

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2005271961 B8**

- (54) Title
Identification of markers in esophageal cancer, colon cancer, head and neck cancer and melanoma
- (51) International Patent Classification(s)
C12Q 1/68 (2006.01) **C07H 21/04** (2006.01)
C07H 21/02 (2006.01) **C12P 19/34** (2006.01)
- (21) Application No: **2005271961** (22) Date of Filing: **2005.07.08**
- (87) WIPO No: **WO06/017151**
- (30) Priority Data
- | | | |
|-------------------|-------------------|--------------|
| (31) Number | (32) Date | (33) Country |
| 60/586,599 | 2004.07.09 | US |
| 60/587,019 | 2004.07.09 | US |
- (43) Publication Date: **2006.02.16**
(44) Accepted Journal Date: **2011.10.13**
(48) Corrigenda Journal Date: **2011.11.24**
- (71) Applicant(s)
University of Pittsburgh - of the Commonwealth System of Higher Education
- (72) Inventor(s)
Hughes, Steven J.;Raja, Siva;Gooding, William E.;Xi, Liqiang;Godfrey, Tony E.
- (74) Agent / Attorney
Davies Collison Cave, 1 Nicholson Street, Melbourne, VIC, 3000
- (56) Related Art
MATSUDA, J.-I. et al, Annals of Surgical Oncology, March 2004, vol. 11, no. 3, pages 250S-254S
RAJA S. et al, General Thoracic Surgery, 2002, Vol. 123, pages 475-483
GODFREY, T.E. et al, Clinical Cancer Research, 2001, Vol. 7, pages 4041-4048

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
16 February 2006 (16.02.2006)

PCT

(10) International Publication Number
WO 2006/017151 A2(51) International Patent Classification:
C12Q 1/68 (2006.01)(21) International Application Number:
PCT/US2005/024194

(22) International Filing Date: 8 July 2005 (08.07.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/586,599 9 July 2004 (09.07.2004) US
60/587,019 9 July 2004 (09.07.2004) US(71) Applicant (for all designated States except US): **UNIVERSITY OF PITTSBURGH - OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION** [US/US]; 200 Gardner Steel Conference Center, Thackeray & O'Hara Streets, Pittsburgh, PA 15260 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **GODFREY, Tony, E.** [GB/US]; 84 Summit Avenue, Bronxville, NY 10708 (US). **XI, Liqiang** [CN/US]; 111 Parker Road South,Plainsboro, NJ 08536 (US). **RAJA, Siva** [US/US]; 304 Lamartine Street, Apt 2, Jamaica Plain, MA 02130 (US). **HUGHES, Steven, J.** [US/US]; 207 Highland Road, Blawnox, PA 15238 (US). **GOODING, William, E.** [US/US]; 2732 Louisiana Avenue, Pittsburgh, PA 15216 (US).(74) Agents: **HIRSHMAN, Jesse, A.** et al.; 2611 Beechwood Boulevard, Pittsburgh, PA 15217 (US).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH,

[Continued on next page])

(54) Title: IDENTIFICATION OF MAKERS IN ESOPHAGEAL CANCER, COLON CANCER, HEAD AND NECK CANCER AND MELANOMA

1 ctctctgtcccc acctctgtctt cctctaggaa cacaggagtt ccagatcaca tcgagttcac
61 catgaattca ctcatgtgaag ccaacaccaa gtctcatgttc gacctgttcc aacagttcac
121 aaaatcaaaa gagaacaaca tcttctattc ccctatcagc atcacatcag cattagggat
181 ggctctctta ggagccaaag acaacactgc acaacagatt aagaaggttc ttcactttga
241 tcaagtcaca gagaacacca caggaaaagc tgcaacatat catgttgata ggctaggaaa
301 tgttcatcac cagtttcaaa agcttctgac tgaattcaac aaatccactg atgcatatga
361 gctgaagatc gccacaagc tcttcggaga aaaaacgtat ctatttttac aggaatatatt
421 agatgccatc aagaaatttt accagaccag tgtggaatct gttgattttg caaatgtctc
481 agaagaaagt cgaagaaga ttaactcctg ggtggaaagt caaacgaatg aaaaaattaa
541 aaaccttaatt cctgaaggta atattggcag caataccaca ttggttcttg tgaacgcaat
601 ctatttcaaa gggcagtgagg agaagaaatt taataaagaa gatactaaag aggaaaaatt
661 ttggccaaac aagaatacat acaagtccat acagatgatg aggcataaca catcttttca
721 ttttgctctg ctggaggatg tacaggccaa ggtcctggaa ataccataca aaggcaaaga
781 tctaagcatg attgtgttgc tgccaaatga aatcgaatgt ctccagaagc ttgaagagaa
841 actcactgct gagaaattga tggaaatggac aagtttgacg aatatgagag agacacgtgt
901 cgattttcac ttacctcggt tcaaatgtga agagagctat gacctcaagg acacgttgag
961 aacctatgga atggtggata tcttcaatgg ggtgcagac ctctcaggca tgaccgggag
1021 ccgcggtctc gtgctatctg gactcctaca caaggccttt gtggagggtta cagaggaggg
1081 agcagaagct gcagctgcca ccgctgtagt aggtatcgga tcatcaccta ctccaactaa
1141 tgaagagttc cattgtaatc accttttctt attcttcata aggcataata agaccaacag
1201 catctctctc tatggcagat tctcctcccc gtagatgcaa ttagtctgtc actccatttg
1261 gaaaatgttc acctgcagat gttctgttaa actgattgct ggcaacaaca gattctcttg
1321 gctcatattt cttttctttc tcatcttgat gatgatcgtc atcatcaaga atttaatatg
1381 taaaatagca tgccctttct tctttctctt aataagccca catataaatg tactttttct
1441 tccagaaaaa ttctccttga ggaaaaatgt ccaaaaatag atgaatcact taataccgta
1501 tcttctaaat ttgaaatata attctgtttg tgacctgttt taaatgaacc aaaccaatc
1561 atactttttc ttgaaattta gcaacctaga aacacacatt tctttgaatt taggtgatgc
1621 ctaaatcctt cttatgtttc taaattttgt gattctataa aacacatcat caataaaaaa
1681 gtgacataaa atca

(57) Abstract: Methods for identifying expression of markers indicative of the presence of esophageal, a squamous cell cancer, a squamous cell cancer of the head and neck, colon cancer and melanoma are provided. Also provided are articles of manufacture useful in such methods and compositions containing primers and probes useful in such methods.



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.

**IDENTIFICATION OF MARKERS IN ESOPHAGEAL CANCER, COLON CANCER,
HEAD AND NECK CANCER, AND MELANOMA****INVENTORS**

Tony Godfrey
Liqiang Xi
William E. Gooding
Steven Hughes
Siva Raja

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. §119(e) to priority United States Provisional Patent Application Nos. 60/586,599 and 60/587,019, both filed on July 9, 2004, each of which is incorporated herein by reference in its entirety.

BACKGROUND**1. Field of the Invention**

Provided are improved cancer diagnostic methods, along with compositions and apparatus useful in conducting those methods.

2. Description of the Related Art

Early detection of cancer typically leads to increased survival rates. Metastatic lesions commonly are detected by histological techniques, including immunohistochemical techniques. Metastasized cells typically infiltrate the lymph nodes, and, thus in most instances, certain sentinel lymph nodes, lymph nodes where metastasized cells typically first infiltrate, are recognized for each cancer type and are analyzed for the presence of lesions, including micrometastases. Trained histologists often can detect metastatic lesions visually after tissue from a sentinel lymph node is sectioned and stained. Highly trained histologists often can visualize micrometastases, but the ability to visualize such lesions varies from histologist-to-histologist.

In many surgical procedures to remove tumors, biopsies of sentinel lymph nodes are taken. The surgical procedure is then halted and the excised lymphatic tissue is then analyzed. Once it is determined that the tumor has metastasized, a second, more radical surgical procedure is performed, removing regional lymphatics. A rapid method for identifying tumors is therefore warranted, not only because more assays can be performed in a given time period, thereby increasing laboratory turnaround, but permitting accurate, intraoperative decisions to be made, rather than conducting a second surgical

procedure. It is therefore desirable to identify useful diagnostics for malignancies, especially that permit rapid and/or intraoperative detection of lymphatic micrometastases.

SUMMARY

The present invention relates to a diagnostic method for detecting the presence of cancer cells in a patient by identifying the expression of certain markers indicative of the presence of cancer cell.

In one embodiment, the present invention relates to a method of identifying the expression of markers indicative of the presence of esophageal cancer cells in a lymph node of a patient. The method comprises determining if an mRNA species specific to one or more of CEA, CK7, CK19, CK20, VIL1, TACSTD1, and PVA is overabundant in an RNA sample prepared from the lymph node. The overabundance of the mRNA species is indicative of the presence of displaced cells of the esophagus in the lymph node.

In another embodiment, the present invention relates to a method of identifying the expression of markers indicative of the presence of cells of squamous cell carcinoma of the head and neck in a lymph node of a patient. The method comprises determining if an mRNA species specific to one or more of CEA, CK19, PTHrP, PVA, TACSTD1 and SCCA1.2 (SCCA1 + SCCA2) is overabundant in an RNA sample prepared from the lymph node. The overabundance of the mRNA species is indicative of the presence of displaced cells of a squamous cell carcinoma of the head and neck in the lymph node.

In still another embodiment, the present invention relates to a method for identifying the expression or markers indicative of the presence of cells of a squamous cell carcinoma in a lymph node of a patient. The method comprises determining if an mRNA species specific to PVA is overabundant in an RNA sample prepared from the lymph node. The overabundance of the mRNA species is indicative of the presence of displaced cells of a squamous cell carcinoma in the lymph node.

In yet another embodiment, the present invention relates to a method for identifying the expression of markers indicative of the presence of colon cancer cells in a lymph node of a patient. The method comprises determining if an mRNA species specific to one or more of CDX1, TACSTD1 and VIL1 is overabundant in an RNA sample prepared from the lymph node. The overabundance of the mRNA species is indicative of the presence of displaced colon cells in the lymph node.

In still another embodiment, the present invention relates to a method for identifying the expression of markers indicative of the presence of melanoma cells in a lymph node of a patient. The method comprises determining if an mRNA species specific to one or more of MAGEA136-plex, MART1, and TYR is overabundant in an RNA sample prepared from the lymph node. The overabundance of the mRNA species is indicative of the presence of melanoma cells in the lymph node.

In yet a further embodiment, the present invention relates to an article of manufacture comprising packaging material and one or more nucleic acids specific to one or more of CEA, CK7, CK19, CK20, VIL1, TACSTD1, and PVA. The packaging material comprises an indicia, for example and without

limitation, a writing, illustration, label, tag, book, booklet and/or package insert, indicating that the one or more nucleic acids can be used in a method of identifying expression of markers indicative of the presence of esophageal cancer cells in a lymph node of a patient.

In a still further embodiment, the present invention relates to an article of manufacture comprising packaging material and one or more nucleic acids specific to one or more of CEA, CK19, PTHrP, PVA, TACSTD1 and SCCA1.2. The packaging material comprises an indicia indicating that the one or more nucleic acids can be used in a method of identifying expression of markers indicative of the presence of cells of a squamous cell carcinoma of the head and neck in a lymph node of a patient.

In another embodiment, the present invention relates to an article of manufacture comprising packaging material and one or more nucleic acids specific to one or more of CDX1, TACSTD1 and VIL1. The packaging material comprises an indicia indicating that the one or more nucleic acids can be used in a method of identifying expression of markers indicative of the presence of colon cancer cells in a lymph node of a patient.

In still another embodiment, the present invention relates to an article of manufacture comprising packaging material and one or more nucleic acids specific to one or more of MAGEA136-plex, MART1 and TYR. The packaging material comprises an indicia indicating that the one or more nucleic acids can be used in a method of identifying expression of markers indicative of the presence of melanoma cells in a lymph node of a patient.

In still another embodiment, the present invention relates to an article of manufacture comprising packaging material and one or more nucleic acids specific to PVA. The packaging material comprises an indicia indicating that the one or more nucleic acids can be used in a method of identifying expression of markers indicative of the presence of cells of a squamous cell carcinoma in a lymph node of a patient.

In yet another embodiment, the present invention relates to a composition comprising one or more primers or probes specific to one or more of CEA, CK7, CK19, CK20, VIL1, TACSTD1, and PVA and RNA extracted from the lymph node of a patient diagnosed with or suspected of having esophageal cancer, or a nucleic acid, or analog thereof, derived from the RNA.

In a further embodiment, the present invention relates to a composition comprising one or more primers or probes specific to one or more of CEA, CK19, PTHrP, PVA, TACSTD1 and SCCA1.2 and RNA extracted from the lymph node of a patient diagnosed with or suspected of having squamous cell carcinoma of the head and neck, or a nucleic acid, or analog thereof, derived from the RNA.

In still a further embodiment, the present invention relates to a composition comprising one or more primers or probes specific to one or more of CDX1, TACSTD1 and VIL1 and RNA extracted from the lymph node of a patient diagnosed with or suspected of having colon cancer, or a nucleic acid, or an analog thereof, derived from the RNA.

In yet a further embodiment, the present invention relates to a composition comprising one or more primers or probes specific to one or more of MAGEA136-plex, MART1 and TYR and RNA extracted from a lymph node of a patient diagnosed with or suspected of having melanoma, or a nucleic acid, or analog thereof, derived from the RNA.

In another embodiment, the present invention relates to a composition comprising one or more primers or probes specific to PVA and RNA extracted from a sentinel lymph node of a patient diagnosed with or suspected of having a squamous cell carcinoma, or a nucleic acid, or analog thereof, derived from the RNA.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a listing of a cDNA sequence of the caudal-type homeo box transcription factor 1 (CDX1) marker (SEQ ID NO: 1).

Figure 2 is a listing of a cDNA sequence for the carcinoembryonic antigen-related cell adhesion molecule 5 (CEA) marker (SEQ ID NO: 2).

Figure 3 is a listing of a cDNA sequence for the cytokeratin 7 (CK7) marker (SEQ ID NO: 3).

Figure 4 is a listing of a cDNA sequence for the cytokeratin 19 (CK19) marker (SEQ ID NO: 4).

Figure 5 is a listing of a cDNA sequence for the cytokeratin 20 (CK20) marker (SEQ ID NO: 5).

Figure 6 is a listing of a cDNA sequence for the melanoma antigen gene family A1 (MAGEA1) marker (SEQ ID NO: 6).

Figure 7 is a listing of a cDNA sequence for the melanoma antigen gene family A3 (MAGEA3) marker (SEQ ID NO: 7).

Figure 8 is a listing of a cDNA sequence for the melanoma antigen gene family A6 (MAGEA6) marker (SEQ ID NO: 8).

Figure 9 is a listing of a cDNA sequence for the melanoma antigen recognized by T cells 1 (MART1) marker (SEQ ID NO: 9).

Figure 10 is a listing of a cDNA sequence for the parathyroid hormone-related protein (PTHrP) marker (SEQ ID NO: 10).

Figure 11 is a listing of a cDNA sequence for the pemphig vulgaris antigen (PVA) marker (SEQ ID NO: 11).

Figure 12 is a listing of a cDNA sequence for the squamous cell carcinoma antigen 1 (SCCA1) marker (SEQ ID NO: 12).

Figure 13 is a listing of a cDNA sequence for the squamous cell carcinoma antigen 2 (SCCA2) marker (SEQ ID NO: 13).

Figure 14 is a listing of a cDNA sequence for the tumor-associated calcium signal transducer 1 (TACSTD1) marker (SEQ ID NO: 14).

Figure 15 is a listing of a cDNA sequence for the tyrosinase (TYR) marker (SEQ ID NO: 15).

Figure 16 is a listing of a cDNA sequence for the villin 1 (VIL1) marker (SEQ ID NO: 16).

Figure 17 is a scatter plot showing the expression levels of CEA, CK7, SCCA 1.2, CK20, TACSTD1, VIL and CK19 in primary tumor, tumor-positive lymph nodes and benign lymph nodes of an esophageal cancer patient.

Figure 18A-O provide scatter plots illustrating the ability of two-marker systems to distinguish between benign and malignant cells in a lymph node of an esophageal cancer patient (negative – gray circle; positive – black circle).

Figures 19 is a scatter plot showing the expression levels of CEA, CK19, PThRP, PVA, SCCA1.2 and TACSTD1 in primary tumor, tumor-positive lymph nodes and benign lymph nodes of a head & neck cancer patient.

Figure 20A-F provides scatter plots illustrating the ability of two-marker systems to distinguish between benign and malignant cells in a lymph node of a head & neck cancer patient (negative – circle; positive – “+”).

Figures 21 is a scatter plot showing the expression levels of MART1, TYR and MAGEA136-plex in primary tumor, tumor-positive lymph nodes and benign lymph nodes of a melanoma patient.

Figures 22A and 22B provide scatter plots illustrating the ability of two-marker systems to distinguish between benign and malignant cells in a lymph node of a melanoma patient (negative – circle; positive – “+”).

Figures 23 is a scatter plot showing the expression levels of CDX1, CEA, CK19, CK20, TACSTD1 and VIL1 in primary tumor, tumor-positive lymph nodes and benign lymph nodes of a colon cancer patient.

DETAILED DESCRIPTION

Provided are methods and compositions useful in identifying esophageal cancer, colon cancer, head and neck cancer and melanoma cells, including micrometastases, in lymph nodes. Early detection of metastases typically is related to patient survival. Very small metastases often go undetected in histological study of lymph node biopsies, resulting in false negative results that result in decreased chances of patient survival. The nucleic acid detection assays described herein are much more discriminating than are histological studies in most instances (a few, excellent histologists are capable of identifying micrometastases in lymph node sections), and are robust and repeatable in the hands of any minimally-trained technician. Although the methods and compositions described herein are necessarily presented comprising expression of specific mRNA markers, this should be understood that it shall not be

deemed to exclude methods and compositions comprising combinations of the specific markers and other markers known in the art.

To this end, a number of molecular markers are identified, that are expressed in certain cancer types, including esophageal cancer, colon cancer, head and neck cancer and melanoma. These markers are markers specific to the tissue from which the particular cancer type arises and typically are not expressed, at least to the same levels, in lymphoid tissue. The presence and/or elevated expression of one or more of these markers in sentinel lymph node tissue is indicative of displaced cells in the lymphoid tissue, which correlates strongly with a cancer diagnosis. As used herein a "squamous cell carcinoma" is a cancer arising, at least in part, from a squamous cell population and/or containing, at least in part, a squamous cell population including, without limitation, cancers of the cervix; penis; head and neck, including, without limitation cancers of the oral cavity, salivary glands, paranasal sinuses and nasal cavity, pharynx and larynx; lung; esophageal; skin other than melanoma; vulva and bladder.

As used herein, the terms "expression" and "expressed" mean production of a gene-specific mRNA by a cell. In the context of the present disclosure, a "marker" is a gene that is expressed abnormally in a lymphatic biopsy. In one embodiment, the markers described herein are mRNA species that are expressed in cells of a specific tumor source at a significantly higher level as compared to expression in lymphoid cells.

Expression levels of mRNA can be quantified by a number of methods. Traditional methods include Northern blot analysis. More recently, nucleic acid detection methods have been devised that facilitate quantification of transcripts. Examples of PCR methods are described in United States Patent Application No. 10/090,326 (US 10/090,326), incorporated herein by reference in its entirety. Other methods for determining expression levels of a given mRNA include isothermic amplification or detection assays and array technologies, as are known in the art, such as, without limitation, those described below.

The improved PCR methods described herein as well as in US 10/090,326, and other nucleic acid detection and amplification methods described herein and as are known in the art permit rapid detection of cancer cells in lymph node tissue. These rapid methods can be used intraoperatively, and also are useful in detecting rare nucleic acid species, even in multiplexed PCR reactions that concurrently detect a more prevalent control nucleic acid.

A typical PCR reaction includes multiple amplification steps, or cycles that selectively amplify a target nucleic acid species. Because detection of transcripts is necessary, the PCR reaction is coupled with a reverse transcription step (reverse transcription PCR, or RT-PCR). A typical PCR reaction includes three steps: a denaturing step in which a target nucleic acid is denatured; an annealing step in which a set of PCR primers (forward and backward primers) anneal to complementary DNA strands; and an elongation step in which a thermostable DNA polymerase elongates the primers. By repeating this step multiple times, a DNA fragment is amplified to produce an amplicon, corresponding to the target

DNA sequence. Typical PCR reactions include 30 or more cycles of denaturation, annealing and elongation. In many cases, the annealing and elongation steps can be performed concurrently, that is at the same temperature, in which case the cycle contains only two steps.

The lengths of the denaturation, annealing and elongation stages may be any desirable length of time. However, in attempting to shorten the PCR amplification reaction to a time suitable for intraoperative diagnosis, the lengths of these steps can be in the seconds range, rather than the minutes range. The denaturation step may be conducted for times of one second or less. The annealing and elongation steps optimally are less than 10 seconds each, and when conducted at the same temperature, the combination annealing/elongation step may be less than 10 seconds. Use of recently developed amplification techniques, such as conducting the PCR reaction in a Rayleigh-Bénard convection cell, also can dramatically shorten the PCR reaction time beyond these time limits (see, Krishnan, My et al., "PCR in a Rayleigh-Bénard convection cell." *Science* 298:793 (2002), and Braun, D. et al., "Exponential DNA Replication by Lominar Convection," *Physical Review Letters*, 91:158103).

As described in US 10/090,326, each cycle may be shortened considerably without substantial deterioration of production of amplicons. Use of high concentrations of primers is helpful in shortening the PCR cycle time. High concentrations typically are greater than about 400nM, and often greater than about 800nM, though the optimal concentration of primers will vary somewhat from assay-to-assay. Sensitivity of RT-PCR assays may be enhanced by the use of a sensitive reverse transcriptase enzyme (described below) and/or high concentrations of reverse transcriptase primer to produce the initial target PCR template.

The specificity of any given PCR reaction relies heavily, but not exclusively, on the identity of the primer sets. The primer sets are pairs of forward and reverse oligonucleotide primers that anneal to a target DNA sequence to permit amplification of the target sequence, thereby producing a target sequence-specific amplicon. PCR primer sets can include two primers internal to the target sequence, or one primer internal to the target sequence and one specific to a target sequence that is ligated to the DNA or cDNA target, using a technique known as "ligation-anchored PCR" (Troutt, A.B., et al. (1992), "Ligation-anchored PCR: A Simple Amplification Technique with Single-sided Specificity," *Proc. Natl. Acad. Sci. USA*, 89:9823-9825).

As used herein, a "derivative" of a specified oligonucleotide is an oligonucleotide that binds to the same target sequence as the specified oligonucleotide and amplifies the same target sequence to produce essentially the same amplicon as the specified oligonucleotide but for differences between the specified oligonucleotide and its derivative. The derivative may differ from the specified oligonucleotide by insertion, deletion and/or substitution of any residue of the specified sequence so long as the derivative substantially retains the characteristics of the specified sequence in its use for the same purpose as the specified sequence.

As used herein, “reagents” for any assay or reaction, such as a reverse transcription and PCR, are any compound or composition that is added to the reaction mixture including, without limitation, enzyme(s), nucleotides or analogs thereof, primers and primer sets, probes, antibodies or other binding reagents, detectable labels or tags, buffers, salts and co-factors. As used herein, unless expressed otherwise, a “reaction mixture” for a given assay or reaction includes all necessary compounds and/or compositions necessary to perform that assay or reaction, even if those compounds or compositions are not expressly indicated. Reagents for many common assays or reactions, such as enzymatic reaction, are known in the art and typically are provided and/or suggested when the assay or reaction kit is sold.

As also described in US 10/090,326, multiplexed PCR assays may be optimized, or balanced, by time-shifting the production of amplicons, rather than by manipulating primer concentrations. This may be achieved by using two primer sets, each primer set having a different T_m so that a two-stage PCR assay can be performed, with different annealing and/or elongation temperatures for each stage to favor the production of one amplicon over another. This time and temperature shifting method permits optimal balancing of the multiplex reaction without the difficulties faced when manipulation of primer concentrations is used to balance the reaction. This technique is especially useful in a multiplex reaction where it is desirable to amplify a rare cDNA along with a control cDNA.

A quantitative reverse transcriptase polymerase chain reaction (QRT-PCR) for rapidly and accurately detecting low abundance RNA species in a population of RNA molecules (for example, and without limitation, total RNA or mRNA), includes the steps of: a) incubating an RNA sample with a reverse transcriptase and a high concentration of a target sequence-specific reverse transcriptase primer under conditions suitable to generate cDNA; b) subsequently adding suitable polymerase chain reaction (PCR) reagents to the reverse transcriptase reaction, including a high concentration of a PCR primer set specific to the cDNA and a thermostable DNA polymerase to the reverse transcriptase reaction, and c) cycling the PCR reaction for a desired number of cycles and under suitable conditions to generate PCR product (“amplicons”) specific to the cDNA. By temporally separating the reverse transcriptase and the PCR reactions, and by using reverse transcriptase-optimized and PCR-optimized primers, excellent specificity is obtained. The reaction may be conducted in a single tube (all tubes, containers, vials, cells and the like in which a reaction is performed may be referred to herein, from time to time, generically, as a “reaction vessel”), removing a source of contamination typically found in two-tube reactions. These reaction conditions permit very rapid QRT-PCR reactions, typically on the order of 20 minutes from the beginning of the reverse transcriptase reaction to the end of a 40 cycle PCR reaction.

The reaction c) may be performed in the same tube as the reverse transcriptase reaction by adding sufficient reagents to the reverse transcriptase (RT) reaction to create good, or even optimal conditions for the PCR reaction to proceed. A single tube may be loaded, prior to the running of the reverse transcriptase reaction, with: 1) the reverse transcriptase reaction mixture, and 2) the PCR reaction mixture to be mixed with the cDNA mixture after the reverse transcriptase reaction is completed. The

reverse transcriptase reaction mixture and the PCR reaction mixture may be physically separated by a solid, or semi-solid (including amorphous, glassy substances and waxy) barrier of a composition that melts at a temperature greater than the incubation temperature of the reverse transcriptase reaction, but below the denaturing temperature of the PCR reaction. The barrier composition may be hydrophobic in nature and forms a second phase with the RT and PCR reaction mixtures when in liquid form. One example of such a barrier composition is wax beads, commonly used in PCR reactions, such as the AMPLIWAX PCR GEM products commercially available from Applied Biosystems of Foster City, California.

Alternatively, the separation of the reverse transcriptase and the PCR reactions may be achieved by adding the PCR reagents, including the PCR primer set and thermostable DNA polymerase, after the reverse transcriptase reaction is completed. Preferably the PCR reagents, are added mechanically by a robotic or fluidic means to make sample contamination less likely and to remove human error.

The products of the QRT-PCR process may be compared after a fixed number of PCR cycles to determine the relative quantity of the RNA species as compared to a given reporter gene. One method of comparing the relative quantities of the products of the QRT-PCR process is by gel electrophoresis, for instance, by running the samples on a gel and detecting those samples by one of a number of known methods including, without limitation, Southern blotting and subsequent detection with a labeled probe, staining with ethidium bromide and incorporating fluorescent or radioactive tags in the amplicons.

However, the progress of the quantitative PCR reactions typically is monitored by determining the relative rates of amplicon production for each PCR primer set. Monitoring amplicon production may be achieved by a number of processes, including without limitation, fluorescent primers, fluorogenic probes and fluorescent dyes that bind double-stranded DNA. A common method is the fluorescent 5' nuclease assay. This method exploits the 5' nuclease activity of certain thermostable DNA polymerases (such as *Taq* or *Tfl* DNA polymerases) to cleave an oligomeric probe during the PCR process. The oligomer is selected to anneal to the amplified target sequence under elongation conditions. The probe typically has a fluorescent reporter on its 5' end and a fluorescent quencher of the reporter at the 3' end. So long as the oligomer is intact, the fluorescent signal from the reporter is quenched. However, when the oligomer is digested during the elongation process, the fluorescent reporter no longer is in proximity to the quencher. The relative accumulation of free fluorescent reporter for a given amplicon may be compared to the accumulation of the same amplicons for a control sample and/or to that of a control gene, such as β -actin or 18S rRNA to determine the relative abundance of a given cDNA product of a given RNA in a RNA population. Products and reagents for the fluorescent 5' nuclease assay are readily available commercially, for instance from Applied Biosystems.

