTHWIPVPTSSTPGTSTVDLGTSGTPS
SLPSPTAAGPLLVPFTLNFTITNLQYEEDMHPGSRKFNTERVLQGLLGPMPFKNTSV
GLLYSGCRLTLRSEKDGAATGVDAICTHRLDPKSPGVDRREQLYWELSQLTNGIKELG
PYTLDRNSLYVNGFTHQTSAPNTSTPGTSTVDLGTSGTPSSLPSPSAGPLLVPFTLN
FTITNLQYEEDMRHPGSRKFNTERVLQGLLKPLFKNTSVGPLYSGCRLTLRSEKDG
AATGVDAICTHRLDPKSPGVDRREQLYWELSQLTNGIKELGPYTLDRNSLYVNGFTHQ
SAPNTSTPGTSTVDLGTSGTPSSLPSPSAGPLLVPFTLNFTITNLQYEEDMHPGSR

Figure 3B (continued)

KFNTTERVLQGLLGPMPFKNTSVGLLYSGCRLTLLRPEKNGAATGMDAICSHRLDPKSP
GLNREQLYWELSQLTHGIKELGPYTLDRNSLYVNGFTHRSSVAPTSTPGTSTVDLGT
GTPSSLPSPTTAVPLLVPTLNFTITNLQYGEDMRHPGSRKFNTTERVLQGLLGPLFK
NSSVGPLYSGCRLISLRSEKDGAATGVDAICTHHLNPQSPGLDREQLYWQLSOMTNGI
KELGPYTLDRNSLYVNGFTHRSSGLTTSTPWTSTVDLGTSGTPSPVPSPTTAGPLLV
FTLNFTITNLQYEDMRHPGSRKFNTTERVLQGLLSPIFKNSSVGPLYSGCRLTSLRP
EKDGAATGMDAVCLYHPNPKRPGLDREQLYWELSQLTHNITELGPYSLDRDSLYVNGF
THQNSVPTTSTPGTSTVYWATTGTPSSFPGHTEPGPLLIPTFNFTITNLHYEENMQH
PGSRKFNTTERVLQGLLKLPLFKNTSVGPLYSGCRLTSLRPEKDGAATGMDAVCLYHPN
PKRPGLDREQLYWELSQLTHNITELGPYSLDRDSLYVNGFTHQNSVPTTSTPGTSTVY
WATTGTPSSFPGHTEPGPLLIPTFNFTITNLHYEENMQHPGSRKFNTTERVLQGLL
PLFKNTSVGPLYSGCRLTLLRPEKHEAATGVDTICTHRVDPGIGPLDRERLYWELSQL
TNSITELGPYTLDRDSLYVNGFNPRSSVPTTSTPGTSTVHLATSGTPSSLPGHTAPV
LLIPTLNFTITNLHYEENMQHPGSRKFNTTERVLQGLLKLPLFKNTSVGPLYSGCRL
LLRPEKHEAATGVDTICTHRVDPGIGPLXEXLYWELSLTXIXELGPYTLDRXSLY
VNGFTHXXSPTTSTPGTSTVXXGTSPTSSXPXTSAGPLLVPTLNFTITNLQYEE
DMHHPGSRKFNTTERVLQGLLGPMPFKNTSVGLLYSGCRLTLLRPEKNGAATGMDAICS
HRLDPKSPGLDREQLYWELSQLTHGIKELGPYTLDRNSLYVNGFTHRSSVAPTSTPGT
STVDLGTSGTPSSLPSPTTAVPLLVPTLNFTITNLQYGEDMRHPGSRKFNTTERVLQ
GLLGPLFKNSSVGPLYSGCRLISLRSEKDGAATGVDAICTHHLNPQSPGLDREQLYWQ
LSOMTNGIKELGPYTLDRNSLYVNGFTHRSSGLTTSTPWTSTVDLGTSGTPSPVPSPT
TAGPLLVPTLNFTITNLQYEDMRHPGSRKFNTTERVLQGLLSPIFKNSSVGPLYSG
CRLTSLRPEKDGAATGMDAVCLYHPNPKRPGLDREQLYWELSQLTHNITELGPYSLDR
DSLYVNGFTHQSSMTTTRTPDSTMHLATSRTPASLSGPTTASPLLVFTINCTITNL
QYEDMRRTGSRKFNTMESVLQGLLKLPLFKNTSVGPLYSGCRLTLLRPKKDGAATGVD
