



Office de la Propriété
Intellectuelle
du Canada

Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An agency of
Industry Canada

CA 2506887 A1 2004/06/03

(21) **2 506 887**

(12) **DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION**

(13) **A1**

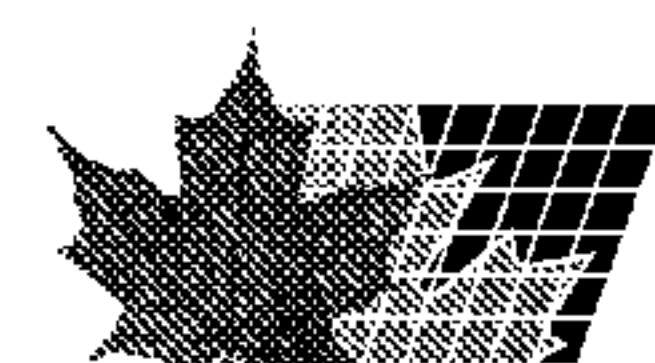
(86) Date de dépôt PCT/PCT Filing Date: 2003/11/21
(87) Date publication PCT/PCT Publication Date: 2004/06/03
(85) Entrée phase nationale/National Entry: 2005/05/20
(86) N° demande PCT/PCT Application No.: GB 2003/005102
(87) N° publication PCT/PCT Publication No.: 2004/046382
(30) Priorité/Priority: 2002/11/21 (0227238.3) GB

(51) Cl.Int.⁷/Int.Cl.⁷ C12Q 1/68
(71) Demandeur/Applicant:
DIAGENIC AS, NO
(72) Inventeurs/Inventors:
SHARMA, PRAVEEN, NO;
SAHNI, NARINDER SINGH, NO;
LOENNEBORG, ANDERS, NO
(74) Agent: RIDOUT & MAYBEE LLP

(54) Titre : PRODUIT ET PROCEDE
(54) Title: PRODUCT AND METHOD

(57) **Abrégé/Abstract:**

The present invention relates to oligonucleotide probes, for use in assessing gene transcript levels in a cell, which may be used in analytical techniques, particularly diagnostic techniques and kits containing the same.



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

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
3 June 2004 (03.06.2004)

PCT

(10) International Publication Number
WO 2004/046382 A3

(51) International Patent Classification⁷: **C12Q 1/68**

(21) International Application Number:
PCT/GB2003/005102

(22) International Filing Date:
21 November 2003 (21.11.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0227238.3 21 November 2002 (21.11.2002) GB

(71) Applicant (*for all designated States except US*): **DIA-GENIC AS** [NO/NO]; Østensjøveien 15B, N-0661 Oslo (NO).

(71) Applicant (*for GB only*): **JONES, Elizabeth, Louise** [GB/GB]; Frank B. Dehn & Co., 179 Queen Victoria Street, London EC4V 4EL (GB).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **SHARMA, Praveen** [IN/NO]; Waldemars Thranes gate 62 D, N-0173 Oslo (NO). **SAHNI, Narinder, Singh** [IN/NO]; Ryenbergveien 52 D, N-0196 Oslo (NO). **LÖNNEBORG, Anders** [SE/NO]; Feltspatveien 14, N-1430 Aas (NO).

(74) Agent: **JONES, Elizabeth, Louise**; Frank B. Dehn & Co., 179 queen Victoria Street, London EC4V 4EL (GB).

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

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

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
22 July 2004

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

(54) Title: PRODUCT AND METHOD

(57) Abstract: The present invention relates to oligonucleotide probes, for use in assessing gene transcript levels in a cell, which may be used in analytical techniques, particularly diagnostic techniques and kits containing the same.



WO 2004/046382 A3

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 276

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 276

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

Product and Method

The present invention relates to oligonucleotide probes, for use in assessing gene transcript levels in a cell, which may be used in analytical techniques, particularly diagnostic techniques. Conveniently the probes are provided in kit form. Different sets of probes may be used in techniques to prepare gene expression patterns and identify, diagnose or monitor different states, such as diseases, conditions or stages thereof. Also provided are methods of identifying suitable probes and their use in methods of the invention.

The identification of quick and easy methods of sample analysis for, for example, diagnostic applications, remains the goal of many researchers. End users seek methods which are cost effective, produce statistically significant results and which may be implemented routinely without the need for highly skilled individuals.

The analysis of gene expression within cells has been used to provide information on the state of those cells and importantly the state of the individual from which the cells are derived. The relative expression of various genes in a cell has been identified as reflecting a particular state within a body. For example, cancer cells are known to exhibit altered expression of various proteins and the transcripts or the expressed proteins may therefore be used as markers of that disease state.

Thus biopsy tissue may be analysed for the presence of these markers and cells originating from the site of the disease may be identified in other tissues or fluids of the body by the presence of the markers. Furthermore, products of the altered expression may be

- 2 -

released into the blood stream and these products may be analysed. In addition cells which have contacted disease cells may be affected by their direct contact with those cells resulting in altered gene expression and their expression or products of expression may be similarly analysed.

However, there are some limitations with these methods. For example, the use of specific tumour markers for identifying cancer suffers from a variety of defects, such as lack of specificity or sensitivity, association of the marker with disease states besides the specific type of cancer, and difficulty of detection in asymptomatic individuals.

In addition to the analysis of one or two marker transcripts or proteins, more recently, gene expression patterns have been analysed. Most of the work involving large-scale gene expression analysis with implications in disease diagnosis has involved clinical samples originating from diseased tissues or cells. For example, several recent publications, which demonstrate that gene expression data can be used to distinguish between similar cancer types, have used clinical samples from diseased tissues or cells (Alon et al. 1999, PNAS, 96, p6745-6750; Golub et al. 1999, Science, 286, p531-537; Alizadeh et al, 2000, Nature, 403, p503-511; Bittner et al., 2000, Nature, 406, p536-540).

However, these methods have relied on analysis of a sample containing diseased cells or products of those cells or cells which have been contacted by disease cells. Analysis of such samples relies on knowledge of the presence of a disease and its location, which may be difficult in asymptomatic patients. Furthermore, samples can not always be taken from the disease site, e.g. in diseases of the brain.

In a finding of great significance, the present inventors identified the previously untapped potential of all cells within a body to provide information

- 3 -

relating to the state of the organism from which the cells were derived. WO98/49342 describes the analysis of the gene expression of cells distant from the site of disease, e.g. peripheral blood collected distant from a cancer site.

5 This finding is based on the premise that the different parts of an organism's body exist in dynamic interaction with each other. When a disease affects one part of the body, other parts of the body are also affected. The interaction results from a wide spectrum of biochemical signals that are released from the diseased area, affecting other areas in the body. Although, the nature of the biochemical and physiological changes induced by the released signals can vary in the different body parts, the changes can be measured at the level of gene expression and used for diagnostic purposes.

10 The physiological state of a cell in an organism is determined by the pattern with which genes are expressed in it. The pattern depends upon the internal and external biological stimuli to which said cell is exposed, and any change either in the extent or in the nature of these stimuli can lead to a change in the pattern with which the different genes are expressed in the cell. There is a growing understanding that by analysing the systemic changes in gene expression patterns in cells in biological samples, it is possible to provide information on the type and nature of the biological stimuli that are acting on them. Thus, for example, by monitoring the expression of a large number of genes in cells in a test sample, it is possible to determine whether their genes are expressed with a pattern characteristic for a particular disease, condition or stage thereof. Measuring changes in gene activities in cells, e.g. from tissue or body fluids is therefore emerging as a powerful tool for disease diagnosis.

- 4 -

Such methods have various advantages. Often, obtaining clinical samples from certain areas in the body that is diseased can be difficult and may involve undesirable invasions in the body, for example biopsy is often used to obtain samples for cancer. In some cases, such as in Alzheimer's disease the diseased brain specimen can only be obtained post-mortem. Furthermore, the tissue specimens which are obtained are often heterogeneous and may contain a mixture of both diseased and non-diseased cells, making the analysis of generated gene expression data both complex and difficult.

It has been suggested that a pool of tumour tissues that appear to be pathogenetically homogeneous with respect to morphological appearances of the tumour may well be highly heterogeneous at the molecular level (Alizadeh, 2000, supra), and in fact might contain tumours representing essentially different diseases (Alizadeh, 2000, supra; Golub, 1999, supra). For the purpose of identifying a disease, condition, or a stage thereof, any method that does not require clinical samples to originate directly from diseased tissues or cells is highly desirable since clinical samples representing a homogeneous mixture of cell types can be obtained from an easily accessible region in the body.

We have now identified a set of probes of surprising utility for identifying one or more diseases. Thus, we now describe probes and sets of probes derived from cells which are not disease cells and which have not contacted disease cells, which correspond to genes which exhibit altered expression in normal versus disease individuals, for use in methods of identifying, diagnosing or monitoring certain conditions, particularly diseases or stages thereof.

Thus the invention provides a set of oligonucleotide probes which correspond to genes in a cell whose expression is affected in a pattern characteristic of a particular disease, condition or

- 5 -

stage thereof, wherein said genes are systemically affected by said disease, condition or stage thereof. Preferably said genes are metabolic or house-keeping genes and preferably are constitutively moderately or highly expressed. Preferably the genes are moderately or highly expressed in the cells of the sample but not in cells from disease cells or in cells having contacted such disease cells.

Such probes, particularly when isolated from cells distant to the site of disease, do not rely on the development of disease to clinically recognizable levels and allow detection of a disease or condition or stage thereof very early after the onset of said disease or condition, even years before other subjective or objective symptoms appear.

As used herein "systemically" affected genes refers to genes whose expression is affected in the body without direct contact with a disease cell or disease site and the cells under investigation are not disease cells.

"Contact" as referred to herein refers to cells coming into close proximity with one another such that the direct effect of one cell on the other may be observed, e.g. an immune response, wherein these responses are not mediated by secondary molecules released from the first cell over a large distance to affect the second cell. Preferably contact refers to physical contact, or contact that is as close as is sterically possible, conveniently, cells which contact one another are found in the same unit volume, for example within 1cm^3 .

A "disease cell" is a cell manifesting phenotypic changes and is present at the disease site at some time during its life-span, e.g. a tumour cell at the tumour site or which has disseminated from the tumour, or a brain cell in the case of brain disorders such as Alzheimer's disease.

- 6 -

"Metabolic" or "house-keeping" genes refer to those genes responsible for expressing products involved in cell division and maintenance, e.g. non-immune function related genes.

5 "Moderately or highly" expressed genes refers to those present in resting cells in a copy number of more than 30-100 copies/cell (assuming an average 3×10^5 mRNA molecules in a cell).

10 Specific probes having the above described properties are provided herein.

Thus in one aspect, the present invention provides a set of oligonucleotide probes, wherein said set comprises at least 10 oligonucleotides selected from:

15 an oligonucleotide as described in Table 1 or derived from a sequence described in Table 1, or an oligonucleotide with a complementary sequence, or a functionally equivalent oligonucleotide.

20 "Table 1" as referred to herein refers to Table 1a and/or Table 1b. Table 1b contains reference to additional clones and sequences as disclosed herein. Similarly Tables 2 and 4 comprise 2 parts, a and b.

25 The invention also provides one or more oligonucleotide probes, wherein each oligonucleotide probe is selected from the oligonucleotides listed in Table 1, or derived from a sequence described in Table 1, or a complementary sequence thereof. The use of such probes in products and methods of the invention, form further aspects of the invention.

30 As referred to herein an "oligonucleotide" is a nucleic acid molecule having at least 6 monomers in the polymeric structure, ie. nucleotides or modified forms thereof. The nucleic acid molecule may be DNA, RNA or PNA (peptide nucleic acid) or hybrids thereof or modified versions thereof, e.g. chemically modified forms, e.g. LNA (Locked Nucleic acid), by methylation or
35 made up of modified or non-natural bases during synthesis, providing they retain their ability to bind

- 7 -

to complementary sequences. Such oligonucleotides are used in accordance with the invention to probe target sequences and are thus referred to herein also as oligonucleotide probes or simply as probes.

5 An "oligonucleotide derived from a sequence described in Table 1" (or any other table) refers to a part of a sequence disclosed in that Table (e.g. Table 1-4), which satisfies the requirements of the oligonucleotide probes as described herein, e.g. in
10 length and function. Preferably said parts have the size described hereinafter.

 Preferably the oligonucleotide probes forming said set are at least 15 bases in length to allow binding of target molecules. Especially preferably said
15 oligonucleotide probes are from 20 to 200 bases in length, e.g. from 30 to 150 bases, preferably 50-100 bases in length.

 As referred to herein the term "complementary sequences" refers to sequences with consecutive
20 complementary bases (ie. T:A, G:C) and which complementary sequences are therefore able to bind to one another through their complementarity.

 Reference to "10 oligonucleotides" refers to 10 different oligonucleotides. Whilst a Table 1
25 oligonucleotide, a Table 1 derived oligonucleotide and their functional equivalent are considered different oligonucleotides, complementary oligonucleotides are not considered different. Preferably however, the at least 10 oligonucleotides are 10 different Table 1
30 oligonucleotides (or Table 1 derived oligonucleotides or their functional equivalents). Thus said 10 different oligonucleotides are preferably able to bind to 10 different transcripts.

 Preferably said oligonucleotides are as described
35 in Table 1 or are derived from a sequence described in Table 1. Especially preferably said oligonucleotides are as described in Table 2 or Table 4 or are derived

- 8 -

from a sequence described in either of those tables. Especially preferably the oligonucleotide (or the oligonucleotide derived therefrom) has a high occurrence as defined in Table 3, especially preferably >40%, e.g. >80 or >90, e.g. 100%.

A "set" as described refers to a collection of unique oligonucleotide probes (ie. having a distinct sequence) and preferably consists of less than 1000 oligonucleotide probes, especially less than 500 probes, e.g. preferably from 10 to 500, e.g. 10 to 100, 200 or 300, especially preferably 20 to 100, e.g. 30 to 100 probes. In some cases less than 10 probes may be used, e.g. from 2 to 9 probes, e.g. 5 to 9 probes.

It will be appreciated that increasing the number of probes will prevent the possibility of poor analysis, e.g. misdiagnosis by comparison to other diseases which could similarly alter the expression of the particular genes in question. Other oligonucleotide probes not described herein may also be present, particularly if they aid the ultimate use of the set of oligonucleotide probes. However, preferably said set consists only of said Table 1 oligonucleotides, Table 1 derived oligonucleotides, complementary sequences or functionally equivalent oligonucleotides, or a sub-set thereof (e.g. of the size as described above), preferably a sub-set for which sequences are provided herein (see Table 1 and its footnote). Especially preferably said set consists only of said Table 1 oligonucleotides, Table 1 derived oligonucleotides, or complementary sequences thereof, or a sub-set thereof.

Multiple copies of each unique oligonucleotide probe, e.g. 10 or more copies, may be present in each set, but constitute only a single probe.

A set of oligonucleotide probes, which may preferably be immobilized on a solid support or have means for such immobilization, comprises the at least 10 oligonucleotide probes selected from those described

- 9 -

hereinbefore. Especially preferably said probes are selected from those having high occurrence as described in Table 3 and as mentioned above. As mentioned above, these 10 probes must be unique and have different sequences. Having said this however, two separate probes may be used which recognize the same gene but reflect different splicing events. However oligonucleotide probes which are complementary to, and bind to distinct genes are preferred.

As described herein a "functionally equivalent" oligonucleotide to those described in Table 1 or derived therefrom refers to an oligonucleotide which is capable of identifying the same gene as an oligonucleotide of Table 1 or derived therefrom, ie. it can bind to the same mRNA molecule (or DNA) transcribed from a gene (target nucleic acid molecule) as the Table 1 oligonucleotide or the Table 1 derived oligonucleotide (or its complementary sequence). Preferably said functionally equivalent oligonucleotide is capable of recognizing, ie. binding to the same splicing product as a Table 1 oligonucleotide or a Table 1 derived oligonucleotide. Preferably said mRNA molecule is the full length mRNA molecule which corresponds to the Table 1 oligonucleotide or the Table 1 derived oligonucleotide.

As referred to herein "capable of binding" or "binding" refers to the ability to hybridize under conditions described hereinafter.

Alternatively expressed, functionally equivalent oligonucleotides (or complementary sequences) have sequence identity or will hybridize, as described hereinafter, to a region of the target molecule to which molecule a Table 1 oligonucleotide or a Table 1 derived oligonucleotide or a complementary oligonucleotide binds. Preferably, functionally equivalent oligonucleotides (or their complementary sequences) hybridize to one of the mRNA sequences which corresponds

- 10 -

to a Table 1 oligonucleotide or a Table 1 derived oligonucleotide under the conditions described hereinafter or has sequence identity to a part of one of the mRNA sequences which corresponds to a Table 1
5 oligonucleotide or a Table 1 derived oligonucleotide. A "part" in this context refers to a stretch of at least 5, e.g. at least 10 or 20 bases, such as from 5 to 100, e.g. 10 to 50 or 15 to 30 bases.

In a particularly preferred aspect, the
10 functionally equivalent oligonucleotide binds to all or a part of the region of a target nucleic acid molecule (mRNA or cDNA) to which the Table 1 oligonucleotide or Table 1 derived oligonucleotide binds. A "target" nucleic acid molecule is the gene transcript or related
15 product e.g. mRNA, or cDNA, or amplified product thereof. Said "region" of said target molecule to which said Table 1 oligonucleotide or Table 1 derived oligonucleotide binds is the stretch over which complementarity exists. At its largest this region is
20 the whole length of the Table 1 oligonucleotide or Table 1 derived oligonucleotide, but may be shorter if the entire Table 1 sequence or Table 1 derived oligonucleotide is not complementary to a region of the target sequence.

25 Preferably said part of said region of said target molecule is a stretch of at least 5, e.g. at least 10 or 20 bases, such as from 5 to 100, e.g. 10 to 50 or 15 to 30 bases. This may for example be achieved by said functionally equivalent oligonucleotide having several
30 identical bases to the bases of the Table 1 oligonucleotide or the Table 1 derived oligonucleotide. These bases may be identical over consecutive stretches, e.g. in a part of the functionally equivalent oligonucleotide, or may be present non-consecutively,
35 but provide sufficient complementarity to allow binding to the target sequence.

Thus in a preferred feature, said functionally

- 11 -

equivalent oligonucleotide hybridizes under conditions of high stringency to a Table 1 oligonucleotide or a Table 1 derived oligonucleotide or the complementary sequence thereof. Alternatively expressed, said
5 functionally equivalent oligonucleotide exhibits high sequence identity to all or part of a Table 1 oligonucleotide. Preferably said functionally equivalent oligonucleotide has at least 70% sequence identity, preferably at least 80%, e.g. at least 90, 95,
10 98 or 99%, to all of a Table 1 oligonucleotide or a part thereof. As used in this context, a "part" refers to a stretch of at least 5, e.g. at least 10 or 20 bases, such as from 5 to 100, e.g. 10 to 50 or 15 to 30 bases, in said Table 1 oligonucleotide. Especially preferably
15 when sequence identity to only a part of said Table 1 oligonucleotide is present, the sequence identity is high, e.g. at least 80% as described above.

Functionally equivalent oligonucleotides which satisfy the above stated functional requirements include
20 those which are derived from the Table 1 oligonucleotides and also those which have been modified by single or multiple nucleotide base (or equivalent) substitution, addition and/or deletion, but which nonetheless retain functional activity, e.g. bind to the
25 same target molecule as the Table 1 oligonucleotide or the Table 1 derived oligonucleotide from which they are further derived or modified. Preferably said modification is of from 1 to 50, e.g. from 10 to 30, preferably from 1 to 5 bases. Especially preferably
30 only minor modifications are present, e.g. variations in less than 10 bases, e.g. less than 5 base changes.

Within the meaning of "addition" equivalents are included oligonucleotides containing additional sequences which are complementary to the consecutive
35 stretch of bases on the target molecule to which the Table 1 oligonucleotide or the Table 1 derived oligonucleotide binds. Alternatively the addition may

- 12 -

comprise a different, unrelated sequence, which may for example confer a further property, e.g. to provide a means for immobilization such as a linker to bind the oligonucleotide probe to a solid support.

5 Particularly preferred are naturally occurring equivalents such as biological variants, e.g. allelic, geographical or allotypic variants, e.g. oligonucleotides which correspond to a genetic variant, for example as present in a different species.

10 Functional equivalents include oligonucleotides with modified bases, e.g. using non-naturally occurring bases. Such derivatives may be prepared during synthesis or by post production modification.

15 "Hybridizing" sequences which bind under conditions of low stringency are those which bind under non-stringent conditions (for example, 6x SSC/50% formamide at room temperature) and remain bound when washed under conditions of low stringency (2 X SSC, room temperature, more preferably 2 X SSC, 42°C). Hybridizing under high
20 stringency refers to the above conditions in which washing is performed at 2 X SSC, 65°C (where SSC = 0.15M NaCl, 0.015M sodium citrate, pH 7.2).

 "Sequence identity" as referred to herein refers to the value obtained when assessed using ClustalW
25 (Thompson et al., 1994, Nucl. Acids Res., 22, p4673-4680) with the following parameters:

Pairwise alignment parameters - Method: accurate,
Matrix: IUB, Gap open penalty: 15.00, Gap extension
penalty: 6.66;

30 Multiple alignment parameters - Matrix: IUB, Gap open penalty: 15.00, % identity for delay: 30, Negative matrix: no, Gap extension penalty: 6.66, DNA transitions weighting: 0.5.

 Sequence identity at a particular base is intended
35 to include identical bases which have simply been derivatized.

 The invention also extends to polypeptides encoded

- 13 -

by the mRNA sequence to which a Table 1 oligonucleotide or a Table 1 derived oligonucleotide binds. The invention further extends to antibodies which bind to any of said polypeptides.

5 As described above, conveniently said set of oligonucleotide probes may be immobilized on one or more solid supports. Single or preferably multiple copies of each unique probe are attached to said solid supports, e.g. 10 or more, e.g. at least 100 copies of each unique
10 probe are present.

One or more unique oligonucleotide probes may be associated with separate solid supports which together form a set of probes immobilized on multiple solid support, e.g. one or more unique probes may be
15 immobilized on multiple beads, membranes, filters, biochips etc. which together form a set of probes, which together form modules of the kit described hereinafter. The solid support of the different modules are conveniently physically associated although the signals
20 associated with each probe (generated as described hereinafter) must be separately determinable.

Alternatively, the probes may be immobilized on discrete portions of the same solid support, e.g. each unique oligonucleotide probe, e.g. in multiple copies,
25 may be immobilized to a distinct and discrete portion or region of a single filter or membrane, e.g. to generate an array.

A combination of such techniques may also be used, e.g. several solid supports may be used which each
30 immobilize several unique probes.

The expression "solid support" shall mean any solid material able to bind oligonucleotides by hydrophobic, ionic or covalent bridges.

"Immobilization" as used herein refers to
35 reversible or irreversible association of the probes to said solid support by virtue of such binding. If reversible, the probes remain associated with the solid

- 14 -

support for a time sufficient for methods of the invention to be carried out.

Numerous solid supports suitable as immobilizing moieties according to the invention, are well known in the art and widely described in the literature and generally speaking, the solid support may be any of the well-known supports or matrices which are currently widely used or proposed for immobilization, separation etc. in chemical or biochemical procedures. Such materials include, but are not limited to, any synthetic organic polymer such as polystyrene, polyvinylchloride, polyethylene; or nitrocellulose and cellulose acetate; or tosyl activated surfaces; or glass or nylon or any surface carrying a group suited for covalent coupling of nucleic acids. The immobilizing moieties may take the form of particles, sheets, gels, filters, membranes, microfibre strips, tubes or plates, fibres or capillaries, made for example of a polymeric material e.g. agarose, cellulose, alginate, teflon, latex or polystyrene or magnetic beads. Solid supports allowing the presentation of an array, preferably in a single dimension are preferred, e.g. sheets, filters, membranes, plates or biochips.

Attachment of the nucleic acid molecules to the solid support may be performed directly or indirectly. For example if a filter is used, attachment may be performed by UV-induced crosslinking. Alternatively, attachment may be performed indirectly by the use of an attachment moiety carried on the oligonucleotide probes and/or solid support. Thus for example, a pair of affinity binding partners may be used, such as avidin, streptavidin or biotin, DNA or DNA binding protein (e.g. either the lac I repressor protein or the lac operator sequence to which it binds), antibodies (which may be mono- or polyclonal), antibody fragments or the epitopes or haptens of antibodies. In these cases, one partner of the binding pair is attached to (or is inherently

- 15 -

part of) the solid support and the other partner is attached to (or is inherently part of) the nucleic acid molecules.

As used herein an "affinity binding pair" refers to two components which recognize and bind to one another specifically (ie. in preference to binding to other molecules). Such binding pairs when bound together form a complex.

Attachment of appropriate functional groups to the solid support may be performed by methods well known in the art, which include for example, attachment through hydroxyl, carboxyl, aldehyde or amino groups which may be provided by treating the solid support to provide suitable surface coatings. Solid supports presenting appropriate moieties for attachment of the binding partner may be produced by routine methods known in the art.

Attachment of appropriate functional groups to the oligonucleotide probes of the invention may be performed by ligation or introduced during synthesis or amplification, for example using primers carrying an appropriate moiety, such as biotin or a particular sequence for capture.

Conveniently, the set of probes described hereinbefore is provided in kit form.

Thus viewed from a further aspect the present invention provides a kit comprising a set of oligonucleotide probes as described hereinbefore immobilized on one or more solid supports.

Preferably, said probes are immobilized on a single solid support and each unique probe is attached to a different region of said solid support. However, when attached to multiple solid supports, said multiple solid supports form the modules which make up the kit. Especially preferably said solid support is a sheet, filter, membrane, plate or biochip.

Optionally the kit may also contain information

- 16 -

relating to the signals generated by normal or diseased samples (as discussed in more detail hereinafter in relation to the use of the kits), standardizing materials, e.g. mRNA or cDNA from normal and/or diseased samples for comparative purposes, labels for incorporation into cDNA, adapters for introducing nucleic acid sequences for amplification purposes, primers for amplification and/or appropriate enzymes, buffers and solutions. Optionally said kit may also contain a package insert describing how the method of the invention should be performed, optionally providing standard graphs, data or software for interpretation of results obtained when performing the invention.

The use of such kits to prepare a standard diagnostic gene transcript pattern as described hereinafter forms a further aspect of the invention.

The set of probes as described herein have various uses. Principally however they are used to assess the gene expression state of a test cell to provide information relating to the organism from which said cell is derived. Thus the probes are useful in diagnosing, identifying or monitoring a disease or condition or stage thereof in an organism.

Thus in a further aspect the invention provides the use of a set of oligonucleotide probes or a kit as described hereinbefore to determine the gene expression pattern of a cell which pattern reflects the level of gene expression of genes to which said oligonucleotide probes bind, comprising at least the steps of:

a) isolating mRNA from said cell, which may optionally be reverse transcribed to cDNA;

b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotide probes or a kit as defined herein; and

c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce said pattern.

The mRNA and cDNA as referred to in this method,

- 17 -

and the methods hereinafter, encompass derivatives or copies of said molecules, e.g. copies of such molecules such as those produced by amplification or the preparation of complementary strands, but which retain the identity of the mRNA sequence, ie. would hybridize to the direct transcript (or its complementary sequence) by virtue of precise complementarity, or sequence identity, over at least a region of said molecule. It will be appreciated that complementarity will not exist over the entire region where techniques have been used which may truncate the transcript or introduce new sequences, e.g. by primer amplification. For convenience, said mRNA or cDNA is preferably amplified prior to step b). As with the oligonucleotides described herein said molecules may be modified, e.g. by using non-natural bases during synthesis providing complementarity remains. Such molecules may also carry additional moieties such as signalling or immobilizing means.

The various steps involved in the method of preparing such a pattern are described in more detail hereinafter.

As used herein "gene expression" refers to transcription of a particular gene to produce a specific mRNA product (ie. a particular splicing product). The level of gene expression may be determined by assessing the level of transcribed mRNA molecules or cDNA molecules reverse transcribed from the mRNA molecules or products derived from those molecules, e.g. by amplification.

The "pattern" created by this technique refers to information which, for example, may be represented in tabular or graphical form and conveys information about the signal associated with two or more oligonucleotides. Preferably said pattern is expressed as an array of numbers relating to the expression level associated with each probe.

- 18 -

Preferably, said pattern is established using the following linear model:

$$y = Xb + f \quad \text{Equation 1}$$

wherein, **X** is the matrix of gene expression data and **y** is the response variable, **b** is the regression coefficient vector and **f** the estimated residual vector. Although many different methods can be used to establish the relationship provided in equation 1, especially preferably the partial Least Squares Regression (PLSR) method is used for establishing the relationship in equation 1.

The probes are thus used to generate a pattern which reflects the gene expression of a cell at the time of its isolation. The pattern of expression is characteristic of the circumstances under which that cells finds itself and depends on the influences to which the cell has been exposed. Thus, a characteristic gene transcript pattern standard or fingerprint (standard probe pattern) for cells from an individual with a particular disease or condition may be prepared and used for comparison to transcript patterns of test cells. This has clear applications in diagnosing, monitoring or identifying whether an organism is suffering from a particular disease, condition or stage thereof.

The standard pattern is prepared by determining the extent of binding of total mRNA (or cDNA or related product), from cells from a sample of one or more organisms with the disease or condition or stage thereof, to the probes. This reflects the level of transcripts which are present which correspond to each unique probe. The amount of nucleic acid material which binds to the different probes is assessed and this information together forms the gene transcript pattern standard of that disease or condition or stage thereof. Each such standard pattern is characteristic of the disease, condition or stage thereof.

- 19 -

In a further aspect therefore, the present invention provides a method of preparing a standard gene transcript pattern characteristic of a disease or condition or stage thereof in an organism comprising at least the steps of:

- a) isolating mRNA from the cells of a sample of one or more organisms having the disease or condition or stage thereof, which may optionally be reverse transcribed to cDNA;
- b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotides or a kit as described hereinbefore specific for said disease or condition or stage thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and
- c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce a characteristic pattern reflecting the level of gene expression of genes to which said oligonucleotides bind, in the sample with the disease, condition or stage thereof.

For convenience, said oligonucleotides are preferably immobilized on one or more solid supports.

The standard pattern for a great number of diseases or conditions and different stages thereof using particular probes may be accumulated in databases and be made available to laboratories on request.

"Disease" samples and organisms as referred to herein refer to organisms (or samples from the same) with an underlying pathological disturbance relative to a normal organism (or sample), in a symptomatic or asymptomatic organism, which may result, for example, from infection or an acquired or congenital genetic imperfection. Such organisms are known to have, or which exhibit, the disease or condition or stage thereof under study.

A "condition" refers to a state of the mind or body of an organism which has not occurred through disease,

- 20 -

e.g. the presence of an agent in the body such as a toxin, drug or pollutant, or pregnancy.

"Stages" thereof refer to different stages of the disease or condition which may or may not exhibit particular physiological or metabolic changes, but do exhibit changes at the genetic level which may be detected as altered gene expression. It will be appreciated that during the course of a disease or condition the expression of different transcripts may vary. Thus at different stages, altered expression may not be exhibited for particular transcripts compared to "normal" samples. However, combining information from several transcripts which exhibit altered expression at one or more stages through the course of the disease or condition can be used to provide a characteristic pattern which is indicative of a particular stage of the disease or condition. Thus for example different stages in cancer, e.g. pre-stage I, stage I, stage II, II or IV can be identified.

"Normal" as used herein refers to organisms or samples which are used for comparative purposes. Preferably, these are "normal" in the sense that they do not exhibit any indication of, or are not believed to have, any disease or condition that would affect gene expression, particularly in respect of the disease for which they are to be used as the normal standard. However, it will be appreciated that different stages of a disease or condition may be compared and in such cases, the "normal" sample may correspond to the earlier stage of the disease or condition.

As used herein a "sample" refers to any material obtained from the organism, e.g. human or non-human animal under investigation which contains cells and includes, tissues, body fluid or body waste or in the case of prokaryotic organisms, the organism itself. "Body fluids" include blood, saliva, spinal fluid, semen, lymph. "Body waste" includes urine, expectorated

- 21 -

matter (pulmonary patients), faeces etc. "Tissue samples" include tissue obtained by biopsy, by surgical interventions or by other means e.g. placenta. Preferably however, the samples which are examined are from areas of the body not apparently affected by the disease or condition. The cells in such samples are not disease cells, e.g. cancer cells, have not been in contact with such disease cells and do not originate from the site of the disease or condition. The "site of disease" is considered to be that area of the body which manifests the disease in a way which may be objectively determined, e.g. a tumour or area of inflammation. Thus for example peripheral blood may be used for the diagnosis of non-haematopoietic cancers, and the blood does not require the presence of malignant or disseminated cells from the cancer in the blood. Similarly in diseases of the brain, in which no diseased cells are found in the blood due to the blood:brain barrier, peripheral blood may still be used in the methods of the invention.

It will however be appreciated that the method of preparing the standard transcription pattern and other methods of the invention are also applicable for use on living parts of eukaryotic organisms such as cell lines and organ cultures and explants.

As used herein, reference to "corresponding" sample etc. refers to cells preferably from the same tissue, body fluid or body waste, but also includes cells from tissue, body fluid or body waste which are sufficiently similar for the purposes of preparing the standard or test pattern. When used in reference to genes "corresponding" to the probes, this refers to genes which are related by sequence (which may be complementary) to the probes although the probes may reflect different splicing products of expression.

"Assessing" as used herein refers to both quantitative and qualitative assessment which may be

- 22 -

determined in absolute or relative terms.

The invention may be put into practice as follows. To prepare a standard transcript pattern for a particular disease, condition or stage thereof, sample
5 mRNA is extracted from the cells of tissues, body fluid or body waste according to known techniques (see for example Sambrook et. al. (1989), Molecular Cloning : A laboratory manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.) from a
10 diseased individual or organism.

Owing to the difficulties in working with RNA, the RNA is preferably reverse transcribed at this stage to form first strand cDNA. Cloning of the cDNA or selection from, or using, a cDNA library is not however
15 necessary in this or other methods of the invention. Preferably, the complementary strands of the first strand cDNAs are synthesized, ie. second strand cDNAs, but this will depend on which relative strands are present in the oligonucleotide probes. The RNA may
20 however alternatively be used directly without reverse transcription and may be labelled if so required.

Preferably the cDNA strands are amplified by known amplification techniques such as the polymerase chain reaction (PCR) by the use of appropriate primers.
25 Alternatively, the cDNA strands may be cloned with a vector, used to transform a bacteria such as E. coli which may then be grown to multiply the nucleic acid molecules. When the sequence of the cDNAs are not known, primers may be directed to regions of the nucleic acid molecules which have been introduced. Thus for
30 example, adapters may be ligated to the cDNA molecules and primers directed to these portions for amplification of the cDNA molecules. Alternatively, in the case of eukaryotic samples, advantage may be taken of the polyA
35 tail and cap of the RNA to prepare appropriate primers.

To produce the standard diagnostic gene transcript pattern or fingerprint for a particular disease or

- 23 -

condition or stage thereof, the above described oligonucleotide probes are used to probe mRNA or cDNA of the diseased sample to produce a signal for hybridization to each particular oligonucleotide probe species, ie. each unique probe. A standard control gene transcript pattern may also be prepared if desired using mRNA or cDNA from a normal sample. Thus, mRNA or cDNA is brought into contact with the oligonucleotide probe under appropriate conditions to allow hybridization.

When multiple samples are probed, this may be performed consecutively using the same probes, e.g. on one or more solid supports, ie. on probe kit modules, or by simultaneously hybridizing to corresponding probes, e.g. the modules of a corresponding probe kit.

To identify when hybridization occurs and obtain an indication of the number of transcripts/cDNA molecules which become bound to the oligonucleotide probes, it is necessary to identify a signal produced when the transcripts (or related molecules) hybridize (e.g. by detection of double stranded nucleic acid molecules or detection of the number of molecules which become bound, after removing unbound molecules, e.g. by washing).

In order to achieve a signal, either or both components which hybridize (ie. the probe and the transcript) carry or form a signalling means or a part thereof. This "signalling means" is any moiety capable of direct or indirect detection by the generation or presence of a signal. The signal may be any detectable physical characteristic such as conferred by radiation emission, scattering or absorption properties, magnetic properties, or other physical properties such as charge, size or binding properties of existing molecules (e.g. labels) or molecules which may be generated (e.g. gas emission etc.). Techniques are preferred which allow signal amplification, e.g. which produce multiple signal events from a single active binding site, e.g. by the catalytic action of enzymes to produce multiple

- 24 -

detectable products.

Conveniently the signalling means may be a label which itself provides a detectable signal. Conveniently this may be achieved by the use of a radioactive or
5 other label which may be incorporated during cDNA production, the preparation of complementary cDNA strands, during amplification of the target mRNA/cDNA or added directly to target nucleic acid molecules.

Appropriate labels are those which directly or
10 indirectly allow detection or measurement of the presence of the transcripts/cDNA. Such labels include for example radiolabels, chemical labels, for example chromophores or fluorophores (e.g. dyes such as fluorescein and rhodamine), or reagents of high electron
15 density such as ferritin, haemocyanin or colloidal gold. Alternatively, the label may be an enzyme, for example peroxidase or alkaline phosphatase, wherein the presence of the enzyme is visualized by its interaction with a suitable entity, for example a substrate. The label may
20 also form part of a signalling pair wherein the other member of the pair is found on, or in close proximity to, the oligonucleotide probe to which the transcript/cDNA binds, for example, a fluorescent compound and a quench fluorescent substrate may be used.
25 A label may also be provided on a different entity, such as an antibody, which recognizes a peptide moiety attached to the transcripts/cDNA, for example attached to a base used during synthesis or amplification.

A signal may be achieved by the introduction of a
30 label before, during or after the hybridization step. Alternatively, the presence of hybridizing transcripts may be identified by other physical properties, such as their absorbance, and in which case the signalling means is the complex itself.

35 The amount of signal associated with each oligonucleotide probe is then assessed. The assessment may be quantitative or qualitative and may be based on

- 25 -

binding of a single transcript species (or related cDNA or other products) to each probe, or binding of multiple transcript species to multiple copies of each unique probe. It will be appreciated that quantitative results will provide further information for the transcript fingerprint of the disease which is compiled. This data may be expressed as absolute values (in the case of macroarrays) or may be determined relative to a particular standard or reference e.g. a normal control sample.

Furthermore it will be appreciated that the standard diagnostic gene pattern transcript may be prepared using one or more disease samples (and normal samples if used) to perform the hybridization step to obtain patterns not biased towards a particular individual's variations in gene expression.

The use of the probes to prepare standard patterns and the standard diagnostic gene transcript patterns thus produced for the purpose of identification or diagnosis or monitoring of a particular disease or condition or stage thereof in a particular organism forms a further aspect of the invention.

Once a standard diagnostic fingerprint or pattern has been determined for a particular disease or condition using the selected oligonucleotide probes, this information can be used to identify the presence, absence or extent or stage of that disease or condition in a different test organism or individual.

To examine the gene expression pattern of a test sample, a test sample of tissue, body fluid or body waste containing cells, corresponding to the sample used for the preparation of the standard pattern, is obtained from a patient or the organism to be studied. A test gene transcript pattern is then prepared as described hereinbefore as for the standard pattern.

In a further aspect therefore, the present invention provides a method of preparing a test gene

- 26 -

transcript pattern comprising at least the steps of:

a) isolating mRNA from the cells of a sample of said test organism, which may optionally be reverse transcribed to cDNA;

5 b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotides or a kit as described hereinbefore specific for a disease or condition or stage thereof in an organism and sample thereof corresponding to the organism and sample thereof under
10 investigation; and

c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce said pattern reflecting the level of gene expression of genes to which said oligonucleotides bind, in said test sample.

15 This test pattern may then be compared to one or more standard patterns to assess whether the sample contains cells having the disease, condition or stage thereof.

Thus viewed from a further aspect the present
20 invention provides a method of diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism, comprising the steps of:

a) isolating mRNA from the cells of a sample of said organism, which may optionally be reverse
25 transcribed to cDNA;

b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotides or a kit as described hereinbefore specific for said disease or
30 condition or stage thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation;

c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce a characteristic pattern reflecting the level
35 of gene expression of genes to which said oligonucleotides bind, in said sample; and

d) comparing said pattern to a standard

- 27 -

5 diagnostic pattern prepared according to the method of the invention using a sample from an organism corresponding to the organism and sample under investigation to determine the presence of said disease or condition or a stage thereof in the organism under investigation.

The method up to and including step c) is the preparation of a test pattern as described above.

10 As referred to herein, "diagnosis" refers to determination of the presence or existence of a disease or condition or stage thereof in an organism.

"Monitoring" refers to establishing the extent of a disease or condition, particularly when an individual is known to be suffering from a disease or condition, for example to monitor the effects of treatment or the development of a disease or condition, e.g. to determine the suitability of a treatment or provide a prognosis.

15 The presence of the disease or condition or stage thereof may be determined by determining the degree of correlation between the standard and test samples' patterns. This necessarily takes into account the range of values which are obtained for normal and diseased samples. Although this can be established by obtaining standard deviations for several representative samples binding to the probes to develop the standard, it will be appreciated that single samples may be sufficient to generate the standard pattern to identify a disease if the test sample exhibits close enough correlation to that standard. Conveniently, the presence, absence, or extent of a disease or condition or stage thereof in a test sample can be predicted by inserting the data relating to the expression level of informative probes in test sample into the standard diagnostic probe pattern established according to equation 1.

35 Data generated using the above mentioned methods may be analysed using various techniques from the most

- 28 -

basic visual representation (e.g. relating to intensity) to more complex data manipulation to identify underlying patterns which reflect the interrelationship of the level of expression of each gene to which the various probes bind, which may be quantified and expressed mathematically. Conveniently, the raw data thus generated may be manipulated by the data processing and statistical methods described hereinafter, particularly normalizing and standardizing the data and fitting the data to a classification model to determine whether said test data reflects the pattern of a particular disease, condition or stage thereof.

The methods described herein may be used to identify, monitor or diagnose a disease, condition or ailment or its stage or progression, for which the oligonucleotide probes are informative. "Informative" probes as described herein, are those which reflect genes which have altered expression in the diseases or conditions in question, or particular stages thereof. Probes of the invention may not be sufficiently informative for diagnostic purposes when used alone, but are informative when used as one of several probes to provide a characteristic pattern, e.g. in a set as described hereinbefore.

Preferably said probes correspond to genes which are systemically affected by said disease, condition or stage thereof. Especially preferably said genes, from which transcripts are derived which bind to probes of the invention, are metabolic or house-keeping genes and preferably are moderately or highly expressed. The advantage of using probes directed to moderately or highly expressed genes is that smaller clinical samples are required for generating the necessary gene expression data set, e.g. less than 1ml blood samples.

Furthermore, it has been found that such genes which are already being actively transcribed tend to be more prone to being influenced, in a positive or

- 29 -

negative way, by new stimuli. In addition, since transcripts are already being produced at levels which are generally detectable, small changes in those levels are readily detectable as for example, a certain
5 detectable threshold does not need to be reached.

In preferred methods of the invention, the set of probes of the invention are informative for a variety of different diseases, conditions or stages thereof. A sub-set of the probes disclosed herein may be used for
10 diagnosis, identification or monitoring a particular disease, condition or stage thereof.

Thus the probes may be used to diagnose or identify or monitor any condition, ailment, disease or reaction that leads to the relative increase or decrease in the
15 activity of informative genes of any or all eukaryotic or prokaryotic organisms regardless of whether these changes have been caused by the influence of bacteria, virus, prions, parasites, fungi, radiation, natural or artificial toxins, drugs or allergens, including mental
20 conditions due to stress, neurosis, psychosis or deteriorations due to the ageing of the organism, and conditions or diseases of unknown cause, providing a sub-set of the probes as described herein are informative for said disease or condition or stage
25 thereof.

Such diseases include those which result in metabolic or physiological changes, such as fever-associated diseases such as influenza or malaria. Other diseases which may be detected include for example
30 yellow fever, sexually transmitted diseases such as gonorrhea, fibromyalgia, candida-related complex, cancer (for example of the stomach, lung, breast, prostate gland, bowel, skin, colon, ovary etc), Alzheimer's disease, disease caused by retroviruses such as HIV,
35 senile dementia, multiple sclerosis and Creutzfeldt-Jakob disease to mention a few.

The invention may also be used to identify patients

- 30 -

with psychiatric or psychosomatic diseases such as schizophrenia and eating disorders. Of particular importance is the use of this method to detect diseases, conditions, or stages thereof, which are not readily detectable by known diagnostic methods, such as HIV which is generally not detectable using known techniques 1 to 4 months following infection. Conditions which may be identified include for example drug abuse, such as the use of narcotics, alcohol, steroids or performance enhancing drugs.

Preferably said disease to be identified or monitored is a cancer or a degenerative brain disorder (such as Alzheimer's or Parkinson's disease).

In particular, a set of oligonucleotide probes, wherein said set comprises at least 10 oligonucleotides selected from:

an oligonucleotide as described in Table 4 or an oligonucleotide derived therefrom or an oligonucleotide with a complementary sequence, or a functionally equivalent oligonucleotide, may be used for diagnosis or identification or monitoring the progression of Alzheimer's disease. Similarly Table 2 probes and Table 2 derived probes and their functional equivalents may be used to diagnose, identify or monitor the progression of breast cancer. Especially preferably the probes used for breast cancer analysis are selected based on their occurrence as set forth in Table 3 and as described hereinbefore.

The diagnostic method may be used alone as an alternative to other diagnostic techniques or in addition to such techniques. For example, methods of the invention may be used as an alternative or additive diagnostic measure to diagnosis using imaging techniques such as Magnetic Resonance Image (MRI), ultrasound imaging, nuclear imaging or X-ray imaging, for example in the identification and/or diagnosis of tumours.

The methods of the invention may be performed on

- 31 -

cells from prokaryotic or eukaryotic organisms which may be any eukaryotic organisms such as human beings, other mammals and animals, birds, insects, fish and plants, and any prokaryotic organism such as a bacteria.

5 Preferred non-human animals on which the methods of the invention may be conducted include, but are not limited to mammals, particularly primates, domestic animals, livestock and laboratory animals. Thus preferred animals for diagnosis include mice, rats,
10 guinea pigs, cats, dogs, pigs, cows, goats, sheep, horses. Particularly preferably the disease state or condition of humans is diagnosed, identified or monitored.

As described above, the sample under study may be
15 any convenient sample which may be obtained from an organism. Preferably however, as mentioned above, the sample is obtained from a site distant to the site of disease and the cells in such samples are not disease cells, have not been in contact with such cells and do
20 not originate from the site of the disease or condition. In such cases, although preferably absent, the sample may contain cells which do not fulfil these criteria. However, since the probes of the invention are concerned with transcripts whose expression is altered in cells
25 which do satisfy these criteria, the probes are specifically directed to detecting changes in transcript levels in those cells even if in the presence of other, background cells.

It has been found that the cells from such samples
30 show significant and informative variations in the gene expression of a large number of genes. Thus, the same probe (or several probes) may be found to be informative in determinations regarding two or more diseases, conditions or stages thereof by virtue of the particular
35 level of transcripts binding to that probe or the interrelationship of the extent of binding to that probe relative to other probes. As a consequence, it is

- 32 -

possible to use a relatively small number of probes for screening for multiple disorders or diseases. This has consequences with regard to the selection of probes, discussed in relation to random identification of probes hereinafter, but also for the use of a single set of probes for more than one diagnosis. Table 9 which represents preferred probes of the invention discloses probes which are informative for both Alzheimer's and breast cancer.

Thus, the present invention also provides sets of probes for diagnosing, identifying or monitoring two or more diseases, conditions or stages thereof, wherein at least one of said probes is suitable for said diagnosing, identifying or monitoring at least two of said diseases, conditions or stages thereof, and kits and methods of using the same. Preferably at least 5 probes, e.g. from 5 to 15 probes, are used in at least two diagnoses.

Thus, in a further preferred aspect, the present invention provides a method of diagnosis or identification or monitoring as described hereinbefore for the diagnosis, identification or monitoring of two or more diseases, conditions or stages thereof in an organism, wherein said test pattern produced in step c) of the diagnostic method is compared in step d) to at least two standard diagnostic patterns prepared as described previously, wherein each standard diagnostic pattern is a pattern generated for a different disease or condition or stage thereof.

Whilst in a preferred aspect the methods of assessment concern the development of a gene transcript pattern from a test sample and comparison of the same to a standard pattern, the elevation or depression of expression of certain markers may also be examined by examining the products of expression and the level of those products. Thus a standard pattern in relation to the expressed product may be generated:

- 33 -

In such methods the levels of expression of a set of polypeptides encoded by the gene to which an oligonucleotide of Table 1 or a Table 1 derived oligonucleotide, binds, are analysed.

5 Various diagnostic methods may be used to assess the amount of polypeptides (or fragments thereof) which are present. The presence or concentration of polypeptides may be examined, for example by the use of a binding partner to said polypeptide (e.g. an
10 antibody), which may be immobilized, to separate said polypeptide from the sample and the amount of polypeptide may then be determined.

"Fragments" of the polypeptides refers to a domain or region of said polypeptide, e.g. an antigenic
15 fragment, which is recognizable as being derived from said polypeptide to allow binding of a specific binding partner. Preferably such a fragment comprises a significant portion of said polypeptide and corresponds to a product of normal post-synthesis processing.

20 Thus in a further aspect the present invention provides a method of preparing a standard gene transcript pattern characteristic of a disease or condition or stage thereof in an organism comprising at least the steps of:

25 a) releasing target polypeptides from a sample of one or more organisms having the disease or condition or stage thereof;

b) contacting said target polypeptides with one or more binding partners, wherein each binding partner is
30 specific to a marker polypeptide (or a fragment thereof) encoded by the gene to which an oligonucleotide of Table 1 (or derived from a sequence described in Table 1) binds, to allow binding of said binding partners to said target polypeptides, wherein said marker polypeptides
35 are specific for said disease or condition thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and

- 34 -

c) assessing the target polypeptide binding to said binding partners to produce a characteristic pattern reflecting the level of gene expression of genes which express said marker polypeptides, in the sample with the disease, condition or stage thereof.

As used herein "target polypeptides" refer to those polypeptides present in a sample which are to be detected and "marker polypeptides" are polypeptides which are encoded by the genes to which Table 1 oligonucleotides or Table 1 derived oligonucleotides bind. The target and marker polypeptides are identical or at least have areas of high similarity, e.g. epitopic regions to allow recognition and binding of the binding partner.

"Release" of the target polypeptides refers to appropriate treatment of a sample to provide the polypeptides in a form accessible for binding of the binding partners, e.g. by lysis of cells where these are present. The samples used in this case need not necessarily comprise cells as the target polypeptides may be released from cells into the surrounding tissue or fluid, and this tissue or fluid may be analysed, e.g. urine or blood. Preferably however the preferred samples as described herein are used. "Binding partners" comprise the separate entities which together make an affinity binding pair as described above, wherein one partner of the binding pair is the target or marker polypeptide and the other partner binds specifically to that polypeptide, e.g. an antibody.

Various arrangements may be envisaged for detecting the amount of binding pairs which form. In its simplest form, a sandwich type assay e.g. an immunoassay such as an ELISA, may be used in which an antibody specific to the polypeptide and carrying a label (as described elsewhere herein) may be bound to the binding pair (e.g. the first antibody:polypeptide pair) and the amount of label detected.

- 35 -

Other methods as described herein may be similarly modified for analysis of the protein product of expression rather than the gene transcript and related nucleic acid molecules.

5 Thus a further aspect of the invention provides a method of preparing a test gene transcript pattern comprising at least the steps of:

- a) releasing target polypeptides from a sample of said test organism;
- 10 b) contacting said target polypeptides with one or more binding partners, wherein each binding partner is specific to a marker polypeptide (or a fragment thereof) encoded by the gene to which an oligonucleotide of Table 1 (or derived from a sequence described in Table 1)
- 15 binds, to allow binding of said binding partners to said target polypeptides, wherein said marker polypeptides are specific for said disease or condition thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and
- 20 c) assessing the target polypeptide binding to said binding partners to produce a characteristic pattern reflecting the level of gene expression of genes which express said marker polypeptides, in said test sample.

A yet further aspect of the invention provides a
25 method of diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism comprising the steps of:

- a) releasing target polypeptides from a sample of said organism;
- 30 b) contacting said target polypeptides with one or more binding partners, wherein each binding partner is specific to a marker polypeptide (or a fragment thereof) encoded by the gene to which an oligonucleotide of Table 1 (or derived from a sequence described in Table 1)
- 35 binds, to allow binding of said binding partners to said target polypeptides, wherein said marker polypeptides are specific for said disease or condition thereof in an

- 36 -

organism and sample thereof corresponding to the
organism and sample thereof under investigation; and

5 c) assessing the target polypeptide binding to said
binding partners to produce a characteristic pattern
reflecting the level of gene expression of genes which
express said marker polypeptides in said sample; and

10 d) comparing said pattern to a standard diagnostic
pattern prepared as described hereinbefore using a
sample from an organism corresponding to the organism
and sample under investigation to determine the degree
of correlation indicative of the presence of said
disease or condition or a stage thereof in the organism
under investigation.

15 The methods of generating standard and test
patterns and diagnostic techniques rely on the use of
informative oligonucleotide probes to generate the gene
expression data. In some cases it will be necessary to
select these informative probes for a particular method,
e.g. to diagnose a particular disease, from a selection
20 of available probes, e.g. the probes described
hereinbefore (the Table 1 oligonucleotides, the Table 1
derived oligonucleotides, their complementary sequences
and functionally equivalent oligonucleotides). The
following methodology describes a convenient method for
25 identifying such informative probes, or more
particularly how to select a suitable sub-set of probes
from the probes described herein.

30 Probes for the analysis of a particular disease or
condition or stage thereof, may be identified in a
number of ways known in the prior art, including by
differential expression or by library subtraction (see
for example WO98/49342). As described hereinafter, in
view of the high information content of most
transcripts, as a starting point one may also simply
35 analyse a random sub-set of mRNA or cDNA species and
pick the most informative probes from that sub-set. The
following method describes the use of immobilized

- 37 -

oligonucleotide probes (e.g. the probes of the invention) to which mRNA (or related molecules) from different samples is bound to identify which probes are the most informative to identify a particular type of sample, e.g. a disease sample.

The immobilized probes can be derived from various unrelated or related organisms; the only requirement is that the immobilized probes should bind specifically to their homologous counterparts in test organisms. Probes can also be derived from commercially available or public databases and immobilized on solid supports or, as mentioned above, they can be randomly picked and isolated from a cDNA library and immobilized on a solid support.

The length of the probes immobilised on the solid support should be long enough to allow for specific binding to the target sequences. The immobilised probes can be in the form of DNA, RNA or their modified products or PNAs (peptide nucleic acids). Preferably, the probes immobilised should bind specifically to their homologous counterparts representing highly and moderately expressed genes in test organisms. Conveniently the probes which are used are the probes described herein.

The gene expression pattern of cells in biological samples can be generated using prior art techniques such as microarray or macroarray as described below or using methods described herein. Several technologies have now been developed for monitoring the expression level of a large number of genes simultaneously in biological samples, such as, high-density oligoarrays (Lockhart et al., 1996, Nat. Biotech., 14, p1675-1680), cDNA microarrays (Schena et al., 1995, Science, 270, p467-470) and cDNA macroarrays (Maier E et al., 1994, Nucl. Acids Res., 22, p3423-3424; Bernard et al., 1996, Nucl. Acids Res., 24, p1435-1442).

In high-density oligoarrays and cDNA microarrays,

- 38 -

hundreds and thousands of probe oligonucleotides or cDNAs, are spotted onto glass slides or nylon membranes, or synthesized on biochips. The mRNA isolated from the test and reference samples are labelled by reverse transcription with a red or green fluorescent dye, mixed, and hybridised to the microarray. After washing, the bound fluorescent dyes are detected by a laser, producing two images, one for each dye. The resulting ratio of the red and green spots on the two images provides the information about the changes in expression levels of genes in the test and reference samples. Alternatively, single channel or multiple channel microarray studies can also be performed.

In cDNA macroarray, different cDNAs are spotted on a solid support such as nylon membranes in excess in relation to the amount of test mRNA that can hybridise to each spot. mRNA isolated from test samples is radio-labelled by reverse transcription and hybridised to the immobilised probe cDNA. After washing, the signals associated with labels hybridising specifically to immobilised probe cDNA are detected and quantified. The data obtained in macroarray contains information about the relative levels of transcripts present in the test samples. Whilst macroarrays are only suitable to monitor the expression of a limited number of genes, microarrays can be used to monitor the expression of several thousand genes simultaneously and is, therefore, a preferred choice for large-scale gene expression studies.

A macroarray technique for generating the gene expression data set has been used to illustrate the probe identification method described herein. For this purpose, mRNA is isolated from samples of interest and used to prepare labelled target molecules, e.g. mRNA or cDNA as described above. The labelled target molecules are then hybridised to probes immobilised on the solid support. Various solid supports can be used for the

- 39 -

purpose, as described previously. Following hybridization, unbound target molecules are removed and signals from target molecules hybridizing to immobilised probes quantified. If radio labelling is performed, PhosphoImager can be used to generate an image file that can be used to generate a raw data set. Depending on the nature of label chosen for labelling the target molecules, other instruments can also be used, for example, when fluorescence is used for labelling, a FluoroImager can be used to generate an image file from the hybridised target molecules.

The raw data corresponding to mean intensity, median intensity, or volume of the signals in each spot can be acquired from the image file using commercially available software for image analysis. However, the acquired data needs to be corrected for background signals and normalized prior to analysis, since, several factors can affect the quality and quantity of the hybridising signals. For example, variations in the quality and quantity of mRNA isolated from sample to sample, subtle variations in the efficiency of labelling target molecules during each reaction, and variations in the amount of unspecific binding between different macroarrays can all contribute to noise in the acquired data set that must be corrected for prior to analysis.

Background correction can be performed in several ways. The lowest pixel intensity within a spot can be used for background subtraction or the mean or median of the line of pixels around the spots' outline can be used for the purpose. One can also define an area representing the background intensity based on the signals generated from negative controls and use the average intensity of this area for background subtraction.

The background corrected data can then be transformed for stabilizing the variance in the data structure and normalized for the differences in probe

- 40 -

intensity. Several transformation techniques have been described in the literature and a brief overview can be found in Cui, Kerr and Churchill
<http://www.jax.org/research/churchill/research/expression/Cui-Transform.pdf>). Normalization can be performed by dividing the intensity of each spot with the collective intensity, average intensity or median intensity of all the spots in a macroarray or a group of spots in a macroarray in order to obtain the relative intensity of signals hybridising to immobilised probes in a macroarray. Several methods have been described for normalizing gene expression data (Richmond and Somerville, 2000, Current Opin. Plant Biol., 3, p108-116; Finkelstein et al., 2001, In "Methods of Microarray Data Analysis. Papers from CAMDA, Eds. Lin & Johnson, Kluwer Academic, p57-68; Yang et al., 2001, In "Optical Technologies and Informatics", Eds. Bittner, Chen, Dorsel & Dougherty, Proceedings of SPIE, 4266, p141-152; Dudoit et al, 2000, J. Am. Stat. Ass., 97, p77-87; Alter et al 2000, supra; Newton et al., 2001, J. Comp. Biol., 8, p37-52). Generally, a scaling factor or function is first calculated to correct the intensity effect and then used for normalising the intensities. The use of external controls has also been suggested for improved normalization.

One other major challenge encountered in large-scale gene expression analysis is that of standardization of data collected from experiments performed at different times. We have observed that gene expression data for samples acquired in the same experiment can be efficiently compared following background correction and normalization. However, the data from samples acquired in experiments performed at different times requires further standardization prior to analysis. This is because subtle differences in experimental parameters between different experiments, for example, differences in the quality and quantity of

- 41 -

mRNA extracted at different times, differences in time used for target molecule labelling, hybridization time or exposure time, can affect the measured values. Also, factors such as the nature of the sequence of transcripts under investigation (their GC content) and their amount in relation to the each other determines how they are affected by subtle variations in the experimental processes. They determine, for example, how efficiently first strand cDNAs, corresponding to a particular transcript, are transcribed and labelled during first strand synthesis, or how efficiently the corresponding labelled target molecules bind to their complementary sequences during hybridization. Batch to batch difference in the printing process is also a major factor for variation in the generated expression data.

Failure to properly address and rectify for these influences leads to situations where the differences between the experimental series may overshadow the main information of interest contained in the gene expression data set, i.e. the differences within the combined data from the different experimental series. Figure 1 provides one such example showing a classification based on Principal Component Analysis (PCA) of combined data from two experimental series where the main goal is to distinguish between Alzheimer/non-Alzheimer patients.

PCA (also known as singular value decomposition) is a technique for studying interdependencies and underlying relationships of a set of variables. The data are modelled in terms of a few significant factors or principal components (PC's), plus residuals. The PC's contain the main phenomena and define the systematic variability present in the data, while the residuals represent the variability interpreted as noise. Details on PCA can be found in Jolliffe (1986, Principal Component Analysis, Springer-Verlag, NY), and Jackson (1991, A User's Guide to Principal Components, Wiley, NY). The results of Figure 1 show that two

- 42 -

clusters are formed representing the data from two experimental series rather than the Alzheimer/non-Alzheimer differentiation. There were eight samples in common between the two series of experiments, which ideally should have fallen on top of, or in near proximity to, each other if appropriately standardized.

We have now found that gene expression data between different experiments can be efficiently standardized by including a subset of samples from one experimental series in the next experimental series and using a direct standardization method (DS), originally described by Wang and Kowalski (Anal. Chem., 1991, 63, p2750 and J. Chemometrics, 1991, 5, p129-145). Although the method of DS is well known in the field of analytical chemistry, it remains undescribed and unused in the field of gene expression data analysis.

In DS, the secondary data representing for example experimental series 2 (secondary measurements, R_2) are corrected to match the data measured on the primary measurements representing data from series 1 (R_1), while the calibration model remains unchanged. In DS, response matrices for both experimental series are related to each other by a transformation matrix F , i.e.

$$R_1 = R_2 F \quad (1)$$

Where F is a square matrix dimensioned gene by gene. From (1), the transformation matrix is calculated as:

$$F = R_2^+ R_1 \quad (2)$$

The transformation matrix F in equation (2) is calculated using a relatively small subset of samples which are measured on both the master primary and the secondary series of data.

Finally, the response of the unknown sample measured on the secondary series $r_{2,un}^T$, is standardized

- 43 -

to the response vector $\hat{r}_{1,un}^T$ expected from the primary series

5

$$\hat{r}_{1,un}^T = r_{21,un}^T \hat{F} \quad (3)$$

From the preceding equation it can be seen that the column i of the transformation matrix contains the multiplication factors for a set of genes measured in the secondary series to obtain the intensity at spot i of the corrected series.

The number of samples that are repeated in the experimental series, R_1 and R_2 , should be equal to their ranks, which in this case is equal to the number of principal components retained for explaining the variation in the R_1 and R_2 . For example, if three principal components are retained for explaining the variation in the data set, a minimum of three samples should be repeated between R_1 and R_2 . The samples that should be repeated between different series should ideally be those that exhibit high leverages in the gene expression pattern. At times, two samples may suffice, while at other times, more than two samples should be ideally be included for good representativity. In some cases, the samples selected can be the same in all the experimental series to be compared (reference samples), while in other cases, representative samples can be selected sequentially by analyzing the expression pattern after each experiment. The selected samples with high leverages are then included in the next experimental series. The results of using Direct Standardization are shown in Figure 1.

Another approach for normalizing and standardizing the gene expression data set is to hybridize each DNA array with target molecules prepared from a test sample and an equal amount of labelled target molecules prepared from representative reference samples. In order to measure the intensity of labelled target

- 44 -

molecules hybridizing to the immobilized probes it is necessary that the labelled molecules are prepared from test and reference samples using different labels, for example, different fluorescent dyes can be used for preparing the labelled material. The labelled molecules prepared from reference samples can be added to the hybridization solution together with the labelled material prepared from test samples. A data file from each array representing the expression pattern of different genes in the test sample and reference samples can then be obtained, normalized and standardized by the direct standardization method as described above. An instant advantage of including the differentially labelled target molecules from reference samples during hybridization is that it enables an efficient comparison of new test samples to the data sets already stored in a database.

Monitoring the expression of a large number of genes in several samples leads to the generation of a large amount of data that is too complex to be easily interpreted. Several unsupervised and supervised multivariate data analysis techniques have already been shown to be useful in extracting meaningful biological information from these large data sets. Cluster analysis is by far the most commonly used technique for gene expression analysis, and has been performed to identify genes that are regulated in a similar manner, and or identifying new/unknown tumour classes using gene expression profiles (Eisen et al., 1998, PNAS, 95, p14863-14868, Alizadeh et al. 2000, supra, Perou et al. 2000, Nature, 406, p747-752; Ross et al, 2000, Nature Genetics, 24(3), p227-235; Herwig et al., 1999, Genome Res., 9, p1093-1105; Tamayo et al, 1999, Science, PNAS, 96, p2907-2912).

In the clustering method, genes are grouped into functional categories (clusters) based on their expression profile, satisfying two criteria: *homogeneity*

- 45 -

- the genes in the same cluster are highly similar in expression to each other; and *separation* - genes in different clusters have low similarity in expression to each other.

5 Examples of various clustering techniques that have been used for gene expression analysis include hierarchical clustering (Eisen et al., 1998, supra; Alizadeh et al. 2000, supra; Perou et al. 2000, supra; Ross et al, 2000, supra), K-means clustering (Herwig et
10 al., 1999, supra; Tavazoie et al, 1999, Nature Genetics, 22(3), p. 281-285), gene shaving (Hastie et al., 2000, Genome Biology, 1(2), research 0003.1-0003.21), block clustering (Tibshirani et al., 1999, Tech report Univ Stanford.) Plaid model (Lazzeroni, 2002, Stat. Sinica,
15 12, p61-86), and self-organizing maps (Tamayo et al. 1999, supra). Also, related methods of multivariate statistical analysis, such as those using the singular value decomposition (Alter et al., 2000, PNAS, 97(18), p10101-10106; Ross et al. 2000, supra) or
20 multidimensional scaling can be effective at reducing the dimensions of the objects under study.

 However, methods such as cluster analysis and singular value decomposition are purely exploratory and only provide a broad overview of the internal structure
25 present in the data. They are unsupervised approaches in which the available information concerning the nature of the class under investigation is not used in the analysis. Often, the nature of the biological perturbation to which a particular sample has been
30 subjected is known. For example, it is sometimes known whether the sample whose gene expression pattern is being analysed derives from a diseased or healthy individual. In such instances, discriminant analysis can be used for classifying samples into various groups
35 based on their gene expression data.

 In such an analysis one builds the classifier by training the data that is capable of discriminating

- 46 -

between member and non-members of a given class. The trained classifier can then be used to predict the class of unknown samples. Examples of discrimination methods that have been described in the literature include

5 Support Vector Machines (Brown et al, 2000, PNAS, 97, p262-267), Nearest Neighbour (Dudoit et al., 2000, supra), Classification trees (Dudoit et al., 2000, supra), Voted classification (Dudoit et al., 2000, supra), Weighted Gene voting (Golub et al. 1999, supra),

10 and Bayesian classification (Keller et al. 2000, Tec report Univ of Washington). Also a technique in which PLS (Partial Least Square) regression analysis is first used to reduce the dimensions in the gene expression data set followed by classification using logistic

15 discriminant analysis and quadratic discriminant analysis (LD and QDA) has recently been described (Nguyen & Rocke, 2002, Bioinformatics, 18, p39-50 and 1216-1226).

A challenge that gene expression data poses to

20 classical discriminatory methods is that the number of genes whose expression are being analysed is very large compared to the number of samples being analysed. However in most cases only a small fraction of these genes are informative in discriminant analysis problems.

25 Moreover, there is a danger that the noise from irrelevant genes can mask or distort the information from the informative genes. Several methods have been suggested in literature to identify and select genes that are informative in microarray studies, for example,

30 t-statistics (Dudoit et al, 2002, J. Am. Stat. Ass., 97, p77-87), analysis of variance (Kerr et al., 2000, PNAS, 98, p8961-8965), Neighbourhood analysis (Golub et al, 1999, supra), Ratio of between groups to within groups sum of squares (Dudoit et al., 2002, supra), Non

35 parametric scoring (Park et al., 2002, Pacific Symposium on Biocomputing, p52-63) and Likelihood selection (Keller et al., 2000, supra).

- 47 -

In the methods described herein the gene expression data that has been normalized and standardized is analysed by using Partial Least Squares Regression (PLSR). Although PLSR is primarily a method used for regression analysis of continuous data (see Appendix A), it can also be utilized as a method for model building and discriminant analysis using a dummy response matrix based on a binary coding. The class assignment is based on a simple dichotomous distinction such as breast cancer (class 1) / healthy (class 2), or a multiple distinction based on multiple disease diagnosis such as breast cancer (class 1) / Alzheimer (class 2) / healthy (class 3). The list of diseases for classification can be increased depending upon the samples available corresponding to other diseases or conditions or stages thereof.

PLSR applied as a classification method is referred to as PLS-DA (DA standing for Discriminant analysis). PLS-DA is an extension of the PLSR algorithm in which the Y-matrix is a dummy matrix containing n rows (corresponding to the number of samples) and K columns (corresponding to the number of classes). The Y-matrix is constructed by inserting 1 in the k th column and -1 in all the other columns if the corresponding i th object of X belongs to class k . By regressing Y onto X , classification of a new sample is achieved by selecting the group corresponding to the largest component of the fitted, $\hat{y}(x) = (\hat{y}_1(x), \hat{y}_2(x), \dots, \hat{y}_k(x))$. Thus, in a -1/1 response matrix, a prediction value below 0 means that the sample belongs to the class designated as -1, while a prediction value above 0 implies that the sample belongs to the class designated as 1.

An advantage of PLSR-DA is that the results obtained can be easily represented in the form of two different plots, the score and loading plots. Score plots represent a projection of the samples onto the principal components and shows the distribution of the

- 48 -

samples in the classification model and their relationship to one another. Loading plots display correlations between the variables present in the data set.

5 It is usually recommended to use PLS-DA as a starting point for the classification problem due to its ability to handle collinear data, and the property of PLSR as a dimension reduction technique. Once this purpose has been satisfied, it is possible to use other
10 methods such as Linear discriminant analysis, LDA, that has been shown to be effective in extracting further information, Indahl et al. (1999, Chem. and Intell. Lab. Syst., 49, p19-31). This approach is based on first decomposing the data using PLS-DA, and then using the
15 scores vectors (instead of the original variables) as input to LDA. Further details on LDA can be found in Duda and Hart (Classification and Scene Analysis, 1973, Wiley, USA).

 The next step following model building is of model
20 validation. This step is considered to be amongst the most important aspects of multivariate analysis, and tests the "goodness" of the calibration model which has been built. In this work, a cross validation approach has been used for validation. In this approach, one or
25 a few samples are kept out in each segment while the model is built using a full cross-validation on the basis of the remaining data. The samples left out are then used for prediction/classification. Repeating the simple cross-validation process several times holding
30 different samples out for each cross-validation leads to a so-called double cross-validation procedure. This approach has been shown to work well with a limited amount of data, as is the case in some of the Examples described here. Also, since the cross validation step
35 is repeated several times the dangers of model bias and overfitting are reduced.

 Once a calibration model has been built and

- 49 -

validated, genes exhibiting an expression pattern that is most relevant for describing the desired information in the model can be selected by techniques described in the prior art for variable selection, as mentioned
 5 elsewhere. Variable selection will help in reducing the final model complexity, provide a parsimonious model, and thus lead to a reliable model that can be used for prediction. Moreover, use of fewer genes for the purpose of providing diagnosis will reduce the cost of
 10 the diagnostic product. In this way informative probes which would bind to the genes of relevance may be identified.

We have found that after a calibration model has been built, statistical techniques like Jackknife
 15 (Effron, 1982, The Jackknife, the Bootstrap and other resampling plans. Society for Industrial and Applied mathematics, Philadelphia, USA), based on resampling methodology, can be efficiently used to select or confirm significant variables (informative probes).

20 The approximate uncertainty variance of the PLS regression coefficients B can be estimated by:

$$S^2B = \sum_{m=1}^M ((B-B_m)g)^2$$

25

where

S^2B = estimated uncertainty variance of B ;

30 B = the regression coefficient at the cross validated rank A using all the N objects;

B_m = the regression coefficient at the rank A using all objects except the object(s) left out in cross validation segment m ; and

g = scaling coefficient (here: $g=1$).

35 In our approach, Jackknife has been implemented together with cross-validation. For each variable the difference between the B -coefficients B_i in a

- 50 -

cross-validated sub-model and B_{tot} for the total model is first calculated. The sum of the squares of the differences is then calculated in all sub-models to obtain an expression of the variance of the B_i estimate for a variable. The significance of the estimate of B_i is calculated using the t-test. Thus, the resulting regression coefficients can be presented with uncertainty limits that correspond to 2 Standard Deviations, and from that significant variables are detected.

No further details as to the implementation or use of this step are provided here since this has been implemented in commercially available software, The Unscrambler, CAMO ASA, Norway. Also, details on variable selection using Jackknife can be found in Westad & Martens (2000, J. Near Inf. Spectr., 8, p117-124).

The following approach can be used to select informative probes from a gene expression data set:

- a) keep out one unique sample (including its repetitions if present in the data set) per cross validation segment;
- b) build a calibration model (cross validated segment) on the remaining samples using PLSR-DA;
- c) select the significant genes for the model in step b) using the Jackknife criterion;
- d) repeat the above 3 steps until all the unique samples in the data set are kept out once (as described in step a). For example, if 75 unique samples are present in the data set, 75 different calibration models are built resulting in a collection of 75 different sets of significant probes;
- e) select the most significant variables using the frequency of occurrence criterion in the generated sets of significant probes in step d). For example, a set of probes appearing in all sets (100%) are more informative than probes appearing in only 50% of the

- 51 -

generated sets in step d).

Once the informative probes for a disease have been selected, a final model is made and validated. The two most commonly used ways of validating the model are cross-validation (CV) and test set validation. In cross-validation, the data is divided into k subsets. The model is then trained k times, each time leaving out one of the subsets from training, but using only the omitted subset to compute error criterion, RMSEP (Root Mean Square Error of Prediction). If k equals the sample size, this is called "leave-one-out" cross-validation. The idea of leaving one or a few samples out per validation segment is valid only in cases where the covariance between the various experiments is zero. Thus, one sample at-a-time approach can not be justified in situations containing replicates since keeping only one of the replicates out will introduce a systematic bias in our analysis. The correct approach in this case will be to leave out all replicates of the same samples at a time since that would satisfy assumptions of zero covariance between the CV-segments.

The second approach for model validation is to use a separate test-set for validating the calibration model. This requires running a separate set of experiments to be used as a test set. This is the preferred approach given that real test data are available.

The final model is then used to identify a disease, condition or stage thereof in test samples. For this purpose, expression data of selected informative genes is generated from test samples and then the final model is used to determine whether a sample belongs to a diseased or non-diseased class or has a condition or stage thereof.

Thus viewed from a yet further aspect the present invention provides a method of identifying probes useful for diagnosing or identifying or monitoring a disease or

- 52 -

condition or stage thereof in an organism, comprising the steps of:

- 5 a) immobilizing a set of oligonucleotide probes, preferably as described hereinbefore, on a solid support;
- b) isolating mRNA from a sample of a normal organism (normal sample), which may optionally be reverse transcribed to cDNA;
- 10 c) isolating mRNA from a sample from an organism, corresponding to the sample and organism of step (b), which is known to have said disease or condition or a stage thereof (diseased sample), which may optionally be reverse transcribed to cDNA;
- 15 d) hybridizing the mRNA or cDNA of steps (b) and (c) to said set of immobilized oligonucleotide probes of step (a); and
- e) assessing the amount of mRNA or cDNA hybridizing to each of said oligonucleotide probes to determine the level of gene expression of genes to which said oligonucleotide probes bind in said normal and diseased samples to generate a gene expression data set for each sample;
- 20 f) normalizing and standardizing said data set of step (e);
- g) constructing a calibration model for classification, preferably using the statistical techniques Partial Least Squares Discriminant Analysis (PLS-DA) and Linear Discriminant Analysis (LDA);
- 30 h) performing JackKnife analysis and identifying those oligonucleotide probes which are required for classification of said disease and normal samples into their respective groups.
- 35

Preferably a model for classification purposes is

- 53 -

generated by using the data relating to the probes identified according to the above described method. Preferably the sample is as described previously. Preferably the oligonucleotides which are immobilized in
5 step (a) are randomly selected as described below or are the probes as described hereinbefore. Such oligonucleotides may be of considerable length, e.g. if using cDNA (which is encompassed within the scope of the term "oligonucleotide"). The identification of such
10 cDNA molecules as useful probes allows the development of shorter oligonucleotides which reflect the specificity of the cDNA molecules but are easier to manufacture and manipulate.

The above described model may then be used to
15 generate and analyse data of test samples and thus may be used for the diagnostic methods of the invention. In such methods the data generated from the test sample provides the gene expression data set and this is normalized and standardized as described above. This is
20 then fitted to the calibration model described above to provide classification.

The method described herein can also be used to simultaneously select informative probes for several related and unrelated diseases or conditions. Depending
25 upon which diseases or conditions have been included in the calibration or training set, informative probes can be selected for the said diseases or conditions. The informative probes selected for one disease or condition may or may not be similar to the informative probes
30 selected for another disease or condition of interest. It is the pattern with which the selected genes are expressed in relation to each other during a disease, condition, or stage thereof, that determines whether or not they are informative for the disease, condition or
35 stage thereof.

In other words, informative genes are selected based on how their expression correlates with the

- 54 -

expression of other selected informative genes under the influence of responses generated by the disease, condition or stage thereof under investigation. In examples 1 and 2 provided hereinafter, 139 informative probes were selected for breast cancer diagnosis and 182 probes were selected for Alzheimer's disease diagnosis by training the gene expression data set of genes representing 1435 or 758 randomly picked cDNA clones for breast cancer/non breast cancer samples, or Alzheimer/non-Alzheimer samples, respectively. Among the probes selected for breast cancer and Alzheimer, about 10 probes were informative both for breast cancer and Alzheimer disease diagnosis.

For the purpose of isolating informative probes or identifying several related and unrelated diseases, conditions and stages thereof simultaneously, the gene expression data set must contain the information on how genes are expressed when the subject has a particular disease, condition or stage thereof under investigation. The data set is generated from a set of healthy or diseased samples, where a particular sample may contain the information of only one disease, condition or stages thereof or may also contain information about multiple diseases, conditions or stages thereof. For example, if the isolation of informative probes for Alzheimer disease, breast cancer and diabetes is sought, whole blood samples can be obtained from an Alzheimer patient who has breast cancer and diabetes. Hence, the method also teaches an efficient experimental design to reduce the number of samples required for isolating informative probes by selecting samples representing more than one disease, condition or stage thereof.

As mentioned previously, in view of the high information content of most transcripts, the identification and selection of informative probes for use in diagnosing, monitoring or identifying a particular disease, condition or stage thereof may be

- 55 -

dramatically simplified. Thus the pool of genes from which a selection may be made to identify informative probes may be radically reduced.

5 Unlike, in prior art technologies where informative probes are selected from a population of thousands of genes that are being expressed in a cell, like in microarray, in the method described herein, the informative probes are selected from a limited number of randomly obtained genes. For example, from a population
10 of 1435 cDNA clones, randomly picked from a human whole blood cDNA library, we were able to select 139 informative probes for breast cancer diagnosis (see Example 1 and Table 2).

Thus in a preferred aspect of the above mentioned
15 method of identifying probes useful for diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism, said set of oligonucleotides which are immobilized in step (a) are randomly selected from a larger set of oligonucleotides, e.g. from a cDNA library or other oligonucleotide pool,
20 which may be, but is preferably not selected from the set provided herein. Preferably said larger set comprises oligonucleotides which correspond to moderately or highly expressed genes. Thus preferably
25 in methods of the invention, the set of oligonucleotides according to the invention are replaced with a set of oligonucleotides which are randomly selected, e.g. from commercially available oligonucleotide or cDNA libraries.

30 As referred to herein "random" refers to selection which is not biased based on the extent of information carried by the transcripts in relation to the disease, condition or organism under study, ie. without bias towards their likely utility as informative probes.
35 Whilst a random selection may be made from a pool of transcripts (or related products) which have been biased, e.g. to highly or moderately expressed

- 56 -

transcripts, preferably random selection is made from a pool of transcripts not biased or selected by a sequence-based criterion. The larger set may therefore contain oligonucleotides corresponding to highly and moderately expressed genes, or alternatively, may be enriched for those corresponding to the highly and moderately expressed genes.

Random selection from highly and moderately expressed genes can be achieved in a wide variety of ways. A strategy used in this work, but not limiting in itself involves randomly picking a significant number of cDNA clones from a cDNA library constructed from a biological specimen under investigation. Since, in a cDNA library, the cDNA clones corresponding to transcripts present in high or moderate amount are more frequently present than transcripts corresponding to cDNA present in low amount, the former will tend to be picked up more frequently than the latter. A pool of cDNA enriched for those corresponding to highly and moderately expressed genes can be isolated by this approach.

To identify genes that are expressed in high or moderate amount among the isolated population for use in methods of the invention, the information about the relative level of their transcripts in samples of interest can be generated using several prior art techniques. Both non-sequence based methods, such as differential display or RNA fingerprinting, and sequence-based methods such as microarrays or macroarrays can be used for the purpose. Alternatively, specific primer sequences for highly and moderately expressed genes can be designed and methods such as quantitative RT-PCR can be used to determine the levels of highly and moderately expressed genes. Hence, a skilled practitioner may use a variety of techniques which are known in the art for determining the relative level of mRNA in a biological sample.

- 57 -

Especially preferably the sample for the isolation of mRNA in the above described method is as described previously and is preferably not from the site of disease and the cells in said sample are not disease cells and have not contacted disease cells.

The following examples are given by way of illustration only in which the Figures referred to are as follows:

Figure 1 shows the effect of Direct Standardization (DS) on the Alzheimer data measured in two different series of experiments in which AD denotes Alzheimer's samples and A,B are non-Alzheimer's samples. The samples in both series have been labelled systematically as (xx_7/xx_8), whereas the corrected samples from series 8 (in b,c,d) have been labelled as (xx_c), thus, for example, AD2-7 denotes Alzheimer disease sample number 2 in experiment series 7. The circled spots represent the samples chosen as the transfer samples. The connecting lines in figures b,c,d show the proximity of the replicated samples after applying DS. The dashed lines in figures a,c,d represent the decision boundary separating the classes. These lines have not been drawn on the basis of any statistical criteria, but serve the purpose of visually separating the classes. All the four figures show scores plot (PC1-PC2) from PCA analysis based on (a) non-standardized data, (b) scores plot after direct standardization using 3 transfer samples, (c) scores plot after direct standardization using 4 transfer sample, (d) scores plot after direct standardization using 8 transfer samples;

Figure 2 shows the projection of normal (including benign) and breast cancer samples onto a classification model generated by PLSR-DA using the data of 44 informative genes, in which PC is the principal components and N and C are normal and breast cancer samples, respectively;

Figure 3 shows the projection of individuals with

- 58 -

and without Alzheimer's disease onto a classification model generated by PLSR-DA using 182 informative genes;

Figures 4, 6 and 8 show projection plots as Figure 2 in which the classification model is generated using 719, 111 and 345 cDNAs, respectively, wherein PC is the principal components, N denotes normal and B denotes breast cancer samples;

Figures 5, 7 and 9 show prediction plots based on 3 principal components using the data of 719, 111 and 345 cDNAs, respectively;

Figure 10 shows a projection plot as Figure 3 in which the classification model is generated using 520 cDNAs; and

Figure 11 is the prediction plot corresponding to Figure 10.

Example 1: Diagnosis of Breast Cancer

Methods

Whole blood was obtained from the arms of breast cancer patients and patients with benign tumours (Ullevål and Haukland hospitals in Norway). All of the patients with breast cancer had a malignant tumour of the breast (disease samples). Healthy blood was collected from the above two hospitals, or collected at a Health station at Ås, Norway or at DiaGenic AS, Norway, from the arms of female donors with no reported signs of breast cancer. The blood from healthy individuals or with benign tumours comprise the normal samples. The blood was either collected in tubes containing EDTA and stored immediately at -80°C or was collected in PAXgene tubes and stored for 12-24 hours at room temperature before finally storing them at -80°C before use. Further details of the breast cancer and benign tumour patients from which blood was taken is provided in Table 5.

- 59 -

mRNA was isolated from the blood of the 29 breast cancer patients and 46 normal donors and used to prepare labelled probes by reverse transcribing in the presence of $\alpha^{33}\text{P}$ -dATP. The first strand cDNA of the normal and diseased samples was bound, separately to 1435 cDNA clones immobilized on a solid support (nylon membrane). These cDNA clones were randomly picked, without any prior knowledge of their gene sequences, from a cDNA library constructed using whole blood of 550 healthy individuals (Clontech, Palo Alto, USA). These methods were conducted as follows.

For amplification of inserts, bacterial clones were grown in microtiter plates containing 150 μl LB with 50 $\mu\text{g/ml}$ carbenicillin, and incubated overnight with agitation at 37°C. To lyse the cells, 5 μl of each culture were diluted with 50 μl H₂O and incubated for 12 min. at 95°C. Of this mixture, 2 μl were subjected to a PCR reaction using 20 pmoles of M13 forward and reverse primer in presence of 1.5 mM MgCl₂. PCR reactions were performed with the following cycling protocol: 4 min. at 95°C, followed by 25 cycles of 1 min. at 94°C, 1 min. at 60°C and 3 min. at 72°C either in a RoboCycler® Temperature Cycler (Stratagene, La Jolla, USA) or DNA Engine Dyad Peltier Thermal Cycler (MJ Research Inc., Waltham, USA). The amplified products were denatured by incubating with NaOH (0.2 M, final concentration) for 30 min. and spotted onto Hybond-N+ membranes (Amersham Pharmacia Biotech, Little Chalfont, UK), using MicroGrid II workstation according to the manufacturer's instructions (BioRobotics Ltd, Cambridge England). The immobilized cDNAs were fixed using a UV cross-linker (Hoefer Scientific Instruments, San Francisco, USA).

In addition to the 1435 cDNAs, the printed arrays also contained controls for assessing background level, consistency and sensitivity of the assay. These were

- 60 -

spotted at multiple positions and included controls such as PCR mix (without any insert); positive and negative controls of SpotReport™ 10 array validation system (Stratagene, La Jolla, USA) and cDNAs corresponding to constitutively expressed genes such as b-actin, g-actin, GAPDH, HOD and cyclophilin. Also, oligonucleotides corresponding to SIX1, b-tubulin, TRP-2, MDM2, Myosin Light C, CD44, Maspin, Laminin, and SRP 19 were included to detect disseminated cancer cells.

10

The total RNA from blood collected in EDTA tubes was purified using Trizol LS Reagent protocol (Invitrogen/Life Technologies). From blood contained in PAXgene tubes, the total RNA was purified according to the supplier's instructions (PreAnalytiX, Hombrechtikon, Switzerland). Contaminating DNA was removed from the isolated RNA by DNAase I treatment using DNA-free kit (Ambion, Inc. Austin, USA). RNA quality was determined visually by inspecting the integrity of 28S and 18S ribosomal bands following agarose gel electrophoresis. The concentration and purity of extracted RNA was determined by measuring the absorbance at 260 nm and 280 nm. mRNA was isolated from the total RNA using Dynabeads as per the supplier's instructions (Dynal AS, Oslo, Norway).

25

Labelling and hybridization experiments were performed in batches. The number of samples assayed in each batch varied from six to nine. In the case of samples that were assayed more than once (replicates), aliquots derived from the same mRNA pool were used for probe synthesis. For probe synthesis, aliquots of mRNA corresponding to 4-5 µg of total RNA were mixed together with oligodT_{25NV} (0.5 µg/ml) and mRNA spikes of SpotReport™ 10 array validation system (10 pg; Spike 2, 1 pg), heated to 70°C to remove secondary structures, and then chilled on ice. Probes were prepared in 35µl

35

- 61 -

reaction mixes by reverse transcription in the presence of 50 μ Ci [α^{33} P] dATP, 3.5 μ M dATP, 0.6 mM each of dCTP, dTTP, dGTP, 200 units of SuperScript reverse transcriptase (Invitrogen, LifeTechnologies) and 0.1 M DTT, labelling for 1.5 hr at 42°C. Following synthesis, the enzyme was deactivated for 10 min. at 70°C and mRNA removed by incubating the reaction mix for 20 min. at 37°C in 4 units of Ribo H (Promega, Madison USA). Unincorporated nucleotides were removed using ProbeQuant G 50 Columns (Amersham Biosciences, Piscataway, USA).

Prior to hybridization, the membranes were equilibrated in 4 x SSC for 2 hr at room temperature and prehybridized overnight at 65°C in 10 ml prehybridisation solution (4 x SSC, 0.1 M NaH₂PO₄, 1 mM EDTA, 8% dextran sulphate, 10 x denhardt's solution, 1% SDS). Freshly prepared probes were added to 5 ml of the same prehybridisation solution, and hybridization continued overnight at 65°C. The membranes were washed at 65°C at increasing stringency (2 x 30 min. each in 2 x SSC, 0.1% SDS; 1 x SSC, 0.1% SDS; 0.1 x SSC, 0.1% SDS) to remove unspecific signals.

The amount of labelled first strand cDNA binding to each spot was assessed and quantified using a PhosphorImager to generate a gene expression data set. The data was generated using Phoretix software version 3 (Non Linear Dynamics, England). Background subtraction was performed on the generated data by subtracting the median of the line of pixels around each spot outline from the total intensity obtained from the respective spots.

The background-subtracted data was then normalized and transformed by selecting out 50 lowest and 50 maximum signals from each membrane. This step was to exclude genes that were expressed with a high degree of

- 62 -

variance. Since the genes varied from membrane to membrane, the expression data from 497 genes were removed from the data set. The values for the remaining 938 genes were then normalised by using different approaches such as external controls, dividing each spot by the median intensity of the observed signal in the respective membrane, range normalizing the data from each membrane, and then log transforming the data obtained.

10

The processed data obtained above was then used to isolate the informative probes by:

15

a) keeping one unique sample (including all repetitions of the selected sample) out per cross validation segment;

b) building a calibration model (cross validated) on the remaining samples using PLSR-DA;

c) selecting the set of significant genes for the model in step b using the Jackknife criterion;

20

d) repeating steps a), b) and c) until all the unique samples were kept out once (hence, in all 75 different calibration models were built (after repeating step b) 75 times), resulting in 75 different sets of significant probes (after repeating step c) 75 times));

25

e) selecting significant variables using the frequency of occurrence criterion amongst the 75 different sets of significant probes.

30

The selected informative probes based on occurrence criterion were used to construct a classification model. The result of the classification model based on probes appearing in at least 90% of the generated sets after the step of isolating informative probes as described above is shown in Figure 2 in which it is seen that the expression pattern of these genes was able to classify most women with breast cancer and women with no breast cancer into distinct groups. In this figure PC1 and PC2

35

- 63 -

indicate the two principal components statistically derived from the data which best define the systemic variability present in the data. This allows each sample, and the data from each of the informative probes to which the sample's labelled first strand cDNA was bound, to be represented on the classification model as a single point which is a projection of the sample onto the principal components - the score plot.

10 The ability of the generated model, based on isolated informative probes, to predict future samples was determined by the double cross-validation approach. The performance of the diagnostic test for breast cancer based on the occurrence criterion is presented in Table 15 6.

Correct prediction of most breast cancer cells was achieved. These included all three samples obtained from women with ductal carcinoma in situ (DCIS), 11/15 samples obtained from women with stage I breast cancer, 20 all five samples obtained from women with stage II breast cancer, and one of two samples obtained from women with stage III breast cancer. Interestingly, two correctly predicted stage I samples were obtained from 25 women having a tumour size of <5 mm in diameter.

The model also correctly predicted the class of most non-cancer samples (41/46), including those that were obtained from women with non-cancerous breast 30 abnormalities.

Confirmation that the gene transcripts are not from cells which are disseminated disease cells has been confirmed by several lines of evidences. Firstly, the 35 informative genes were expressed constitutively at high or moderate levels in blood cells of women irrespective of whether they had cancer or not. Secondly, in the

- 64 -

assay described in this Example, in order to identify transcripts, at least 720 disseminated cells in blood samples would be required. Since, the average number of disseminated cells present in blood during different stages of breast cancer is much lower (organ confined breast cancer, 0.8 cells per ml; invasive breast cancer spread to lymph nodes only, 2.4 cells per ml; and metastatic breast cancer, 6 cells per ml; SD>100%) (29), we believe that the signals being detected originated from peripheral blood cells and could not have originated from disseminated cells. Thirdly, we were not able to detect any signal from the eight cancer markers known to have elevated expression in malignant cancer cells, including cancer cells that are disseminated in the blood.

Example 2: Diagnosis of Alzheimer's disease

Similar experiments were conducted with samples from Alzheimer's patients. In this method 7 patients diagnosed with Alzheimer's Disease at the Memory Clinic at Ullevål University Hospital were used in the trial. The patients were confirmed as having Alzheimer's disease based on the following criteria:

- * A standardized interview with a care-giver using IQCODE, an ADL scale and a scale measuring behaviour of the patient (Green scale).
- * Neuropsychological evaluation using MMSE, Clock drawing test, Trailmaking test A and B (TMT A and B), Kendrick object learning test (visual memory test), part of the Wechsler battery and Benton test.
- * A psychiatric evaluation using scales for detection of depression, MADRS for interviewing the patient and Cornell scale for interviewing the care-giver.
- * A physical examination.
- * Laboratory tests of blood samples to rule out other

- 65 -

diseases.