Equipment and software also are readily available for monitoring amplicon accumulation in PCR and QRT-PCR according to the fluorescent 5' nuclease assay and other QPCR/QRT-PCR procedures, including the Smart Cycler, commercially available from Cepheid of Sunnyvale, California, the ABI

Prism 7700 Sequence Detection System (TaqMan), commercially available from Applied Biosystems. A cartridge-based sample preparation system (GenXpert) combines a thermal cycler and fluorescent detection device having the capabilities of the Smart Cycler product with fluid circuits and processing elements capable of automatically extracting specific nucleic acids from a tissue sample and performing QPCR or QRT-PCR on the nucleic acid. The system uses disposable cartridges that can be configured and pre-loaded with a broad variety of reagents. Such a system can be configured to disrupt tissue and extract total RNA or mRNA from the sample. The reverse transcriptase reaction components can be added automatically to the RNA and the QPCR reaction components can be added automatically upon completion of the reverse transcriptase reaction.

Further, the PCR reaction may be monitored of production (or loss) of a particular fluorochrome from the reaction. When the fluorochrome levels reach (or fall to) a desired level, the automated system will automatically alter the PCR conditions. In one example, this is particularly useful in the multiplexed embodiment described above, where a more-abundant (control) target species is amplified by the first, lower T_m , primer set at a lower temperature than the less abundant species amplified by the second, higher T_m , primer set. In the first stage of the PCR amplification, the annealing temperature is lower than the effective T_m of the first primer set. The annealing temperature then is automatically raised above the effective T_m of the first primer set when production of the first amplicon by the first primer set is detected. In a system that automatically dispenses multiple reagents from a cartridge, such as the GeneXpert system, a first PCR reaction may be conducted at the first T_m and, when the first PCR reaction proceeds past a threshold level, a second primer with a different T_m is added, resulting in a sequential multiplexed reaction.

In the above-described reactions, the amounts of certain reverse transcriptase and the PCR reaction components typically are atypical in order to take advantage of the faster ramp times of some thermal cyclers. Specifically, the primer concentrations are very high. Typical gene-specific primer concentrations for reverse transcriptase reactions are less than about 20 nM. To achieve a rapid reverse transcriptase reaction on the order of one to two minutes, the reverse transcriptase primer concentration was raised to greater than 20 nM, preferably at least about 50 nM, and typically about 100 nM. Standard PCR primer concentrations range from 100 nM to 300 nM. Higher concentrations may be used in standard PCR reactions to compensate for T_m variations. However, the referenced primer concentrations are for circumstances where no T_m compensation is needed. Proportionately higher concentrations of primers may be empirically determined and used if T_m compensation is necessary or desired. To achieve rapid PCR reactions, the PCR primer concentrations typically are greater than 200 nM, preferably greater than about 500 nM and typically about 800 nM. Typically, the ratio of reverse transcriptase primer to PCR primer is about 1 to 8 or more. The increase in primer concentrations permitted PCR experiments of 40 cycles to be conducted in less than 20 minutes.

A sensitive reverse transcriptase may be preferred in certain circumstances where either low amounts of RNA are present or a target RNA is a low abundance RNA. By the term "sensitive reverse transcriptase," it is meant a reverse transcriptase capable of producing suitable PCR templates from low copy number transcripts for use as PCR templates. The sensitivity of the sensitive reverse transcriptase may derive from the physical nature of the enzyme, or from specific reaction conditions of the reverse transcriptase reaction mixture that produces the enhanced sensitivity. One example of a sensitive reverse transcriptase is SensiScript RT reverse transcriptase, commercially available from Qiagen, Inc. of Valencia, California. This reverse transcriptase is optimized for the production of cDNA from RNA samples of <50ng, but also has the ability to produce PCR templates from low copy number transcripts. In practice, in the assays described herein, adequate results were obtained for samples of up to, and even in excess of, about 400 ng RNA. Other sensitive reverse transcriptases having substantially similar ability to reverse transcribe low copy number transcripts would be equivalent sensitive reverse transcriptase for the purposes described herein. Notwithstanding the above, the ability of the sensitive reverse transcriptase to produce cDNA from low quantities of RNA is secondary to the ability of the enzyme, or enzyme reaction system to produce PCR templates from low copy number sequences.

As discussed above, the procedures described herein also may be used in multiplex QRT-PCR processes. In its broadest sense, a multiplex PCR process involves production of two or more amplicons in the same reaction vessel. Multiplex amplicons may be analyzed by gel electrophoresis and detection of the amplicons by one of a variety of methods, such as, without limitation ethidium bromide staining, Southern blotting and hybridization to probes, or by incorporating fluorescent or radioactive moieties into the amplicons and subsequently viewing the product on a gel. However, real-time monitoring of the production of two or more amplicons is preferred. The fluorescent 5' nuclease assay is the most common monitoring method. Equipment is now available (for example, the above-described Smart Cycler and TaqMan products) that permits the real-time monitoring of accumulation of two or more fluorescent reporters in the same tube. For multiplex monitoring of the fluorescent 5' nuclease assay, oligomers are provided corresponding to each amplicon species to be detected. The oligomer probe for each amplicon species has a fluorescent reporter with a different peak emission wavelength than the oligomer probe(s) for each other amplicons species. The accumulation of each unquenched fluorescent reporter can be monitored to determine the relative amounts of the target sequence corresponding to each amplicon.

In traditional multiplex QPCR and QRT-PCR procedures, the selection of PCR primer sets having similar annealing and elongation kinetics and similar sized amplicons are desirable. The design and selection of appropriate PCR primer sets is a process that is well known to a person skilled in the art. The process for identifying optimal PCR primer sets, and respective ratios thereof to achieve a balanced multiplex reaction also is known. By "balanced," it is meant that certain amplicon(s) do not out-compete the other amplicon(s) for resources, such as dNTPs or enzyme. For instance, by limiting the abundance of the PCR primers for the more abundant RNA species in an RT-PCR experiment will allow the detection of less abundant species. Equalization of the T_m (melting temperature) for all PCR primer sets

also is encouraged. See, for instance, ABI PRISM 7700 Sequence Detection System User Bulletin #5, "Multiplex PCR with TaqMan VIC Probes", Applied Biosystems (1998/2001).

Despite the above, for very low copy number transcripts, it is difficult to design accurate multiplex PCR experiments, even by limiting the PCR primer sets for the more abundant control species. One solution to this problem is to run the PCR reaction for the low abundance RNA in a separate tube than the PCR reaction for the more abundant species. However, that strategy does not take advantage of the benefits of running a multiplex PCR experiment. A two-tube process has several drawbacks, including cost, the addition of more room for experimental error and the increased chance of sample contamination, which is critical in PCR assays.

A method has been described in WO 02/070751 for performing a multiplex PCR process, including QRT-PCR and QPCR, capable of detecting low copy number nucleic acid species along with one or more higher copy number species. The difference between low copy number and high copy number nucleic acid species is relative, but is referred to herein as a difference in the prevalence of a low (lower) copy number species and a high (higher) copy number species of at least about 30-fold, but more typically at least about 100-fold. For purposes herein, the relative prevalence of two nucleic acid species to be amplified is more salient than the relative prevalence of the two nucleic acid species in relation to other nucleic acid species in a given nucleic acid sample because other nucleic acid species in the nucleic acid sample do not directly compete with the species to be amplified for PCR resources.

As used herein, the prevalence of any given nucleic acid species in a given nucleic acid sample, prior to testing, is unknown. Thus, the "expected" number of copies of a given nucleic acid species in a nucleic acid sample often is used herein and is based on historical data on the prevalence of that species in nucleic acid samples. For any given pair of nucleic acid species, one would expect, based on previous determinations of the relative prevalence of the two species in a sample, the prevalence of each species to fall within a range. By determining these ranges one would determine the difference in the expected number of target sequences for each species. An mRNA species is identified as "overabundant" if it is present in statistically significant amounts over normal prevalence of the mRNA species in a sample from a normal patient or lymph node. As is abundantly illustrated in the examples and plots provided herein, a person of skill in the art would be able to ascertain statistically significant ranges or cutoffs for determining the precise definition of "overabundance" for any one or more mRNA species.

The multiplex method involves performing a two- (or more) stage PCR amplification, permitting modulation of the relative rate of production of a first amplicon by a first primer set and a second amplicon by a second primer set during the respective amplification stages. By this method, PCR amplifications to produce amplicons directed to a lower abundance nucleic acid species are effectively "balanced" with PCR amplifications to produce amplicons directed to a higher abundance nucleic acid species. Separating the reaction into two or more temporal stages may be achieved by omitting the PCR primer set for any amplicons that are not to be produced in the first amplification stage. This is best

achieved through use of automated processes, such as the GenXpert prototype system described above. Two or more separate amplification stages may be used to tailor and balance multiplex assays, along with, or to the exclusion of tailoring the concentration of the respective primer sets.

A second method for temporally separating the PCR amplification process into two or more stages is to select PCR primer sets with variation in their respective T_m . In one example, primers for a lower copy number nucleic acid species would have a higher T_m (T_{m1}) than primers for a higher abundance species (T_{m2}). In this process, the first stage of PCR amplification is conducted for a predetermined number of cycles at a temperature sufficiently higher than T_{m2} so that there is substantially no amplification of the higher abundance species. After the first stage of amplification, the annealing and elongation steps of the PCR reaction are conducted at a lower temperature, typically about T_{m2} , so that both the lower abundance and the higher abundance amplimers are amplified. It should be noted that T_m , as used herein and unless otherwise noted, refers to "effective T_m ," which is the T_m for any given primer in a given reaction mix, which depends on factors, including, without limitation, the nucleic acid sequence of the primer and the primer concentration in the reaction mixture.

It should be noted that PCR amplification is a dynamic process. When using temperature to modulate the respective PCR reactions in a multiplex PCR reaction, the higher temperature annealing stage may be carried out at any temperature typically ranging from just above the lower T_m to just below the higher T_m , so long as the reaction favors production of the amplicon by the higher T_m primer set. Similarly, the annealing for the lower temperature reaction typically is at any temperature below the T_m of the low temperature primer set.

In the example provided above, in the higher temperature stage the amplicon for the low abundance RNA is amplified at a rate faster than that the amplicon for the higher abundance RNA (and preferably to the substantial exclusion of production of the second amplicon), so that, prior to the second amplification stage, where it is desirable that amplification of all amplicons proceeds in a substantially balanced manner, the amplicon for the lower abundance RNA is of sufficient abundance that the amplification of the higher abundance RNA does not interfere with the amplification of the amplicon for the lower abundance RNA. In the first stage of amplification, when the amplicon for the low abundance nucleic acid is preferentially amplified, the annealing and elongation steps may be performed above T_{m1} to gain specificity over efficiency (during the second stage of the amplification, since there is a relatively large number of low abundance nucleic acid amplicons, selectivity no longer is a significant issue, and efficiency of amplicon production is preferred). It, therefore, should be noted that although favorable in many instances, the temperature variations may not necessarily result in the complete shutdown of one amplification reaction over another.

In another variation of the above-described amplification reaction, a first primer set with a first T_m may target a more-abundant template sequence (for instance, the control template sequence) and a second primer set with a higher T_m may target a less-abundant template sequence. In this case, the

more-abundant template and the less-abundant template may both be amplified in a first stage at a temperature below the (lower) T_m of the first primer set. When a threshold amount of amplicon corresponding to the more abundant template is reached, the annealing and/or elongation temperature of the reaction is raised above the T_m of the first primer set, but below the higher T_m of the second primer set to effectively shut down amplification of the more abundant template.

Selection of three or more sets of PCR primer sets having three or more different T_m s (for instance, $T_{m1} > T_{m2} > T_{m3}$) can be used to amplify sequences of varying abundance in a stepwise manner, so long as the differences in the T_m s are sufficiently large to permit preferential amplification of desired sequences to the substantial exclusion of undesired sequences for a desired number of cycles. In that process, the lowest abundance sequences are amplified in a first stage for a predetermined number of cycles. Next, the lowest abundance and the lesser abundance sequences are amplified in a second stage for a predetermined number of cycles. Lastly, all sequences are amplified in a third stage. As with the two-stage reaction described above, the minimum temperature for each stage may vary, depending on the relative efficiencies of each single amplification reaction of the multiplex reaction. It should be recognized that two or more amplimers may have substantially the same T_m , to permit amplification of more than one species of similar abundance at any stage of the amplification process. As with the two-stage reaction, the three-stage reaction may also proceed stepwise from amplification of the most abundant nucleic acid species at the lowest annealing temperature to amplification of the least abundant species at the highest annealing temperature.

By this sequential amplification method, an additional tool is provided for the “balancing” of multiplex PCR reactions besides the matching of T_m s and using limiting amounts of one or more PCR primer sets. The exploitation of PCR primer sets with different T_m s as a method for sequentially amplifying different amplicons may be preferred in certain circumstances to the sequential addition of additional primer sets. However, the use of temperature-dependent sequencing of multiplex PCR reactions may be coupled with the sequential physical addition of primer sets to a single reaction mixture.

An internal positive control that confirms the operation of a particular amplification reaction for a negative result also may be used. The internal positive controls (IPC) are DNA oligonucleotides that have the same primer sequences as the target gene (CEA or tyrosinase) but have a different internal probe sequence. Selected sites in the IPC's optionally may be synthesized with uracil instead of thymine so that contamination with the highly concentrated mimic could be controlled using uracil DNA glycosylase, if required. The IPCs maybe added to any PCR reaction mastermix in amounts that are determined empirically to give C_t values typically greater than the C_t values of the endogenous target of the primer set. The PCR assays are then performed according to standard protocols, and even when there is no endogenous target for the primer set, the IPC would be amplified, thereby verifying that the failure to amplify the target endogenous DNA is not a failure of the PCR reagents in the mastermix. In this embodiment, the IPC probe fluoresces differently than the probe for the endogenous sequences. A

variation of this for use in RT-PCR reactions is where the IPC is an RNA and the RNA includes an RT primer sequence. In this embodiment, the IPC verifies function of both the RT and PCR reactions. Both RNA and DNA IPCs (with different corresponding probes) may also be employed to differentiate difficulties in the RT and PCR reactions.

The rapid QRT-PCR protocols described herein may be run in about 20 minutes. This short time period permits the assay to be run intraoperatively so that a surgeon can decide on a surgical course during a single operation (typically the patient will remain anesthetized and/or otherwise sedated in a single "operation", though there may be a waiting period between when the sample to be tested is obtained and the time the interoperative assay is complete), rather than requiring a second operation, or requiring the surgeon to perform unneeded or overly broad prophylactic procedures. For instance, in the surgical evaluation of certain cancers, including breast cancer, melanoma, lung cancer, esophageal cancer and colon cancer, tumors and sentinel lymph nodes are removed in a first operation. The sentinel nodes are later evaluated for micrometastases, and, when micrometastases are detected in a patient's sentinel lymph node, the patient will need a second operation, thereby increasing the patient's surgical risks and patient discomfort associated with multiple operations. With the ability to determine the expression levels of certain tumor-specific markers described herein in less than 30 minutes with increased accuracy, a physician can make an immediate decision on how to proceed without requiring the patient to leave the operating room or associated facilities. The rapid test also is applicable to needle biopsies taken in a physician's office. A patient need not wait for days to get the results of a biopsy (such as a needle biopsy of a tumor or lymph node), but can now get more accurate results in a very short time.

As used herein, in the context of gene expression analysis, a probe is "specific to" a gene or transcript if under reaction conditions it can hybridize specifically to transcripts of that gene within a sample, or sequences complementary thereto, and not to other transcripts. Thus, in a diagnostic assay, a probe is specific to a gene if it can bind to a specific transcript or desired family of transcripts in mRNA extracted from a specimen, to the practical exclusion (does not interfere substantially with the detection assay) of other transcripts. In a PCR assay, primers are specific to a gene if they specifically amplify a sequence of that gene, to the practical exclusion of other sequences in a sample.

Table B provides primer and probe sequences for the mRNA quantification assays described and depicted in the Examples and Figures. Figures 1-16 provide non-limiting examples of cDNA sequences of the various mRNA species detected in the Examples. Although the sequences provided in Table B were found effective in the assays described in the examples, other primers and probes would likely be equally suited for use in the QRT-PCR and other mRNA detection and quantification assays, either described herein or as are known in the art. Design of alternate primer and probe sets for PCR assays, as well as for other mRNA detection assays is well within the abilities of one of average skill in the art. For example and without limitation, a number of computer software programs will generate primers and primer sets for PCR assays from cDNA sequences according to specified parameters. Non limiting

examples of such software include, NetPrimer and Primer Premier 5, commercially available from PREMIER Biosoft International of Palo Alto, California, which also provides primer and probe design software for molecular beacon and array assays. Primers and/or probes for two or more different mRNAs can be identified, for example and without limitation, by aligning the two or more target sequences according to standard methods, determining common sequences between the two or more mRNAs and entering the common sequences into a suitable primer design computer program.

As used herein, a “primer or probe” for detecting a specific mRNA species is any primer, primer set and/or probe that can be utilized to detect and/or quantify the specific mRNA species. An “mRNA species” can be a single mRNA species, corresponding to a single mRNA expression product of a single gene, or can be multiple mRNAs that are detected by a single common primer and/or probe combination, such as the SCCA1.2 and MAGEA136-plex species described below.

In the commercialization of the methods described herein, certain kits for detection of specific nucleic acids will be particularly useful. A kit typically comprises one or more reagents, such as, without limitation, nucleic acid primers or probes, packaged in a container, such as, without limitation, a vial, tube or bottle, in a package suitable for commercial distribution, such as, without limitation, a box, a sealed pouch, a blister pack and a carton. The package typically contains an indicia, for example and without limitation, a writing, illustration, label, book, booklet, tag and/or packaging insert, indicating that the packaged reagents can be used in a method for identifying expression of markers indicative of the presence of cancer cells in a lymph node of a patient. As used herein, “packaging materials” includes any article used in the packaging, for distribution of reagents in a kit, including, without limitation, containers, vials, tubes, bottles, pouches, blister packaging, labels, tags, instruction sheets, and package inserts.

One example of such a kit would include reagents necessary for the one-tube QRT-PCR process described above. In one example, the kit would include the above-described reagents, including reverse transcriptase, a reverse transcriptase primer, a corresponding PCR primer set, a thermostable DNA polymerase, such as Taq polymerase, and a suitable fluorescent reporter, such as, without limitation, a probe for a fluorescent 5' nuclease assay, a molecular beacon probe, a single dye primer or a fluorescent dye specific to double-stranded DNA, such as ethidium bromide. The primers may be present in quantities that would yield the high concentrations described above. Thermostable DNA polymerases are commonly and commercially available from a variety of manufacturers. Additional materials in the kit may include: suitable reaction tubes or vials, a barrier composition, typically a wax bead, optionally including magnesium; reaction mixtures (typically 10X) for the reverse transcriptase and the PCR stages, including necessary buffers and reagents such as dNTPs; nuclease- or RNase- free water; RNase inhibitor; control nucleic acid(s) and/or any additional buffers, compounds, co-factors, ionic constituents, proteins and enzymes, polymers, and the like that may be used in reverse transcriptase and/or PCR stages of QRT-PCR reactions.

Components of a kit are packaged in any manner that is commercially practicable. For example, PCR primers and reverse transcriptase may be packaged individually to facilitate flexibility in configuring the assay, or together to increase ease of use and to reduce contamination. Similarly, buffers, salts and co-factors can be packaged separately or together.

The kits also may include reagents and mechanical components suitable for the manual or automated extraction of nucleic acid from a tissue sample. These reagents are known to those skilled in the art and typically are a matter of design choice. For instance, in one embodiment of an automated process, tissue is disrupted ultrasonically in a suitable lysis solution provided in the kit. The resultant lysate solution is then filtered and RNA is bound to RNA-binding magnetic beads also provided in the kit or cartridge. The bead-bound RNA is washed, and the RNA is eluted from the beads and placed into a suitable reverse transcriptase reaction mixture prior to the reverse transcriptase reaction. In automated processes, the choice of reagents and their mode of packaging (for instance in disposable single-use cartridges) typically are dictated by the physical configuration of the robotics and fluidics of the specific RNA extraction system, for example and without limitation, the GenXpert system. International Patent Publication Nos. WO 04/48931, WO 03/77055, WO 03/72253, WO 03/55973, WO 02/52030, WO 02/18902, WO 01/84463, WO 01/57253, WO 01/45845, WO 00/73413, WO 00/73412 and WO 00/72970 provide non-limiting examples of cartridge-based systems and related technology useful in the methods described herein.

The constituents of the kits may be packaged together or separately, and each constituent may be presented in one or more tubes or vials, or in cartridge form, as is appropriate. The constituents, independently or together, may be packaged in any useful state, including without limitation, in a dehydrated, a lyophilized, a glassified or an aqueous state. The kits may take the physical form of a cartridge for use in automated processes, having two or more compartments including the above-described reagents. Suitable cartridges are disclosed for example in United States Patent Nos. 6,440,725, 6,431,476, 6,403,037 and 6,374,684.

Array technologies also can facilitate determining the expression level of two or more genes by facilitating performance of the desired reactions and their analysis by running multiple parallel reactions at the same time. One example of an array is the GeneChip® gene expression array, commercially available from Affymetrix, Inc. of Santa Clara, California. Patents illustrating array technology and uses therefor include, without limitation, United States Patent Nos. 6,040,138, 6,245,517, 6,251,601, 6,261,776, 6,306,643, 6,309,823, 6,346,413, 6,406,844 and 6,416,952. A plethora of other “array” patents exist, illustrating the multitude of physical forms a useful array can take. An “array”, such as a “microarray” can be a substrate containing one or more binding reagents, typically in discrete physical locations, permitting high throughput analysis of the binding of a sample to the array. In the context of the methods described herein, an array contains probes specific to transcripts of one or more of the genes described herein affixed to a substrate. The probes can be nucleic acids or analogs thereof, as are known

in the art. An array also can refer to a plurality of discrete reaction chambers, permitting multiple parallel reactions and detection events on a miniaturized scale.

As mentioned above, PCR-based technologies may be used to quantify mRNA levels in a given tissue sample. Other sequence-specific nucleic acid quantification methods may be more or less suited. In one embodiment, the nucleic acid quantification method is a rolling circle amplification method. Non-limiting examples of rolling circle amplification methods are described in U.S. Patents Nos. 5,854,003; 6,183,960; 6,344,329; and 6,210,884, each of which are incorporated herein by reference to the extent they teach methods for detecting and quantifying RNA species. In one embodiment, a padlock probe is employed to facilitate the rolling circle amplification process. (See Nilsson, M. et al. (2002), "Making Ends Meet in Genetic Analysis Using Padlock Probes," *Human Mutation* 19:410-415 and Schweitzer, B. et al (2001), "Combining Nucleic Acid Amplification and Detection," *Current Opinion in Biotechnology*, 12:21-27). A padlock probe is a linear oligonucleotide or polynucleotide designed to include one target-complementary sequence at each end, and which is designed such that the two ends are brought immediately next to each other upon hybridization to the target sequence. The probe also includes a spacer between the target-complementary sequences that includes a polymerase primer site and a site for binding to a probe, such as a molecular beacon probe, for detecting the padlock probe spacer sequence. If properly hybridized to an RNA template, the probe ends can then be joined by enzymatic DNA ligation to form a circular template that can be amplified by polymerase extension of a complementary primer. Thousands of concatemerized copies of the template can be generated by each primer, permitting detection and quantification of the original RNA template. Quantification can be automated by use, for example and without limitation, of a molecular beacon probe or other probe capable of detecting accumulation of a target sequence. By using padlock probes with different spacers to bind different molecular beacons that fluoresce a different color on binding to the amplified spacer, this automated reaction can be multiplexed. Padlock probe sequences target unique portions of the target RNA in order to ensure specific binding with limited or no cross-reactivity. RCA is an isothermic method in that the amplification is performed at one temperature.

Another isothermic method, for example and without limitation, is nucleic acid sequence-based amplification (NASBA). A typical NASBA reaction is initiated by the annealing of a first oligonucleotide primer to an RNA target in an RNA sample. The 3' end of the first primer is complementary to the target analyte; the 5' end encodes the T7 RNA polymerase promoter. After annealing, the primer is extended by reverse transcription (AMV-RT, for example) to produce a cDNA. The RNA is digested with RNase H, permitting a second primer (sense) to anneal to the cDNA strand, permitting the DNA polymerase activity of the reverse transcriptase to be engaged, producing a double-stranded cDNA copy of the original RNA template, with a functional T7 RNA polymerase promoter at one end. T7 polymerase is then used to produce an additional RNA template, which is further amplified, though in reverse order, according to the same procedure. A variety of other nucleic acid detection

and/or amplification methods are known to those of skill in the art, including variations on the isothermal strand displacement, PCR and RCA methods described herein.

Example 1 – General Materials and Methods

Identification of Potential Markers. An extensive literature and public database survey was conducted to identify any potential markers. Resources for this survey included PubMed, OMIM, UniGene (<http://www.ncbi.nlm.nih.gov/>), GeneCards (<http://bioinfo.weizmann.ac.il/cards>), and CGAP (<http://cgap.nci.nih.gov>). Survey criteria were somewhat flexible but the goal was to identify genes with moderate to high expression in tumors and low expression in normal lymph nodes. In addition, genes reported to be upregulated in tumors and genes with restricted tissue distribution were considered potentially useful. Finally, genes reported to be cancer-specific, such as the cancer testis antigens and hTERT, were evaluated.

Tissues and Pathological Evaluation. Tissue specimens were obtained from tissue banks at the University of Pittsburgh Medical Center through IRB approved protocols. All specimens were snap frozen in liquid nitrogen and later embedded in OCT for frozen sectioning. Twenty 5-micron sections were cut from each tissue for RNA isolation. In addition, sections were cut and placed on slides for H&E and IHC analysis at the beginning, middle (between the tenth and eleventh sections for RNA), and end of the sections for RNA isolation. All three H&E slides from each specimen underwent pathological review to confirm presence of tumor, percentage of tumor, and to identify the presence of any contaminating tissues. All of the unstained slides were stored at -20°C. Immunohistochemistry evaluation was performed using the AE1/AE3 antibody cocktail (DAKO, Carpinteria, CA), and Vector Elite ABC kit and Vector AEC Chromagen (Vecta Laboratories, Burlingame, CA). IHC was used as needed as needed to confirm the H&E histology.

Screening Approach. The screening was conducted in two phases. All potential markers entered the primary screening phase and expression was analyzed in 6 primary tumors and 10 benign lymph nodes obtained from patients without cancer (5 RNA pools with 2 lymph node RNA's per pool). Markers that showed good characteristics for lymph node metastasis detection passed into the secondary screening phase. The secondary screen consisted of expression analysis on 20-25 primary tumors, 20-25 histologically positive lymph nodes and 21 benign lymph nodes without cancer.

RNA Isolation and cDNA Synthesis. RNA was isolated using the RNeasy minikit (Qiagen, Valencia, CA) essentially as described by the manufacturer. The only modification was that we doubled the volume of lysis reagent and loaded the column in two steps. This was found to provide better RNA yield and purity, probably as a result of diluting out the OCT in the tissue sections. Reverse transcription was performed in 100- μ l reaction volumes either with random hexamer priming or sequence-specific priming using a probe indicated in Table C and Superscript II (Invitrogen, Carlsbad, CA) reverse transcriptase. For the primary screen, three reverse transcription reactions were performed, each with 500ng of RNA.

The cDNA's were combined and QPCR was performed using the equivalent of 20ng RNA per reaction. For the secondary screen, the RNA input for primary tumors and positive nodes was also 500ng. For benign nodes however, the RNA input was 2000ng resulting in the equivalent of 80ng RNA per QPCR reaction.

Quantitative PCR. All quantitative PCR was performed on the ABI Prism 7700 Sequence Detection Instrument (Applied Biosystems, Foster City, CA). Relative expression of the marker genes was calculated using the delta- C_T methods previously described and with β -glucuronidase as the endogenous control gene. All assays were designed for use with 5' nuclease hybridization probes although the primary screening was performed using SYBER Green quantification in order to save cost. Assays were designed using the ABI Primer Express Version 2.0 software and where possible, amplicons spanned exon junctions in order to provide cDNA specificity. All primer pairs were tested for amplification specificity (generation of a single band on gels) at 60, 62 and 64°C annealing temperature. In addition, PCR efficiency was estimated using SYBER green quantification prior to use in the primary screen. Further optimization and more precise estimates of efficiency were performed with 5' nuclease probes for all assays used in the secondary screen.

A mixture of the Universal Human Reference RNA (Stratagene, La Jolla, CA) and RNAs from human placenta, thyroid, heart, colon, PCI13 cell line and SKBR3 cell line served as a universal positive expression control for all the genes in the marker screening process.

