US 20030236632A1

(19) **United States**

(12) **Patent Application Publication**

(10) **Pub. No.: US 2003/0236632 A1**

(43) **Pub. Date: Dec. 25, 2003**

Ma et al.

(54) **BIOMARKERS FOR BREAST CANCER**

(76) Inventors: **Xiao-Jun Ma**, San Diego, CA (US); **Dennis C. Sgroi**, Winchester, MA (US); **Mark G. Erlander**, Encinitas, CA (US)

Correspondence Address:
Kawai Lau
Morrison & Foerster LLP
Suite 500
3811 Valley Centre Drive
San Diego, CA 92130-2332 (US)

(21) Appl. No.: **10/282,596**

(22) Filed: **Oct. 28, 2002**

Related U.S. Application Data

(63) Continuation-in-part of application No. 10/028,018, filed on Dec. 21, 2001. Continuation-in-part of application No. 10/211,015, filed on Aug. 1, 2002.

Publication Classification

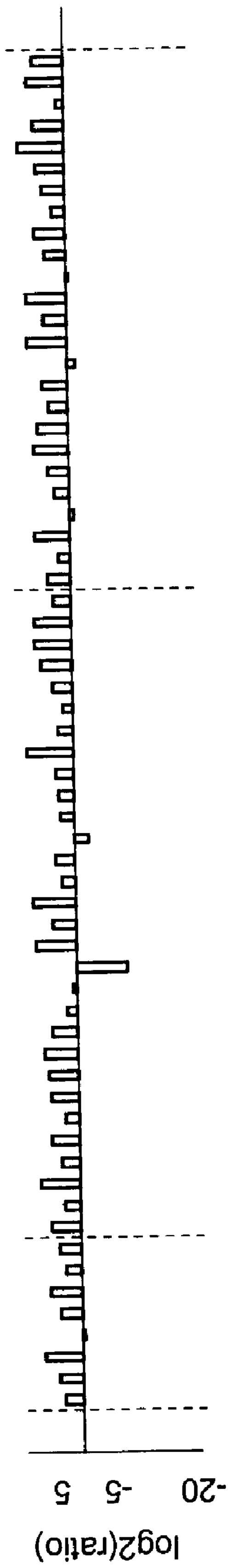
(51) **Int. Cl.**[7] **C12Q 1/68**; G06F 19/00; G01N 33/48; G01N 33/50

(52) **U.S. Cl.** **702/20**; 435/6

(57) **ABSTRACT**

Methods and compositions are provided for the detection of breast cancer based upon the identification of three biomarkers for non-normal breast cells. The biomarkers were identified based upon multiple sampling of reference breast tissue samples from independent cases of breast cancer. Two biomarkers display increased expression in non-normal cells while the third biomarker displays decreased expression in non-normal cells.

Fig. 1



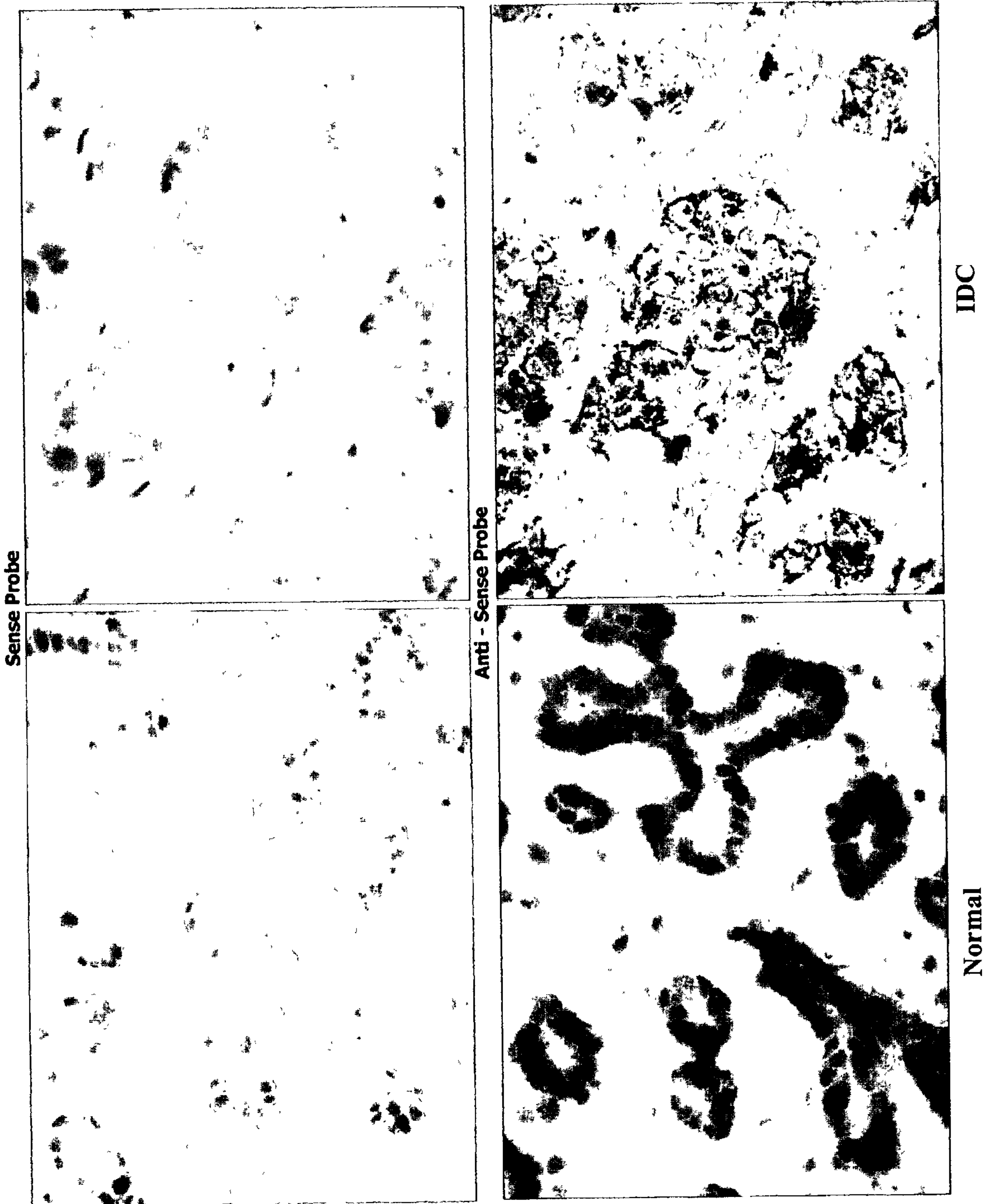


Fig. 2

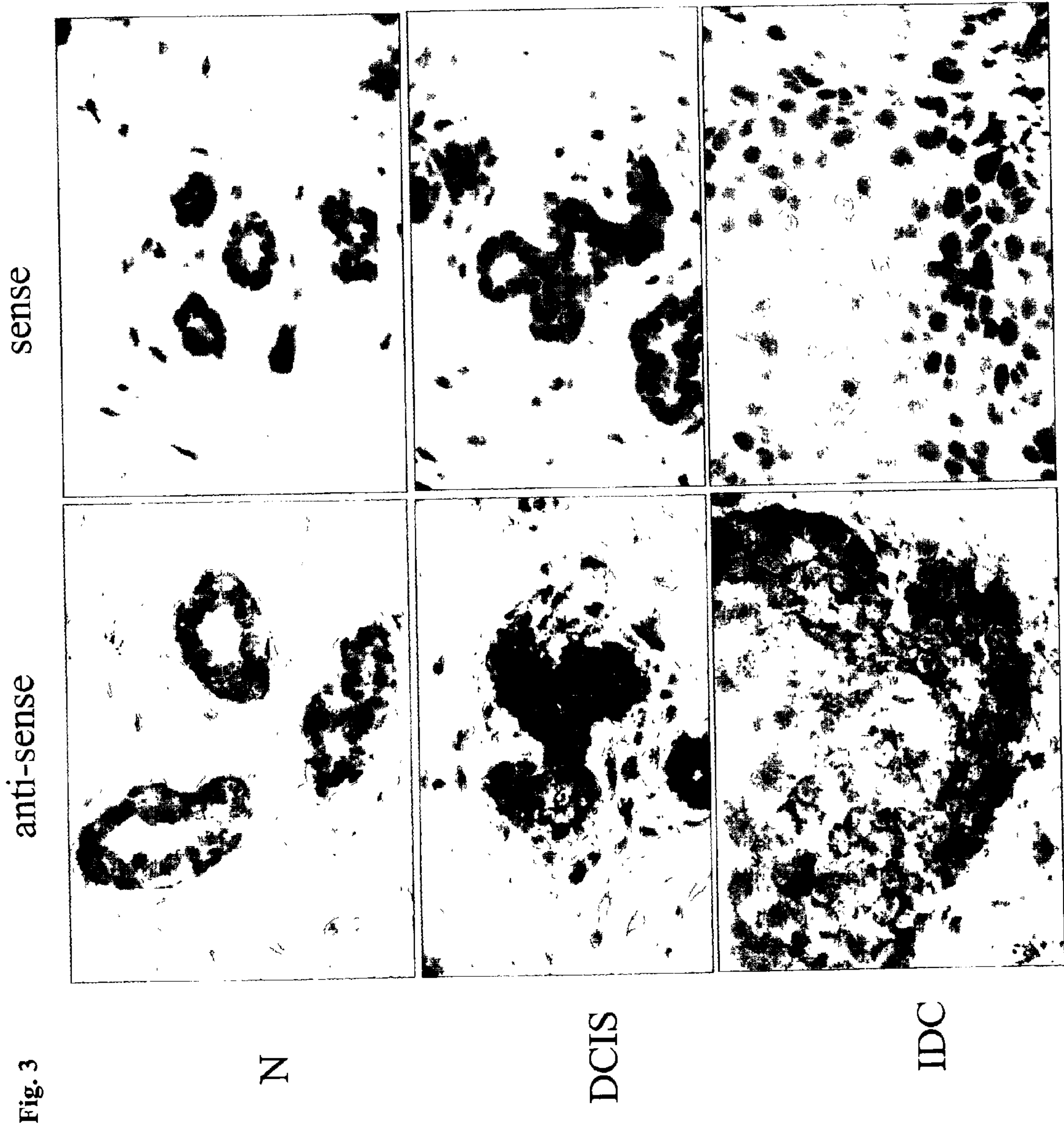


Fig. 3

BIOMARKERS FOR BREAST CANCER

RELATED APPLICATIONS

[0001] This application is a continuation-in-part of U.S. patent applications Ser. Nos. 10/028,018, filed Dec. 21, 2001, and 10/211,015, filed Aug. 1, 2002, which are hereby incorporated in their entireties as if fully set forth.

FIELD OF THE INVENTION

[0002] The invention relates to the identification and use of gene sequences which are differentially expressed in breast cancer. In particular, the invention provides the identities of three sets of sequences that may be used to identify the presence of breast cancer in tissue and cell samples. The expression of these sequences, whether embodied in nucleic acid expression, protein expression, nucleic acid amplification and/or activation, or other formats, are used in the diagnosis and/or treatment of breast cancer as well as for the study and/or determination of prognosis of a patient. When used for diagnosis, the expression levels of these sequences are used to identify the presence of breast cancer and provide guidance as to the treatment thereof.

BACKGROUND OF THE INVENTION

[0003] Breast cancer is by far the most common cancer among women. Each year, more than 180,000 and 1 million women in the U.S. and worldwide, respectively, are diagnosed with breast cancer. Breast cancer is the leading cause of death for women between ages 50-55, and is the most common non-preventable malignancy in women in the Western Hemisphere. An estimated 2,167,000 women in the United States are currently living with the disease (National Cancer Institute, Surveillance Epidemiology and End Results (NCI SEER) program, *Cancer Statistics Review* (*CSR*), www-seer.ims.nci.nih.gov/Publications/CSR1973 (1998)). Based on cancer rates from 1995 through 1997, a report from the National Cancer Institute (NCI) estimates that about 1 in 8 women in the United States (approximately 12.8 percent) will develop breast cancer during her lifetime (NCI's Surveillance, Epidemiology, and End Results Program (SEER) publication *SEER Cancer Statistics Review* 1973-1997). Breast cancer is the second most common form of cancer, after skin cancer, among women in the United States. An estimated 250,100 new cases of breast cancer are expected to be diagnosed in the United States in 2001. Of these, 192,200 new cases of more advanced (invasive) breast cancer are expected to occur among women (an increase of 5% over last year), 46,400 new cases of early stage (in situ) breast cancer are expected to occur among women (up 9% from last year), and about 1,500 new cases of breast cancer are expected to be diagnosed in men (Cancer Facts & Figures 2001 American Cancer Society). An estimated 40,600 deaths (40,300 women, 400 men) from breast cancer are expected in 2001. Breast cancer ranks second only to lung cancer among causes of cancer deaths in women. Nearly 86% of women who are diagnosed with breast cancer are likely to still be alive five years later, though 24% of them will die of breast cancer after 10 years, and nearly half (47%) will die of breast cancer after 20 years.

[0004] Every woman is at risk for breast cancer. Over 70 percent of breast cancers occur in women who have no identifiable risk factors other than age (U.S. General Accounting Office. Breast Cancer, 1971-1991: Prevention, Treatment and Research. GAO/PEMD-92-12; 1991). Only 5 to 10% of breast cancers are linked to a family history of breast cancer (Henderson IC, Breast Cancer. In: Murphy G P, Lawrence W L, Lenhard R E (eds). *Clinical Oncology*. Atlanta, Ga.: American Cancer Society; 1995:198-219).

[0005] The relationship between the expression of cellular factors and breast cancer has been an area of interest. STK15, a centrosomal protein kinase, has been observed as frequently amplified in breast cancer, and its quantitative expression levels positively correlate with tumor grade (Zhou, H. et al. Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation. *Nat Genet* 20, 189-93, (1998)). RAD51 has recently been shown to interact with the tumor suppressor BRCA1 (Chen, J. J., et al. BRCA1, BRCA2, and Rad51 operate in a common DNA damage response pathway. *Cancer Res* 59, 1752s-1756s (1999)), and its expression also positively correlates with tumor grade in breast cancer (Maacke, H. et al. Over-expression of wild-type Rad51 correlates with histological grading of invasive ductal breast cancer. *Int J Cancer* 88, 907-13 (2000)).

[0006] Citation of documents herein is not intended as an admission that any is pertinent prior art. All statements as to the date or representation as to the contents of documents is based on the information available to the applicant and does not constitute any admission as to the correctness of the dates or contents of the documents.

SUMMARY OF THE INVENTION

[0007] The present invention relates to the identification and use of gene sequences identified as differentially expressed in breast cancer. The sequences of two of the genes display increased expression in non-normal (or abnormal) breast cells, such as those that would be identified as atypical ductal hyperplasia (ADH), ductal carcinoma in situ (DCIS), and invasive ductal carcinoma (IDC) by standard pathology techniques based upon cytological criteria. The sequences of the third gene display decreased expression the same non-normal cells.

[0008] The first set of sequences found to be more highly expressed in non-normal breast cells are those of human cysteine-rich intestinal protein 1 (CRIP1 or hCRIP1, also known as human cysteine-rich heart protein or HCRHP). CRIP1 has been mapped to human chromosomal segment 7q11.23 (see Garcia-Barcelo et al. Genomics, 47(3):419-422, 1998).

[0009] The second set of sequences found to be more highly expressed in non-normal breast cells are those of the "hematological and neurological expressed sequence 1" (HN1 or Hn1). Murine HN1 has been identified as being expressed in mouse hemopoietic and brain tissues (see Tang et al. Mamm. Genome, 8:695-696, 1997).

[0010] The set of sequences found to be expressed at lower levels in non-normal breast cells are those expressed with a "second epithelium restricted Ets transcription factor" termed ESE-2 (see Oettgen et al., J. Biol. Chem., 274(41):29439-52, 1999). The coding region sequence of ESE-2b is identical to that of the E74 like factor 5, termed ELF5 (see Zhou et al., Oncogene, 17(21):2719-32, 1998).

[0011] The identified sequences may thus be used in methods of detecting the presence of non-normal breast cells

in a tissue or cell containing sample from a subject. The presence of non-normal breast cells may also be used in methods of diagnosing the presence of breast cancer in a tissue or cell containing sample from a subject. A subject, from which a sample is taken, may be one afflicted with, or suspected of having, breast cancer.

[**0012**] The present invention provides a non-subjective means for detecting the presence of non-normal breast cells. This provides advantages over the use of histomorphological or cytological criteria in standard pathology techniques, which requires some level of interpretation by a pathologist trained in assessing the presence and/or progression of breast cancer. The expression levels of these sequences may also be used as a means to assay small, node negative tumors that are not readily assessed by conventional means.

[**0013**] The expression levels of the identified sequences may be used alone or in combination with other sequences capable of identifying the presence of non-normal cells or of various stages and/or grades of breast cancer. Preferably, the sequences of the invention are used alone or in combination with each other.

[**0014**] The present invention provides means for correlating a molecular expression phenotype with a physiological (cellular) stage or state of a non-normal or abnormal breast cell. This correlation provides a way to molecularly diagnose and/or monitor a cell's status in comparison to normal breast cell phenotypes as disclosed herein. Additional uses of the sequences are in the classification of cells and tissues; determination of diagnosis and/or prognosis. Use of the sequences to identify cells of a sample as non-normal or abnormal may also be used to determine the choice, or alteration, of therapy used to treat such cells in the subject from which the sample originated.

[**0015**] The ability to identify non-normal and abnormal breast cells is provided by the recognition of the relevancy of the level of expression of the identified sequences and not by the form of the assay used to determine the actual level of expression. An assay may utilize a means related to the expression level of the sequences disclosed herein as long as the assay reflects, quantitatively or qualitatively, expression of the sequence. Preferably, however, a quantitative assay means is preferred. Identifying features of the sequences include, but are not limited to, unique nucleic acid sequences used to encode (DNA), or express (RNA), the disclosed sequences or epitopes specific to, or activities of, proteins encoded by the sequences. Alternative means include detection of nucleic acid amplification as indicative of increased expression levels (CRIP1 and HN1 sequences) and nucleic acid inactivation, deletion, or methylation, as indicative of decreased expression levels (ESE-2 and ELF5 sequences). Stated differently, the invention may be practiced by assaying one or more aspect of the DNA template(s) underlying the expression of the disclosed sequence(s), of the RNA used as an intermediate to express the sequence(s), or of the proteinaceous product expressed by the sequence(s). As such, the detection of the amount of, stability of, or degradation (including rate) of, such DNA, RNA and proteinaceous molecules may be used in the practice of the invention.

[**0016**] The practice of the present invention is unaffected by the presence of minor mismatches between the disclosed sequences and those expressed by cells of a subject's sample. A non-limiting example of the existence of such mismatches are seen in cases of sequence polymorphisms between individuals of a species, such as individual human patients within *Homo sapiens*. Knowledge that expression of the disclosed sequences (and sequences that vary due to minor mismatches) is correlated with the presence of non-normal or abnormal breast cells and breast cancer is sufficient for the practice of the invention with an appropriate cell containing sample via an assay for expression.

[**0017**] In one aspect, the invention provides for the identification of the expression levels of the disclosed sequences by analysis of their expression in a sample containing breast cells. In one preferred embodiment, the sample contains single cells or homogenous cell populations which have been dissected away from, or otherwise isolated or purified from, contaminating cells beyond that possible by a simple biopsy. Multiple means for such analysis are available, including detection of expression within an assay for global, or near global, gene expression in a sample (e.g. as part of a gene expression profiling analysis such as on a microarray) or by specific detection, such as quantitative PCR or real time quantitative PCR.

[**0018**] Preferably, the sample is isolated via non-invasive means. The expression of the disclosed sequence(s) in the sample may be determined and compared to the expression of said sequence(s) in reference data of non-normal breast cells. Alternatively, the expression level may be compared to expression levels in normal cells, preferably from the same sample or subject.

[**0019**] When individual breast cells are isolated in the practice of the invention, one benefit is that contaminating, non-breast cells (such as infiltrating lymphocytes or other immune system cells) are not present to possibly affect detection of expression of the disclosed sequence(s). Such contamination is present where a biopsy is used to generate gene expression profiles.

[**0020**] While the present invention has been described mainly in the context of human breast cancer, it may be practiced in the context of breast cancer of an animal known to be potentially afflicted by breast cancer by use of the corresponding sequences of the animal. Preferred animals for the application of the present invention are mammals, particularly those important to agricultural applications (such as, but not limited to, cattle, sheep, horses, and other "farm animals") and for human companionship (such as, but not limited to, dogs and cats).

BRIEF DESCRIPTION OF THE FIGURES

[**0021**] **FIG. 1** shows $\log_2$ plots of the ratio of expression in ADH, DCIS, and IDC cells to normal cells for CRIP1 and ELF5 sequences. The horizontal line is at "0" such that the ratio is "1" and all points above the line represent increases in expression relative to normal breast cells while all points below the line represent decreases in expression relative to normal breast cells.

[**0022**] **FIG. 2** shows the results of in situ hybridization with sense and anti-sense CRIP1 sequences to locate its expression at the cellular level in normal versus IDC cells of the same sample. CRIP1 signal localized to the epithelial cells, and its intensity was markedly increased in the IDC compartment of the same biopsy.

[0023] **FIG. 3** shows the results of in situ hybridization with sense and anti-sense CRIP1 sequences to locate its expression at the cellular level in normal versus DCIS and IDC cells. CRIP1 signal again localized to the epithelial cells, and its intensity was markedly increased in non-normal cells.

DETAILED DESCRIPTION OF THE SPECIFIC EMBODIMENTS

[0024] Definitions of terms as used herein:

[0025] A "sequence" or "gene sequence" as used herein is a nucleic acid molecule or polynucleotide composed of a discrete order of nucleotide bases. The term includes the ordering of bases that encodes a discrete product (i.e. "coding region"), whether RNA or proteinaceous in nature, as well as the ordered bases that precede or follow a "coding region". Non-limiting examples of the latter include 5' and 3' untranslated regions of a gene. It is appreciated that more than one polynucleotide may be capable of encoding a discrete product. It is also appreciated that alleles and polymorphisms of the disclosed sequences may exist and may be used in the practice of the invention to identify the expression level(s) of the disclosed sequences or the allele or polymorphism. Identification of an allele or polymorphism depends in part upon chromosomal location and ability to recombine during mitosis.

[0026] The terms "correlate" or "correlation" or equivalents thereof refer to an association between expression of one or more sequences and a physiologic state of a breast cell to the exclusion of one or more other states by use of the methods as described herein. The invention provides for the correlation between increases in CRIP1 and HN1 sequences and non-normal or abnormal breast cells. Similarly, the invention provides for the correlation between decreases in ESE-2/ELF5 sequences and non-normal or abnormal breast cells. Increases and decreases may be readily expressed in the form of a ratio between expression in a non-normal cell and a normal cell such that a ratio of one (1) indicates no difference while ratios of two (2) and one-half indicate twice as much, and half as much, expression in the non-normal cell versus the normal cell, respectively. Expression levels can be readily determined by quantitative methods as described below.

[0027] For example, increases in CRIP1 expression can be indicated by ratios of or about 1.1, of or about 1.2, of or about 1.3, of or about 1.4, of or about 1.5, of or about 1.6, of or about 1.7, of or about 1.8, of or about 1.9, of or about 2, of or about 2.5, of or about 3, of or about 3.5, of or about 4, of or about 4.5, of or about 5, of or about 5.5, of or about 6, of or about 6.5, of or about 7, of or about 7.5, of or about 8, of or about 8.5, of or about 9, of or about 9.5, of or about 10, of or about 15, of or about 20, of or about 30, of or about 40, of or about 50, of or about 60, of or about 70, of or about 80, of or about 90, of or about 100, of or about 150, of or about 200, of or about 300, of or about 400, of or about 500, of or about 600, of or about 700, of or about 800, of or about 900, or of or about 1000. A ratio of 2 is a 100% (or a two-fold) increase in expression. Similar ratios can be used with respect to increases in HN1 expression. Decreases in ESE-2/ELF5 expression can be indicaed by ratios of or about 0.9, of or about 0.8, of or about 0.7, of or about 0.6, of or about 0.5, of or about 0.4, of or about 0.3, of or about 0.2, of or about 0.1, of or about 0.05, of or about 0.01, of or about 0.005, of or about 0.001, of or about 0.0005, of or about 0.0001, of or about 0.00005, of or about 0.00001, of or about 0.000005, or of or about 0.000001. Non-limiting examples of such ratios are shown in **FIG. 1**.

[0028] A "polynucleotide" is a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides linked by phosphodiester bonds and encompasses the strand of a given sequence as disclosed herein as well as the complementary strand of a given sequence. The term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA and RNA as well as analogs thereof comprising a non-phosphodiester backbone. It also includes known types of modifications including labels known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, and internucleotide modifications such as uncharged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), as well as unmodified forms of the polynucleotide.

[0029] The term "amplify" is used in the broad sense to mean creating an amplification product can be made enzymatically with DNA or RNA polymerases. "Amplification," as used herein, generally refers to the process of producing multiple copies of a desired sequence, particularly those of a sample. "Amplification" may also be used in the context of DNA amplification wherein copies of coding sequences within the cellular genome are increased. "Multiple copies" mean at least 2 copies. A "copy" does not necessarily mean perfect sequence complementarity or identity to the template sequence. Methods for amplifying mRNA are generally known in the art, and include reverse transcription PCR (RT-PCR) and those described in U.S. patent application Ser. No. 10/062,857 entitled "Nucleic Acid Amplification" filed on Oct. 25, 2001 as well as U.S. Provisional Patent Application Nos. 60/298,847 (filed Jun. 15, 2001) and 60/257,801 (filed Dec. 22, 2000), all of which are hereby incorporated by reference in their entireties as if fully set forth.

[0030] By corresponding is meant that a nucleic acid molecule shares a substantial amount of sequence identity with another nucleic acid molecule. Substantial amount means at least 95%, usually at least 98% and more usually at least 99%, and sequence identity is determined using the BLAST algorithm, as described in Altschul et al. (1990), J. Mol. Biol. 215:403-410 (using the published default setting, i.e. parameters w=4, t=17). Alternatively, RNA may be directly labeled as the corresponding cDNA by methods known in the art.

[0031] A "microarray" is a linear or two-dimensional array of preferably discrete regions, each having a defined area, formed on the surface of a solid support such as, but not limited to, glass, plastic, or synthetic membrane. The density

of the discrete regions on a microarray is determined by the total numbers of immobilized polynucleotides to be detected on the surface of a single solid phase support, preferably at least about 50/cm^2, more preferably at least about 100/cm^2, even more preferably at least about 500/cm^2, but preferably below about 1,000/cm^2. Preferably, the arrays contain less than about 500, about 1000, about 1500, about 2000, about 2500, or about 3000 immobilized polynucleotides in total. As used herein, a DNA microarray is an array of oligonucleotides or polynucleotides placed on a chip or other surfaces used to hybridize to amplified or cloned polynucleotides from a sample. Since the position of each particular group of primers in the array is known, the identities of a sample polynucleotides can be determined based on their binding to a particular position in the microarray.

[0032] Because the invention relies upon the identification of sequences that are over- or under-expressed, one embodiment of the invention involves determining expression by hybridization of mRNA, or an amplified or cloned version thereof, of a sample cell to a polynucleotide of a disclosed sequence. Preferred polynucleotides of this type contain at least about 20, at least about 22, at least about 24, at least about 26, at least about 28, at least about 30, at least about 32, at least about 34, at least about 36, at least about 38, at least about 40, at least about 42, at least about 44, or at least about 46 consecutive bases of a sequence that is not found in other human sequences. The term "about" as used in the previous sentence refers to an increase or decrease of 1 from the stated numerical value. Longer polynucleotides may of course contain minor mismatches (e.g. via the presence of mutations) which do not affect hybridization to the nucleic acids of a sample. Such polynucleotides may be label to assist in their detection; alternatively, the nucleic acids to which such polynucleotides will hybridize may be labeled. Such polynucleotides may also be immobilized, such as by attachment to a solid support.

[0033] Even more preferred are polynucleotides of at least or about 50, at least or about 100, at least about or 150, at least or about 200, at least or about 250, at least or about 300, at least or about 350, at least or about 400, at least or about 450, or at least or about 500 consecutive bases of a sequence that is not found in other sequences in the human genome. The term "about" as used in the preceding sentence refers to an increase or decrease of 10% from the stated numerical value. The polynucleotides may of course contain minor mismatches which do not affect hybridization to the nucleic acids of a sample.

[0034] In another embodiment of the invention, all or part of a disclosed sequence may be amplified and detected by methods such as the polymerase chain reaction (PCR) and variations thereof, such as, but not limited to, quantitative PCR (QPCR), reverse transcription PCR (RT-PCR), and real-time PCR, optionally real-time RT-PCR. Such methods would utilize one or two primers that are complementary to portions of a disclosed sequence, where the primers are used to prime nucleic acid synthesis. The newly synthesized nucleic acids are optionally labeled and may be detected directly or by hybridization to a polynucleotide of the invention. The newly synthesized nucleic acids may be contacted with polynucleotides (containing sequences) of the invention under conditions which allow for their hybridization.

[0035] Alternatively, and in another embodiment of the invention, expression of a sequence may be determined by analysis of expressed protein encoded by said sequence in a cell sample of interest by use of one or more antibodies specific for one or more epitopes of the individual products (proteins) in said cell sample. Such antibodies are preferably labeled to permit their easy detection after binding to the gene product. In the case of a protein that may be found in the blood, serum or other bodily fluid, the assay may modified to use such materials in place of a breast cell containing sample.

[0036] The term "label" refers to a composition capable of producing a detectable signal indicative of the presence of the labeled molecule. Suitable labels include radioisotopes, nucleotide chromophores, enzymes, substrates, fluorescent molecules, chemiluminescent moieties, magnetic particles, bioluminescent moieties, and the like. As such, a label is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means.

[0037] The term "support" refers to conventional supports such as beads, particles, dipsticks, fibers, filters, membranes and silane or silicate supports such as glass slides.

[0038] As used herein, a "breast tissue sample" or "breast cell sample" refers to a sample of breast tissue or fluid isolated from an individual, preferably suspected of being afflicted with, or at risk of developing, breast cancer. Such samples are primary isolates (in contrast to cultured cells) and may be collected by a non-invasive means, including, but not limited to, ductal lavage, fine needle aspiration, needle biopsy, the devices and methods described in U.S. Pat. No. 6,328,709, or another suitable means recognized in the art. Alternatively, the "sample" may be collected by an invasive method, including, but not limited to, surgical biopsy.

[0039] "Expression" and "gene expression" include transcription and/or translation of nucleic acid material, such as the sequences of the invention.

[0040] As used herein, the term "comprising" and its cognates are used in their inclusive sense; that is, equivalent to the term "including" and its corresponding cognates.