- * CT scan of the brain.
- * SPECT of the brain.

5 The mean age of the patients was 72.3 with an age range of 69-76. The mean MMSE score was 22.0 (the maximum score attainable being 30).

10 Six age-matched individuals without diagnosed Alzheimer's disease were used as a control. All had been tested with MMSE and had a minimum score of 28 (mean: 28.4). The mean age of the normal control group was 73.0 and the age range 66-81. A sample from a 16-year old individual, with a consequent minimal chance of
15 having Alzheimer's disease, was also included as an additional control.

Using the methods described above (except that hybridization to 758 rather than 1435 cDNA clones was
20 performed), informative probes were selected based on occurrence criterion and used to construct a classification model. The results of the classification model based on probes appearing at least once in the generated sets after the method to isolate informative
25 probes as described above is shown in Figure 3 in which it will be seen that the expression pattern of these genes was able to classify individuals with or without Alzheimer's disease into distinct groups. In this Figure PC1 and PC2 indicate the 2 principal components
30 statistically derived from the data which define the systematic variability present in the data. This allows each sample, and the data from each of the informative probes to which the samples' cDNA was bound, to be represented on the classification model as a single
35 point which is a projection of the sample onto the principal components - the score plot.

- 66 -

The ability of the generated model, based on isolated informative probes, to predict future samples was determined by the double cross-validation. The performance of the diagnostic test for Alzheimer's disease is presented in Table 7.

- 67 -

Appendix A

Partial Least Squares regression (PLSR)

5 Let a multivariate regression model be defined as:

$$Y = XB + F$$

where

10 X a $N \times P$ matrix with N predictor variables (genes);
 Y ($N \times J$) being the J predicted variables. In our case Y
 represents a matrix containing dummy variables;
 B is a matrix of regression coefficients; and
 F is a $N \times J$ matrix of residuals.

15

The structure of the PLSR model can be written as:

$$X = TP^T + E_A, \text{ and}$$

$$Y = TQ^T + F_A, \text{ where}$$

20

where

T ($N \times A$) is a matrix of score vectors which are linear
 combinations of the x-variables;

25 P ($P \times A$) is a matrix with the x-loading vectors p_a as
 columns;

Q ($J \times A$) is a matrix with the y-loading vectors q_a as
 columns;

E_a ($N \times P$) is the matrix for X after A factors; and

F_a ($N \times J$) is the matrix for Y after A factors.

30

The criterion in PLSR is to maximize the explained
 covariance of $[X, Y]$. This is achieved by the loading
 weights vector w_{a+1} , which is the first eigenvector of
 $E_a^T F_a F_a^T E_a$ (E_a and F_a are the deflated X and Y after a
 35 factors or PLS components).

The regression coefficients are given by:

- 68 -

$$B = W(P^T W)^{-1} Q^T$$

A PLSR model with full rank, i.e. maximum number of components, is equivalent to the MLR solutions. Further
5 details on PLSR can be found in Martens & Naes, 1989, Multivariate Calibration, John Wiley & Sons, Inc., USA and Kowalski & Seasholtz, 1991, supra.

- 69 -

Example 3: Validation of Example 1, diagnosis of breast cancer

The results in Example 1 were validated by using the
5 informative probes identified in Example 1 on new breast
cancer and control samples.

Methods

The methods, essentially as described in Example 1, were
10 used. Blood was taken from patients as described in
Table 8. However, blood was collected in PAXgene tubes
and the first strand labelled cDNAs were hybridized to
719 cDNAs spotted on nylon membranes along with other
controls as described in Example 1. After background
15 subtraction using control spots, the data of each
membrane was normalized using the inter quantile range.
The data was analysed as described in Example 1 and the
model validated by cross validation.

20 The 719 cDNAs which were spotted are a subset of the
cDNAs spotted in Example 1 and include 111 cDNAs
described in Table 2 and which were found to be
informative in Example 1.

Results

25 The results are shown in Figures 4 to 9. Figures 4, 6
and 8 are projection plots similar to Figure 2 and show
the projection of normal and breast cancer patients'
samples onto a classification model generated using all
30 719 cDNA. Figure 6 is similar but uses a classification
model generated with the 111 probes common to Example 1.
Figure 8 uses the 345 sequences of the 719 for which
sequence information is provided herein. In each case
classification of normal and breast cancer groups was
35 possible. Figures 5, 7 and 9 show prediction plots
which reflect the ability of the generated models to
correctly diagnose breast cancer. In the 3 prediction

- 70 -

plots shown, the disease samples appear on the x axis at +1 and the non-disease samples appear at -1. The y axis represents the predicted class membership. During prediction, if the prediction is correct, disease samples should fall above zero and non-disease samples should fall below zero. In each case almost all samples are correctly predicted.

10 Example 4: Validation of Example 2, diagnosis of
Alzheimers

The results in Example 2 were validated by using the informative probes identified in Example 2 on new Alzheimer's patient samples.

15 Methods

The methods, essentially as described in Example 2, were used. Twelve female patients diagnosed with Alzheimer's disease at the Memory Clinic at Ullevål University Hospital who were confirmed as having Alzheimer's disease based on the criteria of Example 2 were used in the trial. The mean age of the patients was 72.3 with an age range of 66-83. The mean MMSE score was 22.0 (the maximum score attainable being 30).

25 Sixteen age-matched female individuals without diagnosed Alzheimer's disease were used as the normal control group. All had been tested with MMSE and had a minimum score of 29. The mean age of the normal control group was 74.0 and the age range 66-86.

35 After transfer of the blood to PAXgene tubes, total mRNA was isolated from the blood of the Alzheimer's disease and from the control group donors according to the manufacturers's instructions (PreAnalytiX, Hombrechtikon, Switzerland). The isolated mRNA was labelled during reverse transcription in the presence of

- 71 -

$\alpha^{33}\text{P}$ -dATP, yielding a labelled first strand cDNA.
Hybridization was performed as described previously onto
730 cDNA clones picked from a cDNA library from whole
blood of 550 healthy individuals without knowledge of
5 the gene sequence of the random cDNA clones.

Results

The results are shown in Figures 10 and 11. Figure 10
is a projection plot generated using 520 probes which
10 have been sequenced. Figure 11 is a prediction plot and
shows correct prediction of almost all samples.

72

Table 1a

List of probes informative for disease diagnosis

	Clone ID	Sequence ID	No. of nucleotides	SEQ ID NO: in sequence listing
1	I-01	-	-	-
2	I-02	-	-	-
3	I-13	-	-	-
4	I-21	-	-	-
5	I-24	308	373	11
6	I-28	310	564	13
7	I-30	1180	622	398
8	I-34	313	554	15
9	I-37	-	-	-
10	I-42	-	-	-
11	I-52	-	-	-
12	I-54	1181	155	399
13	I-58	326	554	24
14	I-71	-	-	-
15	I-72	-	-	-
16	I-86	-	-	-
17	I-95	-	-	-
18	II-03	361	622	34
19	II-05	363	628	35
20	II-06	364	528	36
21	II-10	368	329	39
22	II-24	381	534	47
23	II-25	382	444	48
24	II-26	383	566	49
25	II-33	390	523	55
26	II-34	391	566	56
27	II-41	397	534	60
28	II-42	398	512	61
29	II-47	-	-	-

AMENDED SHEET

73

30	II-57	411	505	73
31	II-61	415	596	77
32	II-69	423	387	85
33	II-70	424	420	86
34	II-75	429	535	91
35	II-83	-	-	-
36	II-84	438	577	99
37	II-87	441	552	100
38	II-88	442	606	101
39	II-90	-	-	-
40	II-94	448	329	104
41	III-02	453	747	107
42	III-05	-	-	-
43	III-06	458	682	109
44	III-08	460	536	111
45	III-10	-	-	-
46	III-13	464	615	115
47	III-15	-	-	-
48	III-17	-	-	-
49	III-20	1183	479	401
50	III-23	473	694	119
51	III-26	476	476	122
52	III-35	485	551	130
53	III-39	487	224	131
54	III-40	488	349	132
55	III-43	490	382	500
56	III-44	491	382	134
57	III-53	500	390	142
58	III-56	503	109	144
59	III-57	504	374	145
60	III-60	-	-	-
61	III-60	-	-	-
62	III-61	507	521	148

AMENDED SHEET

74

63	III-63	509	575	150
64	III-68	-	-	-
65	III-74	518	502	155
66	III-80	523	585	158
67	III-82	-	-	-
68	III-85	526	516	161
69	III-89	530	660	165
70	III-92	-	-	-
71	III-96	-	-	-
72	IV-14	684	545	275
73	IV-15	1185	628	402
74	IV-23	-	-	-
76	IV-26	1186	494	403
75	IV-26	-	-	-
77	IV-29	-	-	-
78	IV-31	687	268	278
79	IV-32	688	569	279
80	IV-34	-	-	-
81	IV-35	-	-	-
82	IV-41	-	-	-
83	IV-45	-	-	-
84	IV-53	61	362	498
85	IV-62	-	-	-
86	IV-69	192	286	4
87	IV-80	701	579	291
88	IV-82	-	-	-
89	IV-93	-	-	-
90	IX-10	736	641	314
91	IX-12	-	-	-
92	IX-38	757	583	317
93	IX-39	758	424	318
94	IX-42	-	-	-
95	IX-48	764	626	319

AMENDED SHEET

75

96	IX-77	785	556	325
97	V-01	-	-	-
98	V-02	-	-	-
99	V-03	706	496	296
100	V-04	707	397	297
101	V-06	-	-	-
102	V-07	708	293	298
103	V-11	1188	599	404
104	V-12	711	498	301
105	V-15	-	-	-
106	V-17	-	-	-
107	V-21	-	-	-
108	V-25	-	-	-
109	V-32	-	-	-
110	V-35	-	-	-
111	V-39	-	-	-
112	V-42	-	-	-
113	V-43	-	-	-
114	V-47	-	-	-
115	V-49	-	-	-
116	V-52	-	-	-
117	V-54	-	-	-
118	V-55	77	412	499
119	V-58	-	-	-
120	V-59	-	-	-
121	V-65	-	-	-
122	V-68	-	-	-
123	V-71	-	-	-
124	V-75	-	-	-
125	V-79	-	-	-
126	V-80	726	260	311
127	V-90	-	-	-
128	V-91	-	-	-

AMENDED SHEET

129	V-92	-	-	-
130	V-94	-	-	-
131	VI-02	-	-	-
132	VI-04	865	122	339
133	VI-07	93	405	1
134	VI-09	-	-	-
135	VI-10	-	-	-
136	VI-12	869	667	341
137	VI-14	871	642	343
138	VI-17	-	-	-
139	VI-20	876	115	346
140	VI-21	-	-	-
141	VI-23	878	634	347
142	VI-34	-	-	-
143	VI-41	-	-	-
144	VI-42	-	-	-
145	VI-43	-	-	-
146	VI-44	-	-	-
147	VI-48	891	626	355
148	VI-49	-	-	-
149	VI-50	893	585	356
150	VI-52	-	-	-
151	VI-53	895	560	357
152	VI-55	897	509	359
153	VI-65	-	-	-
154	VI-70	108	550	2
155	VI-71	-	-	-
156	VI-72	-	-	-
157	VI-74	905	655	365
158	VI-76	907	582	367
159	VI-78	-	-	-
160	VI-79	-	-	-
161	VI-84	-	-	-

77

162	VI-87	911	595	370
163	VI-88	912	651	371
164	VI-90	-	-	-
165	VI-93	-	-	-
166	VI-95	915	230	374
167	VI-96	-	-	-
168	VII-02	-	-	-
169	VII-03	1196	412	411
170	VII-06	-	-	-
171	VII-10	-	-	-
172	VII-11	-	-	-
173	VII-15	1199	439	414
174	VII-19	562	580	171
175	VII-21	564	671	173
176	VII-25	-	-	-
177	VII-32	571	457	179
178	VII-36	575	209	182
179	VII-39	576	541	183
180	VII-42	579	502	186
181	VII-43	580	316	187
182	VII-46	583	631	190
183	VII-47	1200	526	415
184	VII-48	1201	613	416
185	VII-59	593	565	199
186	VII-60	-	-	-
187	VII-63	595	98	201
188	VII-66	598	362	204
189	VII-67	-	-	-
190	VII-72	600	595	206
191	VII-73	601	522	207
192	VII-75	-	-	-
193	VII-76	603	624	209
194	VII-77	1203	692	418

AMENDED SHEET

195	VII-80	605	338	210
196	VII-81	606	556	211
197	VII-83	-	-	-
198	VII-86	-	-	-
199	VII-88	-	-	-
200	VII-90	612	576	216
201	VII-91	613	341	217
202	VII-93	615	379	219
203	VIII-01	-	-	-
204	VIII-02	-	-	-
205	VIII-03	-	-	-
206	VIII-06	-	-	-
207	VIII-09	618	598	221
208	VIII-10	-	-	-
209	VIII-15	-	-	-
210	VIII-20	628	419	229
211	VIII-22	-	-	-
212	VIII-26	-	-	-
213	VIII-28	634	511	235
214	VIII-29	635	592	236
215	VIII-30	636	572	237
216	VIII-31	637	482	238
217	VIII-32	638	545	239
218	VIII-33	639	624	240
219	VIII-39	-	-	-
220	VIII-41	645	649	245
221	VIII-42	646	600	246
222	VIII-44	-	-	-
223	VIII-46	649	425	249
224	VIII-48	651	251	251
225	VIII-58	-	-	-
226	VIII-64	663	627	261
227	VIII-65	-	-	-

228	VIII-66	665	345	262
229	VIII-67	666	252	263
230	VIII-74	-	-	-
231	VIII-76	675	591	270
232	VIII-78	-	-	-
233	VIII-82	-	-	-
234	VIII-83	-	-	-
235	VIII-85	-	-	-
236	VIII-87	-	-	-
237	VIII-91	-	-	-
238	VIII-92	-	-	-
239	VIII-93	-	-	-
240	VIII-95	-	-	-
241	X-04	-	-	-
242	X-07	808	641	328
243	X-15	814	132	329
244	X-29	821	370	331
245	X-34	-	-	-
246	X-35	-	-	-
247	X-54	837	603	334
248	X-56	839	71	335
249	X-68	1207	642	421
250	X-72	849	622	336
251	X-94	860	501	337
252	XI-07	-	-	-
253	XI-13	1209	620	423
254	XI-50	-	-	-
255	XI-58	-	-	-
256	XI-81	1212	374	426
257	XII-07	1213	567	427
258	XII-17	-	-	-
259	XII-26	-	-	-
260	XII-27	-	-	-

261	XII-31	-	-	-
262	XII-32	-	-	-
263	XII-35	1214	620	428
264	XII-36	-	-	-
265	XII-52	-	-	-
266	XII-59	1216	484	430
267	XIII-19	1219	559	433
268	XIII-29	-	-	-
269	XIII-52	939	513	378
270	XIII-62	-	-	-
271	XIII-84	-	-	-
272	XIII-92	1221	741	435
273	XV-18	-	-	-
274	XV-22	1099	561	388
275	XV-24	-	-	-
276	XV-25	1224	485	436
277	XV-28	-	-	-
278	XV-34	-	-	-
279	XV-42	-	-	-
280	XV-68	-	-	-
281	XV-74	-	-	-
282	XV-93	-	-	-
283	XV-94	-	-	-
284	XV-96	-	-	-
285	XVI-36	1056	435	382
286	XVI-53	1230	741	439
287	XVI-59	-	-	-
288	XVI-66	1074	689	384
289	XVI-76	1083	198	386
290	XVI-77	1084	198	387
291	XVII-07	-	-	-
292	XVII-08	-	-	-
293	XVII-17	-	-	-

81

294	XVII-28	-	-	-
295	XVII-29	-	-	-
296	XVII-31	1139	503	392
297	XVII-36	-	-	-
298	XVII-39	-	-	-
299	XVII-40	1231	203	440
300	XVII-48	1148	587	393
301	XVII-55	-	-	-
302	XVII-58	-	-	-
303	XVII-67	-	-	-
304	XVII-72	-	-	-
305	XVII-76	1160	650	394
306	XVII-82	-	-	-
307	XVII-87	1165	502	395
308	XVII-95	1172	648	396

Table 1b
List of sequences of probes informative for disease diagnosis

Please see the note at the bottom

Clone ID	Sequence ID	SEQ ID No. in sequencing listing
I-09	298	Missing
I-10	299	6
I-13	1331	444
I-14	1178	397
I-15	300	7
I-16	301	Missing
I-17	302	8
I-19	304	9
I-20	305	Missing
I-22	306	10
I-23	307	Missing
I-24	308	11
I-25	309	12
I-28	310	13
I-30	1180	398
I-31	311	14
I-32	312	Missing
I-34	313	15
I-37	1440	482
I-38	314	16
I-39	315	17
I-40	316	18
I-42	1332	445
I-44	317	Missing
I-45	318	Missing
I-46	319	Missing
I-47	320	Missing
I-48	321	19

Clone ID	Sequence ID	
I-49	322	20
I-53	323	21
I-54	1181	399
I-56	324	22
I-57	325	23
I-58	326	24
I-60	327	25
I-64	328	26
I-67	330	27
I-69	331	28
I-71	332	Missing
I-72	333	Missing
I-73	334	Missing
I-77	335	29
I-79	336	Missing
I-80	337	30
I-81	338	31
I-82	339	32
I-86	1336	447
I-88	1182	400
I-95	1337	448
II-02	360	33
II-03	361	34
II-05	363	35
II-06	364	36
II-07	365	37
II-08	366	38
II-09	367	Missing

- 83 -

II-10	368	39
II-11	369	40
II-12	370	41
II-13	371	42
II-14	372	Missing
II-15	373	43
II-16	374	44
II-17	375	Missing
II-18	376	Missing
II-20	377	Missing
II-21	378	45
II-22	379	Missing
II-23	380	46
II-24	381	47
II-25	382	48
II-26	383	49
II-27	384	50
II-28	385	Missing
II-29	386	51
II-30	387	52
II-31	388	53
II-32	389	54
II-33	390	55
II-34	391	56
II-35	392	Missing
II-37	393	Missing
II-38	394	57
II-39	395	58
II-40	396	59
II-41	397	60
II-42	398	61
II-43	399	62

II-44	400	63
II-46	401	64
II-47	402	65
II-48	403	66
II-49	404	Missing
II-50	405	67
II-52	406	68
II-53	407	69
II-54	408	70
II-55	409	71
II-56	410	72
II-57	411	73
II-58	412	74
II-59	413	75
II-60	414	76
II-61	415	77
II-62	416	78
II-63	417	79
II-64	418	80
II-65	419	81
II-66	420	82
II-67	421	83
II-68	422	84
II-69	423	85
II-70	424	86
II-71	425	87
II-72	426	88
II-73	427	89
II-74	428	90
II-75	429	91
II-76	430	92
II-77	431	93

AMENDED SHEET

- 84 -

II-78	432	94
II-79	433	95
II-80	434	96
II-81	435	97
II-82	436	98
II- 83	437	Missing
II- 84	438	99
II- 85	439	Missing
II- 86	440	Missing
II- 87	441	100
II-88	442	101
II- 89	443	Missing
II- 90	444	Missing
II- 91	445	Missing
II- 92	446	102
II- 93	447	103
II- 94	448	104
II- 95	449	Missing
II- 96	450	105
III-01	452	106
III-02	453	107
III-03	454	108
III-04	455	Missing
III-05	457	Missing
III-06	458	109
III-07	459	110
III-08	460	111
III-09	461	112
III-11	462	113
III-12	463	114
III-13	464	115
III-14	465	Missing

III-15	466	Missing
III-16	467	Missing
III-17	468	Missing
III-18	469	116
III-19	470	Missing
III-20	1183	401
III-21	471	117
III-22	472	118
III-23	473	119
III-24	474	120
III-25	475	121
III-26	476	122
III-27	477	123
III-28	478	124
III-29	479	125
III-31	481	126
III-32	482	127
III-33	483	128
III-34	484	129
III-35	485	130
III-37	486	Missing
III-39	487	131
III-40	488	132
III-42	489	133
III-43	490	500
III-44	491	134
III-45	492	135
III-46	493	136
III-47	494	137
III-48	495	138
III-49	496	139
III-50	497	140

AMENDED SHEET

- 85 -

III-51	498	Missing
III-52	499	141
III- 53	500	142
III- 54	501	Missing
III- 55	502	143
III- 56	503	144
III- 57	504	145
III- 58	505	146
III- 59	506	147
III- 61	507	148
III- 62	508	149
III- 63	509	150
III- 64	510	151
III-65	511	Missing
III-66	512	152
III-67	513	153
III-69	514	Missing
III-70	515	154
III-71	516	Missing
III-73	517	Missing
III-74	518	155
III-75	519	156
III-77	520	Missing
III-78	521	157
III-79	522	Missing
III-80	523	158
III-81	524	159
III-82	1348	451
III-83	525	160
III-85	526	161
III-86	527	162
III-87	528	Missing

III-88	529	163 & 164
III-89	530	165
III- 91	531	Missing
III- 92	1351	452
III- 93	532	166
III- 94	533	167
III- 95	534	168
III- 96	535	Missing
IV-02	681	Missing
IV-04	682	273
IV-13	683	274
IV-14	684	275
IV-15	1185	402
IV-17	685	276
IV-23	1353	454
IV-26	1186	403
IV-28	686	277
IV-31	687	278
IV-32	688	279
IV-35	1355	455
IV-37	g6	497
IV-38	689	280
IV-40	690	281
IV-42	691	282
IV-43	1239	441
IV-44	692	283
IV-47	693	284
IV-53	61	498
IV-55	694	285
IV-56	695	Missing
IV-61	696	286
IV-64	697	287

AMENDED SHEET

- 86 -

IV-65	698	288
IV-69	192	4
IV-72	699	289
IV-73	700	290
IV-80	701	291
IV-82	196	Missing
IV-85	702	292
IV-93	703	293
IV-95	704	294
IV-96	705	295
IX-10	736	314
IX-12	738	Missing
IX-13	739	315
IX-24	747	316
IX-38	757	317
IX-39	758	318
IX-48	764	319
IX-50	766	320
IX-56	768	321
IX-62	773	322
IX-65	776	323
IX-72	782	324
IX-77	785	325
IX-91	796	326
IX-96	801	327
V-01	1361	458
V-03	706	296
V-04	707	297
V-07	708	298
V-08	709	299
V-09	710	300
V-11	1188	404

V1-16	873	344
V1-19	875	345
V-12	711	301
V-17	1364	459
V-18	712	Missing
V-20	713	302
V-24	714	303
V-25	1365	460
V-28	1189	405
V-35	1366	461
V-37	716	Missing
V-38	1190	406
V-39	1109	389
V-40	717	304
V-41	718	305
V-47	1368	463
V-48	719	306
V-49	1369	464
V-55	77	499
V-57	720	307
V-58	1370	465
V-61	721	308
V-64	722	309
V-65	723	Missing
V-68	1448	484
V-71	1495	496
V-74	724	310
V-75	1372	467
V-80	726	311
V-81	727	312
V-87	728	313
V-90	1374	468

AMENDED SHEET

- 87 -

VI-02	340	Missing
VI-03	341	Missing
VI-04	342	Missing
VI-06	343	Missing
VI-07	344	Missing
VI-08	345	Missing
VI-09	346	Missing
VI-11	347	Missing
VI-12	869	341
VI-13	870	342
VI-14	871	343
VI-16	873	344
VI-18	348	Missing
VI-19	349	Missing
VI-20	350	Missing
VI-21	351	Missing
VI-22	352	Missing
VI-23	878	347
VI-24	879	348
VI-25	353	Missing
VI-26	354	Missing
VI-27	355	Missing
VI-31	356	Missing
VI-32	885	351
VI-33	357	Missing
VI-35	358	Missing
VI-39	887	352
VI-43	1382	471
VI-44	1193	409
VI-45	889	353
VI-48	359	Missing
VI-49	892	501

VI-50	893	356
VI-53	895	357
VI-55	897	359
VI-58	899	361
VI-66	903	363
VI-67	904	364
VI-70	108	2
VI-71	1387	472
VI-74	905	365
VI-75	906	366
VI-76	907	367
VI-77	110	3
VI-79	1389	473
VI-80	908	368
VI-85	910	369
VI-87	911	370
VI-88	912	371
VI-90	1390	474
VI-93	1391	475
VI-95	915	374
VI-96	1392	476
VII-02	547	Missing
VII-03	548	Missing
VII-04	549	Missing
VII-05	550	Missing
VII-06	551	Missing
VII-07	552	Missing
VII-08	553	Missing
VII-09	554	Missing
VII-10	555	Missing
VII-11	556	Missing
VII-12	557	Missing

- 88 -

VII-14	558	Missing
VII-15	559	Missing
VII-17	560	169
VII-18	561	170
VII-19	562	171
VII-20	563	172
VII-21	564	173
VII-22	565	174
VII-23	566	175
VII-24	567	176
VII-25	1397	480
VII-26	250	5
VII-27	568	177
VII-28	569	Missing
VII-29	570	178
VII-32	571	179
VII-33	572	180
VII-34	573	Missing
VII-35	574	181
VII-36	575	182
VII-39	576	183
VII-40	577	184
VII-41	578	185
VII-42	579	186
VII-43	580	187
VII-44	581	188
VII-45	582	189
VII-46	583	190
VII-47	1200	415
VII-48	584	Missing
VII-49	585	191
VII-50	586	192

VII-52	587	193
VII-53	588	194
VII-54	589	195
VII-55	590	196
VII-57	591	197
VII-58	592	198
VII-59	593	199
VII-62	594	200
VII-63	595	201
VII-64	596	202
VII-65	597	203
VII-66	598	204
VII-67	1399	481
VII-71	599	205
VII-72	600	206
VII-73	601	207
VII-74	602	208
VII-76	603	209
VII-77	604	Missing
VII-80	605	210
VII-81	606	211
VII-82	607	212
VII-83	608	Missing
VII-84	609	213
VII-86	1453	487
VII-87	610	214
VII-89	611	215
VII-90	612	216
VII-91	613	217
VII-92	614	218
VII-93	615	219
VII-94	616	Missing

AMENDED SHEET

- 89 -

VII-96	617	220
VIII-09	618	221
VIII-10	619	222
VIII-11	620	Missing
VIII-12	621	223
VIII-13	622	224
VIII-15	623	Missing
VIII-16	624	225
VIII-17	625	226
VIII-18	626	227
VIII-19	627	228
VIII-20	628	229
VIII-21	629	230
VIII-22	1455	Missing
VIII-23	630	231
VIII-24	631	232
VIII-25	632	233
VIII-26	1456	489
VIII-27	633	234
VIII-28	634	235
VIII-29	635	236
VIII-30	636	237
VIII-31	637	238
VIII-32	638	239
VIII-33	639	240
VIII-34	640	Missing
VIII-36	641	241
VIII-37	642	242
VIII-38	643	243
VIII-40	644	244
VIII-41	645	245
VIII-42	646	246

VIII-43	647	247
VIII-45	648	248
VIII-46	649	249
VIII-47	650	250
VIII-48	651	251
VIII-50	652	252
VIII-51	653	253
VIII-53	654	254
VIII-54	655	255
VIII-55	656	256
VIII-56	657	257
VIII-57	658	258
VIII-58	659	Missing
VIII-59	660	259
VIII-60	661	260
VIII-61	662	Missing
VIII-64	663	261
VIII-65	664	Missing
VIII-66	665	262
VIII-67	666	263
VIII-68	667	Missing
VIII-69	668	Missing
VIII-70	669	264
VIII-71	670	265
VIII-72	671	266
VIII-73	672	267
VIII-74	673	268
VIII-75	674	269
VIII-76	675	270
VIII-77	676	271
VIII-78	677	Missing
VIII-79	678	Missing

AMENDED SHEET

- 90 -

VIII-80	679	272
X-07	808	328
X-15	814	329
X-20	817	330
X-29	821	331
X-34	825	332
X-46	833	333
X-54	837	334
X-56	839	335
X-68	1207	421
X-72	849	336
X-73	1208	422
X-94	860	337
XI-13	1209	423
XI-37	1460	490
XI-43	1210	424
XI-67	1211	425
XI-81	1212	426
XII-07	1213	427
XII-35	1214	428
XII-36	1215	429
XII-59	1216	430
XII-65	1028	381
XII-92	1217	431
XIII-03	917	375
XIII-04	1218	432
XIII-19	1219	433
XIII-24	926	376
XIII-51	938	377

XIII-52	939	378
XIII-67	947	379
XIII-69	949	380
XIII-88	1220	434
XIII-92	1221	435
XV-22	1099	388
XV-24	1101	Missing
XV-25	1224	436
XV-42	1108	Missing
XV-62	1226	437
XV-64	1118	390
XV-84	1125	391
XVI-19	1228	438
XVI-36	1056	382
XVI-53	1230	439
XVI-60	1071	383
XVI-66	1074	384
XVI-74	1081	385
XVI-76	1083	386
XVI-77	1084	387
XVII-31	1139	392
XVII-40	1231	440
XVII-48	1148	393
XVII-76	1160	394
XVII-87	1165	395
XVII-95	1172	396

AMENDED SHEET

Note

Sequences not available for sequence IDs in Table 1, and corresponding sequence IDs in Tables 2 and 4.

298, 301, 305, 307, 312, 317, 318, 319, 320, 332, 333, 334, 336, 340, 341, 342, 343, 344, 345, 346, 347,
348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 367, 372, 375, 376, 377, 379, 385, 392, 393,
404, 437, 439, 440, 443, 444, 445, 449, 455, 457, 465, 466, 467, 468, 470, 486, 498, 501, 511, 514, 516,
517, 520, 522, 528, 531, 535, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 573, 584,
604, 608, 616, 620, 623, 640, 659, 662, 664, 667, 668, 673, 677, 678, 679, 681, 695, 702, 712, 716, 825,
886, 894, 902, 909, 916, 1101, 1108, 1109, 1177, 1187, 1193, 1204, 1220, 1239, 1255, 1256, 1342, 1347,
1354, 1357, 1362, 1363, 1364, 1373, 1375, 1379, 1403, 1404, 1405, 1406, 1413

- 92 -

Table 2a

List of informative probes for diagnosis of breast cancer

Clone ID	Sequence ID	SEQ ID NO. in sequence listing
I-24	308	11
I-28	310	13
I-30	1180	398
I-52	-	-
I-54	1181	399
II-41	397	60
II-70	424	86
II-87	441	100
III-06	458	109
III-20	1183	401
III-40	488	132
III-57	504	145
III-60	-	-
III-61	507	148
III-89	530	165
IV-14	684	275
IV-15	1185	402
IV-26	1186	403
IV-32	688	279
IV-41	-	-
IV-53	61	498
IV-62	-	-
IV-69	192	4
IV-80	701	291
IV-82	196	missing
IX-10	736	314
IX-12	-	-
IX-38	757	317

Clone ID	Sequence ID	SEQ ID NO. in sequence listing
IX-48	764	319
IX-77	785	325
V-11	1188	404
V-32	-	-
V-39	-	-
V-55	77	499
V-80	726	311
V-94	-	-
VI-07	93	1
VI-34	-	-
VI-41	-	-
VI-48	891	355
VI-49	-	-
VI-52	-	-
VI-55	897	359
VI-65	-	-
VI-70	108	2
VI-72	-	-
VI-78	-	-
VI-84	-	-
VII-03	1196	411
VII-15	1199	414
VII-32	571	179
VII-39	576	183
VII-47	1200	415
VII-48	1201	416
VII-60	-	-
VII-73	601	207

AMENDED SHEET

- 93 -

VII-77	1203	418
VII-90	612	216
VIII-20	628	229
VIII-29	635	236
VIII-30	636	237
VIII-31	637	238
VIII-39	-	-
VIII-44	-	-
VIII-46	649	249
VIII-48	651	251
VIII-66	665	262
VIII-74	-	-
VIII-76	675	270
X-04	-	-
X-07	808	328
X-15	814	329
X-29	821	331
X-34	-	-
X-35	-	-
X-54	837	334
X-56	839	335
X-68	1207	421
X-72	849	336
X-94	860	337
XI-07	-	-
XI-13	1209	423
XI-50	-	-
XI-58	-	-
XI-81	1212	426
XII-07	1213	427
XII-17	-	-
XII-26	-	-

XII-27	-	-
XII-31	-	-
XII-32	-	-
XII-35	1214	428
XII-36	-	-
XII-52	-	-
XII-59	1216	430
XIII-19	1219	433
XIII-29	-	-
XIII-52	939	378
XIII-62	-	-
XIII-84	-	-
XIII-92	1221	435
XV-18	-	-
XV-22	1099	388
XV-24	-	-
XV-25	1224	436
XV-28	-	-
XV-34	-	-
XV-42	-	-
XV-68	-	-
XV-74	-	-
XV-93	-	-
XV-94	-	-
XV-96	-	-
XVI-36	1056	382
XVI-53	1230	439
XVI-59	-	-
XVI-66	1074	384
XVI-76	1083	386
XVI-77	1084	387
XVII-07	-	-

AMENDED SHEET

14-01-2005

- 94 -

XVII-08	-	-
XVII-17	-	-
XVII-28	-	-
XVII-29	-	-
XVII-31	1139	392
XVII-36	-	-
XVII-39	-	-
XVII-40	1231	440
XVII-48	1148	393
XVII-55	-	-
XVII-58	-	-
XVII-67	-	-
XVII-72	-	-
XVII-76	1160	394
XVII-82	-	-
XVII-87	1165	395
XVII-95	1172	396

- 95 -

Table 2b

List of sequences of probes informative for breast cancer

Please see the note at the bottom of Table 1. Some sequences are missing.

Clone ID	Sequence ID	SEQ ID NO. in sequence listing
I-13	1331	444
I-14	1178	397
I-24	308	11
I-25	309	12
I-28	310	13
I-30	1180	398
I-37	1440	482
I-42	1332	445
I-48	321	19
I-54	1181	399
I-60	327	25
I-72	1335	446
I-81	338	31
I-82	339	32
I-86	1336	447
I-88	1182	400
I-95	1337	448
II-02	360	33
II-03	361	34
II-06	364	36
II-07	365	37
II-10	368	39
II-21	378	45
II-23	380	46
II-24	381	47
II-25	382	48
II-27	384	50

Clone ID	Sequence ID	SEQ ID NO. in sequence listing
II-33	390	55
II-34	391	56
II-41	397	60
II-42	398	61
II-46	401	64
II-47	1338	449
II-48	403	66
II-52	406	68
II-57	411	73
II-58	412	74
II-59	413	75
II-60	414	76
II-61	415	77
II-62	416	78
II-64	418	80
II-67	421	83
II-69	423	85
II-70	424	86
II-74	428	90
II-80	434	96
II-82	436	98
II-84	438	99
II-87	441	100
II-88	442	101
II-96	450	105
III-01	452	106
III-02	453	107

III-06	458	109
III-08	460	111
III-12	463	114
III-13	464	115
III-17	1344	450
III-18	469	116
III-20	1183	401
III-21	471	117
III-23	473	119
III-24	474	120
III-25	475	121
III-26	476	122
III-27	477	123
III-28	478	124
III-29	479	125
III-32	482	127
III-33	483	128
III-35	485	130
III-39	487	131
III-40	488	132
III-42	489	133
III-45	492	135
III-46	493	136
III-47	494	137
III-48	495	138
III-56	503	144
III-57	504	145
III-58	505	146
III-59	506	147
III-61	507	148
III-62	508	149
III-63	509	150

III-64	510	151
III-66	512	152
III-67	513	153
III-70	515	154
III-74	518	155
III-75	519	156
III-78	521	157
III-80	523	158
III-81	524	159
III-82	1348	451
III-85	526	161
III-86	527	162
III-88	529	163+164
III-89	530	165
III-92	1351	452
III-93	532	166
III-95	534	168
III-96	1352	453
IV-04	682	273
IV-13	683	274
IV-14	684	275
IV-15	1185	402
IV-17	685	276
IV-23	1353	454
IV-26	1186	403
IV-31	687	278
IV-32	688	279
IV-35	1355	455
IV-37	G6	497
IV-38	689	280
IV-42	691	282
IV-43	1239	441

- 97 -

IV-47	693	284
IV-53	61	498
IV-61	696	286
IV-64	697	287
IV-69	192	4
IV-72	699	289
IV-80	701	291
IV-82	196	Missing
IV-85	702	292
IV-93	1360	457
IV-96	705	295
IX-10	736	314
IX-12	738	Missing
IX-13	739	315
IX-24	747	316
IX-38	757	317
IX-39	758	318
IX-48	764	319
IX-50	766	320
IX-56	768	321
IX-62	773	322
IX-65	776	323
IX-72	782	324
IX-77	785	325
IX-91	796	326
IX-96	801	327
V-01	1361	458
V-03	706	296
V-04	707	297
V-07	708	298
V-08	709	299
V-11	1188	404

V-12	711	301
V-17	1364	459
V-24	714	303
V-25	1365	460
V-28	1189	405
V-35	1366	461
V-38	1190	406
V-39	1109	389
V-41	718	305
V-47	1368	463
V-49	1369	464
V-55	77	499
V-57	720	307
V-58	1370	465
V-61	721	308
V-64	722	309
V-65	1371	466
V-68	1448	484
V-71	1495	496
V-74	724	310
V-75	1372	467
V-80	726	311
V-90	1374	468
VI-03	864	338
VI-04	865	339
VI-07	93	1
VI-08	867	340
VI-09	1378	469
VI-12	869	341
VI-13	870	342
VI-14	871	343
VI-16	873	344

- 98 -

VI-19	875	345
VI-20	876	346
VI-21	1380	470
VI-23	878	347
VI-24	879	348
VI-25	1192	408
VI-26	881	349
VI-32	885	351
VI-39	887	352
VI-43	1382	471
VI-44	1193	409
VI-45	889	353
VI-48	891	355
VI-49	892	501
VI-50	893	356
VI-53	895	357
VI-55	897	359
VI-58	899	361
VI-66	903	363
VI-67	904	364
VI-70	108	2
VI-71	1387	472
VI-74	905	365
VI-75	906	366
VI-76	907	367
VI-77	110	3
VI-79	1389	473
VI-80	908	368
VI-85	910	369
VI-87	911	370
VI-88	912	371
VI-90	1390	474

VI-93	1391	475
VI-95	915	374
VI-96	1392	476
VII-02	1195	410
VII-03	1196	411
VII-06	1394	477
VII-08	1197	412
VII-09	1198	413
VII-10	1395	478
VII-11	1396	479
VII-15	1199	414
VII-17	560	169
VII-19	562	171
VII-21	564	173
VII-22	565	174
VII-23	566	175
VII-24	567	176
VII-25	1397	480
VII-26	250	5
VII-27	568	177
VII-29	570	178
VII-32	571	179
VII-33	572	180
VII-36	575	182
VII-39	576	183
VII-41	578	185
VII-42	579	186
VII-43	580	187
VII-46	583	190
VII-47	1200	415
VII-48	1201	416
VII-49	585	191

- 99 -

VII-54	589	195
VII-57	591	197
VII-58	592	198
VII-59	593	199
VII-62	594	200
VII-63	1202	417
VII-64	596	202
VII-66	598	204
VII-67	1399	481
VII-72	600	206
VII-73	601	207
VII-77	1203	418
VII-80	605	210
VII-82	607	212
VII-86	1453	487
VII-87	610	214
VII-90	612	216
VII-91	613	217
VII-92	614	218
VII-93	615	219
VII-96	617	220
VIII-09	618	221
VIII-10	619	222
VIII-13	622	224
VIII-16	624	225
VIII-20	628	229
VIII-21	629	230
VIII-22	1455	Missing
VIII-23	630	231
VIII-24	631	232
VIII-25	632	233
VIII-26	1456	489

VIII-27	633	234
VIII-28	634	235
VIII-29	635	236
VIII-30	636	237
VIII-31	637	238
VIII-32	638	239
VIII-33	639	240
VIII-34	1204	419
VIII-38	643	243
VIII-40	644	244
VIII-41	645	245
VIII-46	649	249
VIII-48	651	251
VIII-55	656	256
VIII-57	658	258
VIII-59	660	259
VIII-60	661	260
VIII-61	1205	420
VIII-64	663	261
VIII-66	665	262
VIII-73	672	267
VIII-74	673	268
VIII-76	675	270
VIII-80	679	272
X-07	808	328
X-15	814	329
X-20	817	330
X-29	821	331
X-34	825	332
X-46	833	333
X-54	837	334
X-56	839	335

AMENDED SHEET

- 100 -

X-68	1207	421
X-72	849	336
X-73	1208	422
X-94	860	337
XI-13	1209	423
XI-37	1460	490
XI-43	1210	424
XI-67	1211	425
XI-81	1212	426
XII-07	1213	427
XII-35	1214	428
XII-36	1215	429
XII-59	1216	430
XII-65	1028	381
XII-92	1217	431
XIII-03	917	375
XIII-04	1218	432
XIII-19	1219	433
XIII-24	926	376
XIII-51	938	377
XIII-52	939	378
XIII-67	947	379
XIII-69	949	380
XIII-88	1220	434
XIII-92	1221	435
XV-22	1099	388
XV-24	1101	Missing
XV-25	1224	436
XV-42	1108	Missing
XV-62	1226	437
XV-64	1118	390
XV-84	1125	391

XVI-19	1228	438
XVI-36	1056	382
XVI-53	1230	439
XVI-60	1071	383
XVI-66	1074	384
XVI-74	1081	385
XVI-76	1083	386
XVI-77	1084	387
XVII-31	1139	392
XVII-40	1231	440
XVII-48	1148	393
XVII-76	1160	394
XVII-87	1165	395
XVII-95	1172	396

- 101 -

Table 3

List of informative probes (Clone ID) selected for breast cancer diagnosis based on their occurrence critrion during selection

Occurrence*	Clone ID
100%	XI-8, XVI-66, VIII-66, XVI-59, VII-03, XIII-19, XII-35, X-35, XI-50, XII-26, IV-53, XIII-29, XIII-62, I-30, III-06, XV-22, XV-94, VII-15, VII-39, IX-39, XVII-39, III-40, VII-32
90%	I-52, VI-65, VI-34, IV-62, XV-34, XVII-58, V-11, VI-78, XII-36, XIII-92, VIII-29, XVI-53, XVI-77, XI-13, XIII-84, IV-14, XII-31, V-80, VII-48, XVII-29, XVII-72
80%	III-60, VIII-74, IX-12, X-04, XIII-52, VIII-30, IX-38
70%	VI-49, X-29, VIII-48
60%	IV-82, IX-10, VI-52, X-68, VII-77
50%	IV-15
40%	XV-28, II-70, V-55
30%	XVII-17, XVII-67
20%	XI-58, XVI-36, VIII-39, VIII-44, III-61, IV-69, XV-68, X-72
10%	IX-42, IX-77, X-94, XV-96, XVII-55
5%	XII-59, XVI-76, I-54, XV-18, V-94, X-54, VI-07, VII-47, XVII-31, XVII-87, XVII-48
In at least one model	II-41, VI-41, III-57, III-89, VII-73, XV-25, IV-26, X-34, IV-41, VII-90, XV-42, XVII-82, XII-27, VIII-20, I-28, VII-60, VIII-76, III-20, VI-84, XI-07, XVII-28, XII-17, XVII-36, XII-52, XVII-76, VIII-46, VI-70, XV-74, XV-93, VIII-31, II-87, V-39, VI-55, X-07, X-15, XII-07, XVII-07, XVII-08, XVII-95, I-24, IV-32, V-32, VI-48, VI-72, IV-80, IX-48, X-56, XV-24, XII-32, XVII-40

*100% = Genes appearinig in all the 75 cross validated models;

90% = Additional genes appearing in at least 68 out of 75 cross validated models;

5% = Additional genes appearing in at least 4 out of 75 cross validated models and so on.

Table 4a

List of informative probes for diagnosis of Alzheimer disease

Clone ID	Sequence ID	
I-01	-	-
I-02	-	-
I-13	-	-
I-21	-	-
I-34	313	15
I-37	-	-
I-42	-	-
I-58	326	24
I-71	-	-
I-72	-	-
I-86	-	-
I-95	-	-
II-03	361	34
II-05	363	35
II-06	364	36
II-10	368	39
II-24	381	47
II-25	382	48
II-26	383	49
II-33	390	55
II-34	391	56
II-42	398	61
II-47	-	-
II-57	411	73
II-61	415	77
II-69	423	85
II-75	429	91
II-83	-	-

Clone ID	Sequence ID	
II-84	438	99
II-88	442	101
II-90	-	-
II-94	448	104
III-02	453	107
III-05	-	-
III-06	458	109
III-08	460	111
III-10	-	-
III-13	464	115
III-15	-	-
III-17	-	-
III-23	473	119
III-26	476	122
III-35	485	130
III-39	487	131
III-43	490	500
III-44	491	134
III-53	500	142
III-56	503	144
III-60	-	-
III-63	509	150
III-68	-	-
III-74	518	155
III-80	523	158
III-82	-	-
III-85	526	161
III-92	-	-

14-01-2005

- 103 -

III-96	-	-
IV-23	-	-
IV-26	-	-
IV-29	-	-
IV-31	687	278
IV-34	-	-
IV-35	-	-
IV-45	-	-
IV-80	701	291
IV-82	-	-
IV-93	-	-
V-01	-	-
V-02	-	-
V-03	706	296
V-04	707	297
V-06	-	-
V-07	708	298
V-12	711	301
V-15	-	-
V-17	-	-
V-21	-	-
V-25	-	-
V-35	-	-
V-42	-	-
V-43	-	-
V-47	-	-
V-49	-	-
V-52	-	-
V-54	-	-
V-58	-	-
V-59	-	-
V-65	-	-

V-68	-	-
V-71	-	-
V-75	-	-
V-79	-	-
V-80	726	311
V-90	-	-
V-91	-	-
V-92	-	-
VI-02	-	-
VI-04	865	339
VI-09	-	-
VI-10	-	-
VI-12	869	341
VI-14	871	343
VI-17	-	-
VI-20	876	346
VI-21	-	-
VI-23	878	347
VI-41	-	-
VI-42	-	-
VI-43	-	-
VI-44	-	-
VI-48	891	355
VI-49	-	-
VI-50	893	356
VI-53	895	357
VI-71	-	-
VI-74	905	365
VI-76	907	367
VI-78	-	-
VI-79	-	-
VI-87	911	370

AMENDED SHEET

- 104 -

VI-88	912	371
VI-90	-	-
VI-93	-	-
VI-95	915	374
VI-96	-	-
VII-02	-	-
VII-03	-	-
VII-06	-	-
VII-10	-	-
VII-11	-	-
VII-19	562	171
VII-21	564	173
VII-25	-	-
VII-36	575	182
VII-42	579	186
VII-43	580	187
VII-46	583	190
VII-59	593	199
VII-63	595	201
VII-66	598	204
VII-67	-	-
VII-72	600	206
VII-73	601	207
VII-75	-	-
VI-02	-	-
VI-04	866	MISSING
VI-09	-	-
VI-10	-	-
VI-12	873	344
VI-14	875	345
V-17	-	-

VII-91	613	217
VII-93	615	219
VIII-01	-	-
VIII-02	-	-
VIII-03	-	-
VIII-06	-	-
VIII-09	618	221
VIII-10	-	-
VIII-15	-	-
VIII-22	-	-
VIII-26	-	-
VIII-28	634	235
VIII-30	636	237
VIII-32	638	239
VIII-33	639	240
VIII-41	645	245
VIII-42	646	246
VIII-48	651	251
VIII-58	-	-
VIII-64	663	261
VIII-65	-	-
VIII-67	666	263
VIII-78	-	-
VIII-82	-	-
VIII-83	-	-
VIII-85	-	-
VIII-87	-	-
VIII-91	-	-
VIII-92	-	-
VIII-93	-	-
VIII-95	-	-

- 105 -

Table 4b

List of sequences of probes informative for Alzheimer disease

Please see the note to Table 1

Clone ID	Sequence ID	
I-09	298	Missing
I-10	299	6
I-15	300	7
I-16	301	Missing
I-17	302	8
I-19	304	9
I-20	305	Missing
I-22	306	10
I-23	307	Missing
I-24	308	11
I-25	309	12
I-28	310	13
I-31	311	14
I-32	312	Missing
I-34	313	15
I-38	314	16
I-39	315	17
I-40	316	18
I-44	317	Missing
I-45	318	Missing
I-46	319	Missing
I-47	320	Missing
I-48	321	19
I-49	322	20
I-53	323	21
I-56	324	22
I-57	325	23
I-58	326	24

Clone ID	Sequence ID	
I-60	327	25
I-64	328	26
I-67	330	27
I-69	331	28
I-71	332	Missing
I-72	333	Missing
I-73	334	Missing
I-77	335	29
I-79	336	Missing
I-80	337	30
I-81	338	31
I-82	339	32
VI-02	340	Missing
VI-03	341	Missing
VI-04	342	Missing
VI-06	343	Missing
VI-07	344	Missing
VI-08	345	Missing
VI-09	346	Missing
VI-11	347	Missing
VI-18	348	Missing
VI-19	349	Missing
VI-20	350	Missing
VI-21	351	Missing
VI-22	352	Missing
VI-25	353	Missing
VI-26	354	Missing
VI-27	355	Missing

- 106 -

VI-31	356	Missing
VI-33	357	Missing
VI-35	358	Missing
VI-48	359	Missing
II-02	360	33
II-03	361	34
II-05	363	35
II-06	364	36
II-07	365	37
II-08	366	38
II-09	367	Missing
II-10	368	39
II-11	369	40
II-12	370	41
II-13	371	42
II-14	372	Missing
II-15	373	43
II-16	374	44
II-17	375	Missing
II-18	376	Missing
II-20	377	Missing
II-21	378	45
II-22	379	Missing
II-23	380	46
II-24	381	47
II-25	382	48
II-26	383	49
II-27	384	50
II-28	385	Missing
II-29	386	51
II-30	387	52
II-31	388	53

II-32	389	54
II-33	390	55
II-34	391	56
II-35	392	Missing
II-37	393	Missing
II-38	394	57
II-39	395	58
II-40	396	59
II-41	397	60
II-42	398	61
II-43	399	62
II-44	400	63
II-46	401	64
II-47	402	65
II-48	403	66
II-49	404	Missing
II-50	405	67
II-52	406	68
II-53	407	69
II-54	408	70
II-55	409	71
II-56	410	72
II-57	411	73
II-58	412	74
II-59	413	75
II-60	414	76
II-61	415	77
II-62	416	78
II-63	417	79
II-64	418	80
II-65	419	81
II-66	420	82

- 107 -

II-67	421	83
II-68	422	84
II-69	423	85
II-70	424	86
II-71	425	87
II-72	426	88
II-73	427	89
II-74	428	90
II-75	429	91
II-76	430	92
II-77	431	93
II-78	432	94
II-79	433	95
II-80	434	96
II-81	435	97
II-82	436	98
II-83	437	Missing
II-84	438	99
II-85	439	Missing
II-86	440	Missing
II-87	441	100
II-88	442	101
II-89	443	Missing
II-90	444	Missing
II-91	445	Missing
II-92	446	102
II-93	447	103
II-94	448	104
II-95	449	Missing
II-96	450	105
III-01	452	106
III-02	453	107

III-03	454	108
III-04	455	Missing
III-05	457	Missing
III-06	458	109
III-07	459	110
III-08	460	111
III-09	461	112
III-11	462	113
III-12	463	114
III-13	464	115
III-14	465	Missing
III-15	466	Missing
III-16	467	Missing
III-17	468	Missing
III-18	469	116
III-19	470	Missing
III-21	471	117
III-22	472	118
III-23	473	119
III-24	474	120
III-25	475	121
III-26	476	122
III-27	477	123
III-28	478	124
III-29	479	125
III-31	481	126
III-32	482	127
III-33	483	128
III-34	484	129
III-35	485	130
III-37	486	Missing
III-39	487	131

AMENDED SHEET

- 108 -

III-40	488	132
III-42	489	133
III-43	490	500
III-44	491	134
III-45	492	135
III-46	493	136
III-47	494	137
III-48	495	138
III-49	496	139
III-50	497	140
III-51	498	Missing
III-52	499	141
III-53	500	142
III-54	501	Missing
III-55	502	143
III-56	503	144
III-57	504	145
III-58	505	146
III-59	506	147
III-61	507	148
III-62	508	149
III-63	509	150
III-64	510	151
III-65	511	Missing
III-66	512	152
III-67	513	153
III-69	514	Missing
III-70	515	154
III-71	516	Missing
III-73	517	Missing
III-74	518	155
III-75	519	156

III-77	520	Missing
III-78	521	157
III-79	522	Missing
III-80	523	158
III-81	524	159
III-83	525	160
III-85	526	161
III-86	527	162
III-87	528	Missing
III-88	529	163/164
III-89	530	165
III-91	531	Missing
III-93	532	166
III-94	533	167
III-95	534	168
III-96	535	Missing
VII-02	547	Missing
VII-03	548	Missing
VII-04	549	Missing
VII-05	550	Missing
VII-06	551	Missing
VII-07	552	Missing
VII-08	553	Missing
VII-09	554	Missing
VII-10	555	Missing
VII-11	556	Missing
VII-12	557	Missing
VII-14	558	Missing
VII-15	559	Missing
VII-17	560	169
VII-18	561	170
VII-19	562	171

- 109 -

VII-20	563	172
VII-21	564	173
VII-22	565	174
VII-23	566	175
VII-24	567	176
VII-27	568	177
VII-28	569	Missing
VII-29	570	178
VII-32	571	179
VII-33	572	180
VII-34	573	Missing
VII-35	574	181
VII-36	575	182
VII-39	576	183
VII-40	577	184
VII-41	578	185
VII-42	579	186
VII-43	580	187
VII-44	581	188
VII-45	582	189
VII-46	583	190
VII-48	584	Missing
VII-49	585	191
VII-50	586	192
VII-52	587	193
VII-53	588	194
VII-54	589	195
VII-55	590	196
VII-57	591	197
VII-58	592	198
VII-59	593	199
VII-62	594	200

VII-63	595	201
VII-64	596	202
VII-65	597	203
VII-66	598	204
VII-71	599	205
VII-72	600	206
VII-73	601	207
VII-74	602	208
VII-76	603	209
VII-77	604	Missing
VII-80	605	210
VII-81	606	211
VII-82	607	212
VII-83	608	Missing
VII-84	609	213
VII-87	610	214
VII-89	611	215
VII-90	612	216
VII-91	613	217
VII-92	614	218
VII-93	615	219
VII-94	616	Missing
VII-96	617	220
VIII-09	618	221
VIII-10	619	222
VIII-11	620	Missing
VIII-12	621	223
VIII-13	622	224
VIII-15	623	Missing
VIII-16	624	225
VIII-17	625	226
VIII-18	626	227

- 110 -

VIII-19	627	228
VIII-20	628	229
VIII-21	629	230
VIII-23	630	231
VIII-24	631	232
VIII-25	632	233
VIII-28	634	235
VIII-29	635	236
VIII-30	636	237
VIII-31	637	238
VIII-32	638	239
VIII-33	639	240
VIII-34	640	Missing
VIII-36	641	241
VIII-37	642	242
VIII-38	643	243
VIII-40	644	244
VIII-41	645	245
VIII-42	646	246
VIII-43	647	247
VIII-45	648	248
VIII-46	649	249
VIII-47	650	250
VIII-48	651	251
VIII-50	652	252
VIII-51	653	253
VIII-53	654	254
VIII-54	655	255
VIII-55	656	256
VIII-56	657	257
VIII-57	658	258
VIII-58	659	Missing

VIII-59	660	259
VIII-60	661	260
VIII-61	662	Missing
VIII-64	663	261
VIII-65	664	Missing
VIII-66	665	262
VIII-67	666	263
VIII-68	667	Missing
VIII-69	668	Missing
VIII-70	669	264
VIII-71	670	265
VIII-72	671	266
VIII-73	672	267
VIII-74	673	268
VIII-75	674	269
VIII-76	675	270
VIII-77	676	271
VIII-78	677	Missing
VIII-79	678	Missing
VIII-80	679	272
IV-02	681	Missing
IV-04	682	273
IV-13	683	274
IV-14	684	275
IV-17	685	276
IV-28	686	277
IV-31	687	278
IV-32	688	279
IV-38	689	280
IV-40	690	281
IV-42	691	282
IV-44	692	283

AMENDED SHEET

- 111 -

IV-47	693	284
IV-55	694	285
IV-56	695	Missing
IV-61	696	286
IV-64	697	287
IV-65	698	288
IV-72	699	289
IV-73	700	290
IV-80	701	291
IV-85	702	292
IV-93	703	293
IV-95	704	294
IV-96	705	295
V-03	706	296
V-04	707	297
V-07	708	298
V-08	709	299
V-09	710	300
V-12	711	301
V-18	712	Missing
V-20	713	302
V-24	714	303
V-37	716	Missing
V-40	717	304
V-41	718	305
V-48	719	306
V-57	720	307
V-61	721	308
V-64	722	309
V-65	723	Missing
V-74	724	310
V-80	726	311

VI-81	727	312
VI-87	728	313
VI-13	870	342
VI-14	871	343
VI-16	873	344
VI-23	878	347
VI-24	879	348
VI-28	883	350
VI-32	885	351
VI-38	886	Missing
VI-39	887	352
VI-45	889	353
VI-46	890	354
VI-49	892	501
VI-50	893	356
VI-52	894	Missing
VI-53	895	357
VI-54	896	358
VI-55	897	359
VI-57	898	360
VI-58	899	361
VI-63	900	362
VI-65	902	Missing
VI-66	903	363
VI-67	904	364
VI-74	905	365
VI-75	906	366
VI-76	907	367
VI-80	908	368
VI-81	909	Missing
VI-85	910	369
VI-87	911	370

AMENDED SHEET

- 112 -

VI-88	912	371
VI-91	913	372
VI-94	914	373
VI-95	915	374
VI-96	916	Missing
I-13	1177	Missing
I-14	1178	397
I-30	1180	398
I-54	1181	399
I-88	1182	400
III-20	1183	401
IV-15	1185	402
IV-26	1186	403
IV-62	1187	Missing
V-11	1188	404
V-28	1189	405
V-38	1190	406
V-45	1191	407
V-44	1193	409
VII-47	1200	415
I-42	1332	445
I-52	1333	Missing
I-86	1336	447
I-95	1337	448
III-10	1342	Missing
III-60	1347	Missing
III-82	1348	451
III-92	1351	452
IV-23	1353	454
IV-34	1354	Missing
IV-35	1355	455
IV-41	1356	Missing

IV-45	1357	Missing
IV-82	1359	456
V-01	1361	458
V-02	1362	Missing
V-06	1363	Missing
V-17	1364	459
V-25	1365	460
V-35	1366	461
V-42	1367	462
V-47	1368	463
V-49	1369	464
V-58	1370	465
V-75	1372	467
V-79	1373	Missing
V-90	1374	468
V-91	1375	Missing
V-94	1376	Missing
VI-10	1379	Missing
VI-41	1381	Missing
VI-43	1382	471
VI-71	1387	472
VI-72	1388	Missing
VI-79	1389	473
VI-90	1390	474
VI-93	1391	475
VII-25	1397	480
VII-60	1398	Missing
VII-67	1399	481
VIII-22	1403	Missing
VIII-26	1404	Missing
VIII-39	1405	Missing
VIII-44	1405	Missing

- 113 -

I-37	1440	482
V-32	1445	Missing
V-52	1447	483
V-68	1448	484
V-92	1449	485
VI-42	1450	486
VI-78	1452	Missing
VII-86	1453	487
VII-88	1454	488
IV-29	1490	491
V-15	1491	492
V-39	1492	493
V-54	1493	494
V-59	1494	495
V-71	1495	496

Table 5Samples

Diagnosis	No. of women
Normal/Benign	42*
DCIS	3
Invasive cancer	26

*From one woman, whole blood was collected at weeks 1,2,3,4,5 following menstruation. Hence, the number of unique normal/benign samples tested in the experiment is 75.

Information about women with breast cancer

<i>Sample</i>	<i>AGE</i>	<i>Stage</i>	<i>Cancer type</i>	<i>Size hist. (mm)</i>	<i>Nodes</i>
1	51	II	IDC	20	1/7
2	84	II	IDC	22	2/2
3	50	I	DCIS+ 1 IDC	>50 DCIS; 5 x 14	0/7
4	47	I	IDC	15	0
5	69	III	ILC g.2 + tubular adenocarcinoma	50 + 3	1 av 12 + 1 av 7
6	50	II	IDC	24	0
7	65	I	IDC	15	0
8	63	II	IDC	23	0
9	55	I	IDC + DCIS	4	0 av 1
10	52	0	DCIS + small colloid carcinoma foci	50 + 3	0
11	60	II	IDC	24	0
12	54	I	IDC	11	0
13		0	DCIS	20	0
14	49	0	DCIS	9	0
15	48	I	IDC	4	0
16	56	I	IDC	4	0
17	68	I	IDC	14	0
18	68	I	IDC	7	0
19	63	I	IDC	10	0
20	45	I	IDC	19	1
21	57	III	IDC	60	8/20
22	55	II	IDC/DCIS	35 + 55	0
23	71	I	IDC/extensive DCIS	8	0
24	56	I	ODC	9	?

25	66	II	IDC	26	0
26	66	I	IDC	15	?
27	61	I	IDC	9	?
28	?	?	?	?	0
29	65	I	IDC	11	0

Other diseases/conditions present in the women tested

Other diseases/conditions present in the women tested

Disease/condition
Diabetes
Asthma
Ulcerous colitis
Hemochromatose
Crohn's disease
Fibromyalgia
Psoriasis
Atopic eczema
Rheumatism
Allergies

Prior history of cancer in the women tested

Cancer type	No. of women
Breast	3
Colon	2
Stomach	1
Skin	1

Table 6

Number of samples tested by double cross validation and success of the diagnostic test for breast cancer based on selected ionformative genes

Number of samples tested by double cross validation

Number of unique samples tested	75
Number of unique non cancer samples tested	46
Number of cancer samples tested	29

Success of the diagnostic test for breast cancer based on selected informative genes

Occurrence in percentage*	Number of informative probes	Specificity	Sensitivity	Accuracy	False Positive rate	False negative rate	Total error rate
100.00	23	84.78	75.86	81.33	15.22	24.14	18.67
90.00	44	91.30	79.31	86.67	8.70	20.69	13.33
80.00	51	86.96	79.31	84.00	13.04	20.69	16.00
70.00	54	89.13	75.86	84.00	10.87	24.14	16.00
60.00	58	89.13	75.86	84.00	10.87	24.14	16.00
50.00	59	89.13	75.86	84.00	10.87	24.14	16.00
40.00	63	89.13	75.86	84.00	10.87	24.14	16.00
30.00	66	86.96	75.86	82.67	13.04	24.14	17.33
20.00	74	89.13	75.86	84.00	10.87	24.14	16.00
10.00	79	89.13	75.86	84.00	10.87	24.14	16.00
5.00	90	86.96	79.31	84.00	13.04	20.69	16.00
1.33	139	84.78	72.41	80.00	15.22	27.59	20.00

*100% = Genes appearing in all the 75 cross validated models; 90% = Genes appearing in at least 68 out of 75 cross validated models; 5% = Genes appearing in at least 4 out of 75 cross validated models; and so on.

Table 7

Double cross-validation and details of the success of the diagnostic test for Alzheimer disease based on the expression 182 informative genes

Validation Result		Success of diagnostic test		
Total number of samples tested	14	Performance	Description	%
Number of Alzheimer's disease samples tested	7	Accuracy	Percentage of the total number of predictions that were correct	92.9
Number of Alzheimer's disease samples incorrectly predicted	1	Sensitivity	Percentage of positive cases that were correctly identified	85.7
Number of non-Alzheimer's disease samples tested	7	Specificity	Percentage of negatives cases that were correctly predicted	100
Number of non-Alzheimer's disease samples incorrectly predicted	0	False positive rate	Percentage of negatives cases that were incorrectly classified as positive	0.0
		False negative rate	Percentage of positive cases that were incorrectly classified as negative	14.3
		Total error rate	Percentage of the total cases incorrectly predicted	7.1

Table 8

Some relevant features of the blood donors. B, Female donors with breast cancer; N, Female donors with suspected mammogram but no breast cancer; IDC, invasive ductal carcinoma; DCIS, ductal carcinoma in situ; na, not available nd, not determined; ++, no degradation of mRNA and no ribosomal contamination in the sample, +, no degradation of mRNA but ribosomal contamination in the sample.

		AGE	Cancer type/ breast abnormality	Size Hist. (mm)	mRNA Quality
1	B1	na	IDC	5	++
2	B2	49	DCIS	8	nd
3	B3	54	IDC	18	++
4	B4	59	IDC	12	+
5	B5	61	DCIS + micro invasive cancer	15+1.5	++
6	B6	55	IDC	12+17	nd
7	B6		IDC	12+17	nd
8	N1	45	Fibroadenoma	-	nd
9	N2	52	na	-	+
10	N3	55	Cyst	-	++
11	N4	54	na	-	++
12	N5	51	Benign ductal epithelium	-	nd
13	N6	57	Benign	-	nd
14	N7	50	na	-	++
15	N8	52	na	-	+

Table 9

List of sequences of probes informative for both alzheimer and breast cancer disease

Clone ID	Sequence ID	SEQ ID No. in sequence listing
I-24	308	11
I-25	309	12
I-28	310	13
I-48	321	19
I-60	327	25
I-72	333	Missing
I-81	338	31
I-82	339	32
II-02	360	33
II-03	361	34
II-06	364	36
II-07	365	37
II-10	368	39
II-21	378	45
II-23	380	46
II-24	381	47
II-25	382	48
II-27	384	50
II-33	390	55
II-34	391	56
II-41	397	60
II-42	398	61
II-46	401	64
II-47	402	65
II-48	403	66
II-52	406	68
II-57	411	73
II-58	412	74
II-59	413	75
II-60	414	76
II-61	415	77
II-62	416	78
II-64	418	80
II-67	421	83

Clone ID	Sequence ID	SEQ ID No. in sequence listing
II-69	423	85
II-70	424	86
II-74	428	90
II-80	434	96
II-82	436	98
II-84	438	99
II-87	441	100
II-88	442	101
II-96	450	105
III-01	452	106
III-02	453	107
III-06	458	109
III-08	460	111
III-12	463	114
III-13	464	115
III-17	468	Missing
III-18	469	116
III-21	471	117
III-23	473	119
III-24	474	120
III-25	475	121
III-26	476	122
III-27	477	123
III-28	478	124
III-29	479	125
III-32	482	127
III-33	483	128
III-35	485	130
III-39	487	131
III-40	488	132
III-42	489	133
III-45	492	135
III-46	493	136
III-47	494	137

- 120 -

III-48	495	138
III-56	503	144
III-57	504	145
III-58	505	146
III-59	506	147
III-61	507	148
III-62	508	149
III-63	509	150
III-64	510	151
III-66	512	152
III-67	513	153
III-70	515	154
III-74	518	155
III-75	519	156
III-78	521	157
III-80	523	158
III-81	524	159
III-85	526	161
III-86	527	162
III-88	529	163/164
III-89	530	165
III-93	532	166
III-95	534	168
III-96	535	Missing
IV-04	682	273
IV-13	683	274
IV-14	684	275
IV-17	685	276
IV-31	687	278
IV-32	688	279
IV-38	689	280
IV-42	691	282
IV-47	693	284
IV-61	696	286
IV-64	697	287
IV-72	699	289
IV-80	701	291
IV-85	702	292
IV-93	703	293

IV-96	705	295
V-03	706	296
V-04	707	297
V-07	708	298
V-08	709	299
V-12	711	301
V-24	714	303
V-41	718	305
V-57	720	307
V-61	721	308
V-64	722	309
V-65	723	Missing
V-74	724	310
V-80	726	311
VI-03	341	Missing
VI-04	342	Missing
VI-07	344	Missing
VI-08	345	Missing
VI-09	346	Missing
VI-12	869	341
VI-14	871	343
VI-19	349	Missing
VI-20	350	Missing
VI-21	351	Missing
VI-23	878	347
VI-25	353	Missing
VI-26	354	Missing
VI-48	359	Missing
VI-50	893	356
VI-53	895	357
VI-74	905	365
VI-76	907	367
VI-87	911	370
VI-88	912	371
VI-95	915	374
VII-02	547	Missing
VII-03	548	Missing
VII-06	551	Missing
VII-08	553	Missing

VII-09	554	Missing
VII-10	555	Missing
VII-11	556	Missing
VII-15	559	Missing
VII-17	560	Missing
VII-19	562	171
VII-21	564	173
VII-22	565	174
VII-23	566	175
VII-24	567	176
VII-27	568	177
VII-29	570	178
VII-32	571	179
VII-33	572	180
VII-36	575	182
VII-39	576	183
VII-41	578	185
VII-42	579	186
VII-43	580	187
VII-46	583	190
VII-48	584	Missing
VII-49	585	191
VII-54	589	195
VII-57	591	197
VII-58	592	198
VII-59	593	199
VII-62	594	200
VII-63	595	201
VII-64	596	202
VII-66	598	204
VII-72	600	206
VII-73	601	207
VII-77	604	Missing
VII-80	605	210
VII-82	607	212
VII-87	610	214
VII-90	612	216
VII-91	613	217

VII-92	614	218
VII-93	615	219
VII-96	617	220
VIII-09	618	221
VIII-10	619	222
VIII-13	622	224
VIII-16	624	225
VIII-20	628	229
VIII-21	629	230
VIII-23	630	231
VIII-24	631	232
VIII-25	632	233
VIII-28	634	235
VIII-29	635	236
VIII-30	636	237
VIII-31	637	238
VIII-32	638	239
VIII-33	639	240
VIII-34	640	Missing
VIII-38	643	243
VIII-40	644	244
VIII-41	645	245
VIII-46	649	249
VIII-48	651	251
VIII-55	656	256
VIII-57	658	258
VIII-59	660	259
VIII-60	661	260
VIII-61	662	Missing
VIII-64	663	261
VIII-66	665	262
VIII-73	672	267
VIII-74	673	268
VIII-76	675	270
VIII-80	679	272

VII-02	547
VII-03	548
VII-06	551
VII-08	553
VII-09	554
VII-10	555
VII-11	556
VII-15	559
VII-17	560
VII-19	562
VII-21	564
VII-22	565
VII-23	566
VII-24	567
VII-27	568
VII-29	570
VII-32	571
VII-33	572
VII-36	575
VII-39	576
VII-41	578
VII-42	579
VII-43	580
VII-46	583
VII-48	584
VII-49	585
VII-54	589
VII-57	591
VII-58	592
VII-59	593
VII-62	594
VII-63	595
VII-64	596
VII-66	598
VII-72	600
VII-73	601
VII-77	604
VII-80	605
VII-82	607
VII-87	610
VII-90	612
VII-91	613
VII-92	614
VII-93	615
VII-96	617
VIII-09	618
VIII-10	619
VIII-13	622
VIII-16	624
VIII-20	628
VIII-21	629

VIII-23	630
VIII-24	631
VIII-25	632
VIII-28	634
VIII-29	635
VIII-30	636
VIII-31	637
VIII-32	638
VIII-33	639
VIII-34	640
VIII-38	643
VIII-40	644
VIII-41	645
VIII-46	649
VIII-48	651
VIII-55	656
VIII-57	658
VIII-59	660
VIII-60	661
VIII-61	662
VIII-64	663
VIII-66	665
VIII-73	672
VIII-74	673
VIII-76	675
VIII-80	679

- 124 -

Nucleotide sequences

Sequence ID - 93

nt: 405

5 GGATCCTGTGGCCACAGAGCTGCCCCAGCAGACGCTCCGCCCCACCCGGTGATGG
AGCCCCGGGGGACAATCGTGCCTGGGGAGGAGCAGGGTACAGCCCATTCCCCCAG
CCCTGGCTGACCTGGCCTAGCAGTTTGGCCCTGCTGGCCTTAGCAGGGAGACAGGG
GAGCAAAGAACGCCAAGCCGGAGGCCCGAGGCCAGCCGGCCTCTCGAGAGCCAGAG
CAGCAGTTGAATGTAATGCTGGGGACAGGCATGCTGCCGCCAGTAGGGCGGGGACC
CGGACAGCCAGGTGACTACCAGTCCTGGGGACACACTCACCATAAACACATCCCCA
10 GGCAGGACAGATCGGGGAAGGGGTGTGTACCAGGCTATGATTTCTCTTGCAATAAA
ATGTATTATTATT

Sequence ID - 108

nt: 550

15 GGCTTTGACAGAGTGCAAGACGATGACTTGCAAAATGTCGCATCTGGAACGCAACA
TAGANACCATCATCAACACCTTCCACCAATACTCTGTGAAGCTGGGGCACCAGAC
ACCCTGAACCAGGGGGAATTCAAAGAGCTGGTGCGAAAAGATCTGCAAAATTTTCT
CAAGAAGGAGAATAAGAATGAAAAGGTCATAGAACACATCATGGAGGACCTGGACA
CAAATGCAGACAAGCAGCTGAGCTTCGAGGAGTTCATCATGCTGATGGCGAGGCTA
ACCTGGGCCTCCACGAGAAGATGCACGAGGGTGACGAGGGCCCTGGCCACCACCA
20 TAAGCCAGGCCTCGGGGAGGGCACCCCCCTAAGACCACAGTGGCCAAGATCACAGTG
GCCACGGCCACGGCCACAGTCATGGTGGCCACGGCCACAGCCACTAATCAGGAGGC
CAGGCCACCCTGCCTNTACCCAACCAGGGCCCCGGGGCCTGTTATGTCAAACCTGTC
TTGGCTGTGGGGCTAGGGGCTGGGGCCAAATAAAGTCTCTTTCTCC

25 Sequence ID 110

ACGAAGACAGACATCTGTGGAATGATTCACATCCTCTCAAGTTAGGAGGATGGAGG
CCTGCTTCATTAAGAAGCTGGGGGTAGGGTGGGGGTGGGGGAGAACACTTAACAACA
TGGGGACCAGTCAGGGGAATCCCCTTATTTCTGTTTTGCATATGAGGAACCCTAGA
GCAGCCAGGTGAGGCTCTCTAGTTTAATAAAAATCATGGAAAGACTCTTAATGCAG
30 ACTCTTCTTAAGTGTTAATAGGGATTTTTTCAGCTTATTTTGGTTGCAGTTTCCAA
TTTTTAAAAATGTTGAGGTAATCTTTCCACCTTCCCAAACCTAATTCTTGTAGAT
GCATTAGTGTTGAACCAATGCTTTCTCATGTCTCAATTCTTTGTATATGCATTCTT
TTCAGATGTATTAAACAAACAAAACCCTTC

35 Sequence ID - 192

nt: 286

CCGGTAATAGAATAGAAAAGGGAGAGTGTCTTCATGCAATGTGGCATCCTGGATTG
GGTCTCGNNACAAAAACAGGACATTAGTGGGAAAATTGGAAATCTGAAAAAAGTCT

- 125 -

GAATTTTAGTTAATATAACCAATTCAGTCTCTTGGTTTTTGACAGATGTACCATGGT
GATGTAAGATGTTGACCTTGGGGTAGGCTGGGTGAAGGGTATACAGGAACTCTTTG
TACTATCTCTGCAACTTCTCTGTAAATCTAGTATCATTCCAAAATAAAAGTTTATT
TAATTT

5

Sequence ID 250

GTGGAAGTGACATCGTCTTTAAACCCTGCGTGGCAATCCCTGACGCACCGCCGTGA
TGCCCAGGGAAGACAGGGCGACCTGGAAGTCCAACACTTCTTAAGATCATCCAA
CTATTGGATGATTATCCGAAATGTTTCATTGTGGGAGCAGACAATGTGGGCTCCAA
10 GCAGATGCAGCAGATCCGCATGTCCCTTCGCGGGAAGGCTGTGGTGCTGATGGGCA
AGAACACCATGATGCGCAAGGCCATCCGAGGGCACCTGGAAAACAACCCAGCTCTG
GAGAACTGCTGCCTCATATCCGGGGGAATGTGGGCTTTGTGTTTACCAAGGAGGA
CCTCACTGAGATCAGGGACATGTTGCTGGCCAATAAGGTGCCAGCTGCTGCCCGTG
CTGGTGCCATTGCCCCATGTGAAGTCACTGTGCCAGCCCAGAACACTGGTCTCGGG
15 CCCGAGAAGACCTCCTTTTTTCCAGGCTTTAGGTATCACCCTAAAATCTCCAGGGG
CACCATTGAAATCCTGAGTGATGTGCACTGATCAAGACTGG

Sequence ID 299

CAGCGCAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAAG
20 TTTTTTTCTCTTTGAAAGATAGAGATTGNTACAACACTACTTAAAAAATATAGTCAAT
AGGTTACTAAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAAG
ATTTTAAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAA
AGGTTTCTAAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAATG
TATTTAAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAGA
25 AGGGCAATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTT
AAAAGTTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAGAGATTAAAC
CGAAGGTGATTAAAAGACCTTGAAATCCATGACGCANGGAGAATTGCGCATTTAAA
GCCTAGTTACGCATTTACTAAACGCAGACGAAAATGGGAAGATTAATTGGGAGTGG
TAGGATGAAACAATTTTGGAGAAGATAGAAG

30

Sequence ID 300

CTCAAAGGAGAAAAAAAACCTTGTAaaaaaaAGCAAAAATGACAACAGAAAAACAAT
CTTATTCCGAGCATTCAGTAACTTTTTTGTGTATGTACTTAGCTGTACTATAAGT
AGTTGGTTTGTATGAGATGGTTAAAAAGGCCAAAGATAAAAGGTTTCTTTTTTTTTT
35 CCTTTTTTGTCTATGAAGTTGCTGTTTATTTTTTTTTTGGCCTGTTTGATGTATGTGT
GAAACAATGTTGTCCAACAATAAACAGGAATTTTATTTTGCTGAGTTGTTCTAAAA
AAAAAAAAAAAAAAAAAAAA

- 126 -

Sequence ID 302

AGTAGAGACGGGGTTTCACTGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGT
GATCCGGCCACCTCGGCCTCCCGAAAGTGCTGGGATTACAGGCGTGAGCCACGGCG
CCCAGCCCCAGCCTGTCACTTAAACTGATAAACGACAGATTAACAGTAGAAAAATT
5 TTATTTTGCATACATAATGAGGCTTCACAAAAGAGAAGTGAAAACCCAAGTAGGAG
TTTAGGGCTGGGGGCTTATATACCATTTAACAAGGGGTGATAAATTGTAAGAGAAT
AG

Sequence ID 304

10 TCCTTGGTTTTCGATTTGTGGCAACAATCCAGTCTTTTTGTTTTTTTCAGGGATACC
ATATGTAACAGGTGCCATTGTTACTGTAACTTTTTCACACATGCCTTCAGTTTGATG
TCAAAGTCATCATTTAGTGTAACAGCAAGTTATCTGTTAGGCTGCACATCATGAA
CTTTACTTTTAGAAAGTCTTATCTTTTATGCCACAGAAATAGCATTGTGGCTATTAG
TCATGGATGGCAAAGAAATTAATTTTGAGTTGTTTGGATAAAAATGTTTCAGTTGA
15 CTGTAGTGTGTATTGAGAGACACTGCCAGTAAACAAACTCTCTTGGTAGGTGGAAA
TCCCCTAGAAGTTACAGAAAATTGGGAGGAGGTGAACTTAATTAAATAACTTGAAT
TGTTTAGACATATTCAGAGCTTCTTATGACCTTGAAGAAATCACCCAACTTCAAAA
GACCTCGGTTTCTTCATTTGTAAAATTAGGGAGTTTGACTAGATGTGTAAATCTAG
TTGTTAGTTAACTTCTAAGATGTAAAAACCCTCTTGTTTAACAAAAACCTACAAGA
20 TCAAGTTGCTTATCTGAAATCTTTATGAATCAACACTAGTCACTAAGTCTAGCTCG
ACC

Sequence ID 306

CTTTTCCTCCCGCTGTCCCCCACGGAGGGGACTGCTCTCCCCCGCTGCATCCTTTC
25 TGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAACCTC
GACATTCGGGAAGTGTTTTTGAGAAGTCTCGGTCCGTAAGGGAAGTCTTCCAAGTC
CGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTTCGGCCACCCCCCAGATGAGG
CAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTTCTCTCAGCTT
CCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGTCCATGCTGCCAT
30 CAAGGCATTTATTGCAGTGTACTATTTGCTTCCAAAGGATCAGGCCCTGAGAACAA
TGACCTTATTTCTTACAACAGTGTCTGGGTGCGTGCCAGCAGATGCCTCAGATAC
CAAGAGATAACAAAGCTGCAGCTCTTTTGATGCTGACCAAGAATGTGGATTTTGTG
AAGGATGCNCATGAANAAATGGACNAGCTGTG

35 Sequence ID - 308

nt: 373

AAGTGGGTCTTGCCATCCCTGAACTGNAATCATCCCTAACATATTCATACCTGTTT
TCATTTTAAAAGTTGGGTGAGTTTTTTTTATTAGTACATGTATTTCTATCCTACTGA

- 127 -

TTTATTTGCTATATCATCTAATTTAGTTTGAATATTCCATAATTTACTTAATTAGT
CCTGTATGGAGACCTAGCTCTTCTCAGTGTCTACTATTATAAACAATGCTACAGTG
AATATTGGTGNATAAATCCATACNCACCACGTACATATCTTAAGTTCTGGAAGAGA
TATTGCTAAACCAGAAGATAACCTGCATTTAAAATTTGACTGCTAGGGNCAGGGNC
5 ACATTTAATTAAATTAGAACAANGAATGCATAATGNC

Sequence ID 309

CCGGAATCGCGGCCGCGTCGACGAAAATATGTGCCCTGGCCAACTCCACAGGACTA
GTTCTAGGCAATCTGAAGGAAACCAGAAAATGTGAATTTCTCTTCCCTCAAAAAGC
10 TATACTGAAGTAGTATTTAATATTCAAGTACTTGTAATTTGCAGAACAGTACTTT
TTAATTTGACCCATGAATTCTATTTAAATTTGTCACTTAATATTTAGCCAAGAAGC
AAACCATCTAAAAAGATTTCTGGTTTATTTCTCCAACCTCCTAATAAATAGGGTCAC
ATATTTTTTAACTTTTTTTCTAATTTGAAAAGTAATACAGGCATATGGTATTTTTAAA
AATGAAACAACACAAAGGGATATGTTTTGAAAAGTGGTCTTGCCATCCCTGAACTG
15 TAATCATCCCTAACATATTCATACCTGTTTTTCATTTTAAAAGTTGGGTCAGTTTTT
TTATTAGTACATGTATTTCTATCCTACTGATTTATTTGCTATATCATCTAATTTAG
TTTGAATATTCCATAATTTACTTAATTAGTCCTGTATGGAGACCTAGCTCTTCTCA
GTGTCTACTATTATAAACAATGCTACAGTGAATATTGGTGNATAAATCCTACACAC
CACGTAACATATCTTAAGTTCCTGGAAGAGATATTGCTAAACCAGAAGATAACCTG
20 CATTTAAAATTTGACTGCTAGGGTCAGGGTCACATTTAAATTAAATTAGAACAAGG
AATGCATAATGTCTTCGATAGCAATCTATTCAAGGTGCACCGTGGTCACAAAGGAA
AGCAAAACTGTC

Sequence ID - 310 nt:564

CCTGGNCAGAGGCCTCTATCCTGTANTGATAATTGCCATCAAAATTGTCAAAAANG
ATTTAATTTCTATGGGNAATAGTCCTTTTTCTTAGCTTCTGCCNNTCACTTGCTTAT
TTTTTGTGTGGGAATGGGGTTGGATAAACCAATGAACTTTATTATAAACAATCCC
ACCTATATCTANCAAATTTATATTTTCGGTGAAATACAGATATTTGCCTTTCTGGA
GTANTATAGAAGCTGTCAATATGTATCTACTGTACAGTACTAAATAGTATTCATTT
30 ATGAAATGAGTAGTGTTTGGGTGGCTGGGGTTAAGGAAAAATGAGACTTGGAATTG
TAGCTTTTATCCAAGTTTTGAGTATAAATAGGGTTTTGTTTTGTTTTTTTTTAACCT
AAAACTGAAATGCCATATAGAAAAACAGCATTTGTTTTTACAGTTTGTAAGTA
ACTTTTTTAAAGATTTTATCAAAAAGAATTTTGTCTATNGTGAGTAAAAGAAGTTCT
AATAATGGCCTAATCACTGCATTTTTTAAAAAACAAAGTTCAACACAAATGACATTT
35 GTTT

Sequence ID 311

- 128 -

CCTCTCCTCCATCTAAAGGCAACATTCCTTACCCATTAGTCTCAGAAATTGTCTTA
AGCAACAGCCCCAAATGCTGGCTGCCCCCGGCCAAGCATTTGGGGCCGCCATCCTGC
CTGGCACTGGCTGATGGGCACCTCTGTTGGTTCCATCAGCCAGAGCTCTGCCAAAG
GCCCCGCAGTCCCTCTCCCAGGAGGACCCTAGAGGCAATTAAATGATGTCCTGTTC
5 CATTGG

Sequence ID - 313

nt: 554

CCCGGAATCGCGGCCCGCGTCGACAACAAACCTGCATGTTCTGCACATGTATCCAG
GAACTTAAAAAAAAAAAAAGATAGTTTGTGTGTCTTAATTGAATAATAGTAGATTT
10 ATAGATTAAAGATCTATGGGTTTTTAATATGGATTANAAATCTGTGGGTTTTTGAT
ATGGATTANAAATCTGTGGGTTTTTAATATGGATTGGAAATCTGTGGGTTTTTAAT
ATGGATTAAAAACATCTGTGGGTTTTTAATATGGATTAAACATCTGTGGGTTTTT
AATATGGATTAAACATCTGGGTTTTTAATATGGATTAAACATCTGTGGGTTTTTAA
TATGGGTTAAAAATCAAAGAAAATGAACTATTTGCTCCAGTGCAGGAAAATACAG
15 GCAATACTGGATACAATTAGATGGTCAGGAGCGATAACCCGGTTGCCATTGTTTGA
AGAAGAGAATAAGGNGCTAGCATTCCTATCCGTAGATAATTTGACAGCTAGGAAAT
AGGGGGAGTCTTCTATGTAGTTAGTGAAGGCTAAATGAACTATTATATGC

Sequence ID 314

CTTTTCCTCCCGCTGTCCCCACGGAGGGGACTGCTCTCCCCGCTGCATCCTTTC
TGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAACCTC
GACATTCGGGAAGTGTTTTTGAGAAGTCTCGGTCCGTAAGGGAAGTCTTCCAAGTC
CGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTTCGGCCACCCCCCAGATGAGG
CAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTTCTCTCAGCTT
25 CCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGTCCATGCTGCCAT
CAAGGCATTTATTGCAGTGTACTATTTGCTTCCAAAGGATCAGGCCCTGAGAACAA
TGACCTTATTTCTACAACAGTGTCTGGGTTGCGTGCCAGCAGATGCCTCAGATAC
CAAGAGATAACAAAGCTGCAGCTCTTTTGATGCTGACCAAGAATGTGGATTTTGTG
AAGGATGCACATGAAGAAATGGAGCAGGCTGTGGAAGAATGTGACCCTTACTCTGG
30 CCTCTTGAATGATACTGAGGAGAACAACTCTGACAACCACAATCATGAGG

Sequence ID 315

TGGTACAGATACAACTGGACTCTCAGGACAAAACGACACCAGCCAAACCAGCAGC
CCCTCAGCATCCAGCAGCATGAGCGGAGGCATTTTCCTTTTCTTCGTGGCCAATGC
35 CATAATCCACCTCTTCTGCTTCAGTTGAGGTGACACGTCTCAGCCTTAGCCCTGTG
CCCCCTGAAACAGCTGCCACCATCACTCGCAAGAGAATCCCCTCCATCTTTGGGAG
GGGTTGATGCCAGACATCACCAGGTTGTAGAAGTTGACAGGCAGTGCCATGGGGGC

- 129 -

AACAGCCAAAATAGGGGGGTAATGATGTACGGGCCAAGCACTGCCCAGCTGGGGGT
CAATAAAGTTACCCTTGTACTTG

Sequence ID 316

5 CGCCACTTATCCAGTGAACCACTATCACGAAAAAACTCTACCTCTCTATACTAAT
CTCCCTACAAATCTCCTTAATTATAACATTCACAGCCACAGAATAATCATATTAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Sequence ID 321

10 CAGAACAGTACTTTTTTAATTTGACCCATGAATTCTATTTAAATTTGTCACTTAATA
TTTAGCCAAGAAGCAAACCATCTAAAAAGATTTCTGGTTTATTTCTCCAACCTCCTA
ATAAATAGGGTCACATATTTTTTAACTTTTTTCTAATTTGAAAAGTAATACAGGCA
TATGGTATTTTAAAAATGAAACAACACAAAGGGATATGTTTTGAAAAGTGGTTCTT
GCCATCCCTGAACTGTAATCATCCCTAACATATTCATACCTGTTTTCATTTTTAAAA
15 GTTGGGTCAGTTTTTTTTATTAGTACATGTATTTCTATCCTACTGATTTATTTGCTA
TATCATCTAATTTAGTTTGAATATTCCATAATTTACTTAATTAGTCCTGTATGGAG
ACCTAGCTCTTCTCAGTGTCTACTATTATAACAATGCTACAGTGAATATTGGTGT
ATAAATCCATACACACCACGTAACATATCTTAAGTTCCTGGAAGAGATATTGCTAA
ACCAGAAGATAACCTGCATTTAAAATTTTGA CTGCTAGGGTCAGGGTCACATTTAA
20 ATTAAATTAGAACAAGGAATGCATAATGTCTTCGATAGCAATCTATTCCAGGTGCA
CCGTGGTCACAAAGGAAAGCAAACTGTCAATAACTTTCTTCTCA

Sequence ID 322

TAGCATTTGGCCTTTTAAAACATTTGTTTATTTTTTTTTCTGAGAATGGCTAACACA
25 CTTTATTGAGGTTTCGAAATTAATAAAGAAAATAAAAGAAATGTATCTTCATTCATT
CTGTATGTTAGTGTTTTTAATTACCCTTAGAATATATGGATAAAAAATACTATTCTT
TGTCTTGGAGAAGGTAAGAGTCTAGTTAGATGAATAAGGGTTATCTATGTAGAACA
ACTAGAGAATGAGAAGAGAGCTTATGAGATTGAGTACTACGTTATGCAGTAGAGTA
GCACGTCATCTGCTACTGAGTATGGTGTGATAACATTGTGTAACAGGAAAGTATGA
30 TCAATATCTACTTAAAATTAAGGACAATATTAGCACTACATTGCTTTATTTTAAAG
TAAAAATTAGAGAACTAAACACAAGCATTGTAAGTACAATAAAAGCTGATCTTTCT
AGTTAAGCAGAATAATACATGTTCAAGCATCTGCTAAATCATTAAATATAAGAATA
TAGGGGTTTTCTATAATCTTATTTTCTTTGGAAGAGTACCTCATTTTCAAGANGAG
AAGTTTCTAATTGCCACTTCTTTAAAATAAAACAGGGTTTTAATGTTCCCAGCAC
35 AAAAATTAATATCTCTTCAAAAAGTCTCTTGTGATTAAAGTTTGAATCCCTTGTCAT
ACTGCTTCTAATATTGACACTGACCTCCTTAGGTATTTTTTCAGGGGTTATAATCTT
TTCTTAAAGGTATCTTTTTTCAAGAATTGGATACCTTGGGCTT

- 130 -

Sequence ID 323

CGCGTCGACTTTTAAAGTCATCTCTATAGGAAGGTGCTGGGCAGGGATCCCAGAGA
AAGAAAGGGTCCAAGACTCCATTAAGTCCCTGGATGAAGGGCACTGCTACAGCAG
CTAGTACCAGAGACTCTCCTATCTCACGGTTGAGGCAGACCCAGGATAGAATAGAG
5 AATAAAAGGAATGCTTATAGGAAACAATTTTGTATGGAATGCTAGATGGCCAAGCC
TCAGCCTTTGGTCCAGTGCAACCCTTGCCTCGCTTGTCAACAGTGAAAAATTAGTT
TGGTTAGAAGAACCATCTGGAAACACACCAGCTTCTGCTACCTTCATGCTCATTGT
TAAAAAAGATTAAACCAGTGTGAACATTCTGATCTGTTAATTCCAGGGACTGTTTT
CTTTCCAATGGACTGTTTGTGGTAGAATAACCCCCAAAAGCTCAAAGCTAAAATG
10 CATCATCAGTCCTAGTCGGCAGTTCCTTAAGAATGGACTGGCGGCGTGGTTGAGCT
GATATGGAAAAGCTGCACCTTCCTGCAGAAGATCAACTGACCTGCTATCCCACCCC
AAATTCAACCTGAGGTATATTTTCAGTGAAGCAGGTAGCTGTGCTTCTCAAAGCAGA
GAAGCAGTTTTTAAGAACCACAAAAGGTAGAGGAAATCTA