Quantification with SYBER Green (Primary Screen). For SYBR Green I-based QPCR, each 50 μ l reaction contained 1 $\times$ TaqMan buffer A (Applied Biosystems), 300nM each dNTP, 3.5mM MgCl₂, 0.06 units/ μ l Amplitaq Gold (Applied Biosystems), 0.25X SYBR Green I (Molecular Probes, Eugene, OR) and 200nM each primer. The amplification program comprised 2-stages with an initial 95°C *Taq* activation stage for 12 min followed by 40 cycles of 95°C denaturation for 15 s, 60 or 62 or 64°C anneal/extend for 60 s and a 10 second data collection step at a temperature 2-4°C below the T_m of the specific PCR product being amplified (Tom B. Morrison, et al, 1998). After amplification, a melting curve analysis was performed by collecting fluorescence data while increasing the temperature from 60°C-95°C over 20 minutes.

Quantification with 5' Nuclease Probes (Secondary Screen). Probe-based QPCR was performed as described previously (Godfrey *et al.*, *Clinical Cancer Res.* 2001 Dec., 7(12):4041-8). Briefly, reactions were performed with a probe concentration of 200nM and a 60 second anneal/extend phase at 60°C, or 62°C, or 64°C. The sequences of primers and probes (purchased from IDT, Coralville, IA) for genes evaluated in the secondary screen are listed in Table B, below.

Data Analysis. In the primary screen, data from the melt curve was analyzed using the ABI Prism 7700 Dissociation Curve Analysis 1.0 software (Applied Biosystems). The first derivative of the melting curve

was used to determine the product T_m as well as to establish the presence of the specific product in each sample. In general, samples were analyzed in duplicate PCR reactions and the average C_t value was used in the expression analysis. However, in the secondary screen triplicate reactions were performed for each individual benign node and the lowest C_t value was used in the calculation of relative expression in order to obtain the highest value of background expression for the sample.

Cancer tissue-specific studies have been conducted, as described in the Examples below, in which a variety of molecular markers were identified as correlating with pathological states in cancers including esophageal cancer, colon cancer, head and neck cancer and in melanoma. Table A identifies genes used in the following studies. Table B provides PCR primer and TAQMAN probe sequences used in the quantitative PCR and RT-PCR amplifications described herein. Table C provides RT primer sequences as used instead of random hexamer primers. All PCR and RT-PCR reactions were conducted using standard methods. For all figures, T=primary tumor; PN=tumor-positive lymph nodes (by histological screening, that is, by review of H&E stained tissue and, when needed, by IHC, as described above); and BN=benign lymph nodes (by histological screening)

TABLE A

Marker	Accession No./ OMIN No.*	Official Gene Symbol	Official Gene Name	Alternative Gene Symbol	Alias
CDX1	NM_001804/ 600746	CDX1	caudal type homeo box transcription factor 1	NA	NA
CEA	NM_004363/ 114890	CEACAM5	carcinoembryonic antigen-related cell adhesion molecule 5	CEA, CD66e	NA
CK19	NM_002276/ 148020	KRT19	keratin 19	K19, CK19, K1CS, MGC15366	cytokeratin 19; keratin, type I, 40-kd; keratin, type I cytoskeletal 19; 40-kDa keratin intermediate filament precursor gene
CK20	NM_019010/ 608218	KRT20	keratin 20	K20, CK20, MGC35423	cytokeratin 20; keratin, type I; cytoskeletal 20
TACSTD1	NM_002354/ 185535	TACSTD1	tumor-associated calcium signal transducer 1	EGP, KSA, M4S1, MK-1, KS1/4, EGP40, MIC18, TROP1, Ep-CAM, CO17-1A, GA733-2	MK-1 antigen; antigen identified by monoclonal antibody AUA1; membrane component, chromosome 4, surface marker (35kD glycoprotein)
VIL1	NM_007127/ 193040	VIL1	villin 1	VIL, D2S1471	villin-1
CK7	NM_005556/ 148059	KRT7	keratin 7	K7, CK7, SCL, K2C7, MGC3625	Sarcolectin; cytokeratin 7; type II mesothelial keratin K7; keratin, type II cytoskeletal 7; keratin, 55K type II cytoskeletal; keratin, simple epithelial type I, K7
SCCA1	NM_006919/ 600517	SERPINB3	serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 3 carcinoma antigen 1&2	SCC, T4-A, SCCA1, SCCA-PD	squamous cell carcinoma antigen 1

SCCA2	NM_002974/ 600518	SERPINB4	serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 4	PII1, SCCA2, LEUPIN	leupin; squamous cell carcinoma antigen 2; protease inhibitor (leucine- serpin)
PTHrP	NM_002820/ 168470	PTHLH	parathyroid hormone-like hormone	PTHrP, PTHR, HHM,	parathyroid hormone-related protein; PTH-related protein; formerly humoral hypercalcemia of malignancy, included;
PVA	NM_001944/ 169615	DSG3	desmoglein 3 (pemphigus vulgaris antigen)	PVA, CDHF6	pemphigus vulgaris antigen; 130-kD pemphigus vulgaris antigen
MAGEA1	NM_004988/ 300016	MAGEA1	melanoma antigen, family A, 1 (directs expression of antigen MZ2-E)	MAGE1, MGC9326	melanoma antigen MAGE-1; melanoma-associated antigen 1; melanoma-associated antigen MZ2-E
MAGEA3	NM_005362/ 300174	MAGEA3	melanoma antigen, family A, 3	HP8, HYPD, MAGE3, MGC14613	antigen MZ2-D; MAGE-3 antigen; melanoma-associated antigen 3
MAGEA6	NM_005363/ 300176	MAGEA6	melanoma antigen, family A, 6	MAGE6, MAGE3B, MAGE- 3b, MGC52297	MAGE-6 antigen; melanoma-associated antigen 6
MART1	NM_005511/ 605513	MLANA	melan-A	MART1, MART-1	melanoma antigen recognized by T cells 1
TYR	NM_000372/ 606933	TYR	tyrosinase (oculocutaneous albinism IA)	OCA1A, OCA1A	Tyrosinase

* Online Mendelian Inheritance in Man (www.ncbi.nlm.nih.gov).

Table B

Gene	Oligonucleotide	Sequence (5' → 3')	Sequence Listing Reference
CDX1	Forward primer	CGGTGGCAGCGGTAAAGAC	SEQ ID NO: 1, bases 516 to 533
	Reverse primer	GATTGTGATGTAAACGGCTGTAATG	SEQ ID NO: 17
	Probe	ACCAAGGACAAAGTACCGCGTGGTCTACA	SEQ ID NO: 1, bases 538 to 565
CEA	Forward primer	AGACAAATCACAGTCTCTGCGGA	SEQ ID NO: 2, bases 1589 to 1610
	Reverse primer	ATCCTTGTCTCTCCACGGGTT	SEQ ID NO: 18
	Probe	CAAGCCCTCCATCTCCAGCAACAACCT	SEQ ID NO: 2, bases 1617 to 1642
CK19	Forward primer	AGATCGACAAACGCCCGT	SEQ ID NO: 19
	Reverse primer	AGAGCCTGTTCCGTCTCAAA	SEQ ID NO: 20
	Probe	TGGCTGCAGATGACTTCCGAACCA	SEQ ID NO: 4, bases 614 to 637
CK20	Forward primer	CACCTCCCAGAGCCTTGAGAT	SEQ ID NO: 5, bases 915 to 935
	Reverse primer	GGGCTTGGTCTCTCTAGAG	SEQ ID NO: 21
	Probe	CCATCTCAGCATGAAAGAGTCTTTGGAGCA	SEQ ID NO: 5, bases 948 to 977
CK7	Forward primer	CCCTCAATGAGACGGAGTTGA	SEQ ID NO: 3, bases 807 to 827
	Reverse primer	CCAGGAGCGACTGTTGTC	SEQ ID NO: 22
	Probe	AGCTGCAGTCCCAGATCTCCGACACATC	SEQ ID NO: 3, bases 831 to 858
MAGEA136 _{plex} ^A	Forward primer	GTGAGGAGGCAAGGTTYTSAG	SEQ ID NO: 23
	Reverse primer	AGACCCACWGGCAGATCTTCTC	SEQ ID NO: 24
	Probe1	AGGATTCCCTGGAGGCCACAGAGG	SEQ ID NO: 6, bases 80 to 103
MART1	Probe2	ACAGGCTGACCTGGAGGACCAGAGG	SEQ ID NO: 7, bases 90 to 104
	Forward primer	GATGCTCACTTCATCTATGTTACC	SEQ ID NO: 9, bases 66 to 90
	Reverse primer	ACTGTCAGGATGCCGATCC	SEQ ID NO: 25
PTHrP	Probe	AGCGGCTCTTCAGCCGTGGTGT	SEQ ID NO: 26
	Forward primer	GCGGTGTTCTCTGTGAGCTA	SEQ ID NO: 10, bases 356 to 375
	Reverse primer	TCATGGAGGAGCTGATGTTTACA	SEQ ID NO: 27
PVA	Probe	TTCTAGCCCGCCGCTCAAAAGA	SEQ ID NO: 10, bases 409 to 430
	Forward primer	AAAGAAACCCCAATTGCCAAGATTAC	SEQ ID NO: 11, bases 280 to 304
	Reverse primer	CAAAGGGCGGCTGATCGAT	SEQ ID NO: 28
SCCA1.2 ^B	Probe	CCAAGCAACCCAGAAAATCACCTACCG	SEQ ID NO: 11, bases 314 to 340
	Forward primer	AAGTGCAACATATCATGTTGATAGG	SEQ ID NO: 12, bases 267 to 292
	Reverse primer	GGCGATCTTCAGCTCATATGC	SEQ ID NO: 29
	Probe	TGTTTCATCACCCAGTTTCAAAAGCTTCTGACT	SEQ ID NO: 12, bases 301 to 331

TACSTD1	Forward primer	TCAATTGCTCAAAGCTGGCTG	SEQ ID NO: 14, bases 348 to 368
	Reverse primer	GGTTTGTGCTCTTCTCCCAAGTTT	SEQ ID NO: 30
	Probe	AAATGTTTGGTGATGAAGGCAGAAATGAATGG	SEQ ID NO: 14, bases 371 to 402
TYR	Forward primer	ACTTACTCAGCCCAGCATCATTC	SEQ ID NO: 15, bases 1284 to 1306
	Reverse primer	ACTGATGGCTGTTGTACTCCTCC	SEQ ID NO: 31
	Probe	TCTCCTCTTGGCAGATTGTCTGTAGCCGA	SEQ ID NO: 15, bases 1308 to 1336
Villin1	Forward primer	TGGTTCCTGGCTTGGGATC	SEQ ID NO: 16, bases 2152 to 2170
	Reverse primer	TTGCCAGACTCCGCCTTC	SEQ ID NO: 32
	Probe	TCAAGTGGAGTAACACCCAAATCCTATGAGGACC	SEQ ID NO: 16, bases 2174 to 2206

^A A universal primer set designed to recognize transcripts of MAGEA1, MAGEA3 and MAGEA6.
^B A universal primer set designed to recognize transcripts of both SCCA1 AND SCCA2.

Table C

Gene Marker	RT Specific Primer (5' → 3')	Sequence Listing Reference
CEA	GTGAAGGCCACAGCAT	SEQ ID NO: 33
CK20	AACTGGCTGCTGTAACG	SEQ ID NO: 34
MART1	GCCGATGAGCAGTAAGACT	SEQ ID NO: 35
PVA	TGTCAACAACAAAGATTCCA	SEQ ID NO: 36
SCCA1.2	TCTCCGAAGAGCTTGTG	SEQ ID NO: 37
TACSTD1	AGCCCATCATTGTTCTG	SEQ ID NO: 38
TYR	CGTTCCATTGCATAAAG	SEQ ID NO: 40
VIL1	GCTCCAGTCCCTAAGG	SEQ ID NO: 41

Example 2 - Esophageal Cancer

Expression levels of CEA, CK7, CK19, CK20, TACSTD1 and VIL1 were determined by the methods described in Example 1. Figure 17 is a scatter plot showing the expression levels of CEA, CK7, CK19, CK20, TACSTD1 and VIL1 in primary tumor, tumor-positive lymph nodes and benign lymph nodes. Figures 18A-O provide scatter plots illustrating the ability of two-marker systems to distinguish between benign and malignant cells in a lymph node. Tables D and E provide the raw data from which the graphs of Figures 17 and 18A-O were generated. This data illustrates the strong correlation of expression of CEA, CK7, CK19, CK20, TACSTD1 and VIL1 markers, alone or in combination, in sentinel lymph nodes with the presence of malignant cells arising from an esophageal cancer in the sentinel lymph nodes.

Table D -Single Marker Prediction Characteristics for Esophageal Cancer

	Observed Data				Parametric Bootstrap Estimates*		
	Sensitivity	Specificity	AUC	Classification Accuracy	Sensitivity	Specificity	Classification Accuracy
CEA	.95	.95	.98	.95	.93	.93	.93
CK7	.95	.86	.94	.90	.82	.89	.85
CK19	1.0	1.0	1.0	1.0	.99	.94	.97
CK20	1.0	.95	.995	.98	.98	.92	.95
TACSTD1	1.0	1.0	1.0	1.0	.96	.99	.98
Villin1	.95	.95	.98	.95	.92	.93	.92

optimism = .02 - .05

1000 parametric bootstrap samples of lymph node expression levels were generated and a new decision rule based on the most accurate cutoff was formulated each time (total of 1000 decision rules). The bootstrap estimates are the average prediction properties from classifying the *original* 41 lymph nodes 1000 times.

Table E - Two Marker Prediction Characteristics for Esophageal Cancer

	Observed Data			Parametric Bootstrap Estimates*		
	Sensitivity	Specificity	Classification Accuracy	Sensitivity	Specificity	Classification Accuracy
CEA + CK7	.95	1.0	.98	.93	.99	.96
CEA + CK19	.95	1.0	.98	.97	.99	.98
CEA + CK20	.95	1.0	.98	.97	.99	.97
CEA + TACSTD1	1.0	1.0	1.0	.99	1.0	.99
CEA + Villin1	.95	1.0	.98	.95	1.0	.98
CK7 + CK19	1.0	1.0	1.0	.99	.99	.99
CK7 + CK20	.95	1.0	.98	.93	.99	.97
CK7 + TACSTD1	1.0	1.0	1.0	.99	1.0	.99
CK7 + Villin1	.95	1.0	.98	.95	.99	.98
CK19 + CK20	.95	1.0	.98	.97	.99	.98
CK19 + TACSTD1	1.0	1.0	1.0	.99	1.0	.99
CK19 + Villin1	1.0	1.0	1.0	.99	.99	.99
CK20 + TACSTD1	1.0	1.0	1.0	.99	1.0	.99
CK20 + Villin1	.95	1.0	.98	.94	1.0	.97
TACSTD1 + Villin1	1.0	1.0	1.0	.99	1.0	.99

1000 parametric bootstrap samples of 41 lymph node marker pair expression levels were generated. For each new sample a new decision rule was devised to split the region into 2 zones equal prediction probability (see methods) (total of 1000 decision rules). The bootstrap estimates are the average prediction properties from classifying the *original* 41 lymph nodes 1000 times.

Example 3 – Head and Neck Cancer

Figure 19 is a scatter plot showing the expression levels of CEA, CK19, PTHrP, PVA, SCCA1.2 and TACSTD1 in primary tumor, tumor-positive lymph nodes and benign lymph nodes. Figures 20A-F provides scatter plots illustrating the ability of two-marker systems to distinguish between benign and malignant cells in a lymph node. Tables F and G provide the raw data from which the graphs of Figures 19 and 20A-F were generated. This data illustrates the strong correlation between expression of CEA, CK19, PTHrP, PVA, SCCA1.2 and TACSTD1 markers, alone or in combination, in sentinel lymph nodes and the presence of malignant cells arising from a squamous cell carcinoma of the head and neck in the sentinel lymph nodes.

Table F - Single Marker Prediction Characteristics -Head and Neck Cancer

	Observed Data				Non Parametric Bootstrap Estimates*		
	Sensitivity	Specificity	AUC	Classification Accuracy	Sensitivity	Specificity	Classification Accuracy
CEA	1.0	.905	.990	.950	.974	.880	.872
CK19	.895	.905	.917	.900	.867	.880	.872
EGFR	.895	1.0	.945	.947	.873	.979	.925
PTHrP	.947	1.0	.990	.975	.938	.988	.963
PVA	1.0	1.0	1.0	1.0	1.0	1.0	1.0
SCCA1.2	1.0	1.0	1.0	1.0	.998	.985	.991
TACSTD1	1.0	.952	.997	.975	.983	.944	.962
							bias**
							.078
							.028
							.022
							.012
							.000
							.009
							.013

* 500 bootstrap samples of lymph node expression levels were generated and a new decision rule based on the most accurate cutoff was formulated each time (total of 500 decision rules). 500 bootstrap samples of lymph node expression levels were generated and a new decision rule based on the most accurate cutoff was formulated each time (total of 500 decision rules). The optimism in for each bootstrap sample is calculated as the difference between the classification statistic applied to the original data and applied to the bootstrap data. The average over all bootstrap samples is computed and reported as the bias in the values derived from the observed data (Efron's enhanced bootstrap prediction error estimate, see Efron and Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall/CRC Press Boca Raton, 1993).

** bias = enhanced bootstrap estimate of optimism, or the amount that classification accuracy is overestimated when tested on the original data.

Table G – Two Marker Prediction Characteristics for Head & Neck Cancer

	Observed Data			Non Parametric Bootstrap Estimates		
	Sensitivity	Specificity	Classification Accuracy	Sensitivity	Specificity	Classification Accuracy
PVA + TACSTD1	1.0	1.0	1.0	.993	1.0	.997
PVA + PTHrP	1.0	1.0	1.0	1.0	1.0	1.0
PVA + SCCA1.2	1.0	1.0	1.0	1.0	1.0	1.0
TACSTD1 + PTHrP	.947	1.0	.975	.944	1.0	.974
TACSTD1 + SCCA1.2	1.0	1.0	1.0	.984	1.0	.992
PTHrP + SCCA1.2	1.0	1.0	1.0	1.0	1.0	1.0

* 500 bootstrap samples of lymph node expression levels were generated and a new decision rule based on the most accurate cutoff was formulated each time (total of 500 decision rules). 500 bootstrap samples of lymph node expression levels were generated and a new decision rule based on the most accurate cutoff was formulated each time (total of 500 decision rules). The optimism in for each bootstrap sample is calculated as the difference between the classification statistic applied to the original data and applied to the bootstrap data. The average over all bootstrap samples is computed and reported as the bias in the values derived from the observed data (Efron's enhanced bootstrap prediction error estimate, see Efron and Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall/CRC Press Boca Raton, 1993).

**** bias = enhanced bootstrap estimate of optimism, or the amount that classification accuracy is overestimated when tested on the original data.**

Example 4 - Melanoma

Figure 21 is a scatter plot showing the expression levels of MART1, TYR and MAGEA136-plex in primary tumor, tumor-positive lymph nodes and benign lymph nodes. Figures 22A and 22B provide scatter plots illustrating the ability of two-marker systems to distinguish between benign and malignant cells in a lymph node. This data illustrates the strong correlation between expression of MART1, TYR and MAGEA136-plex markers, alone or in combination, in sentinel lymph nodes and the presence of malignant cells arising from melanoma in the sentinel lymph nodes.

Example 5 - Colon Cancer

Figure 23 is a scatter plot showing the expression levels of CDX1, CEA, CK19, CK20, TACSTD1 and VIL1 in primary tumor, tumor-positive lymph nodes and benign lymph nodes. This data illustrates the strong correlation between expression of CDX1, CEA, CK19, CK20, TACSTD1 and VIL1 markers, in sentinel lymph nodes and the presence of malignant cells arising from colon cancer in the sentinel lymph nodes.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

2005271961 02 Aug 2011

- 33 -

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method of identifying expression of markers indicative of the presence of displaced esophageal cancer cells in a lymph node of a patient, the method comprising determining if a first mRNA species specific to CK-19 is overabundant in an RNA sample prepared from the lymph node of the patient, and determining if a second mRNA species selected from the group consisting of, Villin-1, TACSTD1, and CK-20 is overabundant in the RNA sample from the lymph node, the overabundance of the first, and the second mRNA species being indicative of the presence of displaced esophageal cancer cells in the lymph node.
2. The method of claim 1, comprising determining if CK-19 mRNA and Villin-1 mRNA are overabundant in the RNA sample, wherein an overabundance of said first and second mRNA species is indicative of displaced esophageal cancer cells in the lymph node.
3. The method of claim 1, comprising determining if CK-19 mRNA and TACSTD1 mRNA are overabundant in the RNA sample, wherein an overabundance of said first and second mRNA species is indicative of displaced esophageal cancer cells in the lymph node.
4. The method of claim 1, comprising determining if CK-19 mRNA and CK20 mRNA are overabundant in the RNA sample, wherein an overabundance of said first and second mRNA species is indicative of displaced esophageal cancer cells in the lymph node.
5. The method of any one of claims 1-4, comprising quantifying the levels of the one or more mRNA species in the RNA sample.
6. The method according to any one of the preceding claims wherein a nucleic acid amplification assay selected from the group consisting of a PCR assay and an isothermal amplification assay is used to determine if the one or more mRNA species is overabundant in the RNA sample.
7. The method of claim 6, wherein the nucleic acid amplification assay is a reverse transcriptase-polymerase chain reaction (RT-PCR), or a quantitative reverse transcriptase-polymerase chain reaction (QRT-PCR) assay.

2005271961 29 Sep 2011

- 34 -

8. The method of claim 6, wherein the nucleic acid amplification assay is an isothermic assay.

9. The method of claim 8, wherein the nucleic acid amplification assay is an assay selected from the group consisting of rolling circle amplification (RCA) assay, and nucleic acid sequence based amplification (NASBA).

10. An article of manufacture when used in a method according to claim 1, comprising packaging material and a first nucleic acid that can be used to detect or quantify a CK-19 mRNA, and a second nucleic acid that can be used to detect or quantify an mRNA species selected from the group consisting of Villin-1, TACSTD1, and CK-20, wherein the packaging material comprises an indicia for using the nucleic acids in a method for determining if displaced esophageal identifying expression markers indicative of displaced esophageal cancer cells are present in a lymph node of a patient.

11. The article of manufacture of claim 10, wherein the nucleic acids comprise primers for use in a sequence specific nucleic acid detection or amplification assay.

12. The article of manufacture of claim 11, comprising a detectable probe for use in detecting accumulation of a product of the sequence-specific nucleic acid detection or amplification assay utilizing the one or more primers.

13. A composition comprising one or more primers or probes specific to an mRNA species of CK-19 and one or more primers or probes specific to an mRNA species selected from Villin-1, TACSTD1 and CK-20, and RNA from a lymph node of a patient diagnosed with or suspected of having esophageal cancer.

14. The composition of claim 13 comprising one or more primers selected from the group consisting of: SEQ ID NO 19, SEQ ID NO 20, SEQ ID NO 16, SEQ ID NO 32, SEQ ID NO 14, SEQ ID NO 30, SEQ ID NO 5, SEQ ID NO 21, SEQ ID NO 41, SEQ ID NO 38, and SEQ ID NO 34.

15. The composition of claim 13 comprising one or more probes selected from the group consisting of: SEQ ID NO 4, SEQ ID NO 16, SEQ ID NO 14, and SEQ ID NO 5.

- 35 -

16. The method of any one of claim 1 to 9, the article of any one of claims 10 to 12, or the composition of any one of claims 13 to 15, substantially as hereinbefore described.

2005271961 02 Aug 2011

CDX1 cDNA Sequence (SEQ ID NO: 1)

```
1  aggtgagcgg ttgctcgtcg tcggggcggc cggcagcggc ggctccaggg cccagcatgc
61  gcgggggacc ccgcggccac catgtatgtg ggctatgtgc tggacaagga ttcgcccgtg
121  taccgccggc cagccaggcc agccagcctc ggctggggcc cggcaaacta cggccccccg
181  gccccgcccc cggcgcccc gcagtacccc gacttctcca gctactctca cgtggagccg
241  gccccgcgc ccccgacggc ctggggggcg cccttccttg cggccaagga cgactggggc
301  gccgcctacg gcccgggccc cgcggcccct gccgccagcc cagcttcgct ggcatcggg
361  cccctccag actttagccc ggtgccggcg cccctggggc cgggccggg cctcctggcg
421  cagccctcg ggggcccggg cacaccgtcc tcgcccggag cgcagaggcc gacgccctac
481  gagtggatgc ggcgcagcgt ggcggccgga ggcggcggtg gcagcggtaa gactcggacc
541  aaggacaagt accgcgtggt ctacaccgac caccaacgcc tggagctgga gaaggagttt
601  cattacagcc gttacatcac aatccggcgg aaatcagagc tggctgcaa tctggggctc
661  actgaacggc aggtgaagat ctggttccaa aaccggcggg caaaggagcg caaagtgaac
721  aagaagaaac agcagcagca acagcccca cagccgccga tggccacga catcacggcc
781  accccagccg ggccatccct ggggggcctg tgtcccagca acaccagcct cctggccacc
841  tcctctcaa tgctgtgaa agaggagttt ctgccatagc cccatgcca gcctgtgcgc
901  cgggggacct ggggactcgg gtgctgggag tgtggctcct gtgggcccag gaggtctggt
961  ccgagtctca gccctgacct tctgggacat ggtggacagt cacctatcca cctctgcat
1021  ccccttggcc cattgtgtgc agtaagcctg ttggataaag accttcagc tcctgtgttc
1081  tagacctctg ggggataagg gagtccaggg tggatgatct caatctcccg tgggcatctc
1141  aagccccaaa tggttggggg aggggcctag acaaggctcc agggcccacc tcctcctcca
1201  tacgttcaga ggtgcagctg gaggcctgtg tggggaccac actgatcctg gagaaaagg
1261  atggagctga aaaagatgga atgcttgagc agcatgacct gagggaggag gaacgtggtc
1321  aactcacacc tgctcttct gcagcctcac ctctacctgc cccatcata agggcactga
1381  gcccttccca ggctggatac taagcacaaa gcccatagca ctgggctctg atggctgctc
1441  cactgggtta cagaatcaca gccctcatga tcattctcag tgagggtctt ggattgagag
1501  ggaggccctg ggaggagaga agggggcaga gtcttcctta ccaggtttct acacccccgc
1561  caggctgccc atcagggccc agggagcccc cagaggactt tctcggacc aagcagagct
1621  cacagctgga cagggtgtgt atatagagtg gaatctcttg gatgcagctt caagaataaa
1681  ttttcttct cttttcaaa
```

Fig. 1

CEA cDNA Sequence (SEQ ID NO: 2)

```
1  ctcagggcag agggaggaag gacagcagac cagacagtca cagcagcctt gacaaaacgt
61  tcctggaact caagctcttc tccacagagg aggacagagc agacagcaga gaccatggag
121 tctccctcgg cccctcccca cagatggtgc atcccttggc agaggctcct gctcacagcc
181 tcacttctaa ccttctggaa ccgcccacc actgccaagc tcactattga atccacgccg
241 ttcaatgtcg cagaggggaa ggaggtgctt ctacttgtcc acaatctgcc ccagcatctt
301 tttggctaca gctggtacaa aggtgaaaga gtggatggca accgtcaa at tataggatat
361 gtaataggaa ctcaacaagc taccacaggg ccgcataca gtggtcgaga gataatatac
421 cccaatgcat ccctgctgat ccagaacatc atccagaatg acacaggatt ctacacccta
481 cacgtcataa agtcagatct tgtgaatgaa gaagcaactg gccagtccg ggtatacccg
541 gagctgcccc agccctccat ctccagcaac aactccaaac ccgtggagga caaggatgct
601 gtggccttca cctgtgaacc tgagactcag gacgcaacct acctgtggtg ggtaaacaat
661 cagagcctcc cggtcagtcc caggctgcag ctgtccaatg gcaacaggac cctcactcta
721 ttcaatgtca caagaaatga cacagcaagc tacaatgtg aaaccagaa cccagtgagt
781 gccaggcgca gtgattcagt catcctgaat gtcctctatg gcccgatgc cccaccatt
841 tcccctctaa acacatctta cagatcaggg gaaaatctga acctctctg ccacgcagcc
901 tctaaccac ctgcacagta ctcttggttt gtcaatggga ctttcagca atccaccaa
961 gagctcttta tcccacat cactgtgaat aatagtggat cctatactg ccaagcccat
1021 aactcagaca ctggcctcaa taggaccaca gtcacgacga tcacagtcta tgcagagcca
1081 cccaaacct tcatcaccag caacaactcc aaccccgagg aggatgagga tgctgtagcc
1141 ttaacctgtg aacctgagat tcagaacaca acctacctgt ggtgggtaaa taatcagagc
1201 ctcccggtca gtcccaggct gcagctgtcc aatgacaaca ggacctcac tctactcagt
1261 gtcacaagga atgatgtagg acctatgag tgtggaatcc agaacgaatt aagtgttgac
1321 cacagcgacc cagtcacct gaatgtcctc tatggcccag acgacccac catttcccc
1381 tcatacacct attaccgtcc aggggtgaac ctacgcctct cctgccatgc agcctctaac
1441 ccacctgcac agtattcttg gctgattgat ggaacatcc agcaacacac acaagagctc
1501 tttatctcca acatcactga gaagaacagc ggactctata cctgccaggc caataactca
1561 gccagtggcc acagcaggac tacagtcaag acaatcacag tctctgcgga gctgccaag
1621 ccctccatct ccagcaaca ctccaaacct gtggaggaca aggatgctgt ggccttcacc
1681 tgtgaacctg aggctcagaa cacaacctac ctgtggtggg taaatggtca gagcctccca
1741 gtcagtccca ggctgcagct gtccaatggc aacaggacct tcactctatt caatgtcaca
1801 agaaatgacg caagagccta tgtatgtgga atccagaact cagtgagtgc aaaccgcagt
1861 gaccagtcac ccctggatgt cctctatggg ccggacacct ccatcatttc cccccagac
```