AICTHRLDPKSPGLNREQLYWELSKLTNDIEELGPYTLDRNSLYVNGFTHQSSVSTTS
TPGTSTVDLRTSGTPSSLSSPTIMXXXPLLPFTXNXTITNLXXXXMXXPGSRKFNT
TERVLQGLLRPLFKNTSVSSLYSGCRLTLLRPEKDGAATRVDAACTYRPDPKSPGLDR
EQLYWELSQLTHSITELGPYTLDRVSLYVNGFNPRSSVPTTSTPGTSTVHLATSGTPS
SLPGHTXXXPLLPFTXNXTITNLXXXXMXXPGSRKFNTTERVLQGLLKLFRNSSL
EYLYSGCRLASLRPEKDSSAMAVDAICTHRPDPEDLGLDRERLYWELSNLTNGIQELG
PYTLDRNSLYVNGFTHRSSGLTTSTPWTSTVDLGTSGTPSPVPSPTTAGPLLVPTLN
FTITNLQYEDMRHPGSRKFNTTERVLQGLLPLFKNTSVGPLYSGCRLTLLRPEKQE
AATGVDTICTHRVDPGIGPLDRERLYWELSQLTNSITELGPYTLDRDSLYVNGFN
PWSVPTTSTPGTSTVHLATSGTPSSLPGHTAPVPLLIPTLNFTITDLHYEENMQHPGSR
KFNTTERVLQGLLKLPLFKSTSVGPLYSGCRLTLLRPEKHGAATGVDAICTLRDPTGP
GLDRERLYWELSQLTNSVTELGYPYTLDRDSLYVNGFTHRSSVPTTSTPGTSAVHLETS
GTPASLPHTAPGPLLVPTLNFTITNLQYEDMRHPGSRKFSTTERVLQGLLKLPLFK
NTSVSSLYSGCRLTLLRPEKDGAATRVDAVCTHRPDPKSPGLDRERLYWKLSQLTHGI
TELGPYTLDRHSLYVNGFTHQSSMTTTRTPDSTMHLATSRTPASLSGPTTASPLLV
FTINFTITNLRYEENMHHPGSRKFNTTERVLQGLLRPVFKNTSVGPLYSGCRLTLLRP
KKDGAATKVDICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTQDRDSLYVNGF
THRSSVPTTSTPGTSAVHLETSGTPASLPHTAPGPLLVPTLNFTITNLQYEDMRH
PGSRKFNTTERVLQGLLKLPLFKSTSVGPLYSGCRLTLLRPEKGAATGVDTICTHRLD
PLNPGLDREQLYWELSKLTRGIELGPYLLDRGSLYVNGFTHRTSVPTTSTPGTSTVD
LGTSGTPFSLPSAXXXPLLPFTXNXTITNLXXXXMXXPGSRKFNTTERVLQTLG
PMFKNTSVGLLYSGCRLTLLRSEKDGAATGVDAICTHRLDPKSPGVDREQLYWELSQL
TNGIKELGPYTLDRNSLYVNGFTHWIPVPTSSTPGTSTVDLGSPTSSSLPSPTTAGPL
LVPTLNFTITNLKYEEDMHCPSGRKFNTTERVLQSLGPMFKNTSVGPLYSGCRLT
LRSEKDGAATGVDAICTHRLDPKSPGVDREQLYWELSQLTNGIKELGPYTLDRNSLYV
NGFTHQTSAPNTSTPGTSTVDLGTSGTPSSLPSPTXXXPLLPFTXNXTITNLXXXX
MXXPGSRKFNTTEXVLQGLLXPFXNXSVGXLYSGCRLTXLRXEXGAATGXDAICXH
XXXPKXPGLXEXLYWELSLTXIXELGPYTLDRXSLYVNGFTHWIPVPTSSTPGTS
TVDLGSPTSSSLPSPTTAGPLLVPTLNFTITNLKYEEDMHCPSGRKFNTTERVLQSL
LGPMPFKNTSVGPLYSGCRLTSLRSEKDGAATGVDAICTHRVDPKSPGVDREQLYWEL
QLTNGIKELGPYTLDRNSLYVNGFTHQTSAPNTSTPGTSTVXXGTSPTSSXPXTSA
GPLLVPTLNFTITNLQYEDMHHPGSRKFNTTERVLQGLLGPMPFKNTSVGLLYSGCR