[0041] Conditions that "allow" an event to occur or conditions that are "suitable" for an event to occur, such as hybridization, strand extension, and the like, or "suitable" conditions are conditions that do not prevent such events from occurring. Thus, these conditions permit, enhance, facilitate, and/or are conducive to the event. Such conditions, known in the art and described herein, depend upon, for example, the nature of the nucleotide sequence, temperature, and buffer conditions. These conditions also

depend on what event is desired, such as hybridization, cleavage, strand extension or transcription.

[0042] Sequence "mutation," as used herein, refers to any sequence alteration in the sequence of a gene disclosed herein interest in comparison to a reference sequence. A sequence mutation includes single nucleotide changes, or alterations of more than one nucleotide in a sequence, due to mechanisms such as substitution, deletion or insertion. Single nucleotide polymorphism (SNP) is also a sequence mutation as used herein. Because the present invention is based on the relative level of sequence expression, mutations in non-coding regions of genes as disclosed herein may also be assayed in the practice of the invention. "Detection" includes any means of detecting, including direct and indirect detection of gene expression and changes therein. For example, "detectably less" products may be observed directly or indirectly, and the term indicates any reduction (including the absence of detectable signal). Similarly, "detectably more" product means any increase, whether observed directly or indirectly.

[0043] Increases and decreases in expression of the disclosed sequences are defined in the following terms based upon percent or fold changes over expression in normal cells. Increases may be of 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, or 200% relative to expression levels in normal cells. Alternatively, fold increases may be of 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, or 10 fold over expression levels in normal cells. Decreases may be of 10, 20, 30, 40, 50, 55, 60, 65, 70, 75, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 99 or 100% relative to expression levels in normal cells.

[0044] Unless defined otherwise all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs.

[0045] Specific Embodiments

[0046] The present invention relates to the identification and use of three sets of sequences for the detection of non-normal and cancerous breast cells. The differential expression of these sequences in non-normal or abnormal breast cells relative to normal breast cells is used to identify a breast cancer cell as being non-normal or otherwise abnormal. The invention is advantageously used to identify breast cells as being those of ADH, DCIS, and IDC as otherwise determinable standard pathological techniques. The invention may also be applied to the identification of breast cells as being those of atypical lobular hyperplasia (ALH), lobular carcinoma in situ (LCIS), and invasive lobular carcinoma (ILC) as otherwise determinable by standard pathological techniques.

[0047] Other non-limiting examples of non-normal or abnormal cells include malignant cells, atypical cells (including reactive and pre-neoplastic), neoplastic cells, tumor cells, and cancer or cancerous cells.

[0048] The sequences(s) identified by the present invention are expressed in correlation with non-normal breast cells, and thus negatively correlated with normal breast cells. For example, CRIP1, identified by I.M.A.G.E. Consortium CloneID 1323448 and cluster NM_001311 ("The I.M.A.G.E. Consortium: An Integrated Molecular Analysis of Genomes and their Expression," Lennon et al., 1996, Genomics 33:151-152; see also image.llnl.gov) has been found to be useful in discriminations between normal and ADH or DCIS or IDC breast cells.

[0049] In preferred embodiments of the invention, any sequence, or unique portion thereof, of the CRIP1 sequences identified by the cluster, as well as UniGene Homo sapiens cluster Hs.17409, maybe used. The consensus sequence of the I.M.A.G.E. Consortium cluster is as follows, with the assigned coding region (ending with a termination codon) underlined and preceded by the 5' untranslated and/or non-coding region and followed by the 3' untranslated and/or non-coding region:

```
(consensus sequence for CRIP1)
GGCACGAGGGCCCGTGCCGCCCCAGCCGCTGCCGCCTGCACCGGACCCGGAGCCGCC      SEQ ID NO:1

ATGCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCGAGAGGGTGACCTCTCTGGG

CAAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGG

GCCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGG

CCTAAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTGTTGCCCCTAATGCTCTCAGTAAACCTGAACAC

TTGGAAAAAAAAAAAAAAAA
```

[0050] The sequences identified as belonging to the I.M.A.G.E. Consortium and UniGene clusters, with the assigned coding region underlined, follow below. The 5' and 3' untranslated and/or non-coding regions are by reference to the assigned coding region, which is presented as the complementary strand in the 3' to 5' direction for some of the sequences.

(CloneID 5103334)
GGGGATTGGAGATGTTCCCCTCATGGAGGGTGCTGAGGACCTTAGGGTGGGCTGCCAGGCTG SEQ ID NO:2

GGCGGATGCGGGCTAAGTGCACAGGGCCTTGGGCAGAGCTGGCTGCAAGAGGCGGGTACGCC

AGTGGTGGGTAGGCGCCCGCG

TCCTGCAGCGTCTCACCGGGGCCTGTCTGTGCCTCTGCAGCCGAGAGGGTGACCTCTCTGGG

CAAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGG

GCCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGG

CCTAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGAGCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTGTTGCCCCTAATGCTCTCAGTAAACCTGAACAC

TTGGAAAAAAAAAAAA

(CloneID 5777677)
GCACGAGCGCTGGGCTAGGGGCGCGGCTTGAACTCGCCTAAAGAGCTGCGCCCTCTCATCTC SEQ ID NO:3

GCGCCTGCAGCCCGTGCCGCCCCAGCCGCTGCCGCCTGCACCGGACCCGGAGCCGCC

ATGCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCCAGAGGGTGACCTCTCTGGG

CAAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTCGGG

GCCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGG

CCTAAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTCCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTGTTGCCCCTAATGCTCTCAGTAAACCTGAACAC

TTGG

(CloneID 563289)
GCACGAGCGCTGGGCTACGGGCGCGGCTTGAACTCGCCTAAAGAGCTGCGCCCTCTCATCTC SEQ ID NO:4

GCGCCTGCACCCCGTGCCGCCCCAGCCGCTGCCCCCTGCACCGGACCCGGAGCCGCC

ATCCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCGAGAGCGTGACCTCTCTGGG

CAAGGACTGGCATCGGCCCTGCCTGAAGTGCCAGAAATGTGGGAAGACGCTGACCTCTGGGG

GCCACGCTGACCACGAACGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGG

CCTAAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTCTTGCCCCTAATGCTCTCAGTAAACCTGAACAC

T

(CloneID 1627147 in 3' to 5' orientation)
TGGTTTTCCAAGTGTTCAGGTTTACTGAGAGCATTAGGGGCAACAAGGGAGCCCTGCGCATC SEQ ID NO:5

TGGGGACAAGGCCTGGCATTTGCCTGGACAGTGGCCCTGCAAGCAGCCAAGGATGGGGTCTC

CACCACCTGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCCGCCCAAAGCCTTTAGGCCCAAACATGGCTA

CGTAGCAGGGGTGGTTGCAGTAGGGTTTGCCTTCGTGCTCAGCGTGCCCCCCAGAGGTCAGC

GTCTTCCCACATTTCTCGCACTTCAGGCAGGGCCGATGCCAGTCCTTGCCCAGAGAGGTCAC

CCTCTCGGCGAAGTACACCTCCTTGTTGCACTTGGGACACTTGGGCAT

GGCGGCTCCGGGTCCGGTGCAGGCGGCAGCGGCTGGGCCGGCACGGGCTGCAGGCGCGACAC

-continued

TGGGTGGATCCGGCCTGGTCCGCGCCTTTCAGGGACCCCGGGACCCCGCCCCTTGGAGACAC

CGCCCCTC

(CloneID 1627139 in 3' to 5' orientation)
TGGTTTTCCAAGTGTTCAGGTTTACTGAGACCATTAGGGGCAACAAGGGAGCCCTGGGCATC SEQ ID NO:6

TGGGGACAAGGCCTGCCATTTGCCTGGACAGTGGCCCTGCAAGCAGCCAAGGATGGGGTCTC

CACCACCTGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCCCCCCGCCCAAAGCCTTTAGGCCCAAACATGGCTA

CGTAGCAGGGGTGGTTGCAGTAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGGTCAGC

GTCTTCCCACATTTCTCGCACTTCAGGCAGGGCCGATGCCAGTCCTTGCCCAGAGAGGTCAC

CCTCTCGGCCAAGTACACCTCCTTGTTGCACTTGGGACACTTGGGCAT

GGCGGCTCCGGGTCCGCTGCAGGCGGCAGCGGCTGGGGCGGCACGGGCTGCAGGCGCGAGAC

TGGGTGGATCCGGCCTGGTCCGCGCCTTTCA

(CloneID 3296101 in 3' to 5' orientation)
TTGGTTTTCCAAGTGTTCAGGTTTACTGAGAGCATTAGGGGCAACAAGGCAGCCCTGCGCAT SEQ ID NO:7

CTGGGGACAAGGCCTGGCATTTGCCTGGACAGTGGCCCTGCAAGCAGCCAAGGATGGGGTCT

CCACCACCTGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCCGCCCAAAGCCTTTAGGCCCAAACATGGCTG

CGTAGCAGGGCTGGTTGCAGTAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGGTCAGC

GTCTTCCCACATTTCTCGCACTTCAGGCAGGGCCGATGCCAGTCCTTGCCCAGAGAGGTCAC

CCTCTCGGCGAAGTACACCTCCTTGTTGCACTTGGGACACTTGGGCAT

GGCGGCTCCGGGTCCGGTGCAGCCGGCAGCGGCTGGGGCGGCACGCGCTGCATGCGCGAGAC

TGGGTGGATCCGGCCTGGTCCGCGCCTTTCAGGGACCCCGGGACCCCG

(CloneID 3631097)
GCCGCCTGCACCGGACCGGAGCCGCC SEQ ID NO:8

ATGCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCGAGAGGGTGACCTCTCTGGG

CAAAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGTG

GGGCCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTG

GGCCTAAAGGCTTTGGGCGGGGCGGAGCCGAGGAGCCACACTTTCAAGTAA

AACCAGGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCCACTTGTCCAGGGCAAAAT

TGCCAGGCCTTTGTCCCCCACAATGCCCAGGGGCTCCCTTTGTTTGGCCCCCTTAATTGCTC

TCAGTAAACCCTTGAACACTTGGA

(CloneID 2782288)
CAGCCCGTGCCGCCCCAGCCGCTGCCGCCTGCACCGGACCCGGAGCCGCC SEQ ID NO:9

ATGCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCAAGAGGGTGACCTCTCTGGG

CAAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGG

GCCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCACCCATGTTTGGG

CCTAAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTCCAGGGCCACTGTCCAGGCAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTGTTGCCCCTAATGCTCTCAGT

-continued

```
(CloneID 3903337)
CACCGGACCCGGACCCGGCC                                          SEQ ID NO:10

ATGCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCGAGAGGGTGACCTCTCTGGG

CAAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGG

GCCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGG

CCTAAAGGCTTTCGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCCTTGTTGCCCCCTAAATGCTCTCCAGGTAAAACCT

GAAACACTTGGAAAAAAAAAACAAAAAAAAAAA

(CloneID 2063820 in 3' to 5' orientation)
TCCAAGTGTTCAGGTTTACTGAGAGCATTAGGGGCAACAAGGGACCCCTGGGCATCTGGGGA  SEQ ID NO:11

AAAGGCCTGGCATTTGCCTGGACAGTGGCCCTGCAAGCAGCCAACGATGGGGTCTCCACCAC

CTGG

T

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCCGCCCAAAGCCTTTAGGCCCAAACATCGCTA

CGTAGCAGGGGTGGTTGCAGTAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCACAGGTCAGC

GTCTTCCCACATTTCTCGCACTTCAGGCAGGGCCGATGCCAGTCCTTGCCCAGAGAGGTCAC

CCTCTCGGCTGCAGAGGCACAGACAGGCCCCGGTGAGACGCTCCAGGA

CGCCGCGCCTACCCACCACTGGCGTACCCGCCTCTTGCAGCCAGCTCTGCCCAAGGCCCTGT

GCACTTAGCCCGCATC

(CloneID 2568304 in 3' to 5' orientation)
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCAAGGGTTCAGGTTTACTGAAAGCATTAGCGG  SEQ ID NO:12

CAACAAGGGAGCCCTGGGCATTTGGGGACAAGGCCTGGCATTTCCCTGGACAGGGGCCCTGC

AAGCAGCCAAGGATGGGGTCTCCACCACCTGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCCGCCCAAAGCCTTTAGGCCCAAACATCGCTG

CGTAACAGGGGGGGGTTGCANTACGGTTTGCCTTCGGGCTCAGCGTGGCCCCCAAAGGTCAG

CGTCTTCCCACATTTTTCGCACTTCAGGCAGGGCCGATGCCAGTCCTTGCCCAAAAACGTCA

CCCTCTCGGCAAAGTACACCTCCTTGTTGCACTTGGGACACTTGGGCAT

GGCGGCTCCGGGTCCGGTGCAGGCGCCAACGGCTGGGGCGGCA

(CloneID 2329226 in 3' to 5' orientation)
CCAAGTGTTCAGGTTTACTGAGAGCATTAGGGGCAACAAGGGAGCCCTGGGCATCTGGGGAC  SEQ ID NO:13

AAGGCCTGGCATTTGCCTGGACAGTGGCCCTGCAAGCACCCAAGGATGGGGTCTCCACCACC

TGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCCCCCCGCCCAAAGCCTTTAGGCCCAAACATGCCTA

CGTAGCAGGGGTGGTTGCAGTAGGGTTTGCCTTCGTGCTCACCGTGGCCCCCAGAGGTCAGC

GTCTTCCCACATTTCTCGCACTTCAGGCAGGGCCGATGCCAGTCCTTGCCCAGAGACGTCAC

CCTCTCGGCGAAGTACACCTCCTTGTTGCACTTGGGACACTTGGGCAT

GGCGGCTCCGGGTCCGGTGCAGGCGGCACCGGCTGGGGCCGCACGGGCTGCAGGCGCGAGA

(CloneID 5433206)
GGCCCCTGCCGCCCCAGCCGCTGCCGCCTGCACCGGACCCGGAGCCGCC             SEQ ID NO:14

ATGCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCGAGACGGTGACCTCTCTGGG
```

-continued

CAAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGG

GCCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGG

CCTAAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTGTTGCCCCTAATGCTCTCAGTAAACCTGAACAC

TTGGAAACAAACACACAAAAAAAAAAACAAAACAAAAAAAACACAAACCAAAAAACACGAGA

AAAGCACAAAAAAAGAGGCAAAGAGAAAAACGCGCGCGGGATATCAGAGGCAGGAGGGGCGA

AAAGGGGGGAGACAGAGGAGGAACAGCGACAACCGGCACCGCGCGC

(CloneID 5220536)
GCCGTGCCGCCCAGCCGCTGCCGCTGCACCGGACCCGGAGCCGC SEQ ID NO:15

ATGCCCAAGTGTCCAAGTGCAACAAGGAGGTGTACTTCGCCGAGAGGGTGACCTCTCTGGGC

AAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGGG

CCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGGC

CTAAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAAA

CCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCCAGATGCCCAGGGCTTCCCTTTGTTGCCCCTAATGGCTCTCAGTAAACCTTG

AACAACTTGGAAACACCACACCAAAAACCAACACACAGGG

(CloneID 2387987 in 3' to 5' orientation)
CAAGTGTTCAGGTTTACTGAGAGCATTAGGGGCAACAAGGGAGCCCTGGGCATCTGGGGACA SEQ ID NO:16

AGGCCTGGCATTTGCCTGGACAGTGGCCCTGCAAGCAGCCAAGGATGGGGTCTCCACCACCT

GGT

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCCGCCCAAAGCCTTTAGGCCCAAACATGGCTG

CGTAGCAGGGGTGGTTGCAGTAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGGTCAGC

GTCTTCCCACATTTCTCGCACTTTAGGCAGGGCCGATGCCAGTCCTTGCCCAGAGAGGTCAC

CCTCTCGGCGAAGTACACCTCCTTGTTGCACTTGGGACACTTGGGCAT

GGCGGCTCCGGGTCCGGTGCAGGCGGCAGCGGCTGGGGCGGCACGGGCTGCAGGCGCGAGAT

GAGAGGGCGCAGCTCTTTAGGCGAGTTCAAGCCGCGCCCCTAGCCCAGCG

(CloneID 5186252)
TGCAGCCGTCCGCCCAGCCTGGTGCGCTGCACGGACCCGGAGCCGCAT SEQ ID NO:17

GCGACGAGTGTCCAATGTGCAACAAGGAGGTGTACTTCGCCGAGAGGGTGACCTCTCTGGGC

AAAGACTGGCATCTGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGG

GCCACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGG

CCTAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATTGCCAG

GGCCTTGTCCCCCAGATTGCCCAGGGCTCCCTCTTGTTGCCCCCTAAATTGCTCTCAGGTTA

GAACCTTGAAACAGCTTTGCGAACAAGCAAAAGAAGAGAAGGAGGTGCTGATAATAGCAGAC

GAAGTGAGAAGAGAACACCAATCACATGAACACGAATAGGCAGGCCATACAACGCAGGAGAT

ATCGAACCCGGCGATAGTGCCCGGT

-continued

(CloneID 3609948)
CGGACCCGGAGCCGCC SEQ ID NO:18

ATGCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCGAGAGGGTGACCTCTCTGGG

CAAGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGG

GCCACGCTGAGCCAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGGCC

TAAGCCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTGTTGCCCCTAATGCTCTCAGTAAACCTGAACAC

TTGGA

(CloneID 1584777 in 3' to 5' orientation)
CCAAGTGTTCAGGTTTACTGAGAGCATTAGGGGCAACAAGGGAGCCCTGGGCATCTGGGGAC SEQ ID NO:19

AAGGCCTGGCATTTGCCTGGACAGTGGCCCTGCAAGCAGCCAAGGATGGGGTCTCCACCACC

TGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCGCCCAAAGCCTTTAGGCCAAACATGGCTGCG

TAAGAGAGTGGTTGAAGTAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGGTGAGCGTC

TTCCCACATTTCTCGCAGTTCAGGCAGGGCAGAGTGCAGTCCTTGCCCAGAGAGGGGACCTC

TCGGCGAAGAGAACTCCTTGTTGCACTTGGGACACTTGGGCAT

GGCGGCTCCGGGTCCGGTGCAGGCGGC

(CloneID 2130246 in 3' to 5' orientation)
AGCGGCCGCCCTTTTTTTTTTTTTTTTTCCAAGTGTTCAGGTTTACTGAGAGCATTAGGGG SEQ ID NO:20

CAACAAGGCAGCCCTGGGCATCTGGGGACAAGGCCTGGCATTTGCCTGGACAGTGGCCCTGC

AAGCAGCCAAGGATGGGGTCTCCACCACCTGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCCGCCCAAAGCCTTTAGGCCCAAACATGGCTG

CGTAGCAGGGGTCGTTGCAGTACGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGGTCAGC

GTCTTCCCACATTTCTCGCACTTCAGGCAGGGCCGATGCCAGTCCTTGCCCAGAGAGGTCAC

CCTCTCGCCGAAGTACACCTCCTTGTTGCACTTGGGACAC

(CloneID 1908782 in 3' to 5' orientation)
CCAAGTGTTCAGGTTTACTGAGACCATTAGGGGCAACAAGGGAGCCCTGGGCATCTGGGGAC SEQ ID NO:21

AAGGCCTGGCATTTGCCTGGACAGTGGCCCTGCAAGCAGCCAAGGATGGGGTCTCCACCACC

TGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCCGCCCAAAGCCTTTAGGCCCAAACATGGCTG

CGTAGCAGGGGTGGTTGCAGTAGGGTTTGCCTTCGTCCTCAGCGTGGCCCCCACAGGTCAGC

GTCTTCCCACATTTCTCGCACTTCAGGCAGGGCCGATGCCAGTCCTTGCCCAGAGAGGTCAC

TTCTCTCGGCGAAGTCCACCCTCCTTGTTGCACCTTGGGACACCTT

(CloneID 1323448 in 3' to 5' orientation)
GAGCATTAGGGGCAACAAGGGAGCCCTGGGCATCTGGGGACAAGGCCTGGCATTTGCCTGGA SEQ ID NO:22

CAGTGGCCCTGCAAGCAGCCAAGGATGGGGTCTCCACCACCTGGT

TTACTTGAAAGTGTGGCTCTCGGCTCCGCCCCCCCAAAGCCTTTAGCCCAAAACATGGCTGC

GTACGAGGGTGGTTGCAGTAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGGTCAGCGT

CTTCCCACATTTCTCGCACTTCAGCCAGGGCGANTGCCAGTCCTTGCCCAGAGAGGTCACCC

TCTCGGCGAAGTACACCTCCTTGTTGC

-continued

(CloneID 1837171 in 3' to 5' orientation)
TTTGGTTTTCCAAGTGTTCAGGTTTACTGAGAGCATTAGGGGCAACAAGGCACCCCTGGGCA SEQ ID NO:23

TCTGGGGACAAGCCCTGGCATTTGCCTGGACAGTGGCCCTGCAAGCAGCCAAGGATGGGGTC

TCCACCACCTGGT

TTACTTCAAACTGTGGCTCTCGGCTCCGCCCCGCCCAAAGCCTTTAGGCCCAAACATGGCTG

CGTAGCAGGGGTGGTTGCAGTAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGGTCAGC

GTCTTCCCACATTTCTCGCACTTCAGGCAGCGCCGATGCCAGTCCTTGCCCAGAGAGGTCAC

CCTCTCGGCGAAGTACACCTCCTTGTTGCACTTGGGACACTTGGG

(CloneID 2802267)
GCGTCGACCCCCGAGAGGGTCACCTCTCTGGGCAAGGACTGGCATCGGCCCTGCCTGAAGTG SEQ ID NO:24

CGAGAAATGTGGGAAGACGCTGACCTCTGGGGGCCACGCTGAGCACGAAGGCAAACCCTACT

GCAACCACCCCTGCTACGCAGCCATGTTTCGGCCTAAAGGCTTTGGGCGGGGCCGAGCCGAG

AGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCGCTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTGTTGCCCCTAATGCTCT

(CloneID 2847309)
AAGGTAGCGAAGCGAGAGGGTGACCTCTCTGGGCAAGGACTGGCATCGGCCCTGCCTGAAGT SEQ ID NO:25

GCGAGAAATGTGGGAAGACGCTGACCTCTGGGGGCCACGCTGACCACGAAGGCAAACCCTAC

TGCAACCACCCCTGCTACGCAGCCATGTTTGGGCCTAAAGGCTTTGGGCGGGGCGGAGCCGA

GAGCCACACTTTCAAGTAA

ACCAGGTGGTGGACACCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATCCCCAGGGCTCCCTTGTTGCCCCTAATGCTCTCAGTAAACCTGAACAC

TTGGAAAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAATTTCN

(CloneID 5018693)
GCCCAAGTGTCCCAAGTGCAACAAGGAGGTGTACTTCGCCGAGAGGGTGACCTCTCTGGGCA SEQ ID NO:26

AGGACTGGCATCGGCCCTGCCTGAAGTGCGAGAAATGTGGGAAGACGCTGACCTCTGGGGGC

CACGCTGAGCACGAAGGCAAACCCTACTGCAACCACCCCTGCTACGCAGCCATGTTTGGGCC

TAAAGGCTTTGGGCGGGGCGGAGCCGAGAGCCACACTTTCAAGTAA

ACCAGGTGGTGGAGACCCCATCCTTGGCTGCTTGCAGGGCCACTGTCCAGGCAAATGCCAGG

CCTTGTCCCCAGATGCCCAGGGCTCCCTTGTTGCCCCTAATGCTCTCAGTAAACCTGAACAC

TTGGAAAAGCAAAAAAAAAAAAAAAAAA

(CloneID 155219 in 3' to 5' orientation wherein the coding
and 3' untranslated and/or non-coding regions are present or
similar to that in CloneID 563289)
GCATTAGGGGCAACAAGGGAGCCCTGGGCATCTGGGGACAAGGCCTGGCATTTGCCNGGACA SEQ ID NO:27

GNGGCCCTGCAAGCAGCCAAGGATGGGGNCTCCACCACCTGGN

TNACTTGAAAGTGTGGCTCTCGGCTCCGCCCGCCCAAAGCCTTNAGGNCCAAACATGGCTGC

GNANAGGNGNTGGTTGCAGNAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGGTCAGCG

TCTTCCCACATTTCTCGCACTTCAGGCAGGGCCANTGCCANTCCTTNCCCAGAGAGGTCACC

CTCTCGGCGAAGTACACC

-continued

(CloneID 307716 in 3' to 5' orientation wherein the coding
and 3' untranslated and/or non-coding regions are present or
similar to that in CloneID 563289) 307716 (3')
TTTTTTTTTTTTTTCCAAGTGTTCAGGTTTACTGAGAGCATTAGGGGCAACAAGGGAGCCCT SEQ ID NO:28

GGNATCTGGGGACAAGGCCTGGCATTTCCTGGACAGTGGCCCTGCAAGCAGCCAAGGATGGG

GTCTCCACCACCTGGT

TTACTTGAAAGTGTGGCTCTCGCCTCCGCCCGCCCAAAGCCTTTAGNCNCAAACATGGATAT

GGCTACGTAGCAGGGGTGGTTGCANTAGGGTTTGCCTTCGTGCTCAGCGTGGCCCCCAGAGG

TCAGCGTCTTCCCACATTTCTCGCACTTCAGGCAGGGCNAATGCCANTCCTTGCCCAGAGAG

GTCACCCTCTCGGG

[0051] In another set of preferred embodiments of the invention, any sequence, or unique portion thereof, of the HN1 sequences identified by the I.M.A.G.E. Consortium CloneID 471568 and cluster NM_016185, as well as UniGene Homo sapiens cluster Hs.109706, may be used. HN1 sequences are useful in discriminations between normal and DCIS or IDC (of grades I to III) breast cells. Its expression is also increased in grade III relative to grade I breast cancer cells.

[0052] The consensus sequence of the I.M.A.G.E. Consortium cluster is as follows, with the assigned coding region (ending with a termination codon) underlined and preceded by the 5' untranslated and/or non-coding region and followed by the 3' untranslated and/or non-coding region:

(consensus sequence for HN1)
TGCAGCGGTGGTCGGCTGTTGGCTCTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGA SEQ ID NO:29

GCGTTCTCCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTCCCAGGCAGCCTGCGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCCTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCCTTCTGTCTGTTTCCTCCATGCTTGAGAACTGCACAACT

TGAGCCTGACTGTACATCTTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTAAAAAA

AAAAAAAAAAAA

[0053] The sequences identified as belonging to the I.M.A.G.E. Consortium and UniGene clusters, with the assigned coding region underlined, follow below. The 5' and 3' untranslated and/or non-coding regions are by reference to the assigned coding region, which is presented as the complementary strand in the 3' to 5' direction for some of the sequences.