Fig. 2/1

1921 tcgtcttacc tttcgggagc gaacctcaac ctctcctgcc actcggcctc taacccatcc
1981 ccgcagttatt cttggcggtat caatgggata ccgcagcaac acacacaagt tctctttatc
2041 gccaaaatca cgccaaataa taacgggacc tatgcctggt ttgtctctaa cttggctact
2101 ggccgcaata attccatagt caagagcatc acagtctctg catctggaac ttctcctggt
2161 ctctcagctg gggccactgt cggcatcatg attggagtgc tggttggggt tgctctgata
2221 tagcagccct ggtgtagttt cttcatattca ggaagactga cagttgtttt gcttcttcct
2281 taaagcattt gcaacagcta cagtctaaaa ttgcttcttt accaaggata tttacagaaa
2341 agactctgac cagagatcga gaccatccta gccaacatcg tgaaacccca tctctactaa
2401 aaatacaaaa atgagctggg cttggtggcg cgcacctgta gtcccagtta ctcgggaggc
2461 tgaggcagga gaatcgcttg aaccggggag gtggagattg cagtgagccc agatcgcacc
2521 actgcactcc agtctggcaa cagagcaaga ctccatctca aaaagaaaag aaaagaagac
2581 tctgacctgt actcttgaat acaagtttct gataccactg cactgtctga gaatttccaa
2641 aactttaatg aactaactga cagcttcatg aaactgtcca ccaagatcaa gcagagaaaa
2701 taattaattt catgggacta aatgaactaa tgaggattgc tgattcttta aatgtcttgt
2761 ttcccagatt tcaggaaact ttttttcttt taagctatcc actcttacag caatttgata
2821 aaatatactt ttgtgaacaa aaattgagac atttacattt tctccctatg tggtcgctcc
2881 agacttgagg aactattcat gaatatttat attgtatggt aatatagtta ttgcacaagt
2941 tcaataaaaa tctgctcttt gtataacaga aaaa

Fig. 2/2

CK7 cDNA Sequence (SEQ ID NO: 3)

```

1  acggcgagtg  cgcgctcctc  ctgccccgcc  gctaggtcca  tccccggcca  gccaccatgt
61  ccatccactt  cagctccccg  gtattcacct  cgcgctcagc  cgcttctctg  ggccgcggcg
121  cccaggtgcg  cctgagctcc  gctcgccccg  gcggccttgg  cagcagcagc  ctctacggcc
181  tcggcgctc  gcggccgcgc  gtggccgtgc  gctctgccta  tgggggcccc  gtgggcggcg
241  gcatccgcga  ggtcaccatt  aaccagagcc  tgctggcccc  gctgcggctg  gacgccgacc
301  cctccctcca  gcgggtgcgc  caggaggaga  gcgagcagat  caaggccctc  aacaacaagt
361  ttgcctcctt  catcgacaag  gtgcggtttc  tggagcagca  gaacaagctg  ctggagacca
421  agtggacgct  gctgcaggag  cagaagtcgg  ccaagagcag  ccgcctccca  gacatctttg
481  agggcccagat  tgctggcctt  cggggtcage  ttgaggcact  gcaggtggat  gggggccgcc
541  tggagcaggg  gctgcggacg  atgcaggatg  tgggtggagga  cttcaagaat  aagtacgaag
601  atgaaattaa  ccgccgcaca  gctgctgaga  atgagtttgt  ggtcctgaag  aaggatgtgg
661  atgctgccta  catgagcaag  gtggagctgg  aggccaaggt  ggatgccttg  aatgatgaga
721  tcaacttcct  caggaccctc  aatgagacgg  agttgacaga  gctgcagtcc  cagatctccg
781  acacatctgt  ggtgctgtcc  atggacaaca  gtcgctccct  ggacctggac  ggcacatcgc
841  ctgaggtcaa  ggcacagtat  gaggagatgg  ccaaagtcag  ccgggctgag  gctgaagcct
901  ggtaccagac  caagtttgag  accctccagg  cccaggctgg  gaagcatggg  gacgacctcc
961  ggaatacccg  gaatgagatt  tcagagatga  accgggccat  ccagaggctg  caggctgaga
1021  tcgacaacat  caagaaccag  cgtgccaaagt  tggaggccgc  cattgccgag  gctgaggagt
1081  gtggggagct  ggcgctcaag  gatgctcgtg  ccaagcagga  ggagctggaa  gccgccctgc
1141  agcggggcaa  gcaggatatg  gcacggcagc  tgcgtgagta  ccaggaactc  atgagcgtga
1201  agctggccct  ggacatcgag  atcgccacct  accgcaagct  gctggagggc  gaggagagcc
1261  ggttggctgg  agatggagtg  ggagccgtga  atatctctgt  gatgaattcc  actggtggca
1321  gtagcagtgg  cggtggcatt  gggctgaccc  tcgggggaac  catgggcagc  aatgcctga
1381  gcttctccag  cagtgcgggt  cctgggctcc  tgaaggctta  ttccatccgg  accgcatccg
1441  ccagtcgcag  gagtgcccg  gactga

```

Fig. 3

CK19 cDNA Sequence (SEQ ID NO: 4)

```

1  cgcccctgac accattcctc ccttcccccc tccaccggcc gcgggcataa aaggcgccag
61  gtgagggcct cgccgctcct cccgcgaatc gcagcttctg agaccagggg tgctccgtcc
121 gtgctccgcc tcgccatgac ttctacagc tatcgccagt cgtcggccac gtcgtccttc
181 ggaggcctgg gcggcggtc cgtgctttt ggccggggg tcgcctttcg cgcgccagc
241 attcacgggg gctccggcgg ccgcggcgta tccgtgtcct ccgcccgctt tgtgtcctcg
301 tcctcctcgg gggcctacgg cggcggtac ggcggtcctc tgaccgcgtc cgacgggctg
361 ctggcgggca acgagaagct aaccatgcag aacctcaacg accgcctggc ctctacctg
421 gacaaggtgc gcgccctgga ggcgcccaac ggcgagctag aggtgaagat ccgcgactgg
481 taccagaagc aggggcctgg gccctccgc gactacagcc actactacac gaccatccag
541 gacctgcggg acaagattct tggtgccacc attgagaact ccaggattgt cctgcagatc
601 gacaatgccc gtctggctgc agatgacttc cgaaccaagt ttgagacgga acaggctctg
661 cgcattgagc tggaggccga catcaacggc ctgcgcaggg tgctggatga gctgaccctg
721 gccaggaccg acctggagat gcagatcgaa ggctgaagg aagagctggc ctacctgaag
781 aagaacatg aggaggaaat cagtacgctg aggggccaag tgggaggcca ggtcagtgtg
841 gaggtggatt ccgctccggg caccgatctc gccaaagatc tgagtgacat gcgaagccaa
901 tatgaggtca tggccgagca gaaccggaag gatgctgaag cctgggtcac cagccggact
961 gaagaattga accgggaggt cgctggccac acggagcagc tccagatgag caggctccag
1021 gttactgacc tgcggcgcac ccttcagggt cttgagattg agctgcagtc acagctgagc
1081 atgaaagctg ccttggaaga cacttgga gaaacggagg cgcgctttgg agcccagctg
1141 gcgcatatcc aggcgctgat cagcgggtat gaagcccagc tgggcgatgt gcgagctgat
1201 agtgagcggc agaatcagga gtaccagcgg ctcatggaca tcaagtgcg gctggagcag
1261 gagattgcca cctaccgcag cctgctcgag ggacaggaag atcactacaa caatttgtct
1321 gcctccaagg tcctctgagg cagcaggctc tggggcttct gctgtccttt ggagggtgtc
1381 ttctgggtag agggatggga aggaagggac ccttaccccc ggctcttctc ctgacctgcc
1441 aataaaaatt tatggtccaa gggaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
1501 aaaaaaaaaa aaa

```

Fig. 4

CK20 cDNA Sequence (SEQ ID NO: 5)

```

1  caaccatcct gaagctacag gtgctocctc ctggaatctc caatggattt cagtcgcaga
61 agcttccaca gaagcctgag ctctcctctg caggcccttg tagtcagtac agtgggcatg
121 cagcgccctcg ggacgacacc cagcgtttat gggggtgctg gagggccggg catccgcac
181 tccaactcca gacacacggt gaactatggg agcgatctca caggcgggcg ggacctgttt
241 gttggcaatg agaaaatggc catgcagaac ctaaatagacc gtctagcgag ctacctagaa
301 aaggtgcgga ccctggagca gtccaactcc aaacttgaag tgcaaatcaa gcagtgttac
361 gaaaccaacg ccccgagggc tggtcgcgac tacagtgcac attacagaca aattgaagag
421 ctgcgaagtc agattaagga tgctcaactg caaaatgctc ggtgtgtcct gcaaattgat
481 aatgctaaac tggctgctga ggacttcaga ctgaagtatg agactgagag aggaatacgt
541 ctaacagtgg aagctgatct ccaaggcctg aataaggtct ttgatgacct aaccctacat
601 aaaacagatt tggagattca aattgaagaa ctgaataaag acctagctct cctcaaaaag
661 gagcatcagg aggaagtcga tggcctacac aagcatctgg gcaacactgt caatgtggag
721 gttgatgctg ctccaggcct gaaccttggc gtcacatga atgaaatgag gcagaagtat
781 gaagtcatgg ccagaagaa ccttcaagag gccaaagaac agtttgagag acagactgca
841 gttctgcagc aacaggtcac agtgaatact gaagaattaa aaggaactga ggttcaacta
901 acggagctga gacgcacctc ccagagcctt gagatagaac tccagtccca tctcagcatg
961 aaagagtctt tggagcacac tctagaggag accaaggccc gttacagcag ccagttagcc
1021 aacctccagt cgctgttgag ctctctggag gcccaactga tgcagattcg gagtaacatg
1081 gaacgccaga acaacgaata ccatatcctt cttgacataa agactcgact tgaacaggaa
1141 attgctactt accgccgcct tctggaagga gaagacgtaa aaactacaga atatcagtta
1201 agcaccctgg aagagagaga tataaagaaa accaggaaga ttaagacagt cgtgcaagaa
1261 gtagtggatg gcaaggtcgt gtcacatgaa gtcaaagagg tggagaaaaa tatctaaata
1321 gctaccagaa ggagatgctg ctgaggtttt gaaagaaatt tggctataat cttatctttg
1381 ctccctgcaa gaaatcagcc ataagaaagc actattaata ctctgcagtg attagaaggg
1441 gtgggggtggc gggaatccta tttatcagac tctgtaattg aatataaatg ttttactcag
1501 aggagctgca aattgcctgc aaaaatgaaa tccagtgagc actagaatat ttaaaacatc
1561 attactgcca tctttatcat gaagcacatc aattacaagc tgtagaccac ctaatatcaa
1621 tttgtaggta atgttcctga aaattgcaat acatttcaat tatactaaac ctcacaaagt
1681 agaggaatcc atgtaaattg caaataaacc acttttctaat tttttcctgt ttctgaaaaa
1741 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
1801 aaaaaaaaaa aaaaaaaa

```

Fig. 5

MAGEA1 cDNA Sequence (SEQ ID NO: 6)

```

1  cgtagagttc ggccgaagga acctgaccca ggctctgtga ggaggcaagg ttttcagggg
61  acaggccaac ccagaggaca ggattccctg gaggccacag aggagcacca aggagaagat
121 ctgcctgttg gtcttcattg cccagctcct gccacactc ctgcctgctg ccctgacgag
181 agtcatcatg tctcttgagc agaggagtct gcactgcaag cctgaggaag cccttgaggc
241 ccaacaagag gccctgggcc tgggtgtgtg gcaggctgcc gcctcctcct cctctcctct
301 ggtcctgggc accctggagg aggtgcccac tgctgggtca acagatcctc cccagagtcc
361 tcaggggagc tccgcctttc ccaactaccat caacttcact cgacagaggc aaccagtgga
421 ggggttcagc agccgtgaag aggagggggc aagcacctct tgtatcctgg agtccttggt
481 ccgagcagta atcactaaga aggtggctga tttggttggt tttctgctcc tcaaatatcg
541 agccagggag ccagtcacaa aggcagaaat gctggagagt gtcatacaaa attacaagca
601 ctgttttcct gagatcttcg gcaaagcctc tgagtccctg cagctgggtct ttggcattga
661 cgtgaaggaa gcagacccca ccggccactc ctatgtcctt gtcacctgcc taggtctctc
721 ctatgatggc ctgctgggtg ataatcagat catgcccagg acaggcttcc tgataattgt
781 cctgggtcatg attgcaatgg agggcggcca tgctcctgag gaggaaatct gggaggagct
841 gagtgtgatg gaggtgtatg atgggagggg gcacagtgcc tatggggagc ccaggaagct
901 gctcacccaa gatttggtgc aggaaaagta cctggagtac cggcaggtgc cggacagtga
961 tcccgcacgc tatgagttcc tgtgggggtc aagggccctt gctgaaacca gctatgtgaa
1021 agtccttgag tatgtgatca aggtcagtgc aagagttcgc tttttcttcc catccctgcg
1081 tgaagcagct ttgagagagg aggaagaggg agtctgagca tgagttgcag ccaggggccag
1141 tgggaggggg actggggccag tgcaccttcc agggccgcgt ccagcagctt cccctgcctc
1201 gtgtgacatg agggccattc ttcactctga agagagcggc cagtgttctc agtagtaggt
1261 ttctgttcta ttgggtgact tggagattta tctttgttct cttttggaat tgttcaaatg
1321 ttttttttta agggatggtt gaatgaactt cagcatccaa gtttatgaat gacagcagtc
1381 acacagttct gtgtatatag ttttaagggt agagtcttgt gttttattca gattgggaaa
1441 tccattctat tttgtgaatt gggataataa cagcagtggg ataagtactt agaaatgtga
1501 aaaatgagca gtaaaataga tgagataaag aactaaagaa attaagagat agtcaattct
1561 tgctttatac ctcagtctat tctgtaaaat ttttaaagat atatgcatac ctggatttcc
1621 ttggcttctt tgagaatgta agagaaatta aatctgaata aagaattctt cctgttaaaa
1681 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa

```

Fig. 6

MAGEA3 cDNA Sequence (SEQ ID NO: 7)

```
1 gagattctcg ccctgagcaa cgagcgacgg cctgacgtcg gcggaggaa gccggcccag
61 gctcgggtgag gaggcaaggt tctgagggga caggctgacc tggaggacca gaggcccccg
121 gaggagcact gaaggagaag atctgccagt gggctctccat tgcccagctc ctgccacac
181 tcccgctgtg tgccctgacc agagtcatca tgccctctga gcagaggagt cagcactgca
241 agcctgaaga aggccttgag gcccaggag agggcctggg cctggtgggt gcgcaggctc
301 ctgctactga ggagcaggag gctgcctcct cctcttctac tctagttaa gtcaccctgg
361 gggaggtgcc tgctgccgag tcaccagatc ctcccagag tcctcaggga gcctccagcc
421 tccccactac catgaactac cctctctgga gccaatccta tgaggactcc agcaaccaag
481 aagaggagg gccaagcacc ttccctgacc tggagtccga gttccaagca gcaactcagta
541 ggaaggtggc cgagttggtt cattttctgc tcctcaagta tcgagccagg gagccggtca
601 caaaggcaga aatgctgggg agtgctgtcg gaaattggca gtatttcttt cctgtgatct
661 tcagcaaagc ttccagttcc ttgcagctgg tctttggcat cgagctgatg gaagtggacc
721 ccatcgcca cttgtacatc ttggccacct gcctgggcct ctctacgat ggcctgctgg
781 gtgacaatca gatcatgccc aaggcaggcc tcctgataat cgtcctggcc ataatcgcaa
841 gagagggcga ctgtgcccct gaggagaaaa tctgggagga gctgagtgtg ttagaggtgt
901 ttgaggggag ggaagacagt atcttggggg atcccaagaa gctgctcacc caacatttcg
961 tgcaggaaaa ctacctggag taccggcagg tccccggcag tgatcctgca tgttatgaat
1021 tcctgtgggg tccaagggcc ctcggtgaaa ccagctatgt gaaagtccctg caccatatgg
1081 taaagatcag tggaggacct cacatttcct acccaccct gcatgagtgg gttttgagag
1141 agggggaaga gtgagtctga gcacgagttg cagccagggc cagtgggagg gggctctgggc
1201 cagtgcacct tccggggccg catcccttag tttccactgc ctctgtgac gtgaggccca
1261 ttcttcactc tttgaagcga gcagtcagca ttcttagtag tgggtttctg ttctgttgga
1321 tgactttgag attattcttt gtttcctgtt ggagttgttc aaatgttcct tttaacggat
1381 ggttgaatga gcgtcagcat ccaggtttat gaatgacagt agtcacacat agtgctgttt
1441 atatagttta ggagtaagag tcttggtttt tactcaaatt gggaaatcca ttccattttg
1501 tgaattgtga cataataata gcagtggtaa aagtatttgc ttaaaattgt gagcgaatta
1561 gcaataacat acatgagata actcaagaaa tcaaaagata gttgattctt gccttgtacc
1621 tcaatctatt ctgtaaaatt aaacaaatat gcaaaccagg atttccttga cttctttgag
1681 aatgcaagcg aaattaaatc tgaataaata attcttcctc ttcaaaaaaa aaaaaaaaaa
1741 aaaaaaaaaa aaa
```

Fig. 7

MAGEA6 cDNA Sequence (SEQ ID NO: 8)

```

1  agcaacgagc gacggcctga cgtcggcgga gggaagccgg cccaggctcg gtgaggaggc
61  aaggttctga ggggacaggc tgacctggag gaccagaggc ccccgaggga gactgaagg
121 agaagatctg ccagtgggtc tccattgccc agtccttgcc cacactcccg cctgttgccc
181 tgaccagagt catcatgcct cttgagcaga ggagtcagca ctgcaagcct gaagaaggcc
241 ttgaggcccg aggagaggcc ctgggcctgg tgggtgcgca ggctcctgct actgaggagc
301 aggaggctgc ctctctctct tctactctag ttgaagtcac cctgggggag gtgcctgctg
361 ccgagtcacc agatcctccc cagagtcctc agggagcctc cagcctcccc actaccatga
421 actaccctct ctggagccaa tcctatgagg actccagcaa ccaagaagag gaggggccaa
481 gcaccttccc tgacctggag tctgagttcc aagcagcact cagtaggaag gtggccaagt
541 tggttcattt tctgctcctc aagtatcgag ccagggagcc ggtcacaaag gcagaaatgc
601 tggggagtggt cgtcggaaat tggcagtact tctttcctgt gatcttcagc aaagcttccg
661 attccttgca gctggtcttt ggcctcgagc tgatggaagt ggaccccatc ggccacgtgt
721 acatctttgc cacctgcctg ggcctctcct acgatggcct gctgggtgac aatcagatca
781 tgccaagac aggttcctg ataatcatcc tggccataat cgcaaaagag ggcgactgtg
841 cccctgagga gaaaatctgg gaggagctga gtgtgttaga ggtgtttgag gggagggaag
901 acagtatctt cggggatccc aagaagctgc tcaccaata tttcgtgcag gaaaactacc
961 tggagtaccg gcaggtcccc ggcagtgatc ctgcatgcta tgagtccctg tggggtccaa
1021 gggccctcat tgaaaccagc tatgtgaaag tcctgcacca tatggtaaag atcagtggag
1081 gacctcgcat ttcctaccca ctctgcatg agtgggcttt gagagagggg gaagagtgag
1141 tctgagcacg agttgcagcc agggccagtg ggaggggggt tggggccagtg caccttccgg
1201 ggcccatcc cttagtttcc actgcctcct gtgacgtgag gccattctt cactcttga
1261 agcgagcagt cagcattctt agtagtgggt ttctgttctg ttggatgact ttgagattat
1321 tctttgtttc ctgttgaggt tgttcaaagt ttccttttaa cggatgggtg aatgagcgtc
1381 agcatccagg tttatgaatg acagtagtca cacatagtgc tgtttatata gtttaggagt
1441 aagagtcttg ttttttattc agattgggaa atccattcca ttttgtgaat tgtgacataa
1501 taatagcagt ggtaaaagta tttgcttaaa attgtgagcg aattagcaat aacatacatg
1561 agataactca agaaatcaaa agatagttga ttcttgctt gtacctcaat ctattctgta
1621 aatataaaca aatatgcaaa ccaggatttc cttgacttct ttgagaatgc aagcgaaatt
1681 aaatctgaat aaataattaa aaaaaaaaaa aaaaaaaaaa aaa

```

Fig. 8

MART1 cDNA Sequence (SEQ ID NO: 9)

1 agcagacaga ggactctcat taaggaaggt gtcctgtgcc ctgaccctac aagatgccaa
61 gagaagatgc tcacttcac t tatggttacc ccaagaagg gcacggccac tcttacacca
121 cggctgaaga ggccgctggg atcggcatcc tgacagtgat cctgggagtc ttactgctca
181 tcggctgttg gtattgtaga agacgaaatg gatacagagc cttgatggat aaaagtcttc
241 atggtggcac tcaatgtgcc ttaacaagaa gatgccaca agaagggttt gatcatcggg
301 acagcaaagt gtctcttcaa gagaaaaact gtgaacctgt gggtcccaat gctccacctg
361 cttatgagaa actctctgca gaacagtcac caccacctta ttcaccttaa gagccagcga
421 gacacctgag acatgctgaa attatttctc tcacactttt gcttgaattt aatacagaca
481 tctaattgtt ccttttgaa tgggtgtagga aaaatgcaag ccatctctaa taataagtca
541 gtgttaaaat tttagtaggt ccgctagcag tactaatcat gtgaggaaat gatgagaaat
601 attaaattgg gaaaactcca tcaataaatg ttgcaatgca tgatactatc tgtgccagag
661 gtaatgttag taaatccatg gtgttatttt ctgagagaca gaattcaagt gggatttctg
721 gggccatcca atttctcttt acttgaaatt tggctaataa caaactagtc aggttttctg
781 accttgaccg acatgaactg tacacagaat tgttcagta ctatggagtg ctcacaaagg
841 atacttttac aggttaagac aaagggttga ctggcctatt tatctgatca agaactgtc
901 agcaatgtct ctttgtgctc taaaattcta ttatactaca ataatatatt gtaaagatcc
961 tatagctctt tttttttgag atggagtttc gcttttgttg cccaggctgg agtgcaatgg
1021 cgcgatcttg gctcaccata acctccgcct cccaggttca agcaattctc ctgccttagc
1081 ctcttgagta gctgggatta caggcgtgcg ccactatgcc tgactaattt tgtagtttta
1141 gtagagacgg ggtttctcca tgggtggtcag gctgggtctca aactcctgac ctcagggtgat
1201 ctgcccgcct cagcctccca aagtgtgga attacaggcg tgagccacca cgctggctg
1261 gatcctatat cttaggttaag acatataacg cagtctaatt acatttcact tcaaggctca
1321 atgctattct aactaatgac aagtattttc tactaaacca gaaattggta gaaggattta
1381 aataagtaaa agctactatg tactgcctta gtgctgatgc ctgtgtactg ccttaaatgt
1441 acctatggca atttagctct cttgggttcc caaatccctc tcacaagaat gtgcagaaga
1501 aatcataaag gatcagagat tctg

Fig. 9

PTHrP cDNA Sequence (SEQ ID NO: 10)

```
1  cctgcatctt tttggaagga ttctttttat aaatcagaaa gtgttcgagg ttcaaagggt
61  tgcctcggag cgtgtgaaca ttcttcgctt cggttttcaa ctgcctcca acctgcgccg
121  cccggccagc atgtctcccc gcccgtaag cggggctgcc gcctccctgc cgctccggct
181  gccactaacg acccgccctc gccgccacct ggccctcctg atcgacgaca cagcacttg
241  aaacttggtt tcaggggtgtg tggaaatcaac tttccggaag caaccagccc accagaggag
301  gtcccgagcg cgagcggaga cgatgcagcg gagactgggt cagcagtgga gcgtcgcggg
361  gttcctgctg agctacgcgg tgccctcctg cgggcgctcg gtggaggggtc tcagccgccg
421  cctcaaaaaga gctgtgtctg aacatcagct cctccatgac aaggggaagt ccatccaaga
481  tttacggcga cgattcttcc ttcacatct gatcgagaa atccacacag ctgaaatcag
541  agctacctcg gaggtgtccc ctaactccaa gccctctccc aacacaaaaga accaccccg
601  ccgatttggg tctgatgatg agggcagata cctaactcag gaaactaaca aggtggagac
661  gtacaaagag cagccgctca agacacctgg gaagaaaaag aaaggcaagc ccgggaaacg
721  caaggagcag gaaaagaaaa aacggcgaac tcgctctgcc tgggttagact ctggagtgc
781  tgggagtggg ctagaagggg accacctgtc tgacacctcc acaacgtcgc tggagctcga
841  ttcacggtaa caggcttctc tggcccgtag cctcagcggg gtgctctcag ctgggttttg
901  gagcctccct tctgccttgg cttggacaaa ctagaattt tctcccttta tgtatctcta
961  tcgatttgtg agcaattgac agagaataac tcagaatatt gtctgcctta aagcagtacc
1021  cccctaccac acacaccctc gtcctccagc accatagaga ggcgctagag cccattcctc
1081  tttctccacc gtcaccaaac atcaatcctt taccactcta ccaaataatt tcatattcaa
1141  gcttcagaag ctagtgacca tcttcataat ttgctggaga agtgtgtttc ttccccttac
1201  tctcacacct gggcaaaact tcttcagtgt ttttcatttc ttacgttctt tcacttcaag
1261  ggagaatata gaagcatttg atattatcta caaacactgc agaacagcat catgtcataa
1321  acgattctga gccattcaca ctttttatatt aattaaatgt atttaattaa atctcaaatt
1381  tattttaatg taaagaactt aaattatgtt ttaaacacat gccttaaatt tgtttaatta
1441  aatttaactc tggtttctac cagctcatac aaaataaatg gtttctgaaa atgtttaagt
1501  attaaacttac aaggatatag gtttttctca tgtatctttt tgttcattgg caagatgaaa
1561  taatttttct agggtaatgc cgtaggaaaa ataaaacttc acatttatgt ggcttggtta
1621  tccttagctc acagattgag gtaataatga cactcctaga ctttgggatc aaataactta
1681  gggccaagtc ttgggtctga atttatttaa gttcacaacc tagggcaagt tactctgcct
1741  ttctaagact cacttacatc ttctgtgaaa tataattgta ccaacctcat agagtttggg
1801  gtcaactaaa tgagattata tgtggactaa atatctgtca tatagtaaac actcaataaa
1861  ttgcaacata ttaaaaaaaaa a
```