Figure 3B (continued)

LTLRPEKNGATTGMDAICTHRLDPKSPGLXXEXLYWELSXLTXIXELGPYTLDRXS
 LYVNGFTHXXSXPTTSTPGTSTVXXGTSSTPSSXPXXTXXXPLLPFTXNXTITNLXX
 XXXMXXPGSRKFNTTERVLQGLLKPLFRNSSLEYLYSGCRLASLRPEKDSSAMAVDAI
 CTHRPDPEDLGLDRERLYWELSNLTNGIQELGPYTLDRNSLYVNGFTHRSSMPTTSTP
 GTSTVDVGTSGTPSSSPSPTTAGPLLIPTLNFTITNLQYGEDMGHPGSRKFNTTERV
 LQGLLGPVIFKNTSVGPLYSGCRLTSLRSEKDGAATGVDAICIHHLDPKSPGLNRERLY
 WELSQTNGIKELGPYTLDRNSLYVNGFTHRTSVPTTSTPGTSTVDLGTSGTPFSLPS
 PATAGPLLVLFNLFTITNLKYEEDMHRPGSRKFNTTERVLQTLGPMFKNTSVGLLY
 SGCRLTLRSEKDGAATGVDAICTHRLDPKSPGLXXEXLYWELSXLTXIXELGPYTL
 DRXSLYVNGFTHXXSXPTTSTPGTSTVXXGTSSTPSSXPXXTXXXPLLPFTXNXTIT
 NLXXXXMXXPGSRKFNTTERVLQGLLRPVFKNTSVGPLYSGCRLTLRPPKDGAAATK
 VDAICTYRDPKSPGLDREQLYWELSQTLSITELGPYTQDRDSLYVNGFTHRSSVPT
 TSIPGTSAVHLETTGTPSSFPGHTEPGPLLIPTFNFTITNLRYEENMQHPGSRKFNT
 TERVLQGLLTPFKNTSVGPLYSGCRLTLRPEKQEAATGVDICTHRVDPVIGPGLDR
 ERLYWELSQTNSITELGPYTLDRDSLYVDGFNPWSSVPTTSTPGTSTVHLATSGTSP
 PLPGHTAPVPLLIPTLNFTITDLHYEENMQHPGSRKFNTTERVLQGLLKPLFKSTSV
 GPLYSGCRLTLRPEKHGAATGVDAICTLRDPTGPGLDRELYWELSQTNSITELG
 PYTLDRDSLYVNGFNPWSSVPTTSTPGTSTVHLATSGTPSSLPGHTEGPTLVPFTLN
 FTITNLKYEEDMHCPSGRKFNTTERVLQSLHGPMFKNTSVGPLYSGCRLTLRSEKDG
 AATGVDAICTHRLDPKSPGLXXEXLYWELSXLTXIXELGPYTLDRXSLYVNGFTHXX
 SXPTTSTPGTSTVXXGTSSTPSSXPXXTXXXPLLPFTXNXTITNLXXXXMXXPGSR
 KFNTTEXVLQGLLXPFXKXSVGXLYSGCRLTXLRXKXGAATGXDAICXHXXPXKXP
 GLXXEXLYWELSXLTNSITELGPYTLDRDSLYVNGFTHRSSMPTTSPGTSVHLETS
 GTPASLPGHTEGPTLVPFTLNFTITNLQYEEEMRHPGSRKFNTTERVLQGLLKPLFK
 STSVGPLYSGCRLTLRPEKRGAAATGVDICTHRLDPLNPLXXEXLYWELSXLTXIXI
 XELGPYTLDRXSLYVNGFTHXXSXPTTSTPGTSTVXXGTSSTPSSXPXXTXXXPLLP
 FTXNXTITNLXXXXMXXPGSRKFNTTEXVLQGLLXPFXKXSVGXLYSGCRLTXLRX
 EKXGAATGXDAICXHXXPXKXPGLXXEXLYWELSXLTXIXELGPYTLDRXSLYVNGF
 HPRSSVPTTSTPGTSTVHLATSGTPSSLPGHTEGPTLVPFTLNFTITNLHYEENMQH
 PGRKFNTTERVLQGLLGPMPFKNTSVGLLYSGCRLTLRPEKNGAATGMDAICSHRDL
 PKSPGLXXEXLYWELSXLTXIXELGPYTLDRXSLYVNGFTHXXSXPTTSTPGTSTVX
 XGTSSTPSSXPXXTXXXPLLPFTXNXTITNLXXXXMXXPGSRKFNTTEXVLQGLLX
 PFXKXSVGXLYSGCRLTXLRXKXGAATGXDAICXHXXPXKXPGLXXEXLYWELSXL
 TXIXELGPYTLDRXSLYVNGFTHQNSVPTTSTPGTSTVYWATTGTPSSFPGHTEGPT
 LLIPTFNFTITNLHYEENMQHPGSRKFNTTERVLQGLLTPFKNTSVGPLYSGCRLT
 LLRPEKQEAATGVDICTHRVDPVIGPGLXXEXLYWELSXLTXIXELGPYTLDRXSLY
 VNGFTHXXSXPTTSTPGTSTVXXGTSSTPSSXPXXTXXXPLLPFTXNXTITNLXXXX
 MXXPGSRKFNTTEXVLQGLLXPFXKXSVGXLYSGCRLTXLRXKXGAATGXDAICX
 HXXXPXKXPGLXXEXLYWELSXLTXIXELGPYTLDRXSLYVNGFTHRSSVPTTSSPGT
 STVHLATSGTPSSLPGHTEGPTLVPFTLNFTITNLHYEENMQHPGSRKFNTTERVLQ
 GLLKPLFKSTSVGPLYSGCRLTLRPEKHGAATGVDAICTLRDPTGPGGLXXEXLYWE
 LSXLTXIXELGPYTLDRXSLYVNGFTHXXSXPTTSTPGTSTVXXGTSSTPSSXPXXT
 XXXPLLPFTXNXTITNLXXXXMXXPGSRKFNTTEXVLQGLLXPFXKXSVGXLYSG
 CRLTXLRXKXGAATGXDAICXHXXPXKXPGLXXEXLYWELSXLTXIXELGPYTLDR
 XSLYVNGFTHRTSVPTTSTPGTSTVHLATSGTPSSLPGHTEGPTLVPFTLNFTITNL
 QYEEEMRHPGSRKFNTTERVLQGLLSPVIFKNSSVGPPLYSGCRLTSLRPEKDGAATGMD
 AVCLYHPNPKRPGLDREQLYCELSQLTHNITELGPYSLDRDSLYVNGFTHQNSVPTT
 TPGTSTVYWATTGTPSSFPGHTEGPTLVPFTXNXTITNLXXXXMXXPGSRKFNTTE
 XVLQGLLXPFXKXSVGXLYSGCRLTXLRXKXGAATGXDAICXHXXPXKXPGLXXEX
 LYWELSXLTXIXELGPYTLDRXSLYVNGFTHWSSGLTTSTPWTSTVDLGTSGTPSPV
 PSPTTAGPLLVPTLNFTITNLQYEEEMRHPGSRKFNTTERVLQGLLSPVIFKNSSVGP
 PLYSGCRLTLRPEKQEAATGVDICTHRVDPVIGPGLXXEXLYWELSXLTXIXELGPY
 TLDRXSLYVNGFTHXXSXPTTSTPGTSTVXXGTSSTPSSXPXXTXXXPLLPFTXNXT
 ITNLXXXXMXXPGSRKFNTTEXVLQGLLXPFXKXSVGXLYSGCRLTXLRXKXGAA
 TGXDAICXHXXPXKXPGLXXEXLYWELSXLTXIXELGPYTLDRXSLYVNGFTHRSFG
 LTTSTPWTSTVDLGTSGTPSPVPSPTTAGPLLVPTLNFTITNLQYEEEMRHPGSRKF