15 Sequence ID 324

GTTTGTACAGGCAGAATTGGATAGATACAGCCCTACAAATGTATATGCCCTCCCC
TGAAAAAATTGGATGAAAATCTGCACAGCAAAGTGAAACACACAGATAATAGGAA
CAAAATGTAGTTCCCATGTGCCAAACAAAATAAATGAAATCTCTGCATGTTTGCAG
CATATCTGCCTTTTGGGAATGTAATCAAGGNATAATCTTTGGCTAGTGTTATGTGC
20 CTGTATTTTTTTTAAAATGGTACACCAGAAAAGGACTGGCAGTCTACTTCTACCATA
GTTAAACTTCACCCTCTTTAATTTCAACATATTTCTTTGGAAGCAGGAAGAAATG
CTCATAAAGAGGATCAGACCTTCTTTCCCGTGAAACCAGTATTTGGCGCCATATAT
AAGCCTGGTTAAATTGGTCATCTAAAGCTGTCAAATAAGACATTCTGTGAAAGGTA
AACATCGAAACTGGTTATAAGTAAAACCATCAAGCCAACAACAGGGTCTTGAGATA
25 ACCTTTGAAGCTTATTGTCTGGCCTGCACCAGAAGATGTCTGCATTACTCATTGCT
AAAAATGTGTACACAGAACTGCACTAGGATTAATTGGTTCAAGAAGAAATTTAAAC
TTACGTTTGGGTTTCCATACAGCACTCTATTGAATACATGCATCTGAATTTAAGTT
GCAA

30 Sequence ID 325

GACCAGTAATGGCTTTTAAAGAGTCCATTTTGTCAATTGTCTCCCTAGTTAATTACAG
GTGGGGGATCTTTTGCCTCTATTCTCTTCATATTGAAATGAATCATACTCATGTTT
TGTGGAACCTCTTAAAGTTGTAGCTGTCATGATCAGATTTTTTTTATATTTCTCA
GCTTAACTCTGCTACTTGATTTACAGTGACCCATAACCTACTCATCCTTGGTTTAT
35 AGTGACACATAATCTTATCTCTTTATAGAACCTTAAATTTTATCATTATTTTCGCT
TAGAATACAGCATTTCTTTGCTTCTGTTGCTGGTTTGACTTAAGAAATAAGGCAGT
AACTCTGATCAATCAATTATCCATAAGGAAGGGCTTTTCATGGGTTCTATTAATTT

- 131 -

GTTAGTACCCTAAGTATATCTGAAAAATATGTCTATTGAGAGAAGATTTTGGCATT
CCAGATGGTATAGTCTATATATATTTAAAGTTTTGAATTTGCTTATATATACTCAG
CTTTCTTTTTCTAGCATTTTTGCATTTACCTGTTAATTGAAGTATACCCCCCACAT
ATAAAAGTTCCTCTTAAAGACACTGGACTCTTTCTGGGGGGCTAAAATA

5

Sequence ID - 326

nt: 554

CCCGGAATCGCGGCCCCGCGTCGACAACAAACCTGCATGTTCTGCACATGTATCCAG
GAACTTAAAAAAGATAGTTTGTGTGTCTTAATTGAATAATAGTAGATTT
ATAGATTAAAGATCTATGGGTTTTTAATATGGATTANAAATCTGTGGGTTTTTGAT
10 ATGGATTANAAATCTGTGGGTTTTTAATATGGATTGGAAATCTGTGGGTTTTTAAT
ATGGATTAAAAACATCTGTGGGTTTTTAATATGGATTAAACATCTGTGGGTTTTT
AATATGGATTAAACATCTGGGTTTTTAATATGGATTAAACATCTGTGGGTTTTTAA
TATGGGTTAAAAATCAAAAGAAAATGAACTATTTGCTCCAGTGCAGGAAAATACAG
GCAATACTGGATACAATTAGATGGTCAGGAGCGATAACCCGGTTGCCATTGTTTGA
15 AGAAGAGAATAAGNGCTAGCATTCCTATCCGTAGATAATTTGACAGCTAGGAAAT
AGGGGGAGTCTTCTATGTAGTTAGTGAAGGCTAAATGAACTATTATATGC

Sequence ID 327

CGGCTACCGACAGAAGGACTATTTTCATCGCCACCCAGGGGCCACTGGCACACACGG
20 TTGAGGACTTCTGGAGGATGATCTGGGAGGGGAAGTCCCACACTATCGTGATGCTG
ACGGAGGTGCAGGAGAGAGAGCAGGATAAATGCTACCAGTATTGGCCAACCGAGGG
CTCAGTTACTCATGGAGAAATAACGATTGAGATAAAGAATGATACCCCTTTCAGAAG
CCATCAGTATACGAGACTTTCTGGTCACTCTCAATCAGCCCCAGGCCCGCCAGGAG
GAGCAGGTCCGAGTAGTGCGCCAGTTTCACTTCCACGGCTGGCCTGAGATCGGGAT
25 TCCCGCCGAGGGCAAAGGCATGATTGACCTCATCGCAGCCGTGCAGAAGCANCAGC
AGCAGACAGGCAACCACCCCATCACCGTGCACTGCAGTGCCGGAGCTGGGCGAACA
GGTACATTCATAGCCCTCAGCAACATTTTGGAGCGAGTAAAAGCCGAGGGACTTTT
ANATGTATTTCAAGCTGTGAAGAGTTTACGACTTCAGAGACCACATATGGTGCAAC
CCTGGAACAGTATGAAATGTGCTACAAAGTGGTACAAGATTTATTGATATATTTCT
30 GATTATGCTAATTTCAATGAAGATCCTGCCTTAAATATTTTTTAATTTAATGGCAN
AT

Sequence ID 328

CAAGACTCCATCTCAAAAAAAAAAAAAAAAAATCTACAGTGCTGAGTATATAAAATTAT
35 TAACACATTTTACAACAATATGTGTTTGTGGAGTTAAATATTTTTTGTCTTTAAAA
CAGGTAATTTTAGTGCATACTTAATTTGATGATTAAATATGGTAGAATTAAGCATT
TTAAATGTTAATGTTTGTACATTGTTCAAGAAATAAGTAGAAATATATTCCTTTG

- 132 -

TTTTTTTATTTTAAATTTTTGTTCTCTGTAAACTAAAAGAACACGAAGTAATTGGTC
ACAATTACTGGTGTTTAACTGCCAAATATGGGTAAATAAGGGAAAATTTTGTTTAA
TATTTAGTCCTTCTGAGATGGCTTGAATATTTGAATTTTGTTGTACGTCTATACTG
GGTAGTCACAAGTCTTATAAACACTTTAGAGGAAAGATGGATTTTCAGTCTGTATTT
5 TTAAACATCATTTATTTTAAATCTGGTGCTGAAAAATAAGAAAAAAATTAAACTGC
ATTCTGCTGTTCTTCTTTANAAGCATTCTGCGTAAATACTGCTGTAATACTGTCA
TGCAAAGTGTATCCTTTCTTGTCGTATCCTTTTTGGGGCAGTGGTTTTT

Sequence ID 330

10 GCGGGAATCGCGGCCCCGCGTCGACCTCAAAGGAGAAAAAAACCTTGTA AAAAAG
CAAAAATGACAACAGAAAAACAATCTTATTCCGAGCATTCCAGTAAC TTTTTGTG
TATGTACTTAGCTGTACTATAAGTAGTTGGTTTGTATGAGATGGTTAAAAAGGCCA
AAGATAAAAGGTTTCTTTTTTTTTCTTTTTTTGTCTATGAAGTTGCTGTTTATTTT
TTTTGGCCTGTTTGATGTATGTGTGAAACAATGTTGTCCAACAATAAACAGGAATT
15 TTATTTTGCTGAGTTGTTCTAAAAA AAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAATTTTAAAATTTT TAAAATAAAACCTTG GTTAT

Sequence ID 331

GCCGCGTCGACCTGCATGAGCCACAGTTTCTTGACTGGAGGCCATCAACCCTCTTG
20 GTTGAGGCCTTGTTCTGAGCCCTGACATGTGCTTGGGCACTGGTGGGCCTGGGCTT
CTGAGGTGGCCTCCTGCCCTGATCAGGGACCCTCCCCGCTTTCCTGGGCCTCTCAG
TTGAACAAAGCAGCAAAACAAAGGCAGTTTTATATGAAAGATTANAAGCCTGGAAT
AATCAGGCTTTTTTAAATGATGTAATTCCCACTGTAATAGCATAGGGATTTTGGAAG
CAGCTGCTGGTGGCTTGGGACATCANTGGGGCCAAGGGTTCTCTGTCCCTGGTTCA
25 ACTGTGATTTGGCTTTCCCGTGTCTTTCCTGGTGATGCCTTGTTTGGGGTTCTGTG
GGTTTGGGTGGGAAGAGGGCCATCTGCCTGAATGTAACCTGCTAGCTCTCCGAAGC
CCTGCGGGCCTGGCTTGTGTGAGCGTGTGGACAGTGGTGGCCGCGCTGTGCCTGCT
CGTGTTGCCTACATGTCCCTGGCTTGTTGAGGCGCTGCTTCAACCTGCACCCCTCC
TTGTCTCATAGATGCTCCTTTTGACCTTTTCAA AATTAATATGGATGGGAAAGCTC
30 CTATGCCTTTTGGCTTCCTGGTAGAAGGCGGGATGCCCAAGGGTCTGCCTGGGTGT
GGATTGGATGCTTGGGGTGTGGGGGTGGA AACTGTCTTGTGGCCCACTTGGGCCC
C

Sequence ID 335

35 CCCGCGTCGACTTTTAAAGTCATCTCTATAGGAAGGTGCTGGGCAGGGATCCCAGA
GAAAGAAAGGGTCCAAGACTCCATTAACTGCCCTGGATGAAGGGCACTGCTACAGC
AGCTAGTACCAGAGACTCTCCTATCTCACGGTTGAGGCAGACCCAGGATAGAATAG

- 133 -

AGAATAAAAGGAATGCTTATAGGAAACAATTTTGTATGGAATGCTAGATGGCCAAG
CCTCAGCCTTTGGTCCAGTGCAACCCTTGCCTCGCTTGTCAACAGTGAAAAATTAG
TTTGGTTAGAAGAACCATCTGGAAACACACCAGCTTCTGCTACCTTCATGCTCATT
GTTAAAAAAGATTAAACCAGTGTGAACATTCTGATCTGTTAATTCCAGGGACTGTT
5 TTCTTTCCAATGGACTGTTTGTGGTAGAATAACCCCCAAAAGCTCAAAGCTAAAA
TGCATCATCAGTCCTAGTCGGCAGTTCCTTAAGAATGGACTGGCGGCGTGGGTGAG
CTGATTTGGAAAACCTGCCCTTCTGCAAAAAACACTGGCCTGCTTTCCA

Sequence ID 337

10 CAAGACTCCATCTCAAAAAAAAAAAAAAAAAATCTACAGTGCTGAGTATATAAAATTAT
TAACACATTTTACAACAATATGTGTTTGTGGAGTTAAATATTTTTTTGTCTTTAAAA
CAGGTAATTTTAGTGCATACTTAATTTGATGATTAAATATGGTAGAATTAAGCATT
TTAAATGTTAATGTTTGTACATTGTTCAAGAAATAAGTAGAAATATATTCCTTTG
TTTTTTATTTAAATTTTTGTTCCTCTGTAAACTAAAAGAACACGAAGTAATTGGTC
15 ACAATTACTGGTGTTTAACTGCCAAATATGGGTAAATAAGGGAAAATTTTGTTTAA
TATTTAGTCCTTCTGAGATGGCTTGAATATTTGAATTTTGTGTACGTCTATACTG
GGTAGTCACAAGTCTTATAAACACTTTAGAGGAAAGATGGATTTTCAGTCTGTATTT
TTAAACATCATTTATTTTAAATCTGGTGCTGAAAAATAAGAAAAAATTAACTGC
ATTCTGCTGTTCTTCTTTAGAAGCATTCTGCGTAAATACTGCTGTAATACTGTCA
20 TGCAAAGTGTATCCTTTCTTGTCGTATCCTTTTTTGGGGCAGTGGTT

Sequence ID 338

CTGGACTGCATGACCAGATCTGATGGGTGAGACTCAGGTGGCATGGAAGAGCCGAA
AGAGGATACCATATGTGGGTGCCGGGGGGGATAGGTGAGAAGTACTAGAAGGCGGA
25 ATGGAAGGACACTTCTGCTCAGCTCTGTGACACGGGCAGGGACCCTGCAGGGCTCA
GGTCCTTTAACACAGCAGCTTCATTCTAACACCAGCAGCGTTGGAACACACGTACA
AGTATGCAGACTAAGCTCTTGCTTGGCTGATACGGCTTTTTTGGGTTTTTAGAGAAC
ATGCATATATGTTCTCATTCATGGTACATGAACTCAGAAGCCTTACTGCCTATTTT
TGTTAATACTTCTGGGCAAACATTACCACTTACAACCTCACACCAGTTAGAAATCAT
30 TTGTAAAATGTTATTTAATAAAGCCAAAGAACTAAATCATATTTATTTTCCAAGN
TTTCTAAGATCTCTGAAACTAATGAGGTTTTTTTAAATCCCCATTAAGTACTCATCA
CTGCTAGTAAAAGCAGTTGTCTTTACCTTTAATTCCAGTGAGTCCCCTTAAATTTA
TTTTTTATTATCTTTGGCTACATTGCCTTAGACAAAATGTGGTCACCCTAATTTAA
NGGATAAAATTCACATCCTCACAGATTTCTTATTAAGAGGGTCTAANCCTTGAATA
35 ATCANCAGTGGAATGGAAGTCTTCTTTACTGGNTTTNATCCTTTCCCTTTTTTAT
CCCATG

- 134 -

Sequence ID 339

TTTTTTTTTAAATAAAGCTGTCGGCACTCAAGGGTAATTTTCATATCAGTGTGNTCT
ACAAGCTGGGGGAAAATGAGTTCTAATTGTCANAGCTACCAAATCCTTCACCTTTA
GCATAAAGGTTTAAAGATATCACAAAGATGCCAAGTGATTAATAATGTTTTAAACC
5 ACCCCTTTTTCTGTCTGAAAAACAACATAAACAATATTACAACAGTATAGTTACA
GAAGGGTTCTATTTTCATATGTTTTATGCACACTGTGCCTCAAAGGTACTATTTAA
ATATATATACTTTTGAGGGGGTGGCTAATGCAGAAACACCCAAGACCTAAGGAAGA
TACAACCCCATTTCTAGGTGTGAGGTCTAAATGCTTCACACACCCACTTGTGACCT
TTTTTCATGAAGAATCATAACACTGTGCAGTGAGAAACAGTGGCAAAGCAATACTG
10 AAAGCATTTTAAATTATTTACTAGGTTAAAAGGGTGAAGTACTTTAAATACAT
CAAATTCATCAT

Sequence ID 360

GCAAGTGAGAGCCGGACGGGCACTGGGCGACTCTGTGCCTCGCTGAGGAAAAATAA
15 CTAAACATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGC
ATTTTTTGTGCAAACCTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAG
TCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCT
AAAGAGAAAGGAAAATTTGAAGATATGGCAAAGCGGACAAGGCCCGTTATGAAAG
AGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAGAAGTTCAAGGATC
20 CCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTGCTCTGAGTATCGC
CCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAACT
GGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGG
CTGCGAAGCTGAAGGAAAAATACGAAAAGGTA

25 Sequence ID - 361 nt: 622

CTGTNATNGAATCTGCTTGTNACTNAAATGCTAAACTCAATTCTGTAATTCAATAG
GTGCACCTNTCTGAGAAACATANNAGACAATGAGGAAAAGGATTCANCATTCCGTG
GAATTTGTACCATGATCAGTGTGAATCCCAN TGCGTAATCCAAGTAAGATGTTCA
CAAAGATTTGTTTTTAATGTCTAATTAATAAAATTTTAAAGGAAGAAACATTCTAA
30 TACTTTAATTATAAAAAGTTAACTATTTTCAAAGGTATCAAATAACAGTTAAACCT
TTAAAATGTATATTTCTTAATATCTTGAAATTGTAATGCCTTTTTTTTTTCTAAA
TTTTTTTTTGTCATGAAATGAGATAGTAACAGCAGATTGGGACAACAAGGTTATATT
CTTGTCTTGAATCAGGCCATGGCTTCTTTCATCCAAATTTTCAAGACCTCATTTATTT
ACTTTGTCCCTGCCTCCCATCCCTGGATATCANGTTTGTGGATATCTACAGTTAAT
35 AGAGTGACCAAATAGTAGGAATACTGTCTCTCTATTCTGAATAAAATACTTTGAAT
CAGATTTAGAAATAATGAATAAAATACAAATCACCATTTGAAATTGCTCTAATTTTG
AGAGCT

- 135 -

Sequence ID - 363

nt: 628

ATCACNTGAGGCAAGAGTTTGAGCCAGCCTAGCTAACATGGTGAAACCCCATCTCT
ACAAAAATATAAAAATTAGCCTGGGTGGTGATGGGCACCTGTAACCCAGCTACTC
GGGAGGCTGAGGTAGGAGAATCACTTGAACCCGGGAGATGGAGGTTGCAGTGAGCC
5 AAGATCGTGCCACTGCACTCCAGCCTGTGTGACAGAACAAGACTCTGTCTCAAAAA
AAAATAATAATAATAATAATAAAAAAGGAATAACATAGCTAGGAATAAATTTAA
TCAAAGAGGTGAAAGACTTATACACTTAAAACTACAAAAAAAATCACTGAAGGA
ATTATAGACCCAAATAAAAAATAATAAAAAGACATTCTGTGTTTTAGGGAAAGAAG
ACTTAATATTGTTAAGATGTCAATACTACCCAAAGTGATCTACAGATTCAACATAA
10 TCCCTATCAAATTCACACAGCCTACTTTGTAGAAATGGAAAAGCCAATTTTCAA
TTCAGATGGAATTGCGAGGGGTCTGAATAACAAAAACAATCTTGGGGAAAAAAA
CAAAAAACAAAGTCAAAGAACTCACACTTCTCTATTTATAAATTTACTACAAAGTT
ATAGTAATCAAA

15 Sequence ID - 364

nt: 528

TGAACATCCAGCCATGTCATTTCTTCCATTCCCTGCCCTGGAGTAAAGTAGATTTAC
TGAGCTGATGACTTGTGTGCATTTGTACATTGCAACCTTAGCTTACCTCTTGAAGC
ATGTAGAGCATTCATCACCCACCATTCACTTCACTGCCTACTCCCACCACAGCTGTT
TCGTGGTCTGTCTGCTCCCTGTGCCACCCCCACCCCATCAGGTGGGCCTTTTGCAA
20 GTGATGAAGTCACCTGTGGGGGAAGAGCTTTCCTTTCCTCTCCTCAACTCAGAAGG
CCTCTTCCTCTTGCTCAAGAGGGTGCTGCTGCTTTCTGCCTCCTTCCCCGGCCGGC
CTCCATCCCAGTTCACCTTTTCAGAAATGGCCCCTCAGTCAACTCTTCCCTTTTCT
CCTGGCTTTTTTATTTCTCCAGTCTCTTAAGAGTATCCTTAGCTTTAAAAACAATA
ACACAGAGGATGGGTGCAGTGGCTCATGCCTGTAATCCCAGCACTTTGGAGCCTGG
25 GGCGGGCGGATCACTTGAGGNCA

Sequence ID 365

GTCCCGGAATCGCGGCCGCGTCGACCTTTTCTATGCCTGCTATATAAACAGTACCT
TGCAAGATGTCCTGTCTGATATCCACAAAGGGGTATTGTCAACCCCAAGTTCAGAC
30 AGCTTTGTATTCTTCTGTCCCTGGATACATGAATTACTGCCATCTTTACACAGCGC
CCTAAAATACCAACGCGAAGTTACCTGCTCAGCTTGAAGCTGCGCTGTACCCTGGA
ACCAGCACTTCTGCTGAATGACTCAGGATGAAGCCTCGACTTCTCCTTCCCATCCC
ATGCCCAGACCCCAGTGGCTCCTTTCCCAATCTGATCCAGTGACTTTAAGTCCAGC
TGTTGCAACCTGGGCATGAGGAGGAGTGCAAGATGGCTTTGTCCTACCTGGAAAGA
35 GGCTTTCTGGA

Sequence ID 366

- 136 -

CACCATTTACACACAGTGGGTCCTTGAATAGCATCGTTTTATTCAATGTCATTTTG
TTATAACATTGAGAAAAAATTGATTCCCGGCTGGGGCCACTGTCTGTGCACCGT

Sequence ID - 368

nt: 329

5 GAAAGATCTAAAATCGACACCCTAACATCACAATTAAAAGAACTAGAGAAGCAAGA
GCAAATTCAAAGCTAGCAGAAGGCAAGAAATAACTAAGATCAGAGCAGAGCTGAA
AGAGATAGAGACACAAAAAACCATTCAAAAAAAACAATGAATCCAGGAGTTTTTT
TTTTAAAAGATCAACAGAATTGACAGACTGCTAGCAAGACTAATAAAGAAGAGAG
AAGCATCAAATAGACTCAATAAAAAATGATAAAGGGGATATCACCACCAATCCCAC
10 AGAAATACAACTACCATCAGAGAACACTATAAACACCTCTATGCAAAT

Sequence ID 369

GAAAGATCTAAAATCGACACCCTAACATCACAATTAAAAGAACTAGAGAAGCAAGA
GCAAATTCAAAGCTAGCAGAAGGCAAGAAATAACTAAGATCAGAGCAGAGCTGAA
15 AGAGATAGAGACACAAAAAACCATTCAAAAAAAACAATGAATCCAGGAGTTTTTT
TTTTAAAAGATCAACAGAATTGACAGACTGCTAGCAAGACTAATAAAGAAGAGAG
AAGCATCAAATAGACTCAATAAAAAATGATAAAGGGGATATCACCACCAATCCCAC
AGAAATACAACTACCATCAGAGAACACTATAAACACCTCTATGCAAATAAAGTAG
AAAAT

20

Sequence ID 370

GAAAGATCTAAAATCGACACCCTAACATCACAATTAAAAGAACTAGAGAAGCAAGA
GCAAATTCAAAGCTAGCAGAAGGCAAGAAATAACTAAGATCAGAGCAGAGCTGAA
AGAGATAGAGACACAAAAAACCATTCAAAAAAAACAATGAATCCAGGAGTTTTTT
25 TTTTAAAAGATCAACA

Sequence ID 371

GCCCGGAATCGCGGCCGCGTCGACGTAAGCTCGGCTGAATCCACGGTTCAAGAACA
GGAAAGAAGGCCAAGGCATAGGGAGTGGGGCAGTTGGGTGAATATTAGTACCTTTC
30 CCTCAGNTNCATTAATTACCCCTGCCTACTCTGCACAAAAGGATNTAACAACAGTT
TCCTTTTTTAATGGCCAGGTACAGCTGCTTATATGGANGGGCATTNTNAATGATAT
CCTTNATCACTGTCTTAATCATCACATNCTTAAACAATCACTTTATTGTGTTAAG
GAAGATAAAAATGGCTGGGTTCATTTCCGTTCTGGAAGAAATCGANTNAAAAGGT
AACCATTTAATAATGCANAGGGCANTTTCCTGCAGACCCTAATACTGGAAATTTT
35 TAAAAACAAATGAAAACTTCTACTTTTTCTTCTAAGCTTACTTAACCACCCAAAT
TTTCCAGCCACATATCTTCCTAGTCTACAACCTGCCTTTAACTTTAAGAGATGCTCA
AAAAATGTAAATTCTCAAATACATTCTTATTACAATTACTGCTAACCT

- 137 -

Sequence ID 373

CCAGTGTGCTGGGATTACAGGCATGAGCCCTGCACCCAGCCTCTTAAACTGATCAT
ATGATATTGGTTCTCAACCAAGGGTGACTTTGCCCCCAGAGGATACTTGGCAATGT
CTGGAGATACTCAGTTGTCATGACTTGGACAGGTGCTACTGTCACCCAGTGGGTAG
5 AGGTCAGGGATGGTGCTAAACATAGGACAGCTGTCAAGAGAAAAGAATGTACCCAG
CCCCAAATGTCAGTAGGGCTGAGGTTGAGAAACCCAGCTGTAGCTGACGTGTGAAG
GACAGACTGGCCTGGAAGTGTGTTTTCTGCCCCCTTTCCACCCCTGCATATTAGTTA
AGGCCAAAGGAAAAAAGGAATGCAGGAAATGCCCGTTAAAAATCTTCAAACAATA
TAAAATGATCAATTCCACTAAAACCCTTTACACATTTAAGTATAAAGGTATTGGTA
10 GGAAAATTTGTTATTCAGTCTTTTCTCAGTGTGTCATGAAATAATTATTTCTGCTGT
CAGTTT

Sequence ID 374

AAAAAAAAAATCACTGAAGGAATTATAGACCCAAATAAAAATAAATAAAAAGACAT
15 TCTGTGTTTTAGGGAAAGAAGACTTAATATTGTTAAGATGTCAATACTACCCAAAG
TGATCTACAGATTCAACATAATCCCTATCAAATTTCCAACAGCCTACTTTGTAGAA
ATGGAAAAGCCAATTTTCAAATTCAGATGGAATTGCGAGGGGTTCTGAATAACAAA
ACAATCTTGGGGAAAAAAAACAAAAAACAAAGTCAAAGAACTCACACTTCTCTAT
TTATAATTTACTACAAAGTTATAGTAATCAAAGTCGACGCGGCCGCGATTCCGGG

20

Sequence ID 378

CGACTGCGGCTCTTCCTCGGGCAGCGGAAGCGGCGCGGCGGTCTGGAGAAGTGGCCT
AAAACCTTCGGCGTTGGGTGAAAGAAAATGGCCCGAACCAAGCAGACTGCTCGTAAG
TCCACCGGTGGGAAAGCCCCCGCAAACAGCTGGCCACGAAAGCCGCCAGGAAAAG
25 CGCTCCCTCTACCGGCGGGGTGAAGAAGCCTCATCGCTACAGGCCCGGGACCGTGG
CGCTTCGAGAGATTTCGTCGTTATCAGAAGTCGACCGAGCTGCTCATCCGGAAGCTG
CCCTTCCAGAGGTTGGTGAGGGANATCGCCCAGG

30 Sequence ID 380

GCAATTTAATTTTTTAATAACAAAGATACTGTATTTTAACATGGTGAAATATACTTG
GCTAAGTCCAGATTAAAAAAAAGTATCTAGCCCAACAGTACAATTATACAGCT
TTGTACAGAACATTCCATAGATCAACAGAAAATACATTTGAGCGCAAAAATAAAAA
ATATTTAAGGAGAATCTCTAAGCAGCATTTTATTTCTGCAAAAGACATATCTTGTC
35 TGATTAAATATCTACAAGTGCTTTTCTTTCAAATAACATATATTCTTAATAGAC
TAAGTCATTAACAATGACCTGGTAATTCTTTCACTTCAATTTGAATGATTTATAAG
CTAAATCTTCAACCACAAAAGGTTTTTATTTGTATTAAAGATGTTACCACTTTTGA

- 138 -

CAAAAAGCTTAAAATATTTTATATTTCAAAGGAAAATTAGCAACATAACTTTACAA
TATATTCTATGATATTTTGATTGTGAGGGCTACTCTATTTAAACTGATGATCTCT
GTTGTGTTGCTCAGATGCAGGAAAGCAGCAAAA

5 Sequence ID - 381 nt: 534

GACTTANATCTAAATGGACCACATTCTCTACTTAAAAAATGCTATTAACCATGTG
ATCTTCTCAGTCATGAGGTAATCTGGTGACTACCCCTTCCTCAAAGCCAGTTGGGAT
ATTCTTTGAATAGAGTAAACAGTGTTTCTAGGCTGGGAGACACCAGACATAGTTG
AGGACAGAGGTGCTAGAAAATAGGAAGTTTAAAAGCATGTGCGGTGATGCTCAGAG
10 GAGGTAAACCCCAACCTCATGCTCATAGCTTCCAATCATTTTCTCTAGTTCTTAAC
TCTTAAATGTGAGAAATGCTTGAAGATTCTAGTCATCTGAAGAAAGTCTCTTTATT
AAAGATTTTCATAAAAGAGACCAAAGCAGACAAACAGAAAAAGACATCTTGGGGAA
AAAAACAAGGATAATGGGAAGAGAAGGAAAGTTTTAAAAATTATCAATATCCTCAG
GGGGACAAAATATTATATCCTATAAAGACAGATTTTTTATTTTTTAAAAAAATAGAA
15 AGCAAAACAAGCTCCTAAAAATAAAGTTTG

Sequence ID - 382 nt: 444

GTTAAGGAAGTCAGCACTTACATTAAGAAAATTGGCTACAACCCCGACACAGTAGC
ATTTGTGCCAATTTCTGGTTGGAATGGTGACAACATGCTGGAGCCAAGTGCTAACA
20 TGCCTTGGTTCAAGGGATGGAAAGTCACCCGTAAGGATGGCAATGCCAGTGGAACC
ACGCTGCTTGAGGCTCTGGACTGCATCCTACCACCAACTCGTCCAACCTGACAAGCC
CTTGCGCCTGCCTCTCCAGGATGTCTACAAAATTGGTGGTATTGGTACTGTTCTTG
TTGGCCGAGTGGAGACTGGTGTTCTCAAACCCGGTATGGTGGTCACCTTTGCTCCA
GTCAACGTTACAACGGAAGTAAAATCTGTCGAAATGCACCATGAAGCTTTGAGTGA
25 AGCTTTTCCTGGGGACAATGTGGGCTTCAATGTCAAGAATGTGTCTGTCAAG

Sequence ID - 383 nt: 566

CTTTGAAGAACTTTGCCAAATACTTTCTTACCAATCTCATGAGGAGAGGGAACATG
CTGAGAAACTGATGAAGCTGCAGAACCAACGAGGTGGCCGAATCTTCCTTCAGGAT
30 ATCAAGAAACCAGACTGTGATGACTGGGAGAGCGGGCTGAATGCAATGGAGTGTGC
ATTACATTTGGAAAAAATGTGAATCAGTCACTACTGGAACCTGCACAACTGGCCA
CTGACAAAAATGACCCCCATTTGTGTGACTTCATTGAGACACATTACCTGAATGAG
CAGGTGAAAGCCATCAAAGAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGG
AGCGCCCGAATCTGGCTTGGCGGAATATCTCTTTGACAAGCACACCCTGGGAGACA
35 GTGATAATGAAAGCTAAGCCTCGGGCTAATTTCCCATAGCCGTGGGGTGACTTCC
CTGGTCACCAAGGCAGTGCATGCATGTTGGGGTTTCCTTTACCTTTTCTATAAGTT
GTACCAAAACATCCACTTAAGTTCTTTGATTTGTCCATTCCTTCAAATAAAGAAAT

- 139 -

TTGGTA

Sequence ID 384

TTTTGGGGTTTATATATAAGCCTGGTTCTTGCTGAAACTGCTTATGTTGATAACCA
5 GTTAGTGAGTTCCTCTCTATTGACTTGCTGGGAAGTTTATAGAGACATTTTTTATG
CATTCAGAGATTTTCAGTACAAATCTTGAAAAAGGGACATTTAGGCCGGGCGCGGTG
GCTCACATCTGTAACCCTAGCACTCTGGGAGGCTGAGGTGGGTGGATCATGAAGTC
AAGAGATAGAGACCATCCTGGCAAAAATTAGCTGGGCGTGGTGGGGTGCGCCCGTA
GTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAGCCCGGGAGGCGGAG
10 GTTTCATTGAGCCGAGATAGTGCCACTGCACTCCAGCCTGGACAACAGAGCGAGAC
TGTGTCTT

Sequence ID 386

CTAAGGGTTTAAAGATGGAAAGAGGCATTGATGAACAGCTGGGGAAGGAGTAGTTT
15 GAGGTAGATGTGCAGATGGAATGAAGAGAAGGTCTCAAGAAGAGGGTGGAGCCAAA
GAGGGCTGCAGATTTAGAAGGCTAAAGTCTTTAGATGGCTTTGGATAGCCTGTTGT
ATCTTGGACCATGCAGGTTACAGTGGAGCATGGAGTGGGGACAGAAGTGGAGGAAG
GAACCAGGGAACATGGAGTGAGAAGCTAAAGGAAAGTGATGCAGTAGATACATGGC
TCTAAAGTACTCAGGACTTTCAGAGGCTTAAACATAGGGTGACCAACTATCCCCT
20 ATGCCTGATACTAAGGGCATTCCCTGGATGTGGACCTTTCATTCCCCAAATTAGGA
AAGTCTTGGGCATACCAAGACAAGTTGGCCACCCTACTCAAAGTATGTAAGCTAA
CATATCTGTTCTCTAAGAGGTTAAAGCTGGATGGGGATACCAGATGTATGTACGTG
ATGCAGTTAAACAGCAATACAAGGGGGCAAGTCTACCTGATCGGCCAATTCAATGG
GA

25

Sequence ID 387

GAAGCCAAACCAAAGGAGCTTCTACTTCATGATGCCATTTATGTAAAGTTCAGGCA
GAGAAAATCAGTGGTTTAAGAAGTTAGAATAATGATTATCTTTGGAGGGATTGCAA
CTGGAAGAAGTCATGATTGGGATTTCTGGGTCCTAATAGTGCTCTGTGTCTTGATC
30 TGAGTGCCGACTACATGAGTGGTTAGGTTTGCAAAATTCATTGAGTTATGCACTTA
ATGGTGTTGTCTTATTAGAGCTGATGGAGGAGAGAGGGCTTCAATTTGCACAACCTG
AGTAATCAGCTAGGCCCGAGTCACTAGGTGAACAACCTTACTGCTCCAATCAGCCTTA
GAGCAGGAATCAAACCTCATGTCTCAGAAAAGTTATTAATTCAGCTTGTCTTGGGAC
TTCCTTCAGAGTCACTCTTGAATAGCTGAAATAGTAAATGTTAAATCTGTGGATGC
35 AAGTGTGTAAATTATTTTAGTCATCAGCTCTAATAAGATGGCCTTTGGGGAAATGA
GTATAAGGTCACGAAAATGAAATGGCAAGAAGGAGGTCTACTATTTCTTCTGTAAT
ACTGATTTTTTACCCCATCAGGGTCAGTCCCCAGAGGTTGTAAATGTGAAGCTTG-T

- 140 -

CTTTTTCTTTAATAA

Sequence ID 388

CTTTGGACACTAGGAAAAACCTTGTAGAGAGAGTAAAAAATTTAACACCCATAGT
5 AGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACACCCACTACCTA
AAAAATCCCAAACATATAACTGAACTCCTCACACCCAATTGGACCAATCTATCACC
CTATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCCTCCGCATAAG
CCTGCGTCAGATTAAAACACTGAACTGACAATTAACAGCCCAATATCTACAATCAA
CCAACAAGTCATTATTACCCTCACTGTCAACCCAACACAGGCATGCTCATAAGGAA
10 AGGTTAAAAAAAGTAAAAGGAACTCGGCAAATCTTACCCCGCCTGTTTACCAAAAA
CATCACCTCTAGCATCACCGAGTATTAGAGGCACCGCCTGCCCAGTGACACATGTTT
AACGGCCGCGGTACCCTAACCGTGCAAAGGTAGCATAATCACTTGTTCTTAAATTA
GGGACCTGTATGAATGGCTCCACGAGGGTTCAGCTGTCTCTTACTTTTAACCAGTG
AAATTGACCTGCCCGTGAAGAGGCGGGCATAACACAGCAAGACGAGAAGACCCTAT
15 GGAGCTTTAATTTATTAATGCAAACAGTCCTAACAAACCCCAGGTCCTAAACTCCA
AACCTGCATTAAA

Sequence ID 389

CGACCCGGAATTCGCGGCCGCGTCGACTGAGTTCTTGACAAGAGTGTTTTTCCCTT
20 CCCGTCACAGAGTGGGCCCAACGACCTACGGCACTTTGACCCCGAGTTTACCGAAG
AGCCTGTCCCCAACTCCATTGGCAAGTCCCCTGACAGCGTCCTCGTCACAGCCAGC
GTCAAGGAAGCTGCCGAGGCTTTCTTAGGCTTTTCTTATGCGCCTCCACGGACTC
TTTCTCTGAACCCTGTTAGGGCTTGGTTTTAAAGGATTTTATGTGTGTTTCCGAA
TGTTTTAGTTAGCCTTTTGGTGGAGCCGCCAGCTGACAGGACATCTTACAAGAGAA
25 TTTGCACATCTCTGGAAGCTTAGCAATCTTATTGCACACTGTTTCGCTGGAAGCTTT
TTGAAGAGCACATTCTCCTCAGTGAGCTCATGAGGTTTTTCATTTTTATTCTTCCTT
CCAACGTGGTGCTATCTCTGAAACGAGCGTTAGAGTGCCGCCTTAGACGGAGGCAG
GAGTTTCGTTAGAAAGCGGACGCTGTTCT

30 Sequence ID - 390

nt: 523

GAATCCCTAGAAAAAGAGAATTCCCAACTTGATGAGGAAAACCTTAGAACTGCGAAG
GAATGTAGAATCTTTGAAGTGTGCAAGCATGAAAATGGCTCAGCTACAGCTAGAAA
ACAAAGAACTGGAAAGTGAAAAAGAGCAACTTAAGAAGGGTTTGGAGCTCCTGAAA
GCATCTTTCAAGAAAACAGAACGCTTAGAAGTTAGCTACCAGGGTTTAGATATAGA
35 AAATCAAAGACTGCAAAAAACTTTAGAGAACAGCAATAAAAAAATCCAGCAATTAG
AGAGTGAACCTACAAGACTTAGAGATGGAAAATCAAACATTGCAGAAAAACCTAGAA
GAACTAAAAATATCTAGCAAAAGACTAGAACAGCTGGAAAAAGAAAATAAATCATT

AGAGCAAGAGACTTCTCAACTGGAAAAGGATAAGAAACAATTGGAGAAGGAAAATA
AGAGACTCCGACANCAAGCAGAAATTAAAGATCCACATTTGAAGAAAATAATGTGA
AGATTGGAAATTTGGAAAA

CTTTGAAGAACTTTGCCAAATACTTTCTTACCAATCTCATGAGGAGAGGGGAACATG
CTGAGAAACTGATGAAGCTGCAGAACCAACGAGGTGGCCGAATCTTCCTTCAGGAT
ATCAAGAAACCAGACTGTGATGACTGGGAGAGCGGGCTGAATGCAATGGAGTGTGC
ATTACATTTGGAAAAAATGTGAATCAGTCACTACTGGAAC TGACAAACTGGCCA
CTGACAAAATGACCCCCATTTGTGTGACTTCATTGAGACACATTACCTGAATGAG
CAGGTGAAAGCCATCAAAGAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGG
AGCGCCCGAATCTGGCTTGGCGGAATATCTCTTTGACAAGCACACCCTGGGAGACA
GTGATAATGAAAGCTAAGCCTCGGGCTAATTTCCCCATAGCCGTGGGGTGACTTCC
CTGGTCACCAAGGCAGTGCATGCATGTTGGGGTTTCCTTTACCTTTTCTATAAGTT
GTACCAAAACATCCACTTAAGTTCTTTGATTTGTCCATTCCTTCAAATAAAGAAAT
TTGGTA

20 GACCCGGAATCGCGGCCGCGTCGACCATTTTAGCCAAGGTGCCTCTATAGGGGTCA
AGACATCATGTGCCCAGACCTAAGGTCAGGAATGTCATATTTTTCTGTTAAAATCA
TTTTATTTCTGTGTATCTTACCTTTAAATCATTGTGGTTTACTCTGAGATTCTGTA
GTCCTAATATTGTATCATTGTGCTGTCTGCAAACAACCTTGAATCTATTTTGTTTG
CATCTTTTGTTACATGTAACGCAGCTGTACTTTATGTTCTTTGCAACTGTTTCCAT
TATGAGAACGCTGTGCTATTTACAAGGTTACATTTTTCTTGGCCAGGCGAGGTGGT
25 CATGCCTGTGATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACTTGAGGTAA
AGAGTTGAGACCAGCCTGGCTAGCATGGCGAAGCCCAGTCTCTACTAAAAATACAA
AAATTGGCCGGGTGAAATTAGCCGGGCGTGGTGGTGTGTGCTTGTAATCCCAGCTA
CTCGGGAGGCTGAGGCAGGAGAATCGCTTGAATCCGGGAGGCAGAGGTTGCAGTGA
GCCAAGATCANGCCACTGCACTCCACCTCGGGGTCAAGAGCGAAACTCTGTCTCAA

CCGTTTTAGTCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCTGCCTCGGCC
TCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCGTAAATCAGGTTT
TTTAAATGTTTGCCAAACCTTATCACTGACTTTTATAACAAAATTATTTACTATAA
35 TCATTAGGGAATATTTAAGTTCTGCTAATACTTAAAATTGCAGAGTGCTAAAACCA
GCAGTGAGTTTAGAATCAAGCTAAGCTTTATTGTTGCTACTATTTGAGGCATATTA
GTTGACTGGTGTTTCATATGCAAGGCAGTCTACTGGGTGCAACAAGGGTTAGAAGGA

- 142 -

TATTTTAAAAAACTGACCCTATTCTCAGGATGAAAATAATACACTAGTAATAGTC
TGCTCTGTTGGTTAACTCCTCGTAAGGAGGTCAATTAAAATGCTGTAGTGTTGCAA
GGGAAGGAGAGGAAGAATCATATTCCTTCACTAGCAGGATCAAGAAAGCTTTTATA
GAAATATACAAAATCTTCACTTCTTGAAGGATTGGTAAAATTTAATAGCCAACATT
5 GGGCACTTATTCATTCTCTGAGTAAATATTTATTGTCAT

Sequence ID 396

CTTAAATCTAAATGGACCACATTCTCTACTTAAAAAAATGCTATTAACCATGTGAT
CTTCTCAGTCATGAGGTAATCTGGTGACTACCCTTCCTCAAAGCCAGTTGGGATAT
10 TCTTTGAATAGAGTAAAACAGTGTTTCTAGGCTGGGAGACACCAGACATAGTTGAG
GACAGAGGTGCTAGAAAATAGGAAGTTTAAAAGCATGTGCGGTGATGCTCAGAGGA
GGTAAACCCCAACCCTCATGCTCATAGCTTCCAATCATTTTTCTCTAGTTCTTAACTC
TTAAATGTGAGAAATGCTTGAAGATTACTAGTCATCTGAAGAAAGTCTCTTTATTA
AAGATTTTCATAAAAGAGACCAAAGCAGACAAACAGAAAAAGACATCTTGGGGAAA
15 AAAACAAGGATAATGGGAAGAGAAGGAAAGTTTTAAAAATTATCAATATCCTCAGG
GGGACAAAATATTATATCCTATAAAGACAGATTTTTATTTTTTAAAAAAATAGAAA
GCAAAACAAGCTCCTAAAAA

Sequence ID - 397

nt: 534

20 GACCCGGAATCGCGGCCGCGTCGACGGAAGCTCCTGCCCTCCTAAAGCTGAAGCC
AAAGCGAAGGCTTTAAAGGCCAAGAAGGCAGTGTTGAAAGGTGTCCACAGCCACAA
AAAGAAGGAGATCCGCACGTCACCCACCTTCCGGCGGCCGAAGACACTGCGACTCC
GGAGACAGCCCAAATATCCTCGGAAGAGCGCTCCCAGGAGAAACAAGCTTGACCAC
TATGCTATCATCAAGTTTCCGCTGACCACTGAGTCTGCCATGAAGAAGATAGAAGA
25 CAACAACACACTTGTGTTTCATTGTGGATGTTAAAGCCAACAAGCACCAGATTAAAC
AGGCTGTGAAGAAGCTGTATGACATTGATGTGGCCAAGGTCAACACCCTGATTCGG
CCTGATGGAGAGAAGAAGGCATATGTTGACTGGCTCCTGATTACGATGCTTTGGA
TGTTGCCAACAAAATTGGGATCATTTAAACTGAGTCCAGCTGCCTAATTCTGAATA
TATATATATATATATATCTTTTCACCATAA

30

Sequence ID - 398

nt: 512

GGGGAGCCCCCTCTTCCCTCAGTTGTTCTACTCAGACTGTTGCACTCTAAACCTA
GGGAGGTTGAAGAATGAGACCCTTAGGTTTTAACACGAATCCTGACACCACCATCT
ATAGGGTCCCAACTTGGTTATTGTAGGCAACCTTCCCTCTCTCCTTGGTGAAGAAC
35 ATCCCAAGCCAGAAAGAAGTTAACTACAGTGTTTTCTTTGCACCGATCCCCACCC
CAATTCAATCCCGGAAGGGACTTACTTAGGAAACCCTTCTTTACTAGATATCCTGG
CCCCCTGGGCTTGTGAACACCTCCTAGCCACATCACTACAGTACAGTGAGTGACCC

- 143 -

CAGCCTCCTGCCTACCCCAAGATGCCCTCCCCACCCTGACCGTGCTAACTGTGTG
TACATATATATTCTACATATATGTATATTAAACTGCACTGCCATGTCTGCCCTTT
TTTGTGGTGTCTAGCATTAACTTATTGTCTAGGCCAAAGCGGGGGTGGGAGGGGAA
TGCCACAG

5

Sequence ID 399

TTTTGGCATTACTTAATCCAATTATAAAACTGAATTTTTTAAAAACAGCACTTGT
TTTTTCTTCCAAGATTAATTTGAATTTTTTTTATGGACATTAGAAAACATTGCAGTT
TAGTCATAATCAAAAATAAATCTTGAGGCTGGTAGAGCAGCTTTGTTGCTGTTTAT
10 ATTTTTTATTGCTTACTGGATTTTCAGTGTTACCTAGTGCCATCAGTTTGGTATTTTG
CCACCTTGCACATTCAGTGATGTTTGATTTTTCTTTTTCTTTTTTCATATTACT
TTTAAATCCTGAATAGTTTGTGGCAGCTGGAGATCACCTAGTCCACCACTGTCCAA
CATGGCAATGGTAAGTAATATTGAGTAAAGAATAGAAAATTAGTAAAATGCATGGC
TTCAGAATTATAGCAATTTGCAAAATAGGTTAATGGATGAAAATTAGAATGACCAG
15 TTTAACTTTCCCCCAGCAGATTCTTCTGTAAACAATGCCCTTCAAAATAAAGG
AAGAACAAGTGGGTGTTATACCTATGTTATTTGGCTATGTTAGCACAATATGATGG
ACTAATTTGAGAAAAAGCATTTACTTCCTTTACTATTACTTCTTTTCTTTATAGGG
CTAAGTCTGCCTTCTGGGTCTTTGAA

20

Sequence ID 400

GAAGAAGCGCGAAGAGCCGTTAGTCATGCCGGTGTGGTGGCGGCGGCGGAGACTGC
GGGCCCCTAGCTGGGCTCTGCGAGGTGCAAGAAAGCCTTTGAGGTGAAGGTGTATG
AAAGTCATCATAACAGATGTTTTCCAAAACTTGTAGAAGGTGTGAAAAAACTAC
TAGGATCACGCGGCATGTATTGAGCATATAGGTTGCTGTAGATGAATGTTCTTAGC
25 TGTCAATGTTTAAAAATACTTCTGCTTCGTTACCTCAAGTGTGGCATGCAGCATTTT
GGAAGGAAAATTGAAGACGTGTTCAAGAAAACATGAACAGAAGCAAATGATGAAAA
TGAGCATTTTACTTGATGTTGATAACATCACAATAAATTATGGAGAAAAATACATA
TTTGGCTAACTTTTAATTGCTGAACAATAAAGTGTTTTCTTTTAAATCNAAAAA

30

Sequence ID 401

GAAGCCAAACCAAAGGGAGCTTCTACTTCATGATGCCATTTATGTAAAGTTCAGGC
AGAGAAAATCAGTGGTTTAAGAAGTTAGAATAATGATTATCTTTGGAGGGATTGCA
ACTGGAAGAAGTCATGATTGGGATTTCTGGGTCCTAATAGTGCTCTGTGTCTTGAT
CTGAGTGCCGACTACATGAGTGGTTAGGTTTGCAAAATTCATTGAGTTATGCACTT
35 AATGGTGTGTCTTATTAGAGCTGATGGAGGAGAGAGGGCTTCAATTTGCACAACT
GAGTAATCAGCTAGGCCAGTCACTAGGTGAACAACTTACTGCTACCAATCAGCCT
TAGAGCAGGAATCAAACCTCATGTCTCAGAAAAGTTATTAATTCAGCTTGTCTTGGG

- 144 -

ACTTCCTTCAGAGTCACTCTTGAATAGCTGAAATAGTAAATGTTAAATCTGTGGAT
GCAAGTGTGTAAATTATTTTAGTCATCAGCTCTAATAAGATGGCCTTTGGGGAAAT
GAGTATAAGGTCACGAAAATGAAATGGCAAGAAGGAGGTCTACTATTTCTTCTGTA
ATACTGATTTTACCCCATCAGGGTCAGTCCCCAAAGGTTGTAAATGTGAAGCTTG
5 GTCTTTTCTTTA

Sequence ID 402

GACCCCTATTCTCAGGATGAAAATAATACACTAGTAATAGTCTGCTCTGTTGGTTAA
CTCCTCGTAAGGAGGTACAATTAAAATGCTGTAGTGTTGCAAGGGAAGGAGAGGAA
10 GAATCATATTCTTCCTTCACTAGCAGGATCAAGAAAGCTTTTATAGAAATATACAAAAT
CTTCACTTCTTGAAGGATTGGTAAAATTTAATAGCCAACATTGGGCACTTATTCAT
TCTCTGAGTAAATATTTATTGCATGCTTATCTTGTATCAACATTGNGATGAAAGCN
CAAGAATGAAAGAGGAGGGAGAATGTTTANAGAATAAGGCTGAAACACAGATTTTG
TAGGGAGCGTAGGGGAGACTGANAAAACAG

15

Sequence ID 403

AAGACACCTGATAGATTGTCTTGTATTATTTTTCCTTTGCCTTCTTACAATCTCAG
TGATTAGAATTGGGCTGAAAACAATACATCAAATTCTCAGCAAATCCTTATGGGT
TGCTGGATACCGAGGGTTTTTAAGATCTTTAGACTTCACTATATAGAACAAATGTT
20 GAATGGGAATTTTCTTTATTTCTATANCGTTTNG

Sequence ID 405

CCCGGAATCGCGGCCGCGTCGACGATGAGCATTTTTTTCATGTGTCTTTTGGCTGCA
TAAATGTCTTCTTTTGAGAAGTGTCGGTTCATATCCTTTGCCCACTTTTGGATGGG
25 GTTGTTTTTTTCTTGTAAATTTGTTTGAGTTCATTGTAGATTCTGGATATTAGCCC
TTTGTCAGATGAGTAGGTTGCGAAAATTTTCTCCCATTTTGTAGGTTGCCTGTTCA
CTCTGATGGTAGTTTCATTTGCTGTGCAGAAGCTCTTTAGTTTAATTAGATCCCAT
TTGTCAATTTTGGCTTTTGTGTCATTTGCTTTTGGTGTTTTTAGACTTGAAGTCCTT
GCCCATGCCTATGTCCTGAATGGTAATGCCTAGGTTTTCTTCTAGGGTTTTGATGG
30 TTTTAGGTCTAACGTTTCAGTCTTTAATCCATCTTTTAAAAGTCTCTTCACAGTAC
ATGAGTAGTAGTGACACCAATAATGTCAGAGCAGGGAAGTCCCAGGTTCTGCCCAT
CCACAAAAACAACAAATAAGCTGGCAAAAACCTTTAAGAATCAACTTTTGCAGATCT
CTGAAATCTAGTCAAACTTAAACAGAGGAAAGATTAATAAAGACNGGCTGCCTGA
GATAACACTAACACACAC

35

Sequence ID 406

CATCAAATAAATAAATAAATAAATTTTAAAAGTCACAGCATTTGAATTTTTAAATGT

- 145 -

TTGGGATGATAAAGCACCTGCTTATCATGAAGCTANAGAAATTCAATGACACGTTT
GCCAGGGTCTTTGCTAGTGATGTTGGAACAAGTCTGTAATGCTGATGAAACATCAC
TGTTCTGGGCATTATTGCCCCAGAAAGACACTGACTGCAGCTGATGAAACAGCCCTT
CCAAGAATTAAGGATGCCAAAGACCAAATAACTGTGCTGAGATATACTTACGCAGC
5 AGGCATGCATAAGTGTAAACTTGCTGTTATAAGCAAAAGCTTGCGTTCTCACTGTT
TTCAAGGAGTGAATTTTCATACCAATCCATTATTATGCTAATAAAAAGGCATGGATC
ACCAGGGACATCTTTTCAGATTGGTTTCACAAACATTTTGTACCAGCAGCTTGTGC
T TACTGCAGGGAAGCTGACTGGATGATGACTGCAAGATTTTGTATATCTTAACAA
CTGTTGTGCTCATCCTCCAGCTGAAATTCTCATCAAAAATAATGTTTATGGCTCAC
10 ACCTGTAATCTCAACACTTTGGGAGGATTGCCTGACCCAGGAGTTCAAGCCCACCC
TGGGCAACACAGCAAGACCCAACCTNTC

Sequence ID 407

TTTTAAAAATCATAAAACGTTTCTTACAAAAGAGCATTACATTNTGCACACTGCTC
15 TGAACAGATGCCAGGGACATGTGGACTATTGTTACTTTTCTCCTCCCTGTCCCACCCC
CCAAATGTTACAGTGACCACAAAGCAAGGTGTTTACAATAATTACATGGGGGGAAT
TTTTTAAACCACCAACAATAACGAAAAATAAAATCCACTCACTCTGCTGCTGTTTC
AAAATTTCAATGTTAGTTTTTGCACGCCCTTCCCCCCCCCAACCCTGTTTGTAAGG
AACTAAAACATTACATCTGGTGAACAGCAAAGATTTCACTACACCTCAAATGCAGA
20 ACACCTATGAAGCAGAGGAATGTTGGCTTTTTTAAACAGAAGCAGATAAAAAAAAAA
GATGCAGGACTCCTTCAGTTCTTCACTAGTCTTAGAAAACTTTCCAGAATACTGC
TTCACACTATAAAAAAGAAAAAATATCTTGCATTAGAATCCTTCAACATCTGCATA
CTGCTTCACACTGTTTCGTTTCTAGGAGCACTTTGTACAGGACACTTCTGCTTATA
TTTCTTTAATCAGAACTTAGTTGGATGGGCCGGGCATGGTGGCTCACGCCTGTAAT
25 CCCAGCACTTTGGGAGGCCGAGG-GGGTGGATCACC

Sequence ID 408

CCATCTCCAAATTTAGTATTCATTCTGTTTAGCATATTATCAGTTGCCATCTATTT
GTTTTAACTGATTACTTGAATCTGATTAAACATCACAGAAATGGGCTTTGATAAGA
30 ACAATATTGAATAAGAAATTTTAAATAACAAAACAGCTTATAGAAAAATTCAGCAT
AACTTTTCCATCACCTTCACCACCCTTGCCTTTTATTATCCTGTCCTGTATCACTG
CTTTCTGTTAGCAGTGTTGTGTGAGTTAGGATTTGGGCAGGAAAGCAAAAGCAACC
ACCCGTCATTTTCCCAGAATGAAGGGTTTGACGTAGGATGTAGACTTTGTATAGTA
GTTGGGAGAGCTGTGGGAGTGAAGGTCAGGGATGTCACCTACAGAAGTCAGGGAAT
35 CTGCCACCAGAGATCCTGCATCAGAAACAGCCAACAGCGTGCTTCTGAAGAACTAG
TGGGGAAGTGGCTATAATTCTTAGGAATCCCAGCAAGTCCGCACCACTGTCTCAGT
CTACAGCAGTGGAGAAAGGGGTTTCCAGGAGCTCTCTGGAAAGTTCCTGCCACAC

- 146 -

TTTGCAACAATCTTCAGAGGATAATGGGCTTCTCTTCCAGCTTCCACACCCAACAA
GAGTGCCTTTTCATCGGCCAACTCTAACCTGGAACCCTATGGCAGAGGGGATTTAGG
AGACAGTTTGTNATGTCTGTGGAATGCAAATGAANANGTANCAATGCTTANTTGAC
AGCGGNCATACACAAATNTNGAAA

5

Sequence ID 409

GATCCGTNGACT

Sequence ID 410

10 CTCTTCCCAGCCCCTGAGCCCAGCCCCTTCCCAAGTGGTGCCAGACAAAAAACTAC
ATGGCCCTTTTCGTGTCTTGGGGGTGGAAAGGGAGGGATGAATTGGGGTGATAGAAC
CCTGGTGAATTCAGAGTAATCTTTCTTTAGAAAACCTGGTGTTTTCTAAAGAAACAG
GATAGGAGTTTAGAGAAGGCACCAAAGCTTTCACTTTGGTTTGGCACCAGTTTCTA
ACCATCTGTTTTTTCTACCCTAGCTATCTTTTATTGGTAAAATATAAATGTATAAT
15 TATGTTTGTAGAGCTTTACCAAGGAGTTTCCCTCCTTTTTTGTGTTGTTGATTAGCA
AATTTTTGATTCTCCATTTTCCAAAAGTAAGAGACTCCAGCATGGCCTTCTGTTTG
CCCCGCAGTAAAGTAACTTCCATATAAAATGGTATTTGAAAGTGAGAGTTCATGAC
AACAGACCGTTTTTCCATTTTCATCTGTATTTTATCTCCGTGACTCCACTTGTGGGTT
T

20

Sequence ID - 411

nt: 505

TGGAGCTGAAAAATTCCTATTACCTAGGGGCATCACACGCATTGCATTTTCGCCCCG
TGTTTGGGATGATGCTGGTGTAAACCTACTATGCTGCCAGTCATGTAAAAGTATAG
CACACACAATTAGTAGGTAATGCTTGCAAATAATAATGAAAGACTCTGCTACTGGT
25 TTATGTATTTACTATGCTATACTTTTTTGTCACTTTTAGAGTGTACTCCTACTTT
TTTTTTTTTTTTTTTTTTGAGATGGAGTTTCACTCTTGTCTGTAGGCTGGAGCGAAN
TGGCGCGATCTCGGCTTACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTG
CCTCANCTTCCCAGAGTAGCTGAGATTACAGGCATGCACCGCCACGCACGGGTAAAT
TTTGTATTTTTTGGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCACCAACT
30 CCTGACCTCAGGTGACCCGCCTCCTCACCTCCAGAGTGTTGGGATTACAGGNGTGA
G

Sequence ID 412

ATAAAAATTAGCTGGGGGTGATGGGCCCTGTACCCCAGCTACTCGGGAGGTGAGGT
35 AGGAGAATCACTTGAACCCGGGAGATGGAGGTTGCAGTGAGCCAAGATCGTGCCAC
TGCACTCCAGCCTGTGTGACAGAACAAGACTCTGTCTCAAAAAAAAAATAATAATA
TAATAATAATAAAAAGGAATAACATAGCTAGGAATAAATTTAATCAAAGAGGTGAA

- 147 -

AGACTTATACACTTAAAACCTACAAAAAAAATCACTGAAGGAATTATAGACCCAA
ATAAAAATAAATAAAAAGACATTCTGTGTTTTAGGGAAAGAAGACTTAATATTGTT
AAGATGTCAATACTACCCAAAGTGATCTACAGATTCAACATAATCCCTATCAAAAT
TCCAACAGCCTACTTTGTAGAAATGGAAAAGCCAATTTTCAAATTCAGATGGAATT
5 GCGAGGGGTCTGAATAACAAAACACAATCTTGGGGAAAAAAAACAAAAACAAAG
TCAAAGAACTCACACTTCTCTATTTATAATTTACTACAAAGTTATAGNATCAAAGT
CGACGCGCCGCGATCCGGGC

Sequence ID 413

10 CACAGTACTCCATTTTGGGGTCCAAACTGTAATGCTCAAATAATAAATGCTTACA
CGAAAATTATTTATTGAGAATATTCATATAAAAATTACCTAAAGCAAAGTAAAAAA
AGTAAAATCAAGGTGGTATATTTGAAGTGAATGGTGATTGGAAATTTTGTAGCTGTA
ACAAAAGAAAGAAAACAACCTTTTTTTTAAAGCCTCATTCTCTTTTCTTTCAAATG
TACCTTATTCCCACACACTCTTGGGCTGACCTTTATTTTATCAATAAGCTCAATAT
15 TACTTTGTTTAAAATAAGATGCTTCAGCAAAGTCATTCTCTCTTTAACCATATAA
TTTAAAAACTCCTCTTCACGATTGATAGCAAATCAGAAACGTTAGGGCACCAGTG
AGTTGAAAAAACTGGTCTTAAGTTGGAAAACTATTATTAATAATATTATCCTATC
CATCCATATCTATTGAAATTGTCAGGTCCATAATTTCATTTTAATTAATTATAGGA
AAGAAGAAAAGATAATACCCATTTGTTCTAT

20

Sequence ID 414

CTCAGACTCTTTCTGCCCTAATGGCCATTACTATCCAGTCTGTATTGCTACAAGGG
ACCCACTGGTACCCCTTTTAGATTCTATCAAAGGAACAGGGTTTTCTTAGAGGCA
GGCAGCCTGGTGGTATGGCACAGCAGAAGCTTACTGCTAATGAAATGGGAACCTCC
25 CCCTCCCTTGTGGTTTCAGCACAGAACCTGAATGCCAGGAAAAATTCCTGGGCCAA
GAAGCTAAAGCTAAAGAAACCTTCCTTTTTTTCAACGTTTTTTTTTTCTTTCAAATG
TAGGGTCACTTTTGATTGAGGCAAAGGGGTCTACTGTAAGTGGAAAAGACTCACT
CCCCTAACATAAGTTTTCACTGTGGTGGGATGGTGCCGCCCGATATGCTTGATATG
CTTTTCCTTCCACATGTTAAGCTAGGAAACCTAACAGGATGTCAGCAGGGCAGTTA
30 ACTCTGGACTCANAGCCCTCAAGGGCATGTGGCANAACCTCATGGCATNCAAGACC
A

Sequence ID - 415

nt: 596

GTATAATTGATTCTTTTGAACCTAAAGTATAAGACTTCACGATTAGAAAAAATTA
35 TCCAAAGACTAATGTAATTAAGTGAGGAAAAGGTGCTGGAGGAACTGGATAACCAC
ATGGAAATGTATGAACCATGACCTCTATGTCACATACTATATATAAACTTAATTT
GAGGTGTATCACAGAGCTAACTGTGGGGGCTAAAACGTTGAAGCCTTTGGATGGCC

- 148 -

GCACAAGAGATGTCTGCATTCATAACCTTGGGGAGGGTATGAACATTTCTTGGTAA
CATGCAAAAAGCACTAACTGTAAAAGAGAACAGTTGGTCAGTTGAATTTTCATGAAA
CATTGTAAACTTCTGCTAAACAACTGACACCATTAAGAATGTGGAAAAAGGCTGGG
CACAGTGGCTCATGCCTATAATCCCAGCATTTTGGGAGGCCGGGGCGGGAGAATCA
5 CTTGAGGCCAGGAGTTTGAAACCAGCCTGGGCAACATGGCAAGACCCCGACTCTAC
AAAAATATTTTAAAAATTAGTTGGGTGTGGTGATGCACTCCTGTAGTCCTAGCTG
CCAGGANGCTAAGGNGGAAGGATCACTTAACCCTGG

Sequence ID 416

10 CTGGTGGCGGCGGTTCGTGCGGACGCAAACATGCAGATCTTTGTGAAGACCCTCACT
GGCAAAACCATCACCCCTTGAGGTGAGCCCACTGACACCATTGAGAATGTCAAAGC
CAAAATTCAAGACAAGGAGGGTATCCCACCTGACCAGCAGCGTCTGATATTTGCCG
GCAAACAGCTGGAGGATGGCCGCACTCTCTCAGACTACAACATCCAGAAAGAGTCC
ACCCTGCACCTGGTGTTGCGCCTGCGAGGTGGCATTATTGAGCCTTCTCTCCGCCA
15 GCTTGCCCAGAAATACAACCTGCGACAAGATGATCTGCCGCAAGTGCTATGCTCGCC
TTCACCCTCGTGCTGTCAACTGCCGCAAGAAGAAGTGTGGTCACACCAACAACCTG
CGTCCCAAGAAGAAGGTCAAATAAGGTTGTTCTTTCTTGAAGGGCAGCCTCCTGC
CCAGGCCCCGTGGCCCTGGAGCCTCAATAAAGTGTCCCTTTTCATTGACTGGAGCAG

20 Sequence ID 417

GCAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCCGCAGATAAGTTT
TTTTCTCTTTGAAAGATAGAGATTAATACAACCTACTTAAAAAATATAGTCAATAGG
TTACTAAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAAGATT
TTAAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAAAGG
25 TTTCTAAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAATGTAT
TTAAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAGAAGG
GCAATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTTAAA
AGTTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAGAGATTAAACCGA
AGGTGATTAAAAGACCTTGAAATCCATGACGCAGGGAGAATTGCGTCATTTAAAGC
30 CTAGTTAACGCATTTACTAAACGCAGACCAAATGGAAAGATTAATTGGGAGTGGT
AGGA

Sequence ID 418

CCCGGAATCGCGGCCGCGTCGACGGGAGGTGATAGCATTGCTTTTCGTGTAAATTAT
35 GTAATGCAAAATTTTTTTTAATCTTCGCCTTAATACTTTTTTATTTTGTTTTATTTT
GAATGATGAGCCTTCGTGCCCCCCTTCCCCCTTTTTTGTCCCCCACTTGAGATG
TATGAAGGCTTTTGGTCTCCCTGGGAGTGGGTGGAGGCAGCCAGGGCTTACCTGTA

- 149 -

CACTGACTTGAGACCAGTTGAATAAAAGTGCACACCTTATAAAAAA

Sequence ID 419

5 CCCGGAATCGCGGCCGCGTCGACGGGAGGTGATAGCATTGCTTTCGTGTAAATTAT
GTAATGCAAAATTTTTTTAATCTTCGCCTTAATACTTTTTTATTTTGTTTTATTTT
GAATGATGAGCCTTCGTGCCCCCCTTCCCCCTTTTTTGTCCCCCAACTTGAGATG
TATGAAGGCTTTTGGTCTCCCTGGGAGTGGGTGGAGGCAGCCAGGGCTTACCTGTA
CACTGACTTGAGACCAGTTGAATAAAAGTGCACACCTTATAAAA

10 Sequence ID 420

CTTCATTTGAAATGGTTGAATCTGCTGTGTAATAAAGTGGTTCACCATGATTAGG
AACTGAAATTTAGTAGAAGAGGGAAAAGGAGTTAATGTAACAAATTATTTTAGCTA
CAAACCCCGGTAATAGAGCACTTGGGGGATGGGATGGGGTGGGTGAGACAAT
CAGAATGGTAAATTGATTAAATGCTCCTAACCTGTAATTTTGTGCATAGAGCACC
15 CTATGCTGTGGAAATAACTGTTCTTAGATTTTCATTGTAACCTGGACTGTTTCAAGTTG
CCCAGAGGGAAAGAACATTCCTAATTCTAATAAAATAAACTTTTATTTTGTTTA

Sequence ID 421

TGTCATTGAATCTGCTTGTTACTTAAATGCTAAACTCAATTCTGTAATTCAATAGG
20 TGCACCTCTCTGAGAAACATAAGAGACAATGAGGAAAAGGATTCAGCATTCCGTGG
AATTTGTACCATGATCAGTGTGAATCCCAGTGGCGTAATCCAAGTAAGATGTTTAC
AAAGATTTGTTTTTTAATGTCTAATTAATAAAATTTTAAAGGAAGAAACATTCTAAT
ACTTTAATTATAAAAAGTTAACTATTTTCAAAGGTATCAAATAACAGTTAAACCTT
TAAAATGTATATTTCTTAATATCTTGAAATTGTAATGCCTTTTTTTTTTCTCTAAAT
25 TTTTTTTGTCATGAAATGAGATAGTAACAGCAGATTGGGACAACAAGGTTATATTC
TTGTCTTGAATCAGGCCATGGCTTCTTTCATCCAAATTTTCAAGACCTCATTTATTTA
CTTTGTCCCTGCCTCCCATCCCTGGATATCAGTTTGTGGATATCTACAGTTAATAG
AGTGACCAAATAGTAGGAATACTGTCTCTCTATTCTGAATAAAATCTTTGAATCAG
ATTTAGAAATAATGAATAAAATACAAATCAGCCATTGAAATTGCTCTAATTTTGAG
30 AGCTTATGATTTATTCATCTTTGGTTTCCAAGTTCAAGTTATATGTAGACATTTTA
ATT

Sequence ID 422

GCTTCCTAGGTGAGGTCACGAGGAAACCTGCTGGCCAAGTGACCTGGCAGGGTGTG
35 GCCAGTGTGGCCAGGGCCGCGAGCCTGCTTTCCTTCCCTGCAGCAGGAACCTTC
TGGGGCTGTGATCCTGCGATGGTGCCTGGGTGGGAGTGGGGGTGGGGGGCGGGATG
GTCTCCCTACCTGCCAGCTTCTTGGTTTGAGGTGAGGACAGCCCCGGAAGCTCANA

- 150 -

CTTGGCTCCTGTCCATGTACTTGGGGCCATGAGCTCTGCAGGGACCTTGGAAGAN
AGAGACGGGTGGTGTANGGCANGGGAAGGCATTGTCTTCAAACAGGAAAAAGCTGA
NAATGGAAACAGGCGAAACTTACCAAGTGTAACATCACCTGGAACTGAAGGAGGGT
GGGAAGGTTTTTAATTATTTTAAAAATAGAGATGGGGTCTCACTATGTTGCCCAGGC
5 TGGTCTCAAACACTACTGGGCTCAAGTGAACCTCCTTCT

Sequence ID - 423

nt: 387

TGTTTCTCNAGGGCGAGAGGCTGTCTTANAGCACCATTTCTCTGGCCCTNGTCCCAT
GAGAAGGAACCGCACTCAGGAGCCACACTCTCCCACTNCCCTTGCCCANAAAGACTC
10 ACAGAGGGCACGGAGCTGGCTGTGGTGAGAGGAGGTCCANCAAATTCCTGTCTGCA
NAAGGGTTCTGAACACCACCGCCTGGCAGCGTGCTGGAGGAGGGATTCTCTTTTC
CTCACAGCAATTCTGACCAGAAACCTGTCAAATCAGGAATGGCTAAAATAAGACCA
GGGTATGAATGACCATCAGCCACAGTAAAACCAAGGCACAGCTCTCCTGAGCCCAC
CCAAGCTGCTGTGGCCAGACTGGTGACATCACCTCAGGGCAAAAAAAAAA

15

Sequence ID - 424

nt: 420

CGCAGAATGGCTCCCGCAAAGAAGGGTGGCGAGAAGAAAAGGGCCGTTCTGCCAT
CAACGAAGTGGTAACCCGAGAATACACCATCAACATTCACAAGCGCATCCATGGAG
TGGGCTTCAAGAAGCGTGACCTCGGGCACTCAAAGAGATTCGGAAATTTGCCATG
20 AAGGAGATGGGAACTCCAGATGTGCGCATTGACACCAGGCTCAACAAAGCTGTCTG
GGCCAAAGGAATAAGGAATGTGCCATACCGAATCCGTGTGCGGCTGTCCAGAAAC
GTAATGAGGATGAAGATTCACCAAATAAGCTATATACTTTGGTTACCTATGTACCT
GTTACCACTTTCAAAAATCTACAGACAGTCAATGTGGATGAGAACTAATCGCTGAT
CGTCAGATCAAATAAAGTTATAAAATTG

25

Sequence ID 425

GGAAACTGATGCCAGTCAGAAACTCAGATCAAATGAAGGGGTGAAGAGAACCAGAA
TTGATCTCTCTGTAGGAGAATATAAATGACTTTTTTTAAAGTACATATTTTCTGTGA
AAGACAGTTTTTTGTTTAATGCAAAAATGTTAACAATGTTTATATCATGTAGAAGT
30 AAAAGATCGTGAAACAGCACAGAGAACAGTAGTAAGACAGATTGAATTGCACTGTT
GTAAGATGATGAACTTACAATATTAAGTGAAGGTAGACTGTGATAGATTAAGGATA
TATATTGTAATCCCTAGAGCAATTGTCAAAGTGGTACAGGTAAAAAGCCAATAGAG
GTGATAAAATGGAATACTAAAAAATATCAGATGAATAATAAGAAGACAGGAAATG
AGGAACAGTGGAACAGAATGAATAAAAAACAAGACCATTAACCTTAATCATTAATAA
35 TTACTTTAAATGGGTAAACATTATGGTTATAAGGCAGAGATTTTCAGACTAGATA
AAAGAGCAAGCTCCACTATATACTGTCTACAAGAGATATACTTTAAAGTGTATATT
ATATTTAAATATAAAGATTTGGAATAAATAAACCTAAGAATAAGCTTACTAGGGAA

- 151 -

GTGAAAGATCTGTACAACAAGAATTACAAAACACTGCTGAACGAAATCATAGGTGA
CCA

Sequence ID 426

5 GTCCCGGAATCGCGGCCGCGTCGACGTTTCCTCAAAATTTATCTTCCTGTTAATGT
CAGGCATGTATCTCCTTAGCTTGCCACAAATAACTATATATACCACAGACCTTCCT
TTGTAGGGCTAACAGTGTTGCATTGTAAGTGGAGGCCTCATAGATACCTGGCCTTT
TCCTACCTTATTCCAAAGATGGTTGCATCTTATAAATAATGTCATTCTTCAGCAAA
TGGTATGGAAATGAGATTGTAATGTCATTATTTTCCTCTTTAAATAATCAGGACAAC
10 TCATGATACAAAGAGCTCTTCTCTATAAAAGGTGGGACTTTTTTTTTTTAGTAATAG
CAAAAATAAAATTGTACCTCCTTAATCTTCTACAGAAAGATGGATTTTCAATTTCAA
CATTAAGAGGTAGTTTTTAAGAAGCAGTAGAAGTCAGCCTGGGCAGCATGGTGAAAC
CCCGTCTCTACAAAAAAGTTAGCTGGGCTTAGTAGTTGCAATCCCAGCTACTCTGG
AGGCTGAGGTTGGAGATCATCTGANCCTGGGGAGGTCNAGGCTGCAATGATACANT
15 GAGCCCTGATTGTGCCACTCCACCTGGTTGCAGA

Sequence ID 427

TTCCAATCTTCGTGTTCACTTTAAGAACACTCGTGAAACTGCTCAGGCCATCAAGG
GTATGCATATACGAAAAGCCACGAAGTATCTGAAAGATGTCACCTTACAGAAACAG
20 TGTGTACCATTCCGACGTTACAATGGTGGAGTTGGCAGGTGTGCGCAGGCCAAGCA
ATGGGGCTGGACACAAGGTCGGTGGCCCCAAAAGAGTGCTGAATTTTTTGCTGCACA
TGCTTAAAAACGCAGAGAGTAATGCTGAACTTAAGGGTTTAGATGTAGATTCTCTG
GTCATTGAGCATATCCAAGTGAACAAAGCACCTAAGATGCGCCGCCGGACCTACAG
AGCTCATGGTCGGATTAACCCATACATGAGCTCTCCCTGCCACATTGAGATGATCC
25 TTACGGAAAAGGAACAGATTGTTTCCTAAACCAGAAGAGGAGGTTGCCCAGAAGAAA
AAGATATCCCAGAAGAACTGAAGAAACAAAACTTATGGCACGGGAGTAAATTCA
GCATTAAAATAAATGTAATTAAAAGG

30 Sequence ID 428

TGCAGGATCCGTGCACTCTAGATAACATGGCTAGAAAAGAGAATGAAAAAGTTGGA
ATTTTTAATTGCCATGGTATGGGGGGTAATCAGGTTTTCTCTTATACTGCCAACAA
AGAAATTAGAACAGATGACCTTTGCTTGATGTTTCCAACTTAATGGCCCAGTTA
CAATGCTCAAATGCCACCACCTAAAAGGCAACCAACTCTGGGAGTATGACCCAGTG
35 AAATTAACCCTGCAGCATGTGAACAGTAATCAGTGCCTGGATAAAGCCACAGAAGA
GGATAGCCAGGTGCCCAGCATTAGAGACTGCAATGGAAGTCGGTCCCAGCAGTGGC
TTCTTCGAAACGTCACCCTGCCAGAAATATTCTGAGACCAAATTT

- 152 -

Sequence ID - 429

nt: 535

5 CACAGTACTCCATTTTGGGGTCCAAACTGTAATGCTCAAAATAATAAATGCTTACA
CGAAAATTATTTATTGAGAATATTCATATAAAAATTACCTAAAGCAAAGTAAAAAA
AGTAAAATCAAGGTGGTATATTTGAAGTGAATGGTGATTGGAAATTTTGTAGCTGTA
10 ACAAAAAGAAAGAAAACAACCTTTTTTTAAAGCCTCATTCTCTTTTCTTTCAAATG
TACCTTATTCCCACACACTCTTGGGCTGACCTTTATTTTATCAATAAGCTCAATAT
TACTTTGTTTAAAATAAGATGCTTCAGCAAAAGTCATTCTCTCTTTAACCATATAA
TTTAAAAACTCCTCTTCACGATTGATAGCAAAATCAGAAACGTTAGGGCACCAGTG
AGTTGAAAAAACTGGTCTTAAGTTGGAAAACTATTATTAATAATATTATCCTATC
15 CATCCATATCTATTGAAATTGTCAGGTCCATAATTTCAATTTAATTAATTATAGGA
AAGAAGAAAAGATAATACCCATTTGTTCTAT

Sequence ID 430

15 CAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAAGTTTTT
TTCTCTTTGAAAGATAGAGATTAATACAACCTACTTAAAAAATATAGTCAATAGGTT
ACTAAGATATTGCTTAGCGTTAAGTTTTTAAACGTAATTTTAATAGCTTAAGATTTT
AAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAAAGGTT
TCTAAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAATGTATTT
AAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAGAAGGGC
20 AATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTTAAAAG
TTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAAGAGATTAAACCGAAG
GTGATTAAAAGACCTTGAAATCCATGACGCAGGGAGAATTGCGTCATTTAAAGCCT
AGTTAACGCATTTACTAAACGCAGACGAAAATGGAAAGATTAATTGGGAGTGGTAG
GATGAAACAATTTGGAGAAGATAGAAGTTTGAAGTGGAAAACTGGAAGACAGAAGT
25 ACC

Sequence ID 431

30 CGCTGGGTGCCTGCAGCGCCTCCCTTGTCTCATATGGTGTGTCCAGCACTCTATTG
TTGTAAACTGTTGNNTTGNCTGACCTAAATTNTCTTTACTAAACANATTTAATAGT
TNAAAAAANANCA

Sequence ID 432

35 TTTTAAAGTCATCTCTATAGGAAGGTGCTGGGCAGGGATCCCAGAGAAAGAAAGGG
TCCAAGACTCCATTAACTGCCCTGGATGAAGGGCACTGCTACAGCAGCTAGTACCA
GAGACTCTCCTATCTCACGGTTGAGGCAGACCCAGGATAGAATAGAGAATAAAAGG
AATGCTTATAGGAAACAATTTTGTATGGAATGCTAGATGGCCAAGCCTCAGCCTTT
GGTCCAGTGCAACCCTTGCCTCGCTTGTCAACAGTGAAAAATTAGTTTGGTTAGAA

- 153 -

GAACCATCTGGAAACACACCAGCTTCTGCTACCTTCATGCTCATTGTTAAAAAAG
ATTAACCAGTGTGAACATTCTGATCTGTTAATTCCAGGGACTGTTTTCTTTCCAAT
GGACTGTTTGTGTTGGTAGAATAACCCCCAAAAGCTCAAAGCTAAAATGCATCATCAG
TCCTAGTCGGCAGTTCCTTAAGAATGGACTGGCGGCGTGGTTGAGCTGATATGGAA
5 AAGCTGCACCTTCCTGCAGAAGATCAACTGACCTGCTATCCCACCCCCAAATTTCAA
CCTGAGGTATATTTCAATGAAGGCAGGTAGCTGTGCTTCTCAGAGCA

Sequence ID 433

TCCCGGAATCGCGGCCGCGTCGACCCGCCGCCGAGGATTCAGCAGCCTCCCCCTTG
10 AGCCCCCTCGCTTCCCGACGTTCCGTTCCCCCCTGCCCGCCTTCTCCCGCCACCGC
CGCCGCCGCCTTCCGCAGGCCGTTTCCACCGAGGAAAAGGAATCGTATCGTATGTC
CGCTATCCAGAACCTCCACTCTTTTCGACCCCTTTGCTGATGCAAGTAAGGGTGATG
ACCTGCTTCCTGCTGGCACTGAGGATTATATCCATATAAGAATTCAACAGAGAAAC
GGCAGGAAGACCCTTACTACTGTCCAAGGGATCGCTGATGATTACGATAAAAAGAA
15 ACTAGTGAAGGCGTTTAAGAAAAAGTTTGCCTGCAATGGTACTGTAATTGAGCATC
CGGAATATGGAGAAGTAATTCAGCTACAGGGTGACCAACGCAAGAACATATGCCAG
TTCCTCGTAGAGATTGGACTGGCTAAGGACGATCAGCTGAAGGTTTCATGGGTTTTA
AGTGCTTGTGGCTCACTGAAGCTTAAGTGAGGATTTCTTGCAATGAGTAGAATTT
CCCTTCCTCCCTTGTACAGGTTTAAAAACCTCACAGCTTGTATAATGTAACCATT
20 TGGGGTCCGCTTTTAAGTTGGACTAGTGTAACNTTCATGCAATAAACTGAAAAG
ACCATGCTGCTANTC

Sequence ID 434

TTCGGACGCAAGAAGACAGCGACAGCTGTGGCGCACTGCAAACGCGGCAATGGTCT
25 CATCAAGGTGAACGGGCGGCCCTGGAGATGATTGAGCCGCGCACGCTACAGTACA
AGCTGCTGGAGCCAGTTCTGCTTCTCGGCAAGGAGCGATTTGCTGGTGTAGACATC
CGTGTCCGTGTAAAGGGTGGTGGTCACGTGGCCCANATTTATGCTATCCGTCAGTC
CATCTCCAAAGCCCTGGTGGCCTATTACCANAAATATGTGGATGAGGCTTCCAAGA
AGGAGATCAAAGACATCCTCATCCAGTATGACCGGACCCTGCTGGTAGCTGACCCT
30 CGTCGCTGCGAGTCCAAAAAGTTTGGAGGCCCTGGTGCCCGCGCTCGCTACCAGAA
ATCCTACCGATAAGCCCATCGTGACTCAAACTCACTTGTATAATAAACAGTTTTT
GAGGGATTTTAAAA

Sequence ID 435

CTGCAATGTGCAATAGTTGCACCACTGCACTCCAGCCTGGGTGACAGAGTGAGAAC
35 CTATCTCTTAAAAAAGGGAAGAAGAGACATGAGAGGGCCCAAGT
CACTTGCTCACTCACTTTCCGTGTACATGTACCAAGAAAAGGCCATGTGGGAAAGA

- 154 -

GCAAGAAGGCAGCCGCCTTCAAGACAGGAAGAGAGCCCTCACCAGAACTGAGCCA
GAACCTTGGAATTCCAGCCTCCANAACCTGTGAGAAAAGAATTTTCTGTTGTTTCAG
TCCCCCACACTATGGCATTGTTACGGCAGCCTGAGCTAATACTCCTACTTTGTC
CTGCATTTACTTGGTCTTCCAGTTAGTTTTTTAGACTTTGGGAATCAGAGCAGTCA
5 GTTGTGTCAGATTTTAGCTTACAGTTGTCCTACCTGTGCAACTGAAATTTCTTCCATT
TTAAACCAGAGCAGAGTTTTAGAGTCAAAAGAAACCAGATCTTTTAGTGCAGAAGC
TTTCCACTGTATTANAAGTGAGGAAGTTGGT

Sequence ID 436

10 AAAAAAACTCCAGAGAAGTTTATAGAAAGAGATGACATGTAAACCCTGCTGAAAAA
TAGTTTTCATTTGTTAGAATATAATTGTCTTCCACTAAAAAAAGAAAAAAAAGCA
TTTAAGGCTCTAAGATCTCTTGAAGTACCACTTTTCCTGAATCCCAGAGTTTTTAT
GTGCATTATTTTTATGCGTTTGTAGTTTGATATGTTGTATTTATAAGTAGTTTTAG
CTTTCCATTATGAATTCTTCTTTGACCCATGAGTTATTTAGGTAAAGTGTTTAAAAA
15 TTTACAATAGTTTATATATGCAAATATTATGTTGTTAGAGTTGGTTTTTCATGTCAT
TTTTACATATACAGGGGCAGTTTCCCCAACTAAATTGTATATTCCTTAAAGCAGCA
CTCTTAAATTTTATTTCTGTGTCAATTTCTTGNCTGTGTTTCCTGGCATGGAATAC
ATGGCATAAAATTTGTTATGTAATTAAATGAAATATTATTATACTTTCTATTTTTT
AGAAAAAA

20

Sequence ID - 438

nt: 577

GTCGACAGGGATGACATAACTATTAGTGGCAGGTAGTTGTTGGTCACTTTCAACT
CTGGGTTCAAGCGATTCTCCTACCTCAGCCTCCCGAGTAGCTGGGATTACAGGCAT
25 GCACCGCCACACCTAATTTTCTATTCTTAGTAGAGACGGGGTTTCTCCCTGTTGGT
CAGGCTGGTCTCGAACTCCCGACCTCAGGTGATCTGCCTGCCTCAGTCTCCCAAAG
TCCTGGAACCACAGACATGAGCCACCACGCCTGGCCCCTTTTAAAATATTTCTGCT
CATTGATGATGCACCCAGTCACCCAAGTGCTCTGATGGAGATGTATAAGGAGATGA
ATGCTGTTTTTCATGGCTGCTAATAACAATTCTGCAACCCCCCAAATCAAGAA
30 GTAATTTTGACTTTCAAGTCTTATTATTTAAGAAATATATTTTGCAAGACTATAGC
TGCCATAGACCGTGATTCTCTGATGGATCAGACAACTAAAATGAAAACCTCCTG
CAACGTATTCATCATTCTAGATCCCTGAGGAATCGCCACACTGACTTNCACAATGG
GTGAACTGGGTACAGT

35 Sequence ID - 441

nt: 552

AAACAAAATTATTCTCTGAGAGGGAAAGGACATTTGAGGGAAACATCAAATTTCCC
CATAAATAAATGAATGGAGTTTGCAGGAAGGTGAGGGTGAGCAGAGATGTGTGTGG

ACATCTCTGACCATCCATCGCTGTATTCAAATGGATTGTTTTATTCCATTCTGGTC
TCAGGCATGACCACGTCCAGTGAAGACATTTGAGGCAGCACATCTCAGGACCCAGG
CAATAGACTGGCCCCAACTCAGGCTGGACTAAGGTGTGATTAATTCTTTGTTTTTT
GTGTGGAACAGCTCACCTTGTCAGACAGCCTCAGGGCATCTCTGAGACACAGGGGC
AGAAAATGACATTCATCTTTTGAGTCCTCATCCATGGAGTGCTGTGTTTGGGGGGC
TGCATCTGCTGAAGCGAGAACCCCATTTCTGCCACCCCACCAGGATGCCCATTCCTCC
AGGACTTCTCCAACCTTACTATTAGACTAAACCAGAACAAAGCAACAACTGTATTTA
TGCAAGCAAAATTGATGAGAAAATTATATTCAAATAAAGCAAAAATTA

CGGACTCCTGTGCTAATTGTCAGCTTACATATCATTGTATAGAGACTGTTTATTCT
25 GTACCAAAGTACTACATNGAAAATAAACCGGTGACTGTTTTTCT
TCATAAAGTTCTGCGTTTGGCATCTTCACTCTTTCCAAAATGTATCTGTACATCAN
AAATGTCACTATTCCAAGTGTCTTTTTTAGTGTGGCTTTAGTATGGCTTCCTTTTAA
TATTGNACATACATTGNATCTTTGTTTTATGGNAATAAGTAATAAAAATGTAGACT
TCATATTTTGTACAAAATGTCCTATGTACAGAATAAAAAAGTTCATAGAAACAGCC
30 NANAA

AGGCCGAGGCAGGCAGATCNCNTGAGGTCAAGAGTTTGAGACCAGCNTAGCTAACA
TGGTGAAACCCCATCTCTACAAAATATA-AAAATTAGCCTGG-GTGGTGATGGGC
35 ACCTGTAACCCAGCTACTCGGGAGGCTGAGGTAGGAGAATCACTTGAACCCGGGA
GATGGAGGTTGCAGTGAGCCAAGATCGTGCCACTGCACTCCAGCCTGTGTGACAGA
ACAAGACTCTGTCTCAAAAAAAAAATAATAATAATAATAATAATAAAGGAATAAC

- 156 -

ATAGCTAGGAATAAATTTAATCAAAGAGGTGAAAGACTTATACACTTAAAACTACA
AAAAAAAATCACTGAAGGAATTATAGACCCAAATAAAAATAAATAAAAAGACATT
CTGTGTTTTAGGGAAAGAAGACTTAATATTGTTAAGATGTCAATACTACCCAAAGT
GATCTACAGATTCAACATAATCCC'TATCAAATTCCAACAGCCTACTTTGTAGAAA
5 TGGAAAAGCCAATTTTCAAATTCAGATGGAATTGCGAGGGGTNTGAATAACAAAA
CACNATCTTGGGGAAAAAAAACAAAAACAAAGTCAAAGAACTCACACTTCTNTAT
TTATAAATTTACTACAAAGTTATAGTAATCNAA

Sequence ID - 448 nt: 329
10 TACGCACACGAGAACATGCCTCTCGCAAAGGATCTCCTTCATCCCTCTCCAGAAGA
GGAGAAGAGGAAACACAAGAAGAAACGCCTGGTGCAGAGCCCCAATTCCTACTTCA
TGGATGTGAAATGCCCAGGATGCTATAAAATCACCACGGTCTTTAGCCATGCACAA
ACGGTAGTTTTTGTGTGTTGGCTGCTCCACTGTCCTCTGCCAGCCTACAGGAGGAAA
AGCAAGGCTTACAGAAGGATGTTCTTCAGGAGGAAGCAGCACTAAAAGCACTCTG
15 AGTCAAGATGAGTGGGAAACCATCTCAATAAACACATTTTGGGGTTAAA

Sequence ID 450
GAGCAGTGGCATGATCACACCTTACTGCGGCCTCCAACCCCTGAGCTTAAGTGATT
CTCCCGCATTATCCTCCTGAGTAGCTGAGACTACAGGTGCATGCCACCATACTA
20 CTAAATTTGGGTCGGGTGGTGGTGGTGATTTTTTAATATTTTTGTAGAGACAGGGT
CTCACTGTGATGCCCAGGCTGGTCTTGAACCTCCTGGGCTCAAGCAGTCACCCACCT
CAGCCTCCCAAAGCACTGGGATTACAGGTGTGAGCCACCACACTGGCCAGCTTTGT
TTTGTTTTGATGACTAAGCTGCTCTTGCTAAAAGGGCTTCTCTCTGAACTTCCCTA
CCTTTCTTCTGTTTCCCTGGGCTAGGGCTCCATGTTGGCAGTCCTACTCCCAATTA
25 ACCTGGGGCTGTCTGGTTAACCTTTATAAGATCTGCAGTCATTGGGAGACCCGGGG
ACCAGGAATATTGTTGTTGAGGGAGCTACCCTGGAAAGTGGATGGGTGGCCAAAGG

Sequence ID 452
TTTGGCTTTGCCTCTAGGCATTAGATGTTATCTTTGGAGGCATCCTTCTATGAGCA
30 TTCATTTTTGGACCAAGCCTGGATTTACAATTCTATTACTGGCCCAGACTTCATTT
CTATCCAATTTCAATCCACTGTGCTATAGTTTACAACATATAATTTGACTTATAAA
TAATTCCTGACTATGGGTTTAAAGACTGAAAATGGATCAATAGAAACTTTGAAAAT
GTTAACATCTTGATTGCTTTTCTCAGTGTAGAAATGGACAATGTTTAGCTTAAAAA
CTGCATGTTTTTAATGAGATACGGGGTTGAAAGACTTATTCCTGGAATTTATTGTT
35 CTGGAGAAAGCCTGTTGCTATCTGCCATACCTTGGTTTACTTTGTGCAAAATGAGC
TTCTTTTTAAGTAATGAGCTCTTTCCATGTTTCACTTAAATTGCTGTCTTAGACAC
TTCATCAGGGTTCCCTGCTCTGCCTCATTCCCCCTTTTGCTCACTTGCAGCCTTTG