Fig. 10

PVA cDNA Sequence (SEQ ID NO: 11)

```
1  ttttcttaga cattaactgc agacggctgg caggatagaa gcagcggctc acttggactt
61  tttcaccagg gaaatcagag acaatgatgg ggctcttccc cagaactaca ggggctctgg
121 ccatcttcgt ggtggtcata ttggttcatt gagaattgct aatagagact aaaggtcaat
181 atgatgaaga agagatgact atgcaacaag ctaaaagaag gcaaaaacgt gaatgggtga
241 aatttgccaa accctgcaga gaaggagaag ataactcaa aagaaaccca attgccaaga
301 ttacttcaga ttaccaagca acccagaaaa tcacctaccg aatctctgga gtgggaatcg
361 atcagccgcc ttttggaatc tttgttggtg aaaaaaacac tggagatatt aacataacag
421 ctatagtcga ccgggaggaa actccaagct tcctgatcac atgtcgggct ctaaatgccc
481 aaggactaga tgtagagaaa ccacttatac taacgggtta aattttggat attaatgata
541 atcctccagt attttcacaa caaattttca tgggtgaaat tgaagaaaat agtgcctcaa
601 actcactggg gatgatacta aatgccacag atgcagatga accaaaccac ttgaattcta
661 aaattgcctt caaaattgtc tctcaggaac cagcaggcac acccatgttc ctcttaagca
721 gaaacactgg ggaagtccgt actttgacca attctcttga ccgagagcaa gctagcagct
781 atcgtctggg tgtgagtggg gcagacaaag atggagaagg actatcaact caatgtgaat
841 gtaatattaa agtgaaagat gtcaacgata acttcccaat gtttagagac tctcagtatt
901 cagcacgtat tgaagaaaat attttaagtt ctgaattact tcgatttcaa gtaacagatt
961 tggatgaaga gtacacagat aattggcttg cagtatatat ctttacctct gggaatgaag
1021 gaaattgggt tgaaatacaa actgataccta gaactaatga aggcatacctg aaagtgggtga
1081 aggcctctaga ttatgaacaa ctacaaagcg tgaaacttag tattgctgtc aaaaacaaag
1141 ctgaatttca ccaatcagtt atctctcgat accgagttca gtcaacccca gtcacaattc
1201 aggtaataaaa tgtaagagaa ggaattgcat tccgtcctgc ttccaagaca ttactgtgc
1261 aaaaaggcat aagtagcaaa aaattgggtg attatatcct gggaacatat caagccatcg
1321 atgaggacac taacaaagct gcctcaaattg tcaaatatgt catgggacgt aacgatgggtg
1381 gatacctaata gattgattca aaaactgctg aaatcaaatt tgtcaaaaat atgaaccgag
1441 attctacttt catagttaac aaaacaatca cagctgaggt tctggccata gatgaataga
1501 cgggtaaaaa ttctacaggc acggtatatg ttagagtacc cgatttcaat gacaattgtc
1561 caacagctgt cctcgaaaaa gatgcagttt gcagttcttc acctccctg gttgtctccg
1621 ctagaacact gaataataga tacactggcc cctatacatt tgcactggaa gatcaacctg
1681 taaagttgcc tgccgtatgg agtatcacia ccctcaatgc tacctcggcc ctctcagag
1741 ccaggaaca gatacctcct ggagtatacc acatctccct ggtacttaca gacagtcaga
1801 acaatcgggt tgagatgcca cgcagcttga cactggaagt ctgtcagtgat gacaacaggg
1861 gcattctgtg aacttcttac ccaaccacaa gccctgggac caggatatggc aggccgcact
```

Fig. 11/1

1921 cagggaggct ggggcctgcc gccatcggcc tgctgctcct tgggtctcctg ctgctgctgt
1981 tggcccccct tctgctgttg acctgtgact gtggggcagg ttctactggg ggagtgcacag
2041 gtggttttat cccagttcct gatggctcag aaggaacaat tcatcagtgg ggaattgaag
2101 gagcccatcc tgaagacaag gaaatcacia atatttgtgt gcctcctgta acagccaatg
2161 gagccgattt catggaaagt tctgaagttt gtacaaatac gtatgccaga ggcacagcgg
2221 tggaaggcac ttcaggaatg gaaatgacca ctaagcttgg agcagccact gaatctggag
2281 gtgctgcagg ctttgcaaca gggacagtgt caggagctgc ttcaggattc ggagcagcca
2341 ctggagtttg catctgttcc tcagggcagt ctggaaccat gagaacaagg cattccactg
2401 gaggaaccaa taaggactac gctgatgggg cgataagcat gaattttctg gactcctact
2461 tttctcagaa agcatttgcc tgtgctggagg aagacgatgg ccaggaagca aatgactgct
2521 tgttgatcta tgataatgaa ggcgcagatg ccactggttc tcctgtgggc tccgtgggtt
2581 gttgcagttt tattgctgat gacctggatg acagcttctt ggactcactt ggacccaaat
2641 ttaaaaaact tgcagagata agccttgggtg ttgatggtga aggcaaagaa gttcagccac
2701 cctctaaaga cagcggttat gggattgaat cctgtggcca tcccatagaa gtccagcaga
2761 caggatttgt taagtgccag actttgtcag gaagtcaagg agcttctgct ttgtccgcct
2821 ctgggtctgt ccagccagct gtttccatcc ctgaccctct gcagcatggg aactatttag
2881 taacggagac ttactcggct tctgggtccc tcgtgcaacc ttccactgca ggctttgatc
2941 cacttctcac acaaaatgtg atagtgcacg aaagggtgat ctgtccatt tccagtgttc
3001 ctggcaacct agctggccca acgcagctac gagggtcaca tactatgctc tgtacagagg
3061 atccttgctc ccgtctaata tgaccagaat gagctggaat accacactga ccaaatctgg
3121 atctttggac taaagtattc aaaatagcat agcaaagctc actgtattgg gctaataatt
3181 tggcacttat tagcttctct cataaactga tcacgattat aaattaaatg tttgggttca
3241 taccceaaaa gcaatatgtt gtcactccta attctcaagt actattcaaa ttgtagtaaa
3301 tcttaaagtt tttcaaaacc ctaaaatcat attcgc

Fig. 11/2

SCCA1 cDNA Sequence (SEQ ID NO: 12)

```
1  ctctctgccc  acctctgctt  cctctaggaa  cacaggagtt  ccagatcaca  tcgagttcac
61  catgaattca  ctcagtgaag  ccaacaccaa  gttcatgttc  gacctgttcc  aacagttcag
121  aaaatcaaaa  gagaacaaca  tcttctattc  ccctatcagc  atcacatcag  cattagggat
181  ggtcctctta  ggagccaaag  acaacactgc  acaacagatt  aagaaggttc  ttcactttga
241  tcaagtcaca  gagaacacca  caggaaaagc  tgcaacatat  catgttgata  ggtcaggaaa
301  tgttcatcac  cagtttcaaa  agcttctgac  tgaattcaac  aaatccactg  atgcatatga
361  gctgaagatc  gccacaagc  tcttcggaga  aaaaacgtat  ctatttttac  aggaatatat
421  agatgccatc  aagaaatatt  accagaccag  tgtggaatct  gttgattttg  caaatgctcc
481  agaagaaagt  cgaaagaaga  ttaactcctg  ggtggaaagt  caaacgaatg  aaaaaattaa
541  aaacctaatt  cctgaaggta  atattggcag  caataccaca  ttggttcttg  tgaacgcaat
601  ctatttcaaa  gggcagtggg  agaagaaatt  taataaagaa  gatactaaag  aggaaaaatt
661  ttggccaaac  aagaatacat  acaagtccat  acagatgatg  aggcaataca  catcttttca
721  ttttgccctg  ctggaggatg  tacaggccaa  ggtcctggaa  ataccataca  aaggcaaaga
781  tctaagcatg  attgtgttgc  tgccaaatga  aatcgatggg  ctccagaagc  ttgaagagaa
841  actcactgct  gagaaattga  tggaatggac  aagtttgag  aatatgagag  agacacgtgt
901  cgattttcac  ttacctcggg  tcaaagtggg  agagagctat  gacctcaagg  acacgttgag
961  aaccatggga  atggtggata  tcttcaatgg  ggatgcagac  ctctcaggca  tgaccgggag
1021  ccgcggtctc  gtgctatctg  gagtcctaca  caaggccttt  gtggagggtta  cagaggaggg
1081  agcagaagct  gcagctgcca  ccgctgtagt  aggattcgga  tcatcaccta  cttcaactaa
1141  tgaagagttc  cattgtaatc  accctttcct  attcttcata  aggcaaaata  agaccaacag
1201  catcctcttc  tatggcagat  tctcatcccc  gtagatgcaa  ttagtctgtc  actccatttg
1261  gaaaatgttc  acctgcagat  gttctggtaa  actgattgct  ggcaacaaca  gattctcttg
1321  gctcatatatt  cttttctttc  tcatcttgat  gatgatcgtc  atcatcaaga  atttaatgat
1381  taaaatagca  tgccctttct  tctttctctt  aataagccca  catataaatg  tactttttct
1441  tccagaaaaa  ttctccttga  ggaaaaatgt  ccaaaataag  atgaatcact  taataccgta
1501  tcttctaaat  ttgaaatata  attctgtttg  tgacctgttt  taaatgaacc  aaaccaaata
1561  atactttttc  tttgaattta  gcaacctaga  aacacacatt  tctttgaatt  taggtgatac
1621  ctaaatcctt  cttatgtttc  taaattttgg  gattctataa  aacacatcat  caataaaata
1681  gtgacataaa  atca
```

Fig. 12

SCCA2 cDNA Sequence (SEQ ID NO: 13)

```

1  gggagacaca cacagcctct ctgccacct ctgcttcctc taggaacaca ggagttccag
61  atcacatcga gttcaccatg aattcactca gtgaagccaa caccaagttc atgttcgatc
121 tgttccaaca gttcagaaaa tcaaaagaga acaacatctt ctattcccct atcagcatca
181 catcagcatt agggatggtc ctcttaggag ccaaagacaa cactgcacaa caaattagca
241 aggttcttca ctttgatcaa gtcacagaga acaccacaga aaaagctgca acatatcatg
301 ttgataggtc aggaaatgtt catcaccagt ttcaaaagct tctgactgaa ttcaacaaat
361 cactgatgc  atatgagctg aagatcgcca acaagctctt cggagaaaag acgtatcaat
421 ttttacagga atatttagat gccatcaaga aattttacca gaccagtgtg gaatctactg
481 attttgcaaa tgctccagaa gaaagtcgaa agaagattaa ctctgggtg  gaaagtcaaa
541 cgaatgaaaa aattaaaaac ctatttcctg atgggactat tggcaatgat acgacactgg
601 ttcttgtaga cgcaatctat ttcaaagggc agtgggagaa taaatttaaa aaagaaaaa
661 ctaaagagga aaaattttgg ccaaacaaga atacatacaa atctgtacag atgatgaggc
721 aatacaattc ctttaatttt gccttgctgg aggatgtaca ggccaaggtc ctggaaatac
781 catacaaagg caaagatcta agcatgattg tgctgctgcc aaatgaaatc gatgggtctgc
841 agaagcttga agagaaaactc actgctgaga aattgatgga atggacaagt ttgcagaata
901 tgagagagac atgtgtcgat ttacacttac ctcggttcaa aatggaagag agctatgacc
961 tcaaggacac gttgagaacc atgggaatgg tgaatatctt caatggggat gcagacctct
1021 caggcatgac ctggagccac ggtctctcag tatctaaagt cctacacaag gcctttgtgg
1081 aggtcactga ggagggagtg gaagctgcag ctgccaccgc tgtagtagta gtcgaattat
1141 catctccttc aactaatgaa gagttctgtt gtaatcacc cttctatttc ttcataaggc
1201 aaaataagac caacagcatc ctcttctatg gcagattctc atccccatag atgcaattag
1261 tctgtcactc catttagaaa atgttcacct agaggtgttc tggtaaactg attgctggca
1321 acaacagatt ctcttggtc atatttcttt tctatctcat cttgatgatg atagtcatca
1381 tcaagaattt aatgattaaa atagcatgcc tttctctctt tctcttaata agccacata
1441 taaatgtact tttccttcca gaaaaatttc cttgaggaa aaatgtccaa gataagatga
1501 atcatttaat accgtgtctt ctaaaatttg aatataattc tgtttctgac ctgttttaaa
1561 tgaaccaaac caaatcatatc tttctcttca aatttagcaa cctagaaaca cacatttctt
1621 tgaatttagg tgatacctaa atccttctta tgtttctaaa ttttgatgatt ctataaaaca
1681 catcatcaat aaaataatga ctaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
1741 aaaaaaaaaa aacccaaaaa aaaaaaaaaa aaaaaaaaaa aa

```

Fig. 13

TACSTD1 cDNA Sequence (SEQ ID NO: 14)

```

1  cggcgagcga gcaccttcga cgcggtccgg ggacccccctc gtcgctgtcc tcccgaacgcg
61  gaccgcgctg cccagggcct cgcgctgccc ggccgggtcc tcgtgtccca ctcccggcgc
121 acgccctccc gcgagtcccc ggccccctccc gcgccccctct tctcggcgcg cgcgcagcat
181 ggcgcccccg caggtcctcg cgttcgggct tctgcttgcc gcggcgacgg cgacttttgc
241 cgcagctcag gaagaatgtg tctgtgaaaa ctacaagctg gccgtaaaact gctttgtgaa
301 taataatcgt caatgccagt gtacttcagt tgggtgcacaa aatactgtca tttgctcaaa
361 gctggctgcc aaatgtttgg tgatgaaggc agaaatgaat ggctcaaaac ttgggagaag
421 agcaaaacct gaagggggccc tccagaacaa tgatgggctt tatgatcctg actgcatga
481 gagcgggctc tttaaggcca agcagtgcac cggcacctcc acgtgctggt gtgtgaacac
541 tgctggggtc agaagaacag acaaggacac tgaaataacc tgctctgagc gagtgagaac
601 ctactggatc atcattgaac taaaacacaa agcaagagaa aaaccttatg atagtaaaag
661 tttgcggact gcacttcaga aggagatcac aacgcgttat caactggatc caaaatttat
721 cacgagtatt ttgtatgaga ataatgttat cactattgat ctggttcaaa attccttctca
781 aaaaactcag aatgatgtgg acatagctga tgtggcttat tattttgaaa aagatgttaa
841 aggtgaatcc ttgtttcatt ctaagaaaat ggacctgaca gtaaattggg aacaactgga
901 tctggatcct ggtcaaactt taattttatta tgttgatgaa aaagcacctg aattctcaat
961 gcagggctca aaagctggtg ttattgctgt tattgtggtt gtggtgatag cagttgttgc
1021 tggaattggt gtgctgggta tttccagaaa gaagagaatg gcaaagtatg agaaggctga
1081 gataaaggag atgggtgaga tgcataggga actcaatgca taactatata atttgaagat
1141 tatagaagaa gggaaatagc aaatggacac aaattacaaa tgtgtgtgcg tgggacgaag
1201 acatctttga aggtcatgag tttgttagtt taacatcata tatttgtaat agtgaaacct
1261 gtactcaaaa tataagcagc ttgaaactgg ctttaccaat cttgaaattt gaccacaagt
1321 gtcttatata tgcagatcta atgtaaaatc cagaacttgg actccatcgt taaaattatt
1381 tatgtgtaac attcaaatgt gtgcattaaa tatgcttcca cagtaaaatc tgaaaaactg
1441 atttgtgatt gaaagctgcc tttctattta cttgagtcct gtacatacat acttttttat
1501 gagctatgaa ataaaacatt ttaaactg

```

Fig. 14

TYR cDNA Sequence (SEQ ID NO: 15)

```

1  tattgagttc ttcaaacatt gtagcctctt tatggtctct gagaaataac taccttaaac
61 ccataatctt taatacttcc taaactttct taataagaga agctctattc ctgacactac
121 ctctcatttg caaggtcaaa tcatcattag ttttgtagtc tattaactgg gtttgcttag
181 gtcaggcatt attattacta accttattgt taatattcta accataagaa ttaaactatt
241 aatgggtgaat agagtttttc actttaacat aggcctatcc cactgggtggg atacgagcca
301 attcgaaaga aaagtcagtc atgtgctttt cagaggatga aagcttaaga taaagactaa
361 aagtgtttga tgctggaggt gggagtggtg ttatataggt ctgagccaag acatgtgata
421 atcactgtag tagtagctgg aaagagaaat ctgtgactcc aattagccag ttctgcaga
481 ccttgtaggg actagaggaa gaatgtcctt ggctgttttg tactgcctgc tgtggagttt
541 ccagacctcc gctggccatt tccctagagc ctgtgtctcc tctaagaacc tgatggagaa
601 ggaatgctgt ccaccgtgga ggggggacag gaggccctgt ggccagcttt caggcagagg
661 ttctgtcag aatatccttc tgtccaatgc accacttggg cctcaatttc ccttcacagg
721 ggtggatgac cgggagtcgt ggccttcctt cttttataat aggacctgcc agtgctctgg
781 caacttcatg ggattcaact gtggaaactg caagtttggc ttttggggac caaactgcac
841 agagagacga ctcttggtga gaagaaacat cttcgatttg agtgccccag agaaggacaa
901 attttttgcc tacctcactt tagcaaagca taccatcagc tcagactatg tcatccccat
961 agggacctat ggccaaatga aaaatggatc aacacccatg ttttaacgaca tcaatattta
1021 tgacctcttt gtctggatgc attattatgt gtcaatggat gcactgcttg ggggatctga
1081 aatctggaga gacattgatt ttgccatga agcaccagct tttctgcctt ggcatagact
1141 cttcttggtg cgggtgggaa aagaaatcca gaagctgaca ggagatgaaa acttcactat
1201 tccatattgg gactggcggg atgcagaaaa gtgtgacatt tgcacagatg agtacatggg
1261 aggtcagcac ccacaaatc ctaacttact cagcccagca tcattcttct cctcttgga
1321 gattgtctgt agccgatttg aggagtacaa cagccatcag tctttatgca atggaacgcc
1381 cgagggacct ttacggcgta atcctggaaa ccatgacaaa tccagaacct caaggctccc
1441 ctcttcagct gatgtagaat ttgacctgag tttagcccaa tatgaatctg gttccatgga
1501 taaagctgcc aatttcagct ttagaaatac actggaagga tttgctagtc cacttactgg
1561 gatagcggat gcctctcaaa gcagcatgca caatgccttg cacatctata tgaatggaac
1621 aatgtcccag gtacagggat ctgccaacga tcctatcttc cttcttcacc atgcatttgt
1681 tgacagtatt tttgagcagt ggctccgaag gcaccgtcct cttcaagaag tttatccaga
1741 agccaatgca cccattggac ataaccggga atcctacatg gttcctttta taccactgta
1801 cagaaatggt gatctcttta tttcatccaa agatctgggc tatgactata gctatctaca
1861 agattcagac ccagactctt ttcaagacta cattaagtcc tatttgaac aagcgagtcg

```

Fig. 15/1

1921 gatctgggtca tggctccttg gggcggcgat ggtagggggcc gtcctcactg ccctgctggc
1981 agggcttgtg agcttgctgt gtcgtcacia gagaaagcag cttcctgaag aaaagcagcc
2041 actcctcatg gagaaagagg attaccacag cttgtatcag agccatttat aaaaggctta
2101 ggcaatagag tagggccaaa aagcctgacc tcactctaac tcaaagtaat gtccagggtc
2161 ccagagaata tctgctggta tttttctgta aagaccattt gcaaaattgt aacctaatac
2221 aaagtgtagc cttcttccaa ctcaggtaga acacacctgt ctttgtcttg ctgttttcac
2281 tcagcccttt taacattttc ccctaagccc atatgtctaa ggaaaggatg ctatttggtta
2341 atgaggaact gttatttgta tgtgaattaa agtgctctta tttt

Fig. 15/2

VIL1 cDNA Sequence (SEQ ID NO: 16)

```

1  cattctcccc caggctcact caccatgacc aagctgagcg cccaagtcaa aggctctctc
61  aacatcacca ccccggggct gcagatatgg aggatcgagg ccatgcagat ggtgcctgtt
121 ccttccagca cctttggaag cttcttcgat ggtgactgct acatcatcct ggctatccac
181 aagacagcca gcagcctgtc ctatgacatc cactactgga ttggccagga ctcatccctg
241 gatgagcagg gggcagctgc catctacacc acacagatgg atgacttcct gaagggccgg
301 gctgtgcagc accgcgaggt ccagggcaac gagagcgagg ccttccgagg ctacttcaag
361 caaggccttg tgatccggaa agggggcgtg gcttctggca tgaagcacgt ggagaccaac
421 tcctatgacg tccagagggt gctgcatgtc aagggcaaga ggaacgtggt agctggagag
481 gtagagatgt cctggaagag tttcaaccga ggggatgttt tcctcctgga ccttgggaag
541 cttatcatcc agtggaatgg accggaaagc acccgatagg agagactcag gggcatgact
601 ctggccaagg agatccgaga ccaggagcgg ggagggcgca cctatgtagg cgtggtggac
661 ggagagaatg aattggcatc cccgaagctg atggaggtga tgaaccacgt gctgggcaag
721 cgaggggagc tgaaggcggc cgtgcccgac acggtggtgg agccggcact caaggctgca
781 ctcaaactgt accatgtgtc tgactccgag gggaatctgg tggtagggga agtcgccaca
841 cggccactga cacaggacct gctcagtcac gaggactgtt acatcctgga ccaggggggc
901 ctgaagatct acgtgtggaa agggaagaaa gccaatgagc aggagaagaa gggagccatg
961 agccatgcgc tgaacttcat caaagccaag cagtaccac caagcacaca ggtggagggtg
1021 cagaatgatg gggctgagtc ggccgtcttt cagcagctct tccagaagtg gacagcgtcc
1081 aaccggacct caggcctagg caaaaccac actgtgggct ccgtggccaa agtggaacag
1141 gtgaagtctg atgccacatc catgcatgtc aagcctcagg tggctgcccc gcagaagatg
1201 gtagatgatg ggagtgggga agtgcagggt tggcgcatgt agaacctaga gctggtacct
1261 gtggattcca agtggctagg ccacttctat gggggcgact gctacctgct gctctacacc
1321 tacctcatcg gcgagaagca gcattacctg ctctacgttt ggcagggcag ccaggccagc
1381 caagatgaaa ttacagcatc agcttatcaa gccgtcatcc tggaccagaa gtacaatggt
1441 gaaccagtcc agatccgggt cccaatgggc aaggagccac ctcatcttat gtccatcttc
1501 aaggggacga tgggtggtcta ccagggaggc acctcccgaa ctaacaactt ggagaccggg
1561 ccctccacac ggctgttcca ggtccaggga actggcgcca acaacaccaa ggcctttgag
1621 gtcccagcgc gggccaatct cctcaattcc aatgatgtct ttgtcctcaa gacctagtct
1681 tgctgctatc tatggtgtgg gaagggttgt agcggggacg agcgggagat ggccaagatg
1741 gttgctgaca ccattctccg gacggagaag caagtgggtg tgggaaggga ggagccagcc
1801 aacttctgga tggccctggg tgggaaggcc ccctatgcca acaccaagag actacaggaa

```

Fig. 16/1

1861 gaaaacctgg tcatcacccc ccggctcttt gagtggtcca acaagactgg gcgcttcctg
1921 gccacagaga tccctgactt caatcaggat gacttggaag aggatgatgt gttcctacta
1981 gatgtctggg accaggtctt cttctggatt gggaaacatg ccaacgagga ggagaagaag
2041 gccgcagcaa ccactgcaca ggaatacctc aagacccatc ccagcgggcg tgaccctgag
2101 acccccatca ttgtggtgaa gcaggacac gagccccca ccttcacagg ctggttcctg
2161 gcttgggatc ctttcaagtg gagtaacacc aaatcctatg aggacctgaa ggcggagtct
2221 ggcaacctta gggactggag ccagatcact gctgaggtca caagcccaa agtggacgtg
2281 ttcaatgcta acagcaacct cagttctggg cctctgcca tcttccccct ggagcagcta
2341 gtgaacaagc ctgtagagga gctccccgag ggtgtggacc ccagcaggaa ggaggaacac
2401 ctgtccattg aagatttcac tcaggccttt gggatgactc cagctgcctt ctctgctctg
2461 cctcgatgga agcaacaaaa cctcaagaaa gaaaaaggac tattttgaga agagtagctg
2521 tggttgtaaa gcagtaccct accctgattg tagggctca ttttctcacc gatattagtc
2581 ctacaccaat tgaagtgaat ttttgcagat gtgcctatga gcacaaactt ctgtggcaaa
2641 tgccagtttt gtttaataat gtacctatc cttcagaaag atgatacccc aaaaaaaaaa
2701 aa