(Clone ID 4795778)
AGCGGGGCGGCTCCTGCAGCGGTGGTCGGCTGTTGGGTCTGGAGTTTCCCAGCGCCCCTCGG SEQ ID NO:30

GTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

-continued

GCGGCCTCGAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGGCCCCAACCCACAGGGAGTCTCTCCGCAGAGCCTTCTTGGTG

TTGCCCTAACTTGCCAGTGGCCTTTGCTCAG

AGCCTCCTCCTGTGACATGTGAACAATGAAGAGGCCTGCGCCTCCTGCCTTGCCGCCTGCAA

AGCAAAGAAACTGCCTTTTATTTTTTAACCTTAACAAGTAGCCAGATAGTAACAAGACTGGC

TGGCTGATGAGCAAAGCCTTTGCTCTCACGCAGAGGACGGCTTGGATGTACAATGAAACTGG

CTGGAACTAAAGCAGTGAAGCAGGGGAGGCCATCACACTGAAGCGGGTCTTCCTCCAGGAAC

GGGTCCACAAGGGTGTGTCTGAATTACCTGATGCTGTGTGCTGATGCTGGCTCTTGACCATG

GACGGCAAGTTCATCTTAACCGTGCTGTCCTCACACCTGACTGTGCTCCCTTAACATTTCCC

(Clone ID 5239921)
CTCCTGCAGCCGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCT SEQ ID NO:31

TTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGCCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTGTTATCCCTTTGGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGGC

CATGGCATGAGTGTTCCTAGTAGTAGATTGGAGGGAAAAGTTTGTGACCCTTAGTACGGGGT

TTTAAGACGAAATAATTGGGTTCCA

(CloneID 4623018)
GGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCA SEQ ID NO:32

GCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

-continued

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTATTTTTTGCTTTGTGAAGACCAGGTTTCTCTCTGTCCTTGACCTCTGTGGGTCTGTG

GGCCATGGCATGAGTGTTTCCTAGTAGTAGATGGGAGGGAAAGCTTGTGACCCCTTAGTACT

GTGTTTTTACAC

(CloneID 5729213)
CGGTCCGGAATTCTCCGGATGCTGTTGGGTGTGGAGTTTCTACCGCCCCTCGGGTCCGACCC SEQ ID NO:33

TTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGGC

CATGGCATGAGTGTTTTCTAGTAGTAGATTGGAGGGAAAGCTTTGTGACACTTAGTACTGTG

TTTTTAAGAAGAAATAATTTGGTTCCAGATGTGTTANAAGATCTTTTGTACTGGAGTTTTTT

ACACTTTTACTTGGGTTTAACCAGCCTCAACTGGGACGACCATAAACAGTCCACCAGGCACC

GTTTCCTGCCAGGCCCCAAACCCACGGGAAGTCTCTCGCAAAAACCCCTTCCTGGGGGTTGC

CCCTAATTTGCCAAGGGGCCTTTTGCTCCGGAGCCTCCCTCCCGGGGAATTGGGGAAACAAT

GAAAAAGGCCTGTGCCCTCCGTGCCTTG

(CloneID 5192505)
CTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTG SEQ ID NO:34

AGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAACAGAAGCACTTTATGTACTGCT

GTCTTATATCCTTTTGAAGAACAGGGTTTCTCTCTTGTCCTTGACCTCTGGGGTCTGTGGGC

-continued

CCATGGCATGAAGTGGTTTCTAGTAGTTAGATGGCGCGCAAAGCTTTTGTGGACCCTATGTA

ACTGGTGGTTTTTACGACAGACACTACTTTGGCTACCAGATGTGTTAGAGGAACCTTGGCAC

TGGAGGTTTAACCACTTTACTGGGTTTACCAGGCCTTAGTGGCACGGCCCATAACGTCCACA

GCCCCGATCCTGCCAGCCCAA

(CloneID 3945044)

TTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCC SEQ ID NO:35

AGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGCAAGAACAAAATGGCCTCTAATATCTTGGGACACCTGAACAAAATCAAGCTTCTTGCGCC

AAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGAG

AAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATTC

ATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCT

GCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTGG

CGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTGTGGATTGTTCTCATTAAAAAGAAGCACTTTATGTACTGCT

GTCTTTTTTTTTCCTTTTTGAAGAACCAGGTTTCTCTCTGTCTTGACTCTGGGGTCTGTGGG

GCCATGGCATGAGTGTTTCTAGTAGTACATTGGCGGGAAACGCTTTGGGACCCTTTACTACC

TGTGTTTCCACGGAACAAATTACTTTGGGGTCCCCATGTTGTTAA

(Clone ID 5240530)

CTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTGTG SEQ ID NO:36

GGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGCAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTNGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTAATTTTACTTTTGAAGAACAGCTTCTCTCTGTCCTTGACTCTGGGTCTGGGGCCAT

GG

(CloneID 3617626)

TGGCTAGGTACGAGGCTGGGTTTGGGCGGACAGGCGGCAGCGGCGGCTCCTGCAGCGGTGGT SEQ ID NO:37

CGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCC

GGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

-continued

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTCCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAGTCCAGCCTCGTCTTGGGTATAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAATGCACAACTT

GAGCCTGACTGTCATCTCTGGGATTTGTTCATTAACAGAGCCTTATGTAAAAAAAAAAAAAA

AACGCGGATTCTTCGGGGGGGAACAAGAAGAGGCCAGGGATA

(CloneID 4814114)
AGCGGGGGGGCGGCTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCT SEQ ID NO:38

CGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCTCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCACCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTCTTTCGGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGGCCAT

GGCATTAGT

(CloneID 4894047)
TGATGGCAGTAATCTCGGGGTATAAGGTAGCGAGGCTGGGTGGCGCCTGCAGCGGTGGTCGG SEQ ID NO:39

CTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGC

GCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGNGCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTTGTGAACTGGACAAG

TTGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCT

-continued

GTCTATTCAGACCTTTTGAAAGACCAGGTTTCTCTCTGTCCTTGACTCTGGGGCCTGTGGGC

CATGGACTGAGTGTTTCCTAGATAGCAGAT

(CloneID 5531619)
AATTTTCCGAAATCGTCGACCCACGCGTCCGTTTCCAGCGGTGGTCGGCTGTTGGGTGTGGA SEQ ID NO:40

GTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCG

CTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTAAAAAAAAAAA

AAA

(CloneID 5588318)
CGGTCCGGAATTCTCCGGATGCTGTTGGGTGTGGAGTTTCCCTGCGCCCCTCGGGTCCGACC SEQ ID NO:41

CTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTCTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGNGTCTGTGGG

CCATGGCATGAGTGTTTTCTAGTAGTAGATTGGAGGGAAAGCTTTGTGACACTTAGTACTGT

GTTTTTAAGAGGAAATAATTTGGTTCCAGATGTGTTAGAGGATCTTTTGTACTGAAGGTTTT

ACACCTTTACTTTGGGTTTACCAGCCTCCACTGGAACAGACATAAACAGTCCACAGGCAACC

GTTCTGCCCAGCCCCAACCCCAGGGAAGTCTCTCCGCAAAACCTTTCTTGGGGTTGCCCTAA

ACTTGCCAGGGGCCTTTGCTTCAAACCTCCCTCCTGTGAAATGTGTAACAAATGAAAAAGCC

GTGGGCCTCTCTGCCTTTGCCCCCTTGAAAACAAAAAACCGGGCCTTTTTATTTTTTAACCT

TAAAAAGAACCCCGATATTAAAAAAAATGGGTTGGGTGTTGAACAAAACCCTGTCCTTCCC

-continued

(CloneID 5768283)
TGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGA SEQ ID NO:42

GCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATGCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCT

GTGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTG

GGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGC

AGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGAT

ATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGT

GCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCC

CTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGG

CCATGGGCATGAGTGTTTTCTAGAAGTAGATTGGAGGGGAAAGCTTTGTGCGCACTTAAAAC

CGGGGTTTTTAAAAAAAATTAACTTGGGTTCCCGAATGGGTTAGAAGAATCTTTTGGACCTG

AGGTTTTTTAAACCTTTTAATTGGGGCTTAACCAAGCCTCTACTGGACCGAAATCATTAACA

GGCCCCCCGGCTCCGCGTTCCTTGCCAGGGCCCCAACCCACAGGGGGAGTTCTCCTCGGCCC

AAGCCCTTTTTTGGGGGCTCGCCCCCAAAATGTGCAGGGGGGCCTTTGCGCTCAAAACCTCC

C

(CloneID 5396994)
GTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCCCCAGCCTA SEQ ID NO:43

CCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGT

GAGGAACAACAAAATGCCCTCTAATATCCTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTACATCTGAAGGGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAACCCCGTGC

CTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTCAACGCTGTCGTTCTGGTCTGTTTCCTCCATGCTTTGTGAACTGGCACA

ACTTGAGCCTGACTGGTACATCTCTTTGGATTTGTTTCCATTAAAAAGAAGCACTTTATGTA

CTGCTCTCTTTATTAGCCCTTTGGAAGAACCGGTTTCTCTCTGTCCTGTGACTCCTGGGGTC

TGTGGGCCATGGCATGAGGGTCTCTAGTAGTAGATTGAGGGAAAGCTTGTGGCCACTTAATA

CTGGGTTTAAAGACGAAATACTTGTGGTCCAGATGTGTTACAAGGAACCTTGGCACGGAGGT

TATAACACCTTTATGGGGTGTCCCAGGCTCCCTGGACGACCCTAAGCGTCCACGGCGCCGGT

CCGCCGGCCCAC

-continued

(CloneID 3941190)
TTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCGACCCTTTGAGCGTTCTGCTCCGGCGCCA SEQ ID NO:44

GCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTA

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCCGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAAGAAGCACTTTATGTACTGCT

GTCTTTTTTTTCCTTTCGAAGAACAGGTTGCTCTCTGCCTTGACTCTGTGGGTCTGTGGCGC

ATGGATGAGTGTTCCTAGCTAGTAGATGGAGGGCAACTTTGGGGACCCTTAGCGTGGTTTAC

CAGAAATACCTGGGTCCAGTGGCCCAGGCCTTGGCCGGGGTTCCACCTTCTGGGTTTCAGCT

CACGGGAACCATAAACCAAGGCGCCGAGGCCACAGGATCCCGCCTTGTGGCATCCGGGGTGA

CCCGTATGAAAAAGCGCCTCCAAATTTTTAAACACAAGCGCCCGGAGACCAT

(CloneID 5928395)
GCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAG SEQ ID NO:45

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGG

CCATGGCATGAGTGTTTTCTAGTAATAGATTGGAGGGAAACCTTTGTGACACTTAATACTGT

GTTTTTAAGAAGAAATAATTTGGTTCCAAATGTGTTAGGAGGATCTTTTTGTACTGAAGTTT

TTAAAACTTTACTTGGGTTTTACCAAGCCCTCACCTGGGACAGACCCTTAACCGGTCCAACG

GGCCACCGTTTCTTGCCAGGGCCCCAAACCCACAGGGGAAGTCTTCTCGCAAAAACCCTTCC

TGGGGGGTTCCCCCTAACTTGCCCAGGGGCCCCTTTGCTCCAAAAC

(CloneID 4750262)
CGATCGGCCGGACAGGCGGCAGCGGCGGCTCCTGCAGCGGTGGTCGGCTCTTGGGTGTGGAG SEQ ID NO:46

TTTCCCACCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGC

-continued

TCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGCAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTCGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGCCAAGTCCAGCCTCGTCTTGGCTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTGGTGTCATTACACAAGAAGCACTTTATGTACTGCT

GGTCTATAAATATCCTTTGCGACGAACAGGTTCCTCTCTGCCTTGACTCTGGGGCCTGTGGG

CCACTGGCAGTGAGTTGTGACTAGCTAGTAGATTGGGCGGAAACGCTTGGTGACCCTTAGCC

CCTGGTTTCTACGACGCCATACTTGGTCCCGATGTGTTACAGGCCTCTTGCGCCGGAGGCTT

TACACCCTTTACTGCGCTCCCAAGGCTCAGGGGCGACCTAA

(CloneID 5274910)
AGCGGAGCGGCGGCTCCTCCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCT SEQ ID NO:47

CGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCAAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTCACAAAGAAGCACTTTATGTAACTGC

TGTCTTTTTTACCTTTTGAAGAACCGGTTCCTCTCTGTCCTTGACTCTGGGGTCTGGGGGCC

AGGCATT

(CloneID 3140086)
TTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCCCC SEQ ID NO:48

GGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCAAATTTTTCATTAGGTTTTGATGAACCAACAGAACAAGGCG

GTGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTG

GGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGC

AGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGAT

ATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGT

-continued

GCCTTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCC

CCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGGTTTCCTCCATGCTTGTGAACTGGACAAC

TTGAGCCTGACTGTACATCTCTTGGATTGTGTTTCATTCAAAAGACGCACTTTATGTAACAC

AAA

(CloneID 5286160)
AGCGGAGGGGCGGCTCCTGCAGCGGTGGTCGGCTGTGGGTGTGGAGTTTCCGAGCGCCCCTC SEQ ID NO:49

GGGTCCGACCCTTTGAGCGTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCG

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTACAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTCCTTTTGAAGAAC

(CloneID 5215141)
GGTACCGGTCCCGAGTTCTTTGGATGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAG SEQ ID NO:50

CGCCCCTCGGGTCCGACCCTTTGGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGG

CGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAACGGTGAT

ATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGT

GCCTGCTGCGCCTGTGCCCACACCCGGTGGACCCGGACCCAGTTGCCATCCAGAAGAAATCC

CCCTGGCGGCAAGTCCAGCCTCGTCTAGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACTT

GAGCCTGACTGTTACATCTCTTGGATAAGTTTCATTAAAAAGAAGCACTTTATGTACTG

(Clone ID 4845679)
GAGCGGTGGTCGGCTGTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCG SEQ ID NO:51

TTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGA

-continued

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATT

CATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCC

TGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTG

GCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTCTATGTACTGCTG

TCTTTTTTTTTCTTTGGAAGAA

(CloneID 5297866)

AGCGGAGCGTCGGCTCCTCCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCT SEQ ID NO:52

CGGGTCCGACCCTTTGACCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAACCAAGCTCCGGAGACTTCTTAGATCTGAAGGGCGACGGTGATA

TTCATGAACAATGTGGACACCAGACTTCCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCG

TGCCTGCTGCGCCTTGTGCCCAGACCCGGTTGGCCCCGGCCCCCACTCCCATCCAGAAGAAA

TCCCCCTCGCGGCATGTCAGCCTCGTCTTGGGTAA

GCTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTACCTCCATGCTTGTTGAACTTGACA

ACTTCGAGCCTGACGTGTACATCTCTTGGATTGGTTTCATTAAAAAAGAAGCACTTTATGTA

CTGGCTGTCTTAAATAACTTTTGAAAGAAC

(CloneID 5194208)

CCAATCCCTGGATCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGG SEQ ID NO:53

GTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAATAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTCATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAACTCTAGTGGTGGCAGGGAAGACTTGCAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGACAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAACAGAAGCCCGTG

CCTGCTGCGCCTCTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTGTGTGAAACTGGACAA

CTTGAGCCTGACTGTACATCTCGTTGGATTGGTTTCATTAACAAAGAAGCACTTTATGTACT

TGCTGTCTTTTGATGCTTTGGAGAGAACAGGTTACTACTCTGGCCCTTGACTCTGGGGTCTG

TGGGCCATGGCATGAGTGTTTACTACGTAGCACAACTGGCCCGGAAACGCTCCGGTGAGACC

CATTAGCTACGGGGGTTTTAAGAACGCAAATTAATCGCGTACGCGAATGTGGTACAAGGACA

CTGTAGGACACGGACGCTGTAACACACTTAACAAGCGGTTTCACCAGGACCTCAACGGAGAC

CGACCATATAACGGCCACAGGAGCACGGACATGACAGGGCCAAGCC

-continued

(CloneID 5017113)
GCCTGCAGCGGTGGTCGGCTGTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTT SEQ ID NO:54

GAGCGTTCTCCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGCAGACTTCTTAGATCTGAAGGCAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCACCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGCCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGCCAAGTCCAGCCTCGTCTGTGGCTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGACAATTG

AGCCTGACTGTACATCTCTTGGATTGGGTTTCATTAAAAAGAAGCACTTTAAAAAAAAAAAA

GAAAAAAAAAAAAAT

(CloneID 5483414)
TTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCC SEQ ID NO:55

AGCCTACCTCGCTCCTCCGCCCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCCAGTTTT

GCGGCCTCCAGGTGGTCCATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGCCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTCAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAACAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTCCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGG

CCATGGCATGAGTGTTTTCTAGTAGTAGATTGGAGGGAAAGCTTTGTGACACTTAGTACTGT

GTTTTTAGAAGAAATAATTTGGTTCCAAATGTGTTAAAGGATCTTTTGAACTGAGGTTTTTA

AACACTTTACTTGGGTTTCCCAAGCCTCAACTCGACAGAACATAAACAGTCCACAGGCACCG

GTTCCTGCCAGGGCCCAAACCCACAGGAAATCCTTTCGCCAAAAACCCCTTCTTGGGGGTGG

CCCCTACTTGGCCAGTGGCCCTTTGCTCCAAAACCTCCTCCTGGGAACCGGGGGAAAACATT

GAAAAGGCCCTGGGGCCTCCCT

(CloneID 5778829)
GCACGAGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCG SEQ ID NO:56

TTCTGCTCCGGCGCCAGCCTACCTCCCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

CCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

-continued

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGACACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCACGCAGCCTGGGGCAGAGTGAACAGAAGCCCGTC

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAACAAATCCCCC

TGGCGGCAAGTCCAGCCTCG

(CloneID 5778829 in 3' to 5' orientation)
TTTTTTTTTTTTTTTTTTTTTACATAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGA SEQ ID NO:57

TGTACAGTCAGGCTCAAGTTGTGCACTTCACAAGCATGGAGGAAACAGACAGAACGACAGCG

TTCAGGACACTCAGAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCGGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTGGATCCACCACCTGGAGGCCGCAAAACTCGGGAGCTATTCCTGCTGTTGGGGT

CGACTCCCTTGAAGGT

(CloneID 2781859)
GAGCGGCGGCTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGC SEQ ID NO:58

TCCCACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCCACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATCAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCGCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCCTCTTG

(CloneID 2781859 in 3' to 5' orientation)
TTTTTTTACATAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGCCTCA SEQ ID NO:59

AGTTGTGCACTTCACAAGCATGGAGCAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCGGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTGGATCCACCACCTGGAGGCCCCANAACTCGGGAGCTATTCCTGCTGTTGGGGT

CGACTCCCTTGAAGGTCGTGGTTGTGGTCAT

-continued

(CloneID 3940455)
GGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGC SEQ ID NO:60

CTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTC

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGCCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGACACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGACTGAAGAGAAGCCCGTGC

CTGCTTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGGCAAGTCCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTNTCCTCCATGCTTGGTCGAACTCACAAT

TTGAGCCTGGACTGTACATCTCTGGGAATTGGTTTCATTAAAAAGAAGCACTTTAAAAAAAA

AAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAACCGGAGAGATAAAAAGGGGGGGGAGAAAA

GTGGAAAAAAAGACCAGCCGCAAACATTTGCCGGACCCTTTTTGGCGACACAGGGGAAAACG

GGAGCCCCAAAGGAGGTAAATGAAACGGTTCTTTTTTTCTTATTTCAGAAAGCAGAAGAAAA

AGAAATAAAGAAAAGAAAGAAAAAAAAAAGAATGATAAAAGAGGAAAAATAAAAAAAAAGAAA

AAGAGAGAGAATAGAAACTTAAAAAATAAACAGAAAAAAGAA

(CloneID 5459503)
GCTGCAGCGGTGGTCCGCTGTTGGCTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTT SEQ ID NO:61

GAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGCACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAACAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGT

GCCTCCTGCGCCTGTCCCCAGCCCGGTGGCCCGGCCCCAGTCCCATCCAGAACAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTCCCCTTCCCAAACGAAGCACTTTATGTACTGC

TGTCTTTTTTTTTTTCCCTTTGGAACGCCCCCCGGTTCCTCTCTGTCCCTTCACTCCTGGGGT

CTGGGGGCCTG

(CloneID 4539292)
TGCAGCGGTGGTCGGCTGTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAG SEQ ID NO:62

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGCGAAGACTTGGAGTCATCTGGACTGCA

-continued

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGCTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGCCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTGTTTCATTAAAAAGAAGCACTTTATGTAAAAAAAA

AAAAAA

(CloneID 5247820)
CTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCAGCGCCCCTCGGGTCGACCCTTT SEQ ID NO:63

CAGCGTTCTGCTCCGGCGCCACCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGACTTTT

CCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTCGAGTCATCTGGACTGCA

GAGAAGG

AACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATTCATGA

AAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAACGAGAAGCCCGTGCCTGCT

GCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTGCCGG

CAAGTCCAGCCTCGTCTGGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTCGATTTGTTTCATTACAAAGAAGCACTTTATGTACTGCTG

TCTTTATATTTTGTCCTTTGTGAGAACAGGTCCTCTCGTGTCCTTGACTCTGGGGTCTGTGG

C

(CloneID 5397495)
CTCCTGCAGCGCTGGTCGGCTGTTCCGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCT SEQ ID NO:64

TTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTACGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGCACACCTGAACGAAAATCAAGCTTCTTG

GGCCAACTCACCAGGTGCCAACCTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTG

CAGACAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGCACCAAGGT

GATACTTCATGAAAACTGTGGACACAGACTTGCCAGGCACCCTGGGGCAGACTGAACAGAAG

CCCGTGCCTGCTCCCCCTGTGCCCAGCCCGGTGGGCCCTCGCCCAGTGCCATCCACGAACGA

AATCCCCCTGGCTCGCAATGTCCACCCTCGTCTTGGGTTACG

CTCTGACTGTCCTGAACGCTGTCCTACTGTCTGTCTCCTCCATGCTCGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGCATTCGTATCATTAACCAAGAACCACTTTATGTACTCCT

GTCTTTATTTACGCCTTTCGAAGACCG

(CloneID 5200412)
CTCCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTG SEQ ID NO:65

AGCGTTCTGCTCCGGCCCCAGCCTACCTCGCTCCTCGGCCCC

-continued

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACACCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTCCCAAGTCTAGTCGTGGCAGGGAACACTTGGAGTCATCTGCACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTCAAGGGACAACGTGATA

TTCATGAAAATGTGGACACAGACTTCCCAGGCAGCCTGGGGCACAGTGAAGACAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTGTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGGTACATCTCTTGGATCCGTTTCATTAAAAAGAAGCACTTTATGTACTGCT

GTCTTTATTTTCACTTTAGACCAACCAGGTGTCTCTCGTGTCCTGGACTCCTGGGACCTGTG

GGCCATG

(CloneID 5208934)
CTGCAGCGCTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTC SEQ ID NO:66

ACCGTTCTCCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATCAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTC

TCTTTTTTTTCCTTTTGAAGACCAGGTTCTCTCTGTCCTTGACTCTTGGGCTCTGTGGGCCC

C

(CloneID 4156130)
CCCACCCCTGGCCCCCTCTTGGGTGTGGAGTTTCCCACCCCCCCTCGGGTCCGACCCTTTGA SEQ ID NO:67

GCCTTCTCCTCCCGCGCCAGCCTACCTCCCTCCTCCCCCCC

ATGACCACAACCACCACCTTCAACCCAGTCCACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCCCCCTCCAGGTGGTGGATCCAATTTTTCATTACGTTTGATGAACCAACACAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTACTCGTGGCAGGGAAGACTTCGACTCATCTGGACTGCAG

ACAAGGAACTCCTCTCAACCAACCTCCCGAGACTTCTTAGATCTGAACGGAGAAGGTGATAT

TTATGAAAATGTGCACACAGACTTCCCAGGCAGCCTGGGGCACAGTCAAGAGAAGCCCGTGC

CTGCTGCGCCTCTCCCCACCCCCGTCGCCCCGGCCCCACTGCCATCCAGAAGAAATCCCCCT

GGCGGCAACTCCACCCTCCTCTTGGGTTAG

CTCTGACTCTCCTCAACCCTCTCGTTCTGTCTGTTTCCTCCATCCTTGTGAACTGCACAACT

-continued

TGAGCCTCACTGTACATCTCTCTGGATTGTTTCATTAAAAAGAAGCCCTTTATGTACTCCTG

TCTTTATCTTTCCTTTCCAGGACCCGGTTTCCCCCCCCTCCTGGCCTCTGCCTCCTCGGCCC

CTGCCTTCGCTCTTC

(CloneID 4123030)
GCTCCTGCAGCGGTGGTCGGCTGTGGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCC SEQ ID NO:68

TTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGCGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTC

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGCC

AAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGAG

AAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTACATCTGAAGGGAGAAGGTGATATTC

ATGAAAATGTGGACACAGACTTGCCAGGCAGCTGGGGCAGAGTGAAGAGAAGCCCGTCCCTG

CTGCGCCTGTGCCCAGCCCGGTGGCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTGGGCG

GCAAGTCCAGCCTCGTCGTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGACAACTT

CAGCCTGACTGTACATCTCTGGGATTGTTCATTAAAAGAAGCACTTTATGTACTGCTGTCTT

TTTTTTCTTTTGAGGACGGTTCTCCCTGTCCTGACTCTGGTTGGGGGGCAGGGTGAGGGTTC

TAAAAAGGGGG

(CloneID 5459215)
GGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCGACCCTTTGAG SEQ ID NO:69

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGACAGGAATAGCTCCCGAGTTT

TGCGGCCTCCAGGTGCTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCT

GTGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTG

GGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGC

AGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTACATCTGAAGGGAGAAGGTGAT

ATTCATGAAAATGTGGACACACACTTGCCAGGCAGCCTGGCGCAGAGTGAAGAGAAGCCCGT

GCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCGGCCCAGTGCCATCCACAAGAAATCCCCCT

CGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTTCCTCCATGCTTGTGAACTGCACAAT

TGAGACCTCACTGCTACATCTCTGCATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCT

GTCTTTATTTTTCCTTGTGAACA

(CloneID 4337230)
GCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTT SEQ ID NO:70

TGAGCGTTCTGCTCCGCCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGC

GCCAAGTCAGCAGGTCCCAACTCTAGTGGTGCCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGCGAGAAGGTGATA

-continued

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCCTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGACAATTG

AGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTAAAAAAAAAAAAA

AGAGAAAAGAAAAACCGGGGGTTTTTTGGGGGGGGAAAAGAGGGTTAAAGGCGGGGTAGGTT

TTTGCCCTTTTTGAAAGGGGAGACGGGGGCCCAGGGGGGGTAATGGAAGGGGACG

(CloneID 4649355)
GCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAG SEQ ID NO:71

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGCTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCACAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGACCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TCAGCCTGACTGTACATCTCTTGCATCTGTTTCATTAAACAGAAGCACTTTATGTACTGCTG

TCTCACTACAAATACACTATCGAACAACAGGTTACTACTCTGTCCTTGAN

(CloneID 4126162)
CCTGCAGCGGTGGTCGGCTGTTGCGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTT SEQ ID NO:72

GAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGCGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTCAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGCGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGC

CTGCTGCGCCTGTCCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTTCTCTCTGTTTCCTCCATGCTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTGGGATTGTTCATTCAAAAGAAGCACTTTATGTACTGCTGTC

TTTTTTCCCTTGGACACCAGGTTCCCCTCCCTCCTTGACTCCTGGCTTTGGGGCCCGGCCAG

AGTGTTCCTATAACAATGGGCGGAAAACCTTGGGCACCTAGCCCGGTTAAAGAAAAATGGGC

CACTGGTAAGC

(CloneID 2783676 in 3' to 5' orientation)
TTTTTTTACATAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGCTCA SEQ ID NO:73

AGTTGTGCAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAG

-continued

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCGGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCTTCAGATCTAAGAAGTCTCCGGAGCTTGCTTCAGAG

GAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGCCACCACTAGACTTGGCACC

TGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCAAAGATATTAGAGGCCATTT

TGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAATGAAAAATTGGATCCACCA

CCTGGAGGCCGCAAAACTCGGGAGCTATTCCTGCTGTTGGGGTCGACTCCCTTGAAGGTGGT

GGT

(CloneID 2783676)
GCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAG SEQ ID NO:74

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGAAAATGTGGACAC

AGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGCCCA

GCCCGGTGGCCCCGGCCCCAGTGCC

(CloneID 2822776 in 3' to 5' orientation)
TTTTTTTTTTACATAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGC SEQ ID NO:75

TCAAGTTGTGCAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCA

GAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCGGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTGGATCCACCACCTGGAGGCCGCAAAACTCGGGIAAGCATTCCTGCTGTTGG

(CloneID 2822776)
GGCACGAGGGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGCTCCGACCCTTTGAGCGT SEQ ID NO:76

TCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGACTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAACAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAAAAATCCCCC

TGGCGGCAAGTCCAGC

-continued

(CloneID 2781579)
GCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAG SEQ ID NO:77

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAATGAAAAAAACCCGTGC

CTGCTGCCCTGTGCCCAGCCCGTGGCCCCCGGCCCATGCCATCAAAAAAATCCCCCTGCCGC

GAACCAACCTTTGCTG

(CloneID 2781579 in 3' to 5' orientation)
TTTTTTTACATAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGCTCA SEQ ID NO:78

AGTTGTGCAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCGGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTGGATCCACCACCTGGAGGCCGCAAAACTCGGGAGCTATTCCTGCTGTTGGGGT

CGACTCCCTTGAAAGTGGTGGT

(CloneID 5419867)
GCAGCGGTGGTCGGCTGTTGGGTGTGCAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAG SEQ ID NO:79

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTCGATCCAATTTTTCATTAGGTTTTGATGAACCAACACAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAACTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGTTTTCATAGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAG

AAGCCCGTGCCTTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGACCCAGTGCCATCCAGAA

GAAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

AGAGCCTGACTGTACATCTCTATGGATATGTTTCATTAAAACAGAAGCACTTTATGTACTGG

TGTCTTTATTATTTCTTTGGAAGACCAGGTTCTCCTCTGGCCTAGAATCCTGGGGGTCTGAG

G

(CloneID 4328715)
CGCTGATTGGCTGTGGAGTTTCCCAGCGCCCCTCGCGTCCGACCCTTTGAGCGTTCTGCTCC SEQ ID NO:80

GGCGCCAGCCTACCTCCTCCTCGGGCGCC

ATGACCACAACCACCACCTTCAAGGGATTCGATCCCAACAGCACGAATAGCTCCCGAGTTGT

-continued

GCGGACTCCAGCGTGGTGGATCCAATTTTATCATTAGGTTTTGATGAACCAACAGAGCAACC

TGTGAGGAAGAACAAAATGGCCTCTAAATATCTTTGGGACACCTGAAGAAGAATCAAGCGTT

CTTGGGCCAGAGTCAGTCAGCGTGCCAAGTCTAGATGGTGGCAGGGAACGACTTGGAGTCAT

CTGGACTGCATGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGG

AGACCGGTGATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGTGGCAGAGTGAT

GAGAAGCCCGTGCCTGCTTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCA

GAAGAAATCCCCCTGGCGGCAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGGATTTGGTTCCCTTTAAAAAGAAGCACTTTATGTTACT

GCTGTCTTTTTTTTTTCCTTTTGACGGCACGGGTTTCCTCTCGGTCCTTGACTCTGGGGGTC

TGGGGGCCCAGGGCATGAGTGTTCCTTATATGATGGGGGGAGAGCTTGGGGCCCTTGCACCG

GTTTTTCGCGAGCAATATTGGGCCCAGTGGTTAGGCTCTTTGGCCGGGGTTCACCTCTTGGG

TTCCCCCCCGGGGGCCAAGGACGCCGTTTGGTC

(CloneID 4138051)
GCTGCAGCGGTGGTCGGCTGTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTG SEQ ID NO:81

ACCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTCAAGGGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTCCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGC

CTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTG

GCGGCAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTATCTGTCTCTTGTCCTCCATGCTTGTGAACTGCACAA

GTTGAGCCTGACTGTACATCTCTTGGATTCGTTCATTAACAAGAAGCACTTTATGTACTCCT

GTCTTTTATTTTCCCTTTGAAGACCCGGTCCCCCCAGGCCTGACCTGGGCTGGCGCCAGGAA

GAGGTTCTCATAAAAAAGGGGGGAACACTTGGGACACTAGACTCGGGTTTCGCAAAAATCTG

GGTCCAATGTGAAAAGATGGGAGG

(CloneID 4766330)
GCTGCAGCCGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTT SEQ ID NO:82

GAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TCAGCAAGAACAAAATGCCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAACTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGAAGACTTGGAGTCATCTCGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TCATGAACAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCACTGCCATCCAGAAGAAAATCCCC

-continued

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATCCTTGTGAACTGCACAAT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAACAAGAAGCACCTTTAAACACAAA

AAAAAAAA

(CloneID 3958097)
GTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCAGACCCTTTGAGCGTTCTGCTCCGGCGC SEQ ID NO:83

CAGTCCTATCCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGT

GAGGCAAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTG

GGCCAAGTCACGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTCGACTG

CAGACAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGCAGAAGGTG

ATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGCAGTGAACGAGAAGC

CCGTCCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAAT

CCCCCTGGCGGGAAGTCCAGCCTCGTCTTGCGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAGCTGCACAACT

TGAGTCCTGACTGTACATCTCTTGGATTCGTCTCATTGAGAAGAGAAGCACTTTATGTACTG

TGTCTTTTCTGTCCTTTGGCGGCCCGGTTCTCTCTGTCTTGACCCTGGGGCGGTGGGN

(CloneID 346318 in 3' to 5' orientation)
TTTTTTTTTTTTTTTTTTTAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAG SEQ ID NO:84

TCAGGCTCAAGTTCTGCAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGA

CAGTCAGAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGCGGATTTCTTCTGGATGGCACTGGGGCCGG

GGCACCGGGCTCGGGCACAGGCGCANAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTGCC

TGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTCTC

CGGAGCTTGCTTCAGAGGAGTTCCTTCTCTCCAGTCCAGATGACTCCAAGTCTTCCCTGCCA

CCACTNAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCAA

AGATATTAGAGGCCATT

(CloneID 346318)
CGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGC SEQ ID NO:85

CACTANCCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGACTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGCGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGGACTNG

CAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAAGAAAGGT

GATATTCATG

(CloneID 487018 in 3' to 5' orientation)
AAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGCTCAAGTTGTGCAGT SEQ ID NO:86

TCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAC

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGGCACTGGGGGC

-continued

CGGGGCACCGGGCTGNGCCACAGGCGCANAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCT

GCCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGT

CTCCGGAGCTTGCTTCAGAGGACTTCCTTCTCTCCAGTCCAGATGACTCCAAGTCTTCCCTG

CCACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCNGGTGTCCC

AA

(CloneID 487018)
CTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGCGTCCGACCCT SEQ ID NO:87

TTGAGCGTTCTGCTCCGGCGCCACTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTTGGAGTCATCTTGGACTT

GCAGAGAAGGAACTCCTCTTGAAGCAAAGCTTCCGGAGACTTCTTTAGATCTTGAAGGGAGA

AGGTGATATTTG

(CloneID 3509038)
AGCTAGGTACGAGGCCTGGGTGTTGCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCC SEQ ID NO:88

AGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCG

GCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGACTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTCCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGCTCATCTGCACTGC

AGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAACGGGCAGAACGGT

GATATTCATGAAACATCTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGC

CCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAAT

CCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTCACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTCATCTCTTGCATTGTTCAC

(CloneID 5094728)
GCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCCGGTCCGACCCTTTGAG SEQ ID NO:89

CGTTCTGCTCCGGCCCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAACGGAGTCGACCCCAACAGCACGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTCGATGAACCAACAGAACAACCTG

TGACGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

CCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

CAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGTTTTCATAGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAG

AAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGACCCAGTGCCATCCAGAAG

AAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

-continued

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATATGTTTCATTAAACAAGAAGCACTTTATGTAC

(CloneID 5590830)
GTCCGGAATTCTCCGGATCTCCTGCAGCGGTGGTCGGCTTTTCGGTGTGGAGTTTCCCAGCG SEQ ID NO:90

CCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGC

C

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAACAACAAAATGCCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTCG

GCCAACTCAGCAGGTCCCAAGTCTACTCGTGGCAGGCAAGACTTGGACTCATCTGGACTCCA

GACAACGAACTCCTCTCAAGCAACCTCCGGAGACTTCTTAGATCTCAAGGGAGAAGGTCATA

TTCATGAAAATCTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAACCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCCGTGGCCCCGCCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAACTCCAGCCTCGTCTTGGGTTAG

CTCTGACTCTCCTGAACCCTGTCGTTCTCTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTCACTGTACATCTCTTGGAATTGTTTCATTAAAAAGAAGCACTTTATGTACTCCCT

GTCTTTTTTTTTTTTCTTTTGAAAGAACAGGGTTTCTCTCTGCCCTTGACTCTTGGGCTCCG

TGCGCCATGGGCATGACGGGTTTTCCCAGTAGTTACATTGGGAGGGCAAACGTCTCTGTGAC

CACTTAAATCACCGCCGCTTTTAAACAACAAATTAATTTCGGCTCCCTCAAGGGCCGTTATA

ACGTACCCTTTTCTCTCTCTGAGCGCCCTGAAAACTCCTTCACCCCCCGCATGAACGCAGGT

TCCCTACTTGGTTCACCACCTCAACCGCTCTGTACCCCCCAGGGGACCCCTGTCTCCCCCCG

AGGTGCCCCTCCCTCCCCCCCGGCCCATAATGTTTTCTTCAGCCCAACGAAACTGCCA

(CloneID 4550156)
GCGGCTCTTCGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCGACCCTTTGACCCTTCTGCTCC SEQ ID NO:91

GGCGCCACCCTACCTCCCTCCTCGGCGCC

ATCACCACAACCACCACCTTCAAGGCACTCGACCCCAACAGCAGGAATACCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTCCCACACCTGAAGAAAATCAAGCTTCTTGC

GCCAAGTCACCAGGTGCCAACTCTAGTGGTGCCACCGAAGACTTGCAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCACGCAGCCTGCCCCAGAGTGAAGAGAAGCCCGTC

CCTGCTGCCCCTGTGCCCAGCCCGGTGGCCCCCGCCCCACTCCCATCCAGAACAAATCCCCC

TGGCGGCAACTCCAGCCTCCTCTTGGGTTAG

CTCTCACTGTCCTGAACCCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTTCAACTGGCACAA

CTTGGAGCCCTGACTGTACATCTCTTGGATTGTTTCATTAAAAAGAAGCCACTTTAAAAAAA

AAAAAACCCCTGGAGCACCTACGCAGAGACAGGAGACAAAAAAAATCACAACGCGGGGCCGC

CCTTTTGCGGGGGGGCGAAAAACGAGTAAAAAAGGCCCCGGGTCTACAGTGTTGTCCGGACA

CCCACTTTGGTCACAAGGGGGAACGCCCCGCCCACAACGGGGGGAAATAAGCAACCGGGGGA

TCCTA

-continued

(CloneID 5258724)
GCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCCACCCTTT SEQ ID NO:92

GAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTCGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTCGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATT

CATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCC

TGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTG

GCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGGTCTGTTTCCTCCATGCTTGTGAACTGCACAAC

TTGAGCCTGACTGTACATCTCTTGGATCTGTTCATTAAAAAGAAGCACTTTAAGAAAACCAC

ACAAAAAAACTCCGAGGCCACTA

(CloneID 3458066)
TGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCA SEQ ID NO:93

GCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTGGGACACCTGAAGAAAATCAAGCTTCTGGGCCA

AGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGAGA

AGGAACTCCTCTGAAGCAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGAATATTCA

TGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCTG

CTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTGGCG

GCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTCTCGTTCTGTCTGTTTCCTCCATGCTTTGTGAACTGGCACAA

CTTTGAGCCTGAGTGTACATCTCCTGTGGTTGTTTCATTAAACAGAAGCACTTTATGTACTG

TGTCTTTTTTTTGTCCTTCGGAGACCGGTTCTCTCCGGTCCTGACTCTGCGTCTGGGGGCCG

GCTGGGTGTCTAGCGGCACGGCGGGAGCTTGGCCCCTGCCGGCT

(CloneID 3940211)
TGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTAC SEQ ID NO:94

CTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCCACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTGGGAACACCTGAAGAAAATCAAGCTTCTTGGCC

AAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGCAQTCATCTGGACTGCAGAG

AAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATTC

ATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCT

GCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTGCC

GGCAAGTCCAGCCTCGTCTTGGGTTAG

-continued

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGACAATTG

AGCCTGACTGTACATCTCTTCGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTGTC

TTTTTTTTCTTTCGAAGAACAGGTGCCTCTCTCTCCTGACCCTGGGTCTGGGGCATGCATGG

GTGGTTCTAATAGTAATGGGGGAACTGGGACCTTAGACGGGTGTCAGAGACACACGGGTCCA

GGTCACGCCTCGGCGGGGTTACACTCAGGGTOCACCGCATGGGCCACAAACCCCGGGCGCTG

CGGCCCACAGGGGTCCCGCACCGGGGGCCAACGGGGCGGCACCCGGGAAAAAGAGCGCCTCG

CGCAAAAGTGTATACAAGGGGGGAGCGGCGAAGAACTAGACTCCGCGC

(CloneID 5015022)

GGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGC SEQ ID NO:95

TCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCACGAATAGCTCCCGACTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGACGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGOCAGGGAAGACTTGGAGTCATCTGCACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCCGAGACTTCTTAGATCTGAAGGGACGACCGTGAT

ATTCATGAACAATGTGGACACACACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCG

TGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGACCCAGTGCCATCCAGAAGAAATCCC

CCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTCTGAACTGCACAACT

TGCACGCCTGACTGTACATCTCTTGGGATTTGTTTCATTACAAAGAAGCACTTTATGTACAA

ACACCATAAAAAAC

(CloneID 489924)

GGCTGTTGGGTGTCGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCG SEQ ID NO:96

GCGCCACTAGNCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAAGCAGGTGCCAAGTCCAGTNGTTGCAAGGNAGGACTTTGGAGTCATTCTGGGA

CTGCAGAGANGGAACTCCTCTGAAGGCAAGCTCCCGGACACTTTCTTAGGATCTTGAAAGGG

AGAAGGTGATATNCCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGAAGAGTGAA

GAGAAGCCCGTGCCTGCTGCGCCTTGTGCCCAGCCCGGTGGCCCCGGGCCCAGTGCCATCCA

GNAAAATCCCCCTTGGCGGAANGTCCAGCTNG

TTTNGGGTTAGCNCTGATTTCCTAACGNGGT

(CloneID 489924 in 3' to 5' orientation)

ACATAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGCTCAAGTTGTG SEQ ID NO:97

CAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACACTCAGAG

CTAACCCAAGACGACGCTGGACTTGCCGCCAGGGGATTTCTTCTGGATGGCACTGGGGCCGG

GGCACCGGGCTGGGCACAGGCGCANAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTGCCT

GGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAACTCTCC

GGAGCTTGCTTCAGAGCAGTTCCTTCTCTCCAGTCCAGATGACTCCAAGTCTTCCCTGCCAC

-continued

CACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCAAAG

ATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCCTAATG

AAAATTGGATCCANCACCTNGAACCCGCAAACTTCGGGAGCTANTCCTGCTGNTGGGG

(CloneID 3940896)
GCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCGACCCTTTGAGCGTTCTGCTCCGGC SEQ ID NO:98

GCCAGCCTACCTCGCTCCTCGCCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTCGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAACAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGCAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGCAGAAGGTGATA

TTCATGAAAATCTGGACACAGACCTTGCCAGGCAGCCTCGGGCAGACTGAAGAGAAGCCCGT

CCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCC

CTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGACCCTCACTGTACATCTCTTGGATNTGTTTCATTAAAAAGAAGCACTTTATGT

(CloneID 4857040)
GGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACC SEQ ID NO:99

TCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGCCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTCATGAACCAACAGAACAACCTG

TGACGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

CCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAACACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGCTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGCCAGCCTGGGGCACAGTGAAGAGAAGCCCGTG

CCTGCTCCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TCAGCCTGACTGTACATCTCTTGGATTCGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTAATCCCTTTGGAACGACCAGGTTTCTCCTCTGTCCTTGACTCTAGGGTCCTCTG

GGCCATCG

(CloneID 3844688)
TGTGGGTGTGGAGTTTCCCAGCGCCCCTCGCGTCCGACCCTTTGAGCGTTCTGCTCCGGCGC SEQ ID NO:100

CACCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCAACAGCAGGAATAGCTCCCGAGTTTGC

GGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAAGCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATT

CATCAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCC

-continued

TGCTGCGCCTGTGCCCAGCCCGGTGGCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTGGC

GGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTCTACATCTCTTGGATTGTCTCATTGAAACCAGAAGCACTTTATGTACTGCT

GTCTTTCTTTACCCTTTGGACGGAACAGGTTCCTCTGTCCTGA

(CloneID 4797137)
AGCGGAGCGGCGGCTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCT SEQ ID NO:101

CGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTACATCTGAAGGGAGAACGGTCAT

ATTCATGAAAATGTGGACACAGACTTGCCAGCCAGCCTCCGCCAGAGTCAACAGAAGCCCCT

GCCTGCTGCGCCTGTGCCCAGCCCGGTCGCCCCGGCCCCACTGCCATCCACAAGAAATCCCC

CTGGCGCCAACTCCAGCCTCGTCTTGGCTTAC

CTCTCACTCTCCTGAACCCTCTCGTTCTCTCTGTTTCCTCCATGCTTCTGAACTGCACAACT

TGACCCTGACTGTACATCTCTTGGCATTTCTCTCCATTACCAAGAACCACTTTTATGTACCT

CCTCTCTTTCATCCATTATCACTTTCTGAAGAAACACCTTTTCTCCTCTCTCCTTCACTCCT

CCCGCTCTCTGGGCCCTCCCATGACCTGTCCCTAGCACCTCCCTTGGCCCGCGAACGCTTTG

TCACACCTAAGTACTGACCATCTCACGAACCA

(CloneID 4304080)
GGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGCCG SEQ ID NO:102

CCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAACACTTGGAGTCATCTGGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATT

CATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGACAAGCCCGTGCT

GCTGCGCCTGTCCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTGGC

GGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTACTGAACCGCTGTCGTTCCTGTCTGTTTCCTCCATGCTTGGTGAACTGACAA

CTTGAGCCTGATGTACAATCTCTGGGGATTTGTTTCATTAAAAAGAAGCACTTTATGTAAAA

AAAAGTTCACAAAAAAAAATCGGGGCACAAAGAAGGGGGGGAGAAAAGGGGAGGAAGGGGAC

ATGGACATAGGTGTTCTGGACAAGTTTTGGCCAGGCGAAATGGGGCCCCAAAGGCGGGATAT

GAAAGAAGTAAGTA

(CloneID 5928314)
GCGGTGGTCCGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGT SEQ ID NO:103

TCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

-continued

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTC

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGCAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTCCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCACCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAACAAGCACTTTATGTACTGCTC

TCTTTTTTTTTTTCTTTTGAAAACAAGTTTCTCTCTGTCCTTCACTCTTGGGTCTGTGGGG

CCATGGCATGAGTGGTTTTCTAGTACTAGATTGGAGGGAAAAGCTTTGTGGACACTTTAAGA

ACTGGGGTTTTTTAGAAGAAAATAATTTGGGTTTCCGAAGGTGGGTAAAAAGGATCTTTTTG

GGAACTGGAGGGGTTTTTAACAACTTTAATTTGGGGGTTTTACCCAAGCCCTCAAACTTCGG

ACCGAACCCTTTAAACACGTC

(CloneID 6159735)
GGACGCCTGGGCGGCTGTTGGCTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAG SEQ ID NO:104

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGACGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

CAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATCTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACAACAC

CCCAACCCNCCCACCACCCCCCACACCATCAAACAAAGGGGGGGCCGGCTCTAAGAATATCC

CCTCGACGGGGCCCAAAGCCTTACCGCGTAACCCCATCTTTTTCTTGGAACAAAAGGGGTTC

CCCTTATAGGGGGACGCCGCAAATTATAAAACCTTAGGGCACCCTGCCCCCGCTCCCCTTTT

TTTAACAAAACCTCTCGTCGGAATCCCCGC

(CloneID 308502)
GCTCTTGCCTCTGGAGTTTCCCACCCCCCCTCGGGTCCCACCCTTTGACCGTTCTGCTCCGG SEQ ID NO:105

CCCCACTANCCTCCCTCCTCCGCCCC

ATCACCACAACCACCACCTTCAACCGAGTCGACCCCAACAGCAGGAATACCTCCCGAGTTTT

CCGCCCTCCAGGTGGTCCATNCAATTTTTCAATTACGTTTTGATGAACCAACACGACCANCT

CTTAACAACGACCAAATTCCCCTCCTAATAATCTTTCCGACACCTGAACAAAATCAAGCTTC

TTCCCGCCAAGTCAGCACCTCCCAAGTCTAGTCCTCCCAGGGGAAAGACTTCGACTCATCTG

-continued

CACTTGCAGAGAAGGAACTCCTCTTGAAGCAACCTTCCGGAGACTTTCTTTACATCTTGAAA

GCCCACAAAGCTTGATATTCCATGGAAAATGTTGGACACAGACTTTTN

(CloneID 308502 in 3' to 5' orientation)
AAAGTCCTTCTTTTTAATGAAACAAATCCAAGAGATCTACACTCAGGCTCAAGTTGTGCAGT SEQ ID NO:106

TCACAACCATGGAGGAAACAGACACAACGACAGCCTTCACGACAGTCAGAC

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGATTTCTTCTGGATGGCACTGGGGCCGG

GGCACCGGGCTGGGCACAGGCCCACAGGCACCGGCTTCTCTTCACTCTGCCCCAGGCTGCCT

GGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTCTCC

GGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGCCAC

CACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

(CloneID 4302544)
GGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCG SEQ ID NO:107

CCACCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCCACCCCAACACCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAACAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTCCCAGGGAAGACTTGCAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGC

CTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTG

GCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCCTGACTGTACATCTCTTGGATTGTCTCATTAAAAAGAAGCACTTTATGTTAAACAA

AAACAAAAAAAAA

(CloneID 3937350)
TTGGCTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCC SEQ ID NO:108

AGCCTACCTCGCTCCTCCGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACACCAGCAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCACGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGCCAGAGTGAAGAGAACCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGCATTTGTTTCATTAAAAAGAAGCACTTTATGT

(CloneID 3835958)
TTCCCAGCGCCCCTCGGGTCAGGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGC SEQ ID NO:109

TCCTCGGCCCC

-continued

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATACCTCCCGAGTTTG

CGGCCTCCACGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTACATCTGAAGGGAGAAGGTGATATT

CATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCC

TGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTCG

CGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGACAACTT

GAGCCTGACTGTACATCTCTTCGATTTGTTTCATTAAAAAGAAGCACTTTAAAAAAAAAAAA

AAAAAAAAAAAAAACCGGGGGACACAAGGGGGGGGGGGAAAAAG

(CloneID 5925943)
TCGGCTCTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTC SEQ ID NO:110

CGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGCTGCCAAGTCTAGTGGTGGCAGGCAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCACCCTGGGGCAGAGTCAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTANAAAGAAGCACTTTATGTAAAAAAA

AAAAAAAAAAA

(CloneID 3936327)
GTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCCACCCTTTGAGCGTTCTGCTCCGGCGC SEQ ID NO:111

CAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGCAGTCATCTCCACTGCA

GAGAAGCAACTCCTCTGAACCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGCGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTCCCATCCAGAACAAATCCCCCT

GGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCCTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTCTTCATTAAAAAGAAGCACTT

-continued

(CloneID 4100301)
GGCGGCTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCG SEQ ID NO:112

ACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGCAAGAACAAAATCGCCTCTAATATCTTTGGGACACCTCAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCCGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCCGCTGGGGCAGAGTGAAGAGAAGCCCGTGC

TGCTGCGCCCTGTGCCCAGCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTG

GCGGCAAGTCCAGCCTCGTCTGGGTTAG

CCTCTGACTGTCCTGAACGCTGTCGTTCTCTCTGTTTCCTCCATGCTTGTGAACTGCACAAT

TTGAGCTTGACTGTACATCTTCTGGATTTGTGTCATTACCACAGACGCCTTTACAAACAGAA

AAAAAAATCCGGGGCTTCATTTTCGGGGGGGAACAGGGTATGACG

(CloneID 3903688)
GTTGGGTGTGGGAGTGTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCG SEQ ID NO:113

CCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTCGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAG

ACAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGC

CTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGGCAACTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGACAATTG

AGCCTGACTGTACATCTCTGGATTGTTTCATTAAAAAAAGAAGCAC

(CloneID 4797136)
AGCGGAGTCGCCTGTCTTGCTCCGGTAGTTGGCTGTGGACTGACCGCTTTCCCAGCGCCCCT SEQ ID NO:114

CGGGTCCGACCCTTTGATGCGTTGCTGCTCCTGGCAGCCAGCCATACCTCGCTCCTCGGCGG

CC

ATGACCACAAGCGACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATATGCTCCCGAGTT

TTGCGGCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCT

GTGAGGAAGAACAAAATGGTCTCTAATATCTTTGGGACACCTGAAGAAACATCAAGCTTCTT

GGGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTG

CAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGACGCTGA

TATTCATGAAAACTGTGGACACAGACTTGCCAGGCAGCCTGGGGCACGAGTGAAGAGAAGCC

CGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGGCCCAGTGCCATCCAGAAGAAATC

CCCCTGGCGGCAAGTCCAGCCTCCTCTATGGGTTAAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

-continued

TGAGCCTGACTGTACAATCTCTTGGATTGCTCTCATTACAAACAGAAGCACTTTATGTACTG

CTGTACTTCTCATTTTCGTTTTTCCACAACAGGGTTGCCTCTCTGCACTTTGACTCTTTGGG

CTCTGTGCGCCCATGGGCATGACTTT

(CloneID 4120864)
GCTCCTGCAGCGGTGGTCGGCTGTGGGTGTCGAGTTTCCCAGCGCCCCTCGGGTCCGACCCT SEQ ID NO:115

TTCAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGCAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTCGGACACCTGAAGAATCAAAAGCTTCTTGGGCC

AAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGAG

AAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAACGTGATATTC

ATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCACAGTGAAGAGAAGCCCGTGCCT

GCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTGGC

GGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTCACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGGCTGACTGTACATCTCTGGATTGTTCTCCATTAACAAAGAAGCACTTTAGGCAACACC

CACCAAAAAC

(CloneID 5405894)
CTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCT SEQ ID NO:116

TTGAGCGTTCTGCTCCGGCCCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATACCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGCTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTACATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGCTCGCCCCGGCCCCAGTGCCATCCAGAAGAAATTCCCC

TCGCGGCAAGTCCAGCCTCGTCTTGCGTTAG

CTCTGACTGTCCTGAACGCTGTCCTNCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTCGATTCGTTCCCATTAACAAAGAAGCACTTTAAAACAACC

CACACACACAACGCGG

(CloneID 4276593)
GAGCGCCGGCTCCTGCAGCGGTGGTCGGCTGTTGGCTGTGGAGTTTCCCAGCGCCCCTCGGG SEQ ID NO:117

TCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTTGATGAACCAACAGAACAACCT

GTGAGGAAGAACAAAAAATGGCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTT

GGGCCAAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACT

GCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTG

ATATTCATGCACAATGTGGACACAGACTTCCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCC

-continued

CGTGCTGCTCCGCCTGTGCCAGCCCGGTGGTCCGNGCCAGTGCCATCCAGAAGAAATCCCCT

GGCGGCAAGTCCAGCTCGTCTTTGGGTTAG

CTCTGACTGTGCCTGAACGCTGTCGTCTGTCTGTTCCTCCCATGCTTGTGAACTGACAATTG

GGCAGACTGTCATCTCTTGGTATTTGTTCCTTAACAAGAAGCTAAACAAACACAACACACAA

AAAATTTTGGCGGGCTGGGGCACGAGAAGTTTAACACCACTCTGGGCGCGCGCCCCAGGTGG

AGAAGGGCGCCGCGGTCACGGAACACCGGCAGACGCGGGGGGGGTTCCCACAAAAACGGGGG

GCCGGGGGTGAACACCACAGGGGGGGGTCGGTCACACGACCCCACTCAACCCCG

(CloneID 4127376)
GGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACC SEQ ID NO:118

TCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAACAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTCGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTGGTTTCATTAAAGAAAAACACTTTATGTACACAAC

ACACACACAAAAAC

(CloneID 5934901)
CTGTTGGGTCTCGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGC SEQ ID NO:119

GCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGACTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATCTGGACACACACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCACCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAACTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTANAAACAAGCACTTTATGTAAAAAAA

AAAAAAAAAAA

(CloneID 4622938)
CCACGAGGGGGTGTGGAGTTTCCTACCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCC SEQ ID NO:120

GGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCCACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

-continued

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTCGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTCATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGCCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTCTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTCATCTCTTGGATTTGTTTCATTAAA

(CloneID 3909220)
TGTGGAGTTTCCCAGCGCCCCTCCGGTCCGACCCTTTGAGCCTTCTGCTCCGGCGCCAGCCT SEQ ID NO:121

ACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAACCTTCTTGG

CCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTCGACTGCA

GAGAAGGAACTCCTCTGAAGCAACCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGACAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCTCCATGCTTGTGAACTGGACAACTT

GAGCCTGACTGTACATCTCTTGGTTTGTTTCATTAAAAGAAGCACTTTATGGTACTGGCTGG

CCTTTTTTTCTTCCTTTCGGACCAACCGGGTTCCTCTTCTGTCCCTGTACTCTGGGGTCCTG

GGGGCAGGCTTGGTGTTTCTCGTTGCTTGGGGCGCACAGCTTGTGTCACCTTTTCTGGTCTT

TCCAGACCACATTGGGGCCCCGTGTTGGGCTCTGTCGCGGTGTACTTTTGGTTCTGCCCCGG

CACTAACCGGTCTCGGCCCCGGGTTCTCCCTGTGGTCTCGGGGGGCCCCTTAAAGCGTCGGC

AGTTTAAGGG

(CloneID 6061001)
TGCTGTTGGGTCTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCG SEQ ID NO:122

GCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAACAAAATCAACCTTCTTGG

GCCAAGTCAGAAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGCTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGCTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

-continued

CCTTTTTTTTTTTTCTTTTGAAGAACAGGGTTCCTCTCTGTCCTTGACTCTTGGGGTCTCTG

GGGCCATGGCATGAGTGGTTTTCTAGTCAGAAGATTGGAAGGGAAGCTTTTGGGACACTTAA

CCACTGGGGTTTTTTAAGAAGAAATAATTTTGGTTCCCAGAATGTGTTTACAAGGATCT

(CloneID 4604404)
GAGCGGCGGCTCCTGCAGCGGTGGTCGGCTGTTGCGTGTGGAGTTTCCCAGCGCCCCTCGGG SEQ ID NO:123

TCCGACCCTTTGAGCCTTCTGCTCCGGCGCCACCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCCACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAACAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTCAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAACCTGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGACGAGAAGCCCGTG

CCTCCTGCCGCCTGTGCCCCAGCCCGGTGGACCCGGCCCACTGCATCCAGAAGAAAATCCCC

TGGCGGCAGTCCAGCCTCGTACTTGGGTTAG

CTCTGACTGTCTGAACGCTGTCGTACTGTCTGTTTCTCCATGCTTGTGAACTGACAATTGAG

GCTGACTGTACTTCTCCTGGACTTGGGTTGCATTCAAAGAGGAAAGCACCTTTTATTGTACT

TGGATGGGTATCTATAGTATACATATCGTTAGTGAAGAGACAGGGGTTCATATTAGTGTACC

TGGACTTTCTGGGGCATTGTGGCGCACTGGATTGAGGTCTATACTAAGAATAAATGAGAGGA

AACGGTTGTGACAAGTGTGATGGGATTAGAAGGAGACAATAGATGGGTCCAGAATT

(CloneID 4805717)
GAGCGGCGGCTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGG SEQ ID NO:124

TCCGACCCTTTGAGCGTTCTCCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCCACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGCATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGAAGACTTGGAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TTATGAAAAAATGTGGACACAGACTTGCCAGCCAGCCTGGGGCAGAGTGAAGAGAGGCCCGT

GCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCC

CTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTAC

(CloneID 4109734)
GGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC SEQ ID NO:125

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATT

CATGAAAATGTGCACACAGACTTGCCAGGCAGCCTCGGGCAGAGTGAAGAGAAGCCCGTGCC

-continued

TGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTG

GCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTCACAACTT

GAGCCTGACTGTACATCTCTTGGATTTGTTTCACTTAAAAAGAAGCACTTTATGTACTGTGT

CTTTTTTTTTTCCTTTGGACGACCGGGTTTCTCTCTGCCTTGACTCTTGGGTCTGTGGGCCC

TGGCTGAGTGTTCTCATCATGGGGGGCC

(CloneID 4553496)
GCTGCAGCGGTGCTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCGACCCTTTG SEQ ID NO:126

AGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCACCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTACATCTGAAGGGACGGAAGGTGA

TATTCATGAAAATGTGCACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCG

TGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCC

CCTGGCGGCAAGTCCACCCTCGTCTTGGGTTAG

CTCTGACTGTCCTCAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGGCACAAG

TTGAGCCTCACTGTACATCTCTGGATTGGTCTCATTAAACGAGAAGCACTTTATGTACTGTG

TCTTTATTTAGCATTTGAGAACAGGGTGCCTCTTGTCCCTGAATCTGGGTTGGGGGCAGGAT

GAGTGTTACTATAGTAATGCGCGAGGTTGGGCACTAGACGGGTTCAGGAAGAATATTGGGGC

GATGTGTAAGGATTGGGAGGGGTAACAGTAGGGCGTGACAGCGTAATGGAGGCTATAGTCAG

GAGCTGGAGGCCACCAGAGTGGGATTGGTATATGAGGTGAAGCTTGGGTGAAAGGCTCGGAC

GAAGTTAGCAGAA

(CloneID 4156131)
CGGACGGCGTGGGCGGCTGTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGA SEQ ID NO:127

GCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAACGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGC

CAAGTCAGCAGCTGCCAACTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACGTGCAG

AGAAGGAACTCCTCTCAAGCAAGCTCCGGACACTTCTTACATCTGAAGGGAGAAGGTGCTAT

TTATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGCCACAGTGAAGAGAAGCCCGTGC

CTGCTGCGCCTGTGCCCACCCGGTCGCCCCGCCCCAGTCCCATCCAGAAGAAATCCCCCTGG

CGGAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTTCCTCCATGCTTGTGAACTGGACAAA

TTTGAGCCTGATGGTACATCTCTTTGGATTGGTTCATTAACCAAGAAGCACTTTATGTACGC

NGGCCTAGATTCCTGTGAGGACCGGTTCTCTCGTGTCCTGGACCTGGTCGGGGGGCCGGATG

AGGGGTCTAGAGGGATCGGGGGGAACGTGCGGACCTAGACGGGGTCGAACAACACTTCGGGC

CCACTGGGTGCACAGCTTGGCGGGGGTAACCCAACTGGGGCCCGCCCCGGGGACAAAAAGCA

-continued

AGGCGGCGGCGCGACCCAGCGCCGCCACCTCGCAAGAGGGGGACCCCAAGAAAAGACCCCGG

CGACAGTATTACAACCGAGG

(CloneID 5014564)
GCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTT SEQ ID NO:128

GAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAPAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTCAAGGGAGAACGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGC

CTGCTGCGCCTCTGCCCACCCCGGTGGCCCCGGCCCCAGTCGCATCCAGAAGAAATCCCCCT

CGCGGCAAGTCCAGCCTCGCTCTTGGGTTAG

TCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCCACCACT

TGAGCCTGACTGTACATCTCTGTGGATCGTTTCATTAACNACCGACGCACTTTCATGTCACT

GCGTGGTCGTATTGCTTTCCCCCTCTGCGCAGCAACCGGGTTCCTCCTCGGGGCCCTGAACC

TCTTGGGACCGTGTGGCCCATGGCCTGAGCGTTCCTACTCGCAACACTCGCCGGCACACCAT

CGGCGCACCCGCACCGCCCTTCTCCCGCACAACACACTCGGCTCCCAGGTCTCCACCACCTC

CATCCGGCCCCACACACACCATGCCGTCCCCACG

(CloneID 3458462)
CCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTC SEQ ID NO:129

CTCGGCGCC

ATCACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTACGTTTCATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTCAAGAAAATCAAGCTTCTTGCG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAC

AGAAGGAACTCCTCTGAAGCAACCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGC

CTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGGCAAGTCCAGCCTCGTCTTGCGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTCTCTGTTTCCTCCATGCTTGTGAACTGACAATTG

AGCCTGACTGTACATCTCTGGGATTGTTTCATTTACAAAAGAAGCACTT

(CloneID 3910873)
TCCCAGCGCCCCTCGGGTCCGACCCTTTCAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTC SEQ ID NO:130

CTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

-continued

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCCTG

CCTGCTGCGCCTGTGCCCACCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTCCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTCAAAAAGAAGCACTT

(CloneID 5928362)
CCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTT SEQ ID NO:131

GAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

CCCGCCTCCAGGTGGTGCATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGCGAGAAGGTGATA

TTCATCAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAC

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTAAAAAAAAAAA

AANAAAAANTAA

(CloneID 3904371)
CCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCC SEQ ID NO:132

TCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCCAGTTTT

GCGGCCTCCAGGTCGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTCGGCCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCCGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAC

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTCGTTTCATTAAAAAAGAAGC

(CloneID 4578473)
GCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAG SEQ ID NO:133

CGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACACCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTCATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTCAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

-continued

TTCATGAAAATGTGGACACAGACTTGCCAGGCACCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGCCCCCGCCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTCGGTTAG

CTCTGACTGTCTGAACGCTCTCGTTCTGTCTGTNTCCTCCATGCTTGTGAACTCCACAATTG

AGCTGATCTCATCTCTTGGATTTGTTTCATTAAAAGGAGCACTTATGTACTGCTGGCTTTTT

TTTCTTGGACACCCGCTCCCCCTGCCTTGCTCTGGGTCTGGGGCACCCTGGAGGGTCCAGTA

GAATGGGGAGGAACTGCGACCTTGCGGTTTCGAACAACTCGCCCCTGTTACGACTTGGCGGG

GTACCTCCTGGGTCCGCCCTGGCAGAAGCAGGGCTCGCGCCCGAGCCTTGGCACACGGAACC

TAGAAAGGCTCGCAATTCACGCCGAGGGTCGGGGGGCTGGG

(CloneID 4559434)
GGCTCCTGCAGCGGTGGTCCGCTCTCGGTGTGGAGTTTCCCAGCGCCCCTCGGCTCCGACCC SEQ ID NO:134

TTTGAGCGTTCTGCTCCGCCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAACGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTCTT

GCGGCCTCCAGGTGGTCGATCCAATTTTTCTTTTAGGTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAACTCAGCAGGTGCCCAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAACACAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCT

GGCGGCAGTCCAGCCTCGTCTTGGGTTAC

CTCTGACTGTTCTGAACGCTGTCGTCTGTCTGTTTCCTTCCATGCTGTGAAACTGCACAACT

TGGAGCCCTGACGTGTACCTCTCTGGGATCGTTCCATTCACCAGAAGCACTTTTATGTGCTG

CTGCTATCTTACTCCTGTCCGAGACAGGGTTCTCTCGTCCCTTGACTTCGGGGTCGGGGGCC

CCGGCCCTGTCTGCTTCCTGACACATTGGGGGCGAGAGTCTGTCTCCCTGCGACGGGTTCAG

ACAGAATCGGGCCAGTCGGTGGAGCGTGGTCGGGACAACTCGGGAACAGCCAGGAGAATAAC

GACGCTCGGGCGCCCCAGGCCTCGACTGCGCCTCTCCCGCAACCCCAGACACCCCGCGGAA

(CloneID 4547048)
GCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC SEQ ID NO:135

ATGACCACAACCACCACCTTCAACGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTC

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAACAACAAAATGGCCTCTAATATCTTTGGGACACCTCAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGCGAGAAGGTGAT

ATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGCCAGAGTGAAGAGAAGCCCGT

GCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCT

GGCCGGCACAGTCCAGCCTCGTCTCGGGTTAG

CTCTGACTGTCCTGAACGCTGTACGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAAG

CTTGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAACGAAGC

-continued

(CloneID 4276268)
GAGCGGCGGCTCCTGCAGCGGTGCTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGG SEQ ID NO:136

TCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCCCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGC

GCCAAGTCAGCAGGTGCCAAGTCTAGTGCTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGGAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CTGCTGCGCCTGTCCCCAGCCGGTGGCCCCGGCCCAGTGCCATCCAGAAGAAATCCCCCTGG

CGGCAAGTCCACCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTCCTCCATGCTTGTGAACTGCACAATTT

GAGCCTGACTGTCATCTCTGGGTTTGGTTCTTTAAACAGAAGCCTTTGTGCGCGGCTTCTTT

CTCTCGAGACCACCGTTCCCTGCGCTGGCCCGGGCTCGGGGCCGGCCAGGGTTTCCCCCACT

GGCGGCGCCCTCTGGCCCCCCGGTGTCTCCACACAACATGGCGCCCGGTAGGCTCGTCGGCG

TACCCTCGTGGCCCCCGG

(CloneID 3659047)
GTTGGCTCTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGC SEQ ID NO:137

CAGCCTACCTCGTCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGATGTCTGACCCCAACAGCAGGAATAGCTCCCGAGTT

TCGCGGCCTCCAGGTGGTGGATCCAATTTTTCATATAGGTTTTGATGAACCAACAGAACAAC

CTGTGAGGAAGAACAAAATGGCCTCTAATATCTTGGGGACACCTGAAGAAAATCAAGCTTCT

TGGGCCAAGTCATCCAGGTGCCAAGTCTACGTGGTGGCAGGGAAGACTTGCAGTCATCTGGA

CTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAACGGGAGAAG

GTGATATTCATGAAAATGTGGACACGAGACTTGCCAGTGCAGCCTGGGGCACAGTGAAGAGA

AGCCCGTGCCTGCATGCGCCTGTGCCCAGCCCGGTGTGCCCCGTGCCCAGTGCCATCCAGAA

GAAATCCCCCGTGGCGGCAAGTCCAGCCTCGTGATTGGGTTAG

CTCTGACTGTCCTGAACCCTGTCGTACTGTCTTGTTCCTCCATGCTTGTGAACTGCACAACT

TGACCCTGACTGTACA

(CloneID 3634373)
GGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGC SEQ ID NO:138

CTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAACACTTGGAGTCATCTGGACTGCA

GAGAAGGACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAAGAAGGTGATA

TTCATGTTTTCATAGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAG

AAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGCCCCGGCCCCAGTGCCATCCAGAAGA

AATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

-continued

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAATT

GAGCCTGACTGTGCTCTCTGGATTTGTTCCTTAACAGAGCCTTTTGTCTGG

(CloneID 5741030)
CTCCTCCAGCGGTGGTCGGCTGTTGGGTGTGCAGTTTCCCACCGCCCCTCGGGTCCGACCCT SEQ ID NO:139

TTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATACCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACACAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGACTCATCTGCACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAGAAGCACTTTAAAAAAAAAAAAA

AAAAA

(CloneID 3870232)
ATCCCATGACGCCCCTCGGGTCCGACCCTTTGAGACGTTCTGCTCCGGACGCCAAGCCTACC SEQ ID NO:140

TCGCATCCTCGGCGCC

ATGACCACAACCAACCACCTTCAAGGGAGTCGACCCCAACACCAGTGAATACGCTCCCGAGA

TTTTGCGGCCTCCAGTGTGGTGGATCCAATTTTTCATTAGGTTTGTCATGAACCAACGAGAA

CAACCTCTGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGC

TTCTTGGGCCAAGTCACCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTG

GACTGCAGAGAAGGACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAAGGGAGAA

GGTGATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAA

GCCCGTGCCTGCTCCCCCTGTGCCCAGCCCGGTGGACCCGGGCCCAGTGCCATCCAGAAGAA

ATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTCAACGCTGTCGTTCTGTCTGTTCCTCCATGCTTGTGAACTGCACAACTT

GAGCCTTGACTGTACATCTCTTGGACTAGGTTCACTTAAAAGAAGCACTTTATGTACTGCTG

TCTACAGAAAAACACCCCGGCAGACCGGGTACCCCTACAGGACGTGCACCTTGGGCCAGGGG

CCACGGCCAGCCGGGCCACCAGATCCCACACTGGCAGGCAACCCTGCGACACACCACCACCA

CGGGAAACACAACACACACAACAGGCCACCGCACGGTAAGAAGGACCTCAGGCCACGAGCAC

ACACACAAAACAGGGGGCACAAACCACACGTGGACGACCAAACACCACAGCGACCGAAAGGC

GGCCGACAACAGAGACACTCCCCCACACACCGGGTGCCCCAACACCAAACGCATGACAAACA

CGTTCAACACACACACAAGCCCC

(CloneID 3905523)
GTGTGGAGTTTCCCAGCGCGGCCTCCGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCACC SEQ ID NO:141

CTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGACTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTACGTTTTGATGAACCAACAGAACAACCTG

-continued

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCACGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGCAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAACATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAACCCCGT

GCCTGCTGCGCCTGTGCCCAGCACCGGGTGCCCCCGGCCCCAGTCCCATCCAGAAGAAATCC

CCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACCGCTGTCGTTCTGTCTGTTTCCTCCATCCCTTGTGAACTGCACAA

TTGAGCCTGACTGTACCTCTCTTGGATCCGGTTGCCATGCACACAGAAGCGCTGGCGAAAAA

GACACCGAGCCGACCACCGACGC

(CloneID 5517974)
CGGACGCGTGGGCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGA SEQ ID NO:142

GCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAAAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCCATGCTTGTGAACTGCACAAC

TTGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGCACTGCT

GTCTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGGCCTTGACTCTTGGGGTCTGTGG

GCCCTGGCATGAGTGTTTTCCTAGAATAGATTGGAGGGAAAGCTTTGTGACACTTAATACTG

TGGTTTCAAAAAAAAAATAATTTGGTTCCCAATGTGTTAAAAGGAACTTTTGGTACTGGAGG

TTTTTAAAACCTTTAATTTGGGTTTAACCAAGGCCTAACCTGGTACGAACCTTAAAACGTGT

CCCACGGGCCCCCGTTTCCTTGGCACGGGCCCAAACCCCACCGGGGGAGGCCTCTCTCACCA

AAAACCCTTC

(CloneID 4124775)
GCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC SEQ ID NO:143

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATT

CATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGACTGAAGAGAAGCCCGTGCC

TGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTG

GCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTCTCCTGAACGCTCTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

-continued

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAGAAGCACTTTATGTACTGATGT

ACTTCTTATTTTTCCTTTGAAAGAACCAGGTTTCTCTCTGTCCCTGATCTGGTCTGTGGCCC

ATGGCACGGAGTGTTCCTTGTAGTACGATTGCGAGGGCAAGCTTGTGAACTATAGTACGGGT

TTACGACGAACAATATCTGGCTCCCCAGGTCTCCACGACCTTGACGGGCGTTTCACACCTAC

TGGGTTCCCAGCTAGTGGCAGCACAAACGCACAGCCCGGTCGCCAGCCACCCAGAGTTCGGG

ACCCTGGTGTGCTATGCAGGCTGGGACACCCTGGATGTCCAAGGGGCCTGCTGCGGAAAGGT

TTTATTAGCTAAG

(CloneID 5832873 in 3' to 5' orientation)
TTTTTTTTTTTTTTTTTTAAACTCCTTCTTTTTAATGAAACAAATCCAAGAGATCTACAGTC SEQ ID NO:144

ACGCTCAAGTTCTGCAGTTCACAAGCATGCAGCAAACAGACAGAACGACAGCCTTCAGGACA

CTCACAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGCGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCGGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGCAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTGGATCCACCACCTNCGAGGCCGCAAAACTCGGGAGCTATTCCTGCTGTTGGGG

TCGACTCCCTTGAAGGTGGTGGTTGTGGTCAT

GGCGCCGAGGAGCGAGGT

(CloneID 4277604)
GAGCGGCGGCTCCTGCAGCGGTGGTCGGCTGTTGGGTCTGGAGTCTCCCAGCGCCCCTCGGG SEQ ID NO:145

TCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGACTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGAAGACTTGGAGTCATCTGGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTCAAGGCAGAAGGTGATAT

TCATGAAAATCTGGACACAGACTTGCCAGGCAGCTGGGGCAGAGTGAAGAGAAGCCGTGGCT

GCTGCGCCTGTGCCCAGCCGGTGGCCCGGGCCCACTGCATCCAGAAGAAATCCCCTGCGGAA

GTCCAGCCTCGTTTTGGGTTAG

TCTGATGTCCTGAAGCTGTCGTTCTGTCTGTTCCTCAAGGTTGTGACTGACATTGAGCTGAT

GTACTTTCTGGATTGTTCATTAAAAGAGCTTTTGAAAAAAAAAAAAAAAAAATTTGGGGGCC

GGCCCGAAATTTAACATGGGGGGGGGCCGGGTAGAATGGACGGCAGAAACGGGAAGGGGGTC

ACAAAACGGGGAGGTGGAACGGGGGGTAGGTAAGGCGTACAG

(CloneID 5834676 in 3' to 5' orientation)
TTTTTTTTTTTTTTTTTTAAAGAGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTC SEQ ID NO:146

AGGCTCAACTTGTGCAGTTCACAAGCATGGAGGAAACACACAGAACGACAGGGTTCAGGACA

CTCAGAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

-continued

GGGCCACCGGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGATGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGCTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTGGATCCACCACCTGGGAGGCCGCAAAACTCGGGAGCTATTCCTGCTGTTGGGG

TCGACTCCCTTGAAGGTGGTGGTTGTGGTCAT

GGCGCGGAGGAGCGAG

(CloneID 5405407)

GCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGT SEQ ID NO:147

TCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAAGCAACGCTTCGGAAGACTTCCTTAAAATTCGGACAGGACAAC

GGGAGAATATCACAGAGAAATGTGTGGCACACAAGATTGGCCAGGGCACCCGTGGGGCAGAT

GTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGGTGGCCCCGGTCCCAGTTGC

CATCCAGAAGAGATCCCCCTGGCGGCAAGTCCCAGCCTCGTCTTGGGTTAG

CTCTGACTGGTCCTGAACCGCTGTCGTTCTGTCTGTGTCCTCCATGCTTGTGACAGTGCACA

CACTCGAGCCTGACTGTACATCTCTCGGATTCGTTACATTCAGAAACAACGCACTTCTATGT

ACTGCCTGTCCTTTCTTTCCTTCGCTTTTGGCAGGACCCTGGGTTTCTCTCCTGCCCCTGCA

CTCTGGGCCTCCGCTGGCCCCAGGCCCTGTGCTGTCTCCCTACGGCCTCCGCATGCCAGCGA

TTCGCCTGGTCGTGACCCTTAGGCCTGGCGTTTTACCCCCG

(CloneID 4563923)

GCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTT SEQ ID NO:148

GAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGNGTTTTGNATGAACCAANCAGAACAAC

CTGTGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCT

TGGGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACT

GCAGAGAAGGAACTCCTCCCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGG

TGATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGC

CCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAAT

CCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCAT

(CloneID 5533034)

CGGCTGTTGGGTCTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCC SEQ ID NO:149

GGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

-continued

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGCGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTCGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGCCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCCTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTCCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTCTTTTGAAGAACACGGTTTCTCTCTGGCCTTTGACTCTTGGGGTCTGTGG

GGCCATGGCAATGAATGGTTTTCCTAAGTAGTAAGAATTTGGAGGGAAAAGCTTTTGTTGAC

CCCTTAATCAACTGGGGTTTTTTAAGAAGGAAAATTAATTTGGGGGTCCCACAAATGGTGTT

TAAAAAGGAATCTTTTTTGGGACCTGGGGCGGGTTTTTTTAACCCCCTTTTTACTTTGGGGG

GTTTTTACCCAAGGCCCTTCAAA

(CloneID 6060969)

AGCGGTGGTCGGCTGTTCGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCG SEQ ID NO:150

TTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTCGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCACGCAGCCTGGGCCAGAGTGAACAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGCCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAACTCCAGCCTCGTCTTGGGCTAG

CTCTGACTCTCCTGAAACGCTGTCGNTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAAC

TTGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAAGAAGCACTTTATGTACTGC

TGTCTTTTTTTTTTTTTTTCCTTTGAAGAACAGGTTTTCTCTCTGTCCTTTGACTCTTGGGCT

CTGTGGGCCATGGGCATGAATTGTTTTTCTACTACTAGAATTGGAGGGAAAAGCTTTTGGGA

CACTTAANACTGGGGTTTTTTAAGAAGAAAATAAT

(CloneID 3347594)

TGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTAC SEQ ID NO:151

CTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAAATAGCTCCCGAGTTTTGC

GGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATT

CATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCACAGTGAAGAGAAGCCCGTGCC

TGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTG

-continued

CCGGCAACTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCCTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAA

(CloneID 3531894)
TGTGGGTGTGCAGTTTCCAGCGCCCCTCGGGTCAGACCCTTTGAGCGTTCTGCTCCGGCGCC

AGTCTACCTCGCTCCTCCGCGCC

ATGACCACAACCACCACCTTCAAGGCAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTAT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTAGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGCTGCCAAGTCTAGTGCTGGCAGGGAACCACTTGGAGTCATCTGGACTGC

AGAGAACGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGA

TATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCG

TCCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGGCCAGTGCCATCCAGAAGAAATCCCC

CTGGCGGAAGTCCAGCCTCGTCTTGGGTTAC

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTACTCCATGCTTGTGAACTGACCATCC

GAGACTGAATC

(CloneID 4276267)
GAGGGGGGCCTCCTGTAGGGTGGTCCGTGTTGGGTGTGGAGTGTCCCAGCGCCCCTCGGGTC SEQ ID NO:153

CGACCCTTGAGCGTCTGCTCCGGGCCAGCTACCTCGCTCCTCGGCGCC

ATGAGCACAATCCAGCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTT

TGCGGCTCCAGGTGGTGGATCCAATTTTTCATTAGGTGTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGGCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGCAGTCATCTGGACTGCAGA

CGAAGGAACTCCTCTGGAAGCAAGGCTCCGGAGACTTCTTAGATCTGAAGGGAAGAAGGTGA

TATTCATGAACAATGTGGACACAGACTTGTCCAGGCACGCCTGGGGCAAGACTGAAGAGAAG

CCCGTGCTTCTGGGCTGTGCCATCCGGTGGGCCCGGCTCCAGTGCTATCCAGAAGAAATCCC

CCTGGGGGGAAGTCCAGCCTCGTCTTGGGGTTAG

CTCTGACTGTCCTGAACGCTGTCCGTCCTGTCGTCGTTACATCCATGCTTGTTGAACTTGGA

CCAACTTGGAGTCTGATGGGAGATCTCCTGGATTGGTTCATTAACCAAAACGCGGTTAGGGA

ATGGCCGCCCTTGGGGGAGTCTTCTGACAGACGAAGGTTCACCACTGGGCGAATTCGGCGGG

TCGGGGGACAGGAACACAGGGTCCACAGGCAATGGGGCAGACCCATTGTC

(CloneID 4547025)
GGCGGGTCCGACCCTTTGAGCGTTCCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC SEQ ID NO:154

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTG

CCGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAATCAAAAGCTTCTTGCG

CCAAGTCACCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAACAGAAGCCCGTGC

CTGCTGCGCCTGTGCCCAGCCCGGTGGCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTG

GCGGCAAGTCCAGCTCGTTCTTGGGTTAG

-continued

CTCTGACTGTCTGAACGCTGTTCGTTCTGTACTTGTTTCCTCCATGCTTGTGAACTGCACAA

CTTGAGCTGACTGTACGATCTCTAGGATCTGTCTCATTAAAAAGAAGCCACTTATGTACTGG

TGTCTTTTTCACCTTAGGAAAAACAGCTTCCCTCAGGCCTTGACCACTTGAGGTCTGGAGCC

AAGCGCATGAAAGGTCCCTAGGCACCCGATCGGAGGGAAACGCCTGCGACCTTACTACCGCA

TCTCACACGACACAAGTGCCGTCCCACCTGTCAAAGCACACGTGACCGAGGTCCAACGCATA

CCGGTTGCCGCCTAATTGCAGACCAAACGCCCCCACCAGCGCGCCCACACGGACTCGACACT

GCTGCAAGGCGACCGCAACCCCTCACACAAACCCCCCTCCCAACAGATATCCAGCGAACCCA

CCACGCAGCAAAACGTCCATCACAAAACGTC

(CloneID 5314132)
AGCGGAGCGGCGGCTCCTGCAGCGGTGGTCGGCTGTGGGTGTGGAGTTTCCCAGCGCCCCTC SEQ ID NO:155

GGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGT

CAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAG

AGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAACAGAAGCCCGTGC

CTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTCCCATCCACAAGNAATCCCCCT

GGCGGCAAAGTCCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACCGCTGTCCGTTCCTGTCCTGGTTTCCTCCCATGCTTTTTTTTTCT

CGTCATCTCGTACGTGAAGGGTGACAGTCACCGACAATACCGTCCGCTGCCGAGAAGACCTC

AACGCGACGGAGTAACACAACATCAGCATATACCAATGAACCTTCCCGATGGCGTCAGACTC

TCCACGCGACATATACTAGA

(CloneID 5576026)
CTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCT SEQ ID NO:156

TTCAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATCTGGACACAGACTTGCCAGCCAGCCTGGCGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCCCAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAACACTTTATGTACTGCTG

GCCTTTTTTTTTTTCCTTTTGGAAAACAGGTTTCCTCCTCTGTCCTTGAACTCCTGGGTCTT

GGGGGCCCAGGGCCTGAAGGGGTTTTAATTATCAAACTTGGAGGGAAAACCTTTTGGACACT

CTTTACTTGGGCCTCTTCACAAAACACCACCTCGGTCCCCACAAGGGGTTCGACGGCCTCTT

-continued

TTGTCACGAAGGTTTCCCCAATTTAATATTTGGGCGTTCCATAAACCTCCATTGGCGACGGC

CCAATACACCGGCCATACGGCCGCCGCTCCTTGTCTAATCCCGCCCCCGGATGGGAGGCTCT

TCTTCCAGAGGCCCTCATTGGGGCGGCCCCCGCTCTATCACAGC

(CloneID 5165195)
GCAGGACAGGCGGCAGCCGCGGCTCCTGCAGCGGTGGTCTGGCTGTTGGGTGTGGAGTTGGG SEQ ID NO:157

GAGCGCCCCTCGGGTCCGAGCCTTTCAGCTGTTCTGCTCCGGATGCCAGCCTACCTCGCTCC

TCGGCCCC

ATGACCACAACCACCACCTTCAAGGGAGTCAGAACCCAACAGTCAGGAATAGCTCCCGAGTT

TTCCGCGCCTCCAGGTGTGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAA

CCTGTGAGGAAGAACAAAATGGTCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTC

TATGGGCCAAACTCACGCAGCTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTG

GACTGCAGAGAAGGAACTCCTCTGAAGCAAGCATCCGGAGACTTCTTACATCTGAAGGGAGA

AGGTGACTATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAG

AAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCTGGTGGACCCTGTGCCCAAGTGCCATCCAG

AAGAAATCCCCCTGGCGGCAAGTCCAGAACTCGTCTTGGGTAAG

CTCTGAACTGTACCTGAACCGCTCTCGTTCTGTCTGTATCCTCCATGCTTGTCAACTGCACA

ACTTGAGCCTGACGCTGTACATTCTCCTGGGGATTTCGCTTCCCACTTAAAAAAGACGCCAC

TTCACTGTTACCAGCGAAAAAAGGCAGCAGCGAGAACAGAACAAACAATTGGTGAG

(CloneID 3449557)
GTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGC SEQ ID NO:158

CAGCCTACCTCGCTCCTCGCCGCC

ATGACCACAACCACCACCTTCAAGGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTT

TGCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTCGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAACCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCCGTGGCCCCGGCCCAGTGCCATCCAGAAGATATCCCCCT

GGCGGCAAGTCCAGCCTCGTCTTCGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTTGTCTTGTTTCCTCCCATGCTTTGTTGAACCTGC

ACAAACTTGAAGCCCTTGACGTGGTACAATCTCCTTGGAGTTTGTGTCCACTTTAAAACGAA

GCACTTTTTAAATAG

(CloneID 3351305)
GCTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGTGCGCCAG SEQ ID NO:159

CCTACCTCGCTCCTCGCCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAACTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAACGGAGAAGGTCATA

-continued

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTCGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCCCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAATT

TGAGCTGACTGC

(CloneID 3029189)
TTGGCTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCC SEQ ID NO:160

AGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCCGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATCCTT

(CloneID 5465786)
GCTGTTCGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCCACCCTTTGAGCGTTCTGCTCCGG SEQ ID NO:161

CGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAACACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTCCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCCCTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTAAAAACAAAAA

AAAAAAAAACTCGA

(CloneID 2819762)
GGC SEQ ID NO:162

ACGAGGACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGCAAGACTTGCAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAACCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTCTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

-continued

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGGTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTANNAAAAANAA

AAANNAAA

(CloneID 2819762 in 3' to 5' orientation)
TTTTTTTTTTAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGCTCAA SEQ ID NO:163

GTTGTGCAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCCGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTCCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTTGGATCCACCACCTGGAGGCCGCAAAACTCGGGAGCTATTCCTGCTGTTGGGG

TCGACTCCCT

(CloneID 2819763)
GGC SEQ ID NO:164

ACGAGGACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTCCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAAAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCAT

(CloneID 2819763 in 3' to 5' orientation)
TTTTTTTTTTAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGCTCAA SEQ ID NO:165

GTTGTGCAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAG

CTAACCCAAGACGAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCGGGCTGGGCACAGGCGCAGCAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTGGATCCACCACCTGGAGGCCGCAAAACTCGGGAGCTATTCCTGCTGTTGGGGT

CGACTCCCT

(CloneID 3912245)
GGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGC SEQ ID NO:166

CTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

-continued

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACACACTTGCCAGGCAGCTGGGGCAGAGTGAAGAGAAGCCCGTGC

CTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCT

GGGCGGCAAGTCCAGGCTCCGTCTTGGGGTTAG

CTCTGACTGGTCCTGAACGCTGTCGTTCTGTCTGGTGTCCTCCATGCTTTGTGAACTGGGCC

CAATTTGAGGCCTGGAGTGGTGCCTCTCTTGGGTTTTGGTGCCTCTTAAGGAAGCCTTTTTT

GTCCGGCGGCCCTTGCCCTTTGCGGCACCCGGTGTTCTCCTGCCCTACACCGCGGGCGGGGG

CCCGG

(CloneID 5164014)
C SEQ ID NO:167

CGGCGGCCTCCTTACCCTCGCCCTGCCCCCACTCCTTATTCGCACCTACCTCACAGAGTTTT

GCGGCCTCCAGGTCGTGGATCCAATTTTTCATTAGGTTTTCATGAACCAACAGAACACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

CCAAGTCAGCAGGTGCAACTCTAGTGGTCGCAGGGAACACTTGGAGTCATCTGGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATT

CATGAAAATCTGGACACACACTTGCCAGCCAGCCTGGGGCAGACTGAAGAGAAGCCCGTGCC

TGCTCCGCCTGTGCCCACCCCGGTGCCCCCGGCCCCAGTGCCATCCAGAACAAATCCCCCTG

GCGCCAAGTCCAGCTCCTCTTGGGTTAG

CTCTCACTGTCCTCAACGCTGTCCTTCTGTCTCTTTCCTCCATGCTTGTGAACTGCACAACT

TCAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAACAGAAGGCACCCTTTAATAAA

CAAACACAAGAACACAGAACAAAACACAACAAAGAACAACGGCGGG

(CloneID 4284754)
GAGCGGGGCTCCTGCAGCGGTGGTCGGCTGTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTC SEQ ID NO:168

CGACCCTTTCAGCGTTCTGCTCCGGCGCCACCCCTACCTCGCTCCTCGGGCC

ATGGACCACAACCACCACCTTCAAGGGACTCGACCCCAACAGCAGCAATAGCTCCCCAGTTT

GGCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTGTGGCACACCTGAAGAAAATCAAGCTTCTGG

GGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCACCGAAGACTTCGAGTCATCTGGACTCC

ACACAAGGAACTCCTCTGGAGCAAGCTCCGGAGACTTCTTACATCTGAAGGGAGAAGGTGAT

ATTCATGAACAATGTCGACACACACTTGCCACGCCGCCTGCCCGCACGACTCAAGAGAAGCC

CGTGCCTGCTCCCCCTGTGCCCAGCCCGGTGCCCCGGCCCCAGTGCCATCCAAGAATCAAAT

TCCCCTGCGGGCAGAGTCCAGGCCTCGTCTTGGGTTACG

CTCTGACTGTCTGAACCGCTGTCGTTCTGGTCTTGTTTTCTTCCATGCTGTGTCAACGTGGA

CAACTTTGAGGCCTTGATGTACATCTCTGGATGTGTTCACTTAAAAAGAGCAATTATGTATC

GGTGCCTTTTCATTTCCTTGGACAAACCGGTCTCTCTGACTTCAACTGGGGTTGGGGCAGGC

ACCACGATGTTTCAACACAATGGAGCACCN

(CloneID 428337)
GAGCGGCGGCTCCTGCAGCGGTGGTCCCCTGTTGGGTGTGGAGTTTCCCAGCCCCCCTCGCT SEQ ID NO:169

CCGACCCTTTGAGCGTCTGCTCCGCCGCCAGCCTACCTCGCTCCTCGGCCCC

-continued

ATGACCACAACCACCACCTTCAACGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAACAACAAAATGGCCTCTAATATCTTTCCGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGAAGACTTGGAGTCATCTGGACTGCAG

ACAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATAT

TCATGAACCATGTGCACACAAGACTTGCCAGGCAGCCTGGGGGCAGAGTGAAGAAGAAGCCC

GTGCCTGCTTGCGCCTGTGCCCAGCCCGGGTGGCCCCGGCCCCCAGTGCATCCAGAAGAAAT

CCCCCTGGGCGGCAAGTCCAGCCTCGTCTTGGGTAG

CTCTGACTGTCTGACGCTGGCCTCTGGCTGTTCCTCCTGCTTGTCAACTGCACACTTGAGCC

TGACTGGACATCTCTTGGATTCGTTCATAAAAGGAGCTTGCAAAAACAAAAAAAAATTTGCG

CGCCCCCCAAATTTTAACTCTTGGCGCCCCGGTGAATGGCGCGCCAGACCGCACAGGGGTCA

CAAACCGGGCGGCGAAACAGGGGTTGTCGCCACACCCCGG

(CloneID 5936059)
AGCGAGGCCTGGGTGGCGATCCAAGACGAGGCTGCAGCGGTGGTCGTCTGTTGGGTGTGGAG SEQ ID NO:170

TTTCCCAGCGCCCCTCGGGTCCCACCCTTTGAGCGTTCTGCTCCGCCGCCAGCCTACCTCGC

TCCTCCGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGCAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTCATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTCGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGCGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCGGTTTCCCTCCATGCTTGTGAACTGCACAAC

TTGAGCCTGACCTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTAAGGACTGC

TGGCCTTTTTTTTTTTTCTTTTGAAAAACAGGTTTCCCCCCGGCCCTTTGATCCTTGGGGTC

TGGGGGCCCATGGGAATGAAATGGTCTTCTAACTAACAACATTGGGAAGGAAAAAGCCTTGG

GGACACCTCAACAACGTGGGGTTTTTAAAAAGAAAAAAAATTTGGCTTCCCACAGGGGGTTA

AAAGGAATCTTTTTGTACCCGAAGGGTTTTTAAACCCTTTTACCTGGGGGGTTACCCAGGCC

CTCCCACTTGGCACCGGAACCATTTAACCGGTCTCCAGAGGGGCCCCGTTTCCCGGGCCAGA

CCCCCCCACCCTCACAGGGGAAGTCTCTCCTCCCCAAAAAACTCTCCCTTGGGGGGTGGGCT

CCCCTACCTTTCGCCCGTGGGGCCCTTTTGTCACAAA

(CloneID 4280999)
GAGCGGCGCCTCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTCTCCCAGCGCCCCTCGGG SEQ ID NO:171

TCCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTGATGAACCAACAGAACAACCTGT

GAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGG

-continued

CCAACTCAGCAGGTGCCAAGTCTAGTGGGTGGCAGGGGAACGACTTGGAGTCATCTGGACTG

CAGAGAAGGAACTCCTCTGAAGCAACCTCCGGAGACTTCTTAGATCTGAAGGGAGACGGTGA

TATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCG

TGCCTGCTGCGCCTGTGCCCAGCCCGGTCGGCCCGGCCCAGTGCCATCCAGAAGAAAATCCC

CCTGGCGGGCAACTCCCAGCCTTCCGTCTGTGGGGTTAAG

CTCTTGACTGGGTCCCTGAACGCTGTCCGTTCCTGTCTGTTTTCCCTCCCATGGCTGGTGAA

CTGGACCAACTTTGAGCCTGACTTGTACATTCTCCTGGGATTCGTCTCCTTTCACCCAGAAG

GCCTTTATGACACACAACAACACTCCAAACACACCTTGGGCGCCGCGCGCCCCTCACACATT

TTAAACCCTCTGTGCGGCCGGCCCACGGTAGGACAGGGCTAGGCCCGGCAACGGGAGGGACG

GTGGTTACCAAAACCCGGGCGGGCTGCCCCACGGTCCGCTCG

(CloneID 6154219)
GGACGCGTGGGCGGACGCGTGGCCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGA SEQ ID NO:172

CCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAGCAAGCTCCGGAGACTTCTTAGATCTGAAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACCCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTCCACAACT

TGAGCCTGACTGTACATCTCTTGGGGATTTGTTTCATTAAAAAAGAAGCACTTTTATGTACT

GCTGGNCTTTTTTTTTTTTTTCTTTTTTGAAGAACAGGTTTTCTCTCCTGTCCCTTGACTCC

TTTGGGGTCTGTTGGGCCCATGGGGCATGGAGTTGGTTTTCTAAGGTAAAGTAGAAATTGGG

GAGGGGGAAAAGCCTTTTGGTGGACACCCTTTT

(CloneID 4045037)
GGCGGGGCTCACTGACAGCGGTGGTCGGGTGTGGGTCTGGAGTTTCCCAGCTGGCCCCTCGG SEQ ID NO:173

GTGCGACCCTTGTGAGCGTTTCTGCTCCGGTCGCCAGTCGTTACCTCGCTCCTCGGGCC

ATGACCACAAGCACCAGCTTCAAGGGAGTCGATCCCAACAGCAGGAATAGCTCCCGAGTGTA

GCGGTCTCCAGGTGGTGCATCCAATTCACATCAGGTTTGATGAACCAAGAGAACAACCTGTG

AGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAACCTTCTTGGGC

CAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCACGGAAGACTTGGAGTCATCTGGACTGCAGA

GAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGACACGGTGATATT

CATGACAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCCAGAGTGAAGAGAAGCCCGTGC

TGCTGCGCCTGTGCCAGCCCGGNTGGCCCGGCCCAGTGCCATCCAGAAGAATCCCCTTGGGG

AGTCAGCTCGTCTGGGTA

(CloneID 6048028)
CCACGCGTCCGCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGT SEQ ID NO:174

CCGACCCTTTCACCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCC

-continued

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGCACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCCGAAAAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAA

CTCTGACTGTCCTGAACGATGTCGGTCTGGCCTGTTTCCTCCATGCTTGGGAACTGCACAAC

TGGAGCCTGACTGTACATCTCTTGGAATTGGTTTCATTAAAAAGAAACACCTTTAAAAAAAA

AAAAAAAAAGGGGGGGCCGCTCCAAAAGGAT

(CloneID 2518214)
GAGCTCAAGGATCCTTAATTAAATTAATTCCCCCCCCCCCCGAGCGGCGGCTCCTGCAGCGG SEQ ID NO:175

TGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCCGACCCTTTGAGCGTTCTG

CTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGNGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCAGTGGCCCCGGTCCCAGTGCCATCCAGAAGAAATCCC

(CloneID 2820804)
GGCACGAGGCCTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTC SEQ ID NO:176

CGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATC

(CloneID 2519633)
AGCGGCGGCTCCTGCAGCGGTGGTCGGTTGTTTGGTGTGGAGTTTCCCAGCGCCCCTCGGGT SEQ ID NO:177

CCGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAACGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACCAGAACAACCT

GTGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTG

GGCCAAGTCAGCAGGTGCCAACTCTAGTGGTGGCAGGGAACACTTGGAGTCATCTGGACTGC

-continued

AGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAACGTGAT

ATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTG

(CloneID 2518293)
AGCGGATCCTTATATTAAATTAATTCNCTCCCCCCCCCCAGAGCCGGCGNNNGCTCCTCCG SEQ ID NO:178

GAGCGTGTGTGTCGGCTTTGTTGGGATGTGCGAGCTTTCCCCAGCTGCACCCTTCGGAGTCC

CGACGCCTTTTGANNGCGTTCTGCATCCCGGCGCCCAGCCCCTACTCTCGCTCCTCTGGCGC

CC

ATGANCCACTATACCACCACCTTCAACTGTGAGTCGACCCCAACNATCCAGGAATAGCTCCC

GAGTTTTCGCGGCCTCCACCTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGA

ACAACCTGTGACCAAGAACAAAATGGCCTCTAATATCTTTGCGACACCTCAAGAAAATCAAG

CTTCTTGCGCCAAGTCAGCACGTGCCAAGTCTAGTGGTGCCAGGGAACACTTGGAGTCATCT

GGACTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGA

AGGTGATATTCATGAAAATCTGGACACAGACTTGCCAGCCAGCCTGNCGCAGAGTGAAGAGA

AGCCCGTCCCTGCTGCGCCTGTGCCCACCCCAGTGGCCCCGGTCCCAGTGCCATCCAGAAGA

CATCCCCCTTGCGGCAAGTCCAGCCTCGTCTTG

(CloneID 2819423)
GGCACGAGCCCTGCAGCGCTGGTCGGCTGTTGGGTGTGCAGTTTCCCAGCGCCCCTCGCGTC SEQ ID NO:179

CGACCCTTTGAGCGTTCTCCTCCGGCGCCAGCCTACCTCGCTCCTCCGCGCC

ATGACCACAACCACCACCTTTAAGGCACTCGACCCCAACAGCAGGAATAGCTACCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGCAAGAACAAAATCGCCCCTAATATCTTTGGGACACCTGAACAAAATCAACCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTCGTGGCAGGGAAGACTTGGAGTCATCTCGACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCC

(CloneID 2819397)
GGCACGAGGCCTGCAGCGGTGGTCGGCTGTTCGGTGTGGAGTTTCCCAGCGCCCCTCGGGTC SEQ ID NO:180

CGACCCTTTGAGCGTTCTGCTCCGGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGCGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGGCCCCTAATATCTTTGGGACACCTGAAGAATCAAAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGCGAAGACTTGGAGTCATCTGCACTGCA

GAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTATATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGGCC

(CloneID 2783324)
CGAGCCGCGGCTCCTGGAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGG SEQ ID NO:181

GTCCGACCCTTTGAGCCTTCTGCTCCCGCGCCAGCCTACCTCGCTCCTCGGCGCC

ATGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAAAACAAAATGGCCTCTAATATCTTTGCGACACCTGAAGAAAATCAAGCTTCTTGG

-continued

GCCAAGTCAGCACGTGCCAAGTCTAGTGCTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

AAGAATGAACTCCTCTGAACCAAGCTCCGCACACTTCTTACATCTGAAGCGACAAGGTGATA

TTCATTAAAATGTGGACACACACTAGCCATCCAGCCTGTCCCATAGTGAATAGAAGCCCGTG

CCTGCTGCTCCTGTGCCTATTTCAGTTCATCCAGCACACGCACCGTCCACAACAAATCCCGC

TGGGGCAACTACAGGCTTGATT

(CloneID 471568 in 3' to 5' orientation)
TTACATAAAGTGCTTCTTTTTAATGAAACAAATCCAACAGATGTACACTCAGGCTCAAGTTG SEQ ID NO:182

TGCAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAG

CTAACCCAAGACCAGGCTGGACTTGCCGCCAGGGGATTTCTTCTGCATGGCACTCGGGCCGG

GGCACCGGGCTTGGGCACAGCCGCANAGGCACGGGCTTCTCTTCACTCTGCCCCAGGCTGCC

TGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAAGTCT

CCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTTGCAGTNCAGATTGACTCCAAGTCTTCCCTG

CCACCANTAGACTTTGGCANCTGCNGACTTTGGNCCAACAAAGTTGAATTTCTTCAAGNGGT

CCCAAAGATTT

(CloneID 471568)
CTGCAGCGGTGGTCGGCTGTTGGGTGTGGAGTTTCCCAGCGCCCCTCGGGTCTCGACCCTTT SEQ ID NO:183

GAGCGTTCTGCTCCGGCGCCACTACCTCGCTCCTCGGCGCC

ATCACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCAGGAATAGCTCCCGAGTTTT

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TTG

(CloneID 4080434)
GACCCCAACAGCAGGAATAGCTCCCGAGTTTTGCGGCCTCCAGGTGGTGGATCCAATTTTTC SEQ ID NO:184

ATTAGGTTTTGATGAACCAACAGAACAACCTGTGAGGAAGAACAAAATGGCCTCTAATATCT

TTGGGACACCTGAAGAAAATCAAGCTTCTTGGGCCAAGTCAGCAGGTGCCAAGTCTAGTGGT

GGCAGGGAAGACTTGGAGTCATCTGGACTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGG

AGACTTCTTAGATCTGAAGGGAGAAGGTGATATTCATGAAAATGTGNNNNNNNNNNNNNNNN

NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

N

NNNNNNNNNNNNNNNAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTCCTTTGGAAGAACAGGTTTCTCTCTGTCCTTGACTCTGGGGTCTGTGGG

CCATGGCATGACTGTTTCTAGTAGTAATGGGAGGGAAAGCTTGGTGACACTTAGTACTGTGT

TTTTAAGACGAAATAATTGGGTTTCCAGATGTGTCACAGGATCTTCGTACTGAGGTTTAACA

CTTTACTGGGGTTTACCAAGCCTCAATGGAAGACATAACAGTCACAGGCCCGTCTGCAGGCC

AACCAAGGGAGTTCTCCGAGACCTCTGTTCTGCCTAATGCCATGGCTTGTCAAAGCCCTCTG

GCTGGC

(CloneID 4748522)
CGAGTTTTGCGGCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAA SEQ ID NO:185

CAACCTGTGAGGAAGAACAAAATGGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGC

TTCTTGGGCCAAGTCACCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGCAGTCATCTG

-continued

GACTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAA

GGTGATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAA

GCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAA

ATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACCCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTATTCCTTTGGAAGAACAGGTTTCTCTCTGTCCTTGACTCTGTGGGTCTGTG

GGCCATGGCATGAGTGTTTCCTAGTAGTAGATGGGAGGGAAACCTTGGTGACACTTAGTACT

GTGTTTTCAAGAAGAAATAATTGGGTCCCAGATGTGTTAGAGGATCTTGGTACTGAGGTTTT

TAACACTTTACTGTGGCTTTACCAAGCCTCAACTGGACAGACCATAAA

(CloneID 5187032)
CTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGGCCAAGTCAGCAGGTGCCAAG SEQ ID NO:186

TCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGACAAGGAACTCCTCTGAAGC

AAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATTCATGAAAATGTGGACACAG

ACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGCCCAGC

CCGCTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTGGCGGCAAGTCCAGCCTCGT

CTTGGGTTAG

CTCTGACTGTCCTCAACGCTGTCCTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTC

TCTTTTTTTTTTTTCCTTTGGAAGAACAGGTTTCTCTCGTGTCCTTGACTCTGGGGTCTCGG

GCCCAGGGCATCAGTGTTCCATAGTAGTAGATGGGAGGGAAAGCTTGGCGACACTTAGTACG

GGGTTTTCAACAAGAAATAATCGGGTTCCAGAGTGTGTAAGAGGACTCTTGGGTACGCAGGT

TAAAACACTTAACTGGGGTATACCAAGCCTCAACGCGACAGACCATAAACGGTCCACAGGCA

CCGTTCCGTGCAGGGCCCAAGCCAAAGGGAGTCTCTCCGAAGAGCCTTCTGGGAGTAGCCCT

AACTTTGCCAGGTGGCCTTGGGTCAAGAGCCTCCTCCTGTGACCTGGTCA

(CloneID 3895688)
AATATCTTGGGACACCTGAAGAAAATCAAGCTTCTTGCGCCAAGTCAGCAGGTGCCAAGTCT SEQ ID NO:187

AGTGGTGGCAGGGAAGACTTGGACTCATCTGGACTGCAGAGAAGGAACTCCTCTGAAGCAAG

CTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTCATATTCATGAAAATGTGGACACAGACT

TGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCTGCTCCGCCTGTGCCCAGCCCG

GTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTT

GGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTCCTTTGGAAGAACAGCTTTCTCTCTGTCCTTGACTCTGGGGTCTGTGGG

CCAGGGCATGAGTGTCTCTAGTAGTAGATCGGAGGGAAAGCTTGGTGACACACTTTACTGGG

GTTTACCAAGCCTCAACTGACAGACCATAAACAGTCCACAGGCACCGTTCCTGGCAGGCCCA

ACCCAAGGGAGTCTCTCCGAGAGCCTTCTGGGTGTGGCCTAACTTGCCAGGGGCCTGGCTCA

GAGCTCCTCTGTGACTGTGACACTGAGCAG

-continued

(CloneID 3938961)

AATCAAGCTTCTTGGGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGCAAGACTTGGAG SEQ ID NO:188

TCATCTGGACTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAA

GGGAGAAGGTGATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTG

AAGAGAAGCCCGTGCCTCCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCC

AGAACAAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTCCTTGACTCTGTGGGTCTGTGGG

CCAGTGGCATGAGTGTTTCTAGTACTAGATGGGAGGGAAAGCTTGGTCACACTTAGTACTGT

GTTTTTAACGAAGAAATAATTGCTTCCAGATGTGTTAGAGGATCTTTCTACTGAGCTTTTTA

ACACTTTACTGTGGGTTTACCAAGCCTCAATGGACAGACCATAAACACTCCACGAGGCACCG

TTCCTGCCAGGCCCAACCCAAGAGGGAGTCTCTCCGAAGAGCCTTCTGTGGTGTTGGCCCTA

ATTTGCCAGGTGGCCTTGGTCAGAGCTCCTCTGTGACTTGTGGACCATGGAAGGGCTGGGCT

CTTGCTTGCGGCTGGAAACGAAAGAATGGCTTTATTTTTAACCTTAAAGTAGCAGATATGTA

ACAAGATGGTGGCGTGTTACAA

(CloneID 4345766)

GAAGAAAATCAAGCTTCTTGGGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGA SEQ ID NO:189

CTTGGAGTCATCTGGACTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAG

ATCTGAAGGGAGAAGGTGATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGG

CAGAGTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTCGCCCCGGCCCCAGT

GCCATCCAGAAGAAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTCTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGACCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTCCTTTTGAAGAACAGGTTTCTCTCTGTCCTTCACTCTGGGGTCTGTGGG

CCATGGCATGAGTGTTTCCTAGTAGTAGATGGGAGGGAAAGCTTTGTGACACTTAGTACTGT

GTTTTTAAGAAGAAATAATTGGGTCCCAGATGTGTCAGAGGATCTTTGGTACTGAGGTTTTT

AACACCTTTACTGTGCGTTTACCAAGCCTCAACTCGACAGACCATAAACAGTCCACAAGGCA

CCGTCCTGCCAGGGCCGAACCACAGGGAGTCTCTCCGCAGAGGCTTCTGGTGTTGCCTACTT

GCCAGTGGCCTTGTCAGAGCCTCTCCTGTGACATGTGAACACACGAAAAGCCTCGGCCACTG

ACTTGAGTGTGAAGA

(CloneID 3889866)

TGGAGTCATCTGGACTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGAT SEQ ID NO:190

CTGAAGGCAGAAGGTGATATTCATGAAAATGTGGACACACACTTGCCAGOCAGCCTGGGGCA

GAGTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGC

CATCCAGAAGAAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTTCTTTGGAAGAACAGGTTTCTCTCTGTCCTTGACTCTGGGGTCTGTGGG

CCATGGCATGAGTGTTTCCTAGTAGTAGATGGGAGGGAAAGCTTGGTGACACTTAGTACTGT

-continued

GTTTTAAGAAGAAATAATTGGGTTCCAGATGTGTCAGAGGATCTTNGTACTGAGGTTTTAAC

ACTTTACTTGGGTTTACCAAGCCTCAATGGACAGACCATAAACAGTCCACCAGGCACCGTTC

CTGCCAGGCCAACCCAGAGGGAGTCTCTCCGCGAGAGCCTTCTGTGGTGGTGGCCCTAACTT

GCAGTGGCTGGTCAGGAGCCCTCCCCTGTGACCTGTGACCCATGAAAGGCTGGGCTCTGCTG

CCGCTGAAGAAAAACGCTCTTATTTTACCTAAAAGTCCATTTACCGACGCGCGGTGACCCAC

TGTCCCCCAGAGGCGGGGTCACAACGCCCGACACACCGGCACGGGGCCCGAGGCGCCCAGGC

CCGTCGAAACATGTGCGCCGAAACACCGGCGCAGGCCA

(CloneID 4450834)
TCTTAGATCTGAAGGGAGAAGGTGATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGC SEQ ID NO:191

CTGGGGCAGAGTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGC

CCCAGTGCCATCCAGAAGAAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTGGGGTCTGTGGG

CCAGTGGCATGAGTGTTTTCTAGTAGTAGATGGGAGGGAAACCTTGGTCACACTTAGTACTG

TGTTTTTAAGAAGAAATAATTGGGTTCCAGATGTGTTAGAGGATCTTTGGTACTGAGGTTTT

TAACACTTTACTGGCGTTTACCAAGCCTCAACTGGACAGACCATAAACAGTCCACAGGCACC

GTTCCTGCCAGGCCCAACCCACAGGGAGTCTCTCCGAGAGCCTTCTGTGGTGTGGCCTAAAT

TGCAGTGGCTTGTGTCAGAGCTCCTCCTGTGACATGTGACATGAAGAGGGCTCGGCTCCTGG

CTTGGCGGCTGAAAGGAAAGAAACTGGCTTTATTTTTTAACCTTTAAAAGGTTGCGGATAGT

AACAAGATTGGGTGGGTGATGAGAAGGCTTGGTCTCAGCAACAGAGGTGGTTCAATGAAATG

CTGGATAAAAGCGTGACAG

(CloneID 5521426)
TTACCAACAGAACAACCTGTGAGGAAGAACAAAATGGCCTCTAATATCTTNTGGGACACNCT SEQ ID NO:192

GAAGAAAATCAAGCTTCTTGGGCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGA

CTTGGAGTCATCTGGACTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAG

ATCTGAAGGGAGAACGTGATATTCATGAAAATGTGGACACAGACTTGCCAGGCAGCCTGGGG

CAGAGTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGCCCCCAGT

GCCATCCAGAAGAAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAC

CTCTGACTGTCCTGAACGCTCTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGG

CCATGGCATGAGTGTTTTCTAGTACTAGATTGGAGGGAAAGCTTTGTGACACTTAGTACTGT

GTTTTTAAGAAGAAATAATTTGGTTCCAGATCTGTTAGAGGATCTTTTCTACTGAGGTTTTT

AACACTTTACTTGGGTTTACCAAGCCTCAACTGGACAGACCATAAACAGTCCACACGCACCG

TTCCTGCCAGGCCCCAACCCACAGGGAGTCTCTCCGCAAGCCCTTCTTGGTGTTGCCCTAAC

TTGCCAGTGGCCCTTTGCTCAGAGCCTCCCTCCTGTGACATGTGGAAACATGAAAGAGGCCC

TGCGCCCTCCCTGGCTTTTCCCGGCTTGCAAAGAAAAGAAACTGGCCCTTTTTATTTTTTTA

ACCCTTTAAAAAGGAGGCCAGAATAAGTAACCAGAACTTGGCCTGGGCTTGATGTAAACCAA

AGCCCTTTGGTTCTCCCACCCCAAAACGAAAAGGCTTTTGGGT

-continued

(CloneID 5267723)
AGCGGGCGTAGCCATTGGATCCGGAGACCGGATCGATCAGAGTTTTGCGGCCTCCAGGTGGT SEQ ID NO:193

GGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTGTGAGGAAGAACAAAAT

GGCCTCTAATATCTTTGGGACACCTGAAGAAAATCAAGCTTCTTGGCCCAAGTCAGCAGGTG

CCAAGTCTAGTGGTGGCAGCGAAGACTTGGAGTCATCTGGACTGCAGAGAAGGAACTCCTCT

CAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATTCATGAAAATGTGGA

CACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGC

CCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTGGCGGCAAGTCCAGC

CTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAACAAAGAAGCACTTTATGTACTGCT

GTCTTTTTTTTTTTCCTTTGGAAGAACAGGTTTCTCTCTGTCCTTGACTCTGTGGGTCTGTC

GGCCATGGCATGAGTGTTTCCTAGTAGTAGAT

(CloneID 4441495)
CAGGTGCCAAGTCTAGTGGTGGCAGCGAAGACTTGGAGTCATCTGGACTGCAGAGAAGGAAC SEQ ID NO:194

TCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATTCATGAAAA

TCTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCTCCTGCGC

CTGTGCCCAGCCCGGTGGCCCCCGCCCCAGTGCCATCCACAAGAAATCCCCCTGGCGGCAAG

TCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGGC

CATGGCATGAGTGTTTTCTAGTAGTAGATTGGAGGGAAAGCTTTGTGACACTTAGTACTGTG

TTTTTAAGAGAAATAATTTGGTTCCAGATGTGTTAGAGGATCTTTGTACTGAGGTTTTTTAA

CACTTTACTTGGGTTTACCAAGCCTCAATCTGGACAGACCATAAAACAGTCCACAGGCACCN

GTTCCTTGCCAGGGCCCCAACCCACAGGGAGTCTCTCCCGGCAGAGCCTTTCTTGGTGTTTG

CCTTAAATTGGCAAGTCGGCCTTGGTCAAAGCCTCTTCCTGTGCATGTGAACACTGGAAAGA

GGCTGGGGCCTCCTGCCGTTGGCGGCTGGCAAAGAAAAAAGCCGTTTTTTTTACCTTAAAGT

GCCCAAGTTAACAACGGGTGGTGTTGGAAACCTTGTTTCCCCAGAAGGGGGTGTGAAAAGGG

GGGTACACAGGAACAGGGGAAGCACAGGTCTCTCAACAGAGCACG

(CloneID 5395070)
TGACCACAACCACCACCTTCAAGGGAGTCGACCCCAACAGCGAGGAATAGCTCCCGAGTTTT SEQ ID NO:195

GCGGCCTCCAGGTGGTGGATCCAATTTTTCATTAGGTTTTGATGAACCAACAGAACAACCTG

TGAGGAAGAACAAAATGCCCTCTAATATCTTTGGGACACCTCAAGAAAATCAAGCTTCTTGG

GCCAAGTCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCA

GAGAAGGAACTCCTCTCAAGCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATA

TTCATGAAAATGTGGACACAGACTTGCCAGGCACCCTGGGGCAGAGTGAAGAGAAGCCCGTG

CCTGCTGCGCCTGTGCCCAGCCCGGTGGCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCC

TGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTACTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

-continued

TGAGCCTGACTGTACATCTCTTGGATTCGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTATTCTTTCCCTTTTGAAGACCAGGTTTCTCTCTGTCCTTGACTCTGGGCTCTGTGGG

CCATTGGCATGCAGTTGTCTCCTAGTACGTAGCATGGGCGGGAAACGCTTTGTGACCCCTTA

CGTACTCCGCTTCCAAGAACGACATCATGGGGTCGCAGATTGTCGTAGAGGATCTTTGGCTA

CCTGCCGCGTTCTTCACCGCCTTTTAACTGTGTGGTTATCCCAAGCCCTCCACTTGGACAGA

ACCATTAACCATGCTCCACAGCGCACCGTCACTTGCAGGCCCAACCCCCAGGGGGGTCTTCC

GAGAGACCTTTTGGTGTGCCTATTTGCCCAGGGCGTTGTCAACCCCCG

(CloneID 3445695)
GCAGAGTGAAGAGAAGCCCGTGCCTGCTGCGCCTGTGCCCAGCCCGGTCGCCCCGGCCCCTA SEQ ID NO:196

GTGCCATCCAGAAGAAATCCCCCTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTCCTTTGGAAGAACAGGTTTCTCTCTGTCCTTGACTCTGGGGTCTGTGCGC

CATGGCATGAGTGTTTTCTAGTACTAGATTGGCGGGAAAGCTTTGTGACACTTAGTACTGTG

TTTTAAGAAGAAATAATTGGGTCCCAGATGTGTTAGAGGATCTTTGGTACTGAGGTTTTTAA

CACTTTACTTGGGTTTACCAAGCCTCAACTGGACAGACCATAAACAGTCCACAGGCACCGTT

CCTGCCAGGCCCCAACCCACACGGAGTCTCTCCGCAGAGCCTTCTTGGTGTTGCCCTAACTT

CCCAGTGGCCTTGGCTCAGAGCCTCCTCCTGTGACCTGTGAACAATGAAGAGGCCTGTGCCT

CTGCCTTGCCGCTGCAGGCACAGAACTGCCTTTACTTTTTAACCTTAAAAAGTAGCCGATAG

TACAAGCTGGTTGGCTGTTGAAGCCAAGCCTGGGTCTCACGCCGAGTAAGGCTGCCTGTTCC

CTGCACACTGGCGGGACCAAAAGCCGTGAACCAGGGGGCCATCCACTGTAAGCGGGTCTCCC

GGAACGGGCCCCAAGGGTGTTGTTACACTACCCGCGGGGGTGCATAGCTGTGGCTGCCCTCA

CGAAGCTCCTAACAGGTTTGCTCACCGGCGGGGCCTGAAAATC

(CloneID 5519819)
TCAGCAGGTGCCAAGTCTAGTGGTGGCAGGGAAGACTTGGAGTCATCTGGACTGCAGAGAAG SEQ ID NO:197

GAACTCCTCTGAACCAAGCTCCGGAGACTTCTTAGATCTGAAGGGAGAAGGTGATATTCATG

AAAATGTGGACACAGACTTGCCAGGCAGCCTGGGGCAGAGTGAAGAGAAGCCCGTGCCTGCT

GCGCCTGTGCCCAGCCCGGTGCCCCCGGCCCCAGTGCCATCCAGAAGAAATCCCCCTGGCGG

CAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACGCTGTCGTTCTGTCTGTTTCCTCCATGCTTGTGAACTGCACAACT

TGAGCCTGACTGTACATCTCTTGGATTTGTTTCATTAAAAAGAAGCACTTTATGTACTGCTG

TCTTTTTTTTTTTTCTTTTGAAGAACAGGTTTCTCTCTGTCCTTGACTCTTGGGTCTGTGGG

CCATGGCATGAGTGTTTTCTAGTAGTAGATTGGAGGGAAAGCTTTGTGACACTTAGTACTGT

GTTTTTAAGAAGACATAATTTGGTTCCAGATGTGTTAGAGGATCTTTTGTACTGAGGTTTTT

AACACTTTACCTTGGGTTTACCAAGCCTTCCACTGGACAGGACCATAAACAGGTCCACAGGC

CCCCCGTTTCTTGCCCAGGCCCCAACCCACAGGGGAAGTCTCTTCCGCCACAAACCCTTCCT

TTGGTCGCCCGCCCTTACCACCTGCCCCCGGCGGGGCTCTTTATGCTTTGGCGGAGCTCTCT

CCTCTCCTTCATTCGAAGACATGCTGACCAACAACCACGGCCACACCAACGCCCCTCTCTAT

CTTTCTCATTGCGCCCCTGCGCCCTGTATTGAGCGCGCCCTCAGAATTCGCGGGGCATACAG

-continued

GGCGCGCTGTCAATCCCCGAAACTCCAATATCGGTCTCCCCCCCGCTTTATATTAGCGCGCC

CCGTCGCTTGTAACGAGCCTAACTCAATTTTTTTTTCCCGCACTCACGCCCCGCTCTCGGTG

TTACCAACCAACAGAACAGTCCTCCTCCTTCATAATATTCCCCACACTCCCCCTCGCCATCT

TACTCTTTTATTTNCCCCGCGCGCGTGTTAAAAGCCGCAAAGGNNCCCCCACCGCACACACG

TATATTGCCCGGATTTTAATCATGAATTTTCCGCACCGGGCAAGACCACCCCCAGCTGCCCC

ATCGTGTCCCCCGCCGTNCTGCAGAATTCCTCANATCTGGGCCCTCCGCGATAGGTCACTCC

ATCCTCAGGAGAAAACACCCTCTTGTCCCACGCCGTCCACTCTTAATACCGCCGTACCCTCA

TGTCCCCTCCGCCACTTCTTTTTACCCCCTCNGCGCTGTAGTTTTGCGGCGCCCCGACTACA

CTCACACCGATTATTTATTATCCCCCTTCANCCGACGCCGGNNGCCCCGAAAAAATGCCGCC

CCGGATCGATAACAGCTCACCCCCGCCCCCGCCAGGATGCCCCGCCGCCAGCTAAAAGGCGC

TCAACCA

(CloneID 2989794)

CACGAGCACTGATGCCCAAGAGTTAAATGGCCTCTAATATCTTTGGGACACCATGAAGAAAA SEQ ID NO:198

TCAAGACTTCTTGGGCCAAGTCAACAGGTGCCAAGTCTAGTGGTGGAAGGGAAGACTTGTGA

GTCATCTGGACTGCAGAGAAGGAACTCCTCTGAAGCAAGCTCCGGAGACTTCTTAGATCTGA

AGGGAGAAGGTGATATTCATGAAAATGTGAGACACAGACTTGCCAGTGCAGCCTGGCGCAGA

GTGAAGAGAAGCCCTGTGCCTGCTGCTGCCTGTGCCCAGCCCTGGTGGACCCAGGCCCCAGT

GCCATCCAGAAGAAATCCCCACTGGCGGCAAGTCCAGCCTCGTCTTGGGTTAG

CTCTGACTGTCCTGAACTGCTAGTCAGATATCTAGTCTGTAACCTCCATGCTTGTGAACTGC

ACAACTTGAGCCTGACTGTACATCTCTTGGAATTTGTTTCATTAAAAAAAAGCACTTATATG

TACTGCTGTCTTTATTATTAATATACTTTGTGCAAGAACAGGTAAACACATCTGACACATTG

ACCTCTGTGGACTCTGTGAGAGCCATAGGCATTGAGTTGTGAACTAGCTAGATACATGGCGA

GGGAACAGCTTTGTGAACACTAAGAACTCAGATCATAACGACAAAATACATGGGTCCAGATG

TGTAGAGGATCTAAGGAAATGAGGTATAAAACTAACTTGGTTTCCAAGCTCAATGGGCGACA

(CloneID 1637898 in 3' to 5' orientation)

AAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGCTCAAGTTGTGCAGT SEQ ID NO:199

TCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAG

CTAACCCAAGACCAGGCTGGACTTGCCGCCAGGGGGATTTCTTCTGGATGGCACTGGGGCCG

GGGCCACCGGCCTGGGCACAGGCGCAGCAGGCACGGCCTTCTCTTCACTCTGCCCCAGGCTG

CCTGGCAAGTCTGTGTCCACATTTTCATCAATATCACCTTCTCCCTTCAGATCTAAGAAGTC

TCCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCTTCCCTGC

CACCACTAGACTTGGCACCTGCTGACTTGGCCCAAGAAGCTTGATTTTCTTCAGGTGTCCCA

AAGATATTAGAGGCCATTTTGTTCTTCCTCACAGGTTGTTCTGTTGGTTCATCAAAACCTAA

TGAAAAATTGGATCCACCACCTGGAGGCCGCAAAACTCGGGAGCTATTCCTGCTGTTGGGGT

CGACTCCCTTGAAGGTGGT

(CloneID 1132042 in 3' to 5' orientation)

TTTTTTTAAAGTGCTTCTTTTTAATGAAACAAATCCAAGAGATGTACAGTCAGGCTCAAGTT SEQ ID NO:200

GTGCAGTTCACAAGCATGGAGGAAACAGACAGAACGACAGCGTTCAGGACAGTCAGAG

CTAACCCAAGACGAGGCTGGACTTCCCGCCAGGGAGATTTCTTCTGGATGGCACTGGGGCCG

GGGCACCGGGCTGGGCACAGGCGCACGAGGCACGGGCTTCTCTTCACTCTGCCCCACGCTGC

-continued

CTGGCAAGTCTGTGTCCACATTTTCATGAATATCACCTTCTCCCTTCAGATCTAAGAAGTCT

CCGGAGCTTGCTTCAGAGGAGTTCCTTCTCTGCAGTCCAGATGACTCCAAGTCT

[0054] In another set of preferred embodiments of the invention, the sequence, or a unique portion thereof, of the ESE-2/ELF5 sequences identified by the I.M.A.G.E. Consortium CloneIDs 1864302 and 4480123 as well as clusters C019657 and NM_001422 may be used. Similarly, sequences from the UniGene Homo sapiens cluster Hs.11713, may be used. ESE-2/ELF5 sequences are useful in discriminations between normal and ADH or DCIS or IDC breast cells based upon decreased expression in non-normal breast cells. Human ESE-2/ELF5 has been mapped to 11p13-p15.