- 157 -

ACATAATCCTGGGAGGCAATTGGCATCATACATATTTTGCTTTGTAATCTCCTGCT
TTGATTCTGACTGGGACCCAGC

Sequence ID - 453

nt: 747

5 GGATCTAAGACCAGCCTGGCAGCCACCAGATGGTGATTCTAGTCCTGGCTCAGTCA
GTAATAGGTCACCTGACCCCAGAGAAATCAATTCAGCCTCCCCAGGTCCTTGGATTT
CTTTCTGTGAAAATGAAAGCATAGGTAGGAATTTCCCATGGAACAGCTAGCAGAGG
AGAAATATTAAAAGTCAGGAGACTCATGCTATAGTTTTTCATACTTCATTACAACAA
TGTTGTTTAGGACAAGTGAGTTAACCTGTTAGCTTCCTCTATATAAAATGGAAAGT
10 CATTAAAAACCTACATAGCAGGGTTCTTGTGAAGATCAAGTGATAATGTAGGAAGC
ATGTACAAATGTCACATTCCTGCCGTCACGTAATGGTCCTCACAGCTTGAGGTAGCA
TTTAGCATGTGTCATGATTTAGTACAAGGGTTGGCAAACCTGTTGCTCTTGGATTAA
GTCTGGCTCATTGCCTGTTTTTCAAAGAAAAAAATTGTATATGTGTGTATATATGT
TATATATAGGTACACACACATATGTGCTATATATAGCATATATACACACATAATAT
15 ATAAACATGTACATATATAGCATTATATATATACCGTGTATAATATCTCCAGTCCT
CATGACCAGCCATGCTTGTTTACATTTGCATACTCTATGATTGCTTTCATGC
AACAATGGCAGAGTTGAGTGATTGTTTTGCACAGANACTGTATGGCCCACTAAACC
TAAAATATTAATCTCTGCC

20 Sequence ID 454

CTCCTGCCGGGCTCGTGGCGGCTTCTGTCCGCTCCGCGGAGGGAAGCGCCTTCCCC
ACAGGACATCAATGCAAGCTTGAATAAGAAAAACAAATTCTTCCTCCTAAGCCATG
GCATATCAGTTATACAGAAATACTACTTTGGGAAACAGTCTTCAGGAGAGCCTAGA
TGAGCTCATAACAGTCTCAACAGATCACCCCCCAACTTGCCCTTCAAGTTCTACTTC
25 AGTTTGATAAGGCTATAAATGCAGCACTGGCTCAGAGGGTCAGGAACAGAGTCAAT
TTCAGGGGCTCTCTAAATACGTACAGATTCTGCGATAATGTGTGGACTTTTGTACT
GAATGATGTTGAATTCAGAGAGGTGACAGAACTTATTAAAGTGGATAAAGTGAAAA
TTGTAGCCTGTGATGGTAAAAATACTGGCTCCAATACTACAGAATGAATAGAAAAA
ATATGACTTTTTTTACACCATCTTCTGTTATTTCATTGCTTTTGAAGAGAAGCATAGA
30 AGAGACTTTTTTATTTATT

Sequence ID - 458

nt: 682

TGCCACTGAAGATCCTGGTGTCGCCATGGGCCGCCGCCCGCCCGTTGTTACCGGT
ATTGTAAGAACAAGCCGTACCCAAAGTCTCGCTTCTGCCGAGGTGTCCTTGATGCC
35 AAGATTCGCATTTTTTGACCTGGGGCGGAAAAAGGCAAAAGTGGATGAGTTTCCGCT
TTGTGGCCACATGGTGTCAGATGAATATGAGCAGCTGTCCTCTGAAGCCCTGGAGG
CTGCCCCGAATTTGTGCCAATAAGTACATGGTAAAAAGTTGTGGCAAAGATGGCTTC

- 158 -

CATATCCGGGTGCGGCTCCACCCCTTCCACGTCATCCGCATCAACAAGATGTTGTC
CTGTGCTGGGGCTGACAGGCTCCAAACAGGCATGCGAGGTGCCTTTGGAAAGCCCC
AGGGCACTGTGGCCAGGGTTCACATTGGCCAAGTTATCATGTCCATCCGCACCAAG
CTGCAGAACAAGGAGCATGTGATTGAGGCCCTGCGCAGGGCCAAGTTCAAGTTTCT
5 GGCCGCAGAAGATCCACATCTCAAAGAAGTGGGGCTTCACCAAGTTCAATGCTGAT
GAATTTGAAGACATGGTGGCTGAAAAGCGGCTCATCCCANATGGCTGTGGGGTCAA
GTACATCCCCAATCGTGGCCCTCTGGACAAGTGGCGGCCCTGCACTCATGAAGGCT
TTCAATGTGC

10 Sequence ID 459

TCCCGGAATCGCGGCCGCGTCGACCTTGTCCTTGAGCGTCAACCTTCTTTCCCTGA
AGTGGCTGGGGTTCCTGTTTCCTTCTTTGATTGACAACTTGTGTTAACCCTCGCAC
ATCTCTGGGCCAATTTTTTGCTTGTAAGTCTTTCCGGAGACCCCTGGAATTTAAATC
ATTAGCACCGCGCCCTTCCCCGAAGAGTCTTCGAAGGGTTGCCGCTTTTCGGTGGC
15 GCAGTTCTCGCGAGAAGGTGACTTTCTTTCTCGGTATTTCTGTTTCCAGAATCC
TTAGCGCGAGGCGGAAAAAATATTTCTCCCAGCTTGTTGTTGATGCCGCGATTTTGA
CTGAGACTTCTTCCCACGATTTCTGTTTTTTGCTTCTCCAAGGAAAATGGCAGCTCC
CGAGCAGCCGCTTGCGATATCAAGGGGATGCACGAGCTCCTCCTCGCTTTCCCCGC
CTCGGGGCGACCGAACCCTTCTGGTCAGGCACCTGCCGGCTGAGCTTACTGCTGAG
20 GAGAAAGAGGACTTGCTGAAGTACTTCGGGGCTCAGTCTGTGCGGGTCCTGTCAGA
TAAGGGGCGACTGAAACATACAGCTTTTGCCACATTCCCTAATGAAAAAGCAGCTN
TAAAGGCATTGACAAACTNCATCAACTGAAACTTTTAGTCATACTTTAATCG

Sequence ID - 460

nt: 536

25 CAGAGATCAAAATAGGCCTTACACAGTGCGACGCGAATTTAAAAGATTACCCATT
CAGGTGTATGGATTTTGCAGTATTAAAGATGCTGCCTGGAATAGGTCATTATCTTC
TCCAAGTACTCTGTTAAGTCAATGAGTCACATAGAGTATAAGGTTTATTATCTGCT
TTTCTTTCATTAAATAAATCTTTATTGAATTTCTACTACATTAAAAAACCAACCA
AAACAAAACAAACAAAAAAAACACTTCCCTGAGCCATAAAGGAGAAGGTAGTTTTG
30 ACTGGAACCTTGAAGGATGGGTAACTTTTCAGCAGATAAAGATTGAGAGAAGACCT
TCCAGGTAGAGAAAGCAGTGTGGGCACAGGCAAAGATGGAAGAACACACGTGGCTG
TGGGAAACACAGCTAGAAGCCAGTGCGGATAGAGAGTAGGCTATGATGTGCAAAGG
TTANACACTGGGAGAGACAGGTCCATGAGAGTAGCTTGGACTAACACAGGGAGGGT
TTGGAATCCCAACTGGGGAACCTANAAATCAA

35

Sequence ID 461

TAGGAGGCTTATTCAGTATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCT

- 159 -

ACGCCAAAATCCATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCA
CAACACTTTCTCGGCCTATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGC
ATACACCACATGAAACATCCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAG
TAATATTAATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAAGCGAAAAGTCCTA
5 ATAGTAGAAGAACCCTCCATAAACCTGGAGTGACTATATGGATGCCCCCCACCCTA
CCACACATTCGAAGAACCCGTATACATAAAAT

Sequence ID 462

10 TCTTTATCAAGTTGAGAAAGTTCCTCCCCTCTATTCCTAGTTTGCTAAGAGTCCTT
CTATCCTATTTCTTAATGGTTTAGTAGATGACTCTGTGGTACTTTGAAGGTTGTTT
GCAGAATTTCCATGCCATAGGCAATTTACCTTTCTTGACATTTGAAGGATTGATG
TTGGTGCCAAGTATAGAATCTTCACAGAGTCCTCCTGTAGCTTCTAAAGGTTTAGC
TTGAAAATGTTAATTGCTTAACGCTAGTAAGTGAGTGAAAAGCTGGGGATAAATT
TTGTATCTTGCTTATATTTTCAGTTCCACCTCTGTCCNGACNAAACCCCATATAT
15 AA

Sequence ID 463

20 TAGTTTACATATCCCAACCTTTAAAAATATTCCTCTTATTAGCTTTATATTCACTT
TATAGAAGTTGAGTTTTAATTAAATTCCTGGCATCCTGAAGTATGTCACATAGCA
TGTGCTCCTTATAAATATGTTGATATCTCAGAAGACAGCATCCCGGTTTTTCATTTT
ATAAAGTACCATACTTAAGAATGCTGTAATACTTATCTTTTATAACATGTTTCCTT
CGCTTTGCTTGNCTTTTATGNCATCAGTTTTAACTGTTTACTTCATTTAACAGNTT
ACATCATNCAACAGTTTACTTCATTAAACAGTAGGTGGAAAAATAGATGCCAGTCT
ATGAAAATCTTCCCATCTATATCAAATACTTTCAAGGATATACTTT

25

Sequence ID - 464

nt: 615

CGACTTTCAACCATCAAGTGAGGAATACCTTCACATAACTGAGCCTCCCTCTTTAT
CTCCTGACACAAAATTAGAACCCTTCAGAAGATGATGGTAAACCTGAGTTATTAGAA
GAAATGGAAGCTTCTCCACAGAACTTATTGCTGTGGAAGGAAGTGAAGATTCTCCA
30 AGATTTCCAAAACAAAACCTATGGTCAAGTTTCTGGAGAAGCAATCAAGATGTTTC
CCACCATTAACACCTGAGGCTGGAAGTGTATTACAAGTCCCGATGAAATTGAA
TTAGAAGGTGCTACACAGTGGCCACACTCTACTTCTGCTTCTGCCACCTATGGGGT
CGAGGCAGGTGTGGTGCCTTGGCTAAGTCCACAGACTTCTGAGAGGCCACGCTTT
CTTCTTCTCCAGAAATAAACCTGAACTCAAGCAGCTTTAATCAGAGGGCAGGAT
35 TCCACGATAGCAGCATCAGAACAGCAAGTGGCAGCGAGAATTCTTGATTCCAATGA
TCAGGCAACAGTAAACCCTGTGGAATTTAATACTGAGGGTGCAACACCCCATTTTC
CCTTCTGGAGACTTCTAATGAAACANATTCCTGATTGGCATTAATGAANAGTCA

- 160 -

Sequence ID 469

GATTTTAAAAATACATATAGCAAAAATATTACAGGGTCAGGGGAGACAATTAGAA
TGATATAATTCAAAGTGGATTAAAAAAAACCTGTCACCCAGAATACAATACCCAG
CAAAGTTGTCCTTCATAAATGAAAGAAAAATNAAATCTTTNCCNAACNA

5

Sequence ID 471

TCCCGGGAATCTGCAGGATCCGTCGACT

Sequence ID 472

10 GACAGTGCCAGGGCTCTGATATGTCTNTCACANCTTGNAAGTGTGAGACAGCTG
CCTTGTGTGGGACTGAAAGGCAAGATTTGTTCTGCCCCTCCCTTTGTGACTTGAA
GAACCCTGACTTTGTTTCTGCAAAGGCACCTGCATGTGTCTGTGTTCTTGTAGGCA
TAATGTGAGGAGGTGGGGANACCACCCACCCCATGTCCACCATGACCCTCTTNC
CACNCTNACCTGTGCTCCCTCCCCAATCATNTTT

15

Sequence ID - 473

nt: 694

TGGGCTTTGGGCTGGCTGCAGTCTGTCTGAGGGCGGCCGAAGTGGCTGGCTCATTT
AAGATGAGGCTTCTGCTGCTTCTCCTAGNGGCGGCGTCTGCGATGGTCCGGAGCGA
GGCCTCGGCCAATCTGGGCGGCGTGCCAGCAAGAGATTAAAGATGCAGTACGCCA
20 CGGGGCGCTGCTCAAGTTCCAGATTTGTGTTTCTGAGGTTATAGGCGGGTGTTT
GAGGAGTACATGCGGGTTATTAGCCAGCGGTACCCAGACATCCGCATTGAAGGAGA
GAATTACCTCCCTCAACCAATATATAGACACATAGCATCTTTCCTGTCAGTCTTCA
AACTAGTATTAATAGGCTTAATAATTGTTGGCAAGGATCCTTTTGCTTTCTTTGGC
ATGCAAGCTCCTAGCATCTGGCAGTGGGGCCAAGAAAATAAGGTTTATGCATGTAT
25 GATGGTTTTCTTCTTGAGCAACATGATTGAGAACCAGTGTATGTCAACAGGTGCAT
TTGAGATAACTTTAAATGATGTACCTGTGTGGTCTAAGCTGGAATCTGGTCACCTT
CCATCCATGCAACAACCTTGTTCAAATTCTTGACAATGAAATGAAACTCAATGTGCA
TATGGGATTCAATCCCCACCATCGATCATAGCACCCCCTATCAGCACTGNAAACTC
TTTTGCATTAAGGGATCATTGC

30

Sequence ID 474

GGCAGCGCGGGGAGCCCGTCGGCGCCGGCGGGCGGGCCGGTTTCGAAGTTGATGCA
ATCGGTTTAAACATGGCTGAACGCGTGTGTACACGGGACTGACGCAACCCACGTGT
AACTGTCAGCCGGGCCCTGAGTAATCGCTTAAAGATGTTCTACGGGCTTGTTGCT
35 GTTGATGTTTTGTTTTGTTTTGTTTTTGGTCTTTTTTTGTATTATAAAAAATAAT
CTATTTCTATGAGAAAAGAGGCGTCTGTATATTTTGGGAATCTTTTCCGTTTCAAG
CATTAAGAACAACCTTTTAATAAACTTTTTTTTTGATAATGGTTAAAAAAA

- 161 -

A

Sequence ID 475

5 CATAATAAAAAACAATCAACAAACAGGGAATGGAAAGAACTTCCTCAGCATGGTG
AAGGCCACATATGAAAATCCACAGCTAACATCATACTCAATGATGAAAGACTGAA
AGCTTTTCTCCTGAGATCAGGAACAAGACAAAGATGTCACCTTTTGTCACCTTCTAT
TCAACTCATTATTGGAAGTTTTTGCCAGAGCAATTAGGTAAG

10 Sequence ID - 476 nt: 476

CAGAATCTTTTCATAGGCTGAATGTTGCTCCACAATGTGTCCTTTGACTATCTCTG
GCTAATTATTATTTTAATCTCTTCTCAGCTTTTCCAAGAACATAACGTTAACCAAA
GATCTTAGGCCATTCACAACTCTTTTGTAATAAATTAATGTGGATGTGAAACGAGGC
AACAAATCCTGAAGTAGAAAGTTATTCCTGGCCAGGCACGGTGGCTCACGCCTGTA
15 ATCCTGGCACTTTGGGAGGCCGAGGTGGGTGGATCATGAGGACAGGAGATCGAGAC
CATCCTGGCCAACATGATGAAACCCCATCTCTACTAAAATACAAAAAATTAGCTGG
GCATGGTGACGCGTGCCCTGTAGTCCCAGTACTCGGGAGGCTGAGGCAGGGGAATT
GCTTGAACCTCGGAGGTGGGAGGTTGCAGTGTGCCGAGATCACGCTACTGCACTCC
AGCCTGGCAACAGAGCAAGACTCCATCT

20

Sequence ID 477

AAACAGAAAGTTTCTTCTAAAGGCATGATTCAGTTAAGTCATTCTTAAGTGTTAAA
AAATTGTGAAAAATGTGCCTGTAATCCCAACACTTTGGGAGGCCGAGGCAGGCAGA
TCACGAGGTCAGGAGATCAAGACCATCCTGGCTAACAAGGTGAAACCCCGTCTCTA
25 CGAAAAATACCAAAAACATTAGCCGGGCGTGGTTGTGGGCGCCTGTAGTCCCAGCT
ACTTGAGAGGCTGAGGCAGGAGAATG

Sequence ID 478

TTCTTGGGATATTGATGACTACTGTCTGAGAGGTGCTGTGGGGAGATTTTCAGGAT
30 TGTGTGGTCTTTGAGGGGGGTGTTTTTTTAAGACAACATTGACCACTGTCCACTGT
CCACATGATCATTGTAAAATTGCAATGCCGCATGCTAGTTGGTTACATAAGACATA
ATTCCAGTGATTGAAGGTGGTTACACTGTATGGTGGTGTGTTCAAGATGGCACTGG
CATCTTTGAGCAGAGCCTGGCTATGCAGCATCATTTGAGTTTTTTTAAACACCCTAN
AGGTCTGGTTGTTGTTGCTGTTGTCCTTTCCTGTGAAAGTCACAAANANAAGTTACA
35 GTCCAGGTGAACCTGGAGTTTATAGGTTGGTTTTGTTTTCTGNTATATATATATA
TATATATTTTTTTTTTTTTTTTAAACATTTACCTGTAGTGCTGTAGCTGTTGATACTA
TCACCTGCATGCTATTTCTAGTGAGTGCTAAATACAGTATGGTCCAATGACAATAA

- 162 -

CAGCCCATGGTACTGCCAG

Sequence ID 479

5 CATCAGTCTGTTATCCATGCTGACTTTCCGAAGACTTGCAGCTACTGCATTGATAT
CTTTCCTGCCAATAAGCAAAGTGTTGAACACTTCACAAAATATTTTACTGAGGCAG
GCTTGAAAGAGCTTTCAGAATATGTTTCGGAATCAGCAAACCATCGGAGCTCGTAAG
GAGCTCCAGAAAGAACTTCAAGAACAGATGTCCCGTGGTGATCCATTTAAGGATAT
AATTTTATATGTCAAGGAGGAGATGAAAAAAAAACAACATCCCAGAGCCAGTTGTCA
TCGGAATAGTCTGGTCAAGTGTAATGAGCACTGTGGAATGGAACAAAAAAGAGGAG
10 CTTGTAGCAGAGCAAGCCATCAAGCACTTGAAGCAATACAGCCCTCTACTTGCTGC
CTTTACTACTCAAGGTCAGTCTGAGCTGACTCTGTTACTGAAGATTAGGGAGTATT
GCTATGACAACATTCATTTTCATGAAAGCCTTCCANAAAA

Sequence ID 481

15 CACACTTTCATGATAAAAACAGAACCTAGGAATGAAAAGAAATTATAGCAACATAA
TAAAGACCATATATGAGAAGCCACAGCTAACATACTGTATGGTGAAAACTGAAA
GCTCTTCCTCTAAGATCAGGAACAAGGCAAGGATGCCCATTCCTTGCCACTTCTATC
GAACGTAGTACTGGAAGCCCTAGCCAGAACAACTAGGCAATAGAAAGAAATTAAAG
GCATCCATNTCAGAAAGGAAGAANCAAAATGCTGTCTGTTTAANATGACA

20

Sequence ID 482

TTTCTATANAAAAAATTTTTTAAATAATTGTAAAGTTAGATTTAAAATTGTAAA
ATATAAAATCACAAAGGAATGTACCCAATAAAATGTAAATGCNCCATAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

25

Sequence ID 483

CGNTAACGTGCAATCCGCCGCACGCCAGCAAACCTGGACAAACTCCGGGATCTCATC
GAAGCGATTGAGCACCAGTACCAGAGTAATACCGGACTGATGTAACGAGGCGAGTC
GCTCATCCAGCTTGCTGACGTGAGGCAACATCCAGGCCATCGAACGGNTCATCAAG
30 AATCAACAAGTCAGGCTCCGACATCAGCGCCTGACACAGCAGGGTTTTTTCGCGTCT
CGCCAGTGGAAAGGTATTTAAAGCGTCNGTCGAGGAGGGCGGTAATACCGAACTGC
TGCGCCAGTTGCATGCAACGCGGTGCATCCTTTACTTCATCCTGAATGATCTCAGC
CGTAGTGCGTCCGGTGCCATCTTCGCCAGGGCCGAGCATATCGGTGTTATTCCGCT
GCCATTCGTCGCTGACGAGTTTTTGTCAATTGCTCGAAGGAGAGACGAGTGATGTGG
35 GAAAACTGGCTTTGCCGTTTCACTTTCAAAAGCGGGAAGTTCCCCCGCCAGCGCGC
GGGCCAGGGCCCGAT

- 163 -

Sequence ID 484

TTTTTTTTTTTTTATTCTATTAAAAAATGTTNNTGAAAAAAGATACTTAAATTTTAA
AGATAACTNAATTCCTAANGATTTAAAATAATCCAAGCAGAGATGAAAGANCAAAT
GCAAATGCNTAAAAAGACCCCANAGCATTGTTAGCAAAAAGCAAATATAGTTAGCC
5 AAGCATATATATNTCATAAAAGCAATAANAAGGCNTAAAGCAAGTTTGGGGAGAGC
TTATTTAAAACTTGTA AAAATCATT TGAATTTT TAAAGTTT TCAAAC

Sequence ID - 485

nt: 551

TTTGGAACACAAAGTTCCCTTTTTTAGAAGAATAGGTATTGAGCCCTTGAGCGTGGG
10 TAGAAAGATAGAGACAGAGTGATTTGCAAAATAATGGAGGATCATATTTATATATG
AATTTTCACTTATTTGAACTTTCAGATATCANCTTNA AAAANCTTTGGTTTAAGTAA
AGTNTNTTAATGAGACTCCTTGGATGAAAGTAACCAAAACCAGTAAAAATAAGGTA
ATAAGGATGTAATAGTTTCTTATGGACACTCAACAGCTAGAATGCAGTTAGTCTCA
GAAAAGAATTAGAACAAATAACTGGAAGGCCATCAGGAGTCCAAAACCATCACTCT
15 TTTATATTTTATATTTTATTTTCTCTCTTCANATGAGCATTCTCTTTCTATGTCC
ATATGGTANAAGGCGGCAGCTCCATAGATTATGGCTTCAGATGTTACAGTTCCGCT
NAATGCAGGGACAGACTTGCTATCTTTCAGTCCCCTTACATATCCTGGGGAGAGAG
CAAATGATTGACTGGCTTGAGTCAGGTGCCCCGTTCCCTTTCCAATCT

20 Sequence ID - 487 nt:224

GTTTGNTTGTGACCATCTGTACTTGTAATTTCTTTACNTTCATTGGTATGAAAAAT
ATGTTCTTAGAAGCANGAAAAAGAATTCAGNTTTGCTTTGTATACTAAATTAAATG
CTGTAATTTTGATAAAATGAAAAATCTGCTTTATTTGCAACAATTGGTTTCTTCCT
TGACGTCAGCCTCACTCTTGGACTTTGGTATTCAGCCNGNCACCCCTGGGAATTCC

25

Sequence ID - 488

nt: 349

GTGCCTCCCTGTGTGAGTAGCCTAAGGTGCATTGAAAAAGACTGGGATGTGTTTTA
TTTTTTTGTATTAGATAGCATTAACCTTACTGTTGAAGTATTTTTTGGTGGAGTATT
AGTGACAAGCCATTGAGTCTTAAGCCTTACGGCTTCCTATAAAATCACTAATTTTCG
30 TGTGTGTTTGTGTGTAGGTTACGTTATATATAGGATTCGTGTTCCGCGTGGTGGCC
GAAAACGCCCAGTTCCTAAGGGTGCAACTTACGGCAAGCCTGTCCATCATGGTGTT
AACCAGCTAAAGTTTGCTCGAAGCCTTCAGTCCGTTGCAGAGGANCGAGCTGGACN
CCCTGGGGGGGCTC

35 Sequence ID 489

TTAACAGCTGCATAGAGTTTTTAAAGTACATTATATTTTGTGAGACAAGTAAAATA
TCTGTTTTTTCACGCAAAAAAAGCCATGAAATACGTAATTTTTTTAAAGACAAAAAAT

- 164 -

CATCTTTTGAGTTTGCTCTTTGGTTTTCTTCATTCCTTTTGAGGATTGGGAAAAC
AGAAAGATTCTTTGATTTGGGTAATGAAGAGGTAATTTGGGACAGTGTGGTGGTAC
CAGGAAGAAAGAGGATTGGAAAGGCCAGTACTGTTTTAGTTGCTCGGCACTGTTGG
TTTTGTTTTAATGTGGTTGCCCTGTCCACTACATGGTTCTATCAGTAGTGTAATCC
5 ATTTTCAATGTAAAGCTCTTTTAGTTTTTGTTCATAGACATAAATTAATATTTTGAG
AGGCATCCCTCACCTGTTTCAATTTCTTCTGTGTTGAAATGAAGTACTTAAAATTACC
GTTATACATGAACTTTGTGGACTGTAAGATTTGTTATATATGTTCAAATGCCTTTT
AGCTGGCTTTTTTAATTAATATGCCTGTTTTTGAGTGCTTAATACAATGTAATGNGGA
TTGTAAATCATACCTATTTTAAATCATTCTTCCTGTATATTTGNACTCAGAGAGC
10 CTTATTTTATTCTTCCAGC

Sequence ID - 491

nt: 382

TTTTCTTAGAACTTTATTTTTTCTGGCCAGGCGCAGTGGCTCACACCTGTAATCCC
AGCACTTTGGGAGGCCAAGGCAGGTCGATCACCTGAGGTCAGGAGCTCAAGACCAG
15 CCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAATTAGCTGGGCG
TGGTGGCGCATGCCTGTAATCCCANCTACTCAGGAGGCTGAGGCAGGAGAATTGTT
TGAACCCGGGAGGCGGAGGTTGCANTGAGCCGAGATTGCGCCACTGCACTCCAGCC
TGGGCAACAGAGCGAACTCCATCTCAAAAAAAAAAAAAAAAAACAACCTTTATTT
TTTCTGATTTTAAAAGTAATAACTAGTTTGTAGAAACATTAAAAGT

20

Sequence ID 492

ACCCTAAACATAACTTAAAATTTGTTNGGAATTTGAAAGTACAGAATTTTCCTGTA
ATTGAGACTNTTTAAACTTTTGTGGTTGGAGAAGGTATTCTATTTTTTTGAAAATAT
CTGTAAGTTTTATCTAAATAGTAACTCTAAGTATTCTTCCCCTTTACTTACAGCC
25 ACCCTGGGAATCTGAGACTAGAGAAAATAAAGTTTGTCTCTTGTTCTAAGGAGGGT
CTGGTTTAGAAATCTGATTTAGACATAGAAAAATTGCAAGAAGCTTGAGGTGATTG
GAAGATACGATTTTGTATCAAAGNATGTTTCTGTTTTATAGATTTTATTCATCTA
CAACTCCTTATTAATATATTTAAGAAGTCATTAACCCACCATTGATTACTTGATAT
AAAAGGAGAAACGGTGGTAAAAGGTGAAATANAATTTTTTAATTTTTTTTTTTTAA
30 GTTTAGGATTTTTTTTTTAAATTCTAAGAGTTTCTGTCAATTTGGGGACAATCAGAA

Sequence ID 493

TGGGAATCATAATTNGTTAACTGAAGCTNATAAGATGAGAGCATTCANAGAGAAAA
GAACGGAAAGATTGAATATCAGTTTCCCTTCTTTAAAAAAATTGTGGATATGTGAT
35 CTAGCTTCTTGAGCATCACAGTGACTGATTGGCTCGTGGTAATTGATCGCTATGCT
GACAATCTTATCTCCACCTATGTCATTCAATTTTCTAAGAGGCAAAATCCTTAATC
AGGAGGAGAGTTTAGCTCTAGCTAAATTTCCCTTGTCAGCATGCTCCTGCTCCCC

- 165 -

CAACTTGTGGAAACAGCTAAAGGATTGGACTAGGAGCANAAGTTTGGAATGGTTAA
AATGTAGCAACATGTGTTTCCTGAAACAAAATTCCACTATAATAAAAAAAGCATTT
GAATGCTCCCTTGTAATTCTGTTGGAGCTTGTGTCCTTTTTTATGACACAACCATA
ATCAGTGATAGACAGTAGCATAAAGAAGCAAGAGCAAAGCAATTAAGTAATAATAG
5 CACTACAAAAATGTGTGCTGTACTTACCAAACACGACATTTATGAATTATTANATA
GGAATAAGGGGATGGT

Sequence ID 494

10 GACCCAGCCATCTAAATAAGTTRTACATGTTGCGTATTTTTTTTGTAGGGACTTAT
CTTCCGAAGAGGAAAGGTTTATGAAACCTAAAGTAACAATGATAGCTTGGAATCAA
AATGATAGCATTGTTGGCACAGCTGTGAATGATCATGTCCTCAAAGTGTGGAATTC
TTACACTGGACAACCTGCTTCATAACTTAATGGGACATGCTGATGAAGTATTTGTTC
TGGAGACACATCCCTTTGATTCCAGAATTATGTTATCTGCAGGACATGATGGCAGC
15 ATATTTATATGGGATATTACAAAAGGTACCAAGATGAAACATTATTTTAATATGGT
AAGTGAAGTGAGATGTACCTTGATACATGCTTGATAATTTGTTTAGAGTATTTGGG
TTATGCGGCTTACCCAGAAATTGATCTGCTTGTTTTGGCAGTTTGTTTTTACAAAT
CAACATATTCAAAGCCTGCTAAATATTAGACAGCTACATGTATATACGTACATACA
TGAA

20

Sequence ID 495

TTTC

Sequence ID 496

25 CTCGCTGGCGGGAGGCCACGGGCTTTCCACAGCGCGGGGGAACGGGAGGCTGCAGG
ATGGTCAAGCTGACGGCGGAGCTGATCGAGCAGGCGGCGCAGTACACCAACGCGGT
GCGCGACCGGGAGCTGGACCTCCGGGGGTGATCTGGACCTCTGGCATCTCTCAAA
TCGCTGACTTACCTAAGTATCCTAAGAAATCCGGTAACCAATAAGAAGCATTACAG
ATTGTATGTGATTTATAAAGTTCCGCAAGTCATAGTACTGGATTTCCAGAAAGTGA
30 AACTAAAATTTTAATCCAGGTGCTGGTTTGCCAACTGACAAAAAGAAAGGTGGGCC
ATCTCCAGGGGATGTAAAAGCAATCAAGAATGCCATAGCAAATGCTTNAACTCTGG
CTGAAGTGGANAGGCTGAANGGGTTGCTGCAGTCTGGTC

Sequence ID 497

35 GAAGACCTCACATCTGAGAGCTCATCTGCGTTGGCATTTCTGGAGAACGCCCTTTTG
TTTGTAACCTGGATGTACTGTGGTAAAAGATTTACTCGAAGTGATGAATTACAGAGG
CACAGAAGAACACATACAGGTGAGAAGAAATTTGTTTGTCCAGAATGTTCAAAACG

- 166 -

CTTTATGANAAGTGACCACCTTGCCAAACATATTAAAACACACCAGAATAAAAAAG
GTATTCACCTCTANCAGTACAGTGCTGGCATCTGTGGAAGCTGCGCGAGATGATACT
TTGATTACTGCAGGAGGAACAACGCTTATCCTTGCAAATATTCAACAAGGTTCTGT
TTCAGGGATAGGAACTGTTAATACTTCCGCCACCAGCAATCAAGATATCCTTACCA
5 AACTGAAATACCTTTACAGCTTGTCACAGTTTCTGGAAATGAGACAATGGGAGTA
AATATTACACAAATACTTATTCATTGNGGTTATTTTTTATACAGTAGTGAGAAGAAT
ATTGTTCTTAAGTTCTTAGATATCTTTTTTTGGATGTGCAAAAATTTTTGGATTGA
CAGTAACTTGGGTATACATGACACTGAAATGCCTTACTTTGGATGA

10 Sequence ID 499

TGCCTGCGGGCCAGGACCTCGCCCAGCCCATGTTTCATCCAGTCAGCCAACCAGCCC
TCCGANGGGCAGGCCCCCAGGTGACCGGCGACTGAGGGCCTGAGCTGGCAAGGCC
AAGGACACCCAACAACAATTTTTGCCATACAGCCCCAGGCAATGGGCACAGCCTTCC
TCCCCANAGGACCCGGCCGACCTCAGCGCCTCCTGCAGGCTAGGACACTGGTGCAC
15 TACACCCCATGCCTGGGGGCGGAGATTCTCCAGCAGAAAGATGCAATATTTTTTGT
TTCCTTTTTTTCCATTTTTTTCTCTAAGGAATCAATATTTCAATATGTTGAGTGTG
TGTCCAATGCTATGAAATTAAAATATTAAATAACATATTTATGGCATTTCCTTGAA
GAGTGTGGTTGAAGAAATATTTCTCCTTTTGTTTTTCTTTTTTTTTTGNNTGNAC
TGCCACTTCTTTTTAGGAGCAAATCTCCCCAGGGGTGTACGGNATTTCTTGACTCT
20 GGGAACAGCTGCTACCCCCAAGACTTGCCACGTTGTTCTGCCCTCAAATGGAATTA
AGTG

Sequence ID - 500

nt: 390

GGAATATGGTCAGGATCTTCTCCATACTGTCTTCAAGAATGGCAAGGTGACAAAAA
25 GCTATTCATTTGATGAAATAAGAAAAAATGCACAGCTGAATATTGAACTGGAAGCA
GCACATCATTAGGCTTTATGACTGGGTGTGTGTTGTGTGTATGTAATACATAATGT
TTATTGTACANATGTGTGGGGTTTGTGTTTTATGATACATTACAGCCAAATTATTT
GTTGGTTNATGGACATACTGCCCTTTCATTTTTTTCTTTTCCAGTGTTTAGGTGAT
CTCAAATTAAGAAATGCATTTAACCATGTAAANATGANTGCTAAAGTCAGCTTTT
30 TAGGGCCCTTTGCCAATAGGTANTCATTCAATCTGGTATTGATCTTTTCACAAA

Sequence ID 502

ACCCGCCATCTTCCAGTAATTCGCCAAAATGACGAACACAAAGGGAAAGAGGAGAG
GCACCCGATATATGTTCTCTAGGCCTTTTANAAAACATGGAGTTGTTTCCTTTGGCC
35 ACATATATGCGAATCTATAAGAAAGGTGATATTGTAGACATCAAGGGAATGGGTAC
TGTTCAAAAAGGAATGCCCCACAAGTGTTACCATGGCAAACTGGAAGAGTCTACA
ATGTTACCCAGCATGCTGTTGGCATTGTTGTAAACAAACAAGTTAAGGGCAAGATT

- 167 -

CTTGCCAAGAGAATTAATGTGCGTATTGAGCACATTAAGCACTCTAAGAGCCGAGA
TAGCTTCCTGAAACGTGTGAAGGAAAATGATCAGAAAAAGAAAGAAGCCAAAGAGA
AAGGTACCTGGGTTCAACTAAAGCGCCAGCCTGCTCCACCCAGAGAAGCACACTTT
GTGAGAACCAATGGGAAGGAGCCTGAGCTGCTGGAACCTATTCCCTATGAATTCAT
5 GGCATAATAGGTGTTAAAAAAAAAAAAATAAAGGACCTCTGGG

Sequence ID - 503

nt: 109

ACATTTTCCGGNCCTTTTGCCATACACAGTTACAGAGATCAGTCAAATCCATACCA
CCACTGAGATCTCATTATTGCCACAGATGCACAAAATAAATAACCCAAAATC

10

Sequence ID - 504

nt: 374

CCAGCAACGACCCATACCTCAGACCCGACGGCCCGGAGCGGAGCGCGCCCTGCCCT
GGCGCAGCCAGAGCCGCCGGGTGCCCCTGTCAGTTTCTTGGGACATAGGAGCGCAA
AGAAGCTACAGCCTGGACTTACCACCACTAACTGCGAGAGAAGCTAAACGTGTTT
15 ATTTTCCCTTAAATTATTTTGTAAATGGTAGCTTTTTCTACATCTTACTCCTGTTG
ATGCAGCTAAGGTACATTTGTAAAAAGAAAAAAACCAGACTTTTCANACAAACCC
TTTGTATTGTANATAAGAGGAAAAGACTGAGCATGCTCACTTTTTTATATTAATTT
TTACAGTATTTGTAAGAATAAAGCANCATTTGAAATCG

20 Sequence ID 505

GTACAGGAGGTAAATTGGATACCCCATCTAAGGGGATCTGTGAGACCAGGTAGTTA
TTTGGAATGAAAGAGTAAGATATTAAACCAGCCAGCATGTCAACAGGTGGGTGATA
GTCTTGTTCTCACAGACAACAGATGGCCATCATCTTAAAACAACATTTATGTTAAC
25 CAGCAGATAAGGGACTCCTGCATTGTCAGTGGACTTTGAGCCTGAGTTTTTCTACT
TGCATAGGTGAAAGTGGACTGCAATGCTAGTATAAATGCCGTATGATGACTAGTAC
CCCTTAGGGAGCTCCAGTTTGCCTTCCTGGGGAACCACAGACCCCAAGTGTAATTT
CCTGAGGACAGCCCGACTTCT

30 Sequence ID 506

GTTACTGTGAGCCTGTCAGTAGTGGGTACCAATCTTTTGTGACATATTGTCATGCT
GAGGTGNGACACCTGCTGCACTCATCTGATGTAAAACCATCCCANAGCTGGCGAGA
GGATGGAGCTGGGTGGAACTGCTTTGCACTATCGTTTGCTTGGTGTTTGTTTTAA
ACGCACAACCTTGCTTGTACAGTAACTGTCTTCTGTACTATTTAACTGTAAAATGG
35 AATTTTGACTGATTTGTTACAATAATAACTCTGAGATGTGTGAAAAAAAAAAAAA
AAAAAAAAAAAAA

- 168 -

Sequence ID - 507

nt: 521

CTGCGGTGGAGCCGCCACCAAATGCAGATTTTCGTGAAAACCCTTACGGGGAAGA
CCATCACCTTCGAGGTTGAACCCTCGGATACGATAGAAAATGTAAAGGCCAAGATC
CAGGATAAGGAAGGAATTCCTCCTGATCAGCAGAGACTGATCTTTGCTGGCAAGCA
5 GCTGGAAGATGGACGTACTTTGTCTGACTACAATATTCAAAGGAGTCTACTCTTC
ATCTTGTGTTGAGACTTCGTGGTGGTGCTAAGAAAAGGAAGAAGAAGTCTTACACC
ACTCCCAAGAAGAATAAGCACAAAGAGAAAGAAGGTTAAGCTGGCTGTCCTGAAATA
TTATAAGGTGGATGAGAATGGCAAAATTAGTCGCCTTCGTGAGAGTGCCCTTCTG
ATGAATGTGGTGCTGGGGTGTTTATGGCAAGTCACTTTGACAGACATTATTGTGGC
10 AAATGTTGTCTGACTTACTGTTTCAACAAACCAGAAGACAAGTAACTGTATGAGTT
AATAAAAGACATGAACT

Sequence ID 508

AAGCTCATGATTTTAAATGTATTTTCTAATAAACTATACTCCCATTTAAAAATCA
15 CCAATACCTTAATGTTTCAATTATATAAGCTAATTAAAAATAAAGGCTGGGCGTGG
TGGCTCACTTTGGAAGACCGAGGCAGGCAGATCACCTGAGGTCAGGAGTTCGAGAC
CAGCCTGCCCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAATTAGCCAGG
CATGGTGGCACATGCCCCGTAATCCCAGCTACTGGGGAAGCTGAGGCAGGAGAATCA
CTTGAACCTGGGAGGCAGGGGCTGCAGTGAGCCGAGATCATGCCATTGCACTCCAG
20 TCTGGGCAACAATAGTGGAACCTCCATCTCAAAAATAAATAAAAAAATAAAATAAAA
ATAAAATTCAAACCTAAAATAGATGCTCTACTTCAGGAGTGGGCAAATTAATCACC
TGCATCCTTTTTTTTGGGCTTTC

Sequence ID - 509

nt: 575

TTTTTTTCTAAATGGNGATTACTAATATATGTGGAGACTATTAATCTCTTTTCTGT
25 TGCCATTAGTTCATTTTTTCCCCAAAAGCCAATACATGTTTCATTACAAAATGAATT
ATAAAATATAAGTTAAAAGAAAAACATAAAACCCTACAATCTTACCCACCCAGACA
ACTACTATTAATACCTTAGTATTAACATATACACATCATGTATATGTATAAATTTA
TCTTAAACAAAAATAAAATTATTCTTTACATATTGTTTTTAAACCTATTTATCTGG
30 CCAGGTGCCGTGGCTCACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCACGTG
GATCACCTGAGGTCAGGAATTCGAGACCAGCCCAGCCAACATGGTGAAACCCTGTC
TCTAATGGTTTAAATACCAAAAAATTAGCTGGGCATGGTGGCACATGCCTGTAATA
TCAGCTAACATGGGAGGCTGAGGCAGGAGAATCACTTGAACCANGGAGGGGGAGGT
TGCAGTGAGCCGAAATCACACCACTTCACTGCAGCCTGGGCAACAAAGCAAGACTG
35 TCTCAAAAAGAAAAA

Sequence ID 510

- 169 -

CACTGTCATTCCCAGGAGGCTTTGGAGTCAGAACTGGATTCAAATTCTGACTNTAT
GTTGTGTGACTTGGGCCAATAGCTTCTTTNTGTGCCTCAGTTTCTTTAGCTGTAAA
TANACGGGTAGGTCACCCCTTACCCCATAGGTTATGGGGAAAGTTACAGAAAATGG
TCAGCTGGGCNCAGTGGCTCAAGCCTGTGGTCCCAGCNCCTTGGGAGGCCAAGGTG
5 AGCAGATTGCTTGAGCCCAGGAGTTTGACACCAGTNTGGCAACGTGACGAAACCTT
ATCNCTGTGAAAAATACAAAAAATTAGCCAGGCATGGTGGTGTGTGTCTGTGGTTC
CAGCTGCTTGAGAGTTTGAAGTGGGAGGATCACCTGAGCCCAGAAGGTCGAGGCTG
CAGTGAGCTGTGATCGCGTCACTGCACTCCAGCCTGGC-GACAGAGTGAGA-CCCC
T-TTTGAAAAAAAAAAAAAAAAAAAT

10

Sequence ID 512

GTGAGCGGTGGTGGTTTATTCTTCCGTGGAGTTAAGGGCTCCGTGGACATCTCAGG
TCTTCAGGGTCTTCCATCTGGAACATATAAAGTTTCAGAAAACATGTCTCGAAGAT
ATGACTCCAGGACCACTATATTTTCTCCAGAAGGTCGCTTATAACCAAGTTGAATAT
15 GCCATGGAAGCTATTGGACATGCAGGCACCTGTTTGGGAATTTTAGCAAATGATGG
TGTTTTTGCTTGCAGCAGAGAGACNCAACATCCACAAGCTTCTTGATGAAGTCTTTT
TTTCTGAAAAAATTTATAAACTCAATGAGGACATGGCTTGCAGTGTGGCAGGCATA
ACTTCTGATGCTAATGTTCTGACTAATGAACTAAGGCTCATTGCTCAAAGGTATTT
ATTACAGTATCAGGAGCCAATACCTTGTGAGCAGTTGGTTACAGCGCTGTGTGATA
20 TCAAACAAGCTTATACACAATTTGGAGGAAAACGTCCCTTTGGTGTTCATTGCTG
TACATTGGCTGGGATAAGCACTATGGCTTTCAGCTCTATCAGAGTGACCCTAGTGG
AAATTCGGGGGATGGGAAGGCCACATGCATTGGAAATAATANCGCTGCAGCTGTGT
CAATGTTGAAACAAG

25

Sequence ID 513

TTTTTTTTTTATAAACTCCAATCATTTCCAGAGCTACTTAGCTCAGCATCTTTTTT
TTCCACGCTCTTAAGTTGTGTTTATACATTTTTTGATACAGTTAGATTGTTTTTGTG
ACATTCTTCATTCTATCCTGGGATCCCCCAACCACCTAAGTGGATTTTTTTGATAAT
TTGCATGCTTTAAGGATAACTCTTCATTCTGNAAAGGGCTATGGGTTTTTGGCAAAT
30 GCAGAGTCATGTATCCAAGATTACAATATCGCACAGAAGAGTTTCATCACTATATA
AACTCACCAGTCTTCCTCCTATTCAACCATCTCCATGCCTTCTTCCAGCCCTAA
CTCCTTAAAACCACTCATATCTTTACTATTGCTATAGTATTGCCTCTTCCACCATG
TCATATAAATGGAAACATACAGTATTAGTCTTCTCAAACCTAGTTTCTTTTACCTAA
CAACATGCATTTAAGATTCATAGTGTCTTTTAATGACTTGATAGATTATTTCTTTG
35 TAGCTGAATAATATTGCATCTTATAGATGTAACCGTTTGTATATCCATATTTTCTC
ACAGCCTATGACTTGNCTTTTGATTCTCTGAACAGGCCATTACAAAGCAGAAGTT
TTAATTTTTTATAAAGCTAATGNATCAACTT

GATTTACTGTGGGAATTTGCTCATGCAATTATGGAAACCTAGAAGTCCCATATAATAT
GCCATCTTCAAGCTGGAATCCCAGGAAAGCAGGTGGTGTAAATTCTGAGATTGAAGT
CTTGAGAACCGGGGGAGTCAATGGTGTAACTCCCAATCTAGGGCTTAAGGCCCAAG
GACCAGGGCTGCTGGTGTGCAGATGCAAATCCTGGAGTTCAAAGGATTGAGAACCA
35 GGAGCTCTGGTGTCTGAGGGCAGTAGAAGATGGATGTTCCAGCTCAAGAAGGGAAA
GTAAGAATCCGTCCTTCCTCCACTTTTTTGTTCCTATTTCAGATGAGCCCTCAATGGA
CTGAACGATGCTCACCCACACTGTGAGGGCTGGTCTTCTTTATTCAATCCACTGAC

- 171 -

TTAAGTGCTGATCTCTTCTGGAAACACCTTCACAGACACACCCAGAAATAATGTTC
TACCAGCCATGGGCCTGTTACTTAGCCCAGTCAAGTTGACACAGAAAATTAGCTAT
CACAACATCTGTGTGTGTATATACATATGTATTTGCATGTGTGTGTATATATGGNG
TATATATATTCATGTGTGTGTATAT

5

Sequence ID 524

CTTTTGCCAGTAGGCCCCCTGAGTAGGTTCTCTATCTTTTGGCATGACCCCAGAA
GTCTTTGATAACTTCCTTGCTTTCTGATGTGACAAGACATCCAGGGCCAGATTGTC
CATATCCTGCCCCGGATGCACGATGCACTGTTTCTCCAAGAATCCCTGTGTCCTTT
10 GCTGATGATGCCATGATTTTAAGTTCTCTAATATAGTTTTATCTCTTTGTTTCAGA
TAATGCTTTTGTGTTCTCACATGTCCTGCTCTCTCTCTCTCTCATTTTGGTGTT
GATCAGTCTTTCCATAAGATTGTTTATTTCACTAGTCCTTCATTCTTCTTTTTTCT
AAATTTACTCTTCTTGACTAGTATCCTGTCACCTTCTGAGGACTCATATTTTTTGCAA
CTTGAAAATTATTCTTATTTATTTAAGTATATGTTNCTGAAACTCTCATTAGACAC
15 ATTTTG

Sequence ID 525

GTTAAAAAAAGTAAAAGGAACTCGGCAAATCTTACCCCGCCTGTTTACCAAAAACA
TCACCTGGTAGCATCACCAGTATTAGAGGCACCGCCTGCCAGTGACACATGTTTA
20 ACGGCCGCGGTACCCTAACCGTGCAAAGGTAGCATAATCACTTGTTTCCTTAAATAG
GGACCTGTATGAATGGCTCCACNAGGGTTCANCTGTCTCTTACTTTTAACCAGTGA
AATTGACCTGCCCGTGAAGAGGCGGGCATAACACAGCTGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAATTTT

25

Sequence ID - 526

nt: 516

CTTTTCATGGTCTCTTGTTTCATTAATCATCTAAAATCCAAGCNCAGAGAATTCAAT
TTTAGATGGTCTCCAGAGCAGAATTTGATGTATAATCTTAATTACAAATCATAGAT
AATTAATATTGNTTACAAAATCANAATACGATTAGAGGTAGGGATCCTGCACACAC
30 CCTATTTTCTCCCCAGTGTTCTGACCGAGAGACTAATTAATAATTCAAGGAACTT
ACAGTGAATGANAACCCATGGTTTTTGCTTAATTATCAGAACAGCTAGATCTGAGAA
CAGCTGTCTCCACATGGATAGACACTTATTCCACCCATTTGCAGGTAGAATAGCT
GGCAATAATAAGTCCTTCCCATTGGATATGTTGAAAGGTGCCTGCCATGGCATAGT
TGCCACAAGAGAGGAAGAAATGGACACAAATGTAGGCTGTTTTTCAGGGCANAGGGA
35 AGGTGGGAGGAAACCAANTTGCTGGTTTTTCACACACCCTCTGGGGAACACCCATGC
ACCTATGANATG

- 172 -

Sequence ID 527

GACAAAAGCTGAGAGAATTTTTTTCTTGAATATTTGCACTAAAAGATAGGTTAAAA
TTCTTCAGGCTGAAGAGAGCATAACCAGGTGGAGATTTGGATCTACAAAAAGGAAGG
AAGATTTGGAAATGGATTTGGCACCATTGACTCAATTTCCAGAACAAGAAAGCAGG
5 GACAGTTTTGGGAAGCTCAAGACACACTGCCCATGAGCAGCAATTTGGACCTCCTG
CTGCATCCACTGTGCATCAAACACACACTGTACAGACAAAGACTCCCAGGAAAAGA
AGTATAAACATGGACTAACACAGAGATGGGCAAACCTACAGCCTGTGACCCAGCCAC
CTGTTTATGTAGAATCCAAAGTAAGAATCTTTAACTTACACATAAACTT

10 Sequence 529; 660nt

GACAGCAGAGCACACAAGCTTNTAGGACAAGAGCCAGGAAGAAACCACCGGAAGGA
ACCATCTCACTGTGTGTAAACATGACTTCCAAGCTGGCCGTGGCTCTCTTGGCAGC
CTTCCTGATTTCTGCAGCTCTGTGTGAAGGTGCAGTTTTTGCCAAGGAGTGCTAAAG
AACTTAGATGTCAGTGCATAAAGACATACTCCAAACCTTTCCACCCCAAATTTATC
15 AAAGAACTGAGAGTGATTGAGAGTGGACCACACTGCGCCAACACAGAAATTATTGT
AAAGCTTTCTGATGGAAGANAGCTCTGTCTGGACCCCAAGGAAAACCTGGGTGCANA
GGGTGTGGANAAGTTTTTTGAAGAGGGCTGAGAATTCATAAAAAAATTCATTCTCT
GTGGTATCCAAGAATCAGTGAAGATGCCAGTGAAACTTCAAGCAAATCTACTTCAA
CACTTCATGTATTGTGTGGGTCTGTTGTAGGGTTGCCAGATGCAATACAAGATTCC
20 TGGTTAAATTTGAATTTTCAGTAAACAATGAATAGTTTTTTCATTGTACCATGAAATA
TCCAGAACATACTTATATGTAAAGTATTATTTATTTGAATCTACAAAAACAACAA
ATAATTTTTTAGATATAAGGATTTTCTGATATTGCACGGGAGA

Sequence ID 529

25 GACAGCAGAGCACACAAGCTTNTAGGACAAGAGCCAGGAAGAAACCACCGGAAGGA
ACCATCTCACTGTGTGTAAACATGACTTCCAAGCTGGCCGTGGCTCTCTTGGCAGC
CTTCCTGATTTCTGCAGCTCTGTGTGAAGGTGCAGTTTTTGCCAAGGAGTGCTAAAG
AACTTAGATGTCAGTGCATAAAGACATACTCCAAACCTTTCCACCCCAAATTTATC
AAAGAACTGAGAGTGATTGAGAGTGGACCACACTGCGCCAACACAGAAATTATTGT
30 AAAGCTTTCTGATGGAAGANAGCTCTGTCTGGACCCCAAGGAAAACCTGGGTGCANA
GGGTGTGGANAAGTTTTTTGAAGAGGGCTGAGAATTCATAAAAAAATTCATTCTCT
GTGGTATCCAAGAATCAGTGAAGATGCCAGTGAAACTTCAAGCAAATCTACTTCAA
CACTTCATGTATTGTGTGGGTCTGTTGTAGGGTTGCCAGATGCAATACAAGATTCC
TGGTTAAATTTGAATTTTCAGTAAACAATGAATAGTTTTTTCATTGT

35

Sequence ID - 530

nt: 660

GACAGCAGAGCACACAAGCTTNTAGGACAAGAGCCAGGAAGAAACCACCGGAAGGA

- 173 -

ACCATCTCACTGTGTGTAAACATGACTTCCAAGCTGGCCGTGGCTCTCTTGGCAGC
CTTCCTGATTTCTGCAGCTCTGTGTGAAGGTGCAGTTTTTGCCAAGGAGTGCTAAAG
AACTTAGATGTCAGTGCATAAAGACATACTCCAAACCTTTCCACCCCAAATTTATC
AAAGAACTGAGAGTGATTGAGAGTGGACCACACTGCGCCAACACAGAAATTATTGT
5 AAAGCTTTCTGATGGAAGANAGCTCTGTCTGGACCCCAAGGAAAACCTGGGTGCANA
GGGTTGTGGANAAGTTTTTTGAAGAGGGCTGAGAATTCATAAAAAAATTCATTCTCT
GTGGTATCCAAGAATCAGTGAAGATGCCAGTGAACTTCAAGCAAATCTACTTCAA
CACTTCATGTATTGTGTGGGTCTGTTGTAGGGTTGCCAGATGCAATACAAGATTCC
TGGTTAAATTTGAATTTTCAGTAAACAATGAATAGTTTTTTCATTGTACCATGAAATA
10 TCCAGAACATACTTATATGTAAAGTATTATTTATTTGAATCTACAAAAACAACAA
ATAATTTTGTAGATATAAGGATTTTCCTGGATATTGCACGGGAGA

Sequence ID 532

GAATTGTGATAGTTCAGCTTGAATGTCTCTTAGAGGGTGGGCTTTTGTGATGAGG
15 GAGGGGAACTTTTTTTTTTTCTATAGACTTTTTTTCANATAACATCTTCTGAGTCA
TAACCAGCCTGGCAGTATGATGGCCTANATGCAGAGAAAACAGCTCCTTGGTGAAT
TGATAAGTAAAGGCAGAAAAGATTATATGTCATACCTCCATTGGGGAATAAGCATA
ACCCTGAGATTCTTACTACTGATGAGAACATTATCTGCATATGCCAAAAAATTTTA
AGCAAATGAAAGCTACCAATTTAAAGTTACGGAATCTACCATTTTAAAGTTAATTG
20 CTTGTCAAGCTATAACCACAAAAATAATGAATTGATGAGAAATACAATGAAGAGGC
AATGTCCATCTCAAATACTGCTTTTACAAAAGCAGAATAAAAGCGAAAAGAAATG
AAAATGTTACACTACATTAATCCTGGAATAAAAGAAGCCGAAATAAATGAGAGATG
AGTTGGGATCAAGTGGGATTGANGANGCTGTGCTGTGT

25 Sequence ID 533

CTTGAACCTCGGAGGCAGAGGTTGCAGTGAGCCGAGATCACGCCACTGCACTCCAG
CCTCGGGGACAGAGCAAGACTCCATCTCAAAACACACACACACACACACACACA
CACACACACACACAAAAACAGATATACTGAACACAGCACAAGTGGGACATAAGAG
ATTTAAAAGGGTTAGAGATGTAAAATGGATCTAGGAATGGAAACCATAAGGNGGGA
30 TTTATCAACTGGATTCTGCANAATGCTGTTAAGGCCAGATGTTAGCAGGTGTTACA
TAAAAAAGGGATACCATGAGCAAAAGTATTTGAACATGGGCAATGGTTGAAACAAG
TTTAAACAGATTATNTTTATTACCAAATCTCTCAAACCTTTAATATGCTATAAACA
TTGTGAAACAATAAAAAAACTTTCCAAAA

35 Sequence ID 534

GGGAAGGGAGCTATGAGTGTGTGTGTGTGTATGGACTCACTCCCAGGTTACCTG
GCCACAGGTGCAACCTTCCCACACCCTTTACATTTCCCAGAGCCAAGGGAGTTTAA

- 174 -

5 GTTTGCAGTTACAGGCCAGTTCTCCAGCTCTCCATCTTANAGAGACAGGTCACCTT
GCAGGCCTGCTTGCAGGAAATGAATCCAGCAGCCAACTCGAATCCCCCTAGGGCTC
AGGCACTGAGGGCCTGGGGACAGTGGAGCATATGGGTGGGAGACAGATGGAGGGTA
CCCTATTTACAACCTGAGTCAGCCAAGCCACTGATGGGAATATACAGATTTAGGTGC
TAAACCGTTTATTTTCCACGGATGAGTCACAATCTGAAGAATCAAACCTCCATCCT
GAAAATCTATATGTTTCAAACCACTTGCCATCCTGTTAGATTGCCAGTTCCTGGG
ACCAGGCCTCANACTGTGAAAGTA

Sequence ID 560

10 GGCGGAGGTTGCAGTGAGCTGAGATGGCGCCATTGCTCTCCCAGCCTGGGTGACAA
GAGCAAAACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAGCAATTTACTTAAAAAC
ATACAAACACAGAGACAAGTATTTTGGAGAAACAAATACCTTTTTTCATTTTTTATA
CCAATGTAACAATAATCCATTAAACACACCTTTACTAACTGTTTTCTAGGAGTCTG
ATATGATGAGGAAATAGGTAAACCTTTAATAGCCAGTACTAAATTAGAGTGGCACA
15 ACTTTCCTGAGGAAAAAGATGGGTATTTTACTTTTCTGTTTTAGAAAAGTGGCTT
GACAACAGTATGCTTATGTCTTAGAGTTTGAAATTCAAGTTCTTGAACATTATTAA
TGGCTACAATCATTCATACCCACATTGGGCTGTATTCTTGATGAATCCAAAGTGAT
TTTCACCTCAACTCTGAATTTTCAATTCTCCTCTTTTGAATATAATACAACCATCTCA
CTAGAGGAAGCATTTTCAAGTCTTTTCTGATTGGAGATTCATTATTGTTTTAGATAAT
20 GTTTTCATTTGCTTATGGGTATATAAAAAATTTTATCTTAAAAATATTTCTCTCA
TTTAGCTAGCAACATTGTTTTTC

Sequence ID 561

25 CTCAGGGTGATCTCTGAACCCAAACTTGCCCCAAAGAAGGTTGCTCTGTCCTCTCC
ACATCCCCATCTCCTCCCTAGGGCCTTGTTGGGGAGAGGCTCCTCCATCTTTCCCA
AGTCACACCATCGTTTCCTACGTGGTCTGGACAAGAGCAAGAGCACACCTTGTCCC
CACCTTCTCCAGAGCAGCCAGAACCCACCTCAGGTGCCTTCCCCATCCGGTGCAGT
TAAGGCACTTCTGCCAGCACCATGGTATGAGCACTAGACTTGGAGTTAAGATTTGA
GAGCCCCCTCTGTCACTGTGGAAGCTTGAGCATGTTGCTTGATCTCTCTGAACCTT
30 GTGTTTCTCATCTGTGAAAGGTGATAATGTGGGGCTGCTGTGAGATTTAAAGGACA
TAATGCACCTACGGTCCAAGCACTGCCTGGAATACAGCANAAGCTCAACAGATACT
GGACAACCCATCCCCTTAGTAGAGGCACCTAACCATGTGACCCAAGGCAAAAGTGCT
TAAAAAAA

35 Sequence ID - 562

nt: 580

ATTGCATGCAAGTTTGCTGAGCTGAAGGAAAAGATTGATCGCCGTTCTGGTAAAAA
GCTGGAAGATGGCCCTAAATTCTTGAAGTCTGGTGATGCTGCCATTGTTGATATGG

- 175 -

5 TTCCTGGCAAGCCCATGTGTGTTGAGAGCTTCTCAGACTATCCACCTTTGGGTCGC
TTTGCTGTTTCGTGATATGAGACAGACAGTTGCGGTGGGTGTCATCAAAGCACTGGA
CAAGAAGGCTGCTGGAGCTGGCAAGGTCACCAAGTCTGCCCAGAAAGCTCAGAAGG
CTAAATGAATATTATCCCTAATACCTGCCACCCCACTCTTAATCAGTGGTGGAAGA
ACGGTCTCAGAACTGTTTGTTTCAATTGGCCATTTAAGTTTAGTAGTAAAAGACTG
GTTAATGATAACAATGCATCGTAAAACCTTCAGAAGGAAAGGAGAATGTTTTGTGG
ACCACTTTGGTTTTCTTTTTTTCGTGTGGCAGTTTTAAGTTATTAGTTTTTAAAT
CAGTACTTTTTTAATGGAAACAACCTTGACCAAAAATTTGTCACAGAATTTTGAGACC
CATTAAAAAGTTAAATGAG

10

Sequence ID 563

15 GCAACCTGCACAACCCCGCCCTGTTTCGAGGGCCGGAGCCCTGCCGTGTGGGAGCTG
GCCGAGGAGTATCTGGACATCGTGCGGGAGCACCCCTGCCCCCTGTCCTACGTCCG
GGCCACCTCTTCAAGCTGTGGCACCACACGCTGCAGGTGCACCAGGAGCTGCGAG
AGGAGCTGGCCAAGGTGAANACCCTGGAGGGCATCGCTGCTGTGAGCCAGGAGCTG
AAGCTGCGGTGTCAGGAGGAGATATCCAGGCAGGAGGGAGCGAAGCCCACCGGCCA
CTTGCCCTTCCACTGGATCTGCCAGCCCTACATCCGGCCGGGGCCCAGGGAGGGGA
GCAAGGAGAAGGCAGGTGCGCGCAGCAAGCGGGCCCTGGAGGAAGAGGAGGGTGGC
ACGGAGGTCCTGTCCAAGAACAAGCAAAAGAAGCAGCTGAGGAACCCCCACAAGAC
20 CTTGACCCCTCTCTGAACCAAAATATGCAAAGTGTGACCAGTGTGGAAACCCAAA
GGGCAACAGATGTGTGTTTACGCCTGTGCCGCGGNTTG

Sequence ID - 564

nt: 671

25 GGAATAGAATTTTAAATAGTAATAACTGCTTGTTTTTTTTTGTGCAAGTACTTTTAT
ACATAAGATAAACAACAAACCTTACCACCAAACATACCAAATGCACCTCTTTCATA
AGTGAGTTACTAAGATTTCTATACCTGGAATATCATGTATGTTTCATTTACTGGAT
GTTTACATTTTAGGAAGGAAAATAGTTTTGTTTATTTAAACAACCTGAATACTTATA
AACTGTTGTTCTTGGAAAGTTATTTATTTCCATAAAAAATTTGTTCTTTTGTGATGAA
TTTATAATTCCCTAAATGAAGACCAGAAAGTACAAATTGCTGGGAGGAAGAATAGGC
30 TTTATTAATCAACTGATGTCTTGATTTTTCTAAATGGGAAGATTGCTTTATTTTAA
ACACTAATTATGGGAGCAGATTCTTAGCAAACCTTCTTTGGAAAAGTTAATGTTATG
ATGTGCATTAGGCTGCCCCATCGTGTATATAAATGAAGCAGATTTGATTTTGTAT
TCTTACGTTTCTCTGCTTTGTAGTTGTGGCTGTACTTAAAGAAATACAGAATTTCA
TATATTTAAAAATGTTTAAATGTGACCCACAGACATTGTAAATGGATTNAAAAC
35 AACATGAAAAATATTCAACCTAAAAGAATTCTTAACCTCACAAGTGTTTTACTTC

Sequence ID 565

- 176 -

CTTGGTTCCGCGTTCCCTGCACAAAATGCCCGGCGAAGCCACAGAAACCGTCCCTG
CTACAGAGCAGGAGTTGCCGCAGCCCCAGGCTGAGACAGGGTCTGGAACAGAATCT
GACAGTGATGAATCAGTACCAGAGCTTGAAGAACAGGATTCCACCCAGGCAACCAC
ACAACAAGCCCAGCTGGCGGCAGCAGCTGAAATCGATGAAGAACCAGTCAGTAAAG
5 CAAAACAGAGTCGGAGTGAAAAGAAGGCACGGAAGGCTATGTCCAAACTGGGTCTT
CGGCAGGTTACAGGAGTTACTAGAGTCACTATCCGGAAATCTAAGAATATCCTCTT
TGTCA TCACAAAACCAGATGTCTACAAGAGCCCTGCTTCAGATACTTACATAGTTT
TTGGGGAAGCCAAGATCGAAGATTTATCCCAGCAAGCACAACTAGCAGCTGCTGAG
AAATTCAAAGTTCAAGGTGAAGCTGTCTCAAACATTCAAGAAAACACACAGACTCC
10 AACTGTACAAGAGGAGAGTGAAGAGGAAGAGGTCGATGAAACAGGTGTAGAAGTTA
AGGACATAGAATTTGGTCATTGTCACAAAGCAAATGTGTGCGAGAGCA

Sequence ID 566

GTCACCAAGAGCTTGTTGTCAGGTTTTCACTTGCTATTCGCAGAGATTTTTTTTAA
15 AGGCACTATTTGTAGTGTTAAAAGGGTGAATTTATCANAAGGCATAATAATCATAA
ATGTGTATATGCCTAATAATAGAACTTTAAAAGGCATGAAGCAACACTCAAAGGA
TTAAAGGGAGATCATCTCACCCCCTTCTTACCAATTGATAGAATGATCTGATGAAA
ACAGTAAAATAACAACAGATCTGAACACTGTCAACCATCTTGACAAATACTTATGC
CTAGTGTTCCATTATTGGAACACTAAACATGTGGAATGATTTATATCCTACTGCTC
20 AAGGTCATACCAAGGTCTAATTGTAAAATTTCAAAAAATTGCAACCTCAGGCATA
AATGGGTTAATCGACATTTATAGCACACACATGCAACATGTACCAGAGATTCCTTC
TTTTCTATGAACATGGTACTTCCACCAAGATAGACCACATTGTGAACTATAAAACA
AATCTAAAACATTTGAAATGAAGGAAATTATATAAAATATGTTCTCTTGATCTCA
ATGAAATTAAATTAATACTATAT

25

Sequence ID 567

CTCATGGCGGCCAATGTAGGCCCAAACCTTCCTCAAGTCAAACCTCTCCAGGCCAC
CTTCTGCTTCCCGGTGGCATCAACAGGCCAGCTTTGACTTGAGAACAGCCTCTGC
AGGCCCTGCTCTTGCCTCCCAGGGGCTTTTTCCAGGCCAGCTCTTGCCTCATGGC
30 AGCTGCCCCAGGCCAAATTTCTGCCTGCCTGCCAGCAGCCTCAACAGGCACAGCTC
CTCCCTCACAGTGGCCCATTTAGGCCCAAACCTCATGACTGTGAGGCCATTTCCAGGC
CTAGTGCCTGCCTCGTGGCTGACTCTTGAAGCCCAAACCTTCCTCAAATCAGCCTT
TTGCCCAAACCTTCTGTCTACTGTGCGGACTCTACAGGTCAGCCTCTGCCTCACAGTGG
ACCCTCCAGACCCAGATGGTGTCTNCTGTGGCATCCTCAGGCGAAGCTCCTGCCTT
35 TCGGCAGCCTCTCCAGGCCAGCTCCTCCTGCTCCAGCCTTCTCTCCAGGCTCTGA
ACTTTCTCAGGTCTCCCTCTGTTGTCCAAGGCTGGAGTGTAGTAG

- 177 -

Sequence ID 568

TATATATGTAATGCCCTTAACCTAGTGTTTGGCATGATCGTTGCTGAAAGGGAAGC
TTGTGGGTACAGTGTCCCCTCAGAAGCCAAAGCCCAGGGAAGGTCGCCTGCCCAGG
TCAGGCTCCCAGCGAGTTTGTCTGGGGAGGGGGCCATTCATACCTCCAGGTCAGGAC
5 AGAGGCTCGGGCTGAGGGAACCCTACACAGGTCCTGGAAGCAGATCCTTCCTGCCT
AAGCCAGCAGGACAGCTCAACAGGAAGCATCTTCCAGCCACGGGAGGAGAGGCAGC
ACCTTTTTTTGGAACCATAACAGAGCTAAGAATGGTGGTACAAGTAATAGATTCTGTA
CTGGCAACCCCACTTGGTGGAGCAAGTTCTAGGAAAAGGGGGCTGTCCTTGAGTCA
GCCATGGGGTCAGCCACACAGTCACCGCAGCTGCTCTTTGGCACCGGGCGCTGGAA
10 AGACCTAGGATGACACAGCCTGGAAAGAGCTTGGGAAAAGCTCATCTTCCACAGAA
CTACCTGCTATAACCAGCCAGGGCAGGTGCTTATTCCACAACAGCCCTCTGTTGTA
GGCGGCAGTGCCATCCTGAANGTGCCGTGGTACCTTCTGAANACCCAGCTGAGGGC
CTGTAATGGCACTTGCATGCCACATGGNACACCCTTTCCCGGTAA

15 Sequence ID 570

ACCGCGGCCGCGTNAANAAAAAAAAAAAAAAAAAGAATTCCACTTGATCAACTTAATTCC
TTNTCTTTATCTTCCCTCCCTCACTTCCCTTTTCTCCCACCCTCTTTTCCAAGCTG
TTTCGCTTTGCAATATATTACTGGTAATGAGTTGCAGGATAATGCAGTCATAACTT
GTTTTCTCCTAAGTATTTGAGTTCAAACTCCTGTATCTAAAGAAATACGGTTGGG
20 GTCATTAATAAAGAAAATCTTTCTATCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

Sequence ID - 571

nt: 457

25 TTAGAGAGGTGAGGATCTGGTATTTCCCTGGACTAAATTCCCCTTGGGGAAGACGAA
GGGATGCTGCAGTTCCAAAAGAGAAGGACTCTTCCAGAGTCATCTACCTGAGTCCC
AAAGCTCCCTGTCCTGAAAGCCACAGACAATATGGTCCCAAATGACTGACTGCACC
TTCTGTGCCTCAGCCGTTCTTGACATCAAGAATCTTCTGTTCACATCCACACAGC
CAATACAATTAGTCAAACCACTGTTATTAACAGATGTAGCAACATGAGAAACGCTT
30 ATGTTACAGGTTACATGAGAGCAATCATGTAAGTCTATATGACTTCAGAAATGTTA
AAATAGACTAACCTCTAACAACAAATTAAAAGTGATTGTTTCAAGGTGATGCAATT
ATTGATGACCTATTTTATTTTCTATAATGATCATATATTACCTTTGTAATAAAAC
ATTTTCCC

35 Sequence ID 572

CGTCTATTTGNGTTTCTTCTCACAATTGGTAAGTTCTCTGTATTGATTGATGGCTA

- 178 -

AGTTTGATTAGTGTTTTTCTCTAGTTGGTAATTATATTCTAGTATTTTATCATCTT
ATTGTTTACTCAACTNAAAGTGNCACAGAAGAGTTGCCAGGTTTCTCTTTGATATG
AGATCTCTNNTTGATTTGGAATGCAAATCANAAGTGTCATGTTTTGAATAAAGGGA
CCAGATGAC¹TTATAGGTATTCTTTCTCTAAATATAACTAAGGTAAGATTTTTTGTTT
5 TGAGGTACTTAATCTATATAAGTGGTAAAGAATTTACTTGAATTTCTCCAAATTCT
CATGTCTAAAGTCTGATTGATTAAATTCATTCTTGGTATTTTCATTTTGAAAAGAAT
GTAGCTTTAGCAAACCTCTTTGTATAAATGCAGTGGGATTAAGGTCATTTAAAAAA
TTGTTATATCATTTGTATTTTTTAAAATTTACCAGTTTTATTTTTCTTTTTACCCTTT
AGCCCGGCCTCAGAAAGTGTGTTTGTGTCCATTTCTCCCAGCGCACCTCTGCATA
10 TCTCTACCCACTTGTCTATAATTCAGCATCCAGCAGAGGAAAACAAAGTGTTGCGTA
CAGTTCCTCTACTAGCAGCATGCCTCCCCCAGGACAAGTGTA

Sequence ID 574

TTATTGCTGACATAAAAAATGGTGCACATCGGCCAGGGCCCAGGATGAATCAGCCAA
15 TCTGCACCATTTATACATGGAAGTGGAGAACATTGTGCCAATAATCATTTAATATA
TGCCAAATCTTACACGTCTACTCTAAACTGCTCTAATGAAGTTTCAGTGACCTTGA
GGGCTAAAGATTGTTCTTCTGGGTAAGAGCTCTTGGGCTGGTTTTTTCANAGCAGAG
TTCTTGTTGTGGGTAGACTGTGACTAGGTTTACAGCCTTTGTGGAACATTCCGTAT
AACGGCATTTGTGGAAGCAATAACTAGTTCCTATGAAAGAACCAGAGCTGGGAAGAT
20 GGCTGGGAAGCCAGGCCAAAGTGGGGGCAACAGCTTGCTTCTCTTTCTCTTCTCAC
CCTCAGTTTGTATGGGAAAATGGAGATGTCCTCTCCACTTTATCCCACGATATCTA
AATG

Sequence ID - 575

nt: 209

25 CAGGATATCGAGACCATCCCAGACAGCATGGTGAAACTCCGTCTCTACTGGAATAC
AAAAAGTTAGCCGTGTGTGGTGGCACGCGCCTCTAATCCCAGCTATTCGGGAGGCT
TAGGCAGGAGAATTACTTGAACCCGGGAGGCGAAGGTTGCAGTGAGCTGAGATCGC
ACCATTGCACTCCACCCTGG-CGACAGAGCAAGACTCCGTCT

30 Sequence ID - 576

nt: 541

CAGCCAACCCAGAAGGAGCCAGTCTACAACCTATGCCTGATCCTCCTCATGGCAGGC
CACGAAGCATTTGCTGCCATGTGTTGAATTATAAAACCCACATTGCTTTTTTGAACCC
TGTTGCGGGTAAAAATAACCAAATTATCAGTCCTTGGAAACCCAGGCAATCAAGTG
AGTACAAGGTAAAGATAAGTATGGTTTAGAGGAGAAATTATGTTCTTGAAGTGGTG
35 TCCTTTGATGGCAGCGTCAGCCTTGCTAAGTCAGAGTAGAGGGAGCAGTGACCTTA
ATAAGCTTTGGTGAGCATCATGTGCACGCGTGGGTGGGAGTCCCTTTCACTGATGC
TTTTAAAAGTGCTTTTGCAGACCCTGGAAGGGATCCTCCACACATATGAGGTGTGG

- 179 -

GACAGGTAGGCCAGAGAGGATTAGCCCTGCTTTCGAGACTAGAAATCTACAGTCCT
GAAGGAGCAGTAATTAATTGGTACACCTGTCAGGGCCAGCCCCCAGGTCTCCTGGC
TTTTTCCAGGTTTTCTGTCTCACATGATTTTGCTTTT

5 Sequence ID 577

CTTTAATTTTTCAAGTGTTTAAAAACAATTTTATACTTAAGCCAGCCTTGAAGAT
AAGCACAAAATTTACCAGTTTACATTTAAAAACAAACAAAAAACGACAACAACCTC
AAGCACCCGCTCTGTGCATAGCACTATTCTAGGTGCAATAAAAGGGAATCTTAACC
TTAGAAATATGAGTTCACTTTCTGGAATTGTATTATCTCCTTTTCCAGAGAGTAAA
10 AATAAATAAAATCACCATTGTTTACTACAGATCTGCCCCAAACCACATCTGGTTCA
CAGAAAGGCTAATTTCTGCCAAATTAAAGATGTAATGAACTCAGTTCCTGCTTTCC
CAAAAACACGAAAGCAGAATTCCTTTTCACTGAAAAAATAAACAGTTTTCCATGC
AAGGGCAGTTTGCTTCTAATAAGTATTTTTTAAAAAATTTTTTTTTCTCTAGCTT
TTCTTTTAAATTTTCTTCTCTAATATTGCCTTTTCTTGTACAAGGCAGACCAGGTA
15 TCTTTTTATGCTGTTTTTCTTTACTAAGAAAAGTATTGCATCTTGAAGACAAACC
ATTTCCCAGAGTAGTGATAAAAAATAACACTAAAAAACTTTAAAGGTGAGTCACT
TCATCACCTTGATGAAGTAAAAAA

Sequence ID 578

20 GGAAAAAATATTTCCACTTAGATATTTTACATGGTTTTGTTTAAAATTACCATTAC
TTGTTTTTTTAAAAACACATGACCACATATGTATATGTATATCTACCTAAACATTGT
ATCATGGTTTTCAGTATGTTATTCATGTATTACTGGGAGATGCTACCAAGAAACCAA
CCCAAAGAAAATTCTGGAAAATACATTTCTATTTATAGAATAAATGTTTCATTTAT
ATAAAAGCAAAAGAACTTAGAGTTCTAATAAATGGGATGTCTAATAAATTATGAAG
25 TTAGTATTTGAATATATTATATTTTTTATAACTTCCTTGCCAAAGTCCTGATTTAG
TACATTAGAGAACCTGTGTTTCCTCTCTCCTCTACCATTCATCTCTCTTCCATACA
GTCATTTGGGCTTTTTTACTCAAAGAGAATCAAGAAATAATAAGGTATAACAAGCTT
GGCAAAGTGTTGGCTTTTTTAAAAAAAATTTTTTTAATCTCTAGCAGTTTGGTAAT
TTAGCAGCATCATTTATTTGGGATTCTTTTATCTGATTTCAACAGTGAAAAACATC
30 CCTATGATAAAGCCTAATGACCCATTTCCAAAAGATGGAATTGCCCTTCCTAGAAA
ATATGACGGAGAAAAGT

Sequence ID - 579

nt: 502

35 CGAATAGCCAAGTGGTCTGACAAGATCGAGAGTAATGAGGCCCATACTTTAGTACA
GTCTTGAATGGCCAGATGGTGCTGGGCATACCCCAACCAGAGATATGTAAGTCTTT
ATGTTGTCAAAATTTCCCAGAAACATGAATTTCCCACTAAGATTCATTAAGGAAAA
CTAGAATGAAAACAAAACGTTCTTGTATAATAATTCATTANAAAGAAATGAAGAA

- 180 -

GGCCGGGCATGGTGGCTCACGCCTGTAATCCCAGCACTTTGAGAGGCCAAGGTAGG
CAGATCATGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATAGTGAAATCCCGTC
TCTACCAAAAATACAAAAAAATTAGCCGGGCATGGTGGCACACACCTGTCATCCCA
GCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGTGGAGGTTGCA
5 GTGAGCTGAGATTGCACCACTGTACTACAGCCTAGGTGACAGTGCAAGACTCTG

Sequence ID - 580

nt: 316

CCTATGCCAAACTAAAGAAAGCTTGCCTGGCCTACAGGCCTAAAGGTTCAAATGNG
GATTAAAAAACACAGTAGTCACATAAAATGTCTGCTGGCTGGCTGGAATTCCATC
10 ACCTACAATTTACCTGCTTTCAAAACTGTGTTCAACATTGAGAAAACAGAAAACC
ACTTATCTTGAGCTTAATATGGGCTTCTTTTTCTTAACTGTAGAACACTTACTGA
AATATCAAATCAATGGTTAGGATATGTATCCTAGGCAGGCCTAAACCATTAACT
TGGTTTAAGCAACTTTGTATAATTNACCTCCTAAAT

15 Sequence ID 581

CTTCATGAGTGCCCGGTTGCCCAAGTCAAAAACCTGGGAGTGATATAAACTCCCCA
CACATCCAGTCAGTCACTCATCAACTCTATTGATTCTG-CTGCTAAATATATCTCA
ATTGTATTAACTTAAACATATGCATAATACATCTTCTTCTTCACTGCATTTTTGTG
GGCTGCACTTACCTTTCAGGTAACAACAACACTGGCCCCTCTTGCCCTTCTAGTCA
20 GAAGTGCCAAAATGATGAGAGCTAGCCATGACAAACCCACAGCCAACATTACACTG
AATGTGCAAACTGGAAGGGCATCCAAACAGAGGAGG

Sequence ID 582

TAGAATTCTCGCCTGCCTTGGCTTCTCCCTCTAGTTGTTCTTCTCTGTCTTCTGT
25 GGGCTTCTTATTGTCTGCTCACTCCTTCTTCAGTGTCTCTCATGGGCTTCCTTCC
CTTCTCAGCTGATGCCATCACCTGGGGAATCACAGTTACTCAGCAGCACTGGGGCC
TCTCTATCTCTATGCTGGTCATGCCTATGTGTGAGCTGCAGACCCAGTGGAATTTCA
CATTTGTGCATCCCATGCCCAGCCCACCCTCCACCAGCCTCGAATGCAGCTGTTCA
GCCCTACCCCAGTCCTCAGAAAAGTTCCTCTCCCTGGATCCTCTTTTTCTTTCATG
30 AGTGCCCGGTTGCCCAAGTCAAAAACCTGGGAGTGATATAAACTCCCCACACATCC
AGTCAGTCACTCATCAACTCTATTGATTCTGTCTGCTAAATATATCTCAATTGTAT
TAACTTAAACATATGCATAATACATCTTCTTCTTCACTGCATTTTTGTGGGCTGCA
CTTACCTTTCAGGTAACAACAACACTGGCCCCTCTTGCCCTTCTAGTCAGAAGTGC
CAAAATGATGAGAGCTAGCCATGACAAACCCACAGCCAACATTACACTGAATGTGC
35 AAAACTGGAAGGGCATCCAAACAGAGGA

Sequence ID - 583

nt: 631

- 181 -

CTGAGGTGGGAGGATTCCACTCTCACCCATTTCTTCTTTCATTTTCAGTTTCTCCA
GTTAGTAACTGAAGATGTTCTTTGAGTAATTAAGTGAGTGAGAAAATTTTAAAGTG
AGAAATCTATAAAAAGAACCATGTTAACATAAATATTTTCAGTCCTTACAAGTTGGT
ATTGACTTTTCTCATTTGGTAATCTGACTGATTTAATACTGCTCATTTCCAATATCTG
5 GTGATGTAATTCTGGTTATGAATCCTTGTATTAATAACACCTCCTGGGAGGTTTTT
TTTCCCCAACATTACATTCAGAATATTAGAGCTGAAAATACCTTTTTTTAAAGGTTAT
CAGGAGGAGGGAGCTTATGTTTAATGTGGTGGATAAAACTTAACTGCTGGTTAATA
CAATTGTTATTCAGGTGAAATTCCTTAACTTTTCACGTGCAAAGTTTTGTATGTA
TACAGACATTTGGGGAAAAGTTTTATCATCCCTAAAACCGGTTACTGTCCAGAAAA
10 TGATAAGAATCCCTGGGTTCCAAATCCTTCATAAGGTATTTATTCATTTATTTATT
CAACACATTTACTCAATGCCTCCGCTCTGCTGCAACTACACTGACATTCTGCTTCT
AATCTAACCGAAAAT

Sequence ID 585

15 TTTCAAATTGTACAATAACACAAACAACTTTGTTAAGGCCATGTTTTATTTGCTGA
TTAATGGACAAAAGGCAATGTAATTTATTTTCAAGTATTTTCTTGAAAGTCTGTGC
TCATAAAAATCATGAAAAGTTGGAAAGACTGTTAAATCACTGAAACTTCAAATATA
TCTTACACAATCTTGTTTGTACAAAATAACAAGTTAAATATAAACATAAAGCAATC
ATGGTAATTTTATGCAAATCTGTTTTATGTGATCATCAGTTATATATAAAAGTTTC
20 TCAGTTCTGTTATTTGTGAAAAGATCAATACCAGATTGAATGACTACCTATTGGCA
AAGGGCCCTAAAAGCTTACTTTAGCACTCATCTTTTACATGGTTAAATGCATTTTC
CTAATTTGAGATCACCTAAACACTGGAAAAGAAAAAAAATGAAAGGGCAGTATGTC
CATAAACCAACAAATAATTTGGCTGTAATGTATCATAAAACACAAACCCACACAT
CTGTACAATAAACATTATGTATTACATACACACACACACCCAGTCATAAAGCC
25 TAATGATGTGCTGCTTCCAGTTCAATATTCAGCTGTGCATTTTTTTCTTATTTTCATC
AAATGAATAGCTTTTTTGTCACC

Sequence ID 586

30 GTAAACTGTTCTCTCCGAGGGAAAAAATGGAAGTTATCCTCACAGTTCACTGCCGT
GGTATTTCTTCTGTCCCATGCTTTGCATGACTGCCATGGTACAGCCTTGTTTCAAA
CTGTTCACTGTGATCTGTGGGTCTTTGAGTTTCAGTGAGTTTGCTGAAATGTCGAA
GAAGTAGTTCCAAACTTCAATGTTCAATGAAATTTTTGTTCAAGTTTGAAATGGAG
AGAGCAGCTTTAAAAGGTACTAAGCCTTTTACAAATTGGTGAGTACTGGCACATGA
GAT

35

Sequence ID 587

TTTTTTTTTTTTCCTTAAAAGGTAACCCCTAAACACAGCTAAAACCTATGCCATCAGC

- 182 -

TGACTCCAAGGNACACACAGTCCTGTATCTGGAAC TACTGAGTGGCAGGCATCTTT
CTCTGCCTCTGACAGTGGAGTCCCCATCACTGCAGAGCATAGCCAAAGGAGTCAAA
GGTCTCAGCGGGTCACTGCCTTATCAACCCTCACCAGTCCCTTATGTTTTTTAATA
TTTTATAATCTTGACATGACACCAAGATGCTTTAATAAAAAAGCACCTCTAACTCG
5 GTCTTGTATTCACTTACCTTGAGCCTGGGACTTCTCTAGGCTCCTGAGGCAAAAAC
AGGTAGAGGGGAGATGGTGGAACATAAAACACAATTTTGTCTTGGCACCCACCTTGG
CGTCTGTCCCCATGACCAGGTCTTTCAATTCGATGATTTTGTCAATTGATGGAGGAG
CGATATCGTTTCTCAATGATATTATGGGTGTCCGCCTTTCTCCTTCTTTGGGGGG
CTCAAGCTGCTTGACTCCCCCAGGTACCTGCTTAATGGGGCACTTTCTCTTGCCCC
10 ATCATTACAGGCATTGTGGTCAGAATGGTCCCCTGCTGCCACCAGGGTCTA

Sequence ID 588

CTAGTCTTTTCATAGTCTGCATAGAGTCTGGCCATTACCATCAGTTTTTAAGATGT
CCATATTGTGGCCGGGCGCGGTGGCTCACGCCTGGTAGTCCCAGCACTTTGGGAGG
15 CTGAGGCAGGTGGATCATGAGGTGAGGAGATCGAGACCATCCTGGCTAACACGGTG
AAACCCGTCTCTACTAAAAAAATATTAATAAAATTTGGCCAGGCCTGGTGGTGGGCG
CCTGTGGTCCCGGCTGCTTGGGAGGCTGAGGCAGGANAATGGTGTGAACCCGGAAG
TCGGAGGTTGCAGTGAGCCAAGATTGCACCTGGGCAACACAGCGAGACTCCGTCTC
AAAAAAAAAAAAAAAA

20

Sequence ID 589

CAATTATTTATTACCTTTCCATTTGTTTCGCCTGATGATGTGACAATGCATGGTCTT
TGTGCATGCTGCTAGACACTTTTCTTTCCCAGCCGAAAAGTCTATTATGTAATTTT
TACATTCATAATTTTAATGTGGATGATCAGGATTAAATCAAGATATATATCTGGAA
25 CCTCTTATAAATGGAGCACTTAGAAATTTGTTGTTCTGCACTTAACCTAGAGAGAG
AAAAAATGCTTTTCTTTGTGAAAAATCTGAATTCCTGTCCTGACCTTCTGTGATGT
GGAAACCCTAGGCTCTGAGACACACTCTCTGGTGTCTGAGACAGAACCAAAGCAAT
AACGTTGTGATGCCCACAGGCCTGGAGCCAGCTAGCGACCTTGTGCCGCCAGCTG
TCCATGGCCCGTGCAGAGCAGAGGACAGTGAGTGTCTGCACTGAGAACCTTAAACC
30 ACAGTTGAACATACCACACCTGTTTGTCTTAAGCTATAGTGTAATAACAAAGTTT
GGGCTCTGAAAATTTAACTGAAAAGATTTCTTGT

Sequence ID 590

GTGGCAGCAGGCGCAGCCCAGCCTCGAAATGCAGAACGACGCCGGCGAGTTCGTGG
35 ACCTGTACGTGCCGCGGAAATGCTCCGCTAGCAATCGCATCATCGGTGCCAAGGAC
CACGCATCCATCCAGATGAACGTGGCCGAGGTTGACAAGGTCACAGGCAGGTTTAA
TGGCCAGTTTAAACTTATGCTATCTGCGGGGCCATTTCGTAGGATGGGTGAGTCAG

- 183 -

ATGATTCCATTCTCCGATTGGCCAAGGCCGATGGCATCGTCTCAAAGAACTTTTGA
CTGGAGAGAATCACAGATGTGGAATATTTGTCATAAATAAATAATGAAAACCTAAA

Sequence ID 591

5 CAGCAGCAGAAATGTTTGCAAGATAGGCCAAAATGAGTACAAAAGGTCTGTCTTCC
ATCAGACCCAGTGATGCTGCGACTCACACGCTTCAATTCAAGACCTGACCGCTAGT
AGGGAGGTTTATTCANATCGCTGGCAGCCTCGGCTGAGCAGATGCACAGAGGGGAT
CACTGTGCAGTGGGACCACCCTCACTGGCCTTCTGCAGCAGGGTTCTGGGATGTTT
TCAGTGGTCAAATACTCTGTTTAGAGCAAGGGCTCAGAAAACAGAAATACTGTCA
10 TGGAGGTGCTGAACACAGGGAAGGTCTGGTACATATTGGAAATTATGAGCAGAACA
AATACTCAACTAAATGCACAAAGTATAAAGTGTAGCCATGT

Sequence ID 592

TACTCAATGAAAAACCATGATAATTCTTTGTATATAAAATAAACATTTGAAAAAAA
15 AAAAAAA

Sequence ID - 593

nt: 565

CAGGATCAAGGTGAAAAGGAGAACCCCATGCGGGAACTTCGCATCCGCAAACCTCTG
TCTCAACATCTGTGTTGGGGAGAGTGGAGACAGACTGACGCGAGCAGCCAAGGTGT
20 TGGAGCAGCTCACAGGGCAGACCCCTGTGTTTTCCAAAGCTAGATACACTGTCAGA
TCCTTTGGCATCCGGAGAAATGAAAAGATTGCTGTCCACTGCACAGTTCGAGGGGC
CAAGGCAGAAGAAATCTTGGAGAAGGGTCTAAAGGTGCGGGAGTATGAGTTAAGAA
AAAACAACCTTCTCAGATACTGGAACTTTGGTTTTTGGGATCCAGGAACACATCGAT
CTGGGTATCAAATATGACCCAAGCATTTGGTATCTACGGCCTGGACTTCTATGTGGT
25 GCTGGGTAGGCCAGGTTTCAGCATCGCAGACAAGAAGCGCAGGACAGGCTGCATTG
GGGCCAAACACAGAATCAGCAAAGAGGAGGCCATGCGCTGGTTCCAGCAGAAGTAT
GATGGGATCATCCTTCCTGGCAAATAAATTCCCGTTTCTATCCAAAAGAGCAATAA
AAAGT

30 Sequence ID 594

CAGAAGAGTAAGCAAATCTCAAAGCAGCGAAAGGGAAGAACTAAAAAAGGTAGAG
CAGAAATAAGAGAAAATAGAGAAGAGAACAAATTGAGAAAAATAATTGAAACCAAAA
GGTGGTTCTTTGAAAAGCCTAACAAAATGGACACATCTTTAGTTAGAGTGACCAAG
AAAAAAGGGCAGTGACTCAGATTACTTCATTCAAGAGTGAAAGAGGGCACATCACT
35 ACCAATTTACAGAAATAAAAAGGATTATGAGGAAATACTACAGATAATTGATGACA
TTAACTTAGAAGAATATATTTCAAGAAAGACACAACTACTGAAACCGACTCAAGA
AGAAACAGAAAATCTGAACAGACCTATAAAAAAATAGAGATTTAATTGATATTCAGA

- 184 -

AAGTTTCCCAAAAAGAAAAGCACTGGCCAAGATGACTTCACTGGTGAATTCTATCA
AGTGTCAAAGATGAATTACTGACATTCATTCACACTCCTTTAAGAAATAGAAGAGG
GGACATCACTTTTCAAAGCATCGACATTCTAATCATTAGTCCCTTGGTTTCCTGCT
CCCAAAGCCAGGTGATGTATCACAAAAAACCCCTACAGACCCACTGGGCACAATG
5 GCTTTATGCCTAT

Sequence ID - 595 nt: 98

CTTTGCTCGAATNGTCAGATAAGGATTCTGTGAANGGAGATGAGATTTCCATCCAT
GCTGACTTTGANAATACATGTTCCCGAATTGGGGNCCCCAAA

10

Sequence ID 596

CTCAAGTGTTCCCTCAGCTTAGGCTTTGTTTAAATGATCCCACCCAGGGGCGATGG
TAGGGAACAACAGGGTCACTAAACTATTTGGCTGGCTACAACCTCTGGGAAATGGTA
AGACAGGGAAAGGCCATGTTGTTCAATCCCTTGTGCAGATCTAGGGAGAACCGCAG
15 AGAGAACAGTTAGCATTTCTTGTTCAATGAATTATCCTATTAAGAACACTGGATGT