Fig. 16/2

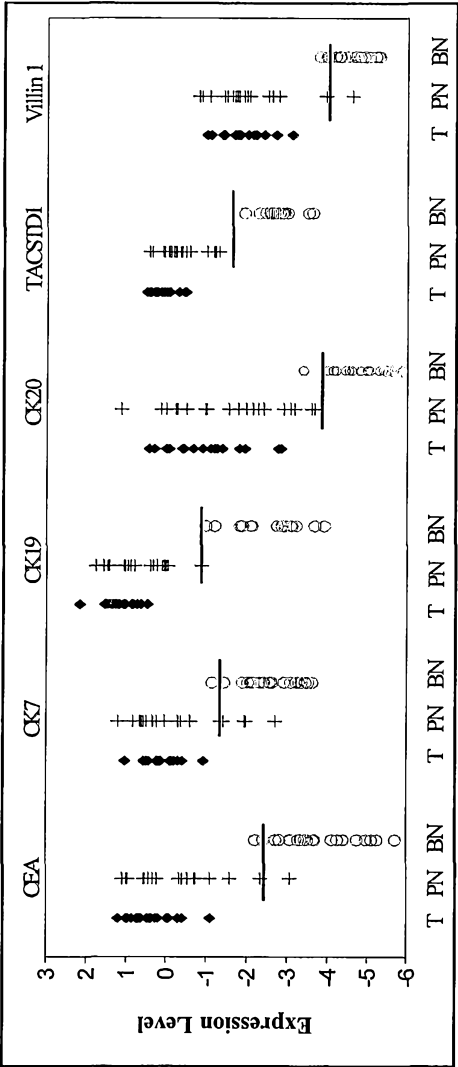


FIG. 17

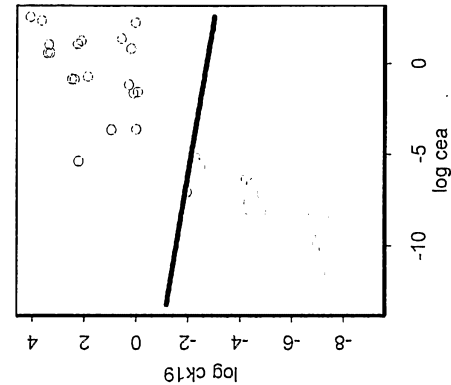


FIG. 18C

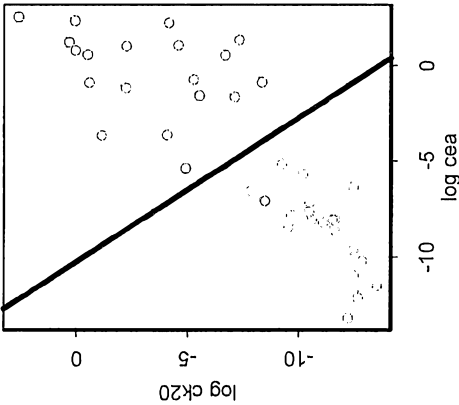


FIG. 18B

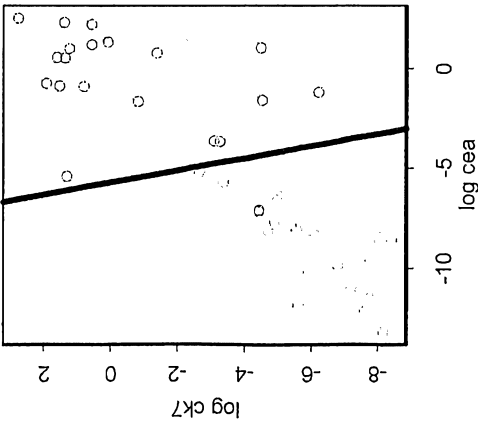


FIG. 18A

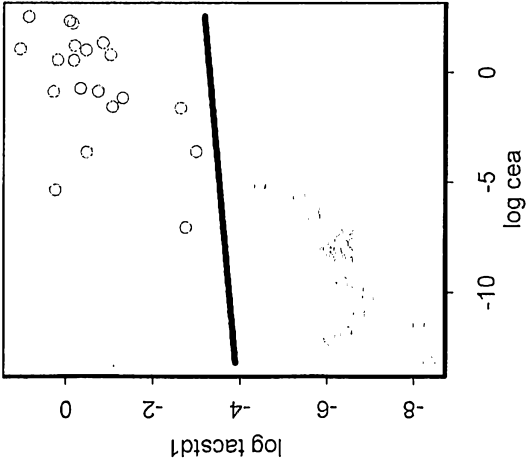


FIG. 18D

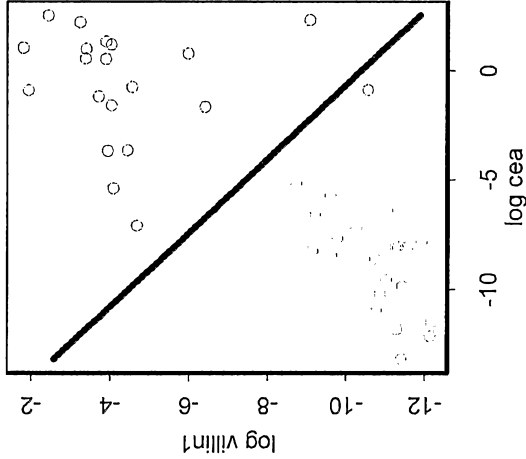


FIG. 18E

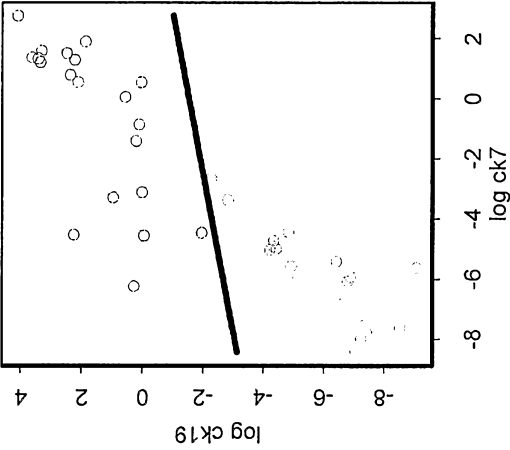


FIG. 18F

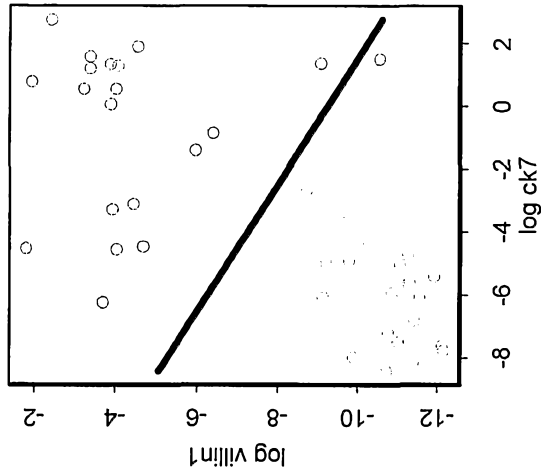


FIG. 18I

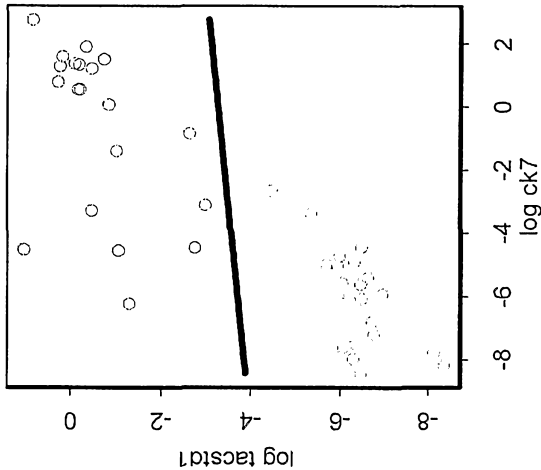


FIG. 18H

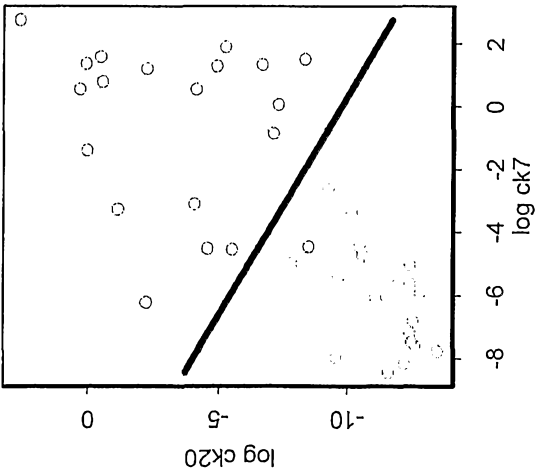


FIG. 18G

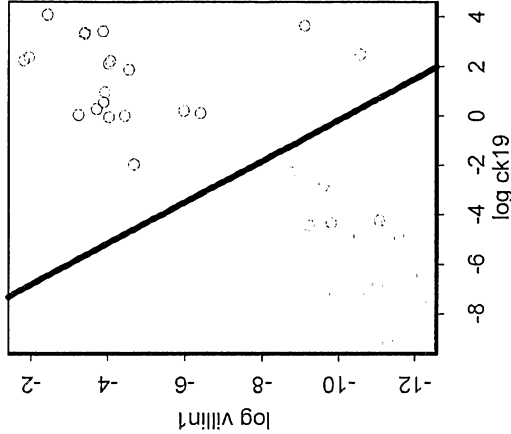


FIG. 18L

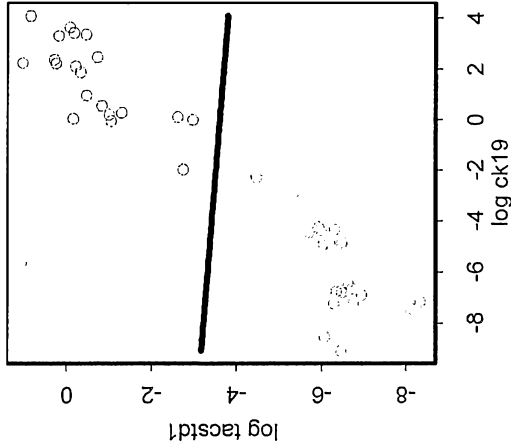


FIG. 18K

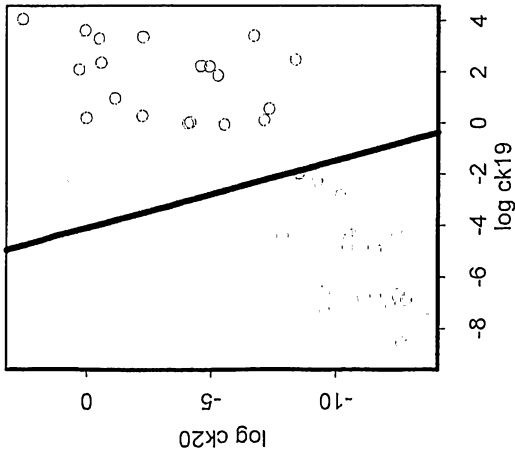


FIG. 18J

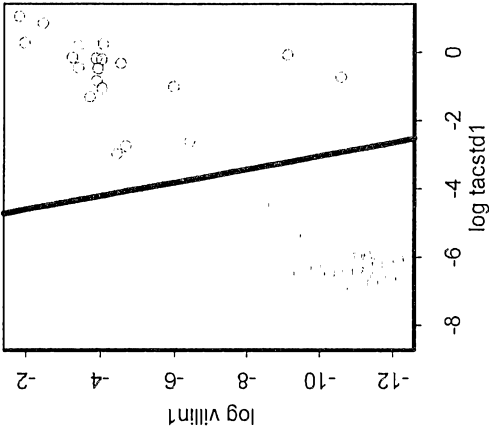


FIG. 18O

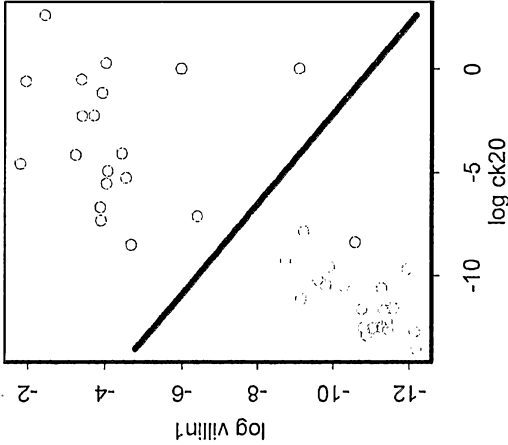


FIG. 18N

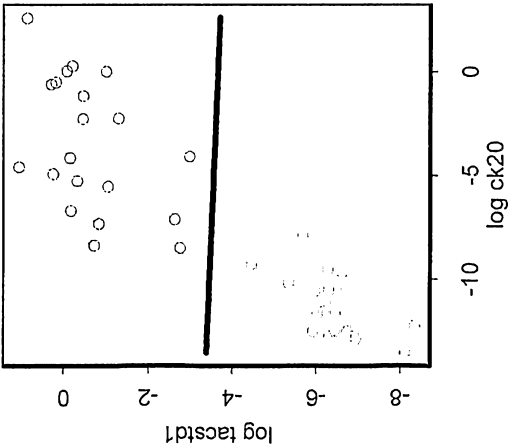


FIG. 18M

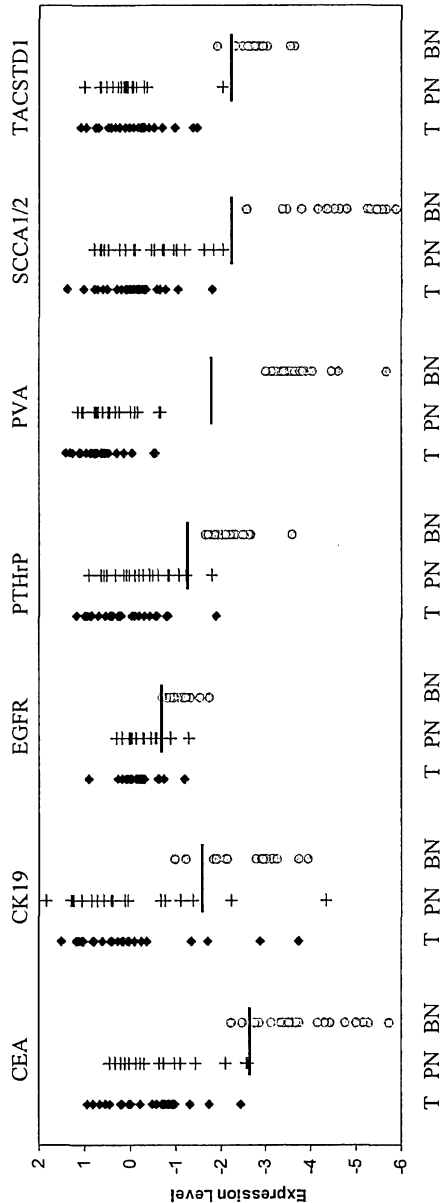


FIG. 19

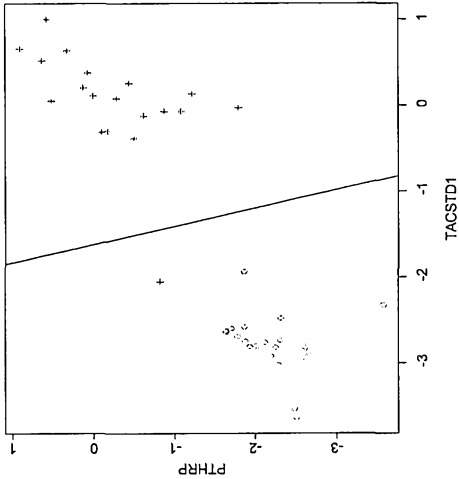


FIG. 20C

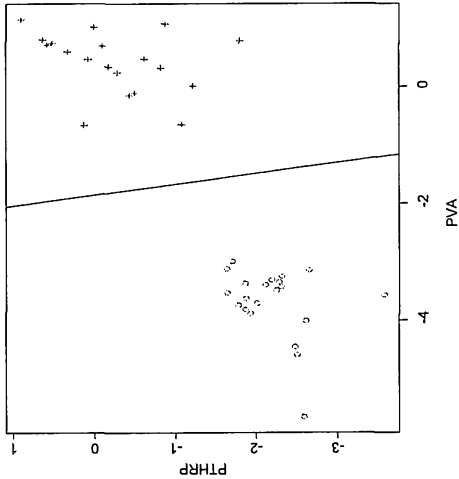


FIG. 20B

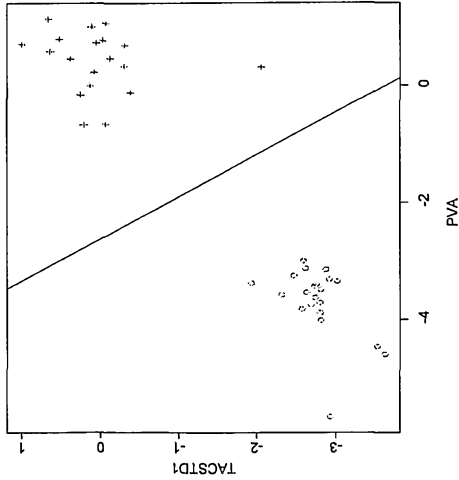


FIG. 20A

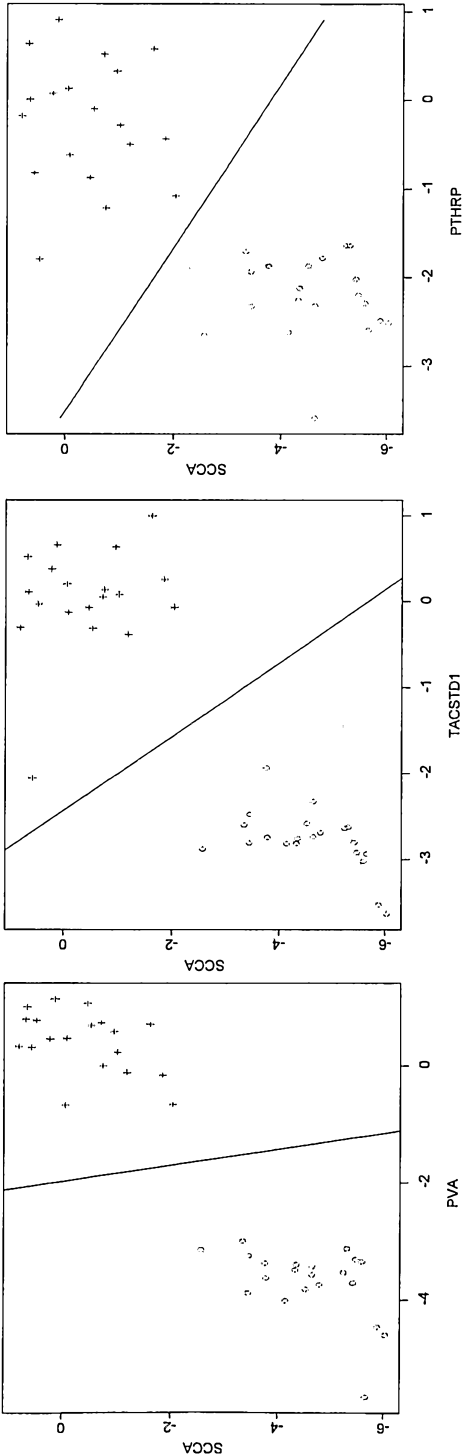


Fig. 20F

Fig. 20E

Fig. 20D

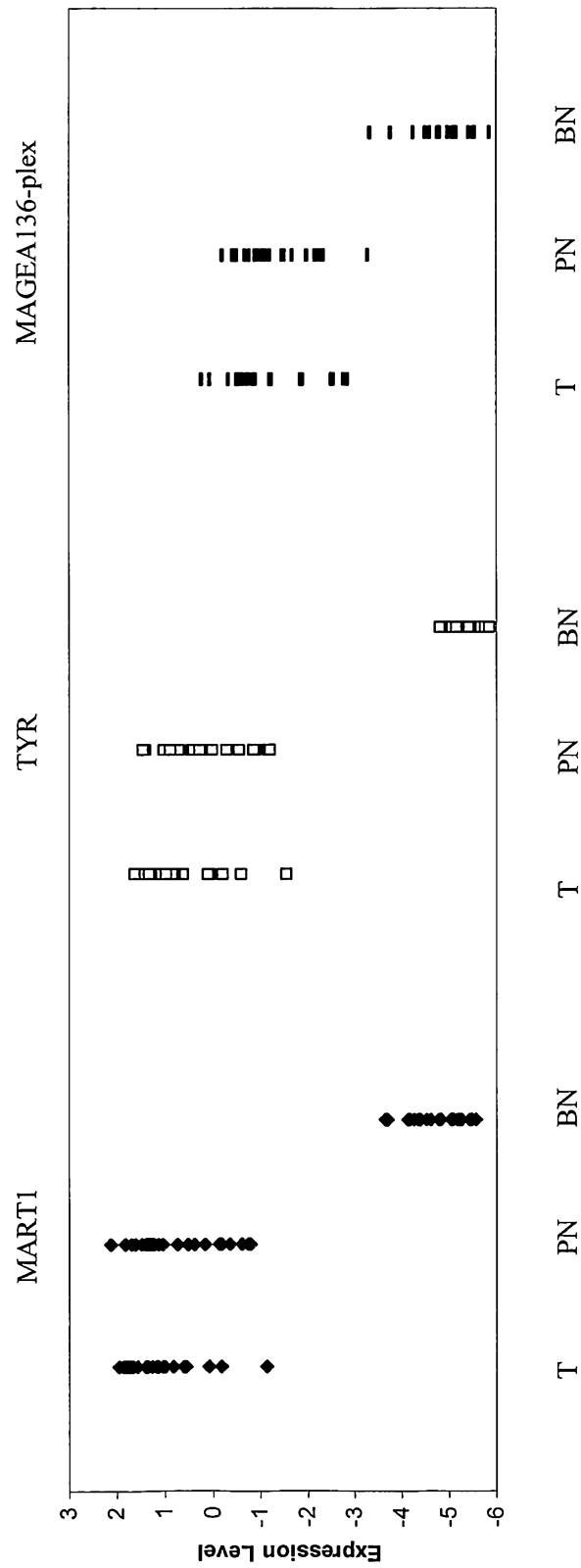


FIG. 21

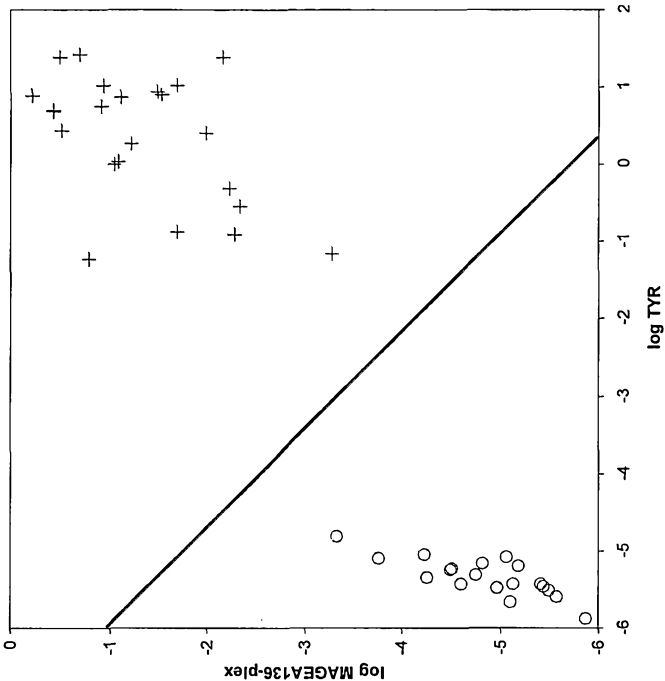


FIG. 22B

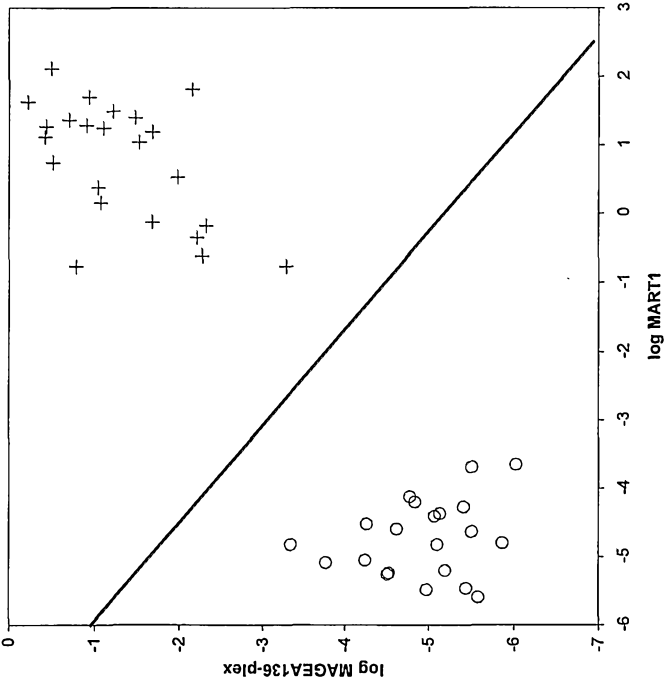


FIG. 22A

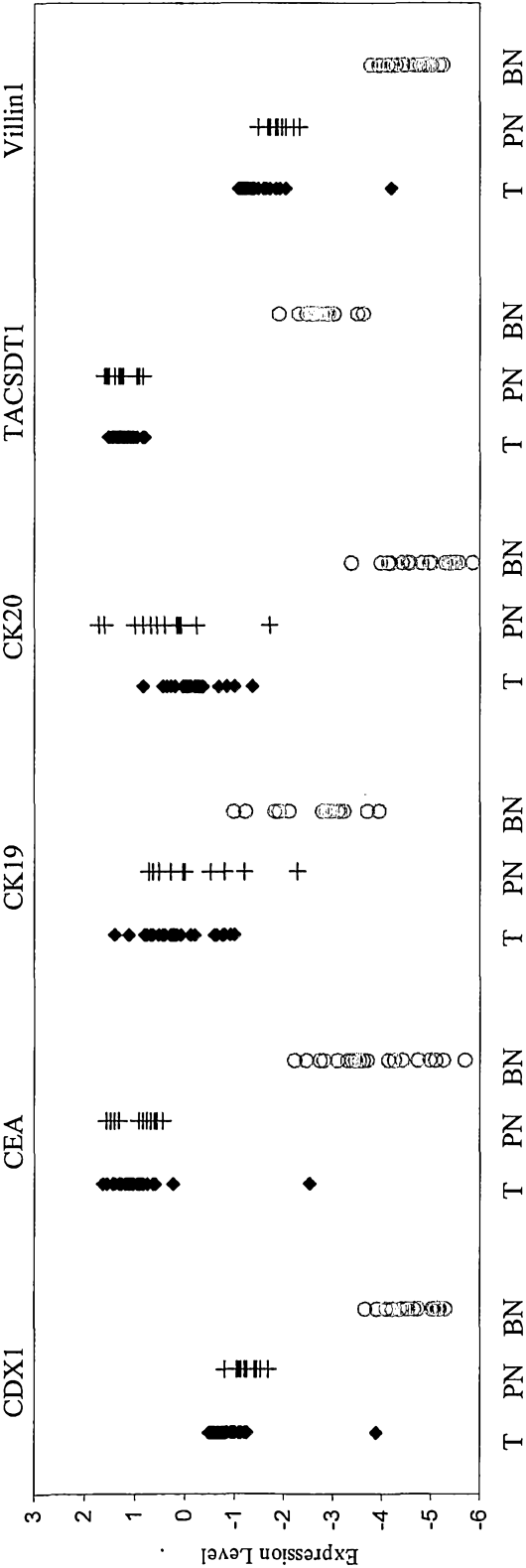


FIG. 23

SEQUENCE LISTING

<110> University of Pittsburgh - of The Commonwealth System of
Higher Education
Godfrey, Tony
Hughes, Steven
Xi, Liqiang
Gooding, William E
Raja, Siva E

<120> Identification of Markers in Esophageal Cancer, Colon Cancer,
Head and Neck Cancer and Melanoma

<130> 030160

<150> 60/586,599
<151> 2004-07-09

<150> 60/587,019
<151> 2004-07-09

<160> 40

<170> PatentIn version 3.3

<210> 1
<211> 1699
<212> DNA
<213> Homo sapiens

<400> 1
aggtgagcgg ttgctcgtcg tcggggcggc cggcagcggc ggctccaggg ccagcatgc 60
gcgggggacc ccgcggccac catgtatgtg ggctatgtgc tggacaagga ttcgcccgtg 120
taccocggcc cagccaggcc agccagcctc ggcctgggcc cggcaaacta cggccccccg 180
gccccgcccc cggcgcccc gcagtacccc gacttctcca gctactctca cgtggagccg 240
gccccgcgc ccccgacggc ctggggggcg cccttcctg cggccaagga cgactgggcc 300
gccgcctacg gccggggccc cgcggcccct gccgccagcc cagcttcgct ggcatcggg 360
ccccctccag actttagccc ggtgccggcg cccctgggc cgggccggg cctcctggcg 420
cagcccctcg ggggcccggg cacaccgtcc tcgccggag cgcagaggcc gacgccctac 480
gagtggatgc ggcgcagcgt ggcggccgga ggcggcgggt gcagcggtaa gactcggacc 540
aaggacaagt accgcgtggt ctacaccgac caccaacgcc tggagctgga gaaggagttt 600
cattacagcc gttacatcac aatccggcg aaatcagagc tggctgcaa tctggggctc 660
actgaacggc aggtgaagat ctggttccaa aaccggcggg caaaggagcg caaagtgaac 720
aagaagaaac agcagcagca acagccccc cagccgccga tggcccacga catcacggcc 780
acccagccg ggccatccct ggggggcctg tgtcccagca acaccagcct cctggccacc 840
tcctctccaa tgctgtgaa agaggagttt ctgccatagc cccatgccca gcctgtgcgc 900

cgggggacct ggggactcgg gtgctgggag tgtggctcct gtgggcccag gaggtctggt 960
 ccgagtctca gccctgacct tctgggacat ggtggacagt cacctatcca ccctctgcat 1020
 ccccttggcc cattgtgtgc agtaagcctg ttggataaag accttccagc tcctgtgttc 1080
 tagacctctg ggggataagg gagtccaggg tggatgatct caatctcccg tgggcatctc 1140
 aagccccaaa tggttggggg aggggcctag acaaggctcc agggcccacc tcctcctcca 1200
 tacgttcaga ggtgcagctg gaggcctgtg tggggaccac actgatcctg gagaaaaggg 1260
 atggagctga aaaagatgga atgcttgacg agcatgacct gaggagggag gaacgtggtc 1320
 aactcacacc tgcctcttct gcagcctcac ctctacctgc ccccatcata agggcactga 1380
 gcccttccca ggctggatac taagcacaaa gcccatagca ctgggctctg atggctgctc 1440
 cactgggtta cagaatcaca gccctcatga tcattctcag tgagggtctt ggattgagag 1500
 ggaggccctg ggaggagaga agggggcaga gtcttcccta ccaggtttct acacccccgc 1560
 caggctgccc atcaggggccc agggagcccc cagaggactt tattcggacc aagcagagct 1620
 cacagctgga caggtgttgt atatagagtg gaatctcttg gatgcagctt caagaataaa 1680
 tttttcttct cttttcaaa 1699