Figure 3B (continued)

NTTERVLQGLLTPLFRNTSVSSLSYSGCRLTLLRPEKDGAATRVDVAVCTHRPDPKSPGL
 XXEXLYWELSXLTXIXELGPYTLDRXSLYVNGFTHXXSXPTTSTPGTSTVXXGTS
 PSSXPXXTXXXPLLPFTXNXTITNLXXXXMXXPGSRKFNTTEXVLQGLLXPXFKNX
 SVGXLYSGCRLTXLRXEKXGAATGXDAICXHXXPKXPGLXXEXLYWELSXLTXIXE
 LGPYTLDRXSLYVNGFTHWIPVPTSSTPGTSTVDLGSSTPSSLPSPPTTAGPLLVFTL
 NFTITNLQYGEDMGHPGSRKFNTTERVLQGLLGPVFKNTSVGPLYSGCRLTSLRSEKD
 GAATGVDAICIHHLDPKSPGLXXEXLYWELSXLTXIXELGPYTLDRXSLYVNGFTHX
 XSXPTTSTPGTSTVXXGTSSTPSSXPXXTXXXPLLPFTXNXTITNLXXXXMXXPGS
 RKFNTTEXVLQGLLXPXFKNXSVGXLYSGCRLTXLRXEKXGAATGXDAICXHXXPKX
 PGLXXEXLYWELSXLTXIXELGPYTLDRXSLYVNGFTHQTFAPNTSTPGTSTVDLGT
 SGTSSLPSPPTAGPLLVFTLNFTITNLQYEDMHHHPGSRKFNTTERVLQGLLGPV
 KNTSVGLLYSGCRLTLLRPEKNGAATRVDVAVCTHRPDPKSPGLXXEXLYWELSXLTX
 IXELGPYTLDRXSLYVNGFTHXXSXPTTSTPGTSTVXXGTSSTPSSXPXXTAPVPLLI
 PFTLNFTITNLHYEENMQHPGSRKFNTTERVLQGLLKPFLKSTSVGPLYSGCRLTLLR
 PEKHGAATGVDAICTLRDPTGGLDRERLYWELSQTNSVTELGPYTLDRDSLYVNG
 FTQRSSVPTTIPGTSAVHLETSGTPASLPGHAPGPLLVFTLNFTITNLQYEDMR
 HPGSRKFNTTERVLQGLLKPFLKSTSVGPLYSGCRLTLLRPEKNGAATGVDTICTHRL
 DPLNPGLDREQLYWELSKLTRGIIELGPYLLDRGSLYVNGFTHRNFPITSTPGTSTV
 HLGTSSTPSSLPRIVPGPLLVFTLNFTITNLQYEEAMRHPGSRKFNTTERVLQGLL
 RPLFKNTSIGPLYSSCRLTLLRPEKDKAATRVDVAICTHHPDPQSPGLNREQLYWEL
 LTHGITELGPYTLDRDSLYVDGFTHWSPITPTSTPGTSIVNLGTSGIPPSLPETXXX
 PLLXPFTXNXTITNLXXXXMXXPGSRKFNTTERVLQGLLKPFLKSTSVGPLYSGCRL
 TLLRPEKDVATRVDVAICTHRPDPKI PGLDRQQLYWELSQTLSITELGPYTLDRDSL
 YVNGFTQRSSVPTTSTPGTFTVQPETSETPSSLPGPTATGPVLLPFTLNFTITNLQY
 EDMHRPGSRKFNTTERVLQGLLMPFKNTSVSSLSYSGCRLTLLRPEKDGAATRVDVAV
 CTHRPDPKSPGLDRERLYWKLSQLTHGITELGPYTLDRHSLYVNGFTHQSSMTTTRTPD
 TSTMHLATSRTPASLSGPTTASPLLVFTINFTITNLRYEENMHHHPGSRKFNTTERVL
 QGLLRPVFKNTSVGPLYSGCRLTLLRPKKGAATKVDAICTYRDPKSPGLDREQLYW
 ELSQLTHSITELGPYTLDRDSLYVNGFTQRSSVPTTIPGTPTVDLGTSGTPVSKPGP
 SAASPLLVFTLNFTITNLRYEENMQHPGSRKFNTTERVLQGLLRSLFKSTSVGPLY
 GCRLTLLRPEKDGATGVDAICTHHPDPKSPRLDREQLYWELSQTLSITELGHYALD
 NDSL FVNGFTHRSSVTTSTPGTPTVYL GASKTPASIFGSAASHLLILFTLNFTITN
 LRYEENMWPGRKFNTTERVLQGLLRPLFKNTSVGPLYSGSRLTLLRPEKDGEATGVD
 AICTHRPDPTGGLDREQLYLELSQLTHSITELGPYTLDRDSLYVNGFTHRSSVPTT
 TGVVSEEPFTLNFTINNLRYMADMGPGLKFNITDNVMKHLSPFQRSSLGARYTG
 CRVIALRSVKNGAETRDLLCTYLQPLSGPLPIKQVFHELSSQTHGITRLGPYSLDK
 DSYLNGYNEPGLDEPPTPKPATTFLPPLSEATTAMGYHLKTLTLNFTISNLQYSPD
 MGKGSATFNSTEGVLQHLRPLFQKSSMGPFFYLGCLISLRPEKDGAATGVDTTCTYH
 PDPVGPGLDIQQLYWELSQTLSHGVTLGFFVLDLDRSLFINGYAPQNLSIRGEYQINFH
 IVNWNLSPDPTSSEYITLLRDIQDKVTTLTKGSQLHDTFRFCLVTNLTMDSVLVTVK
 ALFSSNLDPSLVEQVFLDKTLNASFWHLGSTYQLVDIHVTEMESVYQPTSSSSTQHF
 YLNFTITNLPYSQDKAQPGTTNYQRNKRNIEDALNQLFRNSSIKSYFSDCQVSTFRSV
 PNRHHTGVDSL CNFSPLARRVDRVAIYEEFLRMTRNGTQLQNFTLDRSSVLVDGYSPN
 RNEPLTGNSDLPFWAVILIGLAGLLGLITCLICGVLVTTTRRRKKEGEYNVQQCPGY
 QSHLDLEDLQ