Sequence ID 597

CGGNCGCGGTCGACGCTACTCCTACCTATCTCCCCTTTTATACTAATAATCTTATA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

20

Sequence ID - 598 nt: 362

GGCATGTGCCTGTAGTCCTAGTTGCTGAGGTAAGAGGATTGCTTGAGCCCAAGAGT
TCAAGGCTGCAACAAGCTTTGATTGCGCCACTGCACTCCANCCTTGGCGACAGACT
AAAACGCTGTCTCAAAAAAAAAAACAAAAACGACNAAAAAAAAACAAAACAGAAAAA
25 ATTAACCTTAGGCAATGACAGTCCCTGGCAAATGCTGGGAGGGAGGCAACANTGGTC
AAGGAAGGTAACCCTGAANCAGGACTTGTAAGCAAATAANATTGGGAGGCCAAGG
TGGGTGGATCACNAGGTCAGGAGTTCGAGACCAACCTGGCCAACATAGTGAAACCC
CGTCTTTCTAAAAATACAAAAAAATT

30 Sequence ID 599

GACAAAAGAACCATTTGGATACATAGGTATGGTCTGAGCTATGATATCAATTGGCT
TCCTAGGGTTTATCGTGTGAGCACACCATATATTTACAGTAGGAATAGACGTAGAC
ACACGAGCATATTTACCTCCGCTACCATAATCATCGCTATCCCCACCGGCGTCAA
AGTATTTAGCTGACTCGCCACACTCCACGGAAGCAATATGAAATGATCTGCTGCAG
35 TGCTCTGAGCCCTAGGATTCATCTTTCTTTTCACCGTAGGTGGCCTGACTGGCATT
GTATTAGCAAACCTCATCACTAGACATCGTACTACACGACACGTACTACGTTGTAGC
TCACTTCCACTATGTCCTATCAATAGGAGCTGTATTTGCCATCATAGGAGGCTTCA

- 185 -

TTCACTGATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCTACGCCAAAATC
CATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCACAACACTTTCT
CGGCCTGTCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACACCACAT
GAAACATCCTATCATCTGGAG'

5

Sequence ID - 600

nt: 595

TTCAAATTCTTGNTAANAGTCTTTGTTCTGAATTTTACTTTGTCTGTTATTCCTAT
AGCCTTTCCAATTTTCTTTCGCTTGGAATTTTACGTGATAAGTTTTTCCCCCATTT
TACTTTTANCAACTCTATATTTTTTTAGTTGAGGTTGGGTTTCTTGTAACAGCATA
10 TAATTTGGGTTTTTTAATCCAATCTGAAAATTAATGTCCTTAATTTTGTGTTTATA
CCATTTACACATAATGTACTCATATATAAGGTTTAACTGAAACCTACTATCTTGCT
AGTTGTGCTCTACTTGAATTTTTTTTTTAGTATTCTGTTTTAATTGACCAACATTTG
ACTGTATCTCTTTGTGTAATTCTTTTACAGGTTGCTGTAGGCATGACAATATATAC
ACTTAACTTTTCTCAGTACACTGAGAGTTGAAATTGTAGTACTTCGAGGAAAACAT
15 AGAAAACCTTGCAATGATATCGGTTACATTTTACCACCTCCATATGTTGCAATTATT
AAATGTATTAGATCTGCCTACCTCGAAAACCCATCAGTCTTTTAACTTTGCTCTCA
ATGGTGATTCATATTTTTTAAAAAACTTGAGGCAA

Sequence ID - 601

nt: 522

20 TCGACCGGGTTTGGAGCAGTGCCTTGTTTGCTGTGCAGCGGATACTCTACAGGTAC
ATTTCTTTTTTGGAAACCAAAGGGAGGGATTGACAATATTGATGGTAGATCTTTT
TTCTTTAGCAAGAATTAAGGATTTTGGTGGGTGGGGGGAGGCTTCTGTGGGGACCA
AGACAATGTACTGTCAGTCAGGATTTAAGTCGAACTACCTCATCCCTTGCCCCAGA
GAACAGTTGATCGTGTTTTAAACCAAAGGTGCGGAATGGAGAGAGGGAGGCGGTG
25 CATTGCAGCTTCCGATAGAGCTTTTTTATTTTTTGGATATCAGGAACCAATTTTGAAG
ATTTCTTAAGAAAGTCATTTACATCAGGGACATGAAGAGCAAAGTAGGTATTTTTG
GTCAGTACTTGAATTTGATAGGCTTTATGCAAACAACTCTCCCTCTGCTGGAGTCT
GGCAAGTTTGCTTTTCACTGGACGCTAATTCAAGTGCCATACAAAACCTAAAATAAN
AGTTTTACTTATAACACA

30

Sequence ID 602

CAGAAATCGCAATTGAAGACCAGATTTGTCAAGGTTTGAAACTGACATTTGATACT
ACCTTCTCACCAAACACAGGAAAGAAAAGTGGTAAAATCAAGTCTTCTTACAAGAG
GGAGTGTATAAACCTTGGTTGTGATGTTGACTTTGATTTTGCTGGACCTGCAATCC
35 ATGGTTCAGCTGTCTTTGGTTATGAGGGCTGGCTTGCTGGCTACCAGATGACCTTT
GACAGTGCCAAATCAAAGCTGACAAGGAATAACTTTGCAGTGGGCTACAGGACTGG
GGACTTCCAGCTACACACTAATGTCAATGATGGGACAGAATTTGGAGGATCAATTT

- 186 -

ATCAGAAAGTTTGTGAAGATCTTGACACTTCAGTAAACCTTGCTTGGACATCAGGT
ACCAACTGCACTCGTTTTTGGCATTGCAGCTAAATATCAGTTGGATCCCAGTCTTC
CATTCTGCAAAAGTCAACAACCTCTAGCTTAATTGGAGTAGGCTATACTCAGACTC
TGAGGCCTGGTGTGAAGCTTACACTCTCTGCTCTGGTAGATGGGAAGAGCATTAAAT
5 GCTGGAGGCCACAAGGTTGGGCTCG

Sequence ID - 603

nt: 624

GACACACGAGCATATTTTACCTCCGCTACCATAATCATCGCTATCCCCACCGGCGT
CAAAGTATTTAGCTGACTCGCCACACTCCACGGAAGCAATATGAAATGATCTGCTG
10 CAGTGCTCTGAGCCCTAGGATTCATCTTTCTTTTACCGTAGGTGGCCTGACTGGC
ATTGTATTAGCAAACCTCATCACTAGACATCGTACTACACGACACGTACTACGTTGT
AGCCCACTTCCACTATGTCCTATCAATAGGAGCTGTATTTGCCATCATAGGAGGCT
TCATTCAGTATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCTACGCCAAA
ATCCATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCACAACACTT
15 TCTCGGCCTATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACACCA
CATGAAACATCCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAGTAATATTA
ATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAAGCGAAAAGTCCTAATAGTAGA
AGAACCCTCCATAAACCTGGAGTGACTATATGGATGCCCCCCACCCTACCACACAT
TCGAAGAA

20

Sequence ID - 605

nt: 338

ACCTGAGGCCTCGGTGGGGCCAGTGCGACGCTGGCTTAAGGAGCTGGAGGGGTTC
TAATACACATTTAATTCAGTTTCTCTTCCCTAAGAGGCTGCCGGAGTTGGGGCCTC
CTCCAGCAGAGACCCTCGGACCCCTGCAGGGCCTGGACTTGGGGTGAACAGGGCTT
25 CAGTCAGCGCAAGTATTCCATTTGCATTTGGTAATTTTTCATGCCACCTATTTATG
AATATATAAATCTTTATACCAAATCTATTTTTTAAAACATGGAAAAGTTGCCTTTA
TGGAACCTTGGCAGAGCCAGAGTGTACACATTCCTAAACCATTAAACAGATTTCTA
TA

30

Sequence ID - 606

nt: 556

GGATAATGATACCTCTGACCTTTCTTCCTTTTGGGAAGTACTTGAGTGTGCAGCTG
CATGAGGCCTCAGCAGGAGAGAGATTTTAGGTCCAAGAAGCTATACCAGTAGGACA
AGGCAGGAAAATACTACACTTTCAGGATCAAGCCCCTCTGACTCTCATTTGGAAAC
TGGATGTTTGCTAAGCACCTGCTTCTTAAGGATGCCGAGGGATTTAATGATACTCC
35 CAGAAACCTGGAGAGATTAATGGGGCCTATGGAGAAGTGCTCTGAACTCAGTGTTG
GGACTTGAATAAAATTAACCATTTGTCATGTTTTTCAGAACAACCTAAGCTGTTTTATA
TTTCATGTGCATGAAAGCCCTAGAACTAAGTTGTGTTATTTCCAGAAATGAAATAG

- 187 -

ATCCCACAGTTAGATGATGTGGCCATTAGGAAGTACCAAATTTATAAAAATCACTG
GAGGTCTGTCTGAGCAGTACCTAATAAAATATAGTATACTGAAAGTGAACAGATCT
TTGTCTCTTTCTTTGGCTGCTTGATACTTTATCTGTGTCTGCCGGACAGTGC

5 Sequence ID 607

CAATAAAAGCAGGTTAACCTCAATGATAGCAGTTAAAATGTTCTATCTTATGTATT
TCTTTTAAAGTATTACCATTTATGGTGCTACTGAGCGTTTTCTTTTGGTAAAAAGAAA
AATGCCATGGGCTGCAGTCTTCTTCCATCACTTTTCCCTACCAGGTCCATTAATAT
GCTTATAACACTAGTGCCAGTTATTTTATTTGATAATGCTTATGGTATTTGTATAT
10 TTGTTTGCATTCCAATTTTGTTTAATAATGAGTGTGTAAACTGCATACGTAAATA
AATGTAAATACTAATGTACTGCTGC

Sequence ID 609

TTTTATTACCCAAGTTTTAACCTCTGTCTGGTGATTTGTTGTTGTTGTTGTTGTNG
15 TTGTTGTTGAAGTTCAGGCTGCATGTGGGATAGGTTTGCTCAGGCATACTTCTTAG
GAAGTAGTCACTTGCATGACTGTTTTTGGGATAACTCTTTGAGTATTTGGAGAGGT
CTATTGTAACCTTCTGAAAGGCATTGTTTTTACGTATGAATGTTCTAAAATTCATTC
TAAATGGTCATGAAAAGAAAAGGATTCACATTTTAGAATGGCAATAGTCCCTGAGG
ACTATTATGTCTTTTAGATTTCTGTGGGTTTCTAGGAATGTTAGTGTAACCTTANA
20 TTCCACCTACCTGATTTCTGGATGTGCCTATTGGAACCTTGCTGAGATCTTTTTTTT
TTCCTTAACATGTTGTCCCCTTGACCCGTACTTCGAAACTAAACATATTATTTTAT
TTGCTTACACTTCAGGAGGCAATTGGCAGACACCAGGCCAACAGTCT

Sequence ID 610

25 GCTCTGACCCCAGTTGGAAATGTATCTGTACTTTGTCCGGCTTCCACTCAAGGACC
ATTTATGACATTGCTTGGTGTCAGCTGACAGGGGCTCTGGCCACAGCTTGTGGGGA
TGACGCGATCCGCGTGTTTCAGGAGGATCCCAACTCGGATCCACAGCAGCCCACCT
TCTCCCTGACAGCCCCTTGCATCAGGCCCATTTCCAGGATGTCAACTGTGTGGCC
TGGAACCCCAAGGAGCCAGGGCTACTGGCCTCCTGCAGTGATGATGGGGAGGTGGC
30 CTTCTGGAAGTATCAGCGGCCTGAAGGCCTCTGAGCTACCTCGACTTTGGACAGAG
TAATGACTCCCCAGAAAACGTCATATAAGACTTTACCAGCCCCTGAGAGGACCAGG
AGGAGCATCCTTGACCTTCATTTAACTTGGCTCACTTCTCTTCANACTTGGGTAGA
AGTGCAGAGCCACAAAATTGCTTTCCTTCCCCGCCTTTGACATGAGGCCTTCAGTA
AAG

35

Sequence ID 611

TGCAGGATCCGTCGACT

- 188 -

Sequence ID - 612

nt: 576

5 GAGAAATATAAGATTATGTATAGATCAAATCTACCTCTATTTGGTGTCTGAAAGA
GATGAGGAGAATGGGACAAACTTGGAAAGCTTATTTCAAGATAACATTCCTGAGAA
CTTCCCCAATCTTGCTAGAGAGGCCAACATTAAAATTCAGTAAATGCTGAAAACCTC
CAGTAAGATATTTCTTAAGAAAATTATTTCCAAGATATATACTCATCAAATTATCT
AAGGTCAAATGAAGGAAAAAATTTTATAGGCAGCTAGAGAGAAATGTCAGGTCACC
TACAAAGAGAATGGCATAAGACAAAAAGTAGAACTCCCAGCAGAAACTCTAAAAGC
CAGAAGAGATTAGGGGCCAATATTTAACATTCTGAAAGAAATTCCAACAAGGAATT
TCATATCCAGCCAAACTAAGCTTCATAATTGAAGGAGAAATAAGATATTTTCCAGA
10 CAAGCAAATGCTGATGAAATCCATCACCACCAGACCTGCCTTATAAGAGCTCCTGA
GGGAAGCACTAAATATTGAAAGGGAAGAACTTTATGAACCATTTCAAAAACACATT
TAAGTNCACAAAGCAG

Sequence ID - 613

nt: 341

15 CCTTATTTTACAGGTGAAAAACCACGAATCAGATAGATTTTATTTGCCCCAAGTCA
CATAATATTAAGAACAGGCCAAGTGTGGTGGCTCATGTCTGTAATCTGAGCACTTT
GGGAGGCTAAGGCGGGTGGATTTCTCTGAGCCTAGGAGTTTGAGATCAGCCTGGGCA
ACATGGCGAAACCTCATCTCTACAAAACATACAAAAATTAGTCAGTGTGGTGGTGA
GAGCCTGTAGTCCTGGCTACTCGTGAGGCTGAGGTGGGAGCATCACCTGAGCCTGG
20 GAAGTCGAGGCTGCAGTGGCAACAGAAATGGGTAACCTGGACATCAGAGTGAGACCC
TGTCT

Sequence ID 614

25 CTCACACCTGTAATTCCATTACTTTGGAAGGCTGAGAGAGGAGGATCAGTGGAGCC
CAGGAGTTTGAGACCAGCCTGGGCAATATAGGGAGACCCTGTCTCTACAAAAATGA
AATAGCCAGGCGAGGTGGCATGTGCCTGTGGTCCCAGCTACTTGGGAGACTGAGGT
GGAAGGCTGCCTTGAGCCCAGGAGTTCCAGGCTGCAGTGAGCCATCATTATGCCAC
TGCACTCCAACCTGGGAGACAGAGTGAGAGAGACCCTGTCTCAAACAAACAAACCC
AAAATAGGCCAGGCACAGTGACTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGA
30 AATAGGCGGATCATTTGAGGTCAGGAGTTCAAATTCAAGACCAGCCCGGCCAACAT
GGCAAACACATCTCTACTACAAATAAAAAATTAGTTGGGTGTGGNGGAGCATTC
CTGTAATCACAGCTATTCAGGAGGCTGAGGCATGANAACCGCTTCA

Sequence ID - 615

nt: 379

35 TAAATTTAAAACATTTTAAATTAGCTGGCATGATGGCATGCACCTGTAGTCCTACCT
ACTTGGGAGGCCAAGGCAGGAAGATTGCTTGAGCCCAGGAGTTTGAGCTTACTGTG
AGCTGTGATCACACCACTGCACTCCAGCCTGGGTGACAAAGGAAGACCGTATTTCT

- 189 -

AAAAAATAAAAAATACAAATACAACACTACAACACTAGCACTAGACCAACAGTGACTAT
GTACCATGAACTGAGGAATATTATTAATTCCACCATTTGCATCTGAGGTTAACAAT
ATGTCAATGACTTAAATAACATCATATCTCTGAGAGTAATTTCTCCTATATTTCCA
TGACAAATGTTAGATAATTTTCCATTTTTCATTTCAACAAA

5

Sequence ID 617

TTTTCAGGCATGTCAGAGAAGGGAGGACTCACTAGAATTAGCAAACAAAACCACCC
TGACATCCTCCTTCAGGAACACGGGGAGCAGAGGCCAAAGCACTAAGGGGAGGGCG
CATACCCGAGACGATTGTATGAAGAAAATATGGAGGAACGTGTACATGTTCCGTAC
10 TAAGTCATTTTTCAGGGGATTGAAAGACTATTGCTGGATTTCATGATGCTGACTGGC
GTTAGCTGATTAACCCATGTAAATAGGCACTTAAATAGAAGCAGGAAAGGGAGACA
AAGACTGGCTTCTGGACTTCCTCCCTGATCCCCACTCTTACTCATCACCTGCAGTG
GCCAGAATTAGGGACTCAGAATCAAACCAGTGTAAGGCAGTGCTGGCTGCCATTGC
CTGGTCACATTGAAATTGGTGGCTTCATT

15

Sequence ID - 618

nt: 598

GATTAACTTTCATTTTAAGCTCTTCTCTACTAATTCTGTTTCGTATGTTTATTCATT
TTGCGTTGATCATATTTTGTACACCAGGCACTCTTCTCAGTTTTATATGTGTGTTA
ATTTACTCCTTTCAAGAGCCCTATGATACATGAATTTATCTCCATTTTATAGATGA
20 GGAAATTAAGACCTAGAGTTACTGAACTTGCCCAAGGTTATACAGCTGATGGGTAG
GGCCAGAACTTTGCCTCAGAGAATCTGAATTTCCAAAAATAACCTAAAAGAGAAA
TTTAAGTACTAATTAGTAAGCAAAGAAATGCACATTTAAGGAAGACAGTGCACATT
TAAGGAAGACAGTAACCTTTTATCTATTAGAGAAAAACACACATTCTGTCTTTAAC
ACACACATAAATCTTATATTGGCAGGGATTTTCTTTATTTCAGCAATTATTTATTGG
25 TTGTCTGCTTTGTGGTACACATAAATGCTGGGGATAAACACTTAATAAAATATACT
TCCTTCTCTTGAATATCTTGCACTTTAAGTGGAAGGTAAGTCAACAGAGTAGAGG
TGATATATCCAAGTGATAGACTGTTTCATTGCCAGTAG

Sequence ID 619

GTTGCCTGAGAGTGACCTTTGCATCTGCCTGTCCAGCCAGCATGGAACCAAAGCGG
ATCAGAGAGGGCTACCTTGTGAAGAAGGGGAGCGTGTTCAATACGTGGAAACCCAT
GTGGGTTGTATTGTTAGAAGATGGAATTGAATTCTATAAGAAGAAAAGTGACAACA
GCCCCAAAGGAATGATCCCGCTGAAAGGGAGCACTCTGACTAGCCCTTGTCAAGAC
TTTGGCAAAAGGATGTTTGTGTTTAAGATCACTATGACCAAACAGCAGGACCACTT
35 CTTCCAGGCAGCCTTCCTGGAGGAGAGAGATGCCTGGGTTCGGGATATCAATAAGG
CCATTAAATGCATTGAAGGAGGCCAGAAATTTGCCAGGAAATCTACCAGGAGGTCC
ATTCGACTGCCAGAAACCATTGACTTAGGTGCCTTATATTTGTCCATGAAAGACAC

- 190 -

TGAAAAAGGAATAAAAGAACTGAAT

Sequence ID 621

5 TGGTACTGAACCTACGAGTACACCGACTACGGCGGACTAATCTTCAACTCCTACAT
ACTTCCCCCATTATTCTAGAACCGGCGACCTGCGACTCCTTGACGTTGACAATC
GAGTAGTACTCCCGATTGAAGCCCCCATTCGTATAATAATTACATCACAAGACGTC
TTGCACTCATGAGCTGTCCCCACATTAGGCTTAAAAACAGATGCAATTCCCGGACG
TCTAAACCAAACCACTTTCACCGCTACACGACCGGGGGTATACTACGGTCAATGCT
CTGAAATCTGTGGAGCAAACCACAGTTTCATGCCCATCGTCCTAGAATTAATTCCC
10 CTAAAAATCTTTGAAATAGGGCCCGTATTTACCTATAGCACCCCCTCTACCCCCT

Sequence ID 622

TTTTTCTTGTTTTTGTGTGTCTACCTTGGCATATACTAAAGGAAGGTGTGTATTCA
TTTATTACATGATATCTCTGGGTATAATTATTTACATATATGAATTTGAAAGAAA
15 GATTGAGAGGGATATGTGTGACCTTTGTTTCATTATGATCATTTACATGACTAAAG
ATAAAGATCATATGTCTGATTTTCAGTTTAATGGCAAGTTACTTAAAATAAATGAA
ATATGTTTTTTATTGTTTTTCGTGGGTTTGATGCTTTGTGTTTTATTTCAAGTAACTT
GAGAATGCATTGTGTTTTGGTACTGTTTTTTATGAATATCATTAATAAATTTATTAA
GGAGAGAGTAATTTTGCAATAATATTTTTGATTTATTTGAAAATAAATTCAGAT
20 AAATGAAATAATTGAAATTTTCTAAAGAAGGAATTGAATATATTTTTTACATTTGAA
TGAATAAGGATTAACCTGAACCATTTATATATAGTACTTTCAGAACTGAATGTCTT
AAATGATAAAGCTCTAATTGGTTAAAGTGACTTTCTTTCAAGTCAAAGAACCCAGA
AACTGAATAGATGATCTAACTACTGCCACTGAGGTTTTGGATTAGTGAGTATAAAT
TT

25

Sequence ID 624

TGCAGGATCCGTCGACT

Sequence ID 625

30 GACAATCAGAGCAGATCTTGGGCTTCTGTGGCTCATCTCAGCCCTTTATAACTGGC
CTGAGAAGAGGGTTTATCTACTTGTGCAAGTGGCCCAGAAATCTCACTCGTACATG
AGGCTTTGGAACATCCTTGCAAAGGTACGCTGAAAGCAAATTGCTGTTTTCTGGT
GGTTCTGCACGTTTCCTAACTTTTATCATAGTTTGATTTTCATTATTTAAGAAAAA
ATAAAAAATCCAAGACCATAAGATGGCATTAGATTTTTTACCATTAAATTATTAA
35 TGCCTATTTGGTGCTCATAAAGATTAATCATGTCACGCATGTTTCCAATCTTTCTT
TTGCAGTATATTATTTTCTAAAAATTGTTACATGCAAATTTAAACCAAGATTTATC
AGTA

- 191 -

Sequence ID 626

TTGGAAGAAATAAACCAAGGCAGAAAAATTTTAAATGGCCAAAATAAATTGTATTG
CTAACTTAGATGGCCACAGATGGGGGCAGGGGTGGAGAGAGGAGAAATTGAAAACN
CCACAAAGACCCCGCAATGGCTAGA^ΔCTTGAAATCTCTGGATATTGCAACAATAGC
5 AGCCTCCTTAAGTCAGCAAAAAGATAAAGATTGATCCAATGTTCTATATTACAGAA
CAGAGCAGATTGTCAATATAGCAAATAAAGTTACCGTTGAGTGGACTGCGCTGTNT
AAGCTGCTTGGTTGGCCTTAAGTGCCGACAATTAAGAGATGAAGGCAATGAGAACT
GAAACAAACATTTAAGTTCAAGACCCAGTTTACTGACACTGGGACTATTACTATAT
CTCTTTGGGCCTCAGTTTACTTATCTGTAACATTAAGAGGTTGGATTACATGATGT
10 CTCACGATTCTTTTTTTTTTATTTAGAGATGGGGTTTTTGCTCTGTTGCCCAGGCTGG
AGTGCAGTGGCATGATCATAGCTCACAGCAG

Sequence ID 627

CCAGCCTGTCACTGGCCTGGCCAAGGAGGAGAGACAGGCCAGGGATTCTGGTCCTA
15 ACTCTACTGGCCACACTGTGTGGCCTGAGACCCCCCTTTCCCTCCCAAGCCCCTGC
CTCCGCATCTGCGTGGTGAAGGCCATTGGCCCTCATCGGTGGATCTGCGTTTCCTC
GGGCCTACACTGTCTAGGATTGTGCGGGGCTGGTGAGAGAACAAGATCTCTTCCGT
GTTCAAGGCAGACTTCCTGCCCCCTGCACCCTGCTCTCTCCAGGCCTTGAGGTCA
GTGTGAGCCCCAAGGGCAAGAACACTTCTGGAAGGGAGAGTGGATTTGGCTGGGCC
20 ATCTGGATGGAAGGTAAAAAAAAGAAAATCCCTTGAAAGGAGATTGAGGGAAGTTT

Sequence ID - 628

nt: 419

AAGAGAAAGGACTCAGTGTGTGATCCGGTTTCTTTTTTGCTCGCCCCCTGTTTTTTGT
AGAATCTCTTCATGCTTGACATACCTACCAGTATTATTCCCGACGACACATATACA
25 TATGAGAATATACCTTATTTATTTTTGTGTAGGTGTCTGCCTTCACAAATGTCATT
GTCTACTCCTAGAAGAACCAAATACCTCAATTTTTTGTTTTTGAGTACTGTACTATC
CTGTAAATATATCTTAAGCAGGTTTGTTTTTCAGCACTGATGGAAAATACCAGTGTT
GGGTTTTTTTTTTAGTTGCCAACAGTTGTATGTTTGCTGATTATTTATGACCTGAAA
TAATATATTTCTTCTTCTAAGAAGACATTTTGTTACATAAGGATGACTTTTTTTATA
30 CAATGGGAATAAATTATGGCATTTTTTT

Sequence ID 629

CTGAGAGTCACTGTGTTTTTAGCCAAATCTAAGGGAGAAAATGAATATTGATAGCA
GCATGCTGTAGCCAGCTCCTTAAAGGAAGGATGGTGCCTGGTACAGAGTTAGAGTT
35 AGTGCTTCAGTAAATAATGAATGTGTGCTAGGTAGGTTCTGCTGGGTAGGCTGCAT
GCATTGACCAATTTATTCCTCCTTGTTTCAAAACAGGATTTAAGGGCACTTATATA
TATATATTTTTTTAGTTTTTTTTAATGTAAATGAGAGAATAAAGATATATATATATGT

- 192 -

CTATATATGTATATATGTATATATATGTCTATATGTCTATATGTATATATGTCTAT
ATGTATATATGTGTGTGTGTATATATATATATATATATAAGTTTTCTGTTGCTA
GCATAACAACTACCAGAACTTAGCAACTGAAACAACATGAATTTATCTTACGGT
TCTATAGTTCAGAAGTCTAACGTGTCACCTGGGATGAAATCCAGGTTTCAACAGGAC
5 TGGGTTCCCTTCTAGCTCATTACAGCTACCTGGCTCATTACAGGTTGTNGGCAGAATA
TACTTCCATGAACTGTAGGGCTGAGACCCCGTTCCTTCCTGGCTATCATCTGAAA
ACTTTC

Sequence ID 630

10 AGGCGCAGCCCAGCCTCGAAATGCAGAACGACGCCGGCGAGTTCGTGGACCTGTAC
GTGCCGCGGAAATGCTCCGCTAGCAATCGCATCATCGGTGCCAAGGACCACGCATC
CATCCAGATGAACGTGGCCGAGGTTGACAAGGTCACAGGCAGGTTTAATGGCCAGT
TTAAAACTTATGCTATCTGCGGGGCCATTCGTAGGATGGGTGAGTCAGATGATTCC
ATTCTCCGATTGGCCAAGGCCGATGGCATCGTCTCAAAGAACTTTTGACTGGAGAG
15 AATCACAGATGTGGAATATTTGTCATAAATAAATAATGAAAACCTAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

Sequence ID 631

TNCACTCACACACTCCCAAACCTTAACAAACACATACATGTGCAGCCAACCCAATG
20 GGCCAGCCTCTTTTATGCTCCTCACATGTTTCCTTTAACTGGAATACCCATGACAG
CTCCCTACATAGTTACTTGTAACCTCCTCCTCTCTGTATAAGTTTTCTGAATTTT
TTTGATAAAATTAAGTTGTGCCACCCCTTTATGCTCTCTTANAACCTTTGTTCTGTT
CTCATGGCTGTTCTGCAACGAATCTCATTGTGTTCTCCTACTCAATTACATTCCTG
CGTCTCCCACTAGATGGCAGACTCTTTGAGAGTAGGAGATTCCCTTGTTATCTCTG
25 GATCCCTGGCACTTGCAGAAAGCCTGTTACGTAATAATTGCTCAACAATTAGTTTT
TAAATAAATGAATTATTTTTTAAACGCCAAAATTACAATGATTGTGCATTAAGTGA
AAGATGACCATCTAAAAACATAAAGCCATGCTTCATGACATTGGC

Sequence ID 632

30 GACCATTACAGGGAAATTTTATAAAAAATGCAGATACTGTCTTGAGCAGATCGAAAT
GCCGATGAGGTGGATGCAATTTCTTTTGTGCAAGCAGTGCACGGTGCCCCCCCCCT
CGGGTGTCGGTGCTGTGCCTTAGCTTCCCCAGGTGCCGGGACTCACACCTGCTAGG
GGCTGGGCAAGGCCCGGCTCTGCTTTCTCTGAAGGGCTTGTCCAAGTTCATTGCC
CTGTTACAGGTGGTCAAGACGTCCGGCCGCCTTGACCCAGGCTACCCTTAGCCAAT
35 ATCCTCTGCCCCCTGGGTGGTGGTGGCTGGGCCTCAGGGTGGGCAACGTTAGGGGT
TTGGCGAAAGCCCCGCCCATGGGATTGAGGGACGGGGCTGCACTCCAACCGTCTGC
ACCTGCTCTTCCCCCACCCCTGTGGGACCTCATCTTCACGTGCCATGTGTGCTGAA

- 193 -

GGCCCAGGGCCCAGCAGGGGGCAGTGGCACCTGTTGACGGAAAAGCCGAGGTGCTT
ACCAATGGACCTTCTGGCCCGCCCTCCCCTGTACTTGTCGGGCATTCAGGGCCCCG
ACCTGTGCCTACCCGCA

5 Sequence ID 633

CAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCTGACTCCTGAGG
AGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGT
GAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTC
CTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTC
10 ATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTC
AAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCC
TGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTG
GCAAAGAATTCACCCCACCAGTGCAGGCTGCCTATCANAAAGTGGTGGCTGGTGTG
GGCTAATGCCTGGCCCCACAAGTATCACTAAGCTCGCTTTCTTGCTGTCCAATTTT
15 TATTAAAGGTTCTTTGTTCCCTAAGTCCAACCTACTAACTGGGGGATATTATGAA
GGGCCTTG

Sequence ID - 634

nt: 511

TTTTTTAATTTACCAAATTTGTTGACGTCCTTGATTTGCTGATAGGGACAATA
20 ATTAAATATTTTCCACTTGTTTTTATAAAAAGTGAATGGTGATTTGTTTAACAGA
TGTTGACTTAGCACCTTCTCTCTTTTTTTTTTTTTTTTTTTTGGAGTTGGAGTCTTGC
TCTGTCACCCAGCTGGAGTGCAGTGGCACGATTTCCGGCTCACTGCAACCTCCGCCT
CCCAGGTTCCGGGCGCTTCTCCTGCCTCAGCCTCCCANATAGTTGGGATTACAGGTG
CATGCCGCCACNCCTAGCTAATGTTTTTTGTATCTTGGTANANATGGNGTTTCACC
25 TTGTTGCCCATGCCGCTCTTGAACCTCCTTGGCCTCCCAAAGTGTTAGGATTACAGG
CGTGAGCCACTGTGCCTGGCCCCAATTTANCACCTTACTGGGTGCTGAGGCTGTGA
GCCATAGTAGAATGCATGTGATCCAGGGCCTTGCTGAATTCATGGGCTAATAGGGA
GCCTGAC

30 Sequence ID - 635

nt: 592

TGAGCGTTGGGCTGTAGGTCGCTGTGCTGTGTGATCCCCCAGAGCCATGCCCGAGA
TAGTGGATACCTGTTTCGTTGGCCTCTCCGGCTTCCGTCTGCCGGACCAAGCACCTG
CACCTGCGCTGCAGCGTCGACTTTACTCGCCGGACGCTGACCGGGACTGCTGCTCT
CACGGTCCAGTCTCAGGAGGACAATCTGCGCAGCCTGGTTTTTGGATACAAAGGACC
35 TTACAATAGAAAAAGTAGTGATCAATGGACAAGAAGTCAAATATGCTCTTGGAGAA
AGACAAAGTTACAAGGGATCGCCAATGGAAATCTCTCTTCTATCGCTTTGAGCAA
AAATCAAGAAATTGTTATAGAAATTTCTTTTGAGACCTCTCCAAAATCTTCTGCTC

- 194 -

TCCAGTGGCTCACTCCTGAACAGACTTCTGGGAAGGAACACCCATATCTCTTTAGT
CAGTGCCAGGCCATCCACTGCAGAGCAATCCTTCCTTGTCAGGACACTCCTTCTGN
GAAATTAACCTATACTGCAGAGGTGTCTGTCCCTAAAGAACTGGTGGCACTTATGA
GTGCTATTCGTGATGGAGAAACACCTGACCCA

5

Sequence ID - 636

nt: 572

CTTANAAGAGTTGCTCATTACACCCACGCCCTTGCCCAAGGCTGGCCCACTCAGA
GCGAAACTTAACTTTTGTCTGGATGGGAAGAGAAGTAAGTCTACCCCGAGGTTGCC
ATGTTGAAGAGTGAGAGGTCCAAGTGATTCTGTGCATTGAAACCAAGACACCCAC
10 CCAGAACACTTCTTCCCTCCCTCAGCCCAAACCAAAGGCTGGGGTTCTCATCTCCA
AGTGGCTGTTCTCCAACCTTCCCAAGCCGCTTGCAATCCCCAGACTGGACTACTGT
GGCGGTTAGGTTAGATTTGAAGACGGGGCCCAGGCTGGGTATGAACGGGTGCAGCC
CTCTTCTCCTCTTCCCCCCCCACATCTCTCATGAGAGAGGTAGTGGCATTTCTTCT
CAGGGAGCTTCAATGGGAAGGTCTCGAAAGCTTCAGGAGGAGCAGAATACCAACG
15 CAGGGGGATGGCTGTAACGATCTCACCGTCTCCTAACCTCAGTCCCTTTTTTGAGA
GTGAATGGTGGAGGGTGGGAAGGGACCCAAATTTGTAGATCTCTTTGTCTGGGGG
AGGGGAANGATG

Sequence ID - 637

nt: 482

20 TTAAACAGGCGCAGGGGTAAAAATGAGAATGAATCTGAAAAAGAGAGTTGGTGT
TTAAAGAGGATGGACAAGAGTATGCTCAGGTAATCAAAATGTTGGGAATGGACGA
TTGGAAGCATTGTGTTTTGATGGTGTAAGAGGTTATGCCATATCAGAGGGAAATT
GAGAAAAAGGTTTGGATAAATACATCAGACATTATATTGGTTGGTCTACGGGACT
ATCAGGATAACAAAGCTGATGTAATTTTAAAGTACAATGCAGATGAAGCTAGAAGC
25 CTGAAGGCATATGGCGAGCTTCCAGAACATGCTAAAATCAATGAAACAGACACATT
TGGTCCTGGAGATGATGATGAAATCCAGTTTGACGATATTGGAGATGATGATGAAG
ACATTGATGATATCTAAATTGAACCAAGTGTTTTTACATGACAAGTTCTCTGAGGA
TGGTTCTACAGTTGGGATTTTGGCCATCATCAAC

30 Sequence ID - 638

nt: 545

TTTGAAGGCAAAGAGGGATTAATCTGTGCTGGCATCATGTAAGGAGACTTGATAGA
TAAGAAAAAGCTTTACCTAAGTTTTGAAGAATAGGTTTTTCATAATGGAAAATTTA
AGGGAAAAATCTCCAAAAAAGTGCTACTCAAGTTTTATCCATTTGTATTTCCAACA
CAGCCTAGGACAGTACCTGCACATAGTAGGTGATTAATAAAAATTTAGAAAGCATT
35 AATACTAAAGAGGAAAAATAGCAATGGCAAGAAAACACATGTAGGGAACACATGTA
GCCAAAAAATAATATATAATCAGAGAAATAATAGGACTTCTGGAAAAAAAAGATGA
GATCAGATTGGTTAGGATCTTTACTAACATGACAAGAGCATGAATTTTTTTTCTGT

- 195 -

AGATAATAAGTATGAAAGAATTTTAGCTTAAAAATTAGCATAATTTGGATCCACAT
ATGCAAATCAATGAATGTAATTCATAATATAAACAGAACTAAACACAAAAACCACG
TGATTATCTCAATAGACACAGAAAAGGCCTTCAAAAAAATT

5 Sequence ID - 639 nt: 624

GACACACGAGCATATTTACCTCCGCTACCATAATCATCGCTATCCCCACCGGCGT
CAAAGTATTTAGCTGACTCGCCACACTCCACGGAAGCAATATGAAATGATCTGCTG
CAGTGCTCTGAGCCCTAGGATTCATCTTTCTTTTACCGTAGGTGGCCTGACTGGC
ATTGTATTAGCAAACATCACTAGACATCGTACTACACGACACGTACTACGTTGT
10 AGCCCACTTCCACTATGTCCTATCAATAGGAGCTGTATTTGCCATCATAGGAGGCT
TCATTCACTGATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCTACGCCAAA
ATCCATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCACAACACTT
TCTCGGCCTATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACACCA
CATGAAACATCCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAGTAATATTA
15 ATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAAGCGAAAAGTCCTAATAGTAGA
AGAACCCTCCATAAACCTGGAGTGACTATATGGATGCCCCCACCCTACCACACAT
TCGAAGAA

Sequence ID 641

20 CAAGATGACAAAGAAAAGAAGGAACAATGGTCGTGCCAAAAGGGCCGCGGCCACG
TGCAGCCTATTCGCTGCACTAACTGTGCCCCGATGCGTGCCCAAGGACAAGGCCATT
AAGAAATTCGTCATTCGAAACATAGTGGAGGCCGCAGCAGTCAGGGACATTTCTGA
AGCGAGCGTCTTCGATGCCTATGTGCTTCCCAAGCTGTATGTGAAGCTACATTACT
GTGTGAGTTGTGCAATTCACAGCAAAGTAGTCAGGAATCGATCTCGTGAAGCCCGC
25 AAGGACCGAACACCCCCACCCCGATTTAGACCTGCGGGTGCTGCCCCACGTCCCCC
ACCAAAGCCCATGTAAGGAGCTGAGTTCTTAAAGACTGAAGACAGGCTATTCTCTG
GAGAAAAATAAAATGGAAATTGTACTTAA

Sequence ID 642

30 TGCTTGGCCCTCTACCTCCTGCCCTCTTCCTGTTCATCTCCCAACCACTGCACTCT
TGATTTTTTATACCACACAGAAGGTAAGAAAATTCTAGGAACCCTAAGGATCAATCC
TCTCCATTTTCACTCAAATGCCTGGGGCCCAGCTCTGCAATGACTGACTCCAGGGC
CTCTTTCCTCACTGCCAGCATAGAAGTCAGGGGAGCCAGCTGGGGCCCTGCGGTCAG
GAAGGTTCTCATTTTTTGGAGCATTCCTGAGCCCAGATCATAGGAGCAGCTGTCCC
35 TGGTGGGACACAGGAGTCATGACTCCTACCCTCCACCCTCCACACCCACCAGGCAT
TTAGCAGTCTGTCCTATGCAAGACAGATGAATTCTCAGCCAGGATACCTCAAGGCA
GGCAAAGGTGAGTGGAGGGAAAATTCACAAACATTCAGGGTGTGTGGTGCTGGCAT

- 196 -

CACCATGGCCAAATCCAAGAGGTCTTCCTGGAAGAGGGCCCAAACCTGGAACCAAAA
GAATGCTGTCAGCAGTTGGAATAGAGCTGTGAATT

Sequence ID 643

5

CTTTCCAAGAGGAATCCTCGGCAGATAAACTGGACTGTCCTCTACAGAAGGAAGCA
CAAAAAGGGACAGTCGGAAGAAATTCAAAAGAAAAGAACCCGCCGAGCAGTCAAAT
TCCAGAGGGCCATTACTGGTGCATCTCTTGCTGATATAATGGCCAAGAGGAATCAG
AAACCTGAAGTTAGAAAGGCTCAACGAGAACAAGCTATCAGGGCTGCTAAGGAAGC
10 AAAAAAGGCTAAGCAAGCATCTAAAAAGACTGCAATGGCTGCTGCTAAGGCACCTA
CAAAGGCAGCACCTAAGCAAAAGATTGTGAAGCCTGTGAAAGTTTCAGCTCCCCGA
GTTGGTGGAAAACGCTAAACTGGCAGATTAGATTTTTTAAATAAAGATTGGATTATA
ACTCT

15

Sequence ID 644

CTTTGATAGAGAAGAAAATTCTCCTAGGATACAAGAGCCTCAACATTTTAAAGATT
TTCTGCATCTCAAAAGCGTAGGCTCCTTGCTGGGCAAGGTGAGCCTCTGTGAGTCC
TCATAGGACCGAGCAAATCTGATTCACCCCAGAAAATCCAATATCGAAGCTGAGCT
TTGGCCTGAGCGGGTTCCATTTCTCCCCAGATCCTATTTAGGAAGTGTCTCCTGA
20 CAACCTCCAAAAGGTGCTAACATGCAACGTTCTGAAGGGTTATTGCTCAAAAACAA
GATTTTCTTTGTGGTCAAGACTCTGCGAGCCTCGAACACGATGAATCCGCTCGAAT
GGGCTTGGGCTTTGCCCCGGGTGGCGCACGCTCACACGCTGGAAGCACAGCTTTGAC
GATCTCCACACACGCACAGGCACACACGCCACAGATGATGCCGGCTCATTTCTCAGG
GGGTGTCTAAGTTCTGCTTTAAATATTTACCCCCTAATTGTACAAACAATAGGGGC
25 ATGAGCCTGGTACTCGATAAATGGGGACTTNCTTAAAA

Sequence ID - 645

nt: 649

30

35

CTACAGCCTGGGCAGCGCGCTGCGCCCCAGCACCAGCCGCAGCCTCTACGCCTCGT
CCCCGGGCGGCGTGTATGCCACGCGCTCCTCTGCCGTGCGCCTGCGGAGCAGCGTG
30 CCCGGGGTGCGGCTCCTGCAGGACTCGGTGGACTTCTCGCTGGCCGACGCCATCAA
CACCGAGTTCAAGAACACCCGCACCAACGAGAAGGTGGAGCTGCAGGAGCTGAATG
ACCGCTTCGCCAACTACATCGACAAGGTGCGCTTCCTGGAGCAGCAGAATAAGATC
CTGCTGGCCGAGCTCGAGCAGCTCAAGGGCCAAGGCAAGTCGCGCCTGGGGGACCT
CTACGAGGAGGAGATGCGGGAGCTGCGCCGGCAGGTGGACCAGCTAACCAACGACA
AAGCCCGCGTCGAGGTGGAGCGCGACAACCTGGCCGAGGACATCATGCGCCTCCGG
35 GAGAAATTGCAGGAGGAGATGCTTCAGAGAGAGGAAGCCGAAAACACCCTGCAATC
TTTCAGACAGGAAATCCAGGAGCTGCAGGCTCAGATTCAGGAACAGCATGTCCAAA

- 197 -

TCGATGTGGATGTTTCCAAGCCTGACCTCACGGCTGCCTTGCGTGACGTACGTANC
AATATGAAAGTGTGGCTGCCAAAAACCTTGCAG

Sequence ID - 646

nt: 600

5 GAGATGTCTCGCTCCGTGGCCTTAGCTGTGCTCGCGCTACTCTCTCTTTCTGGCCT
GGAGGCTATCCAGCGTACTCCAAAGATTGAGTTTACTCACGTCATCCAGCAGAGA
ATGGAAAGTCAAATTTCTGAATTGCTATGTGTCTGGGTTTCATCCATCCGACATT
GAAGTTGACTTACTGAAGAATGGAGAGAGAATTGAAAAAGTGGAGCATTGAGACTT
GTCTTTCAGCAAGGACTGGTCTTTCTATCTCTTGTACTACACTGAATTCACCCCCA
10 CTGAAAAAGATGAGTATGCCTGCCGTGTGAACCATGTGACTTTGTCACAGCCCAAG
ATAGTTAAGTGGGATCGAGACATGTAAGCAGCATCATGGAGGTTTGAAGATGCCGC
ATTTGGATTGGATGAATTCCAAATCTGCTTGCTTGCTTTTAAATATTGATATGCT
TATACACTTACACTTTATGCACAAAATGTAGGGTTATAATAATGTTAACATGGACA
TGATCTTCTTTATAATTCTACTTTGAGTGCTGTCTCCATGTTTGATGTATCTGAGC
15 AGGGTGCTCCACAGGTAGCTCTAGGAGGGCTGGCAACTTA

Sequence ID 647

CGAATGTGCAGGTTTGTACATAGGTATATATATGCCATGATGGAAATATTTATTT
TTTTAAGCGTAATTTTGCCAAATAATAAAAACAGAAGGAAATTGAGATTAGAGGGA
20 GGTGTTTAAAGAGAGGTTATAGAGTAGAAGATTTGATGCTGGAGAGGTTAAGGTGC
AATAAGAATTTAGGGAGAAATGTTGTTTATTATTGGAGGGTAAATGATGTGGTGCC
TGAGGTCTGTACGTTACCTCTTAACAATTTCTGTCCTTCAGATGGAACTCTTTAA
CTTCTCGTAAAAGTCATATACCTATATAATAAAGCTACTGATTTCCAAAAA

25 Sequence ID 648

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Sequence ID - 649

nt: 425

CAAAAAACGAAGAAAGTGACGACAGTCTGAGGGACTTATGGGAGATCATCAAGT
30 GAACCACTATATGTGTAATGTAAGTCTTGAATGAGAAGAGAGAAGGAGAAGGAGG
AGAGAGCTTATTTGTAGAAATAATGGCTGAAAACATCCCAAACCTTTCCTTTTTTTG
AGGAAAGAAATAGGCATACAAGTTCAGAAACTCAAGGAACTCCAGAGAGGACAAT
TCTAAAGACACCCCCTCTAACATACATTATAATCAAATTGTCAAAGTAAAATACA
AAGAGAATCTTTTAAATTGACAAGAGAGAAAAGCAGCTGGTCACGTTCAAGGGAGTTC
35 TATAAGAATTTGAGCAGATTTCTCAGCAGAAACCTTGCAGGCCAACAGGCAGTGGG
ATGATACATTCAAAGTGCAAAAAAAAAAAAAAAAAA

- 198 -

Sequence ID 650

CGAGAGTTTACCAGTNGCCTAATAATGCAATAAAAAAATGCTTTGAGATAGCTAACN
GCCCATAAACAAACTCAAATTGCTTATAAAGTTTCTTCCCATGTTCCCATTTGAT
*GAAAAGTCTTACATCACATATAACTGGGAAGCAGGGGTCCCTCCTCAATTTTCAGA
5 CATTTTGAAAGGATGACAGTTCTGTTTGTTAGATGAGTAAACCTCTATATTCATAA
GTTCTAAAATCCTTCATTATGAGGGATTCAAAGTATTTATAAAAACACTGCCCTCT
AAAAATTTCTCAGATCTGAAGTATGGNCTTGGNCCTGAATATACAGTGTTATCCT
ATGTTTAAAAGGGTGATCCAGACATGAGACGCAACTAGTTGGTGCATAAGAAGGCC
CCACTTGGCTATTTTCATATCTACCTACAATTGACCAAAAAAATTTTTTAGGCCAG
10 CAATTATTATTTAGCTTCGCTCTTTCTAGTGCAAGAACTGCAGGCTGGATCAGTA
GTTCAACAGCTAAACAGTCATAAAATAGTCATTGGCATGTTAAATTTCTTTCAATG
CTTCAAAGATAAATTCCAATTCTATTTACTTATTCATTGNGACNGNATTACTAAAC
AGGTAAGGATGGGAATA

15 Sequence ID - 651

nt: 251

CTTTGGGAGGCCGAGGCGGGCGGATCACTTGAGGTCAGGGGTTTCGAGACCAGTCTG
GCCAACATGGTGAAACCCCAACTCTACTAAAAATACAAAAGTTAGCCAAGTGTGGT
GGCAAGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATCACTTTGA
ACCTGGGAGGCGGAGGTTGCAGTGAGCCAAGATCGTGCCACTGCACTTCAGCCTGG
20 GCAACAGAGCAAGATTCCGTCCATCTC

Sequence ID 652

CTTTCTTCAGCCTTGCAGACACCTAAACATCATGTAATTACCTAAGGAATTCCCAA
GTGCCTCTTCCAGGTTATACGTGTAAATAGCTGTTTTTATGCAAGATTAGTTAGAT
25 ACTGCTCTTTACAGGATGAGTGGTGTGTTGTTTGGCTGGGGGGGNCTTAAATGTGT
TTCTAATGTGTGTGTCAAATAATTACCTGTAAACAGACTGCCAATCTGGCTGAAG
CCAATGCTTCTGAAGAAGATAAAATTAAAGCAATGATGTCGCAATCTGGCCATGAA
TACGACCCAATCAATTACATGAAGAAACCTCTAGGTCCACCACCTCCATCTTACAC
GTGTTTCCGTTGTGGTAAACCTGGACATTATATTAAGAATTGCCCAACAAATGGGG
30 ATAAAAACTTTGAATCTGGTCCTAGGATTAAAAAGAGCACTGGAATTCCCAGAAGT
TTCATGATGGAAGTGAAAGATCCTAATATGAAAGGTGCAATGCTTACCAACACTGG
AAAATATGCAATCCAACCTATAGATGCAGAAGCATATGCAATTGGGAAGAAAGAGAA
ACCTCCTTNTTACCAGAGAGCCATCTTNTTTCT

35 Sequence ID 653

GTTGTGACTCGTTGGCATGTGATCTGAAGTTCCTGCCCTGCAGCTGACGAGCCAGT
GTTTCAATAATTAAAAACAACCTCAACTCACTGTCCTCCTGCCTTGAATTTGATCAT

- 199 -

TGCGCTTTGCATGTATGTATCACAATACCACATGTACCCCATAAATATGTACAAAG
ATTATGTGTCAATAAAAAACAAAAATTAAAATCCCAATTTTTTA

Sequence ID 654

5 GTTGCTAGTAGCGGCAGGAAGATGTCAGGCTCACTTTCCTCTGATTCCCGAAATGG
GGGGAACCTCTAACCATAAAGGAATGGTAGAACAGTCCATTCTCGGATCAGAGAA
AAATGCAGACATGGTGTACCTGGATTTTTTTCTGCCCATGAATGTTGCCAGTCAG
TACCTGTCCTCCTTGTTTCTCTATTTTTTGGTTATGAATGTTGGGGTTACCACCTGC
ATTTAGGGGAAAATTGTGTTCTG

10

Sequence ID 655

GTCCCCGGGAATCGCGGCCGCGTCGACGGTTTATTTTCAGTGCTTGAAGATACATT
CACAAATACTTGGTTTGGGAAGACACCGTTTAATTTTAAGTTAACTTGCATGTTGT
AAATGCGTTTTATGTTTAAATAAAGAGGAAAATTTTTTTGAAAAAAAAAAAAAAAAAA
15 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAATTTTT

Sequence ID 656

TAGAGGCCTGAATAGGTAGACAATGGCAGCAGCGTTTTTAATCACAGTCCTATTCA
20 TGCCCTAATTCGGGAGTGATGATTAAAGGACATTAGAGGGAGCACTTTGACATCTG
ATCCTTTGAACTGACGTCTGTGCAGGCTGCACTCCATAGAGCTCACTTGGCCAAAC
TGATTTCTTTAAATAAAGTGCTGTGATTTCCAATGTAGGAAATATTACATTAGAGC
CTATTGAAATGATTAGGAATTGAGGAGCTTTTCTTTAGGTGGGAATGTGGTGTATG
CTGTATACTCACAAAAGTGAGATCATTAATATTGCATGTACTACTTTGAATATCAG
25 GGACCACAGAGAAATAGCATGAGAAACGCCTTCCTGCAGTCATGCACTTAAAATGA
ATATGAACAAAAATGTGGAACCTCTGCTGTCATAGCTCTCCG

Sequence ID 657

GGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAAGTTTTTTT
30 CTCTTTGAAAGATAGAGATTAATACAACCTTTAAAAAATATAGTCAATAGGTTACT
AAGATATTGCTTAGCGTTAAGTTTTTTAACGTAATTTTAATAGCTTAAGATTTTAAG
AGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAAAGGTTTCT
AAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAATGTATTTAAA
AGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAGAAGGGCAAT
35 GCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTA

Sequence ID 658

- 200 -

GACCTTTGAGAAAATTAATTTAAATCCTAGAACTTTGGGTGAACCGAAGAAATTTA
TAATATTTGTTTAGTTAATAACAGATAAAAAGGAAAGATTCAAGCCTATTGGATGA
GAATTTGTACATTATTTTAGAGCTAATAATAATGGTTTTTCAGTTTAGTGAGGATTT
AAAAAATGTTTTTGAATCAAACTTTTTTTTCTTTATAATCCTTTTTTAACTAACTCAG
5 GAAATAAGGTATTATGAAATCCACACACTGTTACCTCCTTAAAGTATGAGGATACT
TCCCCTGTTTGGTCCACTAGTGGCTGATTATTTTGTGGATTATTTGTAATT
TTCTTTTTTAATTCTTCCTTAAAGAGCATGGCATTGAGTCACAGACCTATATTTG
AATCCTGTCATTTACTAGCGTTTTGACCTTGAACAATTATGCTCAGAGTCTCAGTT
TTTTCTTGTAAGTGATGATGATACTACTTAACTCACAGGGTTGTAGTGAAGATCA
10 AATGAGATCATGTCTGTANAACACCCTGCCCGGCACTCAATAAGTATTAATAGGAA
CCCATATACCTC

Sequence ID 660

TGTTTTTATTTTTTAAAAGGTATAAACACCAAAAAAAAAAATTAACATTGTATGAAG
15 ATGGAAAATAAGAAGATGCACTTTCTGTAACCTTGTCTAAGGATTTAAATTACTAA
CTTATGAACTCCAATTTGAATTGAACTTAACTATCGGCTTTCTTACTGGTAAAATT
ATATGGTTTATTTTAAATGCGTACATATTGACCAATGGCCTCTGAAAAGCACATT
TTAGATACTGAAATTGAAGGAAAGAAAATGCATCTTCAAACATTTTTTGGAAATCTC
ACCACATATACTTTGTTANATTTGTGTATTGTAGGGTGTGTTTGTATTTTTTGT
20 ATTGTATATGAACTTTTTTTTAAATGTGACAGTTAAACACATCTTTAAAGCATAGT
CACAGACAAAAGCATACAGTATAAAAATTTCTTGAAAACCTCCTACAATATTATAT
TTGGAGGCAGCTTCAGACTGTTTTATTGG

Sequeunce ID 661

CTCTGGCACACATTAGTTCCTCTTATATTACATTGATATAAGCAAGTCATATGGAT
25 TTATCTGAGTGTAAGGAGAGCTGGAAAAAATAGTTTCTAGCAGGTCAGCCACCTCC
CAGTGAGGGCTGCATACCATAGAAGGGGAGAATGAATTTTGGGAAAACAGGTAATT
ATCTCTGTCACAGAAGGGGATGAAAAGTATGGTAGTTACNCAAGTTANACATCTGT
ATGGAAAATACCACTTGGTTCTACAAATGNGG

30

Sequence ID - 663

nt: 627

GCCTCCCGGGTTCAGGGATTTCTCCTGCCTCAGCCTCCTGAGTGGCTGCATTGCAG
GCACCTGCCACCACGCCTTGCAAATTTTTGTGTTTTTTAGTGAGATGGGGTTTTGC
CATGTTGGCCAGGCTGGTCTCGGACTCCTGACCTCAGGTGATCCGCCCCGCTCAGC
35 CTCCCAGAGGGCTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCCCAAGTTTTGC
ATCTTTTAATGCCCTCTGAACAAATACATAGAGAAAACCTCTCAGAACAATTAAAC
CTGCAGAGCAACAGTGTCTCCATGTCTTAGGTTTCAAGTTTGCCTCTAAAATTCT

- 201 -

AATCCATATTTTTCTACTTCTCAGATAATTTATGTGTGTGTACTCTTCCTAGACGT
ACAAGAGACTTTTAAATGCTAAATATTTGTCAGTGCTTAACAAAACTCAATTTCA
CATTACTCATATTGTTTTTTGTTTTAATTGAATGTGAATTAAATTTTTATTAGTTAT
TTGATTTGGAATGTTATGTATGCCATTAACACTATTAGGGGAATCTCTAGCATTTC
5 TGTATTTTTTAAAGAATTTGATTCTTTTGTANATTCTGCCTGTGTGGCATTTTAAAC
ATGTGTGACAT

Sequence ID - 665

nt: 345

ACCGGCGACATGGCCAAACGTACCAAGAAAGTCGGGATCGTCGGTAAATACGGGAC
10 CCGCTATGGGGCCTCCCTCCGGAAAATGGTGAAGAAAATTGAAATCAGCCAGCACG
CCAAGTACACTTGCTCTTTCTGTGGCAAACCAAGATGAAGAGACGAGCTGTGGGG
ATCTGGCACTGTGGTTCCTGCATGAAGACAGTGGCTGGCGGTGCCTGGACGTACAA
TACCACTTCCGCTGTCACGGTAAAGTCCGCCATCAGAAGACTGAAGGAGTTGAAAG
ACCAGTAGACGCTCCTCTACTCTTTGAGACATCACTGGCCTATAATAAATGGGTTA
15 ATTTATGTA

Sequence ID - 666

nt: 252

ATAATTCAGAACTTCTTCATATGCTCGAGTCTCCAGAGTCACTCCGTTCTAAGGTT
GATGAAGCTGTAGCTGTACTACAAGCCCACCAAGCTAAAGAGGCTGCCCAGAAAGC
20 AGTTAACAGTGCCACCGGTGTTCCAAGTGTTTAAAATTGATCAGGGACCATGAAAA
GAAACTTGTGCTTCACCGAAGAAAAATATCTAAACATCGAAAACTTAAATATTAT
GGAAAAAAACATTGCAAAATATAAAAT

Sequence ID 669

TTACTTTTAACCAGNGAAATTGACCTGCCCCGTGAANAGGCGGGCNTGACACAGCAA
GACGAGAAGACCCTATGGAGCTTTAATTTATTAATGCAAACGGTACCTAACAAACC
CACAGGTCCTAAACTACCAAACCTGCATTA AAAAATTTTCGGTTGGGGCGACCTCGGA
GCAGAACCCAACCTCCGAGCAGTACATGCTAAGACTTCACCAGTCAAAGCGAACTA
CTATACTCAATTGATCCAATAACTTGACCAACGGAACAAGTTACCCTAGGGATAAC
30 AGCGCAATCCTATT

Sequence ID 670

GGCTGATTCCTGAGCTATAAAAGCATAATTGCTTTTATATTTTGGATCATTTTTTAC
TGGGGGCGGACTTGGGGGGGGTTGCATACAAAGATAACATATATATCCAACCTTTCT
35 GAAATGAAATGTTTTTAGATTACTTTTTCAACTGTAAATAATGTACATTTAATGTC
ACAAGAAAAAAATGTCTTCTGCAAATTTTCTAGTATAACAGAAATTTTTGTAGATG
AAAAAAATCATTATGTTTAGAGGTCTAATGCTATGTTTTCATATTACAGAGTGAAT

- 202 -

TTGTATTTTAAACAAAAATTTAAATTTTGAATCCTCTAAACATTTTTGTATCTTTA
ATTGGTTTATTATTAAATAAATCATATAAAAATT

Sequence ID 671

5 CAGGAAGTCACCTGGGATTGGCTGCCTCACCCACTCACAGTGCCATCCCTGCCCCA
GGCCTCCCAGTGGCAATTCCAAACCTGGGTCCCTCCCTGAGCTCTCTGCCTTCTGC
TCTGTCTTTAATGCTACCAATGGGTATTGGGGATCGAGGGGTGATGTGTGGGTTAC
CTGAAAGAACTACACCCTACCTCCACCACCTTACCCTCACCTGGAGAGCAGTTAT
TTCAGAACCATTCTACCTGGCATTTTTATCTTATTTAGCTGACAGACCACCTCCACA
10 GTACATCCACCCTAACTCTATAAATGTTGATGGTAATACAGCATTATCTATCACCA
ATAACCCTTCAGCACTA

Sequence ID 672

CAGGAAGTCACCTGGGATTGGCTGCCTCACCCACTCACAGTGCCATCCCTGCCCCA
15 GGCCTCCCAGTGGCAATTCCAAACCTGGGTCCCTCCCTGAGCTCTCTGCCTTCTGC
TCTGTCTTTAATGCTACCAATGGGTATTGGGGATCGAGGGGTGATGTGTGGGTTAC
CTGAAAGAACTACACCCTACCTCCACCACCTTACCCTCACCTGGAGAGCAGTTAT
TTCANAACCATTCTACCTGGCATTTTTATCTTATTTAGCTGACAGACCACCTCCACA
GTACATCCACCCTAACTCTATAAATGTTGATGGTAATACAGCATTATCTATCACCA
20 ATAACCCTTCAGCACTAGATCCCTATCAGTCCAATGGAAATGTTGGATTANAACCA
GGCATTGTTTCAATANACTCTCGCTCTGTGAACACACATGG

Sequence ID 673

GGGTTTTCTTTCGGAAGCGCGCCTTGTGTTGGTACCCGGGAATTCGCGGCCGCGTC
25 GACTGCTAAACAGAATACTGCTATTTTGAGAGAGTCAAGACTCTTTCTTAAGGGCC
AAGAAAGCCACNTGNNCCCTNNGNCTAATCTGGCTGAGTAGTCAGTTATAAAAGCC
NTAATNGCTTNNTNTTTTGGNNTCNTTTTTTNNCNGGGGNCNGNCTTGGGGGGGGTTG
CNTCCAAAGATANCATNTNTTTCCAACCTTNTNNAANNNAANNNGTTTTAAAATCCCT
TTTCCNCCNGAAAANANNGCCCTTTAAGNGCCNCAAAAAAAAAANNGTNTTCTGCAN
30 NTTTTCTANTATNACAAANNTTTTNGTAGAANAAAAATTTTTTTTTTAGNGGCTACC
CTTTNTTTNTTANNCANNGGAGTTTNTTTTTTACAAAAAAAAAANATTGGGNCCCCT
CCACAACCTTGGGTCTNTAATNGGGGGGTTTTTAAATAAANCNTNTNTAAATCCCC
CNNNNNNNNNCNNNNNNNNNCCNNNNNNNNNNNNNNNCCCNNNNNAAAAAATTTTTNC
TCCCCNCCCTTTTTCTTCCTGCCGGCCCCAATTTAAGCCCNGGCGCTTGGGGCAA
35 ATCCCCCTTTAGNGGGGGGGTTTANAAAAACNGGGGCGGGGNTTTAAACCNCGG
GGNNNGGGGAA

- 203 -

Sequence ID 674

ACCTCTAGCATCACCAGTATTAGAGGCACCGCCTGCCCAGTGACACATG-TTTAAC
GGCCGCGGTACCCTAACCGTGCAAAGGTAGCATAATCACTTGTTCCCTTAATTAGGG
ACCTGTNTGAATGGCTCCACNAGGGTTCACCTGTCTCTTACTTTTAACCAGTGAAA
5 TTGACCTGCC

Sequence ID - 675

nt: 591

GTATAGAAAATAATGTCCCCAGNGCATAGAAAAAATGAGTCTCTGGGCCAGTGAAT
ACAAAACATCATGTGCGAGAATCATTGGAAGATATACAGAGTTCGTATTTTCAGCTTT
10 GTTTATCCTTCCTGTTAAGAGCCTCTGAGTTTTTTAGTTTTTAAAAGGATGAAAAGCT
TATGCAACATGCTCAGCAGGAGCTTCATCAACGATATATGTCAGATCTAAAGGTAT
ATTTTCATTCTGTAATTATGTTACATAAAAGCAATGTAAATCAGAATAAATATGTT
AGACCAGAATAAAATTAATTATATTCTGGTCTTCAAAGGACACACAGAACAGATAT
CAGCAGAATCACTTAATACTTCATAGAACAAAAATCACTCAAACCTGTTTATAAC
15 CAAAGAATTCATGAAAAAGAAAGCCTTTGCCATTTGTCTTAGAAAGTTATTTTTTA
AAAAAAAATCATACTTACTATTAGTATCTATGGAAGTATATGTAACAATTTTTATG
TAAAGGTCATCTTTCTGTGATAGTGAAAAAATATGTCTTTACTAAGTTGAAATGAA
TACTTTCTGNCTTTGCTAATGGATAGTTATT

20 Sequence ID 676

CTCAATTCTACTAAAAAGCCCCCAAGAAAAGCGAATGAGAAAACAGAGTCATCCT
CTGCACAGCAAGTAGCAGTGTACGCCTTAGCGCTTCCAGCTCCAGCTCAGATTCC
AGCTCCTCCTCTTCCTCGTCGTCGTCTTCAGACACCAGTGATTTCAGACTCAGGCTA
AGGGGTCAGGCCAGATGGGGCAGGAAGGCTNCGCAGGACCGGACCCCTAGACCACC
25 CTGCCCCACCTGCCCTTCCCCCTTTGCTGTGACACTTCTTCATCTCACCCCCCCC
TGCCCCCCTCTAGGAGAGCTGGCTCTGCAGTGGGGGAGGGATGCAGGGA

Sequence ID 679

GNANCNTTTCCTNTCGNAAANCGCGCCTTGTGTTGGTACCCGGGAATTCGCGGCCG
30 CGTCGACAAAAAATAAAAAAAAAAAAAAAAAAANTNTAGACTCGANCAAGCTT
ATGCANGCNTGCGGCCGCAATTCGAGCTCGGCCGACTTGGCCAATTCGCCCTATAG
NGAGTCGTATTACAATTCACCTGGCCGTCGTTTTACAACGTCGNGACTGGGAAAACC
CTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGT
AATANCGAANAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGG
35 CGAANGGAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTTGTTA
AATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAA
AAGAATAGACCGAGATAGGGTTGAGNGTTGTTCCAGTTTGAACAANAGTCCACTN

- 204 -

TTAAAGAACGNGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGG
CCCCTACGTGAACCATCNCCCTAATCAAGTTTTTTGGGGTCGAGGNGCCGTAAAG
CACTAAATCGGAACCTTAAAGGGAGCCCCCGATTTAAAGCTTGACGGGGGAAAGCCC
GGCGAACGTGGCGAAA

5

Sequence ID 682

CACCTGCAGTCCAAGTACATCGGCACGGGCCACGCCGACACCACCAAGTGGGAGTG
GCTGGTGAACCAACACCGCGACTCGTACTGCTCCTACATGGGCCACTTCGACCTTC
TCAACTACTTCGCCATTGCGGAGAATGAGAGCAAAGCGCGAGTCCGCTTCAACTTG
10 ATGGAAAAGATGCTTCAGCCTTGTGGACCGCCAGCCGACAAGCCCGAGGAAAAGTG
AAACTTTGCTTAACNACCGAATGGNNGGGGANCTTTTCCAACGNTTTT

Sequence ID 683

TTGGTTTCATACTGNTGGGGNTTGAATGNTCCCTNCAACACTNATGTTGANACTTA
15 ATCCCTAATGNNGCAATACTGAAAGGTGGGGCCTTTGAGATGTGATTGGATCGTAA
GGCTGTGCCTTCATTCATGGGTTAATGGATTAATGGGTATCACAGGAATGGGACT
GGTGGCTTTATAAGAAGAGGAAAAGAGAACTGAGCTTGCATGCCC

Sequence ID - 684

nt: 545

20 GTGGAAGNGACATCGTCTTTAAACCCTGCGTGGCAATCCCTGACGCACCGCCGTGA
TGCCCANGGAAGACAGGGCGACCTGGAAGTCCAACACTTTCCTTAAGATCATCCAA
CTATTGGATGATTATCCGAAATGTTTCATTGTGGGAGCAGACAATGTGGGCTCCAA
GCAGATGCAGCAGATCCGCATGTCCCTTCNCGGGAAGGCTGTGGTGCTGATGGGCA
AGAACACCATGATGCGCAAGGCCATCCGAGGGCACCTGGAAAACAACCCAGCTCTG
25 GAGAACTGCTGCCTCATATCCGGGGGAATGTGGGCTTTGTGTTACCAAGGAGGA
CCTCACTGANATCAGGGACATGTTGCTGGCCAATAAGGTGCCAGCTGCTGCCCCGTG
CTGGTGCCATTGCCCCATGTGAAGTCACTGTGCCAGCCCAGAACACTGGTCTCGGG
CCCGATAAGACCTCCTTTTTTCCAGGCTTTAGGTATCACCACTAAAATCTCCAGGGG
CACCATTGAAATCCTGAGTGATGTGCACTGATCAAGACTGG

30

Sequence ID 685

GGAAAGGGCCATTTTATTGCCTAAAACCACCTGGNTTTTNAGGTAACAGTTCCAAC
ATGTCCTTTTTTTGAATAGCTGTTCTAATTATTATATATTCAGCTGATTAATAGGAG
TACTTGATAGGTGGACTGTGTCAGGTAGCCTCAGGCAATCCTACTTCAACAAGCTG
35 TCAGGGAGCCATGCCATGCTTCTTTATGACATAGGTGAATTTGATAGGCTCACTAG
CAGAACATGGGATCACAAGGTGGAACCNITCCNTTT

- 205 -

Sequence ID 686

GACCCCTTCCTTACACCTTATACAAAAAACTGAAACTGGACCCCTTCCTTACACC
TTATACAAAAATTAACCTCAATTTTATTATGTTGTATTAAATTAAGTTGGGTTTAAT
TAAGATGGATTAAAGACTTAATTATAAGACCTAAAACCATAAAAACCCTAGAAGAA
5 AACCTAGGCCATACCATTTCAGGACACGGGTATGGGCAAAGACTTCATAACTAAAAC
ACCAAAAGCAATGGCAACGAAGTCCAATAGACAAATTGGACCTGATTAACTAAA
GAGCTTCAGCACAGCAGAAGAGACTATCGTCAGAGTGAACAGGCAACCCACAGAAT
GGAAGAAAATTCTTGCAATCTATCCATCTGACAAGGGGCTAATATCCAAAATCTAC
AAAGAACTTAAACAAATTTACAAGGAAAAACACAAACAACCCCATCAAAAAGTGGG
10 CTAAGGATGTGAACAGACACTTCTCAAAGAAAACATTTATGCAGCCAACAAACAT
GAAAAAAAGTTCATCATCACTGCTCATTAGAGACATGCAAATCAAACCACAATGA
GATCCCATCCACACCAGTTAGAATGGCAATCATTA AAAATGT

Sequence ID - 687

nt: 268

15 TTTATGTGTTTTTGTCTTGGGGGGCGCTGGGCCTAGCCCAGAGTAGTGCTTGCTCCC
CCTGCCTTGTCCCACCAGGGAGGCAGCAGACTCAGGCCCTCCATGGTCCTCTTTGT
CATTTTGTGACATGCATTCCTCCTTTTGTGTCATCTTGTTGGGGGGAGGGGATTAAC
CAAAGGCCACCCTGACTTTGTTTTTTGTGGACACACAATAAAAGCCCCGTTTATTG
TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

20

Sequence ID - 688

nt: 569

CTTTAGCCAGCCTGATCAGAAAAAAACAAAAGAAGAGGAAAGACGTAGATTACCAA
CATCAAGAATGTGAGTTATGATATCACTACAGACTCTCCAGGTATTAAAAGCATAA
TTAGAGAATGATATGAGCAGCTATATGCAAATAAGTTCAACATTGGACAAATGGAC
25 AAATTTCTTGAAAGATAAATTATGAAATTTTCATTCTGAAAGAACTACATGACCTTA
ATTGTCTTACATCTATTAAATAAGTGGAATTGTAGTTTAGAAACTTTCCCACAAA
GAAACTCTAGGCCCAGATGGCATCAAAATAATATTCAGATGAATGAAATGGAGAA
AGGATAGCCTTTTCAACAAATGGTGGTGAACAATTGGATTTCCATATGCAAAAAA
ATAGAGATGGACGCAGAGGTGTGTGCTTAGGAGGCTGAGGTGAGAGGATTGTTTGA
30 GGCCAGCCTGGGCAACATAGCAAGACCCCATTTCAAAAACAAAATAAAGAACTTG
TAGCCTTACCTTGTGCCATATTATGAAAATGTATCATAGGCTTAAATGTGAAACGT
AAAACAAA

Sequence ID 689

35 CGCAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAAGTTT
TTTTCTCTTTGAAAGATAGAGATTAATACAACCTACTTAAAAAATATAGTCAATAGG
TTACTAAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAAGATT

- 206 -

TTAAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAAAGG
TTTCTAAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAATGTAT
TTAAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACTAATAGAAGG
GCAATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTTAAA
5 AGTTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAAGAGATTAAACCGA
AGGTGATTAAAAGACCTTGAAATCCATGACGCAGGGAGAATTGC

Sequence ID 690

CGAAAAGCAAATATAACTTGCCACTAACCAAGATCACCTCTGCAAAAAGAAATGAA
10 AACAACTTTTGGCAGGATTCTGTTTCATCTGACAGAATTCAGAAGCAGGAAAAAAA
GCCTTTTAAAAATACCGAGAACATTAAAAATTCGCATTTGAAGAAATCAGCATTTTC
TAACTGAAGTGAGCCAAAAGGAAAATTATGCTGGGGCAAAGTTTAGTGATCCACCT
TCTCCTAGTGTTCTTCCAAAGCCTCCTAGTCACTGGATGGGAAGCACTGTTGAAAA
TTCCAACCAAAACAGGGAGCTGATGGCAGTACACTTAAAAACGCTCCTCAAAGTTC
15 AAAGTTAGATTTCAGATTT

Sequence ID 691

CCGGTCTCTACACAATATATAGAAATCTGGGCATGGTGGTGCCTGGCTGTAGTCTC
AGCTACCTAGTTGGGTGAGGTGGGAGAGTCGCTTGAGTCCTGGAGGTTGAGGCTGT
20 AGTGAGCCAGGGCTGCACCACTGCATTCCAGCCTGGGTAACAGAGTGAGACCCTGT
CTCAAAAAGAAAAAATAATTGCTAATTTTAACAAATCACAAAAGTACTCAGGC
AAGTTGTCTGACTCAAAGCCCTTGAAAAACCATCAAAGACAGTAGAATGTTAAGT
GGTCATTTACGTAAAATAGTGTTTCAATTAATTTTGGTTTCAATTTAGGATAATCATT
TTAAATGAGACTGTATTTGAGACTGTATACACATACATATACATGTTTACACACAT
25 ATACGTACAATATATGTACATTCTATCTAAAAGATCATAACATGTGTGTACATATAT
GTTTTTAAAAGTCAAAGTACATATTAATGGAAACAGTGCTTACATCTCTGGTAGT
GATTTTCTATTAGCAGCAGCCCTACATATGCTGCGTCTCTGAACAGCATGTCAGTG
CCATGACTGTCTAAACATGCAAATATGACTGACAGACTCTTGAGACAGCTTTCACC
TTG

30

Sequence ID 692

AATTCGNGGCCGCGTCNNCCTANGAGGCACCAGGAAATCCCGCGGGGTGGCCCATG
CAGACCAGGCGCACGTGGCTCATGGGGCANAATTGCCAAGGACAGCTCACGACAGT
GCCACCTTCTCACCATTCCAGCCAAGGAGAGATGTGACGTTGGAAGTCTCTGGCA
35 CTTCTGTCAAGCCTCCCCCGCCCCAATTGCCTTGAGATCTCTGCTCTTTGTGAGAG
ATTTGCAAAGACTCACGTTTTTGTGTTTCTCATCATTCATTGTGATACTAAGA
AACTAAGAAGCTTAATGAAAAGAAATAAAATGCCTATGTTGTTGTTCT

- 207 -

Sequence ID 693

CTAGAACCCATGACTCCTAGGTCTTATACTGCAACCACAGTATCAGCAAATAATCT
TTCATAAGGGGATTATTCTCTGATTAACAGGAAATACAGGAATTTAATTTGTGAAC
ACGCTAGGTAGAAGCAGAAACCCAAATCCAAATCCAAATTTAAACATTTAAAATTC
5 ATTCTATAACTAAGATCTAACAGTCATTTTCTTCCCAGTAAGAAATAACCAAAGCA
TGCTAAAAATCACTGGACTAAATTGGTGTCAAACTGCCACATTGCCAGGCATGGG
GGGGTCATACTTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGAAAATTGCTTGA
GGCCAGGAGTTCGAAACCAGCCTGGGCAACACAGTGAGACCCCATCTCCACAAAAA
AAAAAAATTAAAAAACAAACAAAACATTAGCTGGGCATGGTGGTACACGCCTGTA
10 GTCCCAGCTACTCAGGAGCCTGAAGTGAGAGGATCACTGAAGCCCAGGAGGTAGAG
CTATGACTGTAGTGAGCTATGACTGTGCCACTACACTCCACCTGGGTGACAGGGGA
CTC

Sequence ID 694

15 CGACTTCCATTTGTATTAATGGAATACTAAGTCCCTCTGTGATTTCTGAACCAAGC
TATTCCTAGGCCTGAGTTTTATTTTGTGACACAGAAATAAATTANAAGGCCAAGC
GTGGTGGCATGTGCCTGTAGTCCTAGTTGCTGAGGTAAGAGGATTGCTTGAGCCCA
GGAGTTCAAGGCTGCAGCAAGCTTTGATTGCGCCACTGCACTCCAGCCTTGGCGAC
AGACTAAGACGCTGTCTCAAAAAAAAAACAAAAA

20

Sequence ID 696

GGTTATCAATGAGATTAAGAGACAACCTAGAGTAAAAACAAAAGAAAAGAAAAGAAA
NGAAAACAACAGAAGCTCTATTAAGTACCTCTAACCAATACAACAGGTTAAGTGA
TGTTCTCCATTCTGTATATAAAAATCCCAGTGGACACCCACAACACAGGCCTTCAGG
25 CTTGTAGGACACTTTCTAGTTCATCTGAGCACTTTTGTTCCTCAGCAGTTGAGCTGT
ATACTTAGCAACATTTGGTGCTTCCAAACCCATTTGTGCCTGTAGCACTTACTATT
GAAATACATAATTTAATTAAATATTATATAAAGGAATGGAATACGAGTTGGACAAG
AAAAAGAGTTAAATCTGAAGGTTAGGTAAAAAGAGCAACTTCTTTTCTCTGTTTTG
CAGGTTGGCAAAATCATTTAAAAACAATTGGAAGTATTATATGTTCTGCATTAAGT
30 TGTCATTTTACTTAAAAACTAGGCATCAAAGATGATGCATAATAAATTTAGTGTAT
GCAAGAATGACTGCTTGGGACCTCAATATATGAATTCTTAATCCAAGGAAAGTCCT
TGGCCTTACATTTAAAAGTCGGCAAATAAGTGTACGTTTCATT

Sequence ID 697

35 GAACATTTAAAAATAATGCAAATAAGGCTGGGCGTGGGGGCTCACACCTGTAATCC
CAGCACTTTGGGAGGCCGAGGCAGGCAGATCACGAGGTCAGGAGATTGAGACCATC
CTGGCTAACACAGTGAAACCCTGTCTCTACTTAAAAAATAAAAAAATTAGCCAGGC

- 208 -

GTGGTGGTGGGCGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGGT
GTGAACCCGGGAGGCGGAGCTTGCANTGAGCTGAGATCGTGCCACTGCACTCCAGC
CTGAGCGACAGAGCGAGACTCTGTCT

5 Sequence ID 698

TCATTAGAATCCAAGCTTTGAAAATTTCTGATTAATGCTCATGTATTTCTTTATCT
TTGTTTTTTCCTTGTGAAGAAAGACTTTCACCACTGTCTGAGTGATGATGCTGTTGA
TAAGGATGATGTGATGACTACTATATTGCATCTCTCAGGAACAGCTGATGGGAAG
GGAGGGGCTGCTGAGTTCCTTGTCTAGCTAGCAGCACGCTCCTCANAGAGGGGG
10 CCGAGTTACAGACAGCAGCCGCATTTCTCATGCAAAATTAGTTTTAAACTGCTAGTG
TGGGCATCGGTACCTTTTGCCTGGGTGATACCGAAGAATTGTTGAGGATTTAGTAT
GCTCCGTAGAGACAGTTCAGCCAGTCATTTCTGCATTGGAGAGACTTCTCATACTT
TCTTTGAAGACTCATAGAAAGCTGGAT

15 Sequence ID 699

ATTAAGGTTTGTNCCCAACAAGAATAGATGTAATTAGAAAAAANTGNCTTCCTTAC
CTATTGCCTCTGATNTTTACTTGCTTAAATTTTTTTTTTATTGNAAATCCAGAAAAAG
NGGATTTAGAGAACAACACTAACTCCACCTAATCTATGACAGANATGTACAANAN
AGTACCTGTGAAAAATGTGAAAGNATNTGAAAAATGTAACCTTTGGCAGCCTGAGC
20 ATAGTCAACCAGAAAACTATCTGAATTAAAATAATTGGTCCATAGGTACTATTTT
ATTTGGTCCATAAGGATTATTTTTTCAACTTTTTTTTCAAGTGTATTATTATGTCA
TTTCCCACGTAGGTACTGATACCTGAAGACTTTTTTNCACCTTTAACCTTNCTCGT
TGAGGAGCTTTGTANTCTAATAAAAGAGAAATATAAGTAAATGTTAGATATATGGG
NGGATAATGGTAACTATGTGCTTAAAGAGGTATAAAAGAAGGGTAGGGAGCAGATA
25 AGACAAAGGAAGGGCTATATTATAANGAAGAATATTCCAAGTAGGGAAGAGAAAAA
GATATGTTATCCATATAATATTTTATGTGCAGTAGAGAACATGTTCTATAGAANAG
ACAGAAGATG

Sequence ID 700

CTTGAGCCCAGGAATTCCAGCCTGGGCAATATAGTAAGACTCCGTCTCTACAAAAG
ATACAAAAATTAGCCAGATGTGGTGGTGCCTGTAGTTCCAGATACTGGAAAG
ACTGAGGCAGGAGGATTGCTTGAGCATGGGAAGTTGAGGCTGCAATGAGCTGTGAT
TACGCCACTACACTCCAGCCTGGGCAACAGAGTAAGATCTTGTCTCAAAAAAAAAA
TTGAATTCAGCTAAAAATAATAAAATTTTAAAATAATTTTAAAAAGCCCTCAACAG
35 CTTTGTTTTTCTCTCCTTGCCAGCTTCTCTGCAGCCTATAGCCTGCAGGCTGGCTG
CTGCGAGCCAGGACAAGCGGTGGGAAATGCAATCACAGCGTGAAATCTCTGTGTTC
AGAGACACGCAGGAAGCAGGTGAACCATGAAGGGCCAACACATGCCCCCAGTTAGC

- 209 -

AGGGTGTAGAGACCGGGGCAGGGCTTTCTTCTTCCTTCTGGGTATATAAATATCCAT
GTCCTGCCATTTGAAGCTGCAAGTGGCACACATGGATGCTGGACAGGCGCTCGCAC
TTTCTGGGCAGGGCANGGGGCTCAAAGGCAGGACAGCTGGGCAAAGCACCTTGCG
TGGGCCC

5

Sequence ID - 701

nt: 579

CTTTGGAGCTTCTGTCTGTGCTGTGGACCTCAATGCAGATGGCTTCTCAGATCTGC
TCGTGGGAGCACCCATGCAGAGCACCATCAGAGAGGAAGGAAGAGTGTTTGTGTAC
ATCAACTCTGGCTCGGGAGCAGTAATGAATGCAATGGAAACAAACCTCGTTGGAAG
10 TGACAAATATGCTGCAAGATTTGGGGAATCTATAGTTAATCTTGGCGACATTGACA
ATGATGGCTTTGAAGGTAATTAAAATTATCAAATTGGTGCTTGATTTCTGCTTTTA
AAATGGTTTATGGAAGAAAATATGATTAAAGTTTTGTATTGTTTTCTTCCTATAG
AAGATGGAGCCAGAATGGCATGCTAAGTTTTTTCTTTTCTTTAGTGTTATATATGA
CTTCTCCTCAATTGTCACCCATTGATCTTTACCACTGTTAATAATGGATGATATTC
15 AAAATACCTTATTTTCAGTGATTCTAAGGCACCATTGATTAGAACTGCATTATTAT
TTATGTGTCCCTAAAAGCTACCTATTAAGCTGTTACACCCACCATTTTTCTGTAA
GAAAATCCTGATTTTCAGAA

Sequence ID 702

20 GTNNTCCTCTCGGAACGCGCCTTNTGTAGCCAGGTGCTACCAGACCNAATACACGG
TTGTTCCAGCTTGCGCATTACCGATGGCGTAGATATCCGGATCGGAAGTCTGGCA
GGAATCATTAATGACAATAACCCACGCGGAGCAACGTCCAGACCACACTGGGTG
CCAGCTTATCGCGCGGACGGATACCGGTAGAGAAGACGATAAAGTCGACTTCCAGT
TCGCTGCCGTCGGCAAACGCATGGTTTTACGCGCTTCAACACCTTCCTGCACAAT
25 CTCAAGGGTGTTTTTGCTGGTGTGAACGCGCACGCCATACTTTCGATTTTGCGAC
GCAGCTGCTCGCCGCCCATCTGATCAAGCTGTTCTGCCATCAGCATAGGGGCAAAT
TCGATAAC

GTGGGTTTCAATACCTAAGTTTTTCAGCGCGCCTGCGGCTTCCAGACCTAACAGGC
CGCAATTCGAGCTCGGCCGACTTGGCCAATTCGCCCTATAGTGAGTCGTATTACAA
30 TTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAAGAGGC
CCGCACCCGATCGCCCTTTCCAACAGTTGCGCACCTGAATGGCGAATGGAAATTGT
AAGCGTTAATATTTTGTTAAAATTTCGCGT

35 Sequence ID 703

CTGCGCAGACCAGACTTCGCTCGTACTCGTGCGCCTCGCTTCGCTTTTCCTCCGCA
ACCATGTCTGACAAACCCGATATGGCTGAGATCGAGAAATTCGATAAGTCGAAACT

- 210 -

GAAGAAGACAGAGACGCAAGAGAAAAATCCACTGCCTTCCAAAGAAACGATTGAAC
AGGAGAAGCAAGCAGGCGAATCGTAATGAGGCGTGCGCCGCCAATATGCACTGTAC
ATTCCACAAGCATTGCCTTCTTATTTTACTTCTTTTAGCTGTTTAACTTTGTAAGA
TGCAAAGAGGTTGGATCAAGTTTAAATGACTGTGCTGCCCCCTTTCACATCAAAGAA
5 C TACTGACAACGAAGGCCGCGCCTGCCTTTCCCATCTGTCTATCTATCTGGCTGGC
AGGGAAGGAAAGAACTTGCATGTTGGTGAAGGAAGAAGTGGGGTGAAGAAGTGGG
GGTGGGACGACAGTGAAATCTAA

Sequence ID 704

10 CTTGTATTCAAGAACTACTGTAATGCATTAGTGGTCTGGCTTCATTTTGTATGATG
CCAGATCCTTAATTTACCCAGCACAAATCATTTTCAGTAGTTTCCTATGGCTCCTGCA
AAAATGCAAACAGAAACCACCACAGGAACAGCCCCCTTGCTGCCTCCTGTTGCTGAG
GTAGTAGTCGCTAAAGAAAATTGAAGGCTCCTTACAATCTATATTTGAAAAC TAGA
ACTTCTGTAGAAACACACAGATCCCGATCTTAGAAGTTGTACAGGACAATCTGGTA
15 AA ACTGACATAATTGTGATTTTATTAACATGAATTAAAATGCCCAACCAGTGCTTCA
GTGTGACAGTATATTTAAAATAAAAAAGAAATTAAAGGTCATATACTGTACTACTT
TCACAAAGATCCACAGTTTTTGCAAAGACTTGTCATATGTACAATGCTATATATCA
AATGAGAAAAGCTGTAAGCAATTATATACGCAAAGAAATGGCAGTA

20 Sequence ID 705

TTCCAGTCCTTTTCATTTAGTATAAAAGAAATACTGAACAAGCCAGTGGGATGGAAT
TGAAAGAACTAATCATGAGGACTCTGTCCTGACACAGGTCCTCAAAGCTAGCAGAG
ATACGCAGACATTGTGGCATCTGGGTAGAAGAATACTGTATTGTGTGTGCAGTGCA
CAGTGTGTGGTGTGTGCACACTCATTCTTCTGCTCTTGGGCACAGGCAGTGGGTG
25 TAGAGGTAACCAGTAGCTTTGAGAAGCTACATGTAGCTCACCAGTGGTTTTCTCTA
AGGAATCACAAAGGTAAACTACCCAACCACATGCCACGTAATATTTTCAGCCATTCA
GAGGAAACTGTTTTCTCTTTATTTGCTTATATGTTAATATGGTTTTTTAAATTGGTA
ACTTTTATATAGTATGGTAACAGTATGTTAATACACACATACATATGCACACATGC
TTTGGGTCCTTCCATAATACTTTTATATTTGTAAATCAATGTTTTTTGGAGCAATCC
30 CAAGTTTAAGGGAAATATTTTTGTAAA

Sequence ID - 706

nt: 496

CAACCCTCTCTCCTCAGCGCTTCTTCTTTCTTGGTTTGATCCTGACTGCTGTCATG
GCGTGCCCTCTGGAGAAGGCCCTGGATGTGATGGTGTCCACCTTCCACAAGTACTC
35 GGGCAAAGAGGGTGACAAGTTCAAGCTCAACAAGTCAGAACTAAAGGAGCTGCTGA
CCCGGGAGCTGCCCAGCTTCTTGGGGAAAAGGACAGATGAAGCTGCTTTCCAGAAG
CTGATGAGCAACTTGGACAGCAACAGGGACAACGAGGTGGACTTCCAAGAGTACTG

- 211 -

TGTCTTCCTGTCCTGCATCGCCATGATGTGTAACGAATTCTTTGAAGGCTTCCCAG
ATAAGCAGCCCAGGAAGAAATGAAAACCTCCTCTGATGTGGTTGGGGGGTCTGCCAG
CTGGGGCCCTCCCTGTCGCCAGTGGGCACTTTTTTTTTTCCACCCTGGCTCCTTCA
ACACGTGCTTGATGCTGAGCAAAGTTCAATAAAGATTTTGGGAAGTTT

5

Sequence ID - 707

nt: 397

10

CGGATGTGGTGGCAGGCGCCTCTAGTCCCAGCTACTCGGCAGGCTGAGGTAGGAGA
ATGGCTTGAACCCAGGAGGTGGAGCTGACAGTGAGCCGAGATCGCGCCACTGCACT
CCAGCCTGGGCGGCAGAGCGAGACTCCATCTCAAAAAAAAAAAAAAAAAAATAGA
CTTTGAGACCAGCCTGACCAACATAGTGAAACCCGTCCTACTAAAAATACAAAAA
TTACCCGGGCGTGGTGACGGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGACA
GGAGAATCACTTGAACCAGGGAGGCGGAGGTTGTAGTGAAC TGAAATCGTGCCCCCT
GCACTCCAGCCTGGGTAACAAGAGCGAACTCCGTCTCAAAAATAAATAAATAAAT
AAAAT

15

Sequence ID - 708

nt: 293

20

CCAGCTTTTATGGTGTTTAATCTAATACTTAAGCTGCAGTCCCAAATTAGGG
GTCCTTCAGTCTTGAGACTATAAGGGAGCCTCTGCACCCAGGGAAAATGTTACCC
TTTACAGGGGGGAAGGGTAAACCAGTAGGGAATACAGTACAATCCCAACCCTACTG
GGAGGGGCGGGAGGGAGGTGTTGCCGTCCTGTATTAAGTCGATGTTGGGAAACGT
TTTAACATCTGGAGCCTTTGTGGGTGGAAATATGTCTCCAGTTACAACCTCCGCAGT
GGATGTGAAGAAG

Sequence ID 709

25

GGAAGCTACAATGATTTTGGGAATTACAACAATCAGTCTTCAAATTTTGGACCCAT
GAAGGGAGGAAATTTTGGAGGCAGAAGCTCTGGCCCCCTATGGCGGTGGAGGCCAAT
ACTTTGCAAAACCACGAAACCAAGGTGGCTATGGCGGTTCCAGCAGCAGCAGTAGC
TATGGCAGTGGCAGAAGATTTTAATTAGGAAACAAAGCTTANCAGGAGAGGAGAGC
CAGAGAAGTGACAGGGAAGCTACAGGTTACAACAGATTTGTGAACTCAGCCAAGCA
CAGTGGTGGCAGGGCCTAGCTGCTACAAAGAAGACATGTTTTAGACAAATACTCAT
GTGTATGGGCAAAAAACTCGAGGACTGTATTTGTGACTAATTGTATAACAGGTTAT
TTTAGTTTCTGTTCTGTGGAAAGTGTAAGCATTCACAACAAAGGGGTTTTAATGTA
NATT