<210> 2

<211> 2974

<212> DNA

<213> Homo sapiens

<400> 2

ctcagggcag agggaggaag gacagcagac cagacagtca cagcagcctt gacaaaacgt 60
 tcctggaact caagctcttc tccacagagg aggacagagc agacagcaga gaccatggag 120
 tctccctcgg ccctcccca cagatgggtg atccccctgg agaggctcct gctcacagcc 180
 tcacttctaa ctttctggaa cccgcccacc actgccaagc tcactattga atccacgccg 240
 ttcaatgtcg cagaggggaa ggaggtgctt ctacttgctc acaatctgcc ccagcatctt 300
 tttggctaca gctggtacaa aggtgaaaga gtggatggca accgtcaa at tataggatat 360
 gtaataggaa ctcaacaagc taccacaggg ccgcataca gtggctcgaga gataatatac 420
 cccaatgcat ccctgctgat ccagaacatc atccagaatg acacaggatt ctacacccta 480
 cacgtcataa agtcagatct tgtgaatgaa gaagcaactg gccagttccg ggtatacccg 540
 gagctgcca agccctccat ctccagcaac aactccaaac ccgtggagga caaggatgct 600
 gtggccttca cctgtgaacc tgagactcag gacgcaacct acctgtggtg ggtaaacaat 660
 cagagcctcc cgtcagctcc caggctgcag ctgtccaatg gcaacaggac cctcactcta 720
 ttcaatgtca caagaaatga cacagcaagc tacaaatgtg aaaccagaa cccagtgagt 780


```

gccaggcgca gtgattcagt catcctgaat gtctcttatg gcccggatgc cccaccatt      840
tcccctctaa acacatctta cagatcaggg gaaaatctga acctctcttg ccacgcagcc      900
tctaaccacac ctgcacagta ctcttggttt gtcaatggga ctttccagca atccacccaa      960
gagctcttta tccccaacat cactgtgaat aatagtggat cctatacgtg ccaagcccat     1020
aactcagaca ctggcctcaa taggaccaca gtcacgacga tcacagtcta tgcagagcca     1080
cccaaaccct tcatcaccag caacaactcc aaccccgtagg aggatgagga tgctgtagcc     1140
ttaacctgtg aacctgagat tcagaacaca acctacctgt ggtgggtaaa taatcagagc     1200
ctcccggtca gtcccaggct gcagctgtcc aatgacaaca ggaccctcac tctactcagt     1260
gtcacaagga atgatgtagg accctatgag tgtggaatcc agaacgaatt aagtgttgac     1320
cacagcgacc cagtcctcct gaatgtcttc tatggcccag acgacccac catttcccc     1380
tcatacacct attaccgtcc aggggtgaac ctgagcctct cctgccatgc agcctctaac     1440
ccacctgcac agtattcttg gctgattgat gggaacatcc agcaacacac acaagagctc     1500
tttatctcca acatcactga gaagaacagc ggactctata cctgccaggc caataactca     1560
gccagtggcc acagcaggac tacagtcaag acaatcacag tctctgcgga gctgccaag     1620
ccctccatct ccagcaacaa ctccaaaccc gtggaggaca aggatgctgt ggcttcacc     1680
tgtgaacctg aggctcagaa cacaacctac ctgtggtggg taaatggtca gagcctcca     1740
gtcagtccca ggctgcagct gtccaatggc aacaggaccc tcaactctatt caatgtcaca     1800
agaaatgacg caagagccta tgtatgtgga atccagaact cagtgagtgc aaaccgcagt     1860
gaccagtcac ccctggatgt cctctatggg ccggacaccc ccatcatttc cccccagac     1920
tcgtcttacc ttccgggagc gaacctcaac ctctcctgcc actcggcctc taaccatcc     1980
ccgcagtatt cttggcgtat caatgggata ccgcagcaac acacacaagt tctctttatc     2040
gccaaaatca cgccaaataa taacgggacc tatgcctgtt ttgtctctaa cttggctact     2100
ggccgcaata attccatagt caagagcatc acagtctctg catctggaac ttctcctggt     2160
ctctcagctg gggccactgt cggcatcatg attggagtgc tggttggggg tgctctgata     2220
tagcagccct ggtgtagttt cttcatttca ggaagactga cagttgtttt gcttcttctc     2280
taaagcattt gcaacagcta cagtctaaaa ttgcttcttt accaaggata ttacagaaa     2340
agactctgac cagagatcga gaccatccta gccaacatcg tgaaaccca tctctactaa     2400
aaatacaaaa atgagctggg cttgggtggc cgcacctgta gtcccagtta ctcgggaggc     2460
tgaggcagga gaatcgcttg aacccgggag gtggagattg cagtgagccc agatcgaccc     2520
actgcactcc agtctggcaa cagagcaaga ctccatctca aaaagaaaag aaaagaagac     2580
tctgacctgt actcttgaat acaagtttct gataccactg cactgtctga gaatttccaa     2640

```

aactttaatg aactaactga cagcttcatg aaactgtcca ccaagatcaa gcagagaaaa 2700
taattaatth catgggacta aatgaactaa tgaggattgc tgattcttta aatgtcttgt 2760
ttcccagatt tcaggaaact ttttttcttt taagctatcc actcttacag caatttgata 2820
aaatatactt ttgtgaacaa aaattgagac atttacatth tctccctatg tggctgctcc 2880
agacttggga aactattcat gaatatttat attgtatggt aatatagtta ttgcacaagt 2940
tcaataaaaa tctgctcttt gtataacaga aaaa 2974

<210> 3
<211> 1753
<212> DNA
<213> Homo sapiens

<400> 3
cagccccgcc cctacctgtg gaagcccagc cgcccgctcc cgccgataaa aggtgcggag 60
tgtccccgag gtcagcgagt gcgcgctcct cctcgccgc cgctagggtcc atccccggcc 120
agccaccatg tccatccact tcagctcccc ggtattcacc tcgcgctcag ccgccttctc 180
gggcccgggc gccaggtgc gcctgagctc cgctcgcccc ggccgccttg gcagcagcag 240
cctctacggc ctccggcgcct cgcggccgcg cgtggccgtg cgtctgcct atgggggccc 300
ggtgggcgcc ggcatccgcg aggtcaccat taaccagagc ctgctggccc cgctgcggct 360
ggacgcgcac ccctccctcc agcgggtgcg ccaggaggag agcgagcaga tcaagaccct 420
caacaacaag tttgcctcct tcacgcacaa ggtgcggttt ctggagcagc agaacaagct 480
gctggagacc aagtggacgc tgctgcagga gcagaagtcg gccaaagca gccgcctccc 540
agacatcttt gaggcccaga ttgctggcct tcggggtcag cttgaggcac tgcaggtgga 600
tgggggccgc ctggaggcgc agctgcggag catgcaggat gtggtggagg acttcaagaa 660
taagtacgaa gatgaaatta accgccgcac agctgctgag aatgagtttg tgggtgctgaa 720
gaaggatgtg gatgctgcct acatgagcaa ggtggagctg gaggccaagg tggatgccct 780
gaatgatgag atcaacttcc tcaggaccct caatgagacg gagttgacag agctgcagtc 840
ccagatctcc gacacatctg tgggtgctgtc catggacaac agtcgctccc tggacctgga 900
cgccatcatc gctgagggtca aggcacagta tgaggagatg gccaaatgca gccgggctga 960
ggctgaagcc tgggtaccaga ccaagtttga gaccctccag gccaggctg ggaagcatgg 1020
ggacgacctc cggaataccc ggaatgagat ttcagagatg aaccgggcca tccagaggct 1080
gcaggctgag atcgacaaca tcaagaacca gcgtgccaag ttggaggccg ccattgccga 1140
ggctgaggag cgtggggagc tggcgctcaa ggatgctcgt gccaaagcagg aggagctgga 1200
agccgccctg cagcgggcca agcaggatat ggcacggcag ctgcgtgagt accaggaact 1260

```

catgagcgtg aagctggccc tggacatcga gatcgccacc taccgcaagc tgctggaggg 1320
cgaggagagc cggttggctg gagatggagt gggagccgtg aatatctctg tgatgaattc 1380
cactgggtggc agtagcagtg gcggtggcat tgggctgacc ctcgggggaa ccatgggcag 1440
caatgccctg agcttctcca gcagtgcggg tcttgggctc ctgaaggctt attccatccg 1500
gacgcgcatcc gccagtgcga ggagtgcccg cgactgagcc gcctcccacc actccactcc 1560
tccagccacc acccacaatc acaagaagat tcccaccctt gcctcccatg cctgggtccca 1620
agacagtgag acagtctgga aagtgatgtc agaatagctt ccaataaagc agcctcatte 1680
tgaggcctga gtgatccacg tgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1740
aaaaaaaaaa aaa 1753

```

```

<210> 4
<211> 1440
<212> DNA
<213> Homo sapiens

```

```

<400> 4
cgccccctgac accattcctc ccttcccccc tccaccggcc gcgggcataa aaggcgccag 60
gtgagggcct cgccgctcct ccgcggaatc gcagcttctg agaccagggt tgctccgtcc 120
gtgctccgcc tcgcatgac ttctacagc tatcgccagt cgtcgccac gtcgtccttc 180
ggaggcctgg gcggcggtc cgtgcgtttt gggccggggg tcgcctttcg cgcgcccagc 240
attcacgggg gctccggcgg ccgcggcgta tccgtgtcct ccgcccgctt tgtgtcctcg 300
tctcctcggg gggcctacgg cggcggctac ggcgcgctcc tgaccgcgtc cgacgggctg 360
ctggcgggca acgagaagct aaccatgcag aacctcaacg accgcctggc ctccctacctg 420
gacaagggtgc gcgccctgga ggcggccaac ggcgagctag aggtgaagat ccgcgactgg 480
taccagaagc aggggcctgg gccctccgc gactacagcc actactacac gaccatccag 540
gacctgcggg acaagattct tgggtgccacc attgagaact ccaggattgt cctgcagatc 600
gacaatgccc gtctggctgc agatgacttc cgaaccaagt ttgagacgga acaggctctg 660
cgcatgagcg tggaggccga catcaacggc ctgcgcaggg tgctggatga gctgaccctg 720
gccaggaccg acctggagat gcagatcgaa ggctgaagg aagagctggc ctacctgaag 780
aagaaccatg aggaggaaat cagtacgtg aggggccaag tgggaggcca ggtcagtgtg 840
gaggtggatt ccgtccggg caccgatctc gccaaगतcc tgagtacat gcgaagccaa 900
tatgaggtca tggccgagca gaaccggaag gatgctgaag cctgggtcac cagccggact 960
gaagaattga accgggaggt cgctggccac acggagcagc tccagatgag caggtccgag 1020
gttactgacc tgcggcgcac ccttcagggt cttgagattg agctgcagtc acagctgagc 1080

```

atgaaagctg ccttgaaga cacactggca gaaacggagg cgcgctttgg agcccagctg 1140
 gcgcataatcc aggcgctgat cagcgggtatt gaagcccagc tgggcgatgt gcgagctgat 1200
 agtgagcggc agaatacagga gtaccagcgg ctcatggaca tcaagtcgcg gctggagcag 1260
 gagattgccca cctaccgcag cctgctcgag ggacaggaag atcactacaa caatttgtct 1320
 gcctccaagg tcctctgagg cagcaggctc tggggcttct gctgtccttt ggagggtgtc 1380
 ttctgggtag agggatggga aggaagggac ccttaccccc ggctcttctc ctgacctgcc 1440

<210> 5
 <211> 1817
 <212> DNA
 <213> Homo sapiens

<400> 5
 caaccatcct gaagctacag gtgtccctc ctggaatctc caatggattt cagtcgcaga 60
 agcttccaca gaagcctgag ctctccttg caggcccctg tagtcagtag agtgggcatg 120
 cagcgcctcg ggacgacacc cagcgtttat gggggtgctg gaggccgggg catccgcatac 180
 tccaactcca gacacacggt gaactatggg agcgatctca caggcggcgg ggacctgttt 240
 gttggcaatg agaaaatggc catgcagaac ctaaatagacc gtctagcgag ctacctagaa 300
 aaggtgcgga ccctggagca gtccaactcc aaacttgaag tgcaaataca gcagtggtag 360
 gaaaccaacg ccccgagggc tggtcgcgac tacagtgcac attacagaca aattgaagag 420
 ctgcgaagtc agattaagga tgctcaactg caaaatgtc ggtgtgtcct gcaaattgat 480
 aatgctaaac tggctgctga ggacttcaga ctgaagtatg agactgagag aggaatacgt 540
 ctaacagtgg aagctgatct ccaaggcctg aataaggtct ttgatgacct aaccctacat 600
 aaaacagatt tggagattca aattgaagaa ctgaataaag acctagctct cctcaaaaag 660
 gagcatcagg aggaagtcga tggcctacac aagcatctgg gcaacactgt caatgtggag 720
 gttgatgctg ctccaggcct gaaccttggc gtcacatga atgaaatgag gcagaagtat 780
 gaagtcatgg ccagaagaa ccttcaagag gccaaagaac agtttgagag acagactgca 840
 gttctgcagc aacagggtcac agtgaatact gaagaattaa aaggaactga ggttcaacta 900
 acggagctga gacgcacctc ccagagcctt gagatagaac tccagtccca tctcagcatg 960
 aaagagtctt tggagcacac tctagaggag accaaggccc gttacagcag ccagttagcc 1020
 aacctccagt cgctgttgag ctctctggag gcccaactga tgcagattcg gagtaacatg 1080
 gaacgccaga acaacgaata ccatatcctt cttgacataa agactcgact tgaacaggaa 1140
 attgctactt accgccgcct tctggaagga gaagacgtaa aaactacaga atatcagtta 1200
 agcaccctgg aagagagaga tataaagaaa accaggaaga ttaagacagt cgtgcaagaa 1260

gtagtggatg gcaaggtcgt gtcactctgaa gtcaaagagg tggaagaaaa tatctaaata 1320
 gctaccagaa ggagatgctg ctgagggtttt gaaagaaatt tggctataat cttatctttg 1380
 ctccctgcaa gaaatcagcc ataagaaagc actattaata ctctgcagtg attagaaggg 1440
 gtgggggtggc gggaatccta tttatcagac tctgtaattg aatataaatg ttttactcag 1500
 aggagctgca aattgcctgc aaaaatgaaa tccagtgagc actagaatat ttaaaacatc 1560
 attactgcc a tctttatcat gaagcacatc aattacaagc tgtagaccac ctaatatcaa 1620
 tttgtaggta atgttcctga aaattgcaat acatttcaat tataactaac ctcaaaaagt 1680
 agaggaatcc atgtaaattg caaataaacc actttctaata tttttcctgt ttctgaaaaa 1740
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800
 aaaaaaaaaa aaaaaaa 1817

<210> 6
 <211> 1722
 <212> DNA
 <213> Homo sapiens

<400> 6
 cgtagagttc ggccgaagga acctgaccca ggctctgtga ggaggcaagg ttttcagggg 60
 acaggccaac ccagaggaca ggattccctg gaggccacag aggagcacca aggagaagat 120
 ctgcctgtgg gtcttcattg ccagctcct gccacactc ctgcctgctg ccctgacgag 180
 agtcatcatg tctcttgagc agaggagtct gactgcaag cctgaggaag cccttgaggc 240
 ccaacaagag gccctgggcc tgggtgtgtgt gcaggctgcc gcctcctcct cctctcctct 300
 ggtcctgggc accctggagg aggtgcccac tgctgggtca acagatcctc ccagaggtcc 360
 tcagggagcc tccgcctttc ccactacat caacttact cgacagaggc aaccagtgta 420
 gggttccagc agccgtgaag aggagggggc aagcacctct tgtatcctgg agtccttggt 480
 ccgagcagta atcactaaga aggtggctga tttggttggg tttctgctcc tcaaatatcg 540
 agccaggag ccagtcacaa aggcagaaat gctggagagt gtcacaaaa attacaagca 600
 ctgttttcct gagatcttcg gcaaagcctc tgagtccttg cagctggtct ttggcattga 660
 cgtgaaggaa gcagacccca ccggccactc ctatgtcctt gtcacctgcc taggtctctc 720
 ctatgatggc ctgctgggtg ataatcagat catgcccaag acaggcttcc tgataattgt 780
 cctggtcattg attgcaatgg agggcgcca tgctcctgag gaggaatct gggaggagct 840
 gagtgtgatg gaggtgtatg atgggagga gcacagtgcc tatggggagc ccaggaagct 900
 gctcacccaa gatttggtgc aggaaaagta cctggagtac cggcaggtgc cggacagtga 960
 tccgcacgc tatgagttcc tgtgggggtcc aagggccctt gctgaaacca gctatgtgaa 1020

```

agtccttgag tatgtgatca aggtcagtg c aagagttcgc tttttcttcc catccctgcg 1080
tgaagcagct ttgagagagg aggaagaggg agtctgagca tgagttgcag ccagggccag 1140
tgggaggggg actggggccag tgcaccttcc agggccgcgt ccagcagctt cccctgcctc 1200
gtgtgacatg aggcccatc ttcactctga agagagcggg cagtgttctc agtagtaggt 1260
ttctgttcta ttgggtgact tggagattta tctttgttct cttttggaat tgttcaaagt 1320
ttttttttta agggatgggt gaatgaactt cagcatccaa gtttatgaat gacagcagtc 1380
acacagttct gtgtatatag ttttaagggt agagtcttgt gttttattca gattgggaaa 1440
tccattctat tttgtgaatt gggataataa cagcagtgga ataagtactt agaaatgtga 1500
aaaatgagca gtaaaataga tgagataaag aactaaagaa attaagagat agtcaattct 1560
tgctttatac ctgagcttat tctgtaaaat ttttaaagat atatgcatac ctggatttcc 1620
ttggcttctt tgagaatgta agagaaatta aatctgaata aagaattctt cctgttaaaa 1680
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 1722

```

```

<210> 7
<211> 1753
<212> DNA
<213> Homo sapiens

```

```

<400> 7
gagattctcg ccctgagcaa cgagcgacgg cctgacgtcg gcgaggaggaa gccggcccag 60
gctcggtgag gaggcaaggt tctgagggga caggctgacc tggaggacca gaggcccccg 120
gaggagcact gaaggagaag atctgccagt gggctccat tgcccagctc ctgccacac 180
tcccgcctgt tgccctgacc agagtcacat tgccctctga gcagaggagt cagcactgca 240
agcctgaaga aggccttgag gcccgaggag aggcctggg cctggtgggt gcgcaggctc 300
ctgctactga ggagcaggag gctgcctcct cctcttctac tctagttgaa gtcaccctgg 360
gggaggtgcc tgctgccgag tcaccagatc ctcccagag tctcaggga gcctccagcc 420
tccccactac catgaactac cctctctgga gccaatccta tgaggactcc agcaaccaag 480
aagaggaggg gccaaagcacc ttccctgacc tggagtccga gttccaagca gcactcagta 540
ggaaggtggc cgagttgggt cattttctgc tctcaagta tcgagccagg gagccggtca 600
caaaggcaga aatgctgggg agtgtcgtcg gaaattggca gtatttcttt cctgtgatct 660
tcagcaaagc ttccagttcc ttgcagctgg tctttggcat cgagctgatg gaagtggacc 720
ccatcggcca cttgtacatc tttgccacct gcctgggcct ctctacgat ggctgctgg 780
gtgacaatca gatcatgcc aaggcaggcc tcttgataat cgtcctggcc ataatcgcaa 840
gagagggcga ctgtgccctt gaggagaaaa tctgggagga gctgagtggt ttagaggtgt 900

```

```

ttgaggggag ggaagacagt atcttggggg atcccaagaa gctgctcacc caacatttcg      960
tgcaggaaaa ctacctggag taccggcagg tccccggcag tgatcctgca tgttatgaat    1020
tctgttgggg tccaagggcc ctcggtgaaa ccagctatgt gaaagtcctg caccatatgg    1080
taaagatcag tggaggacct cacatttcct acccaccctt gcatgagtgg gttttgagag    1140
agggggaaga gtgagtctga gcacgagttg cagccagggc cagtgggagg gggctctgggc    1200
cagtgcacct tccggggccg catcccttag tttccactgc ctctgtgac gtgaggccca    1260
ttcttcactc tttgaagcga gcagtcagca ttcttagtag tgggtttctg ttctgttga    1320
tgactttgag attattcttt gtttctgtt ggagttgttc aaatgttcct tttaacggat    1380
ggttgaatga gcgtcagcat ccaggtttat gaatgacagt agtcacacat agtgctgttt    1440
atatagttta ggagtaagag tcttggtttt tactcaaatt gggaaatcca ttccattttg    1500
tgaattgtga cataataata gcagtggtaa aagtatttgc ttaaaattgt gagcgaatta    1560
gcaataacat acatgagata actcaagaaa tcaaaagata gttgattctt gccttgtacc    1620
tcaatctatt ctgtaaaatt aaacaaatat gcaaacagg atttccttga cttctttgag    1680
aatgcaagcg aaattaaatc tgaataaata attcttctctc ttcaaaaaaa aaaaaaaaaa    1740
aaaaaaaaaa aaa                                     1753

```

```

<210> 8
<211> 1723
<212> DNA
<213> Homo sapiens

```

```

<400> 8
agcaacgagc gacggcctga cgtcggcgga ggaagccgg cccaggctcg gtgaggagggc      60
aaggttctga ggggacaggc tgacctggag gaccagaggc ccccgaggga gactgaagg      120
agaagatctg ccagtgggtc tccattgccc agtcctgcc cacactccc cctgttgccc      180
tgaccagagt catcatgcct cttgagcaga ggagtcagca ctgcaagcct gaagaaggcc      240
ttgaggcccg aggagaggcc ctgggcctgg tgggtgcgca ggctcctgct actgaggagc      300
aggaggctgc ctctcctct tctactctag ttgaagtcac cctgggggag gtgcctgctg      360
ccgagtcacc agatcctccc cagagtcctc agggagcctc cagcctcccc actaccatga      420
actaccctct ctggagccaa tcctatgagg actccagcaa ccaagaagag gaggggccaa      480
gcaccttccc tgacctggag tctgagttcc aagcagcact cagtaggaag gtggccaagt      540
tggttcattt tctgtcctc aagtatcgag ccagggagcc ggtcacaaag gcagaaatgc      600
tgaggagtggt cgtcggaat tggcagtact tctttcctgt gatcttcagc aaagcttccg      660
attccttgca gctggtcttt ggcatcgagc tgatggaagt ggaccccatc ggccacgtgt      720

```

```

acatctttgc cacctgcctg ggccctctcct acgatggcct gctgggtgac aatcagatca      780
tgcccaagac aggcttcctg ataatcatcc tggccataat cgcaaaagag ggcgactgtg      840
cccctgagga gaaaatctgg gaggagctga gtgtgttaga ggtgtttgag gggagggaag      900
acagtatctt cggggatccc aagaagctgc tcaccaata tttcgtgcag gaaaactacc      960
tgaggtaccg gcaggtcccc ggcagtgatc ctgcatgcta tgagttcctg tggggtccaa     1020
gggccctcat tgaaaccagc tatgtgaaag tcctgcacca tatggtaaag atcagtggag     1080
gacctcgcat ttcctacca ctcctgcatg agtgggcttt gagagagggg gaagagtgag     1140
tctgagcacg agttgcagcc agggccagtg ggaggggggt tgggccagtg caccttcggg     1200
ggcccatcc cttagtttcc actgcctcct gtgacgtgag gccattctt cactctttga     1260
agcgagcagt cagcattctt agtagtgggt ttctgttctg ttggatgact ttgagattat     1320
tctttgtttc ctgttgaggt tgttcaaatg ttccttttaa cggatgggtg aatgagcgtc     1380
agcatccagg tttatgaatg acagtagtca cacatagtgc tgtttatata gtttaggagt     1440
aagagtcttg ttttttattc agattgggaa atccattcca ttttgtgaat tgtgacataa     1500
taatagcagt ggtaaaagta tttgcttaaa attgtgagcg aattagcaat aacatacatg     1560
agataactca agaaatcaaa agatagttga ttcttgccct gtacctcaat ctattctgta     1620
aaattaaaca aatatgcaaa ccaggatttc cttgacttct ttgagaatgc aagcgaaatt     1680
aaatctgaat aaataattaa aaaaaaaaaa aaaaaaaaaa aaa                        1723

```

```

<210> 9
<211> 1524
<212> DNA
<213> Homo sapiens

```

```

<400> 9
agcagacaga ggactctcat taaggaaggt gtcctgtgcc ctgaccctac aagatgccaa      60
gagaagatgc tcacttcata tatggttacc ccaagaaggg gcacggccac tcttacacca     120
cggtgaaga ggccgctggg atcgcatcc tgacagtgat cctgggagtc ttactgctca     180
tcggctgttg gtattgtaga agacgaaatg gatacagagc cttgatggat aaaagtcttc     240
atgttggcac tcaatgtgcc ttaacaagaa gatgccaca agaagggttt gatcatcggg     300
acagcaaagt gtctcttcaa gagaaaaact gtgaacctgt gggtcccaat gctccacctg     360
cttatgagaa actctctgca gaacagtcac caccacctta ttcaccttaa gagccagcga     420
gacacctgag acatgctgaa attatttctc tcacactttt gcttgaattt aatacagaca     480
tctaattgtc tcctttggaa tgggtgtagga aaaatgcaag ccatctctaa taataagtca     540
gtgttaaaat tttagtaggt ccgctagcag tactaatcat gtgaggaaat gatgagaaat     600

```


attaaattgg gaaaactcca tcaataaatg ttgcaatgca tgatactatc tgtgccagag 660
 gtaatgttag taaatccatg gtgttatttt ctgagagaca gaattcaagt ggggtattctg 720
 gggccatcca atttctcttt acttgaaatt tggctaataa caaactagtc aggttttcga 780
 accttgaccg acatgaactg tacacagaat tgttcagta ctatggagtg ctacaaaagg 840
 atacttttac aggttaagac aaagggttga ctggcctatt tatctgatca agaacatgtc 900
 agcaatgtct ctttgtgctc taaaattcta ttatactaca ataatatatt gtaaagatcc 960
 tatagctctt tttttttgag atggagtttc gcttttgttg cccaggctgg agtgcaatgg 1020
 cgcgatcttg gctcaccata acctccgcct cccaggttca agcaattctc ctgccttagc 1080
 ctctgagta gctgggatta caggcgtgcg ccactatgcc tgactaattt tgtagtttta 1140
 gtagagacgg ggtttctcca tgttggtcag gctggtctca aactcctgac ctgaggtgat 1200
 ctgcccgcct cagcctccca aagtgtgga attacaggcg tgagccacca cgctggctg 1260
 gatcctatat cttaggttaag acatataacg cagtctaatt acatttcact tcaaggctca 1320
 atgctattct aactaatgac aagtattttc tactaaacca gaaattggtg gaaggattta 1380
 aataagtaaa agctactatg tactgcctta gtgctgatgc ctgtgtactg ccttaaatgt 1440
 acctatggca atttagctct cttgggttcc caaatccctc tcacaagaat gtgcagaaga 1500
 aatcataaag gatcagagat tctg 1524

<210> 10
 <211> 1881
 <212> DNA
 <213> Homo sapiens

<400> 10
 cctgcatctt ttggaagga ttctttttat aaatcagaaa gtgttcgagg ttcaaagggtt 60
 tgccctggag cgtgtgaaca ttccctcgct cggttttcaa ctgcctcca acctgcgccg 120
 cccggccagc atgtctcccc gcccgtaag cggggctgcc gcctccctgc cgctccggct 180
 gccactaacg acccgccctc gccgccacct ggccctcctg atcgacgaca cacgcacttg 240
 aaacttgttc tcagggtgtg tggaatcaac tttccggaag caaccagccc accagaggag 300
 gtcccagcgc cgagcggaga cgatgcagcg gagactgggt cagcagtga gcgtcgcggg 360
 gttcctgctg agctacgcgc tgccctcctg cgggcgctcg gtggaggggc tcagccgccg 420
 cctcaaaaga gctgtgtctg aacatcagct cctccatgac aaggggaagt ccatccaaga 480
 tttacggcga cgattcttcc ttcaccatct gatcgcagaa atccacacag ctgaaatcag 540
 agctacctcg gaggtgtccc ctaactccaa gccctctccc aacacaaaga accaccccg 600
 ccgatttggg tctgatgatg agggcagata cctaactcag gaaactaaca aggtggagac 660

gtacaaagag cagccgctca agacacctgg gaagaaaaag aaaggcaagc ccgggaaacg 720
 caaggagcag gaaaagaaaa aacggcgaac tcgctctgcc tggtagact ctggagtgc 780
 tgggagtggg ctagaagggg accacctgtc tgacacctcc acaacgtcg tggagctcga 840
 ttcacggtaa caggcttctc tggcccgtag ctcagcggg gtgctctcag ctgggttttg 900
 gagcctccct tctgccttgg cttggacaaa ctagaattt tctcccttta tgtatctcta 960
 tcgatttgt agcaattgac agagaataac tcagaatatt gtctgcctta aagcagtacc 1020
 ccctaccac acacaccct gtctccagc accatagaga ggcgctagag cccattcctc 1080
 tttctccacc gtcacccaac atcaatcctt taccactcta ccaaataatt tcatattcaa 1140
 gcttcagaag ctagtgcaca tcttcataat ttgctggaga agtgtgtttc tcccccttac 1200
 tctcacacct gggcaactt tcttcagtgt tttcatttc ttacgttctt tcacttcaag 1260
 ggagaatata gaagcatttg atattatcta caaacactgc agaacagcat catgtcataa 1320
 acgattctga gccattcaca ctttttattt aattaaatgt atttaattaa atctcaaatt 1380
 tattttaatg taaagaactt aaattatgtt taaacacat gccttaaatt tgtttaatta 1440
 aatttaactc tggtttctac cagctcatalc aaaataaatg gtttctgaaa atgtttaagt 1500
 attaacttac aaggatatag gtttttctca tgtatctttt tgttcattgg caagatgaaa 1560
 taattttct agggtaatgc cgtaggaaaa ataaaacttc acatttatgt ggcttggtta 1620
 tcttagctc acagattgag gtaataatga cactcctaga ctttgggac aaataactta 1680
 gggccaagtc ttgggtctga atttatttaa gttcacaacc tagggcaagt tactctgcct 1740
 ttctaagact cacttacatc ttctgtgaaa tataattgta ccaacctcat agagtttggt 1800
 gtcaactaaa tgagattata tgtggactaa atatctgtca tatagtaaac actcaataaa 1860
 ttgcaacata ttaaaaaaaaa a 1881