Figure 3B (Continued)

Nucleic acid and polypeptide sequences of IL-18Nucleic acid sequence (SEQ ID No:3)

```

1 attctctccc cagcttgctg agccctttgc tcccctggcg actgcctgga cagtcagcaa
61 ggaattgtct cccagtgcac ttgcccctcc tggctgcaa ctctggctgc taaagcggct
121 gccacctgct gcagtctaca cagcttcggg aagaggaaaag gaacctcaga ccttccagat
181 cgcttcctct cgcaacaaac tatttgctgc aggaataaag atggctgctg aaccagtaga
241 agacaattgc atcaactttg tggcaatgaa atttattgac aatacgcttt actttatagc
301 tgaagatgat gaaaacctgg aatcagatta ctttggcaag ctggaatcta aattatcagt
361 cataagaaat ttgaatgacc aagttctctt cattgaccaa ggaaatcggc ctctatttga
421 agatatgact gattctgact gtagagataa tgcaccccgg accatattta ttataagtat
481 gtataaagat agccagccta gaggtatggc tgtaactatc tctgtgaagt gtgagaaaat
541 ttcaactctc tcctgtgaga acaaaaattat ttctttaag gaaatgaatc ctctgataa
601 catcaaggat acaaaaagtg acatcatatt ctttcagaga agtgtcccag gacatgataa
661 taagatgcaa tttgaatctt catcatacga aggatacttt ctagcttgtg aaaagagag
721 agaccttttt aaactcattt tgaaaaaaga ggatgaattg ggggatagat ctataatggt
781 cactgttcaa aacgaagact agctattaaa atttcatgcc gggcgcagtg gctcacgcct
841 gtaatcccag ccctttggga ggctgaggcg ggcagatcac cagaggtcag gtgttcaaga
901 ccagcctgac caacatggtg aaacctcatc tctactaaaa atacaaaaaa ttagctgagt
961 gtagtgacgc atgccctcaa tcccagctac tcaagaggct gaggcaggag aatcacttgc
1021 actccggagg tagaggttgt ggtgagccga gattgcacca ttgcgctcta gcctgggcaa
1081 caacagcaaa actccatctc aaaaaataaa ataaataaat aaacaataa aaaattcata
1141 atgtg