35

Sequence ID 710

TGGATTCCCGTCGTAACTTAAAGGGAACTTTCACAATGTCCGGAGCCCTTGATGT
CCTGCAAATGAAGGAGGAGGATGTCCTTAAGTTCCTTGCAGCAGGAACCCACTTAG

- 212 -

GTGGCACCAATCTTGACTTCCAGATGGAACAGTACATCTATAAAAGGAAAAGTGAT
GGCATCTATATCATAAATCTCAAGAGGACCTGGGAGAAGCTTCTGCTGGCAGCTCG
TGCAATTGTTGCCATTGAAAACCCTGCTGATGTCAGTGTTATATCCTCCAGGAATA
CTGGCCAGAGGGCTGTGCTGAAGTTTGCTGCTGCCACTGGAGCCACTCCAATTGCT
5 GGCCGCTTCACTCCTGGAACCTTCACTAACCAGATCCAGGCAGCCTTCCGGGAGCC
ACGGCTTCTTGTGGTTACTGACCCCAGGGCTGACCACCAGCCTCTCACGGAGGCAT
CTTATGTTAACCTACCTACCATTGCCCTGTGT

Sequence ID - 711 nt: 498
10 GTGGTACATATACACAAAGGAAAACCTATGTAGCCATTAAAAGAAAAGGAACTCCTA
TCATTTGTAACAACATAAATAAATCTGGAGGAGATTAGGCTAAGGTGAAATAAGCC
AGGCACAAAAAGACAACCTACCATATGATCTTACTTATACGTGTGTGGAATCTAAAA
AGGTGGAATTTACAGAAGCAGAGAGTAGAATGGTGATTACCAGAGGCTGGGGAGTG
AGGGCAGGAGGTTGGAGAAATGTTGGTCAAAGGATACAAAGTTTCAGTTATACAGG
15 ATGAATAAGTTCAAGAGATCTATTGTACAACGTGGTGGCTATAGTTGATAACAATG
TATTGTGTTCTTGAAAAATGCTGAGAGAGTAGATTTTAAGTGTTCTCACCACAAAA
CATAAGTATGTGAGGTAATGCATGTGTTAATTANCTTAATTTAGACATTTTCATAAT
GTATTATACATATTTCAAAACCACGTTGTACATGAGAAAGATACACAATT

20 Sequence ID 713
GCCCAGTCGACCCATGTTCTCCTTTCTACACCAGCATTAGACGCTGTCTTCACAGA
TTTGGAAATCCTGGCTGCCATTTTTGCAGCTGCCATCCATGACGTTGATCATCCTG
GAGTCTCCAATCAGTTTCTCATCAACACAAATTCAGAACTTGCTTTGATGTATAAT
GATGAATCTGTGTTGGAAAATCATCACCTTGCTGTGGGTTTCAAACCTGCTGCAAGA
25 AGAACACTGTGACATCTTCATGAATCTCACCAAGAAGCAGCGTCAGACACTCAGGA
AGATGGTTATTGACATGGTGTAGCAACTGATATGTCTAAACATATGAGCCTGCTG
GCAGACCTGAAGACAATGGTAGAAACGAAGAAAGTTACAAGTTCAGGCGTTCTTCT
CCTAGACAACTATACCCGATCGCATTCAGGTCCTTCGCAACATGGTCACTGTGCAG
ACCTGAGCAACCCACCAAGTCCTTG

30
Sequence ID 714
CTGTAACAGAGATTCCTTTTTTCAATAATCTTAATTCAAAGCATTATTAGACTTG
AAAGGGTTTGATAATCTCCCAGTCCTTAGTAAAGATTGAGAGAGGCTGGAGCAGTT
TTCAGTTTTTAAATGAGTCTGCAGTTAATATCAAATGTGAGTTTGGGACTGCCTGGC
35 AACATTTATATTTCTTATTCAGAACCCTTGATGAGACTATTTTTTAAACATACTAGT
CTGCTGATAGAAAGCACTATACATCCTATTGTTTCTTTCTTTCCAAAATCAGCCTT
CTGTCTGTAACAAAAATGTACTTTATAGAGATGGAGGAAAAGGTCTAATACTACAT

- 213 -

AGCCTTAAGTGTTTCTGTCATTGTTCAAGTGTATTTTCTGTAACAGAAACATATTT
GGAATGTTTTTCTTTTCCCCTTATAAATTGTAATTCCTGAAATACTGCTGCTTTAA
AAAGTCCCACCTGTCAGATTATATTATCTAACAATTGAATATTGNAAATATACTTGG
CTTACCTCTCAATAAAAGGGTCTTTTCTATT

5

Sequence ID 717

TCCACCCACCTTGACCTCCCAAAGTGCTGGGATTATAGGCGTGAGCCACCTCGCCC
AGCCCGATACTAGGACTTATGCAGAAAAAACCTTGACATGGAGGAAAGTAAGATCT
AAATAAATACTGTATTCATAGATTAAAAGACTCAGCATAATAAATATACCATTTCT
10 CCCCAGATTGATGTACAGATTTAACACAATTCCTATCAAGATCCCAGCAAGATTTT
TG TAGATATGTAAAAGATTATTCAAAAATGTAAAAGGAAGGACAAAGGACTAGAAT
AGATAAAACAAAATGGAGAAAGATTTAATAGGAATCACTGTAAGTGAATTTTAAGAC
ATACAGAACAATAATAGAACTGCTTGTATTAGTCCATTTTCACGCTGCTGATAAA
GACATACCTGAGATTGGCAATTACAAAGGAAAGANGTTTATTGGCTTACAGTTCCC
15 ATGGCTGGGGAGGCCT

Sequence ID 718

CTCCTCTGGGTTGAAACCCGGGCGCCGCCAAGATGCCGGCTTACCACTCTTCTCTC
ATGGATCCTGATACCAAACCTCATCGGAAACATGGCACTGTTGCCTATCAGAAGTCA
20 ATTCAAAGGACCTGCCCCCAGAGAGACAAAAGATACAGATATTGTGGATGAAGCCA
TCTATTACTTCAAGGCCAATGTCTTCTTCAAAAACCTATGAAATTAAGAATGAAGCT
GATAGGACCTTGATATATATAACTCTCTACATTTCTGAATGTCTGAAGAACTGCA
AAAGTGCAATTCCAAAGCCAAGGTGAGAAAGAAATGTATACGCTGGGAATCACTA
ATTTTCCCATTCCTGGAGAGCCTGGTTTTTCCACTTAACGCAATTTATGCCAAACCT
25 GCAAACAAACAGGAAGATGAAGTGATGAGAGCCTATTTACAACAGCTAAGGCAAGA
GACTGGACTGAGACTTTGTGAGAAAAGTTTTTCGACCCTCAGAATGATAAACCCAGC
AAGTGGNGGGCTTGCTTTGTGAAGAGACAGTTCATGAACAANAGTCTTTCAGGACC
TGGACAGTGAAGGGAGCCCGGGCAGCCA

30 Sequence ID 719

CGNGGCCGCGTNAACTTTTGATCGTCAGCTGGGGCTGGCAGGCACCTAAATGGGAA
GGGTGATAGCAGTGTGTTGGGGGGAGTTTAGGGAACGGTCCTCTACCGATAGAGGC
AGCANCTCATTGGAATTTCCCTCCTGAAGTTGTCTTGCCCCTTGAATCCTGCAGGAA
GGCTGGCAAATGGCCATTTCCCTTCCACTTGAATAGAGACCCATAACTCAAGTATC
35 TGCCCTTAAGACACCACAGGACTGTTCTTCGCGGGCCCTGCCCTTGGAATTTGGGAG
AGGCAGTCCANCTCACCCAACCTAGGCTCTGCANGGGGACCANGAGGGATGGGTGT
GTCCACAGGACCAGCCAGACTGATGAGGGATGCGGCAAGCATATTCTCACCACCTT

- 214 -

CTTTCACGTTTACAACANACCAGCNTTCCCTGTGTGGCAGGGGTACATTGGTCAC
CGAGGACCTANAATCATGGAGTGCTCTGGGGATCCGGGCTTGGA

Sequence ID 720

5 TCAGTGTTGAATTTTGTGTCAGACACTTTCTCTGCATCAATTGGTATGACCATGTGAT
TTTTTTTCTGTAGCCTGTTAATATGGTTAATTTTCAAATATTGAGCTGATTAATTT
TCAAATATTGAGCTCTCCTTGCATCTCTGGAATAAGTACCACTTGGTCGTGGTATA
TATTTCTTTTAATATATTGCTGAATTCTGTTTGATCATGTTTTCTTAAAGACTTTC
GTGTCTGTTTTTCATGATAGATACTGGTCTATAGTTTTGTTGTAATATCTTGGTTTG
10 ATTTTGATATCAGGATAATGCTACCTTAATAGAATGAATTGGAGCCAAGTATGGTG
GCAAATGCCTATAGTCCTAGCTACTCAGGAGGCTGAGGTGGTGGGGACTGCTTGAC
CCANGAGTTCAAATCTAGCTTGGGCAATGTAGCAAGAC

Sequence ID 721

15 TAGAAGGAATGACTATTCATGTCCAAAGTGAATGGTTTTGTGCAGTGAACAACACA
TGGCGAGGTACTAACTGAGAACTTTTTTCATGCTTTATGCCTACCTCTTGTAGTTG
TTGCAGAGCAAATATAAATTGTAATAAGATAGCTAGGCCTTGCAGAAACAACAGA
AAAACCTAAAAAAAATGATATAAGAGCTGGAGTCTAGTATTTATATGAATCTGTG
AGAGATAATTTTTTTTGGTCTCACTGCAATGAACCAAAGCGGCTGAGTTTGGTTTT
20 TAATTGTAGCCATGTATTGAAGGCATCTTTTGTACCAACTCTTGTGGTTCTGTCT
TGAACCATTTGTTAATCACTGTGCTGTAATTAGTATAGCTAAATCTTTTCCTTCCTT
GCTCCTCCCCCAGCCCACCCCGTCTTCCCTTAACATTTTTTTCAGGGGGGGTTGGGA
GTGGTTTCATTTTAATGTGAGTGGATGTTTTGATAGTTGTAAGGAAAAAATGCATT
TCAGACACATTTACACATGAGCTATTTTCTTACACAGTATGTCTTATTGGTAATA
25 AGAATGTAATTCAT

Sequence ID 722

CNTTCCNTAAGAATACAAAAAATTAGCTGGGCGTGGTGGCAGGCGCCTGTAATCCC
ATCTACTCAGGAAGCTGAGGCTGGAGAATCGCTTGAACCCGGGAGGCGGAGGTTGC
30 AGTGAGCAGAGATCACGCCACTGCAGTCCAGCCTGGGCAACAGTGCGAGACTCTGT
CTCAAAAAAAAATAAATAAATTACCTGGGTGTGGCAGCGCGTGCCTGTAATCCCA
GCTACCCAGGAGGCTGAGGCAAGAGAACTGCTTGAACCCAGGAGGCAGAGGTTGCA
TGGAGCTGAGATGGCGCCACTGCACTCCAGTCTGGTGACAGAGTGAG

35 Sequence ID 724

CTCTCTACTAAAAATACAAAAAATTAGCTGGGCACGGNGGTGCATGCCTGTAAACCC
AGCTACCAGGTACTCGGGAGGCTGAGGCAGGAGAAATCGCTTGAACCAGGGAGTCCG

- 215 -

AGGTTGCGGCGAGCTGAGATCATGCCACTGCACTGCGGCCTGGAGACAAGAGCAAG
ACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGACNTCACCTAATTGCAG
NGNGNGGACCTTATTTGGCTNTTAATTCAAACCTATTAAAAATGTGAACN

5 Sequence ID - 726 nt: 260

CGGGGTCTGTACCGGGCTGGCCTGTGCCTATCACCTCTTATGCACACCTCCCACCC
CCTGTATTCCCACCCCTGGACTGGTGGCCCCCTGCCTTGGGGAAGGTCTCCCCATGT
GCCTGCACCAGGAGACAGACAGAGAAGGCAGCAGGCGGCCTTTGTTGCTCAGCAAG
GGGCTCTGCCCTCCCTCCTTCCTTCTTGCTTCTCATAGCCCCGGTGTGCGGTGCAT
10 ACACCCCCACCTCCTGCAATAAAATAGTAGCATCGG

Sequence ID 727

CTGAGTNTAGAAATGATGCCATTAATACTGATTGCAAAAACATTACAACCTCAGTAC
TGCAGCTTTCATTCAAATAGGTTATATGTATAAACTGAGTTCAACAATATTGTATT
15 TGAGATGGTAAAGTTAAAGAAATGCAATAATGTAAATAATACTTAAGAAAATAAGA
TCTCAGGAACTGTATATACTCTGTACTTTTATGCAACTTTATCAGATCATTTTCAG
TATATGCATCAAGGATATAGTGTATATGACATGAACTTTGAGTGCAAAAACCTGTAC
TATGTACCTTTTGTATTATTTTGCTGTCAACATCTAAATAAAGGTTTTTTTGT
TTTTTGTTTTTTTAATTGTTTTGTTTTTAAAGATTGTTTTAATTAATTA
20 ATTTTGTATTAAACAATTGTTTTAATTGTTTTTAAAGTCGCCAGGCTGAGGCAGGT
GAATCACAAGCTTAGGAGTTGGAGGCTAGCCTGCCAACATGGTGAAACCCCGTCTC
TACTAAAAATACAAAAAAATTAACCTGGGTGTGGG

Sequence ID 728

25 CCCATCTGCACCAGTACACAGGCAGGCATTATCATTCTTCACCTACTTTTTTAAATA
GTGGCAACTTGGGATTCATTCTGGTGATTCTGAACCTTGCCTCATAGCTTAAAGTA
TAAAAAAGATTCAAGAGCAGTGAGGTTTGTTCTTTCCAGTGAATGGTGGACTGAGT
GGTGCGAGGTGGAGGGCTAACAAGAGGAAAGAACTACATTCTTCAGAATACAGTGA
TGAAAATTCATTTTGAACTCAAATATTTTCATTTTGGATATTCTCCTGTTTTTAT
30 TAAACCAGTGATTACACCTGGCCATCCCTCTAAATGTTCTAGGAAGGCATGTCTAT
TGTGATTTTGATGAAGACAGAATTATTTTTCTCTGTAGAAACACAGATAACCACTTT
ATCAGGGGAAGTTAGTCAAATGAAATGGAAATTGGTAAATGGACAAAAGCTAGCTA
GTAAAAAGGACGACCCAGCAACATGCTTTAACCCCATTTGTATGTTTGTGGAAAGAG
CATAGTTTAAACATCTTGAGAAATTTGGGACATAAAAGTTTTTCATNGGTAGACAGTT
35 CATGGCAGTATATGAATTGACATAATGGAAATAATCTGATTTTATTTTTTACA
ACATCCTTTCCCC

- 216 -

Sequence ID - 736

nt: 641

5 GGAATTCCAAGTGCTTGGGGATAATGATACCTCTGACCTTTCTTCCTTTTGGGAAG
TACTTGAGTGTGCAGCTGCATGAGGCCTCAGCAGGAGAGAGATTTTAGGTCCAAGA
AGCTATAACCAGTAGGACAAGGCAGGAAAATACTACACTTTCAGGATCAAGCCCCCTC
TGACTCTCATTTGGAAACTGGATGTTTGCTAAGCACCTGCTTCTTAAGGATGCCGA
GGGATTTAATGATACTCCCAGAAACCTGGAGAGATTAATGGGGCCTATGGAGAAGT
GCTCTGAACTCAGTGTGTTGGGACTTGAATAAAATTAACCATTGTCATGTTTTTCAGAA
CAACTAAGCTGTTTTATATTTTCATGTGCATGAAAGCCCTAGAACTAAGTTGTGTTA
TTTCCAGAAATGAAATAGATCCCACAGTTAGATGATGTGGCCATTAGGAAGTACCA
10 AATTTATAAAAATCACTGGAGGTCTGTCTGAGCAGTACCTAATAAAATATAGTATA
CTGAAAGTGAACAGATACTTTGTCTCTTTCTTTGGCTGCTTGATCTTTATCTGTGT
CTGCCGTACAGTGCACCCTTAAAGTATTCTACACCAGTGCTTCTCAAACCTGGAAAT
GTGCATGTAAGTCACCCANGGGTCT

15 Sequence ID 739

TGCATGCCCATAGTCCCAGCTATTTGGGAGGCTGAGGCAGGAAAATCGCTTGAACC
CGGGAGCCAGAGGTTGCAGTGAGCCGAGATCGCACTCCAGCTTGGCGACAGAACAA
GACTCTGTCTCAAAAAAAAAAAAAAAAAAGAAATCTTGGGATCCTGAACCCCTTACTC
GAAGGGCTAAGGTAGCATCTCAGCATGTCTTATTCGAGACTTCGTANAACCAGACC
20 TGCTGTTTGTAGATGTTAATTAATCAAACCTTTCTCTACTCATTCTGGACCAGTTA
AGGTTTTCTCCTTCTCCGTATGAGTTTTGATTTTCGTCCTCCTTGGTTGGAGATCA
CACTTTGGTCTGCTGCTAAGTTGGATGCCTCCCACTGTCTTTCCCTAAGTCTAGGG
CTTCANACCCCAGTGTGGGGAGAGGGACTTTCGTTTCCTGCCCCCTCACCACATCAG
ACACAGGCAGGCAAGAATAAGATGGCCAAAAGGCCGATGAACTTCTTGACCTAGCC
25 TGGGACATTACCTGTTACTAGGTGGACTTCACTGCCTGTGAATGGAAGCTGAAGGG
CTGTTTTTTTGGTTTGTATTTGGACAGGCCAGGCTTANAGAGGGAGAGAACTGGGC
TACTCTTCAGCAGTGATCTTTAAAATGCC

Sequence ID 747

30

CAGAGTGCAAGACGATGACTTGCAAAATGTCGCAGCTGGAACGCAACATAGAGACC
ATCATCAACACCTTCCACCAATACTCTGTGAAGCTGGGGCACCCAGACACCCTGAA
CCAGGGGGGAATTCAAAGAGCTGGTGCGAAAAGATCTGCAAAATTTTCTCAAGAAGG
AGAATAAGAATGAAAAGGTCATAGAACACATCATGGAGGACCTGGACACAAATGCA
35 GACAAGCAGCTGAGCTTCGAGGAGTTCATCATGCTGATGGCGAGGCTAACCTGGGC
CTCCCACGAGAAGATGCACGAGGGTGACGAGGGCCCTGGCCACCACCATAAGCCAG
GCCTCGGGGAGGGCACCCCCTAAGACCACAGTGGCCAAGATCACAGTGGCCACGGC

CACGGCCACAGTCATGGTGGCCACGGCCACAGCCACTAATCAGGAGGCCAGGCCAC
CCTGCCTCTACCCAACCAGGGCCCCGGGGCCTGTTATGTCAAAGTGTCTTGGCTGT
GGGGCTAGGGGCTGGGGCCAAATAAAGTCTCTTTCTC

```
5      Sequence ID - 757                               nt:       583  
GAACCCTGCGGAGGGACTTCAATCACATCAATGTAGA ACTCAGCCTTCTTGGA AAG  
AAAAAAAAGAGGCTCCGGGTTGACAAATGGTGGGGTAACAGAAAGGAACTGGCTAC  
CGTTCGGACTATTTGTAGTCATGTACAGAACATGATCAAGGGTGTTACACTGGGCT  
TCCGTTACAAGATGAGGTCTGTGTATGCTCACTTCCCCATCAACGTTGTTATCCAG  
10    GAGAATGGGTCTCTTGTTGAAATCCGAAATTTCTTGGGTGAAAAATACATCCGCAG  
GGTTCGGATGAGACCAGGTGTTGCTTGTT CAGTATCTCAAGCCCAGAAAGATGAAT  
TAATCCTTGAAGGAAATGACATTGAGCTTGTTTTCAAATTCAGCGGCTTTGATT CAG  
CAAGCCACAACAGTTAAAAACAAGGATATCAGGAAATTTTTGGATGGTATCTATGT  
CTCTGAAAAAGGAACTG TTCAGCAGGCTGATGAATAAGATCTAAGAGTTACCTGGC  
15    TACAGAAAGAAGATGCCAGATGACACTTAAGACCTACTTGTGATATTTAAATGATG  
CAATAAAAGACCTATTGATTTGG
```

```

Sequence ID - 758                               nt:           424
CTTGGCTCCTGTGGAGGCCTGCTGGGAACGGGACTTCTAAAAGGAACCTATGTCTGG
20 AAGGCTGTGGTCCAAGGCCATTTTTGCTGGCTATAAGCGGGGTCTCCGGAACCAAA
GGGAGCACACAGCTCTTCTTAAAATTGAAGGTGTTTACGCCCGAGATGAAACAGAA
TTCTATTTGGGCAAGAGATGCGCTTATGTATATAAAGCAAAGAACAACACAGTCAC
TCCTGGCGGCAAACCAAACAAAACCAGAGTCATCTGGGGAAAAGTAACTCGGGCCC
ATGGAAACAGTGGCATGGTTCGTGCCAAATTCCGAAGCAATCTTCCTGCTAAGGCC
25 ATTGGACACAGAATCCGAGTGATGCTGTACCCCTCAAGGATTTTAAACTAACGAAA
ATCAATAAATAAATGTGGATTTGTGCTCTTGT

```

Sequence ID - 764 nt: 626

30 GATTTTTTTTTTTTTTTTTTGGAGATGGAGTCTTTCTCTGTCGCCCAGGCTGGAGTGCA
GTGGTGAAATCTCGACTCACTGCAACCTCCGTCTCCTGGGTTCAGCAATTCTCCT
GCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCACCAGCCACCACGCCCGGCTAAT
TTTTGTATTTTGTAGTAGAGACAGGTTTTTCACCATGTTGGCTAGGCTGATTTTGAAC
TCATGACCCCCAAGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGAATTACAGGT
GTGAGCTACCACTCCCAGCCAATGATTACATTTATAAGGTAAAATAACTTGTGCCA
35 ATCTGTACAAGTGAATTCAGATTTAAAATTTTAATTGTAAAAAGATATCCAGGTGA
TATTTCTCCCTGAATAATTTAGTTTCCTTTTCTATTTCTTGATATAAAAGTACTCA
GCATTGAAGTAATTGCTATCTTCACATTTCTTCCTATTTGAGCTGTCTAAATAAGT

- 218 -

AGTCCTACATATTTTCCCCCAACACAAAAAACCCAGAAAAGAATTATTTTATACT
GGATTTTTTTTGGTTGTAGCAGGAACCTAAAGGNGCCAATTGTAACATGCATGTTCT
TTTTGGCAAA

5 Sequence ID 766

GTCCATCCTGCAGGCCACAAGCTCTGGATGAGGAACTTGAGGCAAGTCACCAGCCC
CTGATCATTTTCGCCTAAAAGAGCAAGGACTAGAGTTCCTGACCTCCAGGCCAGTCC
CTGATCCCTGACCTAATGTTATCGCGGAATGATGATATATGTATCTACGGGGGCCT
GGGGCTGGGCGGGCTCCTGCTTCTGGCAGTGGTCCTTCTGTCCGCCTGCCTGTGTT
10 GGCTGCATCGAAGAGTAAAGAGGCTGGAGAGGAGCTGGGCCCAGGGCTCCTCAGAG
CAGGAACTCCACTATGCATCTCTGCAGAGGCTGCCAGTGCCCAGCAGTGAGGGACC
TGACCTCAGGGGCAGAGACAAGAGAGGCACCAAGGAGGATCCAAGAGCTGACTATG
CCTGCATTGCTGAGAACAAACCCACCTGAGCACCCCAGACACCTTCCTCAACCCAG
GCGGGTGGACAGGGTCCCCCTGTGGTCCAGCCAGTAAAAACCATGGTCCCCCCT
15 TCTGTGTCTCAGTCCTCTCAGTCATCTCGAGCCTCCGTTCAAAATGATCATCATCA
AAACTTATGTGGCTTTTTTGACCTTTGAATAGGGAATTTTTTAAAATTTTTTAAAAA
TT

Sequence ID 768

20 CCAGCGCAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAA
GTTTTTTTCTCTTTGAAAGATAGAGATTAATACAACCTACTTAAAAAATATAGTCAA
TAGGTTACTAAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAA
GATTTTAAGAGAAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAA
AAGGTTTCTAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAAT
25 GTATTTAAAGAAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAG
AAGGGCAATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTT
TAAAAGTTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAGAGATTAAA
CCGAAGGTGATTAAAAGACCTTGAAATCCATGACGCAGGGAGAATTGCGTCATTTA
AAGCCTAGTTAACGCATTTACTAAACGCAGACGAAAATGGAAAGATTAATTGGGAG
30 TGGTAGGATGAAACAATTTGGAGAAGATAGAAGTTT

Sequence ID 773

GAGGAAAGGGGAGTTAATATTTAGTGGACAGAATTTTCAGTTTTTACAGATGAAAAGA
GTTCTGGAGATAGACGGTGTTGATAGTTGCACAGCAGTGTGAATGTGCTCATTGTT
35 ACCGAACTTAAAAATGTTTAAACATAGTATTATGTGATTTTTTATTTTGCCACTTAAA
AAAAAGAATGAAGTACTGATACATGCTACAACATGGGTGAGCTTTAAATACATTC
TGCTCAGTGAAATAAGCCAGATGCAAAAGATCACATATTATATAATCCACTTATAC

- 219 -

GAGATACCTAGAATAGGCAAATTCATAGAGACAGAAAGTAGAATAGTGGTTCCCAG
GGGCTGGGGACAAGGGGGCAGTGAGAGATTGAGAGTTATTATTAATGCGTACAGAG
TTTCAGTTTGGGCTGATAAAAAAGTTCTGAAGATGGATGGTGATGATGGTTGTACA
TCAATGTGAGTGTAATTACCGCCACTGAACTGCCCTTAAAAACGTTTAAAAGAGTA
5 AATTTTATGTTGNGTATATTTTACCATAAT

Sequence ID 776

TTTTTTTTTTCATAAGAGGCAAGTACAAGAAAAAGCTTAATTACTTTAACTTCTAAG
TAGTTTGGAATCTAAATAAATAGGAGTTACCAAATATATGCGCTTCTGTGAATAGT
10 TTTCCCCCACATGTTTATTTATATTTTTTGCATCTCATCAAACCTAACAGATTCTAA
AGTCTCTGGTGATAATGACAATATCTGCTACGGAGAGACTAGCCTGGGGGAAGAGG
ATCTCCCTGAACAAGGATAGCGGAGTTGCTGCAGCTTTCAAATGAAGCTGGACATT
TAGCTGCGGGGGTAGCACCCCTTTGATCAAGGCAGCCCAAAGATGAGTTTCAGGGAT
GGGACTGACAGAAGAGAAAAAGTTCTTCCCAGCCCTTTCTACTTTTTTCTCTTTGTTT
15 CTCAGGCTTCTGGCCGTCTTCAGTTTTTCACAAGTTTCACTCTCAACCCTAAACAGT
ACTTCTGTGAAGTACCCTTTGGCCCTCGTTTTTCAGCTCCTAAACTCACCTGGAAA
TAGATGTCAATCTAATTTTGGGTCTGACTAGTGCAGTAGGCATTTTTTGGTGA

Sequence ID 782

CTCACACAGAACAAAAATGAATGAGTGTGGCTGTGTGCCACTATCACTGTGTCTAC
AAAAACAGCCAGTGGGCCTGATTTGGCCCTTGGCTGCAGTGCGCCCGTCTCTGTTT
TTGAGGAATAAAATCGCATCATTTTCATATGGCTAATGCAATTTTTTTTCCCATCTGG
AAGCAACATCTGATTGGACTCATCTTGATGGTGCTTGTTACAGTCTCTGTAAATG
GGAGAGGGTCCGAGAATAGCTCTTCCTGTTTTTCATCAGGACTGTTTTTAGGGATGG
25 CAAAGAAGTCAGTGTGTCCAGCCTGTGTCCTCCTCACCACGTGGCTGATTCCTGAA
TCTGCATGTGCANCACNTGCCGTTGTCTGGGGCATGATCTGTGTGA

Sequence ID - 785

nt: 556

CTTTTCTCTGGGTATAGATTTACCCTAGCACCTATCTCATTATATTGAATTTTCCA
30 GCATATTTAAATAAACTATTAATTAGTCACACTATTTCTTAAAAGTCACACTATCA
ACTAATCGTGACCGCAATTATCTAGGGGTGATAATCTGCTGAGTCTACTCTTTAAA
TACACTGGGACCCAGCATATTGAGTTATATTGGCACAGAACTTCACTCTGGGTAT
AGATTTACCCTAGTACCTTGCCGGCAGGATCCTATTATTCATGGTTGTACAAGCAA
GGTTCAGGGAAGAGGCTGGCACAGAGAAGGTACCTGGTAACTGTTGTTTGAGGCTG
35 AATTCAGCTCAACTCAGCTCCAGTAGAGATGGTGTCCCCTTCTCTACCGTGTTGAG
ATAGTGTGCAGTCCCTTCCTAAGGGCTGTTACCCACCGCAATAGGACTTGTCTAGCT
TCAACTTTTAAATTTCTCTGCTCCCGCTGGGACCCACCCGCTTCAAAAATCATCAT

- 220 -

GGNGGNTTTAGCACCAATTTAGTAAACACAAACTGTCTGAAATATTTTGGAT

Sequence ID 796

GAACATTCAAGATAGTGAGAGGAAGAAAAAGATATGGCTGTACGGGACCGAGGTCT
5 CTTCTATTATCGCCTCCTCTTAGTTGGCATTGATGAAGTTAAGCGGATTCTGTGTA
GCCCTAAATCTGACCCTACTCTTGGACTTTTGGAGGATCCGGCAGAAAGACCTGTG
AATAGCTGGGCCTCAGACTTCAACACACTGGTGCCAGTGTATGGCAAAGCCCCTG
GGCAACTATCTCTAAATGCCAGGGGGCAGAGCGTTGTGACCCAGAGCTTCCTAAAA
CTTCATCCTTTGCCGCATCAGGACCCTTGATTCCTGAAGAGAACAAGGAGAGGGTA
10 CAAGAACTCCCTGATTCTGGAGCCCTCATGCTAGTCCCCAATCGCCAGCTTACTGC
TGATTATTTTGGAGAAACTTGGCTTAGCCTTAAAGTTGCTCATCAGCAAGTGTTGC
CTTGGCGGGGAGAATTCCATCCTGACACCCTCCAGATGGCTCTTCAAGTAGTGAAC
ATCCAGACCATCGCAATGAGTAGGGCTGGGTCTCGGCCATGGAAAGCATACTCAG
TGCTCANGATGATACTGGCTGTCTGTTCTTAACAGAACTGCTATTGGAGCCTGGAA
15 ACTCAGAATGCAGATCTTTTGTGAACAAAATGAAGCAAGAACCGGAGACNCTGAAT
AGTTTTATTTCTGTATTAAAAACTGNGATTGGAACAATTGAAGA

Sequence ID 801

CCACTCCACCTTACTACCAGACAACCTTAGCCAAACCATTTACCCAAATAAAGTAT
20 AGGCGATAGAAATTGAAACCTGGCGCAATAGATATAGTACCGCAAGGGAAAGATGA
AAAATTATAACCAAGCATAATATAGCAAGGACTAACCCCTATACCTTCTGCATAAT
GAATTAAGTAGAAATGAGGATTCTGACCTTGACTTTGATATCAGCAAATTGGAACA
GCAGAGCAAGGTGCAAAACACAGGACATGGAAAACCAAGAGAAAAGTCCATAATAG
ACGAGAAATTCTTCCAACCTCTCTGAAATGGAGGCTTATTTAGAAAACAGAGAAAAA
25 GAAGAGGAACGAAAAGATGATAATGATGATGAGTCAGGTAAAAGTTCCAGAAATGT
GAACAACAAAGATTTTTTTTGATCCAGTTGAAAGTGATGAAGACATAGCAAGTGATC
ATGATGATGAGCTGGGTTCAAACAAGATGATGAAATTGCTGAAGAAGAAGCAGAAG
AAGGAAGCATTTCTGAAATATGAATGAAAAAATTACATCTTTAGAAAAAGAGTTA
TTAGAAAAAAGCCTTGGCAGCCGTCNGGGGGAAGTGACGCACAGAAGAGACCAGAG
30 AATAGCTTCCTGGANGAGACCCTGCACTTTACCCATGCTGCTGGATGG

Sequence ID - 808

nt: 641

CCGGGTTTTAGTATTTAACCAAGAGCCTTTTAAATATTGAAAACCCATAGTTCAGA
AAATGTTAGTATTGCTGCCCTTCTTCACATAAATTTTTTTTTTAAATTATACTATTA
35 TTTTGCTTAATTTTATATTGGGTAAAACAACCTTCAAGAAGGTAACTAGGAAAG
AAGACCTTTTTGTTTTATTTTTACTATTTATATATAGAAGACAAATCAGCATTTGG
TGATAGTTTTACATGACCAGTTATCAAACGGTCATAGTATGAAGTGTGCAGTTGTT

- 221 -

CATTATTAGTAAATTATGTTTGATTTTTTAAACTATTTAGTACTAATAGTTGAGATG
AAAACCTGAAGAAAAATGCCAATGTGACGTTTGTGTATAGCTAGCCTTAAAAAACTT
CCCATGTTTTTTAGGTGACTTTTTTCCCCCTCTTAGTACTCTGGAGAAACAATGAAG
ATGGGCCATCTCAATTCCAGATGTAAACAAAAAGTAATTTTTTATTTCAACATTTAA
5 TGTAACCTGCTATTATTGNGGATTCTTGNCTTGNGTATTTTCTTTCCCTTATTCAAG
TAATATAGAATAACTTTCCTTAAAATGATTTGATCCAAGATACGTCATTTCTGTAT
TGGCAAAATGCCNCTATTAAAGTGT

Sequence ID - 814 nt: 132
10 GTTAAAGTGATACATTTTTTATACCAAATGTGTTTATTTTTTTGTGCAAGTAATCCT
TAAAATTGCAATTGTATTAGGTGTAAAATAAAGTTTTTTAAAAAATTAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

Sequence ID 817
15 GACAACCTTAGCCAAACCATTTACCCAAATAAAGTATAGGCGATAGAAATTGAAAC
CTGGCGCAATAGATATAGTACCGTAAGGGAAAGATGAAAAATTATAACCAAGCATA
ATATAGCAAGGACTAACCCCTATACCTTCTGCATAATGAATTAACCTAGAAATAACT
TTGCAAGGAGAGCCAAAGCTAAGACCCCCGAAACCAGACGAGCTACCTAAGAACAG
CTAAAAGAGCACACCCGTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGG
20 CGACAAACCTACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCA
ACTTTAAATTTGCCCCACAGAACCCTCTAAATCCCCTTGTAATTTAACTGTTAGTC
CAAAGAGGAACAGCTCTTTGGACACTAGGAAAAAACCTTGTAGAGAGAGTAAAAAA
TTTAACACCCATAGTAGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTC
AACACCCACTACCTAAAAAAATCCCAAACATATAACTGAACTCCTCACACCCAATT
25 GGACCAATCTATCACCCCTATAGAAGACTAATGTTAGTATAAGTAACATGAAAACAT
TCTTCTNCGCATAAGCCTGCGTCAGATTAAAACACTGAACTGACAATTAA

Sequence ID - 821 nt: 370
AAAGAGCTCCCAAATGCTATATCTATTCAGGGGCTCTCAAGAACAATGGAATATCA
30 TCCTGATTTANAAAATTTGGATGAAGATGGATATACTCAATTACACTTCGACTCTC
AAAGCAATACCAGGATAGCTGTTGTTTCANAGAAAGGATCGTGTGCTGCATCTCCT
CCTTGGCGCCTCATTGCTGTAATTTTGGGAATCCTATGCTTGGTAATACTGGTGAT
AGCTGTGGTCCTGGGTACCATGGCTGGTTTCAAAGCTGTGGAATTCAAAGGATAAA
TTAATGAAGAAAACAAGCGGAGCTGAAGAAGAAAGTACAATATGGTGCTGTCTTCC
35 TAATGAAATAAATTCACTAAATGGACATTAAAAA

Sequence ID 825

- 222 -

AGACTCGAGCAAGCTTATGCATGCATGCGGCCGCAATTCGAGCTCGGCCACTTGGC
CAATTCGCCCTATAGTGAGTCGTATTACAATTCCTGGCCGTCTGTTTTACAACGTC
GTGACTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTT
5 GCGCAGCCTGAATGGCGAATGGAAATTGTAAGCGTTAATATTTTGTAAATTCGC
GTAAATTTTTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAA
TCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGG
AACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGT
CTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGT
10 CGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTAAAGCT
TGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAAAAGCCAAANGGAG
CCGGCGCTAGGGCCTGGCAAGTGTACGGGCACGCTGCGCGTAACCACCCACACCCC
GCCGNGCTTAATGCCCCNTTCAGGGCGCGTNCTGATGCCGNATTTTNTCTTACNCA
TNTGTGCNGGNTT

15

Sequence ID 833

TAAAATAATGGCAAAAAACAAACAAAAACAAGTTCTCTAAACAGAAAGGAAATTA
CTAAAGAAGGAATCTTGAAATAACAGGAAAGAGGAAATACCACAGTAGGCAACATT
ATGGGTAAATAAAACAGACTTTCCTTCTTTAGTTTCCTAAAATATGTTTGATGATT
20 AATGCAAAAATTACAATATTTTCTTATGTAGCACTAAAGGTATGTAGAGAAAATAT
TTAAGATAATTGTACTGTAAGCGGGAGATGACAGTGACATAAAGGCAACGTTTTTA
TACTTCACTCAAACCTTTATGTATTAATGTAATCCATAAAGCAACCAAAAAAGCTAT
ACTAAGTACATTCAAAAACACAATAGATAAACCACAAACAAATTCCTAAAGGATGTAC
AAGTAACCCACTGGAAGCTGCAAAAAATGTAAACAGAACTAAAAACAGAGAATAA
25 ATGAAAAATTAAAAACGAAATGGCAGACTTAGGCCCTAATATACAAATTATCACAT
TAAATATAAATGGTCTAAATACACCAACTGTAAGACAGAGATTAGCAAAGTCGATT
TAAAAACATGACTCAACTACGTGCTGTCTACAAGAACTCACTTCAAATATACCAA
GATAGGAAGGTTGAAAGTAAAACGATGGAAAAAGATGTATCATGTGAACATTAATC
AAAGGAAAGCAGGGGTGGCTATATTAACATCAGGTAAAATAAACTTT

30

Sequence ID - 837

nt: 603

TGAGGNTGGTCATGATGCANAAGCTACTCAAATGCAGTCGGCTTGTCCTGGCTCTT
GCCCTCATCCTGGTTCTGGAATCCTCAGTTCAAGGTTATCCTACGCGGAGAGCCAG
GTACCAATGGGTGCGCTGCAATCCAGACAGTAATTCTGCAAACCTGCCTTGAAGAAA
35 AAGGACCAATGTTTGAACCTACTTCCAGGTGAATCCAACAAGATCCCCCGTCTGAGG
ACTGACCTTTTTTCCAAAGACGAGAATCCAGGACTTGAATCGTATCTTCCCCTTTT
TGAGGACTACTCTGGATCAGGCTTCGGCTCCGGCTCCGGCTCTGGATCAGGATCTG

- 223 -

GGAGTGGCTTCCTAACGGAAATGGAACAGGATTACCAACTAGTAGACGAAAGTGAT
GCTTTCCATGACAACCTTAGGTCTCTTGACAGGAATCTGCCCTCAGACAGCCAGGA
CTTGGGTCAACATGGATTAGAAGAGGATTTTATGTTATAAAAGAGGATTTTCCCAC
CTTGACACCAGGCAATGTAGTTAGCATATTTTATGTACCATGGNTATATGATTAAT
5 CTTGGGACAAAGAATTTTATAGAAATTTTAAACATCTGAAAA

Sequence ID - 839

nt: 71

ATTTATCTAATATTTGGTTTAATAAAATGTGAATAATGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAA

10

Squence 849

nt: 622

15

TTTTTTTTTATTTTTTGAGAATGGAGTCTTGCTCTGCCGTCCAGGCTAGAGTTCAG
TGGTGCGATCTCAGCTCACTGCCACCTCACCTCCTAGGTTCCAGAGATTCTTGTGC
TTCAGCCTCCTCAGTAGTTGAGAATACAGGAACACGCCACCACGCCTAGCTAATTT
TTGTATTTTATAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACCTC
CTGGCCTAAGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTATAGGCGTG
AGTCATTGTCCCCAGCCGGATGTTTTTCATCTTGATTTGCCTTAGTTTCTAAATCTC
ATCCTCTCCATTTTCTCCTGTTAGTAGTCACAGAGAACCAAATTCTGTCAAGTTAT
GAAACTAAAGTCTCTCTTCCACAAGTCTTCCTGTGTTCTGCCTCAAGTGAACCTGA
20 AAGAACATCAGTTTGTGGGAAGGTTGAAGACCGAATGATCTGCTGGGAAATCACTG
AGGCATTGCCATTCTCTTGAGGAATTTCATTTTTCATCGAAGTTTCGGTTTATATCC
CTTTCTTGGTGAGTACTATTGCTGTTATGTAAATTAAATGAGTCGTCATCCTTCTT
NTGAGC

25

Sequence ID - 860

nt: 501

30

GTGAAATCACTTTCATGGATTATTAATGGATTTAAGAGGGCATCAATCAGCTCAAC
TCAAGATTTTCATAATCATTTTTTAGTATTTAGATTGTGCCTCAAAGTTGTAGTACCT
CACAATACCTCCACTGGTTTTCTGTTGTAAAAACCTTCAGTGAGTTTGACCATTGT
GCTCTTGGCTCTTGGGCTGGAGTACCGTGGTGAGGGAGTAAACACTAGAAGTCTTT
AGTACAAAACCTGCTCTAGGGACACCTGGTGATTCTTACACAAGTGATGTTTATATT
TCTCATAAAGAGTCTTCCCTATCCCAAGGTCTTCATGATGCCAGTAGCCATATATG
ATAAATTATGTTTCAGTGATAACTTAGTTATCAGAAATCAGCTCAGTGGTCTTCCCC
GCCATGATTCACATTTGATGAGTTTTTAAAAATCAAAGTGATTTTGAAAATCTCTA
ATGGCTCAGAAAATAAAAAACATCCAGTTTGTGGATGACTATATTTAGATTTCT

35

Sequence ID 864

TTGTGTTTTTAGGACTCCTTATCTAAATTAAGGCAGAGAAGTTACAGTATTTATAT

- 224 -

CTGCATTAAATCTCAATTCCAGAAAAACCTTTTGAAAAATTATTTAATCCTCTGGA
AACTATTGATATGATACAGGAGAAATTTTCAGAAGTTTATTGAATAATTTAATATC
ATTTAATAGGACACTCTGGCTTGTATATAAGCAGATACGTTACTCAGACTTCTTGG
CTGTACTCTAAAATAATATATGTACTAGTCTCCTAAATATTACTAGCTCACCTTTC
5 AAAATGCATACTAATATTTCAATGTCTTTCTTCAATTTGAAAAGCTCTTGAATATC
TACTTGTGATAGCCCTAAGAGCTGAGATAATTATTTCCAGGAGGTTGAATCCCTGA
TTCTTAACTGTTTCAGCAATGCATAAGCAAGAGAGAATATGACATAAGAGGACCATT
TCTACATTAGCCATTTTTTTTTTCAACAAGATACCTATGTGAATACAGGGCACCTGGGA
GGGTAAGTGGAGGACTATTTCTAACTATATTTATAAGCACATACTGATATTGGTGA
10 ATCAAAACCTACAGCAGTGCTTCTCAGATGGGAAGGGAGACAATGTGTAAGGAGAT
CAGGAATTCATTAG

Sequence ID - 865

nt: 122

CCANAATCCACTCTCCAGTCTCCCTCCCCTGACTCCCTCTGCTGTCCTCCCCTCTC
15 ACGAGAATAAAGTGTCAAGCAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

Sequence ID 867

TTTTTTTTTTTTTTTTTTTTTTCAGAGTCACAGATATTGTATAGCTGAGGTAAGCATTT
20 TACAACCTTTTCAGACACAAGTAAGTACATAAATATTATTTTACAACCAACAATNTT
TAATATTTCCACATTGAANAATAGATGTGATAATTAAATCTTTTATAAGGTTTTTAA
AAAGACATGAAACATAAACCTAATTATACATAAAAGAAAAGAATTTTAAACAAGAG
CTTATTGNGATGACATTACTCATAACTTTTACCTTTTAAACCTTTTCTTGGGTAGC
TATTCAAAAGTAAAGACCACAAGTTTTGTGGCCCANATTTCTTATGTTTNGTATAT
25 TTAAGCTCTTTATTTATTGAACAGATGNGTCATTAATTCATTNGGAGCATTACTAT
TATCAGTAAAATTTGATTTTTTTTTTTCCCCTCAGTCATAGGTAAATCAGCTCCACCT
GGAATTTCTAAGGACCCAGTTTTAGTCAATATTTTCAAGTAATCATGACCTCAGAA
ATAGTCTTAATTAAGATAACAAATATTAGCCATCAAAATGGAACCAAGACAAGATT
CTAATGTTTGTAACAGTCAATCCATATTTATGAATATTAGCATATATTGGNGAAT
30 AGTTAAGGCAAAGGGTCTAGCAG

Sequence ID - 869

nt: 667

TTGTGTTTTTTAGGACTCCTTATCTAAATTAAGGCAGAGAAGTTACAGTATTTATAT
CTGCATTAAATCTCAATTCCAGAAAAACCTTTTGAAAAATTATTTAATCCTCTGGA
35 AACTATTGATATGATACAGGAGAAATTTTCAGAAGTTTATTGAATAATTTAATATC
ATTTAATAGGACACTCTGGCTTGTATATAAGCAGATACGTTACTCAGACTTCTTGG
CTGTACTCTAAAATAATATATGTACTAGTCTCCTAAATATTACTAGCTCACCTTTC

- 225 -

AAAATGCATACTAATATTTCAATGTCTTTCTTCAATTTGAAAAGCTCTTGAATATC
TACTTGTGATAGCCCTAAGAGCTGAGATAATTATTTCCAGGAGGTTGAATCCCTGA
TTCTTAACTGTTTCAAGCAATGCATAAGCAAGAGAGAATATGACATAAGAGGACCATT
TCTACATTAGCCATTTTTTTTTCACAAGATACCTATGTGAATACAGGGCACCTGGGA
5 NGGTAAGTGGAGGACTATTTCTAACTATATTTATAAGCACATACTGATATTGNTGA
ATCAAAACCTACAGCAGTGCTTCTCAGATGGGAAGGGAGACAATGTGTAAGGAGAT
CAGGAATTCATTAGTCACCTTTCAGATGGTTTAATGCATACAGCTGTACCG

Sequence ID 870

10 GGAGTTTGAGCAGATCCTTCAGGAGCGGAATGAACTCAAAGCCAAAGTGTTCCCTGC
TCAAGGAGGAACTGGCCTACTTCCAGCGGGAGCTGCTCACAGACCACCGGGTCCCC
GGCCTTCTGCTCGAGGCCATGAAGGTGGCTGTCCGGAAGCAGCGGAAGAAGATCAA
GGCCAAGATGTTAGGGACACCAGAGGAAGCAGAGAGCAGTGAGGATGAGGCTGGCC
CATGGATCCTGCTCTCCGATGACAAGGGAGACCATCCCCACCCCCGGAGTCCAAA
15 ATACAGAGTTTCTTTGGCCTATGGTATCGGGGTAAAGCTGAATCCTCTGAGGATGA
GACCAGCAGCCCTGCACCCAGCAAGCTAGGGGGGAGAAGAGGAGGCCCAACCACAGT
CTCCAGCTCCTGATCCGCCCTGTTCTGCCCTCCACGAACACCTTTGTCTGGGGGCC
TCAGCCGCCCCAGAGGCCTGACTTAGGGGTCTGGCTGTGGAAGGATGTGTGGCCTC
AAATGAGGACAGGGCTCCCGCCTTCACAGCCCTCGCCAGGGGTCTGCCCCAATCCT
20 GGCTGTCATCAGGCAAGGACGGGGTCTCAGC

Sequence ID - 871

nt: 642

GCAAGTCTTCAGTATGTACATTTATCCCCTAGAAGAAGAAAAATTAGTTGTGCATG
AAAAAGAAACATTAACTGCAAAGCTAAATGCTCACACTCTAAATCAGTGCTCTCCA
25 AAGTACAGCAGGCGGGAAAAGAAAATGGTAGATTTTTTTTCTTCCAATTACTTTAAC
TTATTCTTTTTTAATGGACACTTCATACATAAATATATTCACAATATATTAATATAT
ACATAATGTATAAGCATAACATATTGAATGTGCAGTCAAAAAATGTACTAATGGAAT
GCTCTACCAAAACAAGTTCACGTTTCATCTGTAAAATGGGAATAATATTTTTTAAAG
GCATACAGTCTGAACATTTTTTAGATTATTCATAAAATCTATTCAGAAAGTTAACT
30 AAAAAATTTAACGTATGCCTATAACAAATTTTGTACTTAATGTAATTGNTTTTCAT
CCTGAGATCTAATATCCTCGTTTTTAAGTAGAGCCACTTGTTTGCTACAGTTTAGT
CAAAACGTTAACATTAGATGGGTAAAGTAATATGAAATCTTTCTACTACTCCAAA
TAGAAAACAGAACATTAAAAAGATAAAAATTCAAACATACTTACCAGTAGATTTTC
AACTGNGCAAAAGCTCATTGCATGGG

35

Sequence ID 873

GTTTTCCACCGTGAAGAGAACATTTCTCTGGGAATGACAAAGCCCTCAGGAACNG

- 226 -

CTTTTATTTCTATTGGAAGATGCCCATCATACTTCTGGCAGGATAAAATGATAAAT
TTATTTATTCAACAGATGATACTCAATTCCCTGCTGTTTTACTAAAGGTTCTTTAC
GTTTTATAGAAGCTAAATTTACTGTCATAGAAATTGCAATTGTAGATGTTACTGTA
ATCTAGTCAGAATATCCTTATCCTTCTAAAATAAAACTAGTTAAAATTATTAACAT
5 ACGTACTGATATTAATTTTTTAAGTTTAATGCTGCCACGTGCTTCTGCTAAGAACAT
TTATCACTACAAGTGGCAGAAAATTCCAACTCATCAAAACCAAACCTGTTGCTTCT
TCCCTGCTTTTTTCAGAAAATGAGAAAGGATGACTTTATTCCAACATATTCTAAAAG
TATTCCAAGAACACTACCTTTATTCTAAATTCGTTATTTTCACAAAATAAAGGCTG
CAGATTGAAAGATAAAGGATTGCTATTAAAGAACAAAAGAAAACAAAACCGAGAGA
10 GAAGGAGAGCTAGGGAAATCCCTGCANAANAACCGAATANGGTCCCTCTATTCTGG
GCCGGGGCCTGAAACTATGAAACAGGCCAACACAGAATCTTGGCA

Sequence ID 875

CCTCTGACTCGCTCAGCTCACCCACGCTGCTGGCCCTGTGAGGGGGCAGGGAAGGG
15 GAGGCAGCCGGCACCCACAAGTGCCACTGCCCCGAGCTGGTGCATTACAGAGAGGAG
AAACACATCTTCCCTAGAGGGTTCCTGTANACCTAGGGAGGACCTTATCTGTGCGT
GAAACACACCAGGCTGTGGGCCTCAAGGACTTGAAAGCATCCATGTGTGGACTCAA
GTCCTTACCTCTTCCGGAGATGTAGCAAAACGCATGGAGTGTGTATTGTTCCCAGT
GACACTTCANAGAGCTGGTAGTTAGTAGCATGTTGAGCCAGGCCTGGGTCTGTGTC
20 TCTTTTCTCTTTCTCCTTAGTCTTCTCATAGCATTAATACTATTGGGTTCATT
ATTGGAATTAACCTGGTGCTGGATATTTTCAAATTGTATCTAGTGCAGCTGATTTT
AACAATAACTACTGTGTTCTGGCAATAGTGTGTTCTGATTAGAAATGACCAATAT
TATACTAAGAAAAGATACGACTTTATTTTCTGGTAGATAGAAATAAATAGCTATAT
CCATGTACTGNAGTTTTTCTTCAACATCAATGGTCATTGNAATGTTACTGATCATG
25 CATTGGTGAGGNGGTCTGAATGTTCTGACATTAACAATTTTCCAT

Sequence ID - 876

nt: 115

AAACTTTTGTGGCAACAGTGCACTAATTTGGATAATGTTTGTTCCCAATAAATTAA
GAGCCAAATTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
30 AAA

Sequence ID - 878

nt: 634

GCCAGGCTTTGTGAATTACAGGACATTTGAGACAATCGTGAAACAGCAAATCAAGG
CACTGGAAGAGCCGGCTGTGGATATGCTACACACCGTGACGGATATGGTCCGGCTT
35 GCTTTCACAGATGTTTCGATAAAAAATTTTGAAGAGTTTTTTAACCTCCACAGAAC
CGCCAAGTCCAAAATTGAAGACATTAGAGCAGAACAAAGAGAGAGAAGGTGAGAAGC
TGATCCGCCTCCACTTCCAGATGGAACAGATTGTCTACTGCCAGGACCAGGTATAC

- 227 -

AGGGGTGCATTGCAGAAGGTCAGAGAGAAGGAGCTGGAAGAAGAAAAGAAGAAGAA
ATCCTGGGATTTTGGGGCTTTCCAATCCAGCTCGGCAACAGACTCTTCCATGGAGG
AGATCTTTCAGCACCTGATGGCCTATCACCAGGAGGCCAGCAAGCGCATCTCCAGC
CACATCCCTTTGATCATCCAGTTCTTCATGCTCCAGACGTACGGCCAGCAGCTTCA
5 AAAGGCCATGCTGCAGCTCCTGCAGGGACAAGGACACCTACAGCTGGCTCCTGAAG
GAGCGGAGCGACACCAGCGACAAGCGGAAGTTNCTGAAGGAGCGGCTTGCACGGCT
GACGCAGGCTCGGCGCCG

Sequence ID 879

10 GTTGCCGGGTCTGTGATAACTCTGTTTAACATTTTGAGGAACTGTTGAATGGTTT
TTCACAGCAGCTGCCTCATTTTTTATTCCCATCAGCAGTACTTCTTGGTTCTAATA
CCTCCACGTTCTCGCCAACACTTGTTGTTGTCTGTAATTTTCGTTGTTAGCCATCCC
AGTGGGGATGAAGTAGTATCTTACTGTGGTTTTTCAGTTGCGTTTCCCTGATAATTA
ATGATGGTGAACATCTTTTCATGTTCTTGTGTTGGCCATTTGTATGTCTTCTTGGGAA
15 AAAAAAATGTCTGTTCAAATCCTTTACAAAGTATTTATTTTTTATGTCAACAATA
TAACCACTCAGTACACTGCTTTTTANACAATGATCTTTTAAAGGTTTGTTTACAAC
ATTTAGCACTTGAAATTTTAAGGTTATGCCCTCAAAAAAATTGCTGAGGGAGCTAA
GCTATGAAGATGCAAAGGCATAANAATTATACAATGGACTTTGGGGGAATCCAGGG
AAAGGGTGGGAGGGGGGTGANGGA

20

Sequence ID 881

TCGACTCTGATTTTTTTTTCTCCTTCCTCGCAGCCGCGCCAGGGAGCTCGCGGNGC
GCGGCCCCTGTCTCCGGCCCGAGATGAATCCTGCGGCAGAAGCCGAGTTCAACAT
CCTCCTGGCCACCGACTCCTACAAGGTTACTCACTATAAACAATATCCACCCAACA
25 CAAGCAAAGTTTATTCTACTTTGAATGCCGTGAAAAGAAGACAGAAAACCTCCAAA
TTAAGGAAGGTGAAATATGAGGAAACAGTATTTTATGGGTTGCAGTACATTCTTAA
TAAGTACTTAAAAGGTAAAGTAGTAACCAAAGAGAAAATCCAGGAAGCCAAAGATG
TCTACAAAGAACATTTCCAAGATGATGTCTTTAATGAAAAGGGATGGAACCTACATT
CTTGAGAAGTATGATGGGCATCTTCCAATANAAATAAAAGCTGTTCTTGAGGGCTT
30 TGTCATTCCCAGAGGAAATGTTCTCTTCACGGTGGAAAACACAGATCCAGAGTGTT
ACTGGCTTACAAATTGGATTGAGACTATTCTTGTTTCAGTCCTGGTATCCAATCACA
GTGGCCACAAATT

Sequence ID 883

35 TCATTTACATTAATACTCAAAACTGCTCGATTAAGCAGGTGCTGTTCTTATCGCCA
TTTTGCATATGATGAGAAAGGGTAAGGTCACCCAGCTAGTATTTGGCTCACAGCAG
GCCTTAAGACTTGGTTTGTGTGACTCATCAGTCCACGCTCCTAAAACCACTAAGTT

- 228 -

GTTCTACCCTTTAATGTTGAATTAACATTGGATAGTGTTCAAGTTTANATGGGTGG
GTGAGGGCCCAAGGACCTTTCAAACCTCAGATCTCTTATTTAATAACCTGGTCCCAG
ATCCATTCTCTGTCTGAAGAGGAAGTCATCCTTCAGTGGCTATTCATTGTGGGGTT
AAGAGCGCAGACTATGAATTCAGTCTTTTTTGGGTCCCAGTTTGCCAGACCTTGAGT
5 GAGTGCCCCGAGTTTACTTACTTGTAAGGTAGGTGGAGGTAATATAATTAAATAA
ACTTAAAAAACTAATTAAAAACAAAACAAATGAACTAAGGTCTTAGGATATCTGGC
GTCTATTTTGCGCCAAATCACATAATGTCTATTGTTGTGTGTTGGACTATAGGATT
GTCCTTTAACAGGGAAGGGTTTATTTCTGTAATCAAGTCTGTCAATATTATGACCA
TGTTGATAATAGCTACCTTTAATTGAGGGCTTCCATGTGCCAA

10

Sequence ID 885

TCAGTGGAAGGGCAGGTTGAATCAAGGTGAATCAATCTGAAATTGAGCACACCT
GCCTGCCATCGCTGTTCTTCAACTGAGTGCTGCACATCATGGGCTCTGTCTGTGA
GAGAAAAATCCCGGTGCTTGGTGTCTTGCATGACATGGAGTTTTGCATGTAGATC
15 AATTTAAAATGTACCTCTTGTTTACATAATTTGCATAATTTTAAAAGATAATGTTG
CCAAACTTTGGAAATGTTAATGTTCANACTGAAAATCTCCACTACATGTAACCTTC
TTCCTCTGGATCAGTGGCATGGCTTATAATCCCAGCCAGTGGTTTGAACGTTCCTCA
GTGTCAACTGCCATGTGCTCTGCTTCAAGGGGGAAGTGGCTTTTGTGAATTTTTTT
GTACATAAGTATTTGTTACAAATATTTTAGCAAATGCTTCTATTTCTCTTGCTTG
20 TGCATATCTTGGCTGGCGTTACAGAAAAATAGTGTAACATTATTTCTTACCGGG
GAATGAGGGTTTT

Sequence ID 887

AGCACCTGGCACAGAGTAGTAGCTAACACAGATGTTAATTTTGCTGCGTCAAATGT
25 TTTCACCTTTGAATCTCTCTTGAGTATTGTTCTCCTTATTGATTACATGATGACATC
CTGTTTTCTCTCCCTGACCTTTACTGTTTGTGTTAGAAAAAAAAAAAAAAAAAAAAA
AAAAAA

Sequence ID 889

30 CAGAGAGCTTGTTCCCTCCCTCCCTGTGCATGCAAACAAGAGGGCATGGGAGCACA
CAGAGAGATGGCAGCCACCTACAAGCCAAGAGGAGAAGCCTCACAATCAAACCTCTC
GCTGCTGGCGAGAGTCTTGGACTCTGTCTTGGACTTCCAGCCTCCAGACTGTGAGA
AACAAATTTCTGTTGTTTCAGCTTCTCAGTCTCTGGTGTGTTTGTGTTATTGCAGCCTG
AGAACACAGCTGTACNATTATNAGGGAAACAGAAAACACTGATACTTAACAATGCT
35 AATGCAATTATTTATTTGCTTTTTCAGTCTCTACAAAACGTTCTAAAACACTAATCT
AAATATTAACAGTAAAATATTTGCATAACTAATGGAACTAAGAAATCATATGACC
AATATTTCACTTATTGGTAATCTTACTCTACTGATTTCCCCCCAGACTGTGATTTT

- 229 -

TGAACTTCCTTGCCTTTCTCCTGTCTTTCTGNGTTTATTTCATGGAATTCCAGTTAT
CTGGGCTTGAAATTGCAGGCTCTCCTAACTTAAGCAAAATCTGACAGATCAGCAAA
ATGAGATAAATGTTTCTTTTTCTTTCTGACTGCATTAAATCAGATACAACCTCAGC
ATTAAAAAGCTATCTTTGNAAAATGNTGGTACTAATAAATTAGTCTTA

5

Sequence ID 890

CCAGTTCCACATTCAGTGAAGTCATGAACTTGAAATTGGCCATGATCAAAAAGTAT
TTAAATCACAGAAGTTGCAAATGCCACAAATCAAGGTCTTTTTCTCTTGGAGAACC
TGTTAAACATTTACCAACTCACGACCGCCATGCACCCAATACTGCAATAGGTCTAT
10 AGATGCAGATACTGTCTCCATGAATCTTATAGGCTAGAAAGGAAATAGATAAGTAG
TCCTACCAGAAGAACATGATGAAGGCATTTGTGGTAAACAGAATGATGGCCCCCA
AAGATGTCCACATCCTAATCCCTGAAGCCTATGAATATACTACTTTACTTGGCAAA
AGGGACTTTGCCACAGGTTTTTAATTAAGGACCTTGAAATAGAGAGATTATCCTGG
ATAATCCAGATGGCCCCAGTGTAATCCCAAGGGTCCTCACAAAGGGTAGGAAGGAG
15 AGCCAGAGTCAGAGAAGGAGACGTAGCAATGGAGGCAGAGGTCANAGAGAGATCTG
CAGATGCTGCTGTGTTGGCTTTGAAAATGAGGAATGCAGGTGACCTCAANGNGCTA
GATGATGCAAGGAAACAAATAATCTCCTATGAACCCTAGGATGGGCATTATTATGA
GTCCTATTTTATAAACAAGGAACTGACNTCCAGAAAGATAAATGC

20

Sequence ID - 891

nt: 626

GGCAGAGGTTGCAGTGAAGTGAAGATCATGCCATTGCAATCCAGCCTGGGCAACANG
AGTGAGACTCCATCTCAAAAAAAAAAAAAAAAAAAGACAAGAGTNTCCACTCTAAACA
CTTNTATTCAACATAGTCCTGAAAGTCGTAGCCACAGCAATTTAACAAGATAAAGC
AATAAAATGTATTCAAATAGAAAAAGAGGAAGTCAAATTATCTTCACTGGNGATAT
25 AATTCTCTACCTGGGAAACTTCACCGAAAAAGATTTACCAAAGATTTCTAAGCC
TAAATAATGACTTCAGCAAAGTCTCACCATACAAAATCAACATACACAAATGAGTA
GCATTTCTGTGCACCAATAATATTCAAGCTGAGAAAAAAGAACATGGTTCTATTT
ACAATAGCTACAAACAAAAAATATGTACCTAGTAATACATTAAATCAAGGNGGTA
AAATATCTNTACAACAAGAACTACAAAACCTGCTGAAAAAATAGAGACACGCAAA
30 TAAGTAAAAAGGCACTCCATGCTCATGAATTTAAAGAATCAATATAATTAAATGT
CCGNGCTGCCTAAAGCAACTTACAGATTAAAGGCTATTTCTCTCAAACCTATAAATG
CACCTTTTTTA

35 Sequence ID - 893

nt: 585

GTCATTGCTGGGTGGCGCCAGCCCTCAGACTTGCTCTTTGCAGTAGGAAGAAGGC
CTCCCCACATACCTTCCCACACTCATCACCTTAAGCCAGACTCGGTGTCCAGTGAA
TATGACCATCTCTTGCCCATTTTCTAATGAGTGTTTTTCATTAATGAGTTATAAGAA

- 230 -

TGTGGTGGGTAAATCTATGGGCTTTGAACTAGTGAATCAACTTGGTTTCAGAATCT
GGCACTGCTACTTACTAGTGAATTTAAGCAAGTTATTTACCTTTTCAGAGTGTCAG
TTCCCTCATGCATACAAGGAAGATAAAAAATAATGTNTACNAAAGTATTGGAGTAA
TTAATACATGGAGAACTACATGTAAAGCGTTTAGCATGATGTCTGACATATTAAGC
5 ATCCAATATTAGTNGCTTGCAGAATTATTAGTAAAAGAGATTGCTTCTGAAAGCCA
TTCCAATTCTTAAATTTTATAATGCCACATTTGAGGTCACCTGAAGTCGTGTATAA
CATGTGTACATTTTTTGCGATTTATTTTTTCAATTCCCANATTAAAGGCATAGAGAT
ATCCTAGCNANGGACTCCAAGTGTG

10 Sequence ID - 895 nt: 560
GTAATTGCAGCCTGGGCAACGGAGTGAGAGACTGTCTCAGGAAAAAAAAAAGAAAA
AAAACACTAGGAGGTAGTTGAATATATCCTCCATTCCCCATTTGTGGATTAGTTAGT
AAATGGGGCATCTTAGGGTTTAAATATGTCCAGGGTCACTGAGGATCAGATCCTAG
GGTTCCTTTGACTCAAGGCTTTTGTCTCAGCAAACGTCACCTTCCAGCAGGAAGG
15 CTTTCTCAGGCAAGTAGCAGGGTGGCTACTATGTATCGCTTCTTTATTTTTTCTTT
TTTAAATAATGCAGGCACCGTGCGCATAATTTAAAAAATCAGTGCTAAAACCCTT
AAAAAAAAAAAGCTGTTCTCATCTCCTGTCTTTCTTTTTTTTTCTTTTTATTTTTT
TTCTTTTTATTATTATTATACTTTAAGTTTTAGGGTACATGTGCACAACGTGCAGGT
TTGTTACATATGTATACATGTGCCATGTNGGTGAGCTGCACCCATTAACTCGTCAT
20 TTAGCATTAGGTATATCTCCTAATGCTATCCCTCCCCCTCCCCCTTTTTTTTTTT

Sequence ID 896
GGGAATGTCTTAGGCACTGGGACTGTAAGTGCAAAGACCCTGTGGCACAAGGGAAT
GTTAATTATCTACCTTTTCANAACTGGAANAAGGCCTAGCCTAGAGCATTGAAAAC
25 AATAAGGGAAAGGAGGAGTAAGGCTGGANAGATAGGAATGGTTTAAAGTCTTTGTT
AAAAATTTTTTTTAAAAAAATCTTTATCACAAGAAGAGGATTGGCNTGATCAAATTT
GACTTTTAAAAANATTACTTGGGTGGGTCATGATCAAATACTACTTAGGGAGATTA
GTTTANATGATAATGGCATTCTGGACCANAGTGGAGTCAGAGGTGAAAAGAGGTAG
ATATTCCANAATTGAGGGATTTGTGAGGTGAAATCATTTGTTACAGATATTAAAGG
30 ATAAGGAGCTTTGTCAAAGGGGATCTTAAGTTTCTGGTATGGTAACTGGGTTAGAG
AGCCCTGGAACATGACCAGCTTTAAGGGAAGAGAGCTTGAGCTCTGTTCTTGTTAA
GCTCAGTTTGAGATCTTTGTGGAATCAAGTGGAGAGGTCTAAGCAGGGAACTGGCT
TGGCTAGGCTGTAAAGATGAATCTGAGAGTCCCAAGAATATGGTAATTATTAATAA
AAGCCTTAGGTANATGAAATTGTTTTTGGG

35
Sequence ID - 897 nt: 509
GCAAATCTACACATTTGATTAAATGATAGGGAACCTATGCACACACATAATACATAT

- 231 -

AATGCTAGTTTCTTGGTTTTTGATATTGTACCATAGTTATGTAAGATGTAACCATTG
GGGGAAACTGGGTGAAGGCTACATGAGACCTCTCTGTACTTAATCTTTGCAACTTA
TGTGAATCTATAATTATTCCAAAATAAAAAGTTTTTAAAGAACCTAAGTATCCTTAT
TACTGAGGGTCATCGTGCTAGACAGCAAGGTTGGGCCAGAGCTTCTAGTTATTTAA
5 AATACTAAATACCAGCCTGGGCAACATAGCAAGACCCTGCCTCTACAAAAGCAAA
AAAATTAGCTGGGCATGGTGGTACATGCCTGTGGTCCTAGTTACTCTTGGAGGAGT
CTGAGGTGGGGAGCTTGAGCCTAGGAGTTTGAGGCCGCAGTGAGCCTTGATTGTGT
CTCTGTACTCCAGTCTGGGCCACAGAGCAAGACCCGGTCTCTAAAAATAAATAAAT
AAATA

10

Sequence ID 898

ANTGCACTCCAGCTTGGTGACAGAGGGAGACTCCATNTTAAAAAAAAAAAAAAAAAA
AAAAAAAGGGAGTAGCTTGAAGCCACATAGTAGTTAGTGGTAAAGGCCACCCCTTT
TCCCACAACCTCACACCAGCACCACAAGCTAGCCTTTNTAATTTCCAAGCCAGTGCC
15 CTTTCAACGCACACACCCCTGTGTCAGTTCCCTTTCTGCTGCAAGCTCTCTGGAGG
CAGATACTGTTGAGTCCCTGGCCTGCCTATGAGAACGGCTCATGATCTCTATTTCT
TCTGCTTAATGACCATCTCGAAGTAACAAGTTTAGCCTAAAATAAACTTGCTAAGT
TAGCAAAGGAAGTCCTTAGCAGCCACCATTTCTCGATTCCCTCCATCACCTCCCCTG
CCCCTCAACTCCCTCATTTCTCCCAAGATATGGGCTCCAGGCTGGGCGCGGTGGCT
20 CACGCCTATAATCCTAGCACTTTGGGAGGCTGAGGTGAGCAGATCACTGAGGTCAG
GAGTTCG

Sequence ID 899

TCNTTCGGAACGCGCC

25

Sequence ID 900

CTGGAGGGATGGGTAGGATTTTGACAAGAGTGGTTGAAGGTATTCTAATTCACCTTA
GTACCTACATGTGCGAGGCAGCATGAAGGCAAAAAGCCTGGGGCATGTTTCAGAGA
ATAGCAAGTATTCTAGTTTGAGTGGCACCTGGTACGTATATAAGGGAATAGTAAAA
30 GATCTGGCTGGAAAGGAAAAGTAGGGGCAGGTTACGAAGGACCTCTGAAAGTCAGA
CTGTGGAACCTGGAACTTTTATCAGGAAGCAGTAGTTAGTTTTTTCAAGCAAAGCT
AATTAGAGTTGATATTTAGGAGGATGAATCTAACAGTTGTGTGCAAGGATGCCTTC
AAACTGAGTGAGACTAGTACTGGAGACTGGTTAAGAGACTACAACAATAACCTGAG
TAAGAATTAATACAGGCCTGACCTAGTTTTGAGTGAGTAGGATTGGAAACAAGAGT
TTTAGGTATTATAGGATTTATGCATATAAAATGGACTTGACAGAACTTGAAGAAAG
35 AGAAAGTGTCAAAGGACACAGAAAGTGAGGCAGGATATCTTACAATGTTAAAGGA
AAGGAATAATAGAAGTTAC

- 232 -

Sequence ID 903

GGAAACATAAGCTTGTTTCAGTACACTCACGCTGTAGATTAATTCTGATATTACAT
ATCTCCATCAGACTTTGTACCCTCTCTCTTCCATCCCTTACCCTTACCGATTAGGT
TGGTATTACCTAAAAATCCATAGAAAATGTCCAGGTGAATTGCCTTATGCTTTCTA
5 CCCCATAAGGTATAATT

Sequence ID 904

CTCTGTGGTGTGAGAACACAGTGGGTGACCAAGGCTTTCCAGATGAACCCAAGGAA
AGTGAAAAAGCTGATGCTAATAACCAGACAACAGAACCTCAGCTTAAGAAAGGCAG
10 CCAAGTGGAGGCACTCTTCAGTTATGAGGCTACCCAACCAGAGGACCTGGAGTTTC
AGGAAGGGGATATAATCCTGGTGTATCAAAGGTGAATGAAGAATGGCTGGAAGGG
GAGTGCAAAGGGAAGGTGGGCATTTTCCCCAAAGTTTTTGTGAAGACTGCGCAAC
TACAGATTTGGAAAGCACTCGGAGAGAAGTCTAGGATGTTTCACAACTACAAAGC
TGAAGAAAATGAAGCCCTATTACTTGTTTGTAAGATTTAGCACCTTCTGCTGTAT
15 ACTGTACTGAGACATTACAGTTTGGAAGTGTTAACTATTTATTCCTGTAAATTT
TAACCTACTAGACAATGATGTGAGTACCCAGGATGATTTCTGGGGCACAGTGGGT
GAGGAGATGGGGACAGGTGAATGGAGGAGTTAGGGGAGAGGAAAAGTGGATGGAAG
TGTCTGGAAAGGGCACCAAAAAGTCTTCCAGGTCTGATCCTGTTTCTTGCTCTGA
GTGCTAGCTACCACTGTGTCACACTGTAACATN

20

Sequence ID - 905

nt: 655

CTCAGCTCTTGCCCTGGTCACCTTGTGGCTTTTACCATCCTCATCCCCTGTGCCACC
CACATCCTGCCACTTCTGCATGGAGTTGGGGTGGGGCCATTGGAGAAAAGAGGTTA
AACAAGCAGTAATTTACTTGAGTACAGTCTTTGAGCCAATGAAATGCCAGTCATCA
25 TTTCCCAGGGGTACTTGTCATCTTGTCAACAACCCGCTGATAATGCTCCTTCAATG
TGAATAGCAAAAGTAGGGAGAGACGCTGAATGAAGAAGATGCCTACCCCTCAGGAA
GACTGCTGTCCGCCTCCAGGCCTGCATGCACACACCCATGCCCACCTGCACCCCCA
GCACCACGCCCACACTCACTCGCACACACCCACATGCCAGTGTTTTGGGGTTGGCA
GCCTGGACACTGCTGAGGCAAACACAAGTCATCAAGCATAATTCTCATTCTCTCCT
30 TCTGTCTCTGTTTTAGTTACAGGAATTTGGTCAGTTTAGAGGATTTAATAAGTCCG
TGGAATAATTTGTTTTCTGTCTCTTGCTACCCACGTGAAAAGTAAGTGCATGCTTCAT
GATGTGTTTTCCCACTACCTTCCAGGCCAGCCGAGCCCACTGGCCANGGCCTGGCC
CGGTGACCTCGGTTGACACTGTCCTCANGCCACTCACTT

35

Sequence ID 906

CAGAATTTTCATGTTTATGCTGCACAAGGCCTGTATTTTATAATGGTGGCTCTTTTG
GACGATGACTTCCTCGATGGTGAAACTTCCAGTAATCTCCCTCATCATACTGAAAT

- 233 -

GATATCAGTATATCATCAGAACACCATGGAGCTTGTCATTTGAGGGACACAGCTTG
CTTGTGTGCTTGGGAAAGAAGAGGTTTAGCATGGTTTCAGGTCAGTGATGAGTCCA
ATGATCTCTGCAAGTTCCCTTAGCTCTGANAATTCTGATGTCATATGCACTTCTGC
CGCCAGAGTTGCTGCTTACTGGATGCGTAAGAAGAAAAGAAAAAAAAAAAAAAAAA

5

Sequence ID - 907

nt: 582

CTTCCATTGGGGGTAAAGATCAAACTTTAGGCGAGCCAGGTCTGTATCTCCATTCC
TGTCTCTGACTGCTTCCCTGTAGGGATTGTCTGCAAGCGCACACCTGCATTTTCTT
GTCCACAAGTCTATGCTCTAACTCTGTCACCTGCATGGCTGCAAATTAGCTTCCTT
10 CTTCCCTGCCCTCTTCTCTCTAGCTTGGATTTTGAATTTGAATGGCAGGCATGGGAT
GTCCGTGTGTGTGTACTGCTGATGTGTACAGCCGCTTGTTAGCGCTCTCATTGTCT
TCAAATGTAAGTCATTTTGGCTGGGTGCGGTGGCTCATGCGTATAATCCCACGCTT
TGGGAGGCTGAGGTGAGCTGATCATTTGAGGTTAGGAGTTCGAGACCAGCCTGGCC
AACATGGCAAACTCCATCTCTACCAAAAATACAAAATTAGCTGGGTATGGTAGT
15 GCACGCCTGTAATCCCAGCTACTTGGAATGCTGAAGCAGGAGAATTGCCTGAACCC
ANGAGGCGGAGGTTGCGGTGAGCCAAGATCACGCCACTGCACTCCAACCTGGGTGA
CAGAGCAAGGCTGTGTCTCAA

Sequence ID 908

20 ACCTGACTTCAAAC TATACTACGAGGCTACAGTAATCAAACAGCATGGTACTAGT
ACAAAACAGACCAATGGAACAGAATAGAGATCTCAGAAATAAACTGCACATCTA
CAACCATCTGATCTTCAACAAACCTGACAAAACGAGCAATGGGGAAAGGATTCCCT
ATTTAATAAATGGTGCTGGGAGAACTGGCT-AGCCATGTGCAGAAAATTGAACTG
GACCCCTTCCTTACACCTTATACAAAATTAACCTCAAGATGGATTAAAGACTTAAA
25 TGTAGAACCCAAAACGATAAAAACCCTAGAAGAAAATCTAGGCAATATCATTAAGG
ACATAGACATGGGCAAAAATTTTCATGATGAAAACATCAAAGCAATGGCAACAAA
GCAGAACTGACAAATGGGCTTCTGCACAGCAAAGAACTATCGTCAGAGTGAAC
AGACAACCTACAGAATGGGAGACAGTTTTTTGCAATCTATCCATCTGACAAAAGTCT
AATATCCAGAATCTACAAGGAATTTAA

30

Sequence ID 910

CAAAAACAAGAATTACCCGGGCTTGGTGGTGCATGTCTGTAGTCCTATCTACTCA
GGAGGCTGAGGCTGAAGGATCACTTGAGCCCAGGAGTTTGAGGCTGCAGTGAGTGA
GCCATGATCATGCCAGTGTA CTCCAGCCTTGGCAGACTGAGCAAACTTGGTCCCT
35 CGCAAAATGTTGAAGCCCAGTTTTTCACTATTAACCTGTATTTTCAGTTTCCCCATGC
TAACTTTGAAACACTGGGGCTGGCCTGAGGGTATAAAGGCTTATTCAAACCTCAGTA
ATTTAACTTAAAATCCTAAGGAACTTCAAAAAGTGTAATCTAGTCCAAATGGGGC

- 234 -

ATCAATTCTAAAGCATTTGCTTGTTTGAGCAGATTTTCTGTGTCTGAGGTATATAG
ATAACTTATCTTTTTATGACTAAATCCAAGTCCTTAGTTCCTGTTGGAATTCAAAA
TCATATTTAAAAATTGATGCTTTGTTCTATAATTAATGCTTTGATTGTATAAATAA
TAAGTATTCTTCCAAATCCCTTTTTACAGATGATGATTCTGATACCGAGACGTCAA
5 ATGACTTGCCAAAATTTGCAGATGGAATCAAGGCCNGAAACAGAAATCAGAACTAC
CTGGNTCCCAGTCCTGTNCTTAAAATTCTAACTCGAC

Sequence ID - 911

nt: 595

GAGGGTGTAGAAGAGAAGAAGAAGGAGGTTCTGCTGTGCCANAAACCCTTAAGAA
10 AAAGCGAAGGAATTTTCGCAGAGCTGAAGATCAAGCGCCTGAGAAAGAAGTTTGCCC
AAAAGATGCTTCGAAAGGCAAGGAGGAAGCTTATCTATGAAAAANCAAAGCACTAT
CACAAGGAATATAGGCAGATGTACAAANCTGAAATTCTGAATGGCGAGGATGGCAAG
AAAAGCTGGCAACTTCTATGTACCTGCAGAACCCAAATTGGCGTTTGTTCATCAGAA
TCAGAGGTATCAATGGAGTGAGCCCAAAGGTTTCGAAAGGTGTTGCAGCTTCTTCGC
15 CTTTCGTCAAATCTTCAATGGAACCTTTGTGAAGCTCAACAAGGCTTCGATTAACAT
GCTGAGGATTGTAGAGCCATATATTGCATGGGGGTACCCCAATCTGAAGTCAGTAA
ATGAACTAATCTACAAGCGTGGTTATGGCAAATCAATAAGAAGCGAATTGCTTTG
ACAGATAACGCTTTGATTGCTCGATCTCTTGGTAAATACNGCATCATCTGCATGGA
GGATTTGATTCATGAGATCTATACTGTTGGAAAAC

20

Sequence ID - 912

nt: 651

CATTTCCAGAGTTTATGTGAATTGAATTGAACTATGGTTTTATGTTACTGTCAGTA
GAATGAAGTACGAATATTTGAAAAATACACCTTCAACTTCAAAGTGATTCTTGACA
AAAATTATAAGGAATCATTTTGGACACATTTTCTGGTAGAGCCTTGTA AAAATTAA
25 AACCAAGTGTTGTTTTCAAGAAGAACTGTAATACATAATCAGGAATTTGAGTAGGG
AGATTATTTTGTATTTTAAAATTAAAGTGGCTGTGTAGTTTTAACTTTAGTATTGC
AGGTAGAGTAAGCTTACATGATAACAAAAATCTTGGTCTTAGTGACTTAATGATTCT
TGATATTTATTGATTGATTGGTTATCATTCCAAATATTTTAAAAGATAATAGCTGG
CTGGGTGCGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAGGACGGGCG
30 GATCACGAGGTCAGGAGATCAAGACCATCCTGGCTAACACGGTGAAACCCCGTCTC
TACTAAAAATCAAAAAATTAGCCGGGTGTAGTGGCGGGCACCTGTAGTCCCAGCTA
CTCAGGAGGCTGAGGCAGGAGAATGGCATGAACCTGGGAGGCGGAGCTTGCAGTGA
GCTGAAATCGTGCCACTGCCTCCACCTGGCGACAA

35

Sequence ID 913

GTGAGGTGGGGACTTCATTCATTGTCCTATTTCTATCTCCACTTTGTGCCTGGAGA
GCTTTCAGGGGAGGTGGAGGAGGAGGGTCTGCCAAGCTACTGCAACATCTGTCACC

- 235 -

CACTATACCCAGTTACTTGGGGGAGGACAGACACTGTGGTGTTCATTAAAGTTGTTT
GAACCAAAGTGGCGGCTGCATCTTTGTCCCGATGCTAGCCGTGCCGGTCTCCCATC
ATCCGCTCGCCCTCCTTTNCCCTGGGCTGCGCCCACTTGTCTTCCTGGATATTTGG
GGGTGACTCGCCATGCTTGGCACCCCTCTGCTTCCTGGTGCTGCTCTGACTCGAAGA
5 CGGGACAGTCCCTGGTGCACATCCAGGGAAGAGGAGTGTGCGGTAGTTCTTGCAGTA
GGCACTTTATCAGGACCTGACCTGTTGCTGGGTGATTTTAGTCTCTACAAACAGAA
AGCGTTTCAAAGCGTCAGCTGTGGGAGCAGAGTGACCCTTTGCTGATGCTGGGGGG
AGGGGATCTAAATCCTCATTTATCTCT

10 Sequence ID 914

GGCGCCTGCTGGAGGAGGAGAGAGCTCTGCTGGCATGAGCCACAGTTTCTTGACTG
GAGGCCATCAACCCTCTTGGTTGAGGCCTTGTTCTGAGCCCTGACATGTGCTTGGG
CACTGGTGGGCCTGGGCTTCTGAGGTGGCCTCCTGCCCTGATCAGGGACCCTCCCC
GCTTTCCTGGGCCTCTCAGTTGAACAAAGCAGCAAAACAAAGGCAGTTTTATATGA
15 AAGATTANAAGCCTGGAATAATCAGGCTTTTTTAAATGATGTAATTCCCACTGTAAT
AGCATAGGGATTTTGAAGCAGCTGCTGGTGGCTTGGGACATCAGTGGGGCCAAGG
GTTCTCTGTCCCTGGTTCAACTGTGATTTGGCTTTCCCGTGTCTTTCCTGGTGATG
CCTTGTTTGGGGTTCTGTGGGTTTGGGTGGGAAGAGGGCCATCTGCCTGAATGTAA
CCTGCTAGCTCTCCGAAGCCCTGCGGGCCTGCTTGTGTGAACCGTGTGGACAGTGG
20 TGGCCGCGCTGTGCCTGCTCGTGTGCTTACATGTCCCTGGCTGTTGAGGCGCTGC
TTTAACCTGCACCCCTNCCTTG-CTCATANATGCTCCTTTTGA

Sequence ID - 915

nt: 230

TTTGAGACCAGCCTAGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAAA
25 TTAGCCGGGCGTGGCGGCACATGCCTATAATCCCACTTACTTGGGAGGCTGANGTA
GGAGAATCGCTTGAACCCANANAGGCAGAGTTTGCAGTGAGCCGAGATTGTGCCAT
TGCACTCCAGCCTGGGCGACAGAGCGAGACTCCATCTAAAANAAAATAAATGAATA
AAATAA

30 Sequence ID 917

NNCAGATTTTTTTTTTTTTTTTCAGNGTTAGACCATCTTTCATTCCTGGAACAAAC
TTAACTTTCCATGATATGTATTTTTTATACATTGCTGGATTTTATTTGCTAATATT
TACTTAGGATTTAATTTTCTAAGTNGACCTATAATTNTCCTGTATAAAATTGCAT
TTGTCACATTTTAGTATCAAGGTGTCTTANCNCCATGAAATGGATTTANAATGGT
35 TTATGTAANATAAAGTACATTTCTTCTAAAGGTTTGNGTGGATTAACTTTCAAATC
TGCCANAGNGNGTTTTTTTTTCTTTTTTTTTTTTTTTTTCATTTNAAGGGAGNGCAAGT
ANCTTTTCAAATNCTGATTTAATTTTAAAATATTTNCAAGTNTNTTTANAGTTTT

- 236 -

TATTTNTTNTNGAANGTTAACATTTTTATANAAAANGGTNTTATCTTTTTTAAATTC
TTTGACATCAGTTTCTTCANAATTCCTTCTTTTAA

Sequence ID 926

5 GTCATATCTCTTCCCAGGGAAAGCAGGAGCCCTTCTGGAGCCCTTCAGCAGGGTCA
GGGCCCCCTCGTCTTCCCCCTCCTTTCCCAGAGCCATCTTCCCAGTCCACCATCCCCA
TCGTGGGCATTGTTGCTGGCCTGGCTGTCCTAGCAGTTGTGGTCATCGGAGCTGTG
GTCGCTACTGTGATGTGTAGGAGGAAGAGCTCAGGTAGGGAAGGGGTGAGGGGTGG
GGTCTGGGTTTTCTTGTCCCACTGGGGGTTTCAAGCCCCAGGTAGAAGTGTTCCTT
10 GCCTCATTACTGGGAAGCAGCATCCACACAGGGGCTAACGCAGCCTGGGACCCTGT
GTGCCAGCACTTACTCTTTTGTGCAGCACATGTGACAATGAAGGACGGATGTATCA
CCTTGATGGTTGTGGTGTGGGGTCTGATTTTCAGCATTCATGAGTCAGGGGAAGG
TCCCTGCTAAGGACAGACCTTAGGAGGGCAGTTGGTCCAGGACCCACACTTGCTTT
CCTCGTGTTTCCTGATCCTGCCTTGGGTCTGTAG

(15

Sequence ID 938

TGGCCATCCTTTTCCCCCACAACACACCCCCCTTAACCTATCTCTTGGGACTTAGCC
CGACCCTCCCTCTCATTTCCCATTAAGTCTGAGAGGCAAGAGCTAGGTTAGGCAAG
GAGGTGGTTGGCCAGAGATGGGGAACAGCCAGGTGCCCCAGTCCTCTGATTTTTTCC
20 TCCATCCTGCTTACCACCTCCCTGGGTACTTACAGCCTTCTCTTGGGAACAGCCGG
GGCCAGGACTGGGTACCTATGAGCTGAATCAGCATCTCCTCCTGAGTCCCAGGGC
CCCTGCAGTTCCCAGTCTCTTCTGTCCTGCAGCCCTTGCCTCTTTCCCACAGGTTC
CACTTTATATCCACCTTTTTCCTTTTGTTCATTTTTTATTTTTTATTTTTTTATTAT
TAAATGATGTGGTCTATGGAAAAAAAAAATAAAAATCTGACTTAGTTTT

25

Sequence ID - 939

nt: 513

GGAACCCAGTGTATTACCTGCTGGAACCAAGGAACTAACAATGTAGGTACTAGT
GAATACCCCAATGGTTTTCTCCAATTATGCCCATGCCACCAAAACAATAAAACAAAA
TTCTCTAACACTGCAAAGAGTGAGCCATGCCTGTAAACACTGTAAAGAATGTAACA
30 TGTGGGGGACACACAGGGGCAGATGGGATGGTTTAGTTTAGGATTTTATTAGTGCA
TGCCCTACCCTCTGGGGGAACGTCCCATCTGAGGTTTTCTTCTCGGTGGGGGGATT
TAACTTCTGTCCTAGGGAAAACAGTGTCTGATGAGGAGTGTTTCCAACACAGGCTA
CATGAATTCCCCTATACCAGTGCGAAAGCAGCCAGGAGTCCCCGTTGGAAAAGAAC
AATGCCACTCTCTTTTATGTATCTTGGTTCTGCAACTCATTTGTTGTAAGTAGGGT
35 TAATCGAGTATCAGGTTTACAGTATCCTGCCCTTATTATTTTTATGATTCAGTACT
CAAGTTCCA

- 237 -

Sequence ID 947

GAGAGTGAAAAAATTCTGGTACAAATTGGGAAATTAGTATATAACAACATAGTGTT
AAATTCAATGGGAAAAGTTTAATAAGAGGATTTGGTATCAACTGGCTGTCCAAAGA
TAAAAATGGACCGTCCTATCACATACAAAATTGTTTTTTAGATAAAGATTTAAATA
5 CAGGCACTCCTTCATTTGCGTGGTGCACCTTGAGGTGTTGCAGAAATGATGAGAGC
TGAAACTGCAAAGCAATTTTAATACTTTATCTGTTGGAAATCTTATAGTTTTCTG
TGACCGTTAAAATTTTCATTAACTATTAAAAACACCCATGACTGGTCACAAATGT
ATTGGGAAATGGAAAAGAATTAATACACTAAAAATACAAAAAATAGAAAATATTTA
AAATTATCTAAAAATTTGAAACATTAGAAAAATTGAGAACTAGGCAGGGCGTGGTG
10 GCTCACATCTGTAATTTTAGCCCTTTGGGAGGCTGANGCAGGTGGATCACCTGANG
TCAGGAGTTCGAGACCAGCCTGCCAACGTGGGGAAACCCCGTCTCTACTGAAAATA
CAAAAATTANCCGGGCATGGTGGCACAAGCCTGTAATNCTTGCTNACCAGGANGCT
GAGGCAGGAGAATCACTTGAACCCANGANG

15 Sequence ID 949

GTTTCACATGAGAAGGTAGTATTATGTACAGTGACCTTGTTTAAAGTGTCNGTTTA
ATGTTACCACTAAGGCCCTGCCCCAGCTTTATCACCTGAGCACTAACAAGTGCTGT
GTGGAGTTCAGTCCATGCTGGTAACTNTTGAGTATTCAGTGGGTCTTTTAAACAATT
ACCACCGTGGAGGANANAGCAAGGAAGAGAAATGCTGTGATCTTTTNCTGTTTTTA
20 ATTAGNGAAAGAGGGATTANATTAAACAAATGTTACAGAGNTGTGACTNTGATCCC
CCAGNGGTAAGCAATAATTGTANAGACTGGATTNANAAGCCCTGAGAGTTTATTT
TCAACCTATNTATTATAGNNCAATCC