<210> 11
 <211> 3336
 <212> DNA
 <213> Homo sapiens

<400> 11
 ttttcttaga cattaactgc agacggctgg caggatagaa gcagcggctc acttggactt 60
 tttcaccagg gaaatcagag acaatgatgg ggctcttccc cagaactaca ggggctctgg 120
 ccatcttcgt ggtggtcata ttggttcattg gagaattgcg aatagagact aaaggtcaat 180
 atgatgaaga agagatgact atgcaacaag ctaaaagaag gcaaaaacgt gaatgggtga 240
 aatttgccaa accctgcaga gaaggagaag ataactcaaa aagaaacca attgccaaga 300
 ttacttcaga ttaccaagca accagaaaa tcacctaccg aatctctgga gtgggaatcg 360

atcagccgcc	ttttggaatc	tttggtgttg	acaaaaacac	tgagatatatt	aacataacag	420
ctatagtcga	ccgggaggaa	actccaagct	tcctgatcac	atgtcgggct	ctaaatgccc	480
aaggactaga	tgtagagaaa	ccacttatac	taacggttaa	aatttttgat	attaatgata	540
atcctccagt	attttcacaa	caaattttca	tggtgaaat	tgaagaaaat	agtgcctcaa	600
actcactggt	gatgatacta	aatgccacag	atgcagatga	accaaaccac	ttgaattcta	660
aaattgcctt	caaaattgtc	tctcaggaac	cagcaggcac	acccatgttc	ctcctaagca	720
gaaacactgg	ggaagtccgt	actttgacca	attctcttga	ccgagagcaa	gctagcagct	780
atcgtctggt	tgtgagtggg	gcagacaaaag	atggagaagg	actatcaact	caatgtgaat	840
gtaatatata	agtgaagat	gtcaacgata	acttcccaat	gtttagagac	tctcagtatt	900
cagcacgtat	tgaagaaaat	attttaagtt	ctgaattact	tcgatttcaa	gtaacagatt	960
tggtatgaaga	gtacacagat	aattggcttg	cagtataattt	ctttacctct	gggaatgaag	1020
gaaattgggt	tgaatacaaa	actgatccta	gaactaatga	aggcatcctg	aaagtgggtga	1080
aggctctaga	ttatgaacaa	ctacaaagcg	tgaacttag	tattgctgtc	aaaaacaaag	1140
ctgaatttca	ccaatcagtt	atctctcgat	accgagttca	gtcaacccca	gtcacaattc	1200
aggtaataaa	tgtaaagaaa	ggaattgcat	tccgtcctgc	ttccaagaca	tttactgtgc	1260
aaaaaggcat	aagtagcaaa	aaattggtgg	attatatcct	gggaacatat	caagccatcg	1320
atgaggacac	taacaaaagct	gcctcaaatg	tcaaatatgt	catgggacgt	aacgatgggtg	1380
gatacctaata	gattgattca	aaaactgctg	aatcaaat	tgtcaaaaat	atgaaccgag	1440
attctacttt	catagttaac	aaaacaatca	cagctgaggt	tctggccata	gatgaatata	1500
cgggtaaaac	ttctacaggc	acggtatatg	ttagagtacc	cgatttcaat	gacaattgtc	1560
caacagctgt	cctcgaaaaa	gatgcagttt	gcagttcttc	accttccgtg	gttgtctccg	1620
ctagaacact	gaataataga	tacactggcc	cctatacatt	tgcactggaa	gatcaacctg	1680
taaagttgcc	tgccgtatgg	agtatcacia	ccctcaatgc	tacctcgcc	ctcctcagag	1740
cccaggaaca	gatacctcct	ggagtatacc	acatctccct	ggtacttaca	gacagtcaga	1800
acaatcggtg	tgagatgcca	cgcagcttga	cactggaagt	ctgtcagtg	gacaacaggg	1860
gcactctgtg	aacttcttac	ccaaccacia	gccctgggac	caggtatggc	aggccgcact	1920
cagggaggct	ggggcctgcc	gccatcggcc	tgctgctcct	tggtctcctg	ctgctgctgt	1980
tggtcccccct	tctgctgttg	acctgtgact	gtggggcagg	ttctactggg	ggagtgcag	2040
gtggttttat	cccagttcct	gatggctcag	aaggaacaat	tcatcagtg	ggaattgaag	2100
gagcccatcc	tgaagacaag	gaaatcacia	atatttgtgt	gcctcctgta	acagccaatg	2160

gagccgattt catggaaagt tctgaagttt gtacaaatac gtatgccaga ggcacagcgg 2220
 tggaaggcac ttcaggaatg gaaatgacca ctaagcttgg agcagccact gaatctggag 2280
 gtgctgcagg ctttgcaaca gggacagtgt caggagctgc ttcaggattc ggagcagcca 2340
 ctggagttgg catctgttcc tcagggcagt ctggaaccat gagaacaagg cattccactg 2400
 gaggaaccaa taaggactac gctgatgggg cgataagcat gaattttctg gactcctact 2460
 tttctcagaa agcatttgcc tgtgcggagg aagacgatgg ccaggaagca aatgactgct 2520
 tgttgatcta tgataatgaa ggcgcagatg ccaactggttc tcctgtgggc tccgtgggtt 2580
 gttgcagttt tattgctgat gacctggatg acagcttctt ggactcactt ggacccaaat 2640
 ttaaaaaact tgcagagata agccttgggtg ttgatggtga aggcaaagaa gttcagccac 2700
 cctctaaaga cagcgtttat gggattgaat cctgtggcca tcccatagaa gtccagcaga 2760
 caggatttgt taagtgccag actttgtcag gaagtcaagg agcttctgct ttgtccgct 2820
 ctgggtctgt ccagccagct gtttccatcc ctgaccctct gcagcatggt aactatntag 2880
 taacggagac ttactcggct tctggttccc tcgtgcaacc ttccactgca ggctttgatc 2940
 cacttctcac aaaaaatgtg atagtgcag aaagggtgat ctgtcccatt tccagtgttc 3000
 ctggcaacct agctggccca acgcagctac gagggtcaca tactatgctc tgtacagagg 3060
 atccttgctc cgtctaata tgaccagaat gagctggaat accacactga ccaaactctg 3120
 atctttggac taaagtattc aaaatagcat agcaaagctc actgtattgg gctaataatt 3180
 tggcacttat tagcttctct cataaactga tcacgattat aaattaaatg tttgggttca 3240
 taccocaaaa gcaatatgtt gtcactccta attctcaagt actattcaaa ttgtagtaaa 3300
 tcttaaagtt tttcaaaacc ctaaaatcat attcgc 3336

<210> 12

<211> 1694

<212> DNA

<213> Homo sapiens

<400> 12

ctctctgccc acctctgctt cctctaggaa cacaggagtt ccagatcaca tcgagttcac 60
 catgaattca ctcagtgaag ccaacaccaa gttcatgttc gacctgttcc aacagttcag 120
 aaaatcaaaa gagaacaaca tcttctattc ccctatcagc atcacatcag cattagggat 180
 ggtcctctta ggagccaaag acaacactgc acaacagatt aagaaggttc ttcactttga 240
 tcaagtcaca gagaacacca caggaaaagc tgcaacatat catgttgata ggtcaggaaa 300
 tgttcatcac cagtttcaaa agcttctgac tgaattcaac aaatccactg atgcatatga 360
 gctgaagatc gccacaagc tcttcggaga aaaaacgtat ctatttttac aggaatatatt 420

```

agatgccatc aagaaatttt accagaccag tgtggaatct gttgattttg caaatgctcc 480
agaagaaagt cgaaagaaga ttaactcctg ggtggaaagt caaacgaatg aaaaaattaa 540
aaacctaat cctgaaggta atattggcag caataccaca ttggttcttg tgaacgcaat 600
ctatttcaaa gggcagtggg agaagaaatt taataaagaa gatactaaag aggaaaaatt 660
ttggccaaac aagaatacat acaagtccat acagatgatg aggcaatata catcttttca 720
ttttgcctcg ctggaggatg tacaggccaa ggtcctggaa ataccataca aaggcaaaga 780
tctaagcatg attgtgttgc tgccaaatga aatcgatggt ctccagaagc ttgaagagaa 840
actcactgct gagaaattga tggaatggac aagtttgag aatatgagag agacacgtgt 900
cgatttacac ttacctcggt tcaaagtggg agagagctat gacctcaagg acacgttgag 960
aaccatggga atggtggata ttttcaatgg ggatgcagac ctctcaggca tgaccgggag 1020
ccgcggtctc gtgctatctg gagtcctaca caaggccttt gtggagggtta cagaggaggg 1080
agcagaagct gcagctgcca ccgctgtagt aggattcgga tcatcaccta cttcaactaa 1140
tgaagagttc cattgtaatc accctttcct attcttcata aggcaaaata agaccaacag 1200
catcctcttc tatggcagat tctcatcccc gtagatgcaa ttagtctgtc actccatttg 1260
gaaaatgttc acctgcagat gttctggtaa actgattgct ggcaacaaca gattctcttg 1320
gctcatatth cttttctttc tcatcttgat gatgatcgtc atcatcaaga atttaatgat 1380
taaaatagca tgcctttctc tctttctctt aataagccca catataaatg tactttttct 1440
tccagaaaaa ttctccttga ggaaaaatgt ccaaaataag atgaatcact taataccgta 1500
tcttctaaat ttgaaatata attctgtttg tgacctgttt taaatgaacc aaaccaaatc 1560
atactttttc tttgaattta gcaacctaga aacacacatt tctttgaatt taggtgatac 1620
ctaaatcctt cttatgtttc taaattttgt gattctataa aacacatcat caataaaata 1680
gtgacataaa atca 1694

```

```

<210> 13
<211> 1782
<212> DNA
<213> Homo sapiens

```

```

<400> 13
gggagacaca cacagcctct ctgccacct ctgcttctc taggaacaca ggagttccag 60
atcacatcga gttcaccatg aattcactca gtgaagccaa caccaagtgc atgttcgatc 120
tgttccaaca gttcagaaaa tcaaaagaga acaacatctt ctattcccct atcagcatca 180
catcagcatt agggatgggc ctcttaggag ccaaagacaa cactgcacaa caaattagca 240
aggttcttca ctttgatcaa gtcacagaga acaccacaga aaaagctgca acatatcatg 300

```

```

ttgataggctc aggaaatggt catcaccagt ttcaaaagct tctgactgaa ttcaacaaat 360
ccactgatgc atatgagctg aagatcgcca acaagctctt cggagaaaag acgtatcaat 420
ttttacagga atatttagat gccatcaaga aattttacca gaccagtgtg gaatctactg 480
attttgcaaa tgctccagaa gaaagtcgaa agaagattaa ctctctgggtg gaaagtcaaa 540
cgaatgaaaa aattaaaaac ctatttcctg atgggactat tggcaatgat acgacactgg 600
ttcttgtgaa cgcaatctat ttcaaagggc agtgggagaa taaatttaaa aaagaaaaca 660
ctaaagagga aaaatttttg ccaaacaaga atacatacaa atctgtacag atgatgaggc 720
aatacaattc ctttaatttt gccttgctgg aggatgtaca ggccaaggctc ctggaaatac 780
catacaaagg caaagatcta agcatgattg tgctgctgcc aaatgaaatc gatggtctgc 840
agaagcttga agagaaactc actgctgaga aattgatgga atggacaagt ttgcagaata 900
tgagagagac atgtgtcgat ttacacttac ctcggttcaa aatggaagag agctatgacc 960
tcaaggacac gttgagaacc atgggaatgg tgaatatctt caatggggat gcagacctct 1020
caggcatgac ctggagccac ggtctctcag tatctaaagt cctacacaag gcctttgtgg 1080
aggtcactga ggagggagtg gaagctgcag ctgccaccgc tgtagtagta gtcgaattat 1140
catctccttc aactaatgaa gagttctgtt gtaatcacc tttcctattc ttcataaggc 1200
aaaataagac caacagcatc ctcttctatg gcagattctc atcccatag atgcaattag 1260
tctgtcactc cathtagaaa atgttcacct agaggtgttc tggtaaaactg attgctggca 1320
acaacagatt ctcttggtc atatttcttt tctatctcat cttgatgatg atagtcatca 1380
tcaagaattt aatgattaaa atagcatgcc tttctctctt tctcttaata agcccacata 1440
taaatgtact tttccttcca gaaaaatttc ccttgaggaa aaatgtccaa gataagatga 1500
atcatttaat accgtgtctt cttaaattga aatataattc tgtttctgac ctgttttaaa 1560
tgaaccaaac caaatcatc tttctcttca aatttagcaa cctagaaaca cacatttctt 1620
tgaatttagg tgatacctaa atccttctta tgtttctaaa ttttgtgatt ctataaaaca 1680
catcatcaat aaaataatga ctaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1740
aaaaaaaaa aacccaaaaa aaaaaaaaaa aaaaaaaaaa aa 1782

```

<210> 14

<211> 1528

<212> DNA

<213> Homo sapiens

<400> 14

```

cggcgagcga gcaccttcga cgcggtccgg ggacccctc gtcgctgtcc tcccgacgcg 60

```

```

gacccgcgtg cccagggcct cgcgctgcc ggccggctcc tcgtgtcca ctccggcg 120

```

acgccctccc gcgagtcctc ggccccctccc gcgccccctct tctcgggcgcg cgcgcagcat 180
 ggcgcccccg caggctcctcg cgttcgggct tctgcttgcc gcggcgacgg cgacttttgc 240
 cgcagctcag gaagaatgtg tctgtgaaaa ctacaagctg gccgtaaaact gctttgtgaa 300
 taataatcgt caatgccagt gtacttcagt tgggtgcacaa aatactgtca tttgctcaaa 360
 gctggctgcc aatgttttg tgatgaaggc agaaatgaat ggctcaaaac ttgggagaag 420
 agcaaaacct gaagggggccc tocagaacaa tgatgggctt tatgatcctg actgcgatga 480
 gagcgggctc ttaagggcca agcagtgcac cggcacctcc acgtgctggg gtgtgaacac 540
 tgctgggggtc agaagaacag acaaggacac tgaaataacc tgctctgagc gagtgagaac 600
 ctactggatc atcattgaac taaaacacaa agcaagagaa aaaccttatg atagtaaaag 660
 tttgcggact gcacttcaga aggagatcac aacgcgttat caactggatc caaaatttat 720
 cacgagtatt ttgtatgaga ataatgttat cactattgat ctggttcaaa attcttctca 780
 aaaaactcag aatgatgtgg acatagctga tgtggcttat tattttgaaa aagatgttaa 840
 aggtgaatcc ttgtttcatt ctaagaaaat ggacctgaca gtaaatgggg aacaactgga 900
 tctggatcct ggtcaaaact taatttatta tgttgatgaa aaagcacctg aattctcaat 960
 gcagggctca aaagctggtg ttattgctgt tattgtgggt gtgggtgatag cagttgttgc 1020
 tggaattggt gtgctggtta tttccagaaa gaagagaatg gcaaagtatg agaaggctga 1080
 gataaaggag atgggtgaga tgcataggga actcaatgca taactatata atttgaagat 1140
 tatagaagaa gggaaatagc aaatggacac aaattacaaa tgtgtgtgcg tgggacgaag 1200
 acatctttga aggtcatgag tttgttagtt taacatcata tatttgtaat agtgaaacct 1260
 gtactcaaaa tataagcagc ttgaaactgg ctttaccaat cttgaaattt gaccacaagt 1320
 gtcttatata tgcagatcta atgtaaaatc cagaacttgg actccatcgt taaaattatt 1380
 tatgtgtaac attcaaatgt gtgcattaaa tatgcttcca cagtaaaatc tgaaaaactg 1440
 atttgtgatt gaaagctgcc tttctattta cttgagtctt gtacatacat acttttttat 1500
 gagctatgaa ataaaacatt ttaaactg 1528

<210> 15

<211> 2384

<212> DNA

<213> Homo sapiens

<400> 15

tattgagttc ttcaaacatt gtagcctctt tatgggtctct gagaaataac taccttaaac 60

ccataatctt taatacttcc taaactttct taataagaga agctctattc ctgacactac 120

ctctcatttg caagggtcaaa tcatcattag ttttgtagtc tattaactgg gtttgcttag 180

gtcaggcatt attattacta accttattgt taatattcta accataagaa ttaaactatt 240
 aatggtgaat agagtttttc actttaacat aggctatcc cactggtggg atacgagcca 300
 attcgaaaga aaagtcagtc atgtgctttt cagaggatga aagcttaaga taaagactaa 360
 aagtgtttga tgctggagggt gggagtggta ttatataggt ctgagccaag acatgtgata 420
 atcactgtag tagtagctgg aaagagaaat ctgtgactcc aattagccag ttcctgcaga 480
 ccttgtgagg actagaggaa gaatgctcct ggctgttttg tactgcctgc tgtggagttt 540
 ccagacctcc gctggccatt tccctagagc ctgtgtctcc tctaagaacc tgatggagaa 600
 ggaatgctgt ccaccgtgga gcggggacag gagtccctgt ggccagcttt caggcagagg 660
 ttcctgtcag aatatccttc tgtccaatgc accacttggg cctcaatttc cttcacagg 720
 ggtggatgac cgggagtcgt ggccttcctt cttttataat aggacctgcc agtgcctctgg 780
 caacttcatg ggattcaact gtggaaactg caagtgtggc ttttggggac caaactgcac 840
 agagagacga ctcttggtga gaagaaacat cttcgatttg agtgccccag agaaggacaa 900
 attttttgcc tacctcactt tagcaaagca taccatcagc tcagactatg tcatcccat 960
 agggacctat ggccaaatga aaaatggatc aacacccatg tttaacgaca tcaatattta 1020
 tgacctcttt gtctggatgc attattatgt gtcaatggat gcaactgctt ggggatctga 1080
 aatctggaga gacattgatt ttgcccatag agcaccagct tttctgcctt ggcatagact 1140
 cttcttggtg cggtgggaac aagaaatcca gaagctgaca ggagatgaaa acttcactat 1200
 tccatattgg gactggcggg atgcagaaaa gtgtgacatt tgcacagatg agtacatggg 1260
 aggtcagcac ccacaaaatc ctaacttact cagcccagca tcattcttct cctcttgga 1320
 gattgtctgt agccgattgg aggagtacaa cagccatcag tctttatgca atggaacgcc 1380
 cgagggacct ttacggcgta atcctggaaa ccatgacaaa tccagaacct caaggctccc 1440
 ctcttcagct gatgtagaat tttgcctgag tttgaccaa tatgaatctg gttccatgga 1500
 taaagctgcc aatttcagct ttagaaatac actggaagga tttgctagtc cacttactgg 1560
 gatagcggat gcctctcaaa gcagcatgca caatgccttg cacatctata tgaatggaac 1620
 aatgtcccag gtacagggat ctgccaacga tcctatcttc cttcttcacc atgcatttgt 1680
 tgacagtatt tttgagcagt ggctccgaag gcaccgtcct cttcaagaag tttatccaga 1740
 agccaatgca cccattggac ataaccggga atoctacatg gttcctttta taccactgta 1800
 cagaaatggt gatttcttta tttcatcaa agatctgggc tatgactata gctatctaca 1860
 agattcagac ccagactctt ttcaagacta cattaagtcc tatttgaac aagcgagtcg 1920
 gatctggtca tggctccttg gggcggcgat ggtaggggcc gtccctactg ccctgctggc 1980
 agggcttgtg agcttgctgt gtcgtcacia gagaaagcag cttcctgaag aaaagcagcc 2040

actcctcatg gagaaagagg attaccacag cttgtatcag agccatttat aaaaggctta 2100
 ggcaatagag tagggccaaa aagcctgacc tcaactctaac tcaaagtaat gtccagggttc 2160
 ccagagaata tctgctggta tttttctgta aagaccattt gcaaaattgt aacctaatac 2220
 aaagtgtagc cttcttccaa ctcaggtaga acacacctgt ctttgtcttg ctgttttcac 2280
 tcagcccttt taacattttc ccctaagccc atatgtctaa ggaaaggatg ctatttggtta 2340
 atgaggaact gttatttgta tgtgaattaa agtgctctta tttt 2384

<210> 16
 <211> 2702
 <212> DNA
 <213> Homo sapiens

<400> 16
 cattctcccc caggtcact caccatgacc aagctgagcg cccaagtcaa aggctctctc 60
 aacatcacca ccccggggct gcagatatgg aggatcgagg ccatgcagat ggtgcctggt 120
 ccttcagca cctttggaag cttcttcgat ggtgactgct acatcatcct ggctatccac 180
 aagacagcca gcagcctgtc ctatgacatc cactactgga ttggccagga ctcatccctg 240
 gatgagcagg gggcagctgc catctacacc acacagatgg atgacttcct gaagggccgg 300
 gctgtgcagc accgcgaggt ccagggcaac gagagcgagg ccttccgagg ctacttcaag 360
 caaggccttg tgatccggaa agggggcggtg gcttctggca tgaagcacgt ggagaccaac 420
 tcctatgacg tccagaggct gctgcatgtc aagggcaaga ggaacgtggt agctggagag 480
 gtagagatgt cctggaagag tttcaaccga ggggatgttt tcctcctgga ccttggaag 540
 cttatcatcc agtggaatgg accggaaagc acccgatatg agagactcag gggcatgact 600
 ctggccaagg agatccgaga ccaggagcgg ggagggcgca cctatgtagg cgtggtggac 660
 ggagagaatg aattggcatc cccgaagctg atggaggtga tgaaccacgt gctgggcaag 720
 cgcagggagc tgaaggcggc cgtgcccgc acggtgggtg agccggcact caaggctgca 780
 ctcaaactgt accatgtgtc tgactccgag gggaatctgg tggtagaggga agtcgccaca 840
 cggccactga cacaggacct gctcagtcac gaggactggt acatcctgga ccaggggggc 900
 ctgaagatct acgtgtggaa agggaagaaa gccaatgagc aggagaagaa gggagccatg 960
 agccatgcgc tgaacttcat caaagccaag cagtaccac caagcacaca ggtggagggtg 1020
 cagaatgatg gggctgagtc ggccgtcttt cagcagctct tccagaagtg gacagcgtcc 1080
 aaccggacct caggcctagg caaaaccac actgtgggct ccgtggccaa agtggaacag 1140
 gtgaagttcg atgccacatc catgcatgtc aagcctcagg tggctgcca gcagaagatg 1200
 gtagatgatg ggagtgggga agtgcaggtg tggcgcattg agaacctaga gctggtacct 1260

```

gtggattcca agtggctagg ccacttctat gggggcgact gctacctgct gctctacacc 1320
tacctcatcg gcgagaagca gcattacctg ctctacgttt ggcagggcag ccaggccagc 1380
caagatgaaa ttacagcatc agcttatcaa gccgtcatcc tggaccagaa gtacaatggg 1440
gaaccagtcc agatccgggt cccaatgggc aaggagccac ctcatcttat gtccatcttc 1500
aagggacgca tgggtgtcta ccagggaggg acctcccga ctaacaactt ggagaccggg 1560
ccctccacac ggctgttcca ggtccaggga actggcgcca acaacaccaa ggcctttgag 1620
gtcccagcgc gggccaatth cctcaattcc aatgatgtct ttgtcctcaa gaccagctct 1680
tgctgctatc tatgggtgtg gaaggggtgt agcggggacg agcgggagat ggccaagatg 1740
gttgctgaca ccatctcccg gacggagaag caagtgtgtg tggaagggca ggagccagcc 1800
aacttctgga tggccctggg tgggaaggcc ccctatgcca acaccaagag actacaggaa 1860
gaaaacctgg tcatcacccc ccggctcttt gagtgttcca acaagactgg gcgcttcctg 1920
gccacagaga tccctgactt caatcaggat gacttggaag aggatgatgt gttcctacta 1980
gatgtctggg accaggtctt cttctggatt gggaaacatg ccaacgagga ggagaagaag 2040
gccgcagcaa ccactgcaca ggaatacctc aagaccatc ccagcgggcg tgacctgag 2100
accccatca ttgtggtgaa gcagggacac gagccccca ccttcacagg ctggttcctg 2160
gcttgggatc cttcaagtg gagtaacacc aaatcctatg aggacctgaa ggcggagtct 2220
ggcaacctta gggactggag ccagatcact gctgaggtca caagcccaa agtggacgtg 2280
ttcaatgcta acagcaacct cagttctggg cctctgcca tcttccccct ggagcagcta 2340
gtgaacaagc ctgtagagga gctccccgag ggtgtggacc ccagcaggaa ggaggaacac 2400
ctgtocattg aagatttcac tcaggccttt gggatgactc cagctgcctt ctctgctctg 2460
cctcgatgga agcaacaaaa cctcaagaaa gaaaaaggac tattttgaga agagtagctg 2520
tggttgtaaa gcagtaccct accctgattg tagggtctca ttttctcacc gatattagtc 2580
ctacaccaat tgaagtgaat ttttgcagat gtgcctatga gcacaaactt ctgtggcaaa 2640
tgccagtttt gtttaataat gtacctatc cttcagaaag atgatacccc aaaaaaaaaa 2700
aa 2702

```

<210> 17
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 17
 gattgtgatg taacggctgt aatg

24

<210> 18
<211> 20
<212> DNA
<213> Homo sapiens

<400> 18
atccttgtcc tccacgggtt 20

<210> 19
<211> 17
<212> DNA
<213> Homo sapiens

<400> 19
agatcgacaa cgcccgt 17

<210> 20
<211> 20
<212> DNA
<213> Homo sapiens

<400> 20
agagcctgtt ccgtctcaaa 20

<210> 21
<211> 21
<212> DNA
<213> Homo sapiens

<400> 21
gggccttggc ctccctctaga g 21

<210> 22
<211> 19
<212> DNA
<213> Homo sapiens

<400> 22
ccagggagcg actgttgtc 19

<210> 23
<211> 21
<212> DNA
<213> Homo sapiens

<400> 23
gtgaggaggc aaggattytsa g 21

<210> 24
<211> 22
<212> DNA
<213> Homo sapiens

<400> 24
agaccacwg gcagatcttc tc 22

<210> 25
<211> 19
<212> DNA
<213> Homo sapiens

<400> 25
actgtcagga tgccgatcc 19

<210> 26
<211> 23
<212> DNA
<213> Homo sapiens

<400> 26
agcggcctct tcagccgtgg tgt 23

<210> 27
<211> 23
<212> DNA
<213> Homo sapiens

<400> 27
tcatggagga gctgatgttc aga 23

<210> 28
<211> 19
<212> DNA
<213> Homo sapiens

<400> 28
caaaaggcgg ctgatcgat 19

<210> 29
<211> 21
<212> DNA
<213> Homo sapiens

<400> 29
ggcgatcttc agtcatatg c 21

<210> 30
<211> 23
<212> DNA
<213> Homo sapiens

<400> 30
ggttttgctc ttctcccaag ttt 23

<210> 31
<211> 23
<212> DNA
<213> Homo sapiens

<400> 31
actgatggct gttgtactcc tcc 23

<210> 32
<211> 18
<212> DNA
<213> Homo sapiens

<400> 32
ttgccagact ccgccttc 18

<210> 33
<211> 16
<212> DNA
<213> Homo sapiens

<400> 33
gtgaaggcca cagcat 16

<210> 34
<211> 17
<212> DNA
<213> Homo sapiens

<400> 34
aactggctgc tgtaacg 17

<210> 35
<211> 19
<212> DNA
<213> Homo sapiens

<400> 35
gccgatgagc agtaagact 19

<210> 36
<211> 20
<212> DNA
<213> Homo sapiens

<400> 36
tgtcaacaac aaagattcca 20

<210> 37
<211> 18
<212> DNA
<213> Homo sapiens

<400> 37
tctccgaaga gcttggtg 18

<210> 38
<211> 17
<212> DNA

<213> Homo sapiens

<400> 38

agcccatcat tggtctg

17

<210> 39

<211> 17

<212> DNA

<213> Homo sapiens

<400> 39

cgttccattg cataaag

17

<210> 40

<211> 16

<212> DNA

<213> Homo sapiens

<400> 40

gctccagtcc ctaagg

16