[0055] ESE-2/ELF5 related sequences are as follows, with the assigned coding region (ending with a termination codon) underlined and preceded by the 5' untranslated and/or non-coding region and followed by the 3' untranslated and/or non-coding region. The 3' untranslated and/or non-coding regions are presented as the complementary strand in the 3' to 5' direction for some of the sequences.

[0056] SEQ ID NOS:204-209 are found in the 3' untranslated and/or non-coding regions of the disclosed ESE-2/ELF5 sequences.

(CloneID 4480123, ESE-2b/ELF5)

CACAAGGCTACAGGTGTCTTTATTTCCACTGCACGCTGGTGCTGGGAGCGCCTGCCTTCTCT SEQ ID NO:201

TGCCTTGAAAGCCTCCTCTTTGGACCTAGCCACCGCTGCCCTCACGGTA

ATGTTGGACTCGGTGACACACAGCACCTTCCTGCCTAATGCATCCTTCTGCGATCCCCTGAT

GTCGTGGACTGATCTGTTCAGCAATGAAGAGTACTACCCTGCCTTTGAGCATCAGACAGCCT

GTGACTCATACTGGACATCAGTCCACCCTGAATACTGGACTAAGCGCCATGTGTGGGAGTGG

CTCCAGTTCTGCTGCGACCAGTACAAGTTGGACACCAATTGCATCTCCTTCTGCAACTTCAA

CATCAGTGGCCTGCAGCTGTGCAGCATGACACAGGAGGAGTTCGTCGAGGCAGCTGGCCTCT

GCGGCGAGTACCTGTACTTCATCCTCCAGAACATCCGCACACAAGGTTACTCCTTTTTTAAT

GACGCTGAAGAAAGCAAGGCCACCATCAAAGACTATGCTGATTCCAACTGCTTCAAAACAAG

TGGCATCAAAAGTCAAGACTGTCACAGTCATAGTAGAACAAGCCTCCAAAGTTCTCATCTAT

GGGAATTTGTACGAGACCTGCTTCTATCTCCTGAAGAAAACTGTGGCATTCTGGAATGGGAA

GATAGGGAACAAGGAATTTTTCGGGTGGTTAAATCGGAAGCCCTGGCAAAGATGTGGGGACA

AAGGAAGAAAAATGACAGAATGACATATGAAAAGTTGAGCAGAGCCCTGAGATACTACTATA

AAACAGGAATTTTGGAGCGGGTTGACCGAAGGTTAGTGTACAAATTTGGAAAAAATGCACAC

GGGTGGCAGGAAGACAAGCTATGA

TCTGCTCCAGGCATCAAGCTCATTTTATGGATTTCTGTCTTTTAAAACAATCAGATTGCAAT

AGACATTCGAAAGGCTTCATTTTCTTCTCTTTTTTTTTTAACCTGCAAACATGCTGATAAAA

TTTCTCCACATCTCAGCTTACATTTGGATTCAGAGTTGTTGTCTACGGAGGGTGAGAGCAGA

AACTCTTAAGAAATCCTTTCTTCTCCCTAAGGGGATGAGGGGATGATCTTTTGTGGTGTCTT

GATCAAACTTTATTTTCCTAGAGTTGTGGAATGACAACAGCCCATGCCATTGATGCTGATCA

GAGAAAAACTATTCAATTCTGCCATTAGAGACACATCCAATGCTCCCATCCCAAAGGTTCAA

AAGTTTTCAAATAACTGTGGCAGCTCACCAAAGGTGGGGAAAGCATGATTAGTTTGCAGGT

TATGGTAGGAGAGGGTGAGATATAAGACATACATACTTTAGATTTTAAATTATTAAAGTCAA

AAATCCATAGAAAAGTATCCCTTTTTTTTTTTTTTGAGACGGGTTCTCACTATGTTGCCCAG

GGCTGGTCTTGAACTCCTATGCTCAAGTGATCCTCCCACCTCGGCCTCCCAAAGTACTGTGA

TTACAAGCGTGAGCCACGGCACCTGGGCAGAAAAGTATCTTAATTAATGAAAGAGCTAAGCC

ATCAAGCTGGGACTTAATTGGATTTAACATAGGTTCACAGAAAGTTTCCTAACCAGAGCATC

-continued

TTTTTGACCACTCAGCAAAACTTCCACAGACATCCTTCTGGACTTAAACACTTAACATTAAC

CACATTATTAATTGTTGCTGAGTTTATTCCCCCTTCTAACTGATGGCTGGCATCTGATATGC

AGAGTTAGTCAACAGACACTGGCATCAATTACAAAATCACTGCTGTTTCTGTGATTCAAGCT

GTCAACACAATAAAATCGAAATTCATTGATTCCATCTCTGGTCCAGATGTTAAACGTTTATA

AAACCGGAAATGTCCTAACAACTCTGTAATGGCAAATTAAATTGTGTGTCTTTTTTGTTTTG

TCTTTCTACCTGATGTGTATTCAAGTGCTATAACACGTATTTCCTTGACAAAAATAGTGACA

GTGAATTCACACTAATAAATGTTCATAGGTTAAAGTCTGCACTGACATTTTCTCATCAATCA

CTGGTATGTAAGTTATCAGTGACTGACAGCTAGGTGGACTGCCCCTAGGACTTCTGTTTCAC

CAGAGCAGGAATCAAGTGGTGAGGCACTGAATCGCTGTACAGGCTGAAGACCTCCTTATTAG

AGTTGAACTTCAAAGTAACTTGTTTTAAAAAATGTGAATTACTGTAAAATAATCTATTTTGG

ATTCATGTGTTTTCCAGGTGGATATAGTTTGTAAACAATGTGAATAAAGTATTTAACATGTA

AAAAAAAAAAAAAA

(ESE-2a)

GGCTGAGTGGTTTGCTCCTTCCCCTCTCTCTGGGAGGCTGAGCAGGGGTGCCTGGGTTGCTC SEQ ID NO:202

AGGCCATGGGAGCCACACCTGTTATTGCTGCCTCTGATTTGTGTGACACTGAGAAGCCCACA

GGCCTGTCCCTCCAACTCGGTGGACCCTCTCTGTGTGCATTTGGTGTGTGAGCCAGCTCTGA

GAAGGGTTCAGAAGCCACTGGAGGCATCTGGGGACCTCAGCTTCC

ATGCCATCTCTGCCTCACTCCCACAGGGTAATGTTGGACTCGGTGACACACAGCACCTTCCT

GCCTAATGCATCCTTCTGCGATCCCCTGATGTCGTGGACTGATCTGTTCAGCAATGAAGAGT

ACTACCCTGCCTTTGAGCATCAGACAGCCTGTGACTCATACTGGACATCAGTCCACCCTGAA

TACTGGACTAAGCGCCATGTGTGGGAGTGGCTCCAGTTCTGCTGCGACCAGTACAAGTTGGA

CACCAATTGCATCTCCTTCTGCAACTTCAACATCAGTGGCCTGCAGCTGTGCAGCATGACAC

AGGAGGAGTTCGTCGAGGCAGCTGGCCTCTGCGGCGAGTACCTGTACTTCATCCTCCAGAAC

ATCCGCACACAAGGTTACTCCTTTTTTAATGACGCTGAAGAAAGCAAGGCCACCATCAAAGA

CTATGCTGATTCCAACTGCTTGAAAACAAGTGGCATCAAAAGTCAAGACTGTCACAGTCATA

GTAGAACAAGCCTCCAAAGTTCTCATCTATGGGAATTTGTACGAGACCTGCTTCTATCTCCT

GAAGAAAACTGTGGCATTCTGGAATGGGAAGATAGGGAACAAGGAATTTTTCGGGTGGTTAA

ATCGGAAGCCCTGGCAAAGATGTGGGGACAAAGGAAGAAAAATGACAGAATGACGTATGAAA

AGTTGAGCAGAGCCCTGAGATACTACTATAAAACAGGAATTTTGGAGCGGGTTGACCGAAGG

TTAATGTACAAATTTGGAAAAAATGCACACGGGTGGCAGGAAGACAAGCTATGA

TCTGCTCCAGGCATCAAGCTCATTTTATGGATTTCTGTCTTTTAAAACAATCAGATTGCAAT

AGACATTCGAAAGGCTTCATTTTCTTCTCTTTTTTTTTAACCTGCAAACATGCTGATAAAAT

TTCTCCACATCTCAGCTTACATTTGGATTCAGAGTTGTTGTCTACGGAGGGTGAGAGCAGAA

ACTCTTAAGAAATCCTTTCTTCTCCCTAAGGGGATGAGGGGATGATCTTTTGTGGTGTCTTG

ATCAAACTTTATTTTCCTAGAGTTGTGGAATGACAACAGCCCATGCCATTGATGCTGATCAG

AGAAAAACTATTCAATTCTGCCATTAGAGACACATCCAATGCTCCCATCCCAAAGGTTCAAA

AGTTTTCAAATAACTGTGGCAGCTCACCAAAGGTGGGGGAAAGCATGATTAGTTTGCAGGTT

ATGGTAGGAGAGGGTGAGATATAAGACATACATACTTTAGATTTTAAATTATTAAAGTCAAA

AATCCATAGAAAAGTATCCCTTTTTTTTTTGAGACGGGTTCTCACTATGTTGCCCAGGGCTG

-continued

GTCTTGAACTCCTATGCTCAAGTGATCCTCCCACCTCGGCCTCCCAAAGTACTGTGATTACA

AGCGTGAGCCACGGCACCTGGGCAGAAAAGTATCTTAATTAATGAAAGAGCTAAGCCATCAA

GCTGGGACTTAATTGGATTTAACATAGGTTCACAGAAAGTTTCCTAACCAGAGCATCTTTTT

GACCACTCAGCAAAACTTCCACAGACATCCTTCTGGACTTAAACACTTAACATTAACCACAT

TATTAATTGTTGCTGAGTTTATTCCCCCTTCTAACTGATGGCTGGCATCTGATATGCAGAGT

TAGTCAACAGACACTGGCATCAATTACAAAATCACTGCTGTTTCTGTGATTCAAGCTGTCAA

CACAATAAAATCGAAATTCATTGATTCCATCTCTGGTCCAGATGTTAAACGTTTATAAAACC

GGAAATGTCCTAACAACTCTGTAATGGCAAATTAAATTGTGTGTCTTTTTTGTTTTGTCTTT

CTACCTGATGTGTATTCAAGCGCTATAACACGTATTTCCTTGACAAAAATAGTGACAGTGAA

TTCACACTAATAAATGTTCATAGGTTAAAGTCTGCACTGACATTTTCTCATCAATCACTGGT

ATGTAAGTTATCAGTGACTGACAGCTAGGTGGACTGCCCCTAGGACTTCTGTTTCACCAGAG

CAGGAATCAAGTGGTGAGGCACTGAATCGCTGTACAGGCTGAAGACCTCCTTATTAGAGTTG

AACTTCAAAGTAACTTGTTTTAAAAAATGTGAATTACTGTAAAATAATCTATTTTGGATTCA

TGTGTTTTCCAGGTGGATATAGTTTGTAAACAATGTGAATAAAGTATTTAACATGCGAA

(alternative ELF5 encoding sequence)
CAAGGCTACAGGTGTCTTTATTTCCACTGCACGCTGGTGCTGGGAGCGCCTGCCTTCTCTTG SEQ ID NO:203

CCTTGAAAGCCTCCTCTTTGGACCTAGCCACCGCTGCCCTCACGGTA

ATGTTGGACTCGGTGACACACAGCACCTTCCTGCCTAATGCATCCTTCTGCGATCCCCTGAT

GTCGTGGACTGATCTGTTCAGCAATGAAGAGTACTACCCTGCCTTTGAGCATCAGACAGCCT

GTGACTCATACTGGACATCAGTCCACCCTGAATACTGGACTAAGCGCCATGTGTGGGAGTGG

CTCCAGTTCTGCTGCGACCAGTACAAGTTGGACACCAATTGCATCTCCTTCTGCAACTTCAA

CATCAGTGGCCTGCAGCTGTGCAGCATGACACAGGAGGAGTTCGTCGAGGCAGCTGGCCTCT

GCGGCGAGTACCTGTACTTCATCCTCCAGAACATCCGCACACAAGGTTACTCCTTTTTTAAT

GACGCTGAAGAAAGCAAGGCCACCATCAAAGACTATGCTGATTCCAACTGCTTGAAAACAAG

TGGCATCAAAAGTCAAGACTGTCACAGTCATAGTAGAACAAGCCTCCAAAGTTCTCATCTAT

GGGAATTTGTACGAGACCTGCTTCTATCTCCTGAAGAAAACTGTGGCATTCTGGAATGGGAA

GATAGGGAACAAGGAATTTTTCGGGTGGTTAAATCGGAAGCCCTGGCAAAGATGTGGGGACA

AAGGAAGAAAAATGACAGAATGACATATGAAAAGTTGAGCAGAGCCCTGAGATACTACTATA

AAACAGGAATTTTGGAGCGGGTTGACCGAAGGTTAGTGTACAAATTTGGAAAAAATGCACAC

GGGTGGCAGGAAGACAAGCTATGA

TCTGCTCCAGGCATCAAGCTCATTTTATGGATTTCTGTCTTTTAAAACAATCAGATTGCAAT

AGACATTCGAAAGGCTTCATTTTCTTCTCTTTTTTTTTAACCTGCAAACATGCTGATAAAAT

TTCTCCACATCTCAGCTTACATTTGGATTCAGAGTTGTTGTCTACGGAGGGTGAGAGCAGAA

ACTCTTAAGAAATCCTTTCTTCTCCCTAAGGGGATGAGGGGATGATCTTTTGTGGTGTCTTG

ATCAAACTTTATTTTCCTAGAGTTGTGGAATGACAACAGCCCATGCCATTGATGCTGATCAG

AGAAAAACTATTCAATTCTGCCATTAGAGACACATCCAATGCTCCCATCCCAAAGGTTCAAA

AGTTTTCAAATAACTGTGGCAGCTCACCAAAGGTGGGGGAAAGCATGATTAGTTTGCAGGTT

ATGGTAGGAGAGGGTGAGATATAAGACATACATACTTTAGATTTTAAATTATTAAAGTCAAA

AATCCATAGAAAAGTATCCCTTTTTTTTTTGAGACGGGTTCTCACTATGTTGCCCAGGGCTG

-continued

GTCTTGAACTCCTATGCTCAAGTGATCCTCCCACCTCGGCCTCCCAAAGTACTGTGATTACA

AGCGTGAGCCACGGCACCTGGGCAGAAAAGTATCTTAATTAATGAAAGAGCTAAGCCATCAA

GCTGGGACTTAATTGGATTTAACATAGGTTCACAGAAAGTTTCCTAACCAGAGCATCTTTTT

GACCACTCAGCAAAACTTCCACAGACATCCTTCTGGACTTAAACACTTAACATTAACCACAT

TATTAATTGTTGCTGAGTTTATTCCCCCTTCTAACTGATGGCTGGCATCTGATATGCAGAGT

TAGTCAACAGACACTGGCATCAATTACAAAATCACTGCTGTTTCTGTGATTCAAGCTGTCAA

CACAATAAAATCGAAATTCATTGATTCCATCTCTGGTCCAGATGTTAAACGTTTATAAAACC

GGAAATGTCCTAACAACTCTGTAATGGCA

(CloneID 377520)
CAGACACTGGCATCAATTACAAAATCACTGCTGTTTCTGTGATTCAAGCTGTCAACACAATA SEQ ID NO:204

AAATCGAAATTCATTGATTCCATCTCTGGTCCAGATGTTAAACGTTTATAAAACCGGAAATG

TCCTAACAACTCTGTAATGGCAAATTAAATTGTGTGTCTTTTTTGTTTTGTCTTTCTACCTG

ATGTGTATTCAAGCGCTATAACACGTATTTCCTTGACAAAAATAGTGACAGTGAATTCACAC

TAATAAATGTTCATAAGGTTAAAGTCTGCACTGACATTTTCTCATCAATCACTGGTATGTAA

GTTATCAGTGACNGACAGCTAAGGTGNGACTGG

(CloneID 377520 in 3' to 5' orientation)
GAACATGTTAAATACTTTATTCACATTGTTTACAAACTATATCCACCTGGAAAACACATGAA SEQ ID NO:205

TCCAAAATAGATTATTTTACAGTAATTCACATTTTTTAAAACAAGTTACTTTGAAGTTCAAC

TCTAATAAGGAGGTCTTCAGCCTGTACAGCGATTCAGTGCCTCACCACTTGATTCCTGCTCT

GGTGAAACAGAAGTCCTAGGGGCAGTCCACCTAG

(CloneID 81671)
TGATTCAAGCTGTCAACACAATAAAATCGAAATTCATTGATTCCATCTCTGGTCCAGATGTT SEQ ID NO:206

AAACGTTTATAAAACCGGAAATGTCCTAACAACTCTGTAATGGCAAATTAAATTGTGTGTCT

TTTTTGTTTTGTCTTTCTACCTGATGTGTATTCAAGCGCTATAACAC

(CloneID 81671 in 3' to 5' orientation)
GCGGTGAAATACTTTATTCACATTGTTTACAAACTATATCCACCTGGAAAACACATGAATCA SEQ ID NO:207

AAATAGATTATTTTACAGTAATNACATTTTTTAAAACAAGTTACTTTGAAGTCAACTCTAAT

AAGGAGGTCTTCAGCCTGTACAGCGATTCAGTGCCTCACCACTTGATTCCTGCTCTGGTGAA

CAGAAGTCCTAGGGGCAATCCACCTAGCTGTCAGTCACTGATAACTTACATACCAGTGATTG

ATGAGAAAATGTCAGTGCAGACTTTAACCTATGAACATTTATTAGTGTGAATTCACTGTCAC

TATTTTTGTCAAGGAAATACGTGTTATAGCGCTTGAATACACATCAGGTAGAAAGACAAANC

AAAAAAGNCACACAATTTAATTTGCCATTACAGAGTTGTTAGGGCCATTTCCGGTTTTATAA

ACGTTTACATCTGGGCCNGAGATGGGATCAATGAATTTCGNTTTTATTGTGTTGACCGCTTG

NATCACCTCGTGCCGATTTCCTGCGGCCCGGGGNTCCCCTAGTT

(CloneID 2956497 in 3' to 5' orientation)
AGACGGGTTCTCACTATGTTGCCCAGGGCTGGTCTTGAACTCCTATGCTCAAGTGATCCTCC SEQ ID NO:208

CACCTCGGCCTCCCAAAGTACTGTGATTACAAGCGTGAGCCACGGCACCTGGGCAGAAAAGT

ATCTTAATTAATGAAAGAGCTAAGCCATCAAGCTGGGACTTAATTGGATTTAACATAGGTTC

ACAGAAAGTTTCCTAACCAGAGCATCTTTTTGACCACTCAGCAAAACTTCCACAGACATCCT

TCTGGACTTAAACACTTAACATTAACCACATTATTAATTGTTGCTGAGTTTATTCCCCCTTC

TAACTGATGGCTGGCATCTGATATGCAGAGTTAGTCAACAGACACTGGCATCAATTACAAAA

-continued

TCACTGCTGTTTCTGTGATTCAAGCTGTCAACACAATAAAATCGAAATTCATTGATTCCATC

TCTGGTCCAGATGTTAAACGTTTATAAAACCGG

(CloneID 1864302 in 3' to 5' orientation)

TGAACATGTTAAATACTTTATTCACATTGTTTACAAACTATATCCACCTGGAAAACACATGA SEQ ID NO:209

ATCCAAAATAGATTATTTTACAGTAATTCACATTTTTTAAAACAAGTTACTTTGAAGTTCAA

CTCTAATAAGGAGGTCTTCAGCCTGTACAGCGATTCAGTGCCTCACCACTTGATTCCTGCTC

TGGTGAAACAGAAGTCCTAGGGGCAGTCCACCTAGCTGTCAGTCACTGATAACTTACATACC

AGTGATTGATGAGAAAATGTCAGTGCAGACTTTAACCTATGAACATTTATTAGTGTGAATTC

ACTGTCACTATTTTTGTCAAGGAAATACGTGTTATAGCACTTGAATACACATCAGGTAGAAA

GACAAAACAAAAAAGACACACAATTTAATT

(ESE-2 promoter 1, positions 1—1046, and 5' untranslated region
of ESE-2b,
positions 1047 to >1106)

TTCAATCCCACTTCCTCCTTTTGCCACTGGGGAAAAATGAAGCCCAGAGAAGTCCAGGTTAC SEQ ID NO:210

CCAGCCAAGATAGGGGTCTAGGCAGTCACCTATCTTCCTTCAACAATTTTTCCATGATGACA

CCCCGAATCCTCTCACGAACTGCCTTTTGTTTCTGCAGACTATATGCGTATTTGTATTTGTG

TGTTTGTAACCATGCCCGGTGATCCCAAGAAAACATAACAAACCAGTCCTTGCTTTCGCTTT

AAAGCTTGGAGTCTGCCATTTGAATACAACATCTCGGCTGCCCAAGATGGCTAGAAGCAGAA

TGCAAAAAGGCACAAGGGTTATAAATACCTGTCTCATAGATGACCCGGGACACTTGTGCTTT

GCAGCCTAAATTAGGCAGAGTTTCTGTTGTCACGGAGAAGTACTAAAAAGCGGGCAGTTCTC

AGCGAGACACCTTGAGAGGCTGGCATCCACATGAGGAGAGGCCCCATCACTTACATTACACT

CATGAAGCCCAAGAGATGTTAAGCTACATTTTTCTAGGTAGCACAGCTAGGAAATGATGGAC

ACTGAATTTTGAATCTAAAGGCAACTAGCTGCAAAACCTGAGGTCTCAACACTGGGCCTATA

ATGCTCTTCCTTCTCACCACATGGGAAAAACGGAGAAAGACATGATCTTACAAAAGCACTGG

GCTCAGCCAGCCTGGGAGAGAGGCAGCGCAAACCGGCACCAGGGTGGAATCTCTGTGTTTCC

CCTCCAGCTTGCACAGAGGAAAATTCCTCACCAAACGCACGATCCCCGCACAGCTTCCCAGG

GACAAGAATTTTCTGCTTGTTTACTGAGTCCCCTGGGCTGGGAGTGGGGGGTTTGCAGGGCA

GGGGTGAGCTGCGCACAAAAGCAGGATAAAGGTAAACTTTCTGCATATGAGAACCATTTCCC

CCCCTCCATAGGAGCCGTGTCACACTGTATGTCACCGTCATCAAAGGGGCTGTGCGTAAACC

TGAAAAACCAAACGGACCTGTCTGTAGGTGTCACTTATATCACAAGGTACAGGTGTCTTTAT

TTCCACTGCACGCTGGTGCTGGGAGCGCCTGCCTTCTCTTGCCTTGAAAGCC

(ESE-2 promoter 2, positions 1—1879, and 5' untranslated region
of ESE-2a,
positions 1080 to >1975)

GTAAGCGCCTTTTCTTTGCTTTTTTGGATGAATCGTAAGGGGTGTCAGGGCTCAGAGGCTCG SEQ ID NO:211

GCCTGGGGCCTGGGTCTGCTCGTCCATCTCCAGCTCCAGGGCCGGCACCTGCGCTGGGTGCA

GTCCCAGGGCCTCCGGGAGCTGTTTGACACGGATGTGACGGCTTGCCTCAGAGAGACCTGCT

TGCCTCCGCAGGCCGTGTGGCCACCAGCCCTCCTCACAGCCACCACTCCCTCTGGATCCCTT

AGCTGAACCAAGGGGAATTAAGAGAACGCGCCCAGGCGCTCAGTTCTCTATCCTCCCTAAAG

CTTCTTTAGGAAGGGAAGCTGGATTTTACTCAGAAAAATATTCAGCTTCAAGGAAAGCCCCC

CACCTACCCTTCTCCCCTGACTGAATTCCTTTCCCCACCGGCCTGCCCACTGTCCCCCACCC

CCTCGATACAAAGAACTCCCTTATTATAGTTTGGGGCATTTTCAGGAATTTGTACAAAACAG

-continued

TTTTAAAAGAAACCCATTTGCCCTCTGACATCTTCCCCAGAAGAGAATGCTTCCAGAAAAGC

AATGGTTAGCGTTTTTCCACTGTGTTCTCAGGCCTTGTCAGACTCTGCAGTGTGTGGGGTTT

CTGTGGGGTTGCCCCCTGCTGACAGCGACAGGTAGCACAAGCTCCCTCCTCAGGCTGAAGCC

GGGCCTGCGTGTGGATGCCTTTTGTACGCCCGAAAGTGTGTCAGACATCACTAGGTGCAATC

GCTTCATGTTACAGATGAGGAAACTGAGGCCAGAAAGGCTAGATGACTTAGGGAGGCAGCTC

TGTGGTTCAGTCTAGTCCAGAGACCCAGGGAAAGGGGTGGTTTGCCCATCCGAGTGAACGGC

TTGATTCTGTTGTCTTCGTGTTCTCTGCAATACTGTTTTTCATAATCTAAGGTATTCTCGGG

CCGGAGTTTTCAGTTTGGTCATCTCCCTTTTTCAGAAGTCACCAGAGTGGCAAATTTTTAGA

TTCATTGTTGATGGCTGATTTTTCCCTCTAAAACAGGATTGAATATATGTATTATGCCCCGT

CCTTTGCATATATCGGATTGCATGATTTAAAAAAAAAAAAAATCCCCATTTTAGAGATGAGGG

AAACCGTAGTTCGAAGAGGGGCATGAATATTTATATACCACATCTATATGAAAAAAAGCTTA

AGAATTTCAGCAGTCTGGGTAGTAAGAAATGTGCCCCACGTTCGCAGCAGCAGAGGTAGGAT

TTAACCCCTGATATATGTAATCCAAAACTGGTAAGCTTTGCATCGAAGCCTACAGCAAACTA

CATGGCATACCTTTTTGGATCTTCGCTGCACCTTCAGAAAAACAGCTACAAACAAACAAACA

CGCAAACTCCCTAAGCGAGCTAAAGAAAAATGAGGAATCACTTGACAAGAGGTCAGCAAATT

GGTTCTCACAGTTGCTAGGGTTCCTCTCAGAAGAACAAACCCTCAAGCACCCACCCTCAATC

TTTCTTTAGCAGTAAGTGTTGAAATACTCCACTGGGTTCTCCTCAGTGGGCCCCTTCTTAGG

GAGGTCCAAGGGCTGGGAAATTGACCCTCCCTTGCTCACAAGTCTTGGGACGGAGTGGACAT

TTGGGTATCACTGAGGCTTTAAGGAAGGAAAACCGATTTTCTTCCCAATGATACCACTCAGA

GGGTGGGGTTGAGGAAGCCCAATCTTGGCCACTTTTTTCTATATTTTGCACCCTATGGCCTA

GTTCTGCCCAGTGATGATTCGGCCCGTAAACAGCCAATGTGTAGATGCTTAATTGGGCCAAT

TTTTGGTCACATGCCCAGAGTGAAGTTGATGATCACCACCAGAGTCAGGAAGGAATTTTCCT

CCTCTGGCAAACTGGCCAAGGCTGAGTGGTTTGCTCCTTCCCCTCTCTCTGGGAGGCTGAGC

AGGGGTGCCGGGTTGCTCAGGCCATGGGAGCCACACCTGTTATTGCTGCCTCT

[0057] All sequences are provided using conventional representations of a DNA strand starting from the 5' phosphate linked end to the 3' hydroxyl linked end. The above assignment of coding regions is generally by comparison to available consensus sequence(s) and therefore may contain inconsistencies. These have no effect on the practice of the invention because the invention can be practiced by use of shorter segments (or combinations thereof) of sequences unique to each of the three sets described above and not affected by inconsistencies. As a non-limiting example, a segment of CRIP1 composed of a 3' untranslated region sequence and a sequence from the 3' end of the coding region may be used as a probe for the detection of CRIP1 expression without being affected by the presence of any inconsistency in the representations of the coding regions provided above. Similarly, the use of an antibody which specifically recognizes CRIP1 protein to detect its expression would not be affected by the presence of any inconsistency in the representation of the coding regions provided above.

[0058] As will be appreciated by those skilled in the art, some of the above sequences include 3' poly A (or poly T on the complementary strand) stretches that do not contribute to the uniqueness of the disclosed sequences. The invention may thus be practiced with sequences lacking the 3' poly A (or poly T) stretches. The uniqueness of the disclosed sequences refers to the portions or entireties of the sequences which are found only in CRIP1, HN1, and ESE-2/ELF5 nucleic acids. Preferred unique sequences for the practice of the invention are those which contribute to the consensus sequences for each of the three sets. These preferred unique sequences are of the lengths of polynucleotides of the invention as discussed herein.