Sequence ID 1028

25 ACAAGGCTTGGGGGCTGGACTCCCTCTACTGCCTCTGGCCATACCCCTCCTGGAG
ATGGGGTCAAGGCACCAGGACTGA

Sequence ID - 1056

nt: 435

TCGCTTGTAAGCCTGAGACAGCTGCCTGTGTGGGACTGAGATGCAGGATTTCTTC
30 ACACCTCTCCTTTGTGACTTCAAGAGCCTCTGGCATCTCTTTCTGCAAAGGCATCT
GAATGTGTCTGCGTTCCTGTTAGCATAATGTGAGGAGGTGGAGAGACAGCCCACCC
CCGTGTCCACCGTGACCCCTGTCCCCACACTGACCTGTGTTCCTCCCCGATCATC
TTTCCTGTTCCAGAGAAGTGGGCTGGATGTCTCCATCTCTGTCTCAACTTCATGGT
GCGCTGAGCTGCAACTTCTTACTTCCCTAATGAAGTTAAGAACCTGAATATAAATT
35 TGTTTTCTCAAATATTTGCTATGAAGGGTTGATGGATTAATTAAATAAGTCAATTC
CTGGAAGTTGAGAGAGCAAATAAAGACCTGAGAACCTTCCAGA

- 238 -

Sequence ID 1071

NGATATAGTNCCGCATGGGAAAGATGANCAAGGTATAACCNAGCNTNATATAGCAAG
GACTAACCCCCCTGCCTTCTGCATAATGAATTAAGTAGAAATAACTTNGCAAGGAG
AGCCAAAGCTAAGACCCCNAAACCAGACGAGCTACCTAAGAACAGNTAAAAGAGC
5 ACACCCGTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGGCGACAAACCT
ACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAACTTTAAATT
NGCCACAGAACCTCTAAATCCCCTTGTAATTTAACTGTTAGTCCAAAGAGGAA
CAGCTCTTTGGACACTAGGAAAAAACCTTGTAAGAGAGAGTAAAAAATTTAACACCC
ATAGTAGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACACCCACT
10 ACCTAAAAAATCCCAAACATATAACTGAACTCCTNACACCCAATTGGACCAATCTA
TCACCCCTATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCCTCCGC
ATAAGCCTGCN

Sequence ID - 1074

nt: 689

15 GGGAGGCGGAGGCTGCAGTGAGCTGAGATCGTGCCACTTCATTCAGCCTGGGCAA
CAAAGCGAACTCTGTCTCAAAAAAAAAAAAAAAAAAATTTGTTGACTGTTGTAA
TTTAAAGCTTGTCATTTTTTTATTTAGTAATAACACTCATTAGTGTAATCTATGA
TGAACCAGGTTCTGCACAAAGTACCTTATGTTTCATGGCCTCATATCGTCTTCTCCA
AACTCTGCAAGATAGGATTCATCACCACCTTATAGGGAGAGATCTGAAAGTTTAAA
20 ATTGTACCCAAGGTCACACAGCTGGTAAGTGCCAGAGCTGGGATTCGGTAGGGTGT
TCANAGTGCCTCTCCTGCCGTAGGCTTATCACAAAAGTCAAAGTTTGGTCATAAT
AAAGCCTGAAGTTTGGCAGGATTTAAAAATAGTCACCANACTTTTGAGTTGGAGCA
TCCCACCTCACTGCTGTTACCTTCTGTGGCAGGGAGAGTCATCATTTCCATTTCA
GCTTGTGGAATATCTTGTCAATTAACATTCTCATGCAAAAGCCATTTTATGGTGCCC
25 AATGAANATGGTTAAGCTACTGCCCCAAGCCTNTGGAAGCCTTCCTAATTTTGGAC
TTGCACTATGCAAATTGNATAATATTTTCTCTACCCTAAGCCAAATATTTTCTTCA
CTTTTCATTCATTCTAC

Sequence ID 1081

30 CGCCGCCGCGCCGCGCTCGCTCTCCAACGCCAGCGCCGCCTCTCGCTCGCCGAGCT
CCAGCCGAAGGAGAAGGGGGGTAAGTAAGGAGGTCTCTGTACCATGGCTCGTACAA
AGCAGACTGCCCCGAAATCGACCGGTGGTAAAGCACCCAGGAAGCAACTGGCTACA
AAAGCCGCTCGCAAGAGTGCGCCCTCTACTGGAGGGGTGAAGAAACCTCATCGTTA
CAGGCCTGGTACTGTGGCGCTCCGTGAAATTAGACGTTATCAGAAGTCCACTGAAC
35 TTCTGATTCGCAAACCTTCCCTTCCAGCGTCTGGTGCGAGAAATTGCTCAGGACTTT
AAAACAGATCTGCGCTTCCAGAGCGCANCTATCGGTGCTTTGCAGGAGGCAAGTGA
GGCCTATCTGGTTGGCCTTTTTGAAGACACCAACCTGTGTGCTATCCATGCCAAAC

- 239 -

GTGTAACAATTATGCCAAAAGACATCCAGCTAGCACGCCGCATACGTGGAGAACGT
GCTTAAGAATCCACTATGATGGGAAACATTTTCATTCTC

Sequence ID - 1083 nt: 198

5 GCGCGTCGACTTTGTTTAGACATTGAATGACTTTGTTAAAGGCACAATTAATCACA
TTGGTTGTACTCTGNNGACAGCCTTCTTTAAAAAAAAAAAAATAACAATTTAAAACAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAANTTTTAACC

10 Sequence ID - 1084 nt: 198

GCGCGTCGACTTTGTTTAGACATTGAATGACTTTGTTAAAGGCACAATTAATCACA
TTGGTTGTACTCTGNNGACAGCCTTCTTTAAAAAAAAAAAAATAACAATTTAAAACAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAANTTTTAACC

15

Sequence ID - 1099 nt: 561

TGCATGCTTGTGGATTGGAAAACTTTGGGAGACTGATTACTTTTCATTATATATGT
GTCACAGTGAAACAGCTTTTATGTGTCATGTAAGATTACTGCTTGCCTCTCTAAGG
AAGGTCGTGACTGTTTAAATAGACGGGCAAGGTGGAACCTTTTGAAAGATGAGCTT
20 TTGAATATAAGTTGTCTGCTAGATCATGGTTTGTATTGAACTAACAAGGTTTGCAG
ATCTGCTGACTTATATAAAGCTTTTGTATTCTTACTAAGCTTTAAGATTTAAAAAA
TGTTCAATGTTGAAATTTCTGTGGGGCTCTATTTTGTCTTTGGCTTTCTGGTGAGA
GAGTGAGGAAGCATTCCTTCTTCACTAAGTTTGTCTTTCTTGTCTTCTGGATAGA
TTGATTTTAAAGAGACTAAGGGAATTTACAACTAAAGATTTTAGTCATCTGGTGGA
25 AAAGGAGACTTTAAGATTGTTTAGGGCTGGGCGGGGTGACTCACATCTGTAATCCC
AGCACTTTGGGAGGCCAAGGCAGGCAGAACACTTGAAGGAGTTCAAGACCAGCGTG
G

Sequence ID 1109

30 TTTGNCGGTNTTGGANNNNNANAANTTTCTTCCANNCNTNACNTNTTGGTGGNCTA
AATTAANATGGNTTTNGNGGGTTCNTTNCTNNTNNTNNNCATGGGANANAATTNATT
NTCNTNCNNNTTCCTTNCCCTNAANCTACCTTCCCCCNATTTTCTCCCCNTTTCN
TNAATTANCATCCTCTCCNCNTANNTCNANACNTTAATGGCAANACTATCTAATAN
CNANNATAANANCTCCTGTNNNCCACATNTCTTATTNNNCGCNNCANGTTNCANNC
35 CCNCAGAGTNAACTCATCCTCNNCNNAANTTCATATCGTGNNCTNTNNNCNNTNGC
GCGANATATTAANNANACCNGTANNNTNNNANACANNANNTNNGNAANAANCCTTCT
NANNTTTTAGCNTCNNGCNNTAACNNNNNTCTTNGTGNNNNCNCAGCTTTCNCNNC

- 240 -

ATNATNCTNCNNCGAANTNTCANNCNTCTCCNCTTNAATGNNTTCCCATGNATTAA
NTNCCTCGNNNANAGCACTATCGTNNNNGAGNNNATTATNGNCNNTTTACNTCATG
TGGTCCANTNNCGTTNGNCGCNNNNNAATNTTCGTNNNNCNN

5 Sequence ID 1118

GGATTTTAGAGGAAGGCGCTNGGTACATTGGAGAACTGGAGTGGTCTGGAGTTCC
ACGGTGTAGTGGACCAGAGGCCACCTCTCCTGGGCTTCTCAGTGTCTCGCCGGCGG
GGTTCGGCCTGAGCTGGATTGACATAGCCCTTGGCGGATTTAAACAACCTAAACAT
TAAGCAGTACAGCTGCCTCAAACCTTTGGGATTTTCAGAATGACTGACACTGCCGA
10 AGCTGTTCCAAAGTTTGAAGAGATGTTTGCTAGTAGATTACAGAAAATGACAAGG
AGTATCAGGAATACCTGAAACGCCCTCCTGAGTCTCCTCCAATTGTTGAGGAATGG
AATAGCANAGCTGGTGGGAACCAAAGAAACAGAGGCAATCGGTTGCAAGACAACAG
ACAGTTCAGAGGCAGGGACAACAGATGGGGGTGGCCAAGTGACAATCGATCCAATC
AGTGGCATGGACGATCCTGGGGTAACAACCTACCCGCAACACAGACAAGAACCTTAC
15 TATCCCCAGCAATATGGACATTATGGTTACAACCAGCGGCCTCCTTACGGTTACTA
CTGATAGAAATGTTGGCAGCTTTTAGTAAAAGCATTTACTCTGTTACCATGAGAAA

Sequence ID 1125

NGACTGGCTCCCGAAAAGAAGGGTGGCGAGAANAAAAAGGGCCGTCTGCCATGGA
20 CGAAGTGGTAACCCGCGAATACACCATCAACATTNACAAGCGCATCCATGGAGTGG
GCTTCAAGAANCGTGCACCTCGGGCACTCAAAGAGATTCGGAAATTTGCCATGAAG
GAGATGGGAACTCCATATGTGCGCATTGACACCAGGCTCAACAAANCTGTCTGGGC
CAAAGGAATAAGGAATGTGCCATACCGAATCCGTGTGCGGCTGTCCANAAAACGTA
ATGAGGATGAAGATTCACCAAATAAGCTNTATACTTTGGTTACCTATGTACCTGTT
25 ACCACTTTCAAAAATCTACAGACAGTCAATGTGGATGANAAACNAATCGCTGATCGT
CAGATCAAANAAANT

Sequence ID - 1139

nt: 503

CAGCACTGCCAGTGGAGATGGGCGTCACTACTGCTACCCTCATTTACCTGCGCTG
30 TGGACACTGAGAACATCCGCCGTGTGTTCAACGACTGCCGTGACATCATTCAGCGC
ATGCACCTTCGTGAGTACGAGCTGCTCTAAGAAGGGAACCCCCAAATTTAATTAAA
GCCTTAAGCACAATTAATTAAAAGTGAAACGTAATTGTACAAGCAGTTAATCACCC
ACCATAGGGCATGATTAAACAAAGCAACCTTTCCCTTCCCCCGAGTGATTTTGCGAA
ACCCCCTTTTCCCTTCAGCTTGCTTAGATGTTCCAAATTTAGAAAGCTTAAGGCGG
35 CCTACAGAAAAAGGAAAAAAGGCCACAAAAGTTCCTTCTCACTTTCAGTAAAAATA
AATAAAACAGCAGCAGCAAACAAATAAAATGAAATAAAAGAAACAAATGAAATAAA
TATTGTGTTGTGCAGCATTAAAAAAAATCAAAATAAAAATTAAATGTGAGCAAAG

- 241 -

Sequence ID - 1148

nt: 587

TGAAAAATAAAGTTTTTATGTATATTCTACATATGTATATGTTGGTAGAAAGCAAA
AACGCTAGGTAAAAATAAATGTAATACAATTTTAGCTATGAACCAAAAAACCATTT
GTGGTGTGGATGCAAGAAAGTCTGGATGGGTGCAGAGTTCTCCATGTTTCACTTCT
5 GACATTTGAAAATACGCAGTTTGCATTTGATACGTCAAATGTTATTTTAAAGAAA
CCAATAAAATCATTAAAACCGAAAAGGCAGTTTTGCTTGTTTTTACCTTAGTTGGA
GTTATCTGCAATTGCCGTATTAGTGTTTTAAGGAACCTGTAAGTAAGCTCCTTAGT
CCCCTTTAGAGCTACGAAACATGTCAATTTTACTTTTCTCCAGCTTTTTGGAATCT
TATCTAAATTACCATGTAGAGTTCTGCATAGCTTCAAATTCTCTTAGCCAATGTGG
10 TCTGTAAGTGTCTATCGATGAATTTACCGTTAATTGCCGTAGTATACTGTCCTGT
ACCGGATGTGAAGAGGAGCAACTCTGCACAGTGCAGTGGTTGCTCCCATGGTAGGA
ANGAATGGCTTATCAATGGTCGGATTT

Sequence ID - 1160

nt: 650

GGAGGATGGAGCAGTGAGCGGGTCTGGGCGGCTGCTGGCAGCGCCATGGAGACGGT
15 ACAGCTGAGGAACCCGCCGCGCCGGCAGCTGAAAAAGTTGGATGAAGATAGTTTAA
CCAAACAACCAGAAGAAGTATTTGATGTCTTAGAGAACTTGGAGAAGGGTGAGTG
TAAAGAACTATAGGTAGGTCATTGGGTCCCAGTCTTTTTCTGCCCCAGAAGAAG
CAGAAGGATATGAACCTTTCAGCATTTGTTCTAGGTGGGGTGGAAGGTAAATTTACA
20 GCTTGTGATGTCCTTCTTCGCTTTACTCCAATCCCTATTATAGACAGATTTAGTGA
TTCCTGGTCTTTTTTAACACGAAGAATATCTATTGTTTTCTCTTTTGTAGGATCTGT
ATGATTTTATCTACTTAACAGATAGCACTAATTAGATTAAAATTCTATAAGAACT
TTTTAATTTGCTGTTTCATAATTTCTGATTGGTATGCAATAACTGTTTCAATGAAAA
TCAATGTAATTTAGTATTTTAATATTTGCACCTTTGTGAAATATAGTAAATAAATT
25 AAGCACTATCACCACCTTCACAGCTACTTAGGAGATCCACAATCCTGGGTGGGAG
CCAGTGGATTTCTGAAACACAGATTTGTTAATG

Sequence ID - 1165

nt: 502

CTCAAGTGAATCCTGGCTTCTTGGAAGCGCTTGCCTAGACGAGACACAGTGCATAA
30 AAACAACCTTTTGGGGGACAGGTATGTTTTCTTGCAGCTGCGGTTGTAAGGTCTTGG
CAAGACAAGCAGTGTGGCCAGAATTTTGAACTTCTGATGAATGTGTAATGCAAAGG
ACCTTGTACATTTTTTTTGTTCAGGTCTCAAATGAGCACATGAAGAGGTTGCT
GTGAACTTTAAGTGGCCCTACTGCGCAGAAGCATTGAGATGTCAGTTGATGATCT
GTAAGGGAACCTTGCTGATTTGGGAATGTGCTTAGGGAACACACATTCCTTTTGACA
35 GGGTCTGTCACTGGGTGGGTGATGAATTATACAGATGACATGTGCTTTTTTTTCTT
TTTTCAACCTCAATGGTATTCCTACAGGAAATGGATAACCATTTTAACTGTATTTT
TTGCAGCCCGTACCTTCTTGGGAATACAATTGTCTAACTTTTTTATTTTTGGTCT

- 242 -

Sequence ID - 1172

nt: 648

CCACAATAATAAGAGAAAAACAGGAGCAAAAGGATATACAAAACCACCAGAAAACA
AATAACAAAGTGACAGGAGTAAGTCCTTAAGTGGCAATAATAACCATGAATCTAAA
TGGATTCCATTTCCCCTTAAAAGATAAAGACATGCTGAATGGATAAAAAGCTGTC
5 ACCCAGTTATATGCTGCCTACAACAACTCACTTCACCTGTAAACATACATATGGA
TGGAAAGAGAAGGCATGGGAAAAGATACTCTACTCAAATGAAAACAAAACCAAAC
AAAGGTGGCTATTCTTATATGAGATAATACAGACATTAAATCAAAAACCTGGAAACA
AACACAAAGTCATTGTATAATGATGAATTCAATTATATCATGATGAATTCAATTAT
ATCCTCCTTCCTGATCAATTCAGAAAGGAGGATATAATCTTTTAAATATATATAC
10 ACCCAACACCAGAGCATATAAATATGTAAAGGAAGATAAAGGGAGTCCTGTGATCA
AGAATAAATATAACAATTATAAATATTTTATCTAAAGTGATAGACTGTAATA
CAATAATAGGGTGGTGACATTAACACCCCCTCTCACATTGGACTGATCATCTAGAA
GGGAGAAAAAGCTTTATGATTGGAAAAGCCAT

15 Sequence ID 1178

ATTGTGTTGGCCACCCGGGAATTCGCGGCCGCGTCGACCTACGCACACGAGAACAT
GCCTCTCGCAAAGGATCTCCTTCATCCCTCTCCAGAAGAGGAGAAGAGGAAACACA
AGAAGAAACGCCTGGTGACAGAGCCCCAATTCCTACTTCATGGATGTGAAATGCCCA
GGTGAGGAGACGGCTTGCTGTAGTGGGGAAAGCACTGGACCTCAACAGTTGGAAAA
20 TGTTGTAGTGTTAGCTGTCTCGTATCCTTGAAGCTGTGCAGCAGCTTCAGTTTCTT
CGCCTGTGGAAAATATTTTCCCTGATACTCTTAAAATTTGAATGTATGAGACTGGC
AAAGTTTTGCATCTTAGGAGGAGTGATTCATTTACCGTGATCTCTCATCACATTT
CACATACAACCCCTACGTTTTTTTTGTGTTGGGAAACAATGTAATGGATGATGAGTT
GGGCATAAGTGCAGGAAAGACGGGTGTAATAGAGGAAAAAAATGTTATCTGCTTTT
25 CTTTCAGGATGCTATAAAATCACACGGTCTTTAGCCATGCACAAACGGTAGTTTT
GTGTGTTGGCTGCTCCACTGTCCTCTGCCAGCCTACAGGAGGAAAAGCAAGGCTTA
CAGAAGGATGTTCTTCAGGAGGAAGCAGCACTAAAAGCACTCTGAGTCAANATGA
GTGGGAAACCATCTCAATAAACACATTTTGGAT

30 Sequence ID - 1180

nt: 622

CTTTTCCTCCCGCTGTCCCCACGGGAGGGGACTGCTCTCCCCGCTGCATCCTTT
CTGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAACCT
CGACATTCGGGAAGTGTTTTTTGAGAAGTCTCGGTCCGTAAGGGAAGTCTTCCAAGT
CCGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTTCGGCCACCCCCCAGATGAG
35 GCAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTTCTCTCAGCT
TCCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGTCCATGCTGCCA
TCAAGGCATTTATTGCAGTGTACTATTTGCTTCCAAAGGATCAGGCCCTGAGAACA

- 243 -

ATGACCTTATTTCTACAACAGTGTCTGGGTTGCGTGCCAGCAGATGCCTCAGATA
CCAAGAGATAACAAAGCTGCAGCTCTTTTGATGCTGACCAAGAATGTGGATTTTGT
GAAGGATGCACATGAAGAAATGGAGCAGGCTGTGGAAGAATGTGACCCTTACTCTG
GCCTCTTGAATGATACTGAGGAGAACAACCTCTGACANCCACAATCATGAGGATGAT
5 GTGTTG

Sequence ID - 1181

nt: 155

CGCCACTTATCCAGTGAACCACTATCACGAAAAAACTCTACCTCTCTATACTAAT
CTCCCTACAAATCTCCTTAATTATAACATTCACAGCCACAGAACTAATCATATTAA
10 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Sequence ID 1182

CATTGTGTTGGCNCCCGGGAATTCGCGGCCGCGTCGACTTTTTGTGTTGTTTGGAG
CAGAAATACTAAAGAAGATTCGGGGCCGAGTATCCACAGAAGTGGACGCAAGGCTC
15 TCCTTTGATAAAGATGCGATGGTGGCCAGAGCCAGGCGGCTCATCGAGCTCTACAA
GGAAGCTGGGATCAGCAAGGACCGAATTCTTATAAAGCTGTCATCAACCTGGGAAG
GAATTCAGGCTGGAAAGGAGCTCGAGGAGCAGCACGGCATCCACTGCAACATGACG
TTACTCTTCTCCTTCGCCCAGGCTGTGGCCTGTGCCGAGGCGGGTGTGACCCTCAT
CTCCCCATTTGTTGGGCGCATCCTTGATTGGCATGTGGCAAACACCGACAAGAAAT
20 CCTATGAGCCCCTGGAAGACCCTGGGGTAAAGAGTGTCACTAAAATCTACAACACTAC
TACAAGAAGTTTAGCTACAAAACCATTTGTCATGGGCGCCTCCTTCCGCAACACGGG
CGAGATCAAAGCACTGGCCGGCTGTGACTTCCTCACCATCTCACCCAAGCTCCTGG
GAGAGCTGCTGCAGGACAACGCCAAGCTGGTGCCTGTGCTCTCAGCCAAGGCGGCC
CAAGCCAGTGACCTGGAAAAAATCCACCTGGATGAGAAGTCTTTCCGTTGGTTGCA
25 CAACGAGGACCAGATGGCTGTGGAGAAG

Sequence ID - 1183

nt: 479

CGTGGCAGCCATCTCCTTCTCGGCATCATGGCCGCCCTCAGACCCCTTGTGAAGCC
CAAGATCGTCAAAAAGAGAACCAAGAAGTTCATCCGGCACCAGTCAGACCGATATG
30 TCAAAATTAAGCGTAACTGGCGGAAACCCAGAGGCATTGACAACAGGGTTCGTAGA
AGATTCAAGGGCCAGATCTTGATGCCCAACATTGGTTATGGAAGCAACAAAAAAC
AAAGCACATGCTGCCCAGTGGCTTCCGGAAGTTCCTGGTCCACAACGTCAAGGAGC
TGGAAGTGCTGCTGATGTGCAACAAATCTTACTGTGCCGAGATCGCTCACAATGTT
TCCTCCAAGAACCGCAAAGCCATCGTGGAAAGAGCTGCCCAACTGGCCATCAGAGT
35 CACCAACCCCAATGCCAGGCTGCGCAGTGAAGAAAATGAGTAGGCAGCTCATGTGC
ACGTTTTCTGTTTAAATAAATGTAAAAACTG

- 244 -

Sequence ID - 1185

nt: 628

CTTTGATTACCTTTGAGTATTAGGTTGAAAGCTTCTCTGTGCTTGATTGAACATTG
TGATGATGTTGATTGGGTCATGTCAGATTTAGACAGTGTTGTGTTTAAGATAAATG
TTTAATGGCTCTTAGCAGTGTTTCATGCCTCCCCCTTTTCCCCTGATACTTTAAAAAC
5 AGAATATACAGAAAAGGGGAGTTGGGTGAAGAATCACCATATTCTCATTACCAGAG
TAGTGTCTACCAGCTGTTTTTCACATTTTTCTGTTTCCTTCTGTCCTTGGAATCCTT
TTTTTAGATCCTTGTAATACTAGTAAAGATATCCACTCTGTGTTGTAAGCATTTT
TCCATTTTGCTCCATGGTCTTCATAATGCCCTGTGGTCCTTTATTAAGGGGATGCA
CCATGTAGAGGTGAAAGGCTTTCCTTGACTTGGCCACCATTTCTGTATTTTCCTTA
10 GAGGAGGAGGTTTCCAACATTTCTTTTTTAGAGACAGAGTCTCGTTCTGACACGCA
GGCAGGAGTGCAGTGGCATGATAACAGCTCACTGCAGCCTCGAACTCCTGGGCTCA
AGTTATCCTCCCACCTCAGCTTCCTGAGTAGCTAGGACTGCAGGTGCCTGCCACCA
CACCCAGCTAAT

15 Sequence ID - 1186

nt: 494

CAGCCCTCCGTCACCTCTTCACCGCACCTCGGACTGCCCCAAGGCCCCCGCCGCC
GCCTCCAGCGCCGCGCAGCCACCGCCGCGCGCCGCTCTCCTTAGTCGCCGCCA
TGACGACCGCGTCCACCTCGCAGGTGCGCCAGAACTACCACCAGGACTCAGAGGCC
GCCATCAACCGCCAGATCAACCTGGAGCTCTACGCCTCCTACGTTTACCTGTCCAT
20 GTCTTACTACTTTGACCGCGATGATGTGGCTTTGAAGAACTTTGCCAAATACTTTC
TTCACCAATCTCATGAGGAGAGGGGAACATGCTGAGAACTGATGAAGCTGCAGAA
CCAACGAGGGTGGCCGAATCTTCCTTCAGGATATCAAGAAACCAGACTGTGATGAC
TGGGAGAGCGGGCTGAATGCAATGGAGTGTGCATTACATTTGGAAAAAATGTGAA
TCAGTCACTACTGGAAGTGCACAACTGGCCACTGACAAAAATGAC

25

Sequence ID - 1188

nt: 599

GGGAGACAAGCCCAGCCTTTCGGCGAGNATACGTCTAACCTGTGCAACAGCCACT
ACATTACTTCAAAGTGAAGATCCTTCCTTTTGAGGGAGCAAGTCCTTCCCTTTCATT
TTTTCCAGTCTTCCTCCCTGTGTATTCTCATTCTCATGATTATTATTTTAGTGGGGGC
30 GGGGTGGGAAAGATTACTTTTTCTTTATGTGTTTGACGGGAAACAAAAGTAA
AATCTACAGTACACCACAAGGGTCACAATACTGTTGTGCGCACATCGCGGTAGGGC
GTGGAAAGGGGCAGGCCANAGCTACCCGCAGAGTTCTCAGAATCATGCTGAGAGAG
CTGGAGGCACCCATGCCATCTCAACCTCTTCCCCGCCCCGTTTTACAAAGGGGGAGG
CTAAAGCCCAGAGACAGCTTGATCAAAGGCACACAGCAAGTCAGGGTTGGAGCAGT
35 AGCTGGAGGGACCTTGTCTCCCAGCTCAGGGCTCTTTCCTCCACACCATTCAGGTC
TTTCTTTCCGAGGCCCTGTCTCAGGGTGAGGTGCTTGAGTCTCCAACGGCAAGGG
ACAAGTACTTCTTGATACCTGGGATACTGTGCCAGAG

- 245 -

Sequence ID 1189

GGGAGACAAGCCCAGCCTTTTCGGCGAGATACGTCTAACCCCTGTGCAACAGCCACTA
CATTACTTCAAACCTGAGATCCTTCCTTTTGAGGGAGCAAGTCCTTCCCTTTTCATTT
TTTCCAGTCTTCCTCCCTGTGTATTTCATTCTCATGATTATTATTTTAGTGGGGGCG
5 GGGTGGGAAAGATTACTTTTTCTTTATGTGTTTGACGGGAAACAAAACCTAGGTAAA
ATCTACAGTACACCACAAGGGTCACAATACTGTTGTGCGCACATCGCGGTAGGGCG
TGGAAAGGGGCAGGCCAGAGCTACCCGCAGAGTTCTCAGAATCATGCTGAGAGAGC
TGGAGGCACCCATGCCATCTCAACCTCTTCCCCGCCCGTTTTACAAAGGGGGAGGC
TAAAGCCCAGAGACAGCTTGATCAAAGGCACACAGCAAGTCAGGGTTGGAGCAGTA
10 GCTGGAGGGACCTTGTCTCCCAGCTCAGGGCTCTTTCCTCCACACCATTTCAGGTCT
TTCTTTCCGAGGCCCTGTCTCAGGGTGAGGTGCTTGAGTCTCCAACGGCAAGGGA
ACAAGTACTTCTTGATACCTGGGATACTGTGCCCAGAGCCTCGAGGAGGT

Sequence ID 1190

GTTTAAATTTGACAAACTAAAGCTNATNACTGCTATAAGAGTAATAACTGCTCATT
15 TTCCATAACTCATTCTTAAAGTTTTAGTAATGTAAAAGTTATTTTTTTTGCAGTAAG
TTATAATGATAGAAGCTTACATGTTTTTTTCATGCCTCATCTGTTTCCCCTTAAAC
TATAATTATCAGTAAAGTCCTGTGGTATTTTTTCAATTTGTAAGAACTAGGCTATA
TATACATTGGGAAAAACAGCCTTCATTTGTCAATGCACCTAGTGTTCCAAAGGTTTC
20 TGGTAATTGTGTGCTATTGCTTTTTTGTGACTTGCAAAAAAAAAAAAAAAAAAATT
ACTATGACTTGNGGTAGCCCTGCAACCTTCGGAAGTGCTTAGCCCAGTCTGACCAT
ACATTTATATTTTANAATGCTTAGGTAAATAAATAATATGCCTAAACCCAATGCTAT
AAGATACTATATAATATCTCATAATTTTAAAAATCACTGTTTTTGTATAATAATAAA
ACAAGGCAGGCAAGCTGTTCTACAATGACTGTTGGTAAGGGTGCTGAGGAAGAAAA
25 ACAAACAATCTTGATTTCAGGGATAGTGAATAGACAAAAAATGTCCTAATCAATGAA
GCTGTGTGATGATTCTGATTGACAGAGA

Sequence ID 1191

GTGCAAAGTGTTATATCCACTTTCAACAAAGAGAGAAGCTGAAAAGCTAACCCAAT
30 GTTAATTTTGGATCACACACATTCAGTGTAGACTTTAAGATTTTACTTCTGTTGGA
GTAGCTATATTATTTCTAGTTAAAAAACTCTCTATATACATATTTATTTGTTTTTC
TACTTGTTTAATATTTTTTCTCTTCCAATTAGGAACTCAATATGGAATAAAAAATAT
TTAAATGTATTTTACTCAAACGTGTGTGTATATATGTTTGTGTGCATGATAAGGAG
AGTGAGAGCAAGAGTAAGAGAGAGAGAGCACGCATAGATGGAAGCACACATTTAAT
35 GTCTATGAAATGAGAAAACATTAAGGCTAAGATATTTTTCTTCTGAACTAGCAGA
TTGTATCAATGGCTGGTCACTTAAATTAATCAGTTTGTAAGATATTTAAAAGGTA
TGTCTACCTTCTTGCAATTAATTTGATTATGTTCTAATGGCATGGCAAGAGAAATG

- 246 -

AAAGAAGATAACTAAAAGTTAAAAGTCGTTGCATGTTTTTGTTCAGCATACCCTT
CTTTCAGGCTACCGAATAACCTTGATTGACATTGGATTAGTAGAGAATACCTCAT
TGGTAGAGCATATCGCAGCANCTACACTAGAAAACAT

5 Sequence ID 1192

GTCTGGAACTCCAGACCTCAGGTGATACCCCTGCCTCAGCCTCCCAATGTGCTGGG
ATTACAGCTGTGAAGCCACCGCGCCCGGCTGCTGTGATAGTTGAGATGTAAACCAA
AAATAAAATTCTAAGCCACCCAATCCGACTGAATGGACCCTTCCTGTTGAGCAAGG
ACATTCCAAAGTAAACTGAAAAGACCAGCTTAGGCCATGATGGGAAGGGGAGGTGT
10 CAACATGCCTCATTCTACCTTCCTCCCTCTGGAATCCAGACACAACCTGACCAGCAT
TAACATTAAAACAGAGATCTTAAGCTGGGCACGGTGGCTCATGCCTGTAATCCCAG
CACTTTGGGAGGCCAAGGTGGGATCACCTGAGGTCGGAAGTTCAAGACCAGCCTGG
CCGGTATGGTGAAGCCATGTCTCTACTGAAAATGCAAATTTGGCCGGACATTGTGG
TGCA

15

Sequence ID 1193

TNCNTTTTTTTTCCCNCGGGAAAGCGCGCCATTGTGTTGGTCCCCGGGAATTCGCG
GCCGCGTCGACGAGAAATGGCTTGAACCCAGTAGGCAGAGGTTGTAGTGAGCCAG
AATNGGNCACCTGCACNTTTANCCNTGGGTGACAAAANTGAAAACCTTTGTCTNAAA
20 AAAAAAAAAAAAAAAAAAATTTAANTNAAATNAAAAANCCTTTNCNTTNTTTTTTNAAAN
NGGGGGGGGNNTTTTTNNGGNTTNGNNNTGGTAAAAANTNNNTTTTTTTTTTTTA
GGGGCCNANNCCCCNTTTTANAAAANCCNGNTTTTNAAAAAANTTTTTTNCCNCN
NTTNGGGGGGGGGGNTTTTNANCNNTNTTNGGGGGGGGNCCCCNTTANNACCNNC
AAANTTTTTANTTTTTTGNNAANNCCCCCTTTTTTNNTTTTTTTTGNNGGGGGG
25 GGGNNGCCCCCNCCTTTNNGGGGGGGGGNTTNNGNAAAANNACTTTTNAAAANNA
AGGGNNGGGGGNANATNNCCCCCCCNGGNTTTTTTTTTTTAAAAANTNAANNNGGGG
GGGNNNCTNANTNGGGGCNCCCANNGGGGGNTTANAANNATTTTCTNCCCAAACCC
CCNGNTTTTATNNCCCCCCCCCCCCCNCNNNNGAANGGGNGGNCCNTTTTTTTTATT
TTTNNGGNGGGNAAAAAANTTTNAAAAANNANNTNTTTTTTCCCCCCCCCCCCCNC
30 TTTTNGGNAAANCCNNGGGGGGNTCCTTTTTNAAANNNCCCCCAAAAAAANTTT
TTTTNTTNTNTTTTTTCTCTNNGGGGNCCNNANTTNTANANTTTTNCNCNAAAAAAA
ANGGGNCCCCTTTTTTTTNCNGGNNGGNNCCCAAAANNTTTTTTTTNAAAAAAA
AAAA

35 Sequence ID 1195

GTTCGTGACNTTCGGAGCTACCTGACAGAGCAGAGTCAACCAGGNTCTGCCCAAAG
AGAGTGTTAGGCCTGAGCTTGAGAGCCCTGGAGAGACGTGTGCACAAAATGTGACC

- 247 -

TGAGGCCCTAGTCTAGCAAGAGGACATAGCACCCCTCATCTGGGAATAGGGAAGGCA
CCTTGCAGAAAATATGAGCAATTTGATATTAACATCTTCAATGTGCCATAGA
CCTTCCCACAAAGACTGTCCAATAATAAGAGATGCTTATCTATTTTA

5 Sequence ID - 1196

nt: 412

GTCGACGCGGCCGCGGTCTGCTGGAGNCGATCAACTCTAGGCTCCAACCTCGTTATGA
AAAGTGGGAAGTACGTCCTGGGGTACAAGCAGACTCTGAAGATGATCAGACAAGGC
AAAGCGAAATTGGTCATTCTCGCTAACAACTGCCCAGCTTTGAGGAAATCTGAAAT
AGAGTACTATGCTATGTTGGCTAAAACCTGGTGTCCATCACTACAGTGGCAATAATA
10 TTGAACTGGGCACAGCATGCGGAAAATACTACAGAGTGTGCACACTGGCTATCATT
GATCCAGGTGACTCTGACATCATTAGAAGCATGCCAGAACAGACTGGTGAAAAGTA
AACCTTTTCACCTACAAAATTTACCTGCAAACCTTAAACCTGCAAATTTTCCTT
TAATAAAATTTGCTTGTTTT

15 Sequence ID 1197

CCGCCAACATGGGCCGCGTTCGCACCAAACCGTGAAGAAGGCGGCCCGGGTCATC
ATAGAAAAGTACTACACGCGCCTGGGCAACGACTTCCACACGAACAAGCGCGTGTG
CGAGGAGATCGCCATTATCCCCAGCAAAAAGCTCCGCAACAAGATAGCAGGTTATG
TCACGCATCTGATGAAGCGAATTCAGAGAGGCCAGTAAGAGGTATCTCCATCAAG
20 CTGCAGGAGGAGGAGAGAGAAAGGAGAGACAATTATGTTCTGAGGTCTCAGCCTT
GGATCAGGAGATTATTGAAGTAGATCCTGACACTAAGGAAATGCTGAAGCTTTTGG
ACTTCGGCAGTCTGTCCAACCTTCAGGTCACTCAGCCTACAGTTGGGATGAATTTT
AAAACGCCTCGGGGACCTGTTTGAATTTTTTCTGTAGTGCTGTATTATTTTCAATA
AATCTGGGACAA

25

Sequence ID 1198

CAGAGGTGGGAGGATTGCTTCAGTTCAAGAGTTTGAGACCAGCCTGGGTAAACATGG
CGAAACCCTGTCTTTACAAAAAATGCAAACCTTTGCCGCATGTGTTGGGGTGCGCC
TGTAGTCCCAGCTTCTCGGGAGGCTGAGGTGGGGGGACCACCTGAGCCATGGAGGT
30 TGAGGCTGCAGTGAGCCGTGATACCACCACTGTACTCTAGCCTGGGCCATAGAGTG
AGACACCCTGCCTCAGAAATA

Sequence ID - 1199

nt: 439

CCCATCCCCTCGACCGCTCGCGTCGCATTTGGCCGCCTCCCTACCGCTCCAAGCCC
35 AGCCCTCAGCCATGGCATGCCCCCTGGATCAGGCCATTGGCCTCCTCGTGGCCATC
TTCCACAAGTACTCCGGCAGGGAGGGTGACAAGCACACCCTGAGCAAGAAGGAGCT
GAAGGAGCTGATCCAGAAGGAGCTCACCATTGGCTCGAAGCTGCAGGATGCTGAAA

- 248 -

TTGCAAGGCTGATGGAAGACTTGGACCGGAACAAGGACCAGGAGGTGAACTTCCAG
GAGTATGTCACCTTCCTGGGGGCCTTGGCTTTGATCTACAATGAAGCCCTCAAGGG
CTGAAAATAAATAGGGAAGATGGAGACACCCTCTGGGGGTCTCTCTGAGTCAAAT
CCAGTGGTGGGTAATTGTACAATAAATTTTTTT⁴TTTGGTCAAATTTAA

5

Sequence ID - 1200

nt: 526

CTGGAGACGACGTGCAGAAATGGCACCTCGAAAGGGGAAGGAAAAGAAGGAAGAAC
AGGTCATCAGCCTCGGACCTCAGGTGGCTGAAGGAGAGAATGTATTTGGTGTCTGC
CATATCTTTGCATCCTTCAATGACACTTTTGTCCATGTCACTGATCTTTCTGGCAA
10 GGAAACCATCTGCCGTGTGACTGGTGGGATGAAGGTAAAGGCAGACCGAGATGAAT
CCTCACCATATGCTGCTATGTTGGCTGCCCAGGATGTGGCCCAGAGGTGCAAGGAG
CTGGGTATCACCGCCCTACACATCAAACCTCCGGGCCACAGGAGGAAATAGGACCAA
GACCCCTGGACCTGGGGGCCAGTCGGCCCTCANAGCCCTTGCCCGCTCGGGTATGA
AGATCGGGCGGATTGAGGATGTCACCCCATCCCCTCTGACAGCACTCGCAGGAAG
15 GGGGGTCGCCGTGGTTCGCCGTCTGTGAACAAGATTCCTCAAATATTTTCTGTAA
TAAATTGCCTTCATGTAAACTG

Sequence ID - 1201

nt: 613

CTTAAGTATGCCCTGACAGGAGNATGAAGTAAAGAAGATTTGCATGCAGCGGTTCA
20 TTAAAATCGATGGCAAGGTCCGAACCTGATATAACCTACCCTGCTGGATTCATGGAT
GTCATCAGCATTGACAAGACGGGAGAGAATTTCCGTCTGATCTATGACACCAAGGG
TCGCTTTGCTGTACATCGTATTACACCTGAGGAGGCCAAGTACAAGTTGTGCAAAG
TGAGAAAGATCTTTGTGGGCACAAAAGGAATCCCTCATCTGGTGACTCATGATGCC
CGCACCATCCGCTACCCCGATCCCCTCATCAAGGTGAATGATACCATTGAGATTGA
25 TTTAGAGACTGGCAAGATTACTGATTTTCATCAAGTTCGACACTGGTAACCTGTGTA
TGGTGACTGGAGGTGCTAACCTAGGAAGAATTGGTGTGATCACCAACAGAGAGAGG
CACCCCTGGATCTTTTGACGTGGTTCACGTGAAAGATGCCAATGGCAACAGCTTTGC
CACTCGACTTTCCAACATTTTTGTTATTGGCAAGGGCAACAAACCATGGATTTCTC
TTCCCCGAGGAAAGGGTATCCGCCTCACCATTGCTGAAGAGAGAGACAAAAGA

30

Sequence ID 1202

GGAATTCGCGGCCGCGTCGACCTCTGCTCGAATTGACAGAAAAGGATTCTGTGAAG
AGTGATGAGATTTCCATCCATGCTGACTTTGAGAATACATGTTCCCGAATTGTGGT
CCCCAAAGCTGCCATTGTGGCCCGCCACACTTACCTTGCCAATGGCCAGACCAAGG
35 TGCTGACTCAGAAGTTGTCATCAGTCAGAGGCAATCATATTATCTCAGGGACATGC
GCATCATGGCGTGGCAAGAGCCTTCGGGTTCAGAAGATCAGGCCTTCTATCCTGGG
CTGCAACATCCTTCGAGTTGAATATTCCTTACTGATCTATGTTAGCGTTCCTGGAT

CCAAGAAGGTCATCCTTGACCTGCCCCTGGTAATTGGCAGCAGATCAGGTCTAAGC
AGCAGAACATCCAGCATGGCCAGCCGAACCAGCTCTGAGATGAGTTGGGTAGATCT
GAACATCCCTGATACCCAGAAAGCTCCTCCCTGCTATATGGATGTCATTCCTGAAG
ATCACCGATTGGAGAGCCCAACCACTCCTCTGCTAGATGACATGGATGGCTCTCAA
5 GACAGCCCTATCTTTATGTATGCCCCTGAGTTCAGTTCATGCCACCACCGACTTA
TACTGAGGTGGATCCCTGCATCCTCAACAACAATGTGCAGTGAGCAT

nt: 692

10 TGCAGAGGGGTCCATACGGCGTTGTTCTGGATTCCCGTTCGTAACCTTAAAGGGGAAAC
TTTCACAATGTCCGGAGCCCTTGATGTCCTGCAAATGAAGGAGGAGGATGTCCTTA
AGTTCCTTGCAGCAGGAACCCACTTAGGTGGCACCAATCTTGACTTCCAGATGGAA
CAGTACATCTATAAAAGGAAAAGTGATGGCATCTATATCATAAATCTCAAGAGGAC
CTGGGAGAAGCTTCTGCTGGCAGCTCGTGCAATTGTTGCCATTGAAAACCCTGCTG
ATGTCAGTGTTATATCCTCCAGGAATACTGGCCAGAGGGCTGTGCTGAAGTTTGCT
15 GCTGCCACTGGAGCCACTCCAATTGCTGGCCGCTTCACTCCTGGAACCTTCACTAA
CCAGATCCAGGCAGCCTTCCGGGAGCCACGGCTTCTTGTTGGTTACTGACCCCAAGG
CTGACCACCAGCCTCTCACGGAGGCATCTTATGTTAACCTACCTACCATTTGCGCTG
TGTAACACAGATTCTCCTCTGCGCTATGTGGACATTGCCATCCCATGCAACAACAA
GGGAGCTCACTCAGTGGGTTTAATGTGGTGGATGCTGGCTCGGGAAGTTCTGCGCA
20 TGCGTGGCACCATTTCCCGTGAACACCCATGGGAGGTCATGCCTGATCTGTACTTC
TACAGAGATCCTGAAGAGAT

25 TTTTTTTTTTTTCTCTGCGGGAAAGCGCGCCATTGTGTTGGTACCCGGGAAATTCTG
CGGCCGCGTCGACACAGGCCCCAGCATCAAGATCTGGGATTTAGAGAGGAAAGATC
ATTGTAGATGAACTGAAGCAAGAAGTTATCAGTACCAGCAGCAAGGCAGAACCACC
CCAGTGCACCTCCCTGGCCTGGTCTGCTGATGACACAGGTTGGGCNNGGNNCNCNNGG
GGNGGNNNNNGNNNNGCNGNNGGNNCNGNNNNCNNNNNGCNNNNNGNNNNTNNNCNNN
GNNCNNNNNNNNNNNNNNNNNNNGNTCNNNGNNGCNGGGGGCCNGGNCGNCGCGGNCGCG
30 NNTNNNNNGGGTNCNNNCNCNNNGGCGCGC

[illegible]

- 250 -

TNNNNNTNNCNNCNANTAAANNCANNTCNANNNNNANNNNAATTACTTNNANGTNNNTC
ACN

Sequence ID - 1207

nt: 642

5 ACGAGAAGCCAGATACTAAAGAGAAGAANCCCGAAGCCAAGAAGGTTGATGCTGGT
GGCAAGGTGAAAAAGGGTAACCTCAAAGCTAAAAAGCCCAAGAAGGGGAAGCCCCA
TTGCAGCCGCAACCCTGTCCTTGTGAGAGGAATTGGCAGGTATTCCCGATCTGCCA
TGTATTCCANAAAGGCCATGTACAAGAGGAAGTACTCAGCCGCTAAATCCAAGGTT
GAAAAGAAAAAGAAGGAGAAGGTTCTCGCAACTGTTACAAAACCAGTTGGTGGTGA
10 CAAGAACGGCGGTACCCGGGTGGTTAAACTTCGCAAAATGCCTAGATATTATCCTA
CTGAAGATGTGCCTCGAAAGCTGTTGAGCCACGGCAAAAACCCCTTCAGTCAGCAC
GTGAGAAAACCTGCGAGCCAGCATTACCCCCGGGACCATTTCTGATCATCCTCACTGG
ACGCCACAGGGGCAAGAGGGTGGTTTTCTGAAGCAGCTGGCTAGTGGCTTATTAC
TTGTGACTGGACCTCTGGTCCTCAATCGAGTTCCTCTACGAAGAACACACCAGAAA
15 TTTGTGATTGCCACTTCAACCAAAATCGATATCAGCAATGTAAAAATCCCAAACA
TCTTACTGATGCTTACTTCAAAAAGA

Sequence ID 1208

CCCTATACCTTCTGCATAATGAATTANCTAGAAATAACTTTGCAAGGGAGAGCCAA
20 AGCTAAGACCCCCGAAACCAGACGAGCTACCTAAGAACAGCTAAAAGAGCACACCC
GTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGGCGACAAACCTACCGAG
CCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAACTTTAAATTTGCCCA
CAGAACCCTCTAAATCCCCTTGTAATTTAACTGTTAGTCCAAAGAGGAACAGCTC
TTTGGACACTAGGAAAAAACCTTGTAGAGAGAGTAAAAAATTTAACACCCATAGTA
25 GGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACACCCACTACCTAA
AAAATCCCAAACATATAACTGAACTCCTCACACCCAATTGGACCAATCTATCACCC
TATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCCTCCGCATAAG

Sequence ID - 1209

nt: 620

30 CTCTCCTGTCAACAGCGGCCAGCCTCCCAACTACGAGAATGCTCAAGGAGGAGCAG
GAAGTGGCTATGCTGGGGGCGCCCCACAACCCTGCTCCCCGACGTCCACCGTGAT
CCACATCCGCAGCGAGACCTCCGTGCCCGACCATGTCGTCTGGTCCCTGTTCAACA
CCCTCTTCATGAACACCTGCTGCCTGGGCTTCATAGCATTGCGCTACTCCGTGAAG
TCTAGGGACAGGAAGATGGTTGGCGACGTGACCGGGGGCCAGGCCTATGCCTCCAC
35 CGCCAAGTGCCTGAACATCTGGGCCCTGATTTTGGGCATCTTCATGACCATTTCTGC
TCGTGATCATCCCAGTGTTGGTCGTCCAGGCCAGCGATAGATCAGGAGGCATCAT
TGAGGCCAGGAGCTCTGCCCCGTGACCTGTATCCCACGTACTCTATCTTCCATTCT

- 251 -

CGCCCTGCCCCCAGAGGCCAGGAGCTCTGCCCTTGACCTGTATTCCACTTACTCCA
CCTTCCATTCTCTCGCCCTGTCCCCACAGCCGAGTCCTGCATCAGCCCTTTATCCTC
ACACGCTTTTCTACAATGGCATTCAATAAAGTGTATATGTTTCTGGTGCTGCTGTG
ACTT

5

Sequence ID 1210

TTCGTAATTAGAATACTGTTTGGACTTGCTCAACAAGCACCTTATCTTAACAAAAA
GTAACCTTATAGAAAAGGGAGACATTCATTTAACTTCAAGCCCATATTATTCTTAAA
AGCTGACTCTTGAAATAGTATTTATTGAGTCATAGTGGAGTCATGGGACTTTTTTAA
10 GGGCCGGAAGGGACTATTTAGATCATCCAGTCCCACCCTGTCATTTTATGGAGGAG
GAAACTGAGGCCTAGATAAGATAACCAGTTAGTGGGTCCACTGACCTTTAGGACAG
TAGTCTATCCGTAAGAGACAACATGGAGAAAGAAATACAACGTTTTTATAGTGAAT
TATCATCTTACAAAGAATATTCTTCCCATATCGCACTTTTAAAAAGTGGGTACCTT
AGTCAAATAGGAGAAAAAACCACTTGAGTAGTTTCATCCTCAGGTTTTAGGTGAGG
15 AAAGTACTCAGATTAAATAACTTTAAGCACACAGAGCCTGAATGATAGTCTTA
TTTGAGCTCATCTGTGCTTTTAAATGTGTACTACGTTAGGTGTTTTCACTTGCATTT
CCTTTAGTCTTATTTGAGCTCATCTGTGCTTTTAAATGTGTACTACGTTAGGTGTTT
TCACTTGCATTTCTTGTGTTGACGTTGACAATAAATCGTGAAGCTGCCTTATCTAA
GGAAGTCCTAAAGTAAATCATTGGAACACA

20

Sequence ID 1211

CCATTGTGTTGGNACCCGGGAATTCGCGGCCGCGTCGACGGAGTTTTACCTTATTA
CACTTTAATCTCTGGATTTACCCCATCTCATTCTCTTTTAGGAAAAGTGTGTA
TGTGGTGGCTTTGATGGTTCTCATGCCATCAGTTGTGTGGAAATGTATGATCCAAC
25 TAGAAATGAATGGAAGATGATGGGAAATATGACTTCACCAAGGAGCAATGCTGGGA
TTGCAACTGTAGGGAACACCATTTATGCAGTGGGAGGATTCGATGGCAATGAATTT
CTGAATACGGTGGAAAGTCTATAACCTTGAGTCAAATGAATGGAGCCCCTATACAAA
GATTTTCCAGTTTTTAACAAATTTAAGACCCTCTCAAACCTAACAGGCTTAGTGATGT
AATTATGGTTAGCAGAGGTACACTTGTGAATAAAGAGGGTGGGTGGGTATAGATGT
30 TGCTAACAGCAACACAAAGCTTTTGCATATTGCATACTATTAAACATGCTGTACAT
ACTTTTTGGGTTTATTTGGAAAGGAATGCAAAGATGAAGGTCTGTTTTGTGTACTT
TTAAGACTTTGGTTATTTTACTTTTTGGAAAAGAATAAACCAAGAATTGATTGGGC
ACATCATTTCAAGAAG

35

Sequence ID - 1212

nt: 374

AGAGCAGCAGCCATGGCCCTACGCTACCCTATGGCCGTGGGCCTCAACAAGGGCCA

- 252 -

CAAAGTGACCAAGAACGTGAGCAAGCCCAGGCACAGCCGACGCCGCGGGCGTCTGA
CCAAACACACCAAGTTCGTGCGGGACATGATTCGGGAGGTGTGTGGCTTTGCCCCG
TACGAGCGGCGCGCCATGGAGTTACTGAAGGTCTCCAAGGACAAACGGGGCCCTCAA
ATTTATCAAGAAAAGGGTGGGGACGCACATCCGCGCCAAGAGGAAGCGGGAGGAGC
5 TGAGCAACGTACTGGCCGCCATGAGGAAAGCTGCTGCCAAGAAAGACTGAGCCCCT
CCCCTGCCCTCTCCCTGAAATAAAGAACAGCTTGACAG

Sequence ID - 1213

nt: 567

GAATTATTGACTTTGAATTGCATTTTCAGTACCATGAAGTCAAAGTCAGTGGTGTAT
10 TTGCTCATTTGTTTCAATTCTTTCTTTTCCACCAACATTACTGCCTGCAGAGCCAGAG
GTGAGTGCAGAAATCCTGTCAATTCGTCACTTGTGGACAACCTGCAGCTTGCCACA
GCCTACAGTTCCACCACTGTGACCTCTGAAAACCTCCTGAACAAAAGGAAGGAGAC
TTGGAAATCCTGAATGGGCTTGGAGACATTAAGGGAGAAGTGCCTCCCTGGACCAA
GGCAGAATTCAATAGAACCAGCAAGAAATTTTCCTATGAATGGGAAAGCAGGTGGC
15 AGGGGGCAGGGGTGGAAAAGCTTTGTACAGGAATTGTGGAAAAGCTTTTGCATTAT
CTCTAGTCTGAAAGTCACATTTCTCAGTTCCTTTCCACTCTCTTCTGTCAACTTGC
TGTGAGTAAATGACATCTGTCACCTGTGACACGGGCCAGGGACTATCACCATATGG
CCCCACACATTATCTAGTACCAGCCTGCCTGGGCCATGCCTTTTCCAGTCACTGT
ACCAGCC

20

Sequence ID - 1214

nt: 620

CTCTCCTGTCAACAGCGGCCAGCCTCCCAACTACGAGAATGCTCAAGGAGGAGCAG
GAAGTGGCTATGCTGGGGGCGCCCCACAACCCTGCTCCCCGACGTCCACCGTGAT
CCACATCCGCAGCGAGACCTCCGTGCCCCGACCATGTCGTCTGGTCCCTGTTCAACA
25 CCCTCTTCATGAACACCTGCTGCCTGGGCTTCATAGCATTCGCCTACTCCGTGAAG
TCTAGGGACAGGAAGATGGTTGGCGACGTGACCGGGGGCCAGGCCTATGCCTCCAC
CGCCAAGTGCCTGAACATCTGGGCCCTGATTTTGGGCATCTTCATGACCATTCCTGC
TCGTCATCATCCCAGTGTTGGTCGTCCAGGCCCAGCGATAGATCAGGAGGCATCAT
TGAGGCCAGGAGCTCTGCCCCGTGACCTGTATCCCACGTACTCTATCTTCCATTCCCT
30 CGCCCTGCCCCCAGAGGCCAGGAGCTCTGCCCTTGACCTGTATTCCACTTACTCCA
CCTTCCATTCCCTCGCCCTGTCCCCACAGCCGAGTCCTGCATCAGCCCTTTATCCTC
ACACGCTTTTCTACAATGGCATTCAATAAAGTGTATATGTTTCTGGTGCTGCTGTG
ACTT

35

Sequence ID 1215

CACAAGATAGAATGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTTAA
GTGACAGTGCCATAGTTTGGACAGTACCTTTCAATGATTAATTTTAATAGCCTGTG

- 253 -

AGTCCAAGTAAATGATCACTTTATTTGCTAGGGAGGGAAGTCCTAGGGTGGTTTCA
GTTTCTCCCAGACATAACCTAAATTTTACATCAATCCTTTTAAAGAAAATCTGTAT
TTCAAAGAATCTTTCTCTGCAGTAAATCTCGCAGGGGAATTTGCACTATTACACTT
GAAAGTTGTTATTGTTAACCTTTTCGGCAGCTTTTAATAGGAAAGTTAAACGTTTT
5 AAACATGGTAGTACTGGAAATTTTACAAGACTTTTACCTAGCACTTAAATATGTAT
AAATGTACATAAAGACAAACTAGTAAGCATGACCTGGGGAAATGGTCAGACCTTGT
ATTGTGTTTTTTGGCCTTGAAAGTAGCAAGTGACCAGAATCTGCCATGGCAACAGGC
TTTAAAAAAGACCCTTAAAAAGACACTGTCTCAACTGTGGTGTAGCACCAGCCAG
CTCTCTGTACATTTGCTAGCTTGTAGTTTTTCTAAGACTGAGTAAACTTCTTATTTT
10 TAGAAAGTGGAGGTCTGGTTTGTAACCTTTCCTTGTACTTAATTGGGTAAAAGT

Sequence ID - 1216 nt: 484
CAACCTTAGCCAAACCATTTACCCAAATAAAGTATAGGCGATAGAAATTGAAACCT
GGCGCAATAGATATAGTACCGCAAGGGAAAGATGAAAAATTATAACCAAGCATAAT
15 ATAGCAAGGACTAACCCCTATACCTTCTGCATAATGAATTAAGTAACTAGAAATAACTTT
GCAAGGAGAGCCAAAGCTAAGACCCCCGAAACCAGACGAGCTACCTAAGAACAGCT
AAAAGAGCACACCCGTCTATGTAGCAAATAGTGGGAAGATTTATAGGTAGAGGCG
ACAAACCTACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAAC
TTTAAATTTGCCCACAGAACCCTCTAAATCCCCTTGTAATTTAACTGTTAGTCCA
20 AAGAGGAACAGCTCTTTGGACACTAGGAAAAAACCTTGTAAGAGAGAGTAAAAAATT
TAACACCCATAGTAGGCCTAAAAGCAGCCACCAATT

Sequence ID 1217
GACAGGCGGGGGCCAGCGGCCGGGTGAAGGCCGGGTGGCTCTGTGAATCAAAGGA
25 GAGTCCCAGAAAACCTGTGACTGTTGAAGAAAATTCATCTGTGAATTTTTATATTC
AAGGAGTCAGTATTTATATTCATCTTTTAACTGGGAAGATTTATATTTTACTTTA
AACTTCTTGATAATAATTTACAATGAATGGACACAGTGATGAAGAAAGTGTTAGA
AACAGTAGTGGAGAATCAAGGTAAGTAAGCACTTTGTTATCAATTGTTTACTATGA
AGAGAGTTGAAAACCTTGACTTTTTTTCTTTATTGTTATTGTTGTTATTTAGTTTTCC
30 TCATAGGTAGCAGAGTTTTTCAGGTTTTCTCTTAGCTATCCAAATACTAAAAAAT
TCTGATATACGAACCTTTTTTTCATAATACAGGTTTTTAATTATATTTTTTCATTCAGA
TACACAGTAGATCTTAAATATAGAAAGTTTTTGTTTACTTAAATCTATTTGGAAGT
TTATATTTGAGCTAATAATTAAGCTGGAGCATGTATAATAGATTTAAATTGTTTTG
ACTGTTAGTGAAATTT

35

Sequence ID 1218
CTCACTTGGTGGGTGAGCCTCCAATGACTACACCCAAGGAGGATTTAACACAGGGA

- 254 -

TTTTATGACTTGCAACAAGTCAGGAGGACATGGGGTTGGGGTAGTTCAGCAGTGCC
TGTCTGAACAAAGGTGAAAATTGGGCTTTTATTGGGCTGATCAAGGGGGAGTAAAG
GCAGCCAGGAGCAGTCGCCTGTCATGCTTCTACCTATATTGCATGTATAGAAAAGG
GAAAATAAACTCCTTCCTGGGCAGGGTTTTAGTATGCTAAGGAGGGGAGTTATTCA
5 ACTTCAATCCAACCTCAAGCATCAGCATTGCTGCGTCCATCCCAGTTTTTGTTTTGCT
GGGGCTGAACTTCTTCCTATAACTTTTTTGAAACAACAAGAACTCAAGGTGTGACAG
TTACAAGTGGGCCCTTTTTTCACAGTGTGTACCTAAACACGTGAGGACCCTGGATTA
CAGAATGACAGACTCGAAGTGACTCAAGTTCGGTTGTTTCATCTTTAGATGGTAAA
GATGGCTGTACGTACTATCCTTGCTTATTTCCAATCTATTGTTTAACTCTTGTAT
10 ATGTAATACCGCAGAGGCTAGAGATACAACCTTTGACCAAATGAGTGAATTCAAGT
AATCCATTACTAATGTGATCTGGAAACAAACATGGTGTTGAATGTGCATATGT

Sequence ID - 1219

nt: 559

CTTGGCAGCTCCGTTATGTGCCAGCTCTTTGCAAGGGCATACTGGGAAATGAGTG
15 GAGATAAAGGACCCAATCATAAGCATTTTACAGTATGGATACCCCATTTTAAAAAG
GTAAACTGAGGCACAATGCAATTTTTTTTTTTTTTTTAAAGGAGTTTATTTGAGCAA
CAGTGATTCATGAATCAGGCAGCACCAACCAGAAGGAGGCTTTGCTGAANAAGGA
TGAGGGACAAGCATTTATAAAGTGAATGTAGATGTAATACAAAGAAAATATTTGAA
CCGGGTGCGGTGGCTTACACTTGTAATCCCAACACTTTGGGAGGCCAAGGCGGGCA
20 GATCACAAGATCAAGAGATCGAGACCATCCTGGTCAACATGGTGAAACCCCATCTN
TACTAAAAAATACAAAAATTANCTGGGCGTGGTGGTGCCTGTAGTCCCAGCT
ACTTGGGCGGCTGAGGCAGGANAATTGCTTGAACCCGGGAGGTGGAGGTTGCAGTA
AGCCGAGATTGCACCATTGCACTACTCCAGCCTGGTGACAGAGAGAGACTCCATC

25 Sequence ID 1220

GANNNGTGCGATANNATGNNTGTCTTTTTTTTAAAGTNTTTCNNATNGNAGNGAAN
CCCCCNNANNTNNCATAANGAGAGATNACTACNGTACANATAGNGNCANACNGATA
GTAGTANCAANATTGTNTTAGCTANATNANTCAATAGATATCNAGATANAANAANA
NCNNGGATATACAGCGATGTNTNANNGGNNNNNNNANGGAACGAACATCNACNTTA
30 ANNATAAGCTNGNGGAGAGAGACANGTANGTTATANANNAGAATNGNAGTAGGNGT
GATCATAATAGNNNNNANNTANTATATANGATNTTANTGNNCTNTNNTNNGTTTAT
CNNNAATNTCTATNCTNGAGAGNAGCNNNATNNNNAGGCGANGANATTGGGNNNTN
CTCNTNATAGANANCTGGTGTCNNANAANTACNTCATCTATTNANCTCTCACNANA
TGGNANNATANAGNAGNGNNNTNNANAGGANTANGCATAGNGNNTNNCTNAAACAA
35 AANNNATAAGANNTCTCGNNAANANGGGCCTNTNNTNTAGCGAGGNNTTANTTTNT
ATANTTNTTCNCTCTTNNAATANNTANGATANATGANCTNGNNGTGATANATANNN
NNTACNGTNAANNTNTANTCNTATAATAGATANAAATATAGGATNTTNCTCTGGCN

- 255 -

GGTNGAANANTTNNTNCNNTTTTNAATAATGNTGTTAGNGACNGNGNTNTNANANNN
NNTTAGAAAGGTACTCTATATACTNNTATGNTNCGGCNNATAATANAACAGATGTT
TGTATNAATATNAAANAAGGTCNNTTTCGNCAAGAGAANNNTGNCTGGTNATAGAA
TTAGCATAANTTANNTANTATGATNNANTNNTNCTACNANTNTTAGCNNTTNGCAG
5 NAGTCATTNNGNATNTATNNGNNTANTAGTNANTTGGGNCTNNTNCAGANTATAT
TNTGNGAANATGAANNTACGNANTCCTNNGNANTATNATNNTGANTANGANAANCN
ANANNTNTTNTANNANTGNCTATANATTGCCNNGATANATTNTNNAATGAANCGA
TAGCCCGCNCTAAGGANNTNNGTNANNTAAANNTCTCAGATAANNTACNTNTTNT
TATTAANCNANNATCACANTATANCNGNGACANNNGCGANANTATATGTATGNNAN
10 TATNACNGNTCCNNNCCGNGAANNTANTCNTANNAGGCATTTCNGNNGAGCTNTTCT
NCTAGACNATTTNNANTGAAANNATGCNGNNAACGACNNNCTTNAANTTNTGT
CTACANTCCGCNNTNTTTNTACAGATNGCAGNTAAGNNNANTNANNGCTCTCANCT
NGCTNNNACT

15 Sequence ID - 1221 nt: 741
AAGCAGAANTNTCTCTAAAAACATTATCTCCTTAAAATCTTGAGGTGCATATNAGA
GCCACAGGCAATCTCTGACATATAAAATTGCAGTACAGGCCTTTCAAATTTGGCAT
TTCCTGGTACAATAACAACCAAGATATATAATAACTGTACAGTGCCTAGACAT
TCCAGTAAGAACCATTATTTTCTTTAATGTAGAATGATTAATACATATTCTACAAG
20 GGGCAGTAAGGTTAGTAATTCTATAGGGTATGTCCCGACATAATTTTCAAATTGTA
CAATAACACAAACAACCTTTGTTAAGGCCATGTTTTATTTGCTGATTAATGGACAAA
AGGCAATGTAATTTATTTTCAAGTATTTTCTTGAAAGTCTGTGCTCATAAAAATCA
TGAAAAGTTGGAAAGACTGTTAAATCACTGAAACTTCAAATATATCTTACACAATC
TTGTTTGTACAAAATACAAGTTAAATATAAACATAAAGCAATCATGGTAATTTTA
25 TGCAAATCTGTTTTATGTGATCATCAGTTATATATAAAAGTTTCTCAGTTCTGTTA
TTTGTGAAAAGATCAATACCAGATTGAATGACTACCTATTGGCAAAGGGCCCTAAA
AAGCTTACTTTAGCACTCATCTTTTACATGGTTAAATGCATTTCTTAATTTGAGAT
CACCTAAACACTGGAAAAGAAAAAAATGAAAGGGCAGTATGTCCATAAACCAACA
AATAATTTGGCTG

30 Sequence ID - 1224 nt: 485
CGAAATTTCTTGTGACACAGAGGAAGGGCAAAGGTCTGAGCCCAGAGTTGACGGA
GGGAGTATTTTCAAGGTTCACTTCAGGGGCTCCCAAAGCGACAAGATCGTTAGGGAG
AGAGGCCCAGGGTGGGGACTGGGAATTTAAGGAGAGCTGGGAACGGATCCCTTAGG
35 TTCAGGAAGCTTCTGTGCAAGCTGCGAGGATGGCTTGGGCCGAAGGGTTGCTCTGC
CCGCCGCGCTAGCTGTGAGCTGAGCAAAGCCCTGGGCTCACAGCACCCCAAAAGCC
TGTGGCTTCAGTCCTGCGTCTGCACCACACAATCAAAGGATCGTTTTGTTTTGTT

- 256 -

TTTAAAGAAAGGTGAGATTGGCTTGGTTCTTCATGAGCACATTTGATATAGCTCTT
TTTCTGTTTTTCCTTGCTCATTTTCGTTTTGGGGAAGAAATCTGTACTGTATTGGGA
TTGTAAAGAACATCTCTGCACTCAGACAGTTTACAGA

5 Sequence ID 1226

GGTTTTTATACTTGCCATGAAACTGTTCTTTGGGATATTATTTTGTTCAGGTTCCC
CACTTGGACAGCAGAGGGGGTGACTCTGCCCATCCCTGCCACTGGTAGCCAGGCGG
GCAATGTCTGCTAGCAGTCTGCTTCTGTCTGAACTCAGCCAGCAGAGGCAAACCTCC
CGGTTCCCCGAGAAACACTCTGAAGGCAGGGTGGGTGACTCCACCCACCACCGCCT
10 CTCCTAGCCATGCAGGCCATGTCTGCTAGAGCTTCCAGCGCAGTGGTCCTAATTCT
GTCTGAATCCGGCTGAGGGGTGCAGCCTCCTGTTACTGCCCAGGGAAACACCCAGA
TGGCAGGGTGGGTGACTCCAACCACCTCTGCCTGTGGTAGCCAGATGGGCCACACC
TGCTAGAGCTTCCAGCCCAGCAGTCCCGCTACTCTGTGGGTGGGTGCCATCCCCTG
TTCCTCTGGGAAGCACCCAGACAGCTGATTACGTGACCCACCCACTTCTGCAGAT
15 CCTAGCTGAGCAGGACTTGCTGGTTTGGACAATGCCCAAGCAGGGAAGAGCCCTCA
TTCTCTTATCACTGACAGAGGTGAGATGTCCGANTTTGTANGCTGGTGGAGGAGTG
AGGTGGAGGAGGTATGCCTCT

Sequence ID 1228

20 GTTATTCAGGTATCCATCAAAATTTTATAAGAGGGCCGGAAACATCGGCTCACACC
TGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGTTCACTTGAGGTCAGGAGTT
CGAGACCAGCCTGGCCAACATGGCAAAACCCCGTCACTATTAAAAATACAAAACAT
TAGCTGGGTGTAGTGGCAGGTGCCTGTAATCCCAGCTATTCGGGAGGCCTAGGAAG
GAAAATGGCTTGAACCTGGGGGTGGAGGTTGGAGTGAGGCAAGATCACACCACTGC
25 ACTCCAGCCTGGGCGACAGAGCGAGACTCCATCTCAAAAGAAGAAAAAAAAAAACAA
CAAAAAAACCTTTATCAGATTATCAGAGGTTATCACTACAGAGGGAGGTAAAATTG
GAGGGAAAAGGGTACAAATTTATTTTAC

Sequence ID - 1230

nt: 741

30 AAGCAGAANTNTCTCTAAAAACATTATCTCCTTAAAATCTTGAGGTGCATATNAGA
GCCACAGGCAATCTCTGACATATAAAATTGCAGTACAGGCCTTTCAAATTTGGCAT
TTCAGTGGTACAATAACAACAACCAAGATATATAATAACTGTACAGTGCCTAGACAT
TCCAGTAAGAACCATTATTTTCTTTAATGTAGAATGATTAATACATATTCTACAAG
GGGCAGTAAGGTTAGTAATTCTATAGGGTATGTCCCGACATAATTTTCAAATTGTA
35 CAATAACACAAACAACCTTTGTTAAGGCCATGTTTTTATTTGCTGATTAATGGACAAA
AGGCAATGTAATTTATTTTCAAGTATTTTCTTGAAAGTCTGTGCTCATAAAAATCA
TGAAAAGTTGGAAAGACTGTTAAATCACTGAAACTTCAAATATATCTTACACAATC

- 257 -

TTGTTTGTACAAAAATACAAGTTAAATATAAACATAAAGCAATCATGGTAATTTTA
TGCAAATCTGTTTTATGTGATCATCAGTTATATATAAAAGTTTCTCAGTTCTGTTA
TTTGTGAAAAGATCAATACCAGATTGAATGACTACCTATTGGCAAAGGGCCCTAAA
AAGCTTACTTTAGCACTCATCTTTTACATGGTTAAATGCATTTCTTAATTTGAGAT
5 CACCTAAACACTGGAAAAGAAAAAAATGAAAGGGCAGTATGTCCATAAACCAACA
AATAATTTGGCTG

Sequence ID - 1231

nt: 203

TTGAGGAAGGGTCTACTGTCTTTTTTAAATGGCACAATTTTAAAGAGGTTTGAGAGGT
10 ACAGTCCCTTAACCTGCCACGGGAGAGGGGGCCCCCAAACCTTCTTCCCCCCCACACT
TCTGGTTTTCTGTGTGGAGGGGGAGCAGGGATATCTAAGCTGTGGTGTGAAAGGGT
AGGAGAGATGCTGGAGGTGGGGGTGCTGTGTTCTA

Sequence ID 1239

TTTCCTCGGGAAGCGCGCCATTGTGTTGGTACCCGGAATTCGCGGCCGCGTCGAC
15 ATTTTTTTTTTTTTTTTTTTTTTTTAGAATGATTAACAATTTATTGAGTTTTATTTATC
TACAAAATATAGCAATACAGNGAACTTCACCAAATCCTAAATATTCAGTACCTGA
ACTGGCTACAACACCGNGTGCACACCCAGTTCCTGCAGAATCTCTTGCAGATATGG
GAGAGTCAGCCAGTGAAAAGATCCATTTCTTGGGAATCCTTGTCAACAAGACCAGT
20 TCAGAAATCCAGGATATATAGAAGCCTACTGTAATTTAAAAACAGTAACAAAAACC
CCAACAAAACCCAAATCAACAAAGACCAAGATAAAGGNGTGATAAACATTAATTGT
AATGGTTTTCTTTTACATGCAATACATGCATTTTAAAATCACTAAGAAACACGAAA
TTTTGTAGAGCAAAGTTTGNGTTTCACGTAAGTGCAAATGAATATATATTTTATTT
TTTATACTATTAAATTATATATATTTTTTTCCATACAAAAGCACACAGTGTTAATCT
25 ATAAAATGACATCCAAGTGGATGATGATTGTTTTTGCATGTCCCCCTGCTTAGATT
TTTTTAAAATATATAGTCAAAAATTAACATCCTTCTTTAAAATAACAGAAGGGAAA
AANGGGCAAAAAAAAAAATCTAGACTCGAGCAAGCTTATGCATGCATGCGGCCGCA
ATTCGANCTCGGNCGACTTGGCCAATTGCGCCTATAGNGAGTCGNATTACAATTCA
CTGGGCCGNCGNTTTACAACGTCGNGACTGGGAAAACCCTGGCGTTACCCNNCTNA
30 TCGNCTTGNAACAATNCCCNTTTNGCCAGNGGG

Sequence ID 1255

TCACTTCGTATNGAANCTGTTTGGACTTGCTCAACAAGACCTTATCTTAACAAAAA
GTAAC TTATAGAAAAGGGAGACATTCATTTAACTTCAAGCCCATATTATTCTTAAA
35 AGCTGACTCTTGAAATAGTATTTATTGAGTCATAGTGGAGTCATGGGACTTTTTTAA
GGGCCGGAAGGGACTATTTAGATCATCCAGTCCCACCCTGTCATTTTATGGAGGAG
GAAACTGAGGCCTAGATAAGATAACCAGTTAGTGGGTCCACTGACCTTTAGGACAG

- 258 -

TAGTCTATCCGTAAGAGACAACATGGAGAAAGAAATACAACGTTTTTATAGTGAAT
TATCATCTTACAAAGAATATTCTTCCCATATCGCACTTTTAAAAAGTGGGTACCTT
AGTCAAATAGGAGAAAAAACCACCTTGAGTAGTTTCATCCTCAGGTTTTAGGTGAGG
AAACTGATACTCAGATTAAATAACTTTAAGCACACAGAGCCTGAATGATAGTCTTA
5 TTTGAGCTCATCTGTGCTTTTAATGTGTACTACGTTAGGTGTTTTCACTTGCAATTT
CCTTTAGTCTTATTTGAGCTCATCTGTGCTTTTAATGTGTACTACGTTAGGTGTTT
TCACTTGCAATTTCTTTGTTTGACGTTGACAATAAATCGTGAAGCTGCCTTATCTAA
GNAGTCCTAAAGTAAATCATTGGAACACATGTANCCAGTTTGTTGTTTTTTATTTGC
CAGGTNTCAAATATAACTGAAAACCCATGCTAACTGACTNATTTTAAAAGNTGTNT
10 GGGGCATGAAANGATTGCTCTGCCTGGGCGGGNGGTTNANCCTGNGTCCCCCNTTT
NGGAGNCCACCANGANGCGATATTTNAGGGNNGATTCNAAACCCCTGGCACGNGN
NAACCCCNTTTTTTAAANANAAAANANCGGNNG

Sequence 1256

15 TTGTGTTGGTACCCGGGAATTCGCGGCCGCGTCGACGGAGTTTTACCTTATTACAC
TTTAATCTCTGGATTTACCCCATCTCATTTCTCTTTTAGGAAAAGTGTGTATGT
GGTGGCTTTGATGGTTCTCATGCCATCAGTTGTGTGGAAATGTATGATCCAAC TAG
AAATGAATGGAAGATGATGGGAAATATGACTTCACCAAGGAGCAATGCTGGGATTG
CAACTGTAGGGAACACCATTTATGCAGTGGGAGGATTCGATGGCAATGAATTTCTG
20 AATACGGTGGAAGTCTATAACCTTGAGTCAAATGAATGGAGCCCCTATACAAAGAT
TTTCCAGTTTTTAACAAATTTAAGACCCTCTCAAAC TAACAGGCTTAGTGATGTAAT
TATGGTTAGCAGAGGTACACTTGTGAATAAAGAGGGTGGGTGGGTATAGATGTTGC
TAACAGCAACACAAAGCTTTTGCATATTGCATACTATTAAACATGCTGTACATACT
TTTTGGGTTTATTTGGAAAGGAATGCAAAGATGAAGGTCTGTTTTGTGTACTTTTA
25 AGACTTTGGTTATTTTACTTTTTTGAAAAGAATAAACCAAGAATTGATTGGGCACA
TCATTTCAAGAAGTCCCCTCTCCTCCACATTTGTTTTTGCCAATTTGCACATTAAAT
GACTCTTCCCTCAAATGTGTACTATGGGGTAAAAGGGGTAGGGNTTAAANATGTAA
ACAGTTGGGTTTTTTAAGGGNCCTTTTTTCATAACTGGAACACTCTNTACAAGGNTN
CTTNTTAAATAAATAACTTGACTTTTTTTGTTTTNTAAANGNANCTTCNTGCTTCCA
30 TAAAAAATAAATTTAANTNGNCANCTNTGCTGCTGCGNCCANTTNGCTNGNCCNT
GGCATTCCCTAGGGANGNTNAATANTGGCNNNTTAACNNGGCNGNAACNNNNNCCA
NT

Sequence ID 1331

35 GGGCGATGCATGCTTTATTAAGGCTCTTGTTTCACCTGGCAGTGTACTGTATCAAC
GTATAATACAGAAAAAAATCTCTTTAAGGTCCTCCTTCACAAAGACATAGAGTGA
AACTCCCTTTACATGTCAGTATTTGTTCAACACTTTAGGCAACTTGACTGTCAGTG

- 259 -

TTAAAATGGAAAACAGGAAAATGGAAAATCTGACCAATTCTGCCACCTTGAGACT
TTCATATAGACCTTGCACAACAATTGTATAGATCACACACCGGCTGTATTTAATAT
GTAACATTTTCACACATATTAAAGATACAGAAGTATTAAAAAACCCCAATGTTAA
TGTATTTGCTTAAAAGGCACAAGTTTCACATATCTGTCTAGCTATCTGTTGGTAAT
5 ACAGAAAGTATACTACTTTTTTAAAAAAGTGGGCAGAATTCTTGTGTATGTATATT
TGTGTGTACAGTATGTGTATGTGTGTATATATATATATTATATATATAGATAATAT
ATAAATATTTTTTTTAAAGGAGAACTAGAATGTTTAGCTAGAAAATTCCACAGCCT
GTGAAGAAATATTTCAAATGGCCATAAAGGAGGTAAAAATGAAAACCATAACCTA
ACTTTTATAGAGGCTTTTATCTTTAATTTAACGATGTGCGGAGGACTTTCTTGCTTG
10 AATCTGTTCCGGGCTGTCTGCTCTGTCCATCAAATGGGCAGGTCTGGGAATGAGGC
ACCTTCGGCCGTTT CAGAAGTGGCCTGAACAGAATGCTGGAACCCAGGCTGGACTCG
GAC

Sequence ID 1332

15 CAAACCTGCATGTTCTGCACATGTATCCAGGAACCTTAAAAAAAAAAAAAGATAGTT
TGTGTGTCTTAATTGAATAATAGTAGATTTATAGATTAAAGATCTATGGGTTTTTA
ATATGGATTAGAAATCTGTGGGTTTTTGATATGGATTAGAAATCTGTGGGTTTTTA
ATATGGATTGGAAATCTGTGGGTTTTTAATATGGATTAAAAACATCTGTGGGTTT
TTAATATGGATTAAACATCTGTGGGTTTTTAATATGGATTAAACATCTGGGTTTTT
20 AATATGGATTAAACATCTGTGGGTTTTTAATATGGGTAAAAATCAAAGAAAATG
AACTATTTGCTCCAGTGCAGGAAAATACAGGCAATACTGGATACAATTAGATGGTC
AGGAGCGATAACCCGGTTGCCATTGTTTGAAGAAGAGAATAAGGTGCTAGCATTCC
TATCCGTAGATAATTTGACAGCTAGGAAATAGGGGGAGTCTTCTATGTAGTTAGTG
AAGGCTAAATGAACTATTATATGCAGTTATCGTAGAAGAGTACTCAAAAAAATCTG
25 TAAAAAATAAAGAAAGGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTG
GGAGGCCGAGGCGGGTGGATCATGAGGTCAGGAGATCGAGACCATCCTGGCTACCA
NGGTGAAACCCCCGTCT

Sequence ID 1335

30 CAAGACTCCATCTCAAAAAAAAAAAAAAATCTACAGTGCTGAGTATATAAAATTAT
TAACACATTT CACAACAATATGTGTTTGTGGAGTTAAATATTTTTTGTCTTTAAA
CAGGTAATTTTAGTGCATACTTAATTTGATGATTAAATATGGTAGAATTAAGCATT
TTAAATGTTAATGTTTGTACATTGTTCAAGAAATAAGTAGAAATATATTCCTTTG
TTTTTTATTTAAATTTTTGTTCTCTGTAAACTAAAAGAACACGAAGTAATTGGTC
35 ACAATTACTGGTGTTTAACTGCCAAATATGGGTAAATAAGGGAAAATTTTGTTTAA
TATTTAGTCCTTCTGAGATGGCTTGAATATTTGAATTTTGTGTACGTCTATACTG
GGTAGTCACAAGTCTTATAAACACTTTAGAGGAAAGATGGATTT CAGTCTGTATTT

- 260 -

TTAAACATCATTTATTTTAAATCTGGTGCTGAAAAATAAGAAAAAATTAAACTGC
ATTCTGCTGTTCTTCTTTAGAACATTCCTGCGTAAATACTGCTGTAATACTGTCA
TGCAAAGTGTATCCTTTCTTGTCTATCCTTTTTGGGGCAGTGTTTTTTTGT
TTCTAGAAATGTTTGTCTTCCCCCACCTGTTGATCCAGGTTAAGGAATACTTTT
5 TTACACTTTATTCAA

Sequence ID 1336

CTTTTCCTCCCGCTGTCCCCACGGAGGGGACTGCTCTCCCCCGCTGCATCCTTTC
TGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAACCTC
10 GACATTCGGGAAGTGTTTTTGAGAAGTCTCGGTTCGGTAAGGGAAGTCTTCCAAGTC
CGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTTCGGCCACCCCCCAGATGAGG
CAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTTCTCTCAGCTT
CCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGTCCATGCTGCCAT
CAAGGCATTTATTGCAGTGTACTATTTGCTTCCAAAGGATCAGGCCCTGAGAACAA
15 TGACCTTATTTCTTACAACAGTGTCTGGGTTGCGTGCCAGCAGATGCCTCAGATAC
CAAGAGATAACAAAGCTGCAGCTCTTTTGATGCTGACCAAGAATGTGGATTTTGTG
AAGGATGCACATGAAGAAATGGAGCAGGCTGTGGAAGAATGTGACCCTTACTCTGG
CCTCTTGAATGATACTGAGGAGAACAACTCTGACAACCACAATCATGAGGATGATG
TGTTGGGGTTTCCCAGCAATCAGGACTTGTATTGGTCAGAGGACGATCAAGAGCTC
20 ATAATCCCATGCCTTGCGCTGGTGAGAGCATCCAAAGCCTGCCTGAAGAAAA

Sequence ID 1337

CAAGAACTCTGGGACATTTGCAAAGGGTATGGCATATGTGTAATGGGAATACCAGA
GGAGAGGAAAGACAGGAAGTCAAAAAAGAATTTTTCCAAATTAATGATAGGTTCC
25 AAACCACAGATGCAGGAAGCTTAAACACCAACAGGATAAATAAAACAAAATCTACG
CTTAAGCATATCATACTTAACCTGCAGAAAATTACAGACAAAGAAAAAACACCAGA
GGGGAAGCTGGCAGAAACATACCACCTATAGCGGAAGAAGAATAAGAATTACATCA
GACTTCCCTTCAGAAATCTTGCAAACAAAAGATGTAGCACAATATTTAAAGTATT
AAAGGAGGCCGGGCCCCGGTGGCTCGGGCCTGTAATCCTAACACTTTGGGAGGCTGA
30 GGCAGGAGGACCATGAGGTCAGGAGATCGAGACCATCCTGGTGATGGTGATACCCC
ATCTCTACTAAAAATACAAAAAATTAACCGGGCATGGTGACACGCACCTGTAATCC
CAGCTACTTGGGAGGCTGAAGCAGGAGAATCGTTTGAGCCCAGGAGGTGGAGGTTG
CAGTGAGCCGAGATCACATCACTGCACGCCTGGGCAACAGAGCGAGACTCCATCTC
AAAAAA

35

Sequence ID 1338

CGACCCGTTTTAGTCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCTGCCTC

- 261 -

GGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCGTAAATCAG
GTTTTTTTAAATGTTTGCCAAACCTTATCACTGACTTTTATAACAAAATTATTTACT
ATAATCATTAGGGAATATTTAAGTTCTGCTAATACTTAAAATTGCAGAGTGCTAAA
ACCAGCAGTGAGTTTAGAATCAAGCTAAGCTTTATTGTTGCTACTATTTGAGGCAT
5 ATTAGTTGACTGGTGTTCATATGCAAGGCAGTCTACTGGGTGCAACAAGGGTTAGA
AGGATATTTTTTAAA⁵AACTGACCCTATTCTCAGGATGAAAATAATACACTAGTAAT
AGTCTGCTCTGTTGGTTAACTCCTCGTAAGGAGGTACAATTAAAATGCTGTAGTGT
TGCAAGGGAAGGAGAGGAAGAATCATATTCCTTCACTAGCAGGATCAAGAAAGCTT
TTATAGAAATATACAAAATCTTCACTTCTTGAAGGATTGGTAAAATTTAATAGCCA
10 ACATTGGGCACTTATTCATTCTCTGAGTAAATATTTATTGCATGCTTATCTTGTAT
CAAGCATTGTGATGAAAGCACAAGAATGAAAGAGGAGGGAGAATGTTTAGAGAATA
AGGGCTGAAACACAGATTTTGTAGGGAGCGTAGGGGAGACTGANAAGACAGGTTCA
GGTTAGTAAGGGCGCTCATATTTTGACCCTGAATGTTAACTATGTGCACATCATGC
TAGCTATTCTAAATCAGGCATTTTCAAATGGAAGCAGGCACTGACATTTT

15

Sequence ID 1344

CGTGAAGGGTCTTTATGTATTAGTATTAGAGTGATCTTTTGATTATTTTCCTCACT
ATAAGGAAATTATTTTCCTCAGGATGAGCTGCCATAACATTCCACTGTCTGATGGCA
ATTTTAAAGCCTGAAATTGAAGCCCATGGCTAGGCTATGAGAACCCTAGTTCGTAT
20 AGTAAAGTTGATATCTTCTGGATGTATACTAATTTTAGGCTTTATTTTAA⁵AACTGC
TGGAAACTGAACTTAGACAAAAGTATTTTCAGGACATCATTTACAATGTTTAGCC
CTAAAGAGTCAAGCTGTGGGATTCTGAGTCTTTCATATGTTACAGCAGAACTTAA
AAGCAAGAGGAAATTGGCTGGGCACAGTGGCTCTGTAATCCCAGCACTTTGGGAGG
CTGAGGTGGGTGGATCATGAGGTCAAGAGATTGAGACCATCCTAGCCAACATGGTG
25 AAACCCCATCTCTACTAAAAATACAAAATTAGCTGGGCGTGGTGGCACACGCCTG
TAATCCCAGCTAGTCAGGAGGCTGAGGCAGGAGAATATCTTGA⁵ACTTGGGAGGCAG
AGGTTGCAGTGAGCCAAGATTACATCACTGCACTCCAGCCTGGTGACAGAGCGAGA
CTCCGACT

30 Sequence ID 1348

CTGAAACTGCACTGAACCCACAGGTAGGTTACATCACAGGACAGAAATCTGAGGAG
CTGGAGAAAGCAAAGAATAAAGGATGGGCTGACACCAGAAGGAATTAAAGGAATT
TTTATACTGAACTTCAATTACTTGTTCA⁵TTTGAAGTTTGT⁵TTTTTTAATGAACGTT
TTTGCTGTTACTTAAATATAGTGTTTGAAGTGTTTCAAATGTATTCAAGTTGGG
35 ATTTTCCATATTTTACTACAGTTCTGTCTTAGTATGTTTACCATAAAACACTTATC
ATTAAAGCTCACAAAGTGCTTTTTTTGTAATATGAGGATAAAATGAAGCCATATAAG
AATTTTTTTTATATCTGTACATTTA⁵ACCCACATTTGAGCTTTAGCCAAAATATATAG

- 262 -

CTTTTTTTTTTCTGACCTGGCCAACGTATTATCCAGCAAACATCAACTGAAGCAAT
ATGGAAACACTTCCAAATGTTTGCCAATAATGCTATTAAGTGACTGATGTCAACAT
TAGTTACATGGCAAACCTAAAGAGGCATTATACATTTTTTAAAACACACTAACATATA
ACTGTAGATAATGTAAGGTTTATTTATATGCATATTTTCATAGTATATTTAAATGTT
5 TAAATATAAAAAAGGGTTTTTAAACACTTTTAAATTTTTTATCTTTGATTTTTTTTTAT
TGATATCTCTTTCCAGGCTACTAATAAAATTGCCAGAACTAACTATCAGGTAAAG
GTTAAGGCATCAATTGACAAGTAAGTTTTCTAATTTTCGTTTTGAATTACAATTCCA
AATGTAAGACTTTTAAAAATGAATGGCCTTTATTTTATAGAATAATTTTGACCTTT
TAAATTTACTTATCTAACATTATATAATGAATGTACTTCAAATATTTGACTTTGAA
10 GTCAACATTAACAAATTCATGGATCCTAATTAAAATTTACTATAAAACTGGAATCA
TTTATTACTTCCTT'

Sequence ID 1351

TTTTTTTTTTTTTAAAGAGATGGGTCTCACTATGTTGCCCATAAATGTTTATGAG
15 ATTAAGTTCATCTTTTTTATCTGAGTAGTATTTTATTGTATGAATATAACCACCATT
TATTTATCTGTTGGTTATTTCCAGTTTTGGGCTATAATCCAAAATGCTTTTTTCAA
ACAATAGGCTATATATCATTAATGTCCGTTTATCAGCAGTATAAAATATCTTACCA
TAAATATTAATAAAAGAAGCATTCATATATAAAATATAGATATTTCAAACCCTACA
GAGGGCCTTTTAATGATTAAATATTTTGTCTTACAAAAGGTCCAGGTAATTACA
20 CCCATGAGGTTAACCTGCCTTAGTGCAGGACTTAAAATAAGGCTTCTCCTGCCATC
TCTCTCCATTTGTAGAATGTGAAATTCTTTAAAATGCATCCTATATTAGGAATACT
ATAGCTGTGCACTGGTGTTTGTTCTCTTCTTTAAACTCGGGACCGTATATATCTGC
TCAAATTGCCCAAGTATACATATGCTGCACTCCATCAAGTGTCAGGCCACATTCTA
TCAGCACAGCGTGACTGCCTATCAGTGACAATATAAGTGAGCTCTATTTGGATCCC
25 TCTTACCCTACCTTTTATATTTATGACAGCATTATCATAAACTCCAATATTCTTC
AATAACTTACATGTTTGTTGTAGGATAAAATTATTACCCTCAATGAACTACAT

Sequence ID 1352

ACCAGCTTCTTCACAGGTTCCACGAGTCATGTCAACACAGCGTGTTGCTAACACAT
30 CAACACAGACAATGGGTCCACGTCCTGCAGCTGCAGCCGCTGCAGCTACTCCTGCT
GTCCGCACCGTTCCACAGTATAAATATGCTGCAGGAGTTCGCAATCCTCAGCAACA
TCTTAATGCACAGCCACAAGTTACAATGCAACAGCCTGCTGTTTCATGTACAAGGTC
AGGAACCTTTGACTGCTTCCATGTTGGCATCTGCCCCCTCCTCAAGAGCAAAAGCAA
ATGTTGGGTGAACGGCTGTTTCCTCTTATTCAAGCCATGCACCCTACTCTTGCTGG
35 TAAAATCACTGGCATGTTGTTGGAGATTGATAATTCAGAACTTCTTCATATGCTCG
AGTCTCCAGAGTCACTCCGTTCTAAGGTTGATGAAGCTGTAGCTGTACTACAAGCC
CACCAAGCTAAAGAGGCTGCCCAGAAAGCAGTTAACAGTGCCACCGGTGTTCCAAC

- 263 -

TGTTTAAAATTGATCAGGGACCATGAAAAGAACTTGTGCTTCACCGAAGAAAAAT
ATCTAAACATCGAAAACTTAAATATTATGGAAAAAAACATTGCAAAATATAAAA
TAAATAAAAAAAGGAAAGGAACTTTGAACCTTATGTACCGAGCAAATGCCAGGTC
TAGCAAACATAATGCTAGTCCTAGATTACTTATTGATTTAAAA