```

Figure 4APolypeptide sequence (SEQ ID No:4)

```

MAAEPVEDNCINFMAMKFDNTLYFIAEDDENLESDFGKLESKLSVIRNLNDQVLFIDQGNRPLFEDM
TDSDCRDNAPRTIFIISMYSKDSQPRGMAVTISVKCEKISTLSCENKIISFKEMNPPDNIKDTSKDIIF
QRSVPGHDNKMVFESSYEGYFLACEKERDLFKLILKKEDELGDRSIMFTVQNE

```

Figure 4B

Nucleic acid and polypeptide sequences of FGF-2

Nucleic acid sequence (SEQ ID No:5)

```

1  cggccccaga aaaccgagc gagtaggggg cggcgcgag gagggaggag aactgggggc
61  ggggaggct ggtgggtgtc gggggtggag atgtagaaga tgtgacgccc cggcccggcg
121 ggtgccagat tagcggacgc gctgcccgcg gttgcaacgg gatcccgggc gctgcaactt
181 gggaggcggc tctccccagg cggcgtccgc ggagacaccc atccgtgaac cccaggctccc
241 gggccgcccg ctcgcccgcg accaggggcc ggcggacaga agagcggccg agcggctcga
301 ggctggggga ccgcccggcg ggcccgcgcg tgcccggcgg gaggctgggg ggcccggggc
361 ggggcccgtc cccggagcgg gtcggaggcc ggggcccggg ccgggggacg gcggtcccc
421 gcgcccgtcc agcggctcgg ggatcccggc cgggcccgcg agggaccatg gcagcccggg
481 gcatcaccac gctgcccgcg ttgcccgagc atggcggcag cggcgccttc ccgcccggcc
541 acttcaagga cccaagcgg ctgtactgca aaaacggggg cttcttctcg cgcattccac
601 ccgacggccc agttgacggg gtcccggaga agagcgcacc tcacatcaag ctacaacttc
661 aagcagaaga gagaggagt gtgtctatca aaggagtgtg tgctaaccgt tacctggcta
721 tgaaggaaga tggaagatta ctggcttcta aatgtgttac ggatgagtgt ttctttttg
781 aacgattgga atctaataac tacaatactt accggtcaag gaaatacacc agttggtatg
841 tggcactgaa acgaactggg cagtataaac ttggatccaa aacaggacct gggcagaag
901 ctatactttt tcttccaatg tctgctaaga gctgatttta atggccacat ctaattctcat
961 ttcacatgaa agaagaagta tatttttaga atttgttaat gagagtaaaa gaaaataaat
1021 gtgtatagct cagtttggat aattggtcaa acaatttttt atccagtagt aaaatatgta
1081 accattgtcc cagtaaagaa aaataacaaa agttgtaaaa tgtatatctc cccttttata
1141 ttgcatctgc tgttaccagg tgaagcttac ctagagcaat gatctttttc acgcatttgc
1201 tttatttcgaa aagaggcttt taaaatgtgc atgttttaga acaaaaattc ttcatggaaa
1261 tcatatacat tagaaaatca cagtcagatg tttaatcaat ccaaaaatgc cactatttct
1321 tatgtcattc gttagtctac atgtttctaa acatataaat gtgaatttaa tcaattcctt
1381 tcatagtttt ataattctct ggcagttcct tatgatagag tttataaaac agtccctgtg
1441 aaactgctgg aagttcttcc acagtcaggt caattttgtc aaacccttct ctgtaccocat
1501 acagcagcag cctagcaact ctgctggtga tgggagttgt attttcagtc ttcgccaggt
1561 cattgagatc catccactca catcttaagc attcttctctg gcaaaaattt atgggtgaatg
1621 aatattggct taggcggcag atgatataca tatctgactt cccaaaagct ccaggatttg
1681 tgtgctggtg ccgaatactc aggacggacc tgaattctga ttttatacca gtctcttcaa
1741 aaacttctcg aaccgctgtg tctcctacgt aaaaaaagag atgtacaaat caataataat
1801 tacactttta gaaactgtat catcaaagat tttcagttaa agtagcatta tgtaaaggct
1861 caaaacatta ccctaacaaa gtaaagtttt caatacaaat tctttgcctt gtggatatca
1921 agaaatccca aaatatattc ttaccactgt aaattcaaga agcttttgaa atgctgaata
1981 tttctttggc tgctacttgg aggcttatct acctgtacat ttttggggct agctcttttt
2041 aacttcttgc tgctcttttt cccaaaaggt aaaaatatag attgaaaagt taaaacattt
2101 tgcattgctg cagttccttt gtttcttgag ataagattcc aaagaactta gattcatttc
2161 ttcaacaccg aaatgctgga ggtgtttgat cagttttcaa gaaacttgga atataataaa
2221 ttttataatt caacaaaggt tttcacattt tataaggttg atttttcaat taaatgcaaa
2281 tttgtgtggc aggattttta ttgccattaa catatttttg tggtgctttt ttctacacat
2341 ccagatggtc cctctaactg ggctttctct aattttgtga tgttctgtca ttgtctccca
2401 aagtatttag gagaagccct ttaaaaagct gccttctctc accactttgc tggaaagctt
2461 cacaattgtc acagacaaaag atttttgttc caatactcgt ttgaccteta ttttcttgt
2521 ttgtcaaata gtaaatgata tttgcccctg cagtaattct actggtgaaa aacatgcaaa
2581 gaagaggaag tcacagaaac atgtctcaat tcccatgtgc tgtgactgta gactgtctta
2641 ccatagactg tcttaccatc ccctggata tgctcttgtt ttttccctct aatagctatg
2701 gaaagatgca tagaaagagt ataatgtttt aaaacataag gcattcatct gccatttttc
2761 aattacatgc tgacttccct tacaattgag atttgcccac aggttaaaca tggttagaaa
2821 caactgaaag cataaaagaa aaatctaggc cgggtgcagt ggctcatgcc tatattccct
2881 gcactttggg aggccaaaagc agggagatcg cttgagccca ggagttcaag accaacctgg
2941 tgaacccccg tctctacaaa aaaacacaaa aaatagccag gcatggtggc gtgtacatgt