[0059] To determine the (increased or decreased) expression levels of the above described sequences in the practice of the present invention, any method known in the art may be utilized. In one preferred embodiment of the invention, expression based on detection of RNA which hybridizes to polynucleotides containing the above described sequences is used. This is readily performed by any RNA detection or amplification+detection method known or recognized as equivalent in the art such as, but not limited to, reverse transcription-PCR (optionally real-time PCR), the methods disclosed in U.S. patent application Ser. No. 10/062,857 entitled "Nucleic Acid Amplification" filed on Oct. 25, 2001 as well as U.S. Provisional Patent Application Nos. 60/298,

847 (filed Jun. 15, 2001) and 60/257,801 (filed Dec. 22, 2000), the methods disclosed in U.S. Pat. No. 6,291,170, and quantitative PCR. Methods to identify increased RNA stability (resulting in an observation of increased expression) or decreased RNA stability (resulting in an observation of decreased expression) may also be used. These methods include the detection of sequences that increase or decrease the stability of mRNAs containing the CRIP1, HN1, and ESE-2/ELF5 sequences disclosed herein. These methods also include the detection of increased mRNA degradation.

[0060] In particularly preferred embodiments of the invention, polynucleotides having sequences present in the 3' untranslated and/or non-coding regions of the above disclosed sequences are used to detect expression or non-expression of CRIP1, HN1, and ESE-2/ELF5 sequences in breast cells in the practice of the invention. Such polynucleotides may optionally contain sequences found in the 3' portions of the coding regions of the above disclosed sequences. Polynucleotides containing a combination of sequences from the coding and 3' non-coding regions preferably have the sequences arranged contiguously, with no intervening heterologous sequence(s).

[0061] Alternatively, the invention may be practiced with polynucleotides having sequences present in the 5' untranslated and/or non-coding regions of the above CRIP1, HN1, and ESE-2/ELF5 sequences in breast cells to detect their levels of expression. Such polynucleotides may optionally contain sequences found in the 5' portions of the coding regions. Polynucleotides containing a combination of sequences from the coding and 5' non-coding regions preferably have the sequences arranged contiguously, with no intervening heterologous sequence(s).

[0062] Preferred polynucleotides contain sequences from 3' or 5' untranslated and/or non-coding regions of at least about 20, at least about 22, at least about 24, at least about 26, at least about 28, at least about 30, at least about 32, at least about 34, at least about 36, at least about 38, at least about 40, at least about 42, at least about 44, or at least about 46 consecutive nucleotides. The term "about" as used in the previous sentence refers to an increase or decrease of 1 from the stated numerical value. Even more preferred are polynucleotides containing sequences of at least or about 50, at least or about 100, at least about or 150, at least or about 200, at least or about 250, at least or about 300, at least or about 350, or at least or about 400 consecutive nucleotides. The term "about" as used in the preceding sentence refers to an increase or decrease of 10% from the stated numerical value.

[0063] Sequences from the 3' or 5' end of the above described coding regions as found in polynucleotides of the invention are of the same lengths as those described above, except that they would naturally be limited by the length of the coding region. The 3' end of a coding region may include sequences up to the 3' half of the coding region. Conversely, the 5' end of a coding region may include sequences up the 5' half of the coding region. Of course the above described sequences, or the coding regions and polynucleotides containing portions thereof, may be used in their entireties.

[0064] Polynucleotides combining the sequences from a 3' untranslated and/or non-coding region and the associated 3' end of the coding region are preferably at least or about 100, at least about or 150, at least or about 200, at least or about 250, at least or about 300, at least or about 350, or at least or about 400 consecutive nucleotides.

[0065] In another embodiment of the invention, polynucleotides containing deletions of nucleotides from the 5' and/or 3' end of the above disclosed sequences may be used. The deletions are preferably of 1-5, 5-10, 10-15, 15-20, 20-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-125, 125-150, 150-175, or 175-200 nucleotides from the 5' and/or 3' end, although the extent of the deletions would naturally be limited by the length of the disclosed sequences and the need to be able to use the polynucleotides for the detection of expression levels.

[0066] In yet another embodiment of the invention, polynucleotides containing portions of the above disclosed sequences including the 3' end may be used in the practice of the invention. Such polynucleotides would contain at least or about 50, at least or about 100, at least about or 150, at least or about 200, at least or about 250, at least or about 300, at least or about 350, or at least or about 400 consecutive nucleotides from the 3' end of the disclosed sequences.

[0067] The invention thus also includes polynucleotides used to detect CRIP1, HN1, and ESE-2/ELF5 expression in breast cells. The polynucleotides may comprise a shorter polynucleotide consisting of sequences found in the above provided SEQ ID NOS in combination with heterologous sequences not naturally found in combination with CRIP1, HN1, and ESE-2/ELF5 sequences. As a non-limiting example, a polynucleotide of the invention may comprise a polynucleotide consisting of the sequence of SEQ ID NO: 29, with a deletion of one or more nucleotides from the 5' and/or 3' end, in combination with one or more non-HN1 sequences.

[0068] Other polynucleotides for use in the practice of the invention include those that have sufficient homology to those described above to detect expression by use of hybridization techniques. Such polynucleotides preferably have about or 95%, about or 96%, about or 97%, about or 98%, or about or 99% identity with CRIP1, HN1, or ESE-2/ELF5 sequences as described herein. Identity is determined using the BLAST algorithm, as described above. The other polynucleotides for use in the practice of the invention may also be described on the basis of the ability to hybridize to polynucleotides of the invention under stringent conditions of about 30% v/v to about 50% formamide and from about 0.01M to about 0.15M salt for hybridization and from about 0.01M to about 0.15M salt for wash conditions at about 55 to about 65° C. or higher, or conditions equivalent thereto.

[0069] In a further embodiment of the invention, a population of single stranded nucleic acid molecules comprising one or both strands of a human CRIP1 or HN1 sequence is provided as a probe such that at least a portion of said population may be hybridized to one or both strands of a nucleic acid molecule quantitatively amplified from RNA of a non-normal or abnormal breast cell. The population may be only the antisense strand of a human CRIP1 or HN1 sequence such that a sense strand of a molecule from, or amplified from, a non-normal or abnormal breast cell may be hybridized to a portion of said population. The population preferably comprises a sufficiently excess amount of said one or both strands of a human CRIP1 or HN1 sequence in comparison to the amount of expressed (or amplified) nucleic acid molecules containing a complementary CRIP1 or HN1 sequence from a normal breast cell. This condition

of excess permits the increased amount of nucleic acid expression in a non-normal or abnormal cell to be readily detectable as an increase.

[0070] Alternatively, the population of single stranded molecules is equal to or in excess of all of one or both strands of the nucleic acid molecules amplified from a non-normal or abnormal breast cell such that the population is sufficient to hybridize to all of one or both strands. Preferred non-normal cells are ADH, DCIS, or IDC cells. The single stranded molecules may of course be the denatured form of any CRIP1 and/or HN1 sequence containing double stranded nucleic acid molecule or polynucleotide as described herein.

[0071] The population may also be described as being hybridized to CRIP1 or HN1 sequence containing nucleic acid molecules at a level of at least twice as much as that by nucleic acid molecules of a normal breast cell. As in the embodiments described above, the nucleic acid molecules may be those quantitatively amplified from a breast cell such that they reflect the amount of expression in said cell.

[0072] The population is preferably immobilized on a solid support, optionally in the form of a location on a microarray. A portion of the population is preferably hybridized to nucleic acid molecules quantitatively amplified from a non-normal or abnormal breast cell by real time PCR. The real time PCR may be practiced by use of amplified RNA from a breast cancer cell, as long as the amplification used was quantitative with respect to CRIP1 and/or HN1 containing sequences.

[0073] In another embodiment of the invention, expression based on detection of DNA status may be used. Detection of the ESE-2/ELF5 DNA as methylated, deleted or otherwise inactivated, may be used as an indication of decreased expression as found in non-normal breast cells. This may be readily performed by PCR based methods known in the art. The status of the promoter regions (SEQ ID NOS: 210 and 211) of the ESE-2/ELF5 may also be assayed as an indication of decreased expression of ESE-2/ELF5 sequences. A non-limiting example is the methylation status of sequences found in the promoter region.

[0074] Conversely, detection of the DNA of a sequence as amplified may be used for as an indication of increased expression as found in non-normal breast cells. This may be readily performed by PCR based, fluorescent in situ hybridization (FISH) and chromosome in situ hybridization (CISH) methods known in the art.

[0075] A preferred embodiment using a nucleic acid based assay to determine expression is by immobilization of one or more of the sequences identified herein on a solid support, including, but not limited to, a solid substrate as an array or to beads or bead based technology as known in the art. Alternatively, solution based expression assays known in the art may also be used. The immobilized sequence(s) may be in the form of polynucleotides as described herein such that the polynucleotide would be capable of hybridizing to a DNA or RNA corresponding to the sequence(s).

[0076] The immobilized polynucleotide(s) may be used to determine the state of nucleic acid samples prepared from sample breast cell(s) for which the pre-cancer or cancer status is not known or for confirmation of a status that is already assigned to the sample breast cell(s). Without limiting the invention, such a cell may be from a patient suspected of being afflicted with, or at risk of developing, breast cancer. The immobilized polynucleotide(s) need only be sufficient to specifically hybridize to the corresponding nucleic acid molecules derived from the sample.

[0077] In embodiments where only one or a few sequences are to be analyzed, the nucleic acid derived from the sample breast cancer cell(s) may be preferentially amplified by use of appropriate primers such that only the sequences to be analyzed are amplified to reduce contaminating background signals from other sequences present in the breast cell. Alternatively, and where the disclosed sequences are to be analyzed in combination with other sequences or where very few cells (or one cell) is used, the nucleic acid from the sample may be globally amplified before hybridization to the immobilized polynucleotides. Of course RNA, or the cDNA counterpart thereof may be directly labeled and used, without amplification, by methods known in the art.

[0078] Sequence expression based on detection of a presence, increase, or decrease in protein levels or activity may also be used. Detection may be performed by any immunohistochemistry (IHC) based, bodily fluid based (where a CRIP1, HN1, and/or ESE-2/ELF5 polypeptide is found in a bodily fluid, such as but not limited to blood), antibody (including autoantibodies against the protein where present) based, ex foliate cell (from the cancer) based, mass spectroscopy based, and image (including used of labeled ligand where available) based method known in the art and recognized as appropriate for the detection of the protein. Antibody and image based methods are additionally useful for the localization of tumors after determination of cancer by use of cells obtained by a non-invasive procedure (such as ductal lavage or fine needle aspiration), where the source of the cancerous cells is not known. A labeled antibody or ligand may be used to localize the carcinoma(s) within a patient.

[0079] Antibodies for use in such methods of detection include polyclonal antibodies, optionally isolated from naturally occurring sources where available, and monoclonal antibodies, including those prepared by use of CRIP1, HN1, and/or ESE-2/ELF5 polypeptides as antigens. Such antibodies, as well as fragments thereof (including but not limited to Fab fragments) function to detect or diagnose non-normal or cancerous breast cells by virtue of their ability to specifically bind CRIP1, HN1, or ESE-2/ELF5 polypeptides to the exclusion of other polypeptides to produce a detectable signal. Recombinant, synthetic, and hybrid antibodies with the same ability may also be used in the practice of the invention. Antibodies may be readily generated by immunization with a CRIP1, HN1, or ESE-2/ELF5 polypeptide, and polyclonal sera may also be used in the practice of the invention.

[0080] Antibody based detection methods are well known in the art and include sandwich and ELISA assays as well as Western blot and flow cytometry based assays as non-limiting examples. Samples for analysis in such methods include any that contain CRIP1, HN1, or ESE-2/ELF5 polypeptides. Non-limiting examples include those containing breast cells and cell contents as well as bodily fluids (including blood, serum, saliva, lymphatic fluid, as well as mucosal and other cellular secretions as non-limiting examples) that contain the polypeptides.

[0081] The above assay embodiments may be used in a number of different ways to identify or detect the presence of non-normal breast cells or breast cancer in a breast cancer cell sample from a patient. In some cases, this would reflect a secondary screen for the patient, who may have already undergone mammography or physical exam as a primary screen. If positive from the primary screen, the subsequent needle biopsy, ductal lavage, fine needle aspiration, or other analogous methods may provide the sample for use in the assay embodiments described herein. The present invention is particularly useful in combination with non-invasive protocols, such as ductal lavage or fine needle aspiration, to prepare a breast cell sample. The current analysis of ductal lavage samples is by cytological examination by a trained pathologist who classifies the samples in terms that are at least partly subjective: unsatisfactory (too few cells), benign (including fibrocystic change), atypical (or mild atypia), suspicious (or marked atypia), or malignant.

[0082] The present invention provides a more objective set of criteria, in the form of gene expression levels of discrete gene sequences, to discriminate (or delineate) between normal and non-normal breast cells.

[0083] In one embodiment of the invention, the isolation and analysis of a breast cancer cell sample may be performed as follows:

[0084] (1) Ductal lavage or other non-invasive procedure is performed on a patient to obtain a sample.

[0085] (2) Sample is prepared and coated onto a microscope slide. Note that ductal lavage results in clusters of cells that are cytologically examined as stated above.

[0086] (3) Pathologist or image analysis software scans the sample for the presence of atypical cells.

[0087] (4) If atypical cells are observed, those cells are harvested (e.g. by microdissection such as LCM).

[0088] (5) RNA is extracted from the harvested cells.

[0089] (6) RNA is assayed for the expression of CRIP1, HN1, and/or ESE-2/ELF5 sequences.

[0090] A specific example of the above method would be performing ductal lavage following a primary screen, observing and collecting non-normal cells (or cells suspected of being non-normal) for analysis. Alternatively, the sample may permit the collection of both normal and non-normal cells (or cells suspected of being non-normal) for analysis. The expression levels of CRIP1, HN1, and/or ESE-2/ELF5 sequences in each of these two populations may be compared to each other. This approach can be significantly more powerful than one using the non-normal cells only approach because it utilizes information from the normal cells and the differences between normal and non-normal cells to determine the status of the non-normal cells from the sample.

[0091] While many clinical settings focus on identification of the highest stage or grade of breast cancer, the detection of non-normal breast cells of any stage or grade is also important to identify with confidence the presence of, or susceptibility to, breast cancer as early as possible. With use of the present invention, skilled clinicians will be apprised of the presence of non-normal cells quickly and may begin treatment or additional testing based on such information.

[0092] The present invention may also be used with solid tissue biopsies. As a non-limiting example, a solid biopsy may be collected and prepared for visualization followed by determination of increased CRIP1 and/or HN1 expression to identify or diagnose the presence of non-normal cells. One preferred means is by use of in situ hybridization with polynucleotide or protein identifying probe(s) for assaying expression of said gene(s). An analogous method may be used to detect decreased expression of ESE-2/ELF5 sequences.

[0093] In an alternative method, the solid tissue biopsy may be used to extract molecules followed by analysis for expression of the disclosed sequence(s). This provides the possibility of leaving out the need for visualization and collection of only those cells suspected of being non-normal. This method may of course be modified such that only cells suspected of being non-normal are collected and used to extract molecules for analysis. This would require some form of selection as a prerequisite to gene expression analysis.

[0094] In a further modification of the above, both normal cells and cells suspected of being non-normal are collected and used to extract molecules for analysis of sequence expression. The approach, benefits and results are as described above using non-invasive sampling.

[0095] In a further alternative to all of the above, the sequence(s) identified herein may be used as part of a simple PCR or array based assay simply to determine the presence of non-normal cells in a sample from a non-invasive sampling procedure. If normal expression levels of the disclosed sequences are identified, no further examination may be necessary. If non-normal expression levels are detected, a more comprehensive analysis may follow.

[0096] The detection of sequence expression from samples may be by use of a single microarray able to assay expression of the disclosed sequences as well as other sequences, including sequences known not to vary in expression levels between normal and non-normal breast cells, for convenience and improved accuracy.

[0097] Other uses of the present invention include providing the ability to identify breast cancer cell samples as being non-normal for further research or study. This provides a particular advantage in many contexts requiring the identification of non-normal or cancerous cells based on objective genetic or molecular criteria rather than cytological observation.

[0098] The materials and methods of the present invention are ideally suited for preparation of kits produced in accordance with well known procedures. The invention thus provides kits comprising agents (like the polynucleotides and/or antibodies described herein as non-limiting examples) for the detection of expression of the disclosed sequences. Such kits, optionally comprising the agent with an identifying description or label or instructions relating to their use in the methods of the present invention, are provided. Such a kit may comprise containers, each with one or more of the various reagents (typically in concentrated form) utilized in the methods, including, for example, prefabricated microarrays, buffers, the appropriate nucleotide

triphosphates (e.g., dATP, dCTP, dGTP and dTTP; or rATP, rCTP, rGTP and UTP), reverse transcriptase, DNA polymerase, RNA polymerase, and one or more primer complexes of the present invention (e.g., appropriate length poly(T) or random primers linked to a promoter reactive with the RNA polymerase). A set of instructions will also typically be included.

[0099] The methods provided by the present invention may also be automated in whole or in part. All aspects of the present invention may also be practiced such that they consist essentially of a subset, or subregion, of the disclosed sequences to the exclusion of material irrelevant to the identification of non-normal or cancerous breast cells.

[0100] To identify changes in sequence expression in non-normal or cancerous breast cells, abnormal epithelium from ADH, DCIS and IDC and phenotypically normal epithelium (henceforth referred to as normal) from 36 breast cancer patients and 3 healthy mammoplasty reduction patients were isolated via laser capture microdissection (LCM). The resulting 300 independently microdissected samples were used to interrogate a microarray containing approximately 12,000 human genes.

[0101] One important advantage provided by LCM use is the ability to procure both normal and diseased cell populations from the same biopsy. Therefore, the expression level of each gene in a disease state (ADH or DCIS or IDC) is represented as the ratio to the patient-matched normal, which highlights differences due to disease state as opposed to the genetic background of a particular patient. Unsupervised hierarchical clustering revealed one sequence cluster demonstrating increased expression in a majority of the diseased samples. CRIP1 was included in the cluster and identified as displaying increased expression levels in ADH with persistence in DCIS and IDC samples. Its increased expression may thus be a potential biomarker for the detection of breast cancer including the pre-malignant stage of ADH.

[0102] Sequences of another cluster demonstrated decreased expression in all three pathological stages. The epithelium-specific transcription factor ELF5 was included, and loss of expression of ELF5 in ADH may be an important first step in the initiation of breast malignancy.

[0103] Additional experiments resulted in the identification of increased expression of HN1 sequences in grade III DCIS and IDC cells.

[0104] Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

EXAMPLES

Example I

[0105] Materials and Methods

[0106] Clinical specimen collection and clinicopathological parameters. All breast specimens were obtained from the Massachusetts General Hospital between 1998 and 2001. Thirty-six breast cancer patients were selected, 31 of which were diagnosed with two or more pathological stages of breast cancer progression, and 5 of which were diagnosed with pre-invasive disease only. Three healthy women who underwent elective mammoplasty reduction were selected as disease-free normal controls. Tissue specimens that demonstrated one or more pathological lesions (ADH, DCIS and IDC) were selected for the study. Cases of ADH were selected as proliferative epithelial lesions that possessed some, but not all, of the features of carcinoma in situ (Page, D. L. et al. (1992)) and most closely resemble those lesions described as CAPSS (Oyama, T. et al. and Fraser, J. L. et al.). DCIS and IDC were classified (histological grade) according to the European classification (Holland, R. et al.) and by the Nottingham combined histological grade (Elston, C. W. et al.), respectively. ER and PR expression were determined by immunohistochemical staining (negative when none of the tumor cell nuclei showed staining), and Her-2 expression determined by immunohistochemistry or FISH. This study was approved the Massachusetts General Hospital human research committee in accordance with NIH human research study guidelines.

[0107] LCM and RNA isolation and amplification. Each component (Normal, ADH, DCIS or IDC) was laser capture microdissected in triplicate (from consecutive tissue sections) as described (Sgroi et al.) using a PixCell II LCM system (Arcturus Engineering Inc., Mountain View, Calif.). Total RNA was extracted from the captured cells using the Picopure™ RNA Isolation Kit (Arcturus). T7-based RNA amplification was carried out using the RiboAmp™ kit (Arcturus). Briefly, the RNA from each sample was primed with an oligo-dT primer containing a T7 promoter sequence, reverse transcribed and then converted to double stranded cDNA. The cDNA templates were then used in an in vitro transcription reaction using T7 RNA polymerase to generate amplified RNA (aRNA). To obtain enough aRNA for a microarray experiment, a second round of RNA amplification was performed on all samples. To serve as reference in microarray hybridizations, a human universal reference RNA from Stratagene (La Jolla, Calif.) was amplified identically.

[0108] Fabrication of microarrays. Sequence-verified human cDNA clones were obtained from Research Genetics (Huntsville, Ala.). cDNA clones (from the I.M.A.G.E. Consortium via Research Genetics) inserts were amplified by PCR, gel-purified, and spotted onto a 1×3-inch SuperAmine™ (TeleChem International, Sunnyvale, Calif.) glass microscope slide using an OmniGrid™ robotic arrayer (GeneMachines, San Carlos, Calif.). As used herein, the I.M.A.G.E. Consortium CloneID, or the IMAGE CloneID, lists the identifiers of the cDNA clones on the microarrays according to the I.M.A.G.E. Consortium and Research Genetics (www.resgen.com/). This provides a unique single identifier for each clone. Descriptive names of clones (or genes) use the UniGene symbols and titles (www.ncbi.nlm.nih.gov/UniGene/).

[0109] Probe labeling and hybridization. cDNA was transcribed from aRNA in the presence of 5-(3-aminoallyl)-2'-deoxyuridine 5'-triphosphate (aminoallyl dUTP) using Stratagene's FairPlay kit™ (La Jolla, Calif.). Cy3 or Cy5 monoreactive dye (Amersham, Piscataway, N.J.) was conjugated onto purified cDNA and the residual dye was removed using QiaQuick PCR Purification columns (Qiagen, Valencia, Calif.). Each Cy5-labeled cDNA was hybridized together with the Cy3-labeled reference probe to a microarray in 40

μL hybridization solution (5×SSC, 0.1 μg/μL COT I, 0.2%SDS, 50% formamide) at a concentration of 25 ng/μL per channel for 17 hrs at 42° C. in >60% relative humidity.

[0110] Washing, scanning and image analysis. After hybridization, slides were washed as follows: 1×SSC, 0.2% SDS at 42° for 5 min (two times), 1×SSC, 0.2% SDS at 55° C. for 5 min, 0.1×SSC, 0.2% SDS at 55° C. for 5 min and 0.1×SSC at RT for 2 min. Washed slides were scanned using ScanArray 5000 (PerkinElmer, Billerica, Mass.), and Cy5/Cy3-signals were quantitated using ImaGene 4.2 (BioDiscovery, Los Angeles, Calif.).

[0111] Data processing. Fluorescent intensities of Cy5 and Cy3 channels on each slide were subjected to spot filtering and normalization. Spots flagged by ImaGene were excluded from further analysis. Normalization was performed using a robust nonlinear local regression method (Yang, Y. H. et al.). The normalized ratios of Cy5/Cy3 were used to represent the relative gene expression levels in the experimental samples. Measurements from replicate samples were averaged after normalization.

[0112] Cluster and discriminant analysis. Hierarchical cluster analysis was performed in GeneMaths (v1.5, Applied-Maths, Austin, Tex.) using the cosine correlation coefficient as a measure of similarity between two genes or samples and complete linkage. Linear discriminant analysis with variance was performed within GeneMaths.

Example II

[0113] Analysis of Over and Under Expression

[0114] Quantitative real-time PCR analysis of CRIP1 and ELF5 was conducted to confirm their over and under expression in non-normal breast cells.

[0115] For the non-amplified RNA RT-PCR validation study, independently laser captured (~40,000) normal breast epithelial cells from case 215, and ~40,000 abnormal epithelial cells from DCIS (from cases 89, 178, 179) or IDC (from cases 97, 169, 170) were used. Total RNA was isolated and converted to double-stranded cDNA. For studies using amplified RNA (aRNA), 2 mgs of aRNA from each microdissected sample was converted into double-stranded cDNA. In all cases (cDNA derived from non-amplified and amplified RNA), the double-stranded cDNA was quantitated with PicoGreen (Molecular Probes) using a spectrofluorometer (Molecular Devices) and quantitative analysis of gene expression performed (RT-PCR) was performed with an ABI 7900HT (Applied Biosystems, Foster City, Calif.) as described (Sgroi et al., 1999).

[0116] Each reaction was performed in triplicate using 2.5 ng of double stranded cDNA from each sample as template. The relative standard curve method was used for linear regression analysis of unknown samples and data presented as fold change between samples. The sequences of the PCR primer pairs and fluorogenic probe (5' to 3'), respectively, that were used for each gene are as follows:

```
CRIP1:
CCTGCTACGCAGCCATGTT,

GGATGGGTCTCCACCACCT,

VIC-CGGAGCCGAGAGCCACACTTTCAAGT-TAMRA;
```

-continued

```
ELF5:
TGATTCCTGCTCTGGTGAAACA,

ACATTTTCTCATCAATCA CTGGTATGT,

VIC-CAGTCCACCTAGCTGTCAGTCACTGATA-TAMRA;
```

[0117] In agreement with microarray results used to initial identify sequences that were over and under expressed in non-normal breast cells, RT-PCR demonstrated over-expression of CRIP1 (>2-fold) in 7 of 8 ADH, 27 of 30 DCIS, and 23 of 25 IDC cases, and under-expression of ELF5 (>2-fold) in 7 of 8 ADH, 28 of 30 DCIS, and 25 of 25 IDC cases (**FIG. 1**).

[0118] In addition, we performed in-situ hybridization for CRIP1 to confirm its cellular specificity. As expected from the use of LCM, CRIP1 signal localized to the epithelial cells, and its intensity was markedly increased in the IDC compartment of the same biopsy (**FIGS. 2 and 3**), thus verifying the microarray-derived results at the level of cellular resolution.

[0119] All references cited herein, including patents, patent applications, and publications, are hereby incorporated by reference in their entireties, whether previously specifically incorporated or not.

[0120] Having now fully described this invention, it will be appreciated by those skilled in the art that the same can be performed within a wide range of equivalent parameters, concentrations, and conditions without undue experimentation. This application is intended to cover any variations, uses, or adaptations of the invention, following in general the principles of the invention, that include such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth.

We claim:

1. A method to determine the presence of non-normal or abnormal breast cells in a sample from a human subject comprising assaying said sample for increased expression of one or more human CRIP1 or HN1 sequences.

2. The method of claim 1 wherein said assaying is for increased expression of human CRIP1 sequences.

3. The method of claim 1 wherein said assaying is for increased expression of human HN1 sequences.

4. The method of claim 1 wherein said sample is from a subject afflicted with, or suspected of having, breast cancer.

5. The method of claim 1 wherein said sample is obtained by solid tissue biopsy or a non-invasive procedure.

6. The method of claim 5 wherein said non-invasive procedure is selected from ductal lavage, fine needle aspiration, or a needle biopsy.

7. The method of claim 5 wherein microdissection is used to isolate breast cells from said sample before assaying for nucleic acid expression.

8. The method of claim 1 wherein said assaying is by hybridization to a polynucleotide comprising sequences of at least 24 nucleotides from the 3' untranslated region, the coding region, or the 5' untranslated region, of human CRIP1.

9. The method of claim 1 wherein said assaying is by hybridization to a polynucleotide comprising sequences of at least 24 nucleotides from the 3' untranslated region, the coding region, or the 5' untranslated region, of human HN1.

10. The method of claim 1 wherein said assaying is by PCR amplification of said sequences.

11. The method of claim 10 wherein said PCR is quantitative PCR.

12. The method of claim 1 wherein said non-normal cells are ADH, DCIS, or IDC cells.

13. A method to determine the presence of non-normal or abnormal breast cells in a sample from a human subject comprising assaying said sample for decreased expression of human ESE-2/ELF5 sequences.

14. The method of claim 13 wherein said sample is from a subject afflicted with, or suspected of having, breast cancer.

15. The method of claim 13 wherein said sample is obtained by solid tissue biopsy or a non-invasive procedure.

16. The method of claim 15 wherein said non-invasive procedure is selected from ductal lavage, fine needle aspiration, or a needle biopsy.

17. The method of claim 16 wherein microdissection is used to isolate breast cells from said sample before assaying for nucleic acid expression.

18. The method of claim 17 wherein said assaying is by hybridization to a polynucleotide comprising sequences of at least 24 nucleotides from the 3' untranslated region, the coding region, or the 5' untranslated region, of human ESE-2/ELF5.

19. The method of claim 13 wherein said assaying is by PCR amplification of said ESE-2/ELF5 sequence.

20. The method of claim 19 wherein said assaying is by quantitative PCR.

21. The method of claim 13 wherein said assaying is for inactivation or methylation of ESE-2/ELF5 sequences.

22. The method of claim 13 wherein said assaying comprises detection of increased mRNA degradation.

23. The method of claim 13 wherein said non-normal cells are ADH, DCIS, or IDC cells.

24. A polynucleotide comprising

a segment consisting of a fragment of an HN I sequence selected from SEQ ID NOS: 29-83 of between 24 and 500 nucleotides and

one or more non-HN1 nucleic acid molecules.

25. A polynucleotide consisting of a fragment of an HN1 sequence selected from SEQ ID NOS: 29-83 of between 24 and 500 nucleotides.

26. A population of singled stranded nucleic acid molecules comprising one or both strands of a human CRIP1 or HN1 sequence wherein at least a portion of said population is hybridized to one or both strands of a nucleic acid molecule quantitatively amplified from RNA of a non-normal or abnormal breast cell.

27. The population of claim 26 wherein the population is immobilized on a solid support.

28. The population of claim 27 wherein said solid support is a microarray.

29. The population of claim 26 wherein said nucleic acid molecules amplified from a non-normal or abnormal breast cell are amplified by quantitative real time PCR (RT-PCR).

30. The population of claim 29 wherein said quantitative RT-PCR is of amplified RNA of said breast cancer cell.

31. The population of claim 26 wherein said population of single stranded molecules is equal to or in excess of all of one or both strands of the nucleic acid molecules amplified from a non-normal or abnormal breast cell such that the population is sufficient to hybridize to all of one or both strands.

32. The population of claim 26 wherein said population of single stranded molecules comprising sequences of at least 24 nucleotides from the 3' untranslated region, the coding region, or the 5' untranslated region, of human CRIP1.

33. The population of claim 26 wherein said population of single stranded molecules comprising sequences of at least 24 nucleotides from the 3' untranslated region, the coding region, or the 5' untranslated region, of human HN1.

34. The population of claim 26 wherein said non-normal cells are ADH, DCIS, or IDC cells.

* * * * *