5

Sequence ID 1353

ACATTCTGGAAAAGGCAAAAGGGAGGAAGAACTGATTAGTGGTTAGCCCAGGGTTA
GAGTTGGGGAGAGGATATAATGAGGGAACTTTTGTGGATTCTGTACCATGATTATG
ATTACACAAACCTATGCATACATTGAAACACATAGAACTATACGTTGAAAAAAGTG
10 AATCTGCCTGTATGTAAATTTAAAAGAAAAATATTTTTTTTAAAAAACAGATGCTT
CTTAACACATTATCATCTATGTCAGTTTAACAGTTAGTAGACTTAGGCCAGGTGTC
ATGGCTCACTCCTGTAATCCCAGTGCTTTGGGAGTCTGAGGTGGGACGATCTCTTG
AGACTAGGAGGGAGTTTGAGACAAACCTAGGCAATGTAATGAGACTCTTTCTCTAC
AAAAAATTTTAAAGTTATCTGGACATGGTGGTGCCTGCCTGTAGTCCCAGCTACTT
15 GGGAGGCTGAGGTGGGAGGATTCCTTGAGCCCAGAAGTTCAAGGCTACAGTGTGCT
ATGATAGAGCCACTGCACTCCAGCCTGGGCAACCAGGTGAGACCTTGTCTCTAAAA
TGAATAAATAAAT

Sequence ID 1355

20 TGGTCTTTCACCCAGCCAGGGAGAAGGTTCTTCGCTCAGTATGAAGAAAAGCAACC
CAAACTCTCAATCTGATTTGTTTTTGTATTATGTCGATGCCCTGTAGTTTGAAAGT
GAAGTAAAGATTTAGAATTCACCTAAGTCCAAAGGAAAACACGTGGTTTTTTAAAGC
CATTAGGTAAAAAAAGTTCTCAATAAAGGCATTACAATTTTTTTAGGTTTAGAAAGA
TGGACTTTTCTGATAAATCTTGGCAGACATCTAAAAAAAACCATATTTTTTCACA
25 AGAAAATGCAAGTTACTTTTTTTGGAAATAATACTCACTGATTATGGATAAAATGG
AATATTTTCAGATACTATATTGGCTGTTTCAAATAGTACTATTCTTTAAACTTGT
AATTTTTGCTAAGTTATTTGTCTTTGTTGTATCTATAAATATGTAAAAAATATTTA
AATAGATGTACCTGTTTTGCTTTCACACTTAATAAAAAATTTTTTTTTTGT

Sequence ID 1359

30 CGGGATCCCTAGTATAACACATTCAGTGTTCCCCTTTCAGTCTTACTACTTTGACC
GCGATGATGTGGCTTTGAAGAACTTTGCCAAATACTTTCTTCACCAATCTCATGAG
GAGAGGGAACATGCTGAGAACTGATGAAGCTGCAGAACCAACGAGGTGGCCGAAT
CTTCCTTCAGGATATCAAGAAACCAGACTGTGATGACTGGGAGAGCGGGCTGAATG
35 CAATGGAGTGTGCATTACATTTGGAAAAAATGTGAATCAGTCACTACTGGAAGTGC
ACAACTGGCCACTGACAAAAATGACCCCCATGTGAGTATTGGAACCCCAGGAAAT
AAATGGAGGAAATCATTTGCCTTAGGGATTGGGAAAGCTGCCCACTAACTGTCTTC

- 264 -

CCCATTTGTTTTGTCAGTTGTGTGACTTCATTGAGACACATTACCTGAATGAGCAGGT
GAAAGCCATCAAAGAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGGAGCGC
CCGAATCTGGCTTGGCGGAATATCTCTTTGACAAGCACACCCTGGGAGACAGTGAT
AATGAAAGCTAAGCCTCGGGCTAATTTCCCCATAGCCGTGGGGTGACTTCCCTGGT
5 CACCAAGGCAGTGCATGCATGTTGGGGTTTCCTTTACCTTTTCTATAAGTTGTACC
AAAACATCCACTTAAGTTCTTTGATTTGTACCATTCTTTCAAATAAGAAATTTGG
TACC

Sequence ID 1360

10

TGCGCAGACCAGACTTCGCTCGTACTCGTGCGCCTCGCTTCGCTTTTCTCCGCAA
CCATGTCTGACAAACCCGATATGGCTGAGATCGAGAAATTCGATAAGTCGAAACTG
AAGAAGACAGAGACGCAAGAGAAAAATCCACTGCCTTCCAAAGAAACGATTGAACA
GGAGAAGCAAGCAGGCGAATCGTAATGAGGCGTGCGCCGCCAATATGCACTGTACA
15 TTCCACAAGCATTCGCTTCTTATTTTACTTCTTTTAGCTGTTTAACTTTGTAAGAT
GCAAAGAGGTTGGATCAAGTTTAAATGACTGTGCTGCCCCCTTTCACATCAAAGAAC
TACTGACAACGAAGGCCGCGCCTGCCTTTCCCATCTGTCTATCTATCTGGCTGGCA
GGGAAGGAAAGAACTTGCATGTTGGTGAAGGAAGAAGTGGGGTGGAAGAAGTGGGG
TGGGACGACAGTGAAAT

20

Sequence ID 1361

TATAAATACACTCCGGGATGATTTACCCCCGGAGGTCAGCTAGTAAAATACATGAG
TAGAATTCCTTAAAGTATGTGATAATTGCTCATCACTATCCAAGTGTGACATAAAT
CATAAAAAGAATTGACAAAATCAGGGTCGCAAAGAGAATTGAAAAAAATCTGTCAC
25 AACCAAAATTTAAATTGACCTCTGTCCTAGAGTATGAGAGCCACACTGAACAGAAA
AACCAGATAAATCTTTTATAAAATATTCATTTGCAGCCCCATTAACGTTGCTTGTC
ACCCACCTCCCCATGTCCTTGGACAAACTGAATGTATAGTAACATCATCCCAGGC
CAGGCGCGGTGGCTCATGCCTGTAATCCAGCACTTTGTGAGGCTAAGGCAGGCAG
ATCAGGAGGTCAGGAGTTCAGGACCAGCCTGGCCAAAAGGTGAAACTCCGTCTCT
30 ACTAACAATACAAAATTAGCTGGGTGCGGTAGTAGGCGCCTGTAATCCCAGCTAC
TCGGGAGGCTGAGGCAGGAGAATTGCTCAAACCCGGAAGGTGGAGGTTGCAGTGAG
CTGAGATCGTGCCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCTGTCTCGGG
GAGGGGGGTGGCGGAGATAAAGAAATAACATCATCTTATACTGTCAAGCTCAAGGT
GTCTGCAGCCTTATCTTCAGGGGAAGTTGTGTCTTTCTCAGGGAAGATACAGATTT
35 CAATTTAGAGCAAGACAGAGAGAAGTTACATTCAGAGAGGAAAATGCAGTAGTCTA
ACTG

- 265 -

Sequence ID 1364

GCGGCCGCGCTCTTTTCAATTTTAAAGAAAGTTTGTTCATTTTCAGTAATTT
CTGCTTTGATCTTCCTTATGTCCTCCTATTGAGTTGATCAGCTTTCTTTATTCTTG
CCTTTTCTCCTCTGTGTGCCCTTTCTATTAACGTATTTACCCTTAGGCTGGGCACA
5 ATGGCTGATGCCTGTAATCCCTGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTA
AGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTGGTCTCTACTAAA
AACACAAAATTAGCCAGGCATGGTGGTGTGCACCTGTAATCCCAGCTACTCAGGA
GGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGCGGAGATTGTGCCAAAGCACTC
CAGCCTGGGCAACAAAATGAGACTTTGTGTC

10

Sequence ID 1365

CACCAGGCTGTCTTCAGATACTTCATACAGAAATGAGCCTCCCTGTGGGGTCCTCT
TCCCTCCTTCAGCCTGTCCATCAACACAGCATTGCGGGATCCTTACCATGGCATCC
AGCCCTGGAGATGCTTCAGGAAAGTTGCAGGTCCATGCTGCAGGACAGGCTCAGAT
15 CAGCAGAGACGCATCTCACATCGGGCTGTGAAATTCAAGTTGAGCTGCAATTGGCA
ATGAGAA

Sequence ID 1366

GTTATTCAGTACGACCGTGCCCCGGTTATGAGGTTGTACCAGAAAGCAAGTATTCA
20 CTATGCACACTATTCACCGCTCACCCTAGCATTGAAGCCAGCCTGTAGCCTGAAAG
CCTTTGCTTTGAGGGCAGGTCTTTCCCCAAAATGCAGACACGAAGGTGCAAAGTGA
AGCTGCCAGTCTTGCAAAAGATGTAACCTTGTACGAAGGCCACGAGTGGCAGGGAG
AGCTGTCCCACATTTGCGGAAGTGGCTATGTGAGGACGGGGGAGGCGGGTCCCTTA
GAGATGAGACAATCATAAGGGGAGATATCAGAGAAAATCGTAAGGGGAGCAGATGG
25 TTGTCAAGAGAATAGGCTGACCATCGAAGGACTGGCAGAAGCTTTTCAGAAAACCAC
TGGACGGCTGGGCACAGTGGCTTAGGCCTGTAATCCCAGCACTTTGGGAGGCTGAC
GCAGGTGAATCACTTGAGGTCAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAAA
CCCCATCTCTACAGAAAATATAAAAATTAGCCAGGCGTGGTGGCACAAGCCTAGAA
TCCCAGCTACTTGGGAGGCTGAGGCAGGCGAATGGCTTGAACCCAGGAGTCAGAGG
30 CTGCAGTGAGTCGAGATTGTTCCACTGCACTCCAGCCTGGGTGACAGTGCAAGACT
CCTTCCAAAAA

Sequence ID 1367

TTCGTGAGTGATGGCGTCCCGGGTTGCTTGCCGGTGCTGGCCGCCGCCGGGAGAGC
35 CCGGGGCAGAGCAGAGGTGCTCATCAGCACTGTAGGCCCGGAAGATTGTGTGGTCC
CGTTCCTGACCCGGCCTAAGGTCCCTGTCTTGACAGCTGGATAGCGGCAACTACCTC
TTCTCCACTAGTGCAATCTGCCGATATTTTTTTTTTTGTTATCTGGCTGGGAGCAAGA

- 266 -

TGACCTCACTAACCAGTGGCTGGAATGGGAAGCGACAGAGCTGCAGCCAGCTTTGT
CTGCTGCCCTGTACTATTTAGTGGTCCAAGGCAAGAAGGGGGAAGATGTTCTTGGT
TCAGTGCGGAGAGCCCTGACTCACATTGACCACAGCTTGAGTCGTCAGAACTGTCC
TTTCCTGGCTGGGGAGACAGAATCTCTAGCCGACATTGTTTTGTGGGGAGCCCTAT
5 ACCCATTACTGCAAGATCCCGCCTACCTCCCTGAGGAGCTGAGTGCCCTGCACAGC
TGGTTCCAGACACTGAGTACCCAGGAACCATGTCAGCGAGCTGCAGAGACTGTACT
GAAACAGCAAGGTGTCCTGGCTCTCCGGCCTTACCTCCAAAAGCAGCCCCAGCCCA
GCCCCGCTGAGGGAAGGGCTGTCACCAATGAGCCTGAGGAGGAGGAGCTGGCTACC
CTATCTGAGGAGGAGATTGCTATGGCTGTTACTGCTTGGGAGAANGGCCTAGAAAG
10 TTTTGCCCCCGCTGCGGCCCCAGCANAATCCAGTGTTGCCTGTGGCTGGAGAAAGG
AATGTGCTCATCACCCAGTGCCCTCCNTTACGTCAACAATGTCCCCACCTTGGGAA
CATCATTGGTTGTGTGCTCAGTGCCCGATGTCTT

Sequence ID 1368

15 CAGTGAGCCAAGATCACACCACTGCACTCCAGCCTGGACAACAGAACGAGACTCCA
TATCAAAAAAATTAAATTAAAATATAATAAATTTCTTGCCGGGCGCAGTGGCTCAC
ACCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAAGTCAGGAGA
TTGAGACCATCCTGGCTAATACAGTGAAACCCCGTCTCTACTATAAATACAAAAAA
TTAGCTGGGCATGGTGGCGGGCGTCTGTAGTCCCAGCTACTCAGGAGTCTGAGGCA
20 GGAGAATGGTGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGTGCCACT
GCAATCCAGCCTGGGCAGCAGAACGAGACTCCATCTCAAATAAATAAATAAATAAA
ATGAATTTTCAGCTAGAAGAGCCTTATTCCATTTTCCTTTTTTATTAAACATCTGGCA
TAAGTTGGTAAGTATGTGAAGTTTATCATATATTCTTATGCGAATTATTATTTTCG
CCTTTTTTTTTTATAATTCTGTCTGGGATTTGAATAGTAGAGTTTGAATTCAGGAAG
25 GACACCTGTGATAGGACAATAAAAT

Sequence ID 1369

CTGATTGCAAAAACATTACAACCTCAGTACTGCGGCTTTCATTCAAATAGGTTATAT
GTATAAACTGAGGTTCAACAATATTGTATTTGAGATGGGAAAGTTAAAGAAATGCA
30 ATAATGTAAATAATACTTAAGAAAATAAGATCTCAGGAAACTGTGTATACTCTGTA
CTTTTATGCAACTTTATCAGATCATTTTCAGTATATGCATCAAGGATATAGTGTATA
TGACATGAACTTTGAGTGCAAAAACCTGTACTATGTACCTTTTGTATTATTTTGCTGT
CAACATCTAAATAAAGGTTTTTTTG

35 Sequence ID 1370

CGAAAGGACTACAGAGCCCCGAATTAATACCAATAGAAGGGCAATGCTTTTATAGATT
AAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTTAAAAGTTGTAGGTGATTAA

- 267 -

AATAATTTGAAGGCGATCTTTTAAAAAGAGATTAAACCGAAGGTGATTAAAAGACC
TTGAAATCCATGACGCAGGGAGAATTGCGTCATTTAAAGCCTAGTTAACGCATTTA
CTAAACGCAGACGAAAATGGAAAGATTAATTGGGAGTGGTAGGATGAAACAATTTG
GAGAAGATAGAAGTTTGAAGTGGAAAACCTGGAAGACAGAAGTACGGGAAGGCGAAG
5 AAAAGAATAGATAAGATAGGGAAATTAGAAGATAAAAACATACTTTTAGAAGAAAA
AAGATAAATTTAAACCTGAAAAGTAGGAAG

Sequence ID 1371

GTCCAGNAGAAAGTTCAGTGACTTGTCAGAGCTGCAGGTCTTAAGAGGCTGAAAT
10 CTCGCCTCTGCCTCGAGGCTGCGGTTCCACTGACCCATACTACTTGCCTTCAGGAA
AGAGAAATGGTGTAGGAAGGCTGTGGATGAAGACGCTTACATTCATGAAGGATTTG
GATAGGCGAACATGAGCTTTTCCACCAAATTTTCAGAATTTTAAGAAATGCCTTAAA
TTATTTCTTAAAAATCAATTTGGGGCAGACGAGAAGTTCTGATAATAGTTTTTAGG
GAACATGATAAAATTCTGACCTTAGAAGTGGTATACCAGTTTGAGAAGAAGAACAA
15 GCTATAAACGGTGTAGATAACATTCACGGCTATTTAAGAAAGAGTTACTAAGGGAA
ACCAGAATGACTTAAGAGTGTTACTCTTCTTTTTCTGAGAGAACAATAGCATCATC
TCAGAAAGCCTTTTCATGCCATTAATAGGTAAGAATCTGGGCTTCTTGGACCATGGG
TTAGACTTTCTTACAAAACCATAATATGCATTTCTAGCAAAATTTATGCTATTAC
ATTTCTTATCTCAACAAAGACTGGTAAATTCAGTACTTATTCCTCAATTTTCCTA
20 CCCTTAAATGGGGATATTCTGCCTCTCCAAGGAATGCTGGGAACAAGCAAGTCCT
CATGTTAGGGGTCTTTGAGTTTTTCATGGAAGTTTAGGTTATTTATATGATGACATA
GTTGTCAACTTACTTTTCAGGATGGACTTTTCTTTTGTGAGTTTGTGACCTAAATAC
AATAGTTGTTATGCATGTCCAGTTTATGGAAGTACCACTGCAATANCAG

25 Sequence ID 1372

CAGTGCAGCCAAGTATCACACCACTGCACTCCAGTCCTGGACAACAGAAACGANTA
CTCCATATCAAAAAAATTAAATTAAANGATAATAAATTTCTTGCCGGGCGCAGTGG
CTCACACCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAAGTCA
GGAGATTGAGACCATCCTGGCTAATACAGTGAAATCCCCGTCTCTACTATAAATAC
30 AAAAAATTAGCTGGGCATGGTGGCGGGCGTCTGTAGTCCCAGCTACTCAGGAGTCT
GAGGCAGGAGAATGGTGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGT
GCCACTGCAATCCAGCCTGGGCAGCAGAACGAGACTCCATCTCAAATAAATAAATA
AATAAAATGAATTTTCAGCTAGAAGAGCCTTATTCATTTTTCTTTTTATTAAACAT
CTGGCATAAGTTGGTAAGTATGTGAAGTTTATCATATATTCTTATGCGAATTATTA
35 TTTTCGCCTTTTTTTTTTATAATTCTGTCTGGGATTTGAATAGTAGAGTTTGAATTC
AGGAAGGACACCTGTGATAGGACAATAAAATCTA

- 268 -

Sequence ID 1374

GAAAGCACATATGATATACATGTGTGTCATATGTATTATTTTGTGTTGCCATCTGAG
TCTTCAAAATTTGTTACAGAATACCTGCATATTAATATTTCAAGGTATGGATTAAT

5 Sequence ID 1378

CTGAGTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Sequence ID 1380

CCAAACCCAACTGGTCCAGTAGGATACTCACCTTACAGGGGGCGTCTCAAGAGTCT
10 CACAGTTCCCTTGGGTCTTAAGAGACTCACTGTTGGACCAGGCGTGGTGACTCACG
CCTGTAAAACCAGCACTTTGGGAGGCCGAGGCGGGCGGATCAGTTGAGGTCAAGAG
TTCAAGACCAGCCTGACCAAGGTGCTGAAACCCCGTCTCTACTAAAAATACAAAAA
TTAGCCAGGCATGGTGGTGTGCGCCTGTAATCCCAGCTACTCCAGAGGCTGAGGCA
GGAGAATCTCTTGAACCCAGGAGGTGGAGGTTGCAGTGAGTCGAGATCATGCCACT
15 GCACTCCAGCCTGGGTGACAGAGCGAGACTCCGTCTTAGAAAAAAAAAAAAAAAAAA
AAAAGAACCTCACAGTTCAGCAGGGTTCTAGCATGAGACAATGAGGACAAGGGTAG
GTGAGCAGGTGGAAAGAGTGAGAACAGGTCAATTGTGATGGAGAAAATAATAAGA
CAGAAAAGGCAGAAGACTGCCTGGCAGAAGACCTGTCCCAGCAGATACAAAATAC
AGACAACAGGAGCCAGCATAGACCCTTGACCTGTGTAAGTCTTTCTCAGGCCTTCT
20 TTTAAGTAGAAACATGCCTTTGAAAAAAGTTTTAATAAACAGGAAAATCATAAAT
CCCTATTTACATAAATAATATATCCTGGTCTTATTCTTAAAACCATTGATTTTTCA
CGGCTCATTAANAAAGCTGGGCGAGGTGGCTCACGCCCCGTCATCCTAGCACTTTGG
GAGGCCGAGGCGGGCANATCACAAGGTGAGGAGTTGGGAGACCAGCCTGACCAACA
CGGTGAAACCCAGTCTCTACTAAAAATACAAAATTANCTGGGGGTGGTGGTGTGT
25 GCCTGTAATCCAAGCTACTCGGGAGGCTGAGGCAGGA

Sequence ID 1382

CTTACTACCTCCAACATGAAACAAGCAGCCCCGCACTTCTCGAAGGTCTGAGTTAC
TTGGAATCGTTTTACCACATGATGGACAGAAGGAATATTTTCAATATCTCTGAAAA
30 CCTCAAGCGTTACCTTCTTCAGTATTTTAAGCCAGTGATTGACAGGCAAAGCTGGA
GTGACAAGGGCTCAGTCTGGGACAGGATGCTCCGCTCGGCTCTCTTGAAGCTGGCC
TGTGACCTGAACCATGCTCCTTGCATCCAGAAAGCTGCTGAACTCTTCTCCCAGTG
GATGGAATCCAGTGGAATAATAATATACCAACAGATGTTTTAAAGATTGTGTATT
CTGTGGGTGCTCAGACAACAGCAGGATGGAATTACCTTTTAGAGCAATATGAACTG
35 TCAATGTCAAGTGCTGAACAAAACAAAATTCTGTATGCTTTGTCAACGAGCAAGCA
TCAGGAAAAGTTACTGAAGTTAATTGAACTAGGAATGGAAGGAAAGGTTATCAAGA
CACAGAACTTGGCAGCTCTCCTTCATGCGATTGCCAGACGTCCAAGGGGCAGCAA

- 269 -

CTAGCATGGGATTTTGTAGAGAAAATTGGACCCATCTTCTGAAAAAATTTGACTT
GGGCTCATATGACATAAGGATGATCATCTCTGGCACAACAGCTCACTTTTCTTCCA
AGGATAAGTTGCAAGAGGTGAACTATTTTTTTGAATCTCTTGAGGCTCAAGGATCA
CATCTGGATATTTTTTCAAACGTCTCTGGAAACGATAACCAAAAATATAAAATGGCT
5 GGAGAAGAATCTTCCGACTCTGAGGACTTGGCTAATGGTTAATACTTAAATGGTCA
ATAGAAAAAGTAGGCTGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGA

Sequence ID 1387

AAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
10 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCAGT
GTTAAAGTAGGTTTGTCTGACGCGGCCACGAATTTCCCGGGGACCAA

Sequence ID 1389

TTTTTTTTTTTTTTTTTGGGAGTCAGTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
15 GNTTTTCGGAAACGGAGTCTCGCTTTCTCGCCCACTCTGGAGTGGNGCAGTGGGGN
GGTCTCAGCTCACACAGCCTCCACCTCCTGGGCCCAAGCGATCCTNTCACCTCAG
CCTCCTGCGTAGCTGGGACTACAGGCGTGCACCACCATTTCCAGGTAATTTTTGTA
TTTTTTGTANANACAGGGTTTCACTGTTGTTGCCAGGCTGGTCTCGAACTCCTGC
TTCAGTCTGCCANAATGCTGGATTCTAGGCGTGAGCCACCGNGCCTGGCCCAAAG
20 TTAATTTCTTACAGAAGCAAAGCTTTAATGCATTTTACTGAATGCTTATAGCTTT
GTAGATACTGAAAAGAGTATGAGCGTCACATACAGACACATNTAACAGCACTGCCT
CCAACCAGCCCCTACCCACTGGTCAGGNGAGTAANAATCAAAATTCTTTTCTGNGA
GTGGAACGGAAATTTTCATCTCTCCTCCTCAGGCAAGTAGTTAANAGGCTGGNGGGA
GTCATGGCCCCATTTTGTTCAAAATACAAGCTCCACAGGAACAAAAGGCTGAACTG
25 CTCACCTCCCAACTGATGAACCTCGTCTTTGTTCCATGTCAAAGGGGCCTTTGTGT
TACTGCAGCAGAACTCCAGCTATCAAACCATCAGGCACCAAAAGTAAACTCCTT
TCTCTAAAAAGACCTCTCTTTACCTGAGCCTTTCAATGCATCTTTGCCCCCANATA
ATCCTGGATGAGATAATCCCCAGAGGAANACCAGCGCTTGCCTAGTGAAATTATAC
TATGAGACAAGGGTAAAAGACCTCAAANACCGGGTTGGCAGGTAAGGGAGTAGGGN

30

Sequence ID 1390

TCNGTGGCACCCGTTTCCGGCACCTTCAGACTCTGAAGAGCCACCTGCGAATCCAC
ACAGGAGAGAAACCTTACCATGTACGTAAGCCTCTTGAGGCCGCTCTCTGACCTGC
GGGGATGTGGAGGGCAGGGAAGGAGGTGGAGCGCAGGGAAGGAGGTGGAGCAGGGA
35 GGCAGTGGAACGTGTTTGCTCCCATCTCAAGCACACAGTGGGGCAACCACTACGCTA
ATGGTTGGAAGACCTAGATCTGGGCCCAATGGCCAGACACCCTGCTTGACCTTGGC
CCAAGCATTAGGGGACTCATCTTTAAATGAGGGTATGGGACTAGATGATCTGGGC

- 270 -

CTTAGGAGAGGAGT

Sequence ID 1391

CGGCTNCTACCCCTGCGGAGATCACACTGACCTGGCAGTGGGATGGGGAGGACCAAA
5 CTCAGGACACCGAGCTTGTGGAGACCAGGCCAGCAGGAGATGGAACCTTCCAGAAG
TGGGCAGCTGTGGTGGTGCCTTCTGGAGAAGAGCAGAGATACACGTGCCATGTTCA
GCACGAGGGGCTGCCGGAGCCCCTCACCCCTGAGATGGAAGCCGTCTTCCCAGCCCA
CCATCCCCATCGTGGGCATCGTTGCTGGCCTGGCTGTCCTGGCTGTCCTAGCTGTC
CTAGGAGCTATGGTGGCTGTTGTGATGTGTAGGAGGAAGAGCTCAGGTGGAAAAGG
10 AGGGAGCTGCTCTCAGGCTGCGTCCAGCAACAGTGCCCAGGGCTCTGATGAGTCTC
TCATCGCTTGTAAGCCTGAGACAGCTGCCTGTGTGGGACTGAGATGCAGGATTTTC
TTCACACCTCTCCTTTTGTGACTTCAAGAGCCTCTGGCATCTCTTTCTGCAAAGGCA
TCTGAATGTGTCTGCGTTCCTGTTAGCATAATGTGAGGAGGTGGAGAGACAGCCCA
CCCCCGTGTCCACCGTGACCCCTGTCCCCACACTGACCTGTGTTCCCTCCCCGATC
15 ATCTTTCCTGTTCCAGAGAAGTGGGCTGGATGTCTCCATCTCTGTCTCAACTTCAT
GGTGCGCTGAGCTGCAACTTCTTACTTCCCTAATGAAGTTAAGAACCTGAATATAA
ATTTGTTTTTCTCAAATATTTGCTATGAAGGGTTGATGGATTAATTAAATAAGTCAA
TTCCTGGAAGTTGAGAGAGCAAATAAAGACCTGAGAACCTTCCANAATCCG

20 Sequence ID 1392

TGAAACAAAATGAATTTNTATGGGTAAGAGAGGGTAATATTTTAGAGTTGTGTTAC
AAACTACAAATTTTTTATTAAATTAATAAATCAGAATACTAAATCCATGTGTTTTT
TTCTTTCTTAAAAAATATCTTTTGGCTGGGCACGGTAGCTCATGGCTGTAATCCCA
GCACTTTGGGAGGCTGAGGTGGGTGGATCGCCTGATGTCAGGAGTTCAAGACCAGC
25 CTGGTCAACATGTTGAAACCCCATCTCTACTAAAAATATAAAAATTAGCCGGTGTG
GTGGTGGGCGCCTGTAATCCCAGCTACTCAGGAGGCTAAGGCAGGAGAATTGCGTG
AACCAGGAGTTCAGTGATGTAGCGGGGAGCTGAGATTGTGCCACTACACTCCAGC
CTGGATGACAGAGTGAGACTCCATCTCAAAAAAAAAAAAAAAAAAAAAA

30 Sequence ID 1394

GCATAATGTGAGGAGGTGGAGAGACAGCCACCCCCGTGTCCACCGTGACCCCTGT
TCCCATGCTGACTTGTGTTTCCTCCCCAGTCATCTTTCCTGTTCCAGAGAGGTGGG
GCTGGATGTCTCCATCTCTGTCTCAACTTTATGTGCACTGAGCTGCAACTTCTTAC
TTCCCTACTGAAAATAAGAATCTGAATATAAATTTGTTTTTCTCAAATATTTGCTAT
35 GAGAGGTTGATGGATTAATTAAATAAGTCAATTCCTGGAATTTGAGAGAGCAAATA
AAGACCTGAGAACCTTCCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

- 271 -

Sequence ID 1395

CTTACCATGTCAGTGCACAGAAATGCTGTCTTGGGATGTAGGAAAAATAAATCCAC
AAAAGCTACCAAGTTTGAAGGGGACCATGAGTCTTCAGGCTGGAGCTTCCAAACCA
GATGAAAACCCCAACAATTAACCTGCAGTTTAAGATCCAGCAGCTGGCCATTTCTGG
5 ACTCAAGGTGAATCGTCTGGATATGTATGGAGAAAAGTACAAACCCTTTAAGGGCA
TAAAATACATGACCAAAGCTGGGAAGTTCCAAGTTCGAACCTGAAGGGAGCATTTG
CTGAGGGAATAGTCTTGACATTTTTTTCATTTCTTACTTGTCTAAAAGTAAAAAAA
AATATCAGCCTGTCTCCTAGGTCAGTCCCCTCCTGGACCCACCCGCTCCCTTTTTTT
CCTTAGCCTTCAGTGCCATGGAATAATCAAGGGAGGAAAAGGTCACCAGGGAGAA
10 CTGGACAGAACTGAAACACAGCAACACCAGTTCTCAAGGACAAGGTGTGTGATGGG
GGTAGGAAGCTTGGTGCTTATGTAACCATTTTAAACGTGGTTTCTATAGGAAAGAC
CAACATTTGTTTAGCTTGCTTGGCTTTAATTATCTAAAGCCAATGAAAGACTTCTT
TGTTGATTTTTTTAAGATAGAAAGATT

15 Sequence ID 1396

CAAACACTATGTTATTTTATGAANAAGACTTGAACATCTATGGATTTTGGTATTTG
CAAGGGGTGAATGGGGTATTTGCAAGCAGTGAATGAGGAGGCCTGGAACCAATCTT
CTGCTGATATTGAGGCACAACCTGAAAAAGGTATATTACTTAAATCTCTTATTGTAT
TGTAAACTGTATAAGTAATGAAATTAAAAGGCAGAAATTGTCAGACTGAATAAAAT
20 GAAAAGACCAAACAATATGCTGCTTACAAGAAACACAATTCAAATATAAGGACACA
ATTAGTTTAAAGGAAAAGAACTGGAAAAGATATACCATGATAACACAAGTCAGAAG
AAAGCTGCTGTGGATATATTAATATGAGATGTAGATTTTCAGAGCAGTGAATATTGC
CAGGCATAAAGAAAGTTATTACATAATAATTAAGGTATCAGTTCATCAAGAAGATG
TAATAACCCCTAAGTATTTATACAACCTAATATCAGAGCTTCAAATACATGAAGCAA
25 AAACCAGTGGAATTGATAGGAGAAACACACAATTACACAATTATAGTCAGAATTTT
CAACATATCTTTCTCAATGGAGAAAACAACCTAGACAGGAAATCATTAAGGATATAG
ATGATTTAAATTATATGATCAACTACCTGGACGTAATTGGCATTTTATGGAACACTG
CACCACCAACAGCAGAGTACATATTATTTTCAAGTACACAGAAAACAGTTACCAAT
ATAGACCATTTTCTGGGTCATAAAACACATCTCAATAAATGTAAAACAATTAATGT
30 TATATAAAGTATGTGCTCTGACCNCAAAGGAATTAGAGATCAATAAAAGAACATCT
TTGAAAAATCTCACNTATTTAAAACTAATAACTCACTTCTAAATAACTCCTGTNT
CAAGAGAATNAAANGG

Sequence ID 1397

35

CCCAGCCTCACTGCGCCCCGTCAGGCCAGGCAGCTGCCCTCAGGGTCTGCCAAGGT
GGGGGTCAAGGGCCATGGGGGCAGGTAGCTCTGCCTGCAAAGCCCACAAGCATGTC

- 272 -

AGATCACCTGGGCTGCAGACAGACAAACACCTGAGCTGTTCTGAATACCTTCAGGT
TCCTGGCCTCGCTGAGCAAGTGCAGAAATTTTACCTTCAAGGATCAGGGTTTTTTC
TGTTTGTGTTGTTTTTTTAAACACACACATATGTGAACAAAGAGTATGCGTTTGTACTG
GCAGAAGAAGCGTCTGGTAAGACAACCAGCAAGTTAACAATGGTCACCTCCAGAAA
5 TGGGCTGGGTAAACCAAAGAATTTTTTTTGTGTTTTTTTGTGAGTCAGGGTCTA
GCTCTGTCACCCAGGCTGGAACGCCTGGTGTGATCACGGCTCACTGCAGCCTTGA
CCTCCCTGGCTCAAGCAATCCTCCCAGCTCAGCCTCCTGAGTCGTTGGGACTACAG
GCACGTGCCACCACGCCTGACACATTTTTTTAAATTTTTGTAGAGACAGTGTTTCAC
CATGTTGCCCAGGCAGGTCTCAAACCTCCTGGGCTCAAGTGGTCCTCCAGCTTCAGC
10 CTCCCAAAGTGCTAGGATTATAGGTGTGAGCCACAGTGCCCAGCCCCGTAGTGGAG
AATTTCTGTTGAATGAACCAAAGCAACTGCCAACCTCTCCATGCACCATGTGTTT
CAGAGGAGAAAGCACAGTGAAGAATGCAGTGTGTTCTGAGGTCCTGTCACCCCTGA
GGCTGTGTGTGTCCTTTGCCAAATTAAAGAGTCTTACTGAATGCGGTGCATCCAGG
AGACAGGCCNAGGTTTGGACTGGTAAAAAAAAA

15

Sequence ID 1399

CAGACACCTGGNAGAACGGGAAGGAGACGCTGCAGCGCGCGGACCCCCCAAAGACA
CATGTGACCCACCACCCCATCTNTGACCATGAGGCCACCCTGAGGTGCTGGGCCCT
GGGCTTCTACCCTGCGGAGATCACACTGACCTGGCAGCGGGATGGCGAGGACCAAA
20 CTCAGGACACCGAGCTTGTGGAGACCAGACCAGCAGGAGACAGAACCTTCCAGAAG
TGGGCAGCTGTGGTGGTGCCTTCTGGAGAAGAGCAGAGATACACATGCCATGTACA
GCATGAGGGGCTGCCGAAGCCCCTCACCTGAGATGGGAGCCATCTTCCCAGTCCA
CCGTCCCCATCGTGGGCATTGTTGCTGGCCTGGCTGTCCTAGCAGTTGTGGTCATC
GGAGCTGTGGTCGCTGCTGTGATGTGTAGGAGGAAGAGTTCAGGTGGAAAAGGAGG
25 GAGCTACTCTCAGGCTGCGTCCAGCGACAGTGCCCAGGGCTCTGATGTGTCTCTCA
CAGCTTGAAAAGCCTGAGACAGCTGTNTTGTGAGGGACTGAGATGCAGGATTTCTT
CACGCCTCCCCCTTGTGACTTCAAGAGCCTCTGGCATCTCTTTCTGCAAAGGCACC
TGAATGTGTCTGCGTCCTTGTTAGCATAATGTGAGGAGGTGGAGAGACAGCCCACC
CTTGTGTCAACTGTGACCCCCTGTTCCCATGCTGACCTGTGTTTCCTCCCCAGTCA
30 TCTTTTTTGTTCNCAATAGGTGGGGCCTGGATGTCTCCATCTCTGTNTCA

Sequence ID 1440

TTATAAGGTACTTTTAAGGTATTTTAGTTGTCTTAGTCTATATTTCTGTACTCACC
35 TTTCTTTATCCACTCATCAGTTGATGGGCATGTAGGTTGGTTCCATATCTTTGCAA
TTCTGAATTGTGCTGTGATCAGGTGTCTTTTTAGTATAATGATTTACTCTCCTTTG
GGTAGATACCCAGTAGTGGGATTGCTGGATCGAATGGTTTTTATAATTTTCTATTT

- 273 -

TACCACAGTTTCTCTCTGCATTTTTCTCTTTGACCACTAACCATGTGAAATTCTC
ATATTGACCTTTATAATGATCATGAACTCTTAGTATCATTGGGAAGGCCACATTTG
CCACTTATGATTGTAAACCTTATCCTCCATTTTTCTCTGTTATTGTTGGTGCAAAA
GCACCTATTATACCAGGACTTTAAAAATCAGTCTGATAAGTCTTTGATAAGTCTAA
5 TAATAATAACTGATAAGTCCATTGAATTTGCTTCTGATTACTTTTTCTTTAGTAGC
TAAACATGTATGTACTCCTATGATTACAATGAACACTCCTCTCCATTTAAATTAAT
TATTTACATTGATGAAATAGCAAAATGTTAATGACTAAATACTGTCTTGGTTTTTT
CGTTCCAGGTCAGTCAATATTAACCTTCTTATAATTTTCTTTTTTTTTCTTT

10 Sequence ID 1447

GCAAGGACTAACCCCTATACCTTCTGCATAATGAATTAACCTAGAAATAACTTTGCA
AGGAGAGCCAAAGCTAAGACCCCCGAAACCAGACGAGCTACCTAAGAACAGCTAAA
AGAGCACACCCGTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGGCGACA
AACCTACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAACTTT
15 AAATTTGCCACAGAACCCCTCTAAATCCCCTTGTAATTTAACTGTTAGTCCAAAG
AGGAACAGCTCTTTGGACACTAGGAAAAAACCTTGTAGAGAGAGTAAAAAATTTAA
CACCCATAGTAGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACAC
CCACTACCTAAAAAATCCCAAACATATACTGAACTCCTCACACCCAATTGGACCA
ATCTATCACCCCTATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCC
20 TCCGCATAAGCCTGCGTCAGATTAAACACTGAACTGACAATTAACAGCCCAATAT
CTACAATCAACCAACAAGTCATTATTACCCTCACTGTCAACCCAACACAGGCATGC
TCATAAGGAAAGGT

Sequence ID 1448

25 GGCCACCGGGTGCAAGGTCAGGGCTGGGGTGGAGGCTGGGAAGCCCAGGGCTTGGC
CCACTGTGGCCGCCTTGTGTGGTCACTGCTTTCTCTGGGCCTGCTGTGAGCTCCCTC
TAGGACCCCAGGCCTGTCTGGTGGGTCACTGTGACCACCACCTTGCACAGCACCTG
GCGCGTGGCAGGTGCTCAAACATTACTTGTTTCGGAATGAACTTCATCTTGCTCTT
GGCTTTTTGACTAATGCTGTGGAACATCTGACTAATTAGTGACTCTTTGGGGCCCC
30 CAGTTTCCCAGCTATAAAGTGGTAATATTAAGATAATAATTCGGCCGGGCGCGGTG
GCTCACGCCTGTAATCCCAGCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACGAG
GTCAGAAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCATCTCTACTAAAAA
TACAAAAAATTANCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCANGAG
GCTGANGCAGGAGAATGGTGTGAACCCGGGAGGCAGAGGTTGCAGTGAACCAAGAT
35 CGNNCCACTGCACTCCAGCCTGGGCAACAGAGCGAGACTCCATCTTAAAAAA

Sequence ID 1449

- 274 -

AATCAGGGCCGCGAGTGTGTTCTGCGCCTGCCCAGAGCTGACTCCTGATTTAACCGC
TGGCGTAACCGCGGGTTGCACGCATGCGTGCTGAAAAGCCTTTCACCCTCACGTGG
TTTCTTTTTTAACCAGTCATCAAGCGAGGCTCGCGCGCAGGCCCCGCGTTGGAAAA
TGGCGGGGAAGCTGAAACCTCTGAATGTGGAGGCGCCAGAAGCTGCTGAGGAGGCT
5 GAAGGTAGTGAGGGCAAGTGGGCTGCACTCCTTTCTCTCCAACCAGGGCAGAAAGG
AGGGAGGATTCGTCCCATTACAATAATGAAATAATGATATTCTAATTTTTTTTAAAT
AAAATGTTAAGCCTTTTGTTATTGAA

Sequence ID 1450

10 GGAAANCATGAGGCTTCGGGAGCCGCTCCTGAGCGGCAGCGCCGCGATGCCAGGCG
CGTCCCTACAGCGGGCCTGCCGCCTGCTCGTGGCCGTCTGCGCTCTGCACCTTGGC
GTCACCCTCGTTTACTACCTGGCTGGCCGCGACCTGAGCCGCCTGCCCCAACTGGT
CGGAGTCTCCACACCGCTGCAGGGCGGCTCGAACAGTGCCGCCGCCATCGGGCAGT
CCTCCGGGGAGCTCCGGACCGGAGGGGCGCCGCCGCCCTCCTNTAGGCGCCTCC
15 TCCCAGCCGCGCCCGGGTGGCGACTCCAGCCCAGTCGTGGATTCTGGCCCTGGCCC
CGCTAGCAACTTGACCTCGGTCCCAGTGCCCCACACCACCGCACTGTGCTGCCCCG
CCTGCCCTGAGGAGTCCCCGCTGCTTGGTAAGGACTCGGGTCGGCGCCAGTCGGAG
GATTGGGACCCCCCGGATTTCCCCGACAGGGTCCCCCANACATTCCTCAGGCTG
GCTCTTCTACGACAGCCAGCCTCCCTCTTCTGGATCAGAGTTTTTAAATCCCANACA
20 GAGGCTTGGGACTGGATGGGAGAGAAGGTTTGCGAGGTGGGTCCCTGGGGAGTCCT
GTTGGAGGCGTGGGGCCGGGACCGCACAGGGAAGTCCCGAGGCCCTCTAGCCCCA
AAACCANAGAAGGCCTTGGAGACTTCCTTGCTGTGGCCCGAGGCTNAGGAAGTTTT
GGAGTTTTGGGTCTGCTTANGGCTTCNAGCAGCCTTGCACTGAGAACTTTGGTAGG
GACCTCGAGTAATCCACTCCNTTTTNGGGACTGACGTGAGGCTCCCGGTGGGGAAA
25 GANACTGACCTNTC

Sequence ID 1453

CCGACCTGTCTCGCTCCGTGGCCTTAGCTGTGCTCGCGCTACTCTCTCTTTCTGGC
CTGGAGGCTATCCAGCGTACTCCAAAGATTCAGGTTTACTCACGTCATCCAGCAGA
30 GAATGGAAAGTCAAATTTCTGAATTGCTATGTGTCTGGGTTTCATCCATCCGACA
TTGAAGTTGACTTACTGAAGAATGGAGAGAGAATTGAAAAAGTGGAGCATTCAGAC
TTGTCTTTCAGCAAGGACTGGTCTTTCTATCTCTTGTACTACACTGAATTCACCCC
CACTGAAAAAGATGAGTATGCCTGCCGTGTGAACCATGTGACTTTGTACAGCCCA
AGATAGTTAAGTGGGATCGAGACATGTAAGCAGCATCATGGAGGTTTGAAGATGCC
35 GCATTTGGATTGGATGAATTCCAAATCTGCTTGCTTGCTTTTTAATATTGATATG
CTTATACACTTACACTTTATGCACAAAATGTAGGGTTATAATAATGTTAACATGGA
CATGATCTTCTTTATAATTCTACTTTGAGTGCTGTCTCCATGTTTGATGTATCTGA

- 275 -

GCAGGTTGCTCCACAGGTAGCTCTAGGAGGGCTGGCAACTTAGAGGTGGGGAGCAG
AGAATTCTCTTATCCAACATCAACATCTTGGTCAGATTTGAACTCTTCAATCTCTT
GCACTCAAAGCTTGTTAAGATAGTTAAGCGTGCATAAGTTAACTTCCAATTTACAT
ACTCTGCTTAGAATTTGGGGGAAAATTTAGAAATATAATTGACAGGATTATTGGAA
5 ATTTGTTATAATGAATGAAACATTTTTTGT CATATAAGATTCATATTTACTTCTTAT
ACA

Sequence ID 1454

TAAATAGGGAATCCTTTCCCCATTGCTTGTTTTTCTCAGGTTTGTCAAAGATCAGA
10 TAGTTGTAGATATGCGACGTTATTTCTGAGGGCTCTGTTCTGTTCCATTGATCTAT
ATCTCTGTCACATGCACACGTATGTTTGTGTGGCACTATTACAGTGGCAAAGAC
TTGGAACCAACCCAAATGTCCAACAATGATAGACCGGGTTAAGAAAATGCGGCACA
TATACACCATGGAATACTATGTAGCCATAAAAAATGATGAGTTCGTGTCCTTTGTA
GGGACATGGATGAAATTGGAAATCATCATTCTCAGTAAACTATCGCAGGAACAAAA
15 AACCAAACACTGCATATTCTCACTCATAGGTGGGAATTGAACAGTGGGAACACATG
GACACAGGAAGGGGAACATCACACTCTGAGGACTGTTGTGGGGTGGGGGGAGGGAG
GAGGGATAGCATTGGGAGATATACCTAGTGCTGGATGACGAGTTAGTGGGTGCAGC
GCACCAGCATGTCACATGTATACATATGTAACCTAACCTGCACATTGTGCACATGTA
CCCTAAAACTTAAGGTAT

20

Sequence ID 1456

CCGCAACAAACACGGGAGTGCAGATATCGCTGCGATGGGCTGATTTCTTTTATTTG
GGTATATACCCAGCAGTGGGATTGCTGGATTGTATGGTAGCTCTATTAGTTTTTTG
25 AGGAACCTCCAAACTGTTCTNCATAGTGGTTGTACTCATTTACATTCCCACTGTGA
ACCCTGAAAATTTGAGGCAGGTCTCAGTTAAATTAGAAAGTTGATTTTGCCAAGTT
GGGGACACGCACTCGTGACACAGCCTCAGGAGGAACTGATGACATGTGCCCAGGTG
GTCAGAGCACAGCTTGGTTTTTATACATTTTAGGGAAACCTGAGCCATCAATCAACA
TACGTAAAATGGGCCGGGCACAGCAGCTCAAGCTGTAATCCCAGCACTCTGGGAGG
30 CCGAGGCGGGTGGATCACTTGAGGTCAGGAGTTGAGACCAGCCTGGCCAACATGG
TGAAACCCCGTCTCTATTAAAAATACAAAGCTTAGCTGGATGTGGTGGCGCATGCC
TGTAGTCCCAGCTGCTCTAGGAGGCTGAGGCATGAGAATTGCTTGAACCTGGGAGG
CAGAGGCTGCAGTGAGCCGAGATCGAGCCACTATACTCCAGCCTGGTCAACAGAGT
GAGACCCTGTCT

35

Sequence ID 1460

CCACAACCTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCTGACTCCTG

- 276 -

AGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGT
GGTGAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCCTTGGACCCAGAGGTTCTTTGA
GTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGG
CTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAAC
5 CTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGA
TCCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACT
TTGGCAAAGAATTCACCCCACCAGTGCAAGGCTGCCTATCAGAAAGTGGTGGCTGGT
GTGGCTAATGCCCTGGCCCAACAAGTATCACTAAGCTCGCTTTCTTGCTGTCCAATT
TCTATTAAAGGTTCTTTTGTTCCCTAAGTCCAATACTAACTGGGGGATATTATG
10 AAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTG

Sequence ID 1490

ATGGGCATCTCTCGGGACAACCTGGCACAAGCGCCGCAAAACCGGGGGCAAGAGAAA
GCCCTACCACAAGAAGCGGAAGTATGAGTTGGGGCGCCAGCTGCCAACACCAAGA
15 TTGGCCCCCGCCGCATCCACACAGTCCGTGTGCGGGGAGGTAACAAGAAATACCGT
GCCCTGAGGTTGGACGTGGGGAATTTCTCCTGGGGCTCANAGTGTTGTACTCGTAA
AACAAGGATCATCGATGTTGTCTACAATGCATCTAATAACGAGCTGGTTCGTACCA
AGACCCTGGTGAAGAATTGCATCGTGCTCATCGACAGCACACCGTACCGACAGTGG
TACGAGTCCCCTATGCGCTGCCCCCTGGGCCGCAAGAAGGGAGCCAAGCTGACTCC
20 TGAGGAAGAAGAGATTTTAAACAAAAACGATCTAAAAAATTCAGAAGAAATATG
ATGAAAGGAAAAAGAATGCCAAAATCAGCAGTCTCCTGGAGGAGCAGTTCCAGCAG
GGCAAGCTTCTTGCGTGCATCGCTTCAAGGCCGGGACAGTGTGGCCGAGCAGATGG
CTATGTGCTAGAGGGCAAAGAGTTGGAGTTCTATCTTAGGAAAATCAAGGCCCGCA
AAGGCAAATAAATCCTTGTTTTGTCTTCACCCATGTAATAAAGGTGTTTATTGTTT
25 TTGTT

Sequence ID 1491

CTTNCACATACTGATTGATGTCTCATGTCTCTCTAAATGTGTAAAACCAAGCTGT
GCCCCAACCACTTGGGNACATGTGGNGAGGACCTCCTGAGGCTGTGTCATGGGCA
30 CACCTTAACCCTGGGAAAATAAACTTTCTAACTGACTTGAGAGCTGTCTCAGATA
TTCTGAGCTTACAGTTATTGTGAAATCATTTTAATTATAAATTAAGTGGAGATTTA
CTTAAAATCATGTGTAGAAGTAGCCTGTGATATAGTCCTAGATACATACATTATCA
TCTTATGTATCTTCCCTCCCTCTTCCAGGTTCTGATAAAAACAGATGAAATCTGAA
AGACCATGACAGTAGTATTTTGAAAATGACAGTATTTGAAATTAAAAAATTGTAAA
35 AGTGTTCTGTTCTATCACTGCCAAAGGATAAGTTACAAATTGGTTCTTGGAACGTA
ATATGTACTATGTGCTTGCTATTTAATAATTTACCAGTCTTAGTCTTTTTTTATTCA
GACTAATTTTACCTTTTTTTTAACTATGACTCTTTAGTTATAGTAGTACAAAAAAG

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 276

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 276

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

- 280 -

Claims:

1. A set of oligonucleotide probes, wherein said set comprises at least 10 different oligonucleotides,
5 wherein each oligonucleotide is selected from:

an oligonucleotide having a sequence as set forth in SEQ
ID No. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14,
15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28,
10 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42,
43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56,
57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70,
71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84,
85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98,
15 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109,
110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120,
121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131,
132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142,
143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153,
20 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164,
165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175,
176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186,
187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197,
198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208,
25 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219,
220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230,
231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241,
242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252,
253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263,
30 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274,
275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285,
286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296,
297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307,
308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318,
35 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329,
330, 331, 332, 333, 334, 335, 336, 337, 339, 341, 342,
343, 344, 345, 346, 347, 348, 351, 352, 353, 355, 356,

- 281 -

357, 359, 361, 363, 364, 365, 366, 367, 368, 369, 370,
371, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383,
384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394,
395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405,
5 406, 409, 411, 414, 415, 416, 418, 421, 422, 423, 424,
425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435,
436, 437, 438, 439, 440, 441, 444, 445, 447, 448, 451,
452, 454, 455, 458, 459, 460, 461, 463, 464, 465, 467,
468, 471, 472, 473, 474, 475, 476, 480, 481, 482, 484,
10 487, 489, 490, 496, 497, 498, 499, 500 or 501,

or an oligonucleotide derived from said sequence, or an
oligonucleotide with a complementary sequence, or a
functionally equivalent oligonucleotide.

15

2. A set of oligonucleotide probes as claimed in claim
1 wherein said oligonucleotide probes are each selected
from:

20 an oligonucleotide having a sequence as set forth in SEQ
ID No. 1, 2, 3, 4, 5, 11, 12, 13, 19, 25, 31, 32, 33,
34, 36, 37, 39, 45, 46, 47, 48, 50, 55, 56, 60, 61, 64,
66, 68, 73, 74, 75, 76, 77, 78, 80, 83, 85, 86, 90, 96,
98, 99, 100, 101, 105, 106, 107, 109, 111, 114, 115,
25 116, 117, 119, 120, 121, 122, 123, 124, 125, 127, 128,
130, 131, 132, 133, 135, 136, 137, 138, 144, 145, 146,
147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157,
158, 159, 161, 162, 163, 164, 165, 166, 168, 169, 171,
173, 174, 175, 176, 177, 178, 179, 180, 182, 183, 185,
30 186, 187, 190, 191, 195, 197, 198, 199, 200, 202, 204,
206, 207, 210, 212, 214, 216, 217, 218, 219, 220, 221,
222, 224, 225, 229, 230, 231, 232, 233, 234, 235, 236,
237, 238, 239, 240, 243, 244, 245, 249, 251, 256, 258,
259, 260, 261, 262, 267, 268, 270, 272, 273, 274, 275,
35 276, 278, 279, 280, 282, 284, 286, 287, 289, 291, 292,
295, 296, 297, 298, 299, 301, 303, 305, 307, 308, 309,
310, 311, 314, 315, 316, 317, 318, 319, 320, 321, 322,

- 282 -

323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333,
334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344,
345, 346, 347, 348, 349, 351, 352, 353, 355, 356, 357,
359, 361, 363, 364, 365, 366, 367, 368, 369, 370, 371,
5 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384,
385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395,
396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406,
408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418,
419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429,
10 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440,
441, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453,
454, 455, 457, 458, 459, 460, 461, 463, 464, 465, 466,
467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477,
478, 479, 480, 481, 482, 484, 487, 489, 490, 496, 497,
15 498, 499 or 501,

or an oligonucleotide derived from said sequence, or an
oligonucleotide with a complementary sequence, or a
functionally equivalent oligonucleotide.

20

3. A set of oligonucleotide probes as claimed in claim
1 wherein said oligonucleotide probes are each selected
from:

25 an oligonucleotide having a sequence as set forth in SEQ
ID No. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,
19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,
33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46,
47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60,
30 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74,
75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88,
89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101,
102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112,
113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123,
35 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134,
135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145,
146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156,

- 283 -

157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167,
168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178,
179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189,
190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200,
5 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211,
212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222,
223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233,
235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245,
246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256,
10 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267,
268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278,
279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289,
290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300,
301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311,
15 312, 313, 339, 341, 342, 343, 344, 345, 346, 347, 348,
350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360,
361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371,
372, 373, 374, 397, 398, 399, 400, 401, 402, 403, 404,
405, 406, 407, 409, 415, 445, 447, 448, 451, 452, 454,
20 455, 456, 458, 459, 460, 461, 462, 463, 464, 465, 467,
468, 471, 472, 473, 474, 475, 480, 481, 482, 483, 484,
485, 486, 487, 488, 491, 492, 493, 494, 495, 496, 500 or
501,

25 or an oligonucleotide derived from said sequence, or an
oligonucleotide with a complementary sequence, or a
functionally equivalent oligonucleotide.

30 4. A set of oligonucleotide probes as claimed in any
one of claims 1 to 3, wherein each probe in said set
binds to a different transcript.

5. A set as claimed in any one of claims 1 to 4
consisting of from 10 to 500 oligonucleotide probes.

35

6. An oligonucleotide probe wherein said probe is
selected from:

- 284 -

an oligonucleotide having a sequence as set forth in SEQ
ID No. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14,
15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28,
29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42,
5 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56,
57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70,
71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84,
85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98,
99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109,
10 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120,
121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131,
132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142,
143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153,
154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164,
15 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175,
176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186,
187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197,
198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208,
209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219,
20 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230,
231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241,
242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252,
253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263,
264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274,
25 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285,
286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296,
297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307,
308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318,
319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329,
30 330, 331, 332, 333, 334, 335, 336, 337, 339, 341, 342,
343, 344, 345, 346, 347, 348, 351, 352, 353, 355, 356,
357, 359, 361, 363, 364, 365, 366, 367, 368, 369, 370,
371, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383,
384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394,
35 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405,
406, 409, 411, 414, 415, 416, 418, 421, 422, 423, 424,
425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435,

AMENDED SHEET

- 285 -

436, 437, 438, 439, 440, 441, 444, 445, 447, 448, 451,
452, 454, 455, 458, 459, 460, 461, 463, 464, 465, 467,
468, 471, 472, 473, 474, 475, 476, 480, 481, 482, 484,
487, 489, 490, 496, 497, 498, 499, 500 or 501,

5

or an oligonucleotide derived from said sequence, or a
complementary sequence thereof.

10 7. A set of oligonucleotide probes as claimed in any
one of claims 1 to 5, or an oligonucleotide probe as
claimed in claim 6, wherein each of said oligonucleotide
probes is from 15 to 200 bases in length.

15 8. A set of oligonucleotide probes as claimed in any
one of claims 1 to 5 or 7 or an oligonucleotide probe as
claimed in claim 6 or 7, wherein the transcript to which
said probe binds is derived from a gene which is
constitutively moderately or highly expressed.

20 9. A set of oligonucleotide probes as claimed in any
one of claims 1 to 5, 7 or 8 or an oligonucleotide probe
as claimed in any one of claims 6 to 8, wherein said
probes are immobilized on one or more solid supports.

25 10. A set of oligonucleotide probes or an
oligonucleotide probe as claimed in claim 9, wherein
said solid support is a sheet, filter, membrane, plate
or biochip.

30 11. A polypeptide encoded by the mRNA sequence to
which an oligonucleotide as defined in claim 6 binds.

12. An antibody to a polypeptide as defined in claim
11.

35

13. A kit comprising a set of oligonucleotide probes
immobilized on one or more solid supports as defined in

- 286 -

claim 9 or 10.

14. A kit as claimed in claim 13 wherein said probes
are immobilized on a single solid support and each
5 unique probe is attached to different region of said
solid support.

15. A kit as claimed in claim 13 or 14 further
comprising standardizing materials.

10

16. The use of a set of probes as described in any one
of claims 1 to 5 or 7 to 10 or a kit as described in any
one of claims 13 to 15 to determine the gene expression
pattern of a cell which pattern reflects the level of
15 gene expression of genes to which said oligonucleotide
probes bind, comprising at least the steps of:

a) isolating mRNA from said cell, which may
optionally be reverse transcribed to cDNA;

b) hybridizing the mRNA or cDNA of step (a) to a
20 set of oligonucleotides or a kit as defined in any one
of claims 1 to 5, 7 to 10 or 13 to 15; and

c) assessing the amount of mRNA or cDNA hybridizing
to each of said probes to produce said pattern.

25 17. A method of preparing a standard gene transcript
pattern characteristic of a disease or condition or
stage thereof in an organism comprising at least the
steps of:

a) isolating mRNA from the cells of a sample of one
30 or more organisms having the disease or condition or
stage thereof, which may optionally be reverse
transcribed to cDNA;

b) hybridizing the mRNA or cDNA of step (a) to a
set of oligonucleotides or a kit as defined in any one
35 of claims 1 to 5, 7 to 10 or 13 to 15 specific for said
disease or condition or stage thereof in an organism and
sample thereof corresponding to the organism and sample

- 287 -

thereof under investigation; and

5 c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce a characteristic pattern reflecting the level of gene expression of genes to which said oligonucleotides bind, in the sample with the disease, condition or stage thereof.

18. A method of preparing a test gene transcript pattern comprising at least the steps of:

10 a) isolating mRNA from the cells of a sample of said test organism, which may optionally be reverse transcribed to cDNA;

15 b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotides or a kit as defined in any one of claims 1 to 5, 7 to 10 or 13 to 15 specific for a disease or condition or stage thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and

20 c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce said pattern reflecting the level of gene expression of genes to which said oligonucleotides bind, in said test sample.

25 19. A method of diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism, comprising the steps of:

a) isolating mRNA from the cells of a sample of said organism, which may optionally be reverse transcribed to cDNA;

30 b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotides or a kit as defined in any one of claims 1 to 5, 7 to 10 or 13 to 15 specific for said disease or condition thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation;

35 c) assessing the amount of mRNA or cDNA

- 288 -

hybridizing to each of said probes to produce
a characteristic pattern reflecting the level
of gene expression of genes to which said
oligonucleotides bind in said sample; and
5 d) comparing said pattern to a standard
diagnostic pattern prepared as described in
claim 17 using a sample from an organism
corresponding to the organism and sample under
investigation to determine the degree of
10 correlation indicative of the presence of said
disease or condition or a stage thereof in the
organism under investigation.

20. A method as claimed in any one of claims 17 to 19
15 wherein said mRNA or cDNA is amplified prior to step b).

21. A method as claimed in any one of claims 17 to 20
wherein the oligonucleotides and/or the mRNA or cDNA are
labelled.

20 22. A method as claimed in any one of claims 17 to 21
wherein said probes are as defined in claim 3 and said
disease is Alzheimer's disease.

25 23. A method as claimed in any one of claims 17 to 21
wherein said probes are as defined in claim 2 and said
disease is breast cancer.

30 24. A method as defined in any one of claims 17 to 23,
wherein said set of oligonucleotides as defined in any
one of claims 1 to 5, 7 to 10 or 13 to 15 are replaced
with a set of oligonucleotides which are randomly
selected, preferably from a cDNA library.

35 25. A method of preparing a standard gene transcript
pattern characteristic of a disease or condition or
stage thereof in an organism comprising at least the

steps of:

a) releasing target polypeptides from a sample of one or more organisms having the disease or condition or stage thereof;

5

b) contacting said target polypeptides with one or more binding partners, wherein each binding partner is specific to a marker polypeptide (or a fragment thereof) encoded by the gene to which an oligonucleotide having a sequence as set forth in SEQ ID No. 1, 2, 3, 4, 5, 6, 7,

10

8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,

22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35,

36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49,

50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63,

64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77,

15

78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91,

92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104,

105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115,

116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126,

127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137,

20

138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148,

149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159,

160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170,

171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181,

182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192,

25

193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203,

204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214,

215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225,

226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236,

237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247,

30

248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258,

259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269,

270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280,

281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291,

292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302,

303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313,

35

314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324,

325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335,

- 290 -

336, 337, 339, 341, 342, 343, 344, 345, 346, 347, 348,
351, 352, 353, 355, 356, 357, 359, 361, 363, 364, 365,
366, 367, 368, 369, 370, 371, 374, 375, 376, 377, 378,
379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389,
5 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400,
401, 402, 403, 404, 405, 406, 409, 411, 414, 415, 416,
418, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430,
431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441,
444, 445, 447, 448, 451, 452, 454, 455, 458, 459, 460,
10 461, 463, 464, 465, 467, 468, 471, 472, 473, 474, 475,
476, 480, 481, 482, 484, 487, 489, 490, 496, 497, 498,
499, 500 or 501 (or derived from said sequence) binds,
to allow binding of said binding partners to said target
polypeptides, wherein said marker polypeptides are
15 specific for said disease or condition thereof in an
organism and sample thereof corresponding to the
organism and sample thereof under investigation; and
c) assessing the target polypeptide binding to said
binding partners to produce a characteristic pattern
20 reflecting the level of gene expression of genes which
express said marker polypeptides, in the sample with the
disease, condition or stage thereof.

26. A method of preparing a test gene transcript
25 pattern comprising at least the steps of:
a) releasing target polypeptides from a sample of
said test organism;
b) contacting said target polypeptides with one or
more binding partners, wherein each binding partner is
30 specific to a marker polypeptide (or a fragment thereof)
encoded by the gene to which an oligonucleotide as
defined in claim 25 binds, to allow binding of said
binding partners to said target polypeptides, wherein
said marker polypeptides are specific for said disease
35 or condition thereof in an organism and sample thereof
corresponding to the organism and sample thereof under
investigation; and

- 291 -

c) assessing the target polypeptide binding to said binding partners to produce a characteristic pattern reflecting the level of gene expression of genes which express said marker polypeptides; in said test sample.

5

27. A method of diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism comprising the steps of:

10 a) releasing target polypeptides from a sample of said organism;

b) contacting said target polypeptides with one or more binding partners, wherein each binding partner is specific to a marker polypeptide (or a fragment thereof) encoded by the gene to which an oligonucleotide as
15 defined in claim 25 binds, to allow binding of said binding partners to said target polypeptides; wherein said marker polypeptides are specific for said disease or condition thereof in an organism and sample thereof corresponding to the organism and sample thereof under
20 investigation; and

c) assessing the target polypeptide binding to said binding partners to produce a characteristic pattern reflecting the level of gene expression of genes which express said marker polypeptides in said sample; and

25 d) comparing said pattern to a standard diagnostic pattern prepared as described in claim 25 using a sample from an organism corresponding to the organism and sample under investigation to determine the degree of correlation indicative of the presence of said disease
30 or condition or a stage thereof in the organism under investigation.

28. A method as claimed in any one of claims 17 to 27 wherein said pattern is expressed as an array of numbers
35 relating to the expression level associated with each probe.

- 292 -

29. A method as claimed in any one of claims 17 to 28 wherein said organism is a eukaryotic organism, preferably a mammal.

5 30. A method as claimed in claim 29 wherein said organism is a human.

31. A method as claimed in any one of claims 17 to 30 wherein the data making up said pattern is
10 mathematically projected onto a classification model.

32. A method as claimed in any one of claims 17 to 31 wherein said disease is cancer or a degenerative brain disorder.

15 33. A method as claimed in any one of claims 17 to 32 wherein said sample is tissue, body fluid or body waste.

34. A method as claimed in any one of claims 17 to 33 wherein said sample is peripheral blood.
20

35. A method as claimed in any one of claims 17 to 34 wherein the cells in the sample are not disease cells, have not been in contact with such cells and do not
25 originate from the site of the disease or condition.

36. A method as claimed in any one of claims 19 to 35 for the diagnosis, identification or monitoring of two or more diseases, conditions or stages thereof in an
30 organism, wherein said pattern produced in step c) is compared to at least two standard diagnostic patterns prepared as described in claim 17 or 25, wherein each standard diagnostic pattern is a pattern generated for a different disease or condition or stage thereof.

35 37. A method of identifying probes useful for diagnosing or identifying or monitoring a disease or

- 293 -

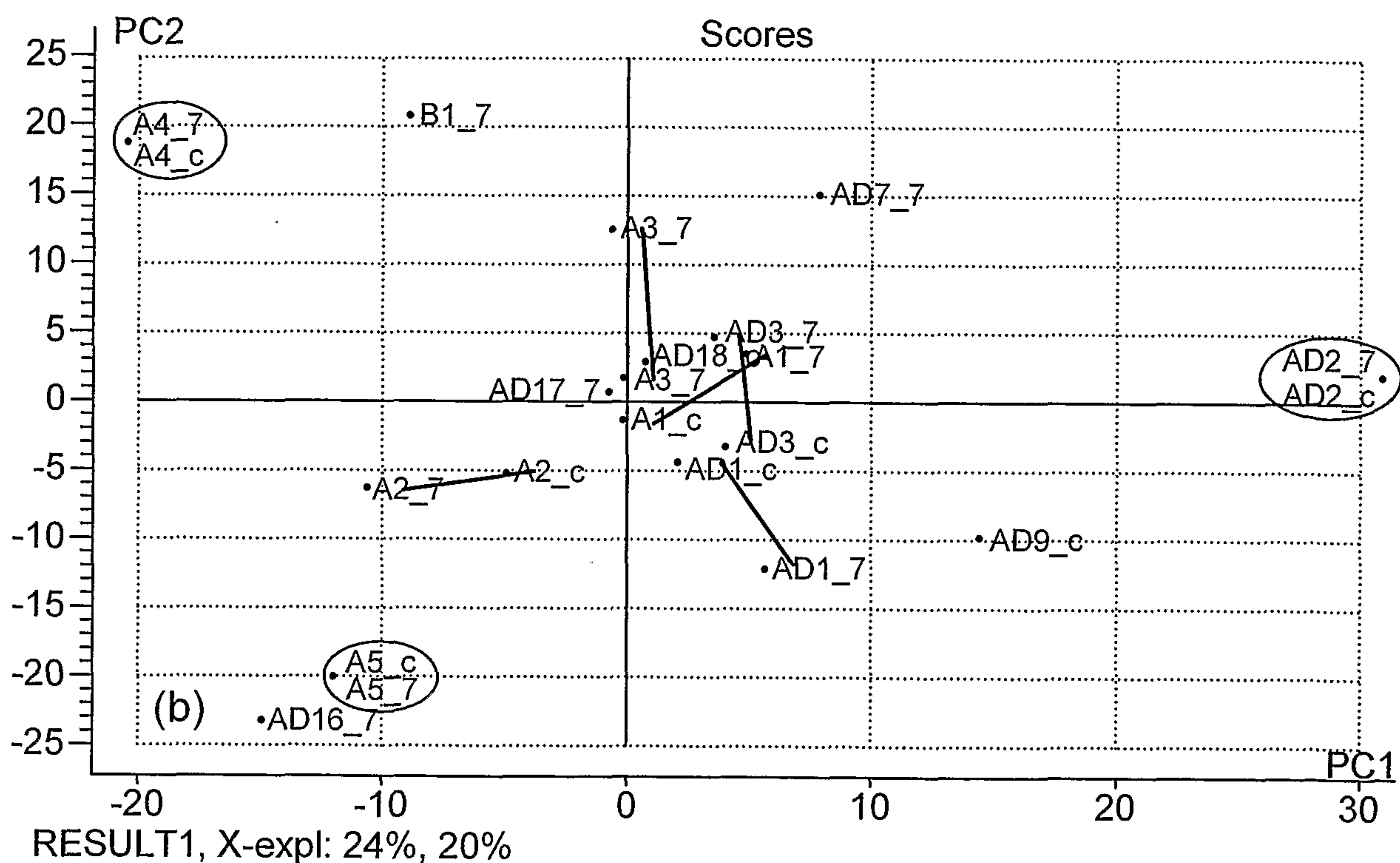
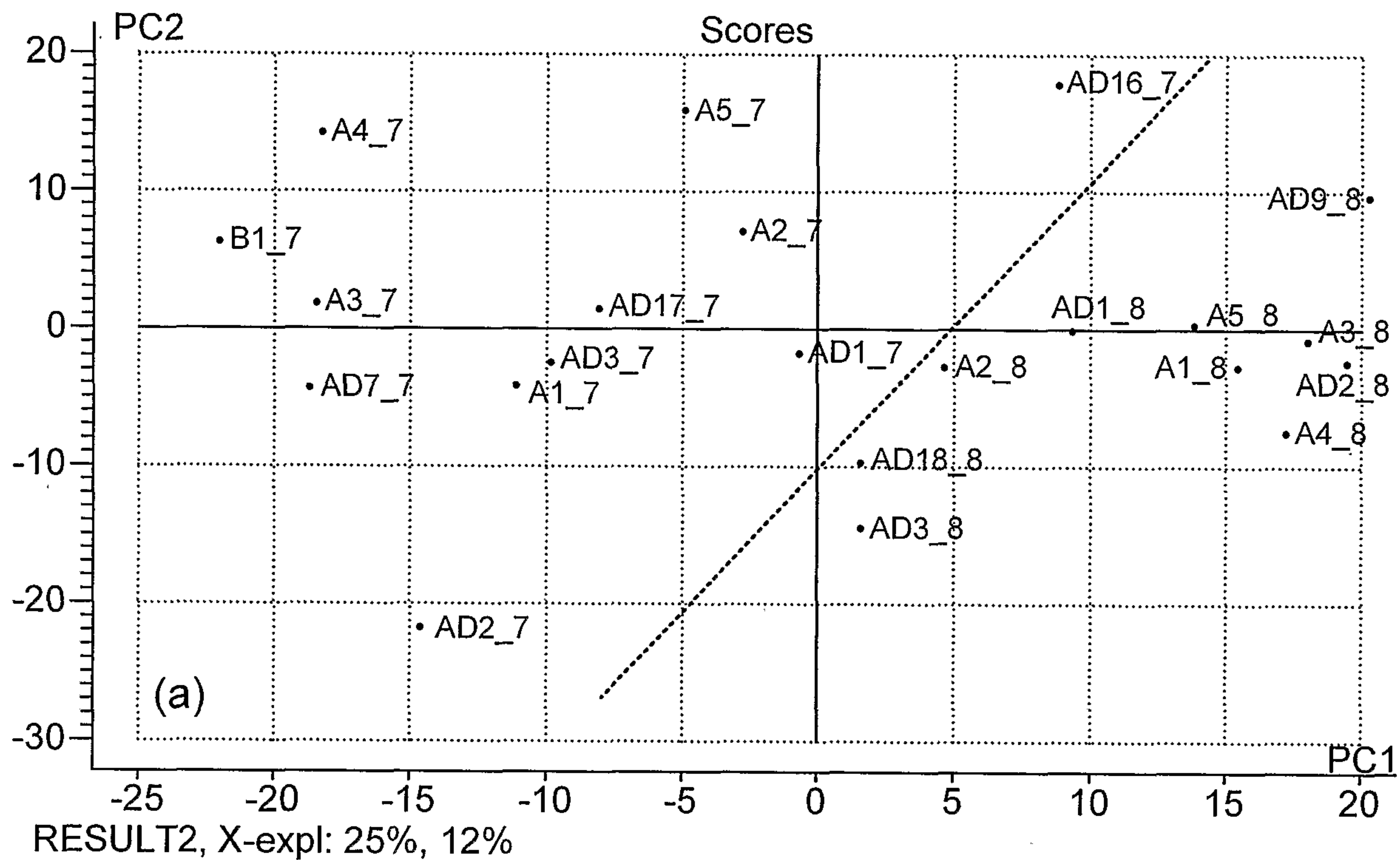
condition or stage thereof in an organism, comprising the steps of:

- 5 a) immobilizing a set of oligonucleotide probes, preferably as described hereinbefore, on a solid support;
- b) isolating mRNA from a sample of a normal organism (normal sample), which may optionally be reverse transcribed to cDNA;
- 10 c) isolating mRNA from a sample from an organism, corresponding to the sample and organism of step (b), which is known to have said disease or condition or a stage thereof (diseased sample), which may optionally be reverse transcribed to cDNA;
- 15 d) hybridizing the mRNA or cDNA of steps (b) and (c) to said set of immobilized oligonucleotide probes of step (a); and
- e) assessing the amount of mRNA or cDNA hybridizing to each of said oligonucleotide probes to determine the level of gene expression of genes to which said oligonucleotide probes bind in said normal and diseased samples to generate a gene expression data set for each sample;
- 20 f) normalizing and standardizing said data set of step (e);
- g) constructing a calibration model for classification, preferably using the statistical techniques Partial Least Squares Discriminant Analysis (PLS-DA) and Linear Discriminant Analysis (LDA);
- 30 h) performing JackKnife analysis and identifying those oligonucleotide probes which are required for classification of said disease and normal samples into their respective groups.
- 35

1 / 12

FIG. 1

Effect of Direct Standardization (DS) on the Alzheimer
Data measured in two different series of experiments



2 / 12

FIG. 1 CONT'D

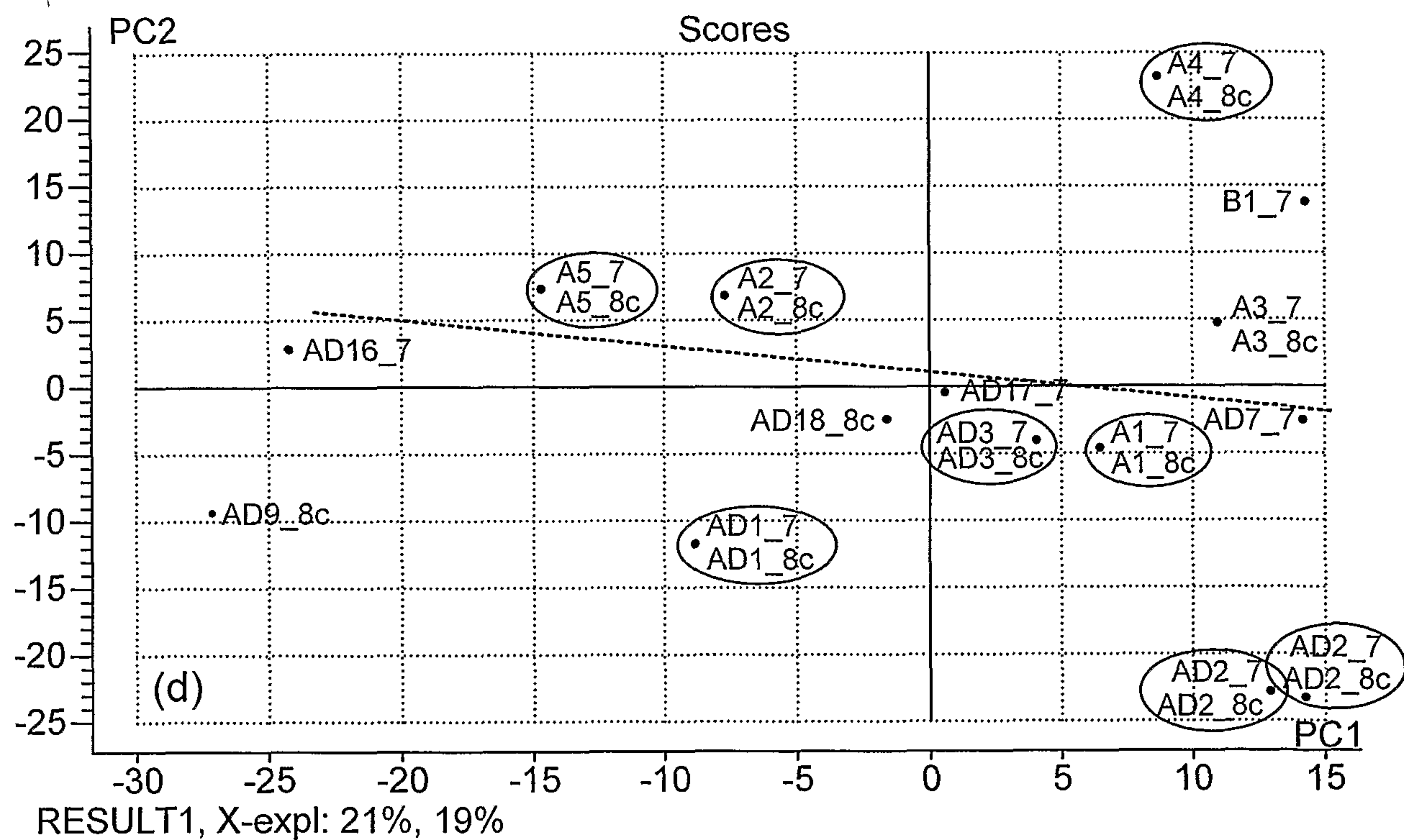
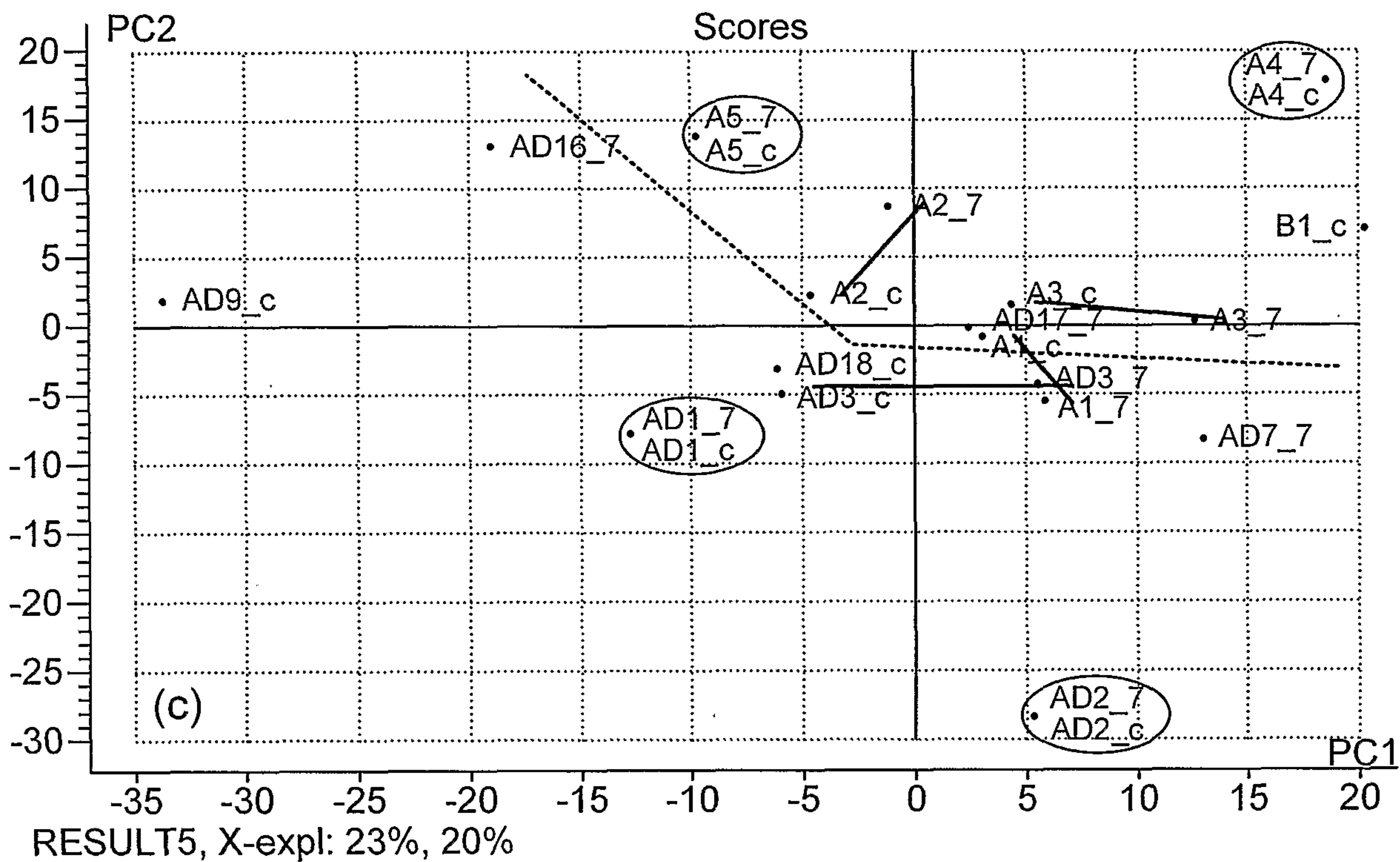
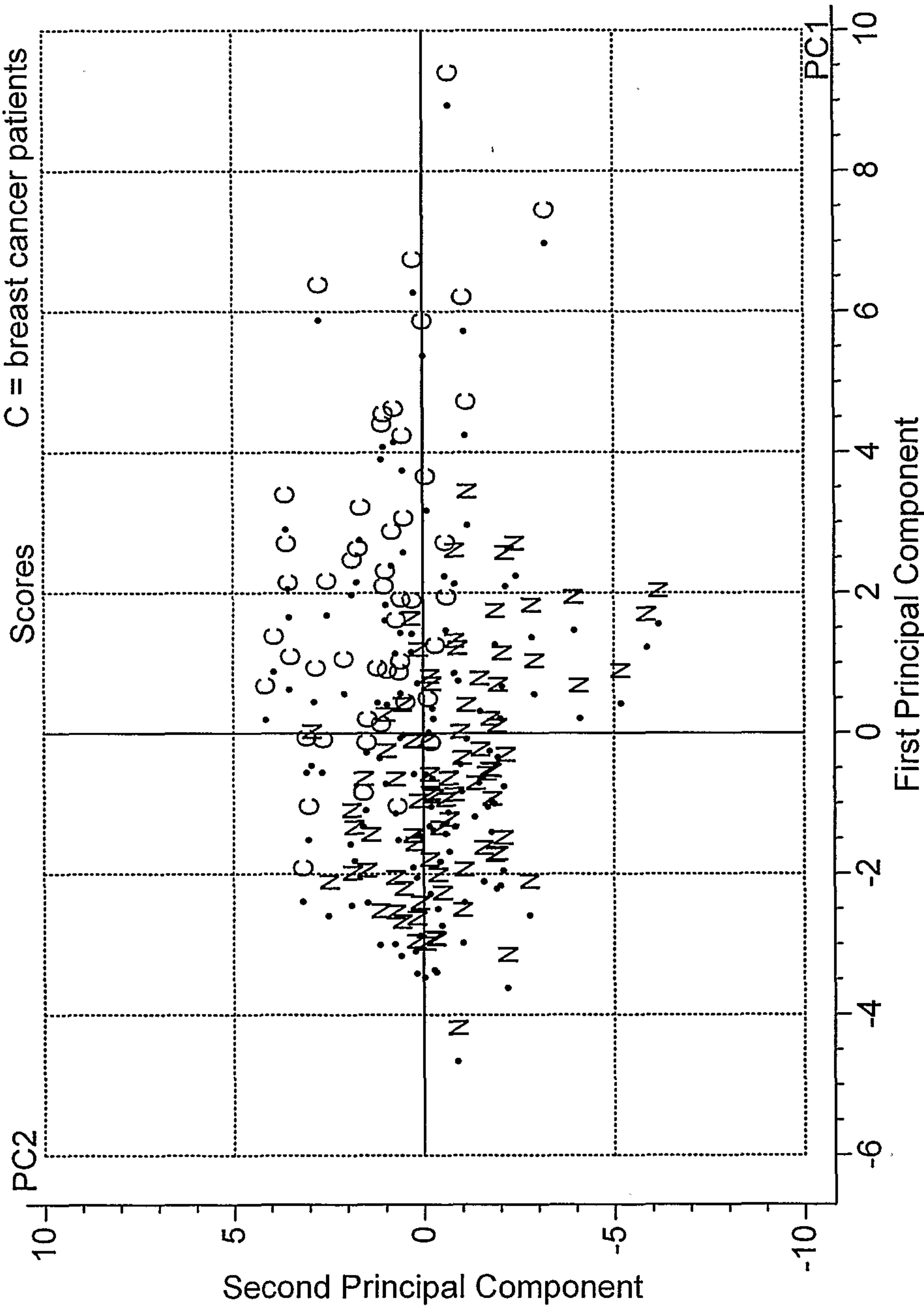


FIG. 2

Projection of normal (including benign) and breast cancer samples onto a classification model generated by PLSR-DA using the data of 44 informative genes

PC = Principal Components
N = normal
C = breast cancer patients



4 / 12

FIG. 3

Projection of individuals with and without Alzheimer's disease onto a
classification model generated by PLSR-DA using 182 informative genes.

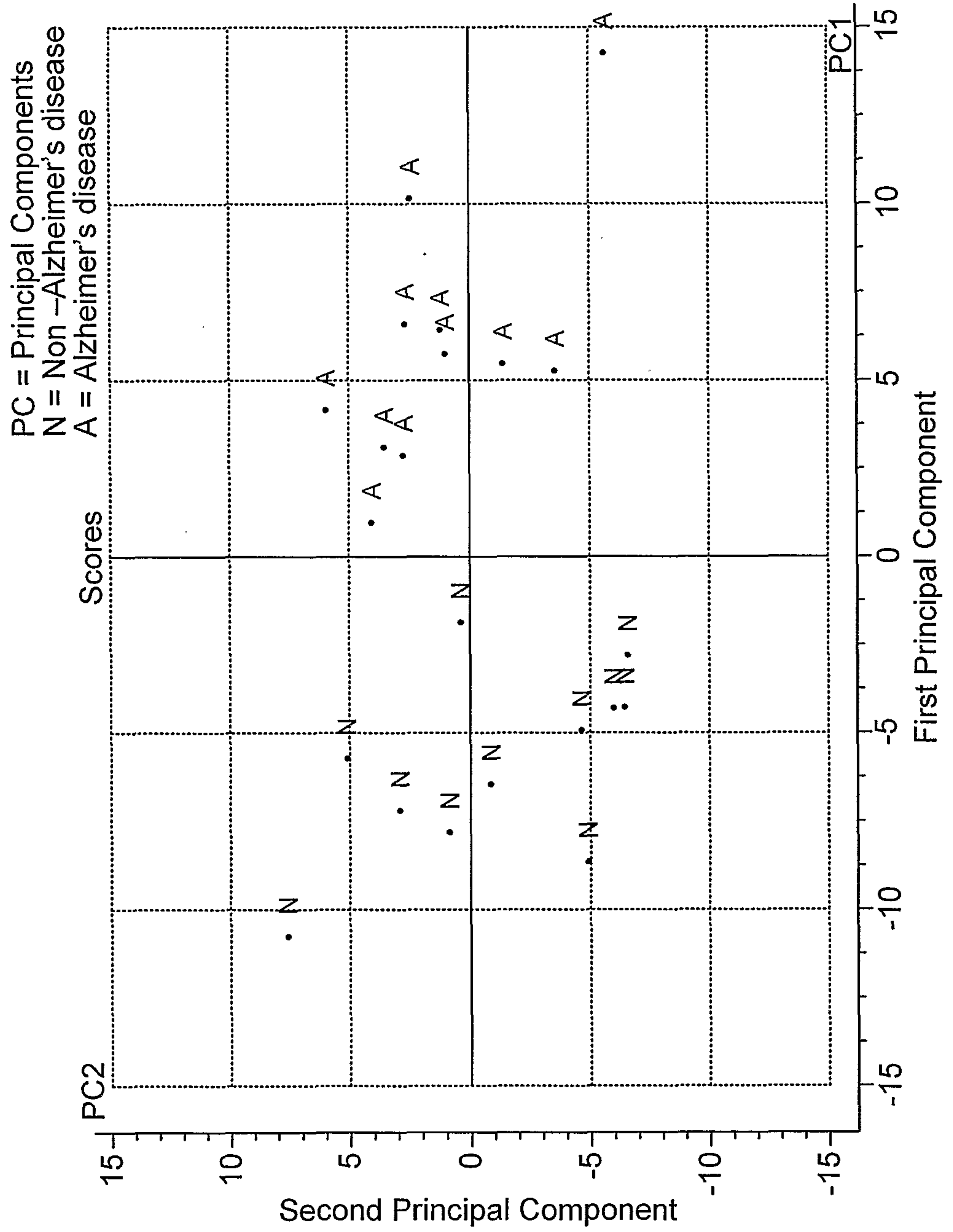


FIG. 4

Projection of normal (women with suspected mammogram but no breast cancer) and breast cancer samples onto a classification model generated by PLSR-DA using the data of 719 cDNAs

PC = Principal Components
N = normal
B = breast cancer patients

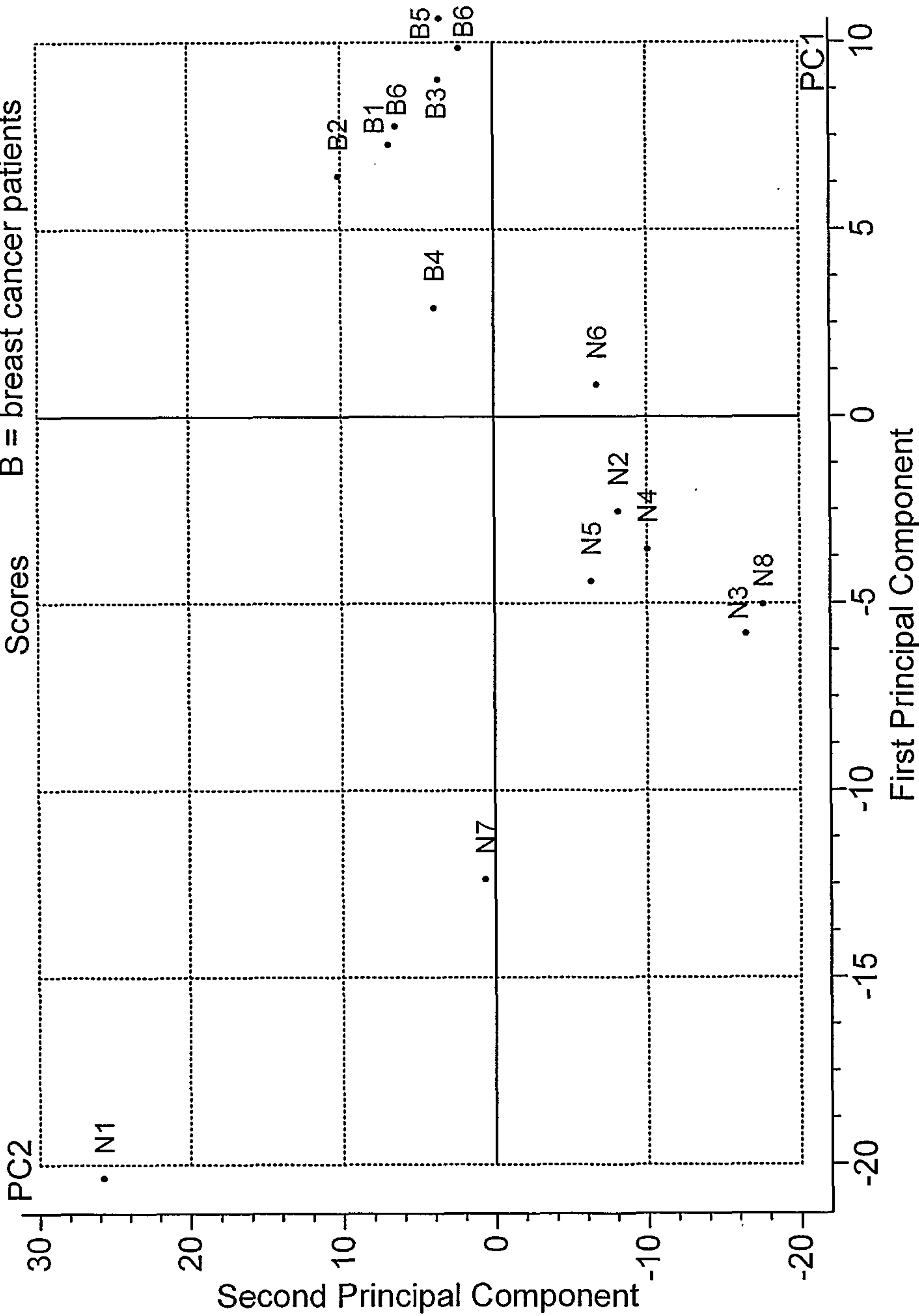


FIG. 5
Prediction of normal (women with suspected mammogram but no breast cancer) and breast cancer samples based on 3 principal components using the data of 719 cDNAs