```

Figure 5A

3001 ggtctcagat acttgggagv ctgaggtggv agggttgacv acttgaggct gagaggtcaa
 3061 ggttgcagtv agccataatc gtgccactgc agtccagcct aggcaacaga gtgagacttt
 3121 gtctcaaaaa aagagaaatt ttccttaata agaaaagtaa tttttactct gatgtgcaat
 3181 acatthgtta ttaaattttat tatttaagat ggtagcacta gtcttaaaat gtataaaaata
 3241 tcccctaaca tgtttaaatg tccattttta ttcattatgc tttgaaaaat aattatgggg
 3301 aaatacatgt ttgttattaa atttattatt aaagatagta gcactagtct taaatttgat
 3361 ataacatctc ctaacttggt taaatgtcca tttttattct ttatgcttga aaataaatta
 3421 tggggatcct atttagctct tagtaccact aatcaaaaagt tccggcatgta gctcatgatc
 3481 tatgctgtht ctatgtcgtg gaagcaccgg atgggggtag tgagcaaatc tggcctgctc
 3541 agcagtcacc atagcagctg actgaaaatc agcactgcct gagtagthtt gatcagthta
 3601 acttgaatca ctaactgact gaaaattgaa tgggcaaaata agtgctthtt tctccagagt
 3661 atgcgggaga cccttccacc tcaagatgga tatttcttcc ccaaggattt caagatgaa
 3721 tgaattthtt aatcaagata gtgtgcttta ttctgttga tttttatta ttttaataata
 3781 ctgtaagcca aactgaaata acatthgtct tttttataggt ttgaagaaca taggaaaaac
 3841 taagaggtth ttgtthttatt tttgtctgat aagagatag tttaaatag ttgtattgth
 3901 ttgtthtagtt acaggacaat aatgaaatgg agthttatatt tghtatttct atthttgtht
 3961 atthtaataat agaattagat tgaataaaaa tataatggga aataatctgc agaatgtggg
 4021 tttcctggth tttcctctga ctctagtgca ctgatgatct ctgataaggc tcagctgctt
 4081 tatagthctc tggctaathc agcagatact ctctcctcca tgggtaatac gattthttta
 4141 gaaggcagth tgtcaathtt aatctthgtg ataccttht actcttaggg tattattthta
 4201 tacaaaagcc ttgaggattg cattctattt tctatagac cctcttgata tttaaaaaac
 4261 actatggata acaattcttc atthacctag tattatgaaa gaatgaagga gttcaaacaa
 4321 atgtgthttc cagthtaacta gggthtactg tttgagccaa tataaatgth taactgthttg
 4381 tgatggcagth atctctaaag tacattgcat gthttcctaa atacagagth taaataatth
 4441 cagthaattct tagatgathc agcttcatca ttaagaatat cthttgthtt atgttgagth
 4501 agaathgcct tcatatagac atagthcttc agacctctac tgtcagthtt cattcttagc
 4561 tgctthtcagg gthttatgaa tthtcaggca aagctthta tttatactaa cttaggaagt
 4621 atggctaath ccaacggcag tthttthctt cttaattcca catgactgag gcataatgta
 4681 tctctgggta ggtgagthgt tgtgacaacc acaagcactt tthttthttt taagaaaaaa
 4741 aaggtagthga atthtttaatc atctggactt taagaaggat tctggagtht acttaggctc
 4801 gaaathatat ataththgct tggaaatgth tthttcttca attacatcta caagthaagta
 4861 cagctgaaat tcagaggacc cataagagth cacatgaaaa aatcaattc atthgaaag
 4921 gcaagatgca ggagagagga agccttgcaa acctgcagac tgctthttgc ccaatagata
 4981 ttgggthaagg ctgcaaaaaca taagctthaa tagctcacat gctctgctct cacgtggcac
 5041 cagthggatag tgtgagagaa ttaggctgta gaacaaatgg ccttctctth cagcattcac
 5101 accactacaa aatcatctth tatatcaaca gaagaataag cataaactaa gcaaaaggthc
 5161 aataagthacc tgaaccaag atthggctaga gatathctt aatgcaatcc atthtctgat
 5221 ggatthgtac gagthggcta tataatgtht gthtggtatt ttgattthgt taaagthttt
 5281 aaaaatcaag cthtaagthc atggacatth ttaataaaaa tattthaaga caathtagaa
 5341 aatthcctta atatcattgt tggcthaata gaatagggga catgcatatt aagghaaagg
 5401 tcatggagaa ataaththgg tatcaaacaa atacattgat ttgtcatgat acacattgaa
 5461 tthgatccaa tagththaagg aatagthtag aaaaththgg thctatthtt cgattthctg
 5521 taaatcagth acataaataa thcttagctt atthttatatt thctthgtct aaatactgag
 5581 ctcagthaagt tgtthtaggg gathatthct cagthtgagac thtcttatat gacaththac
 5641 tatgthttga thtctgact atthaaaaata aatagthaga acaaththca taaagthgaag
 5701 aatthataaa tcactgctth ataactgact thattatatt taththcaaag thcaththaaa
 5761 ggctactatt catctctgt gatggaatgg tcaggaatth gthttctcat agththaattc
 5821 caacaacaat atthagthga thcaaaaata cththaatgc taaactthac tgatgthtat
 5881 ccaaagctth thctthtcag acagathaat ccagaagcag thcataacag aagaathaggt
 5941 ggtatgthcc taatgatath atthtacta atggaataaa ctgthaatth agaaathatg
 6001 ctgthaatha tatcagctct gagghaath ctgaaatgth cagactcagth cghaacaath
 6061 tggaaaath aaaththtt thttagctat aaagcaagaa agthaaacaa thaatthctc
 6121 caacaththtt aagccaatha aaaaataaaa agatacacac caathctthc thcaggctct
 6181 gacaggctth ctgghaaact ccacathatt thcaactgca gthataagthc agaaaaataa
 6241 gthaacataa ththcactaa cacacacata tgtagathth caaaaatcca cthataathg
 6301 gthcaagthgg thgagaath atthththtag aathgcatgc aaaaaththc thagctthcat
 6361 cththctccc thgtthctth ththththgg ggagctggta actgatgaaa ththththcca
 6421 cththctct thcagghaata thaagthgth thththgth acgtgataca thctgthga
 6481 atgaaacath ggagghaaac atctactgaa thctgthaat ththththth thctgctgag
 6541 thaaactatga acagatagaa gaaththaca gatgctgcta thaaathagth gaaaaataa
 6601 atthcatcac thaaathatg ththththaaa thctatthct atathgthatt thtaathcaga
 6661 tghatthct thathththc ththththgt gthththgath ththththth atghaaththc
 6721 ththcatgag thgthgath aaaaathgath agthththgt thctthgtct cghaaaaaaa
 6781 aaaaaaaaa aaaaaaaaa aaa

Figure 5A (continued)

Polypeptide sequence (SEQ ID No:5)

MVGVGGGDVEDVTPRPGGCQISGRAARGCNGIPGAAWEAALPRRRPRRHPSVNP
AAGSPRTRGRRTEERPSGSRLGDRGRGRALPGGRLGGRGRGRAPERVGGGRGRGTAA
PRAAPAARGSRPGPAGTMAAGSITTLPALPEDGGSGAFPPGHFKDPKRLYCKNGGFFL
RIHPDGRVDGVREKSDPHIKLQLQAEERGVSISIKGVCANRYLAMKEDGRL LASKCVTD
ECFFFERLESNNYNTYRSRKYTSWYVALKRTGQYKLGSKTGPGQKAILFLPMSAKS

Figure 5B

专利名称(译)	诊断卵巢癌的方法及其试剂盒		
公开(公告)号	EP1924710A4	公开(公告)日	2008-10-29
申请号	EP2006790705	申请日	2006-09-15
[标]申请(专利权)人(译)	VAL CHUM发EQ 麦吉尔大学		
申请(专利权)人(译)	VAL-CHUM , S.E.C. 麦吉尔大学		
当前申请(专利权)人(译)	VAL-CHUM , S.E.C. 麦吉尔大学		
[标]发明人	LE PAGE CECILE MES MASSON ANNE MARIE PROVENCHER DIANE TONIN PATRICIA HUDSON THOMAS		
发明人	LE PAGE, CÉCILE MES-MASSON, ANNE-MARIE PROVENCHER, DIANE TONIN, PATRICIA HUDSON, THOMAS		
IPC分类号	C12Q1/68 C12Q1/04 G01N33/574 G01N33/53		
CPC分类号	C12Q1/6886 C12Q2600/112 C12Q2600/136 C12Q2600/158 G01N33/57449 G01N2500/10		
优先权	60/716941 2005-09-15 US		
其他公开文献	EP1924710B1 EP1924710A2		
外部链接	Espacenet		

摘要(译)

一种诊断卵巢癌的方法，包括提供来自受试者（受试者样品）的生物样品，并检测受试者样品中的每种标志物FGF-2，CA125和IL-18的表达水平。检测来自受试者（受试者样品）的生物样品中的每种标志物CA 125，FGF-2和IL-18的表达水平，以及在本发明的方法中使用所述标志物的说明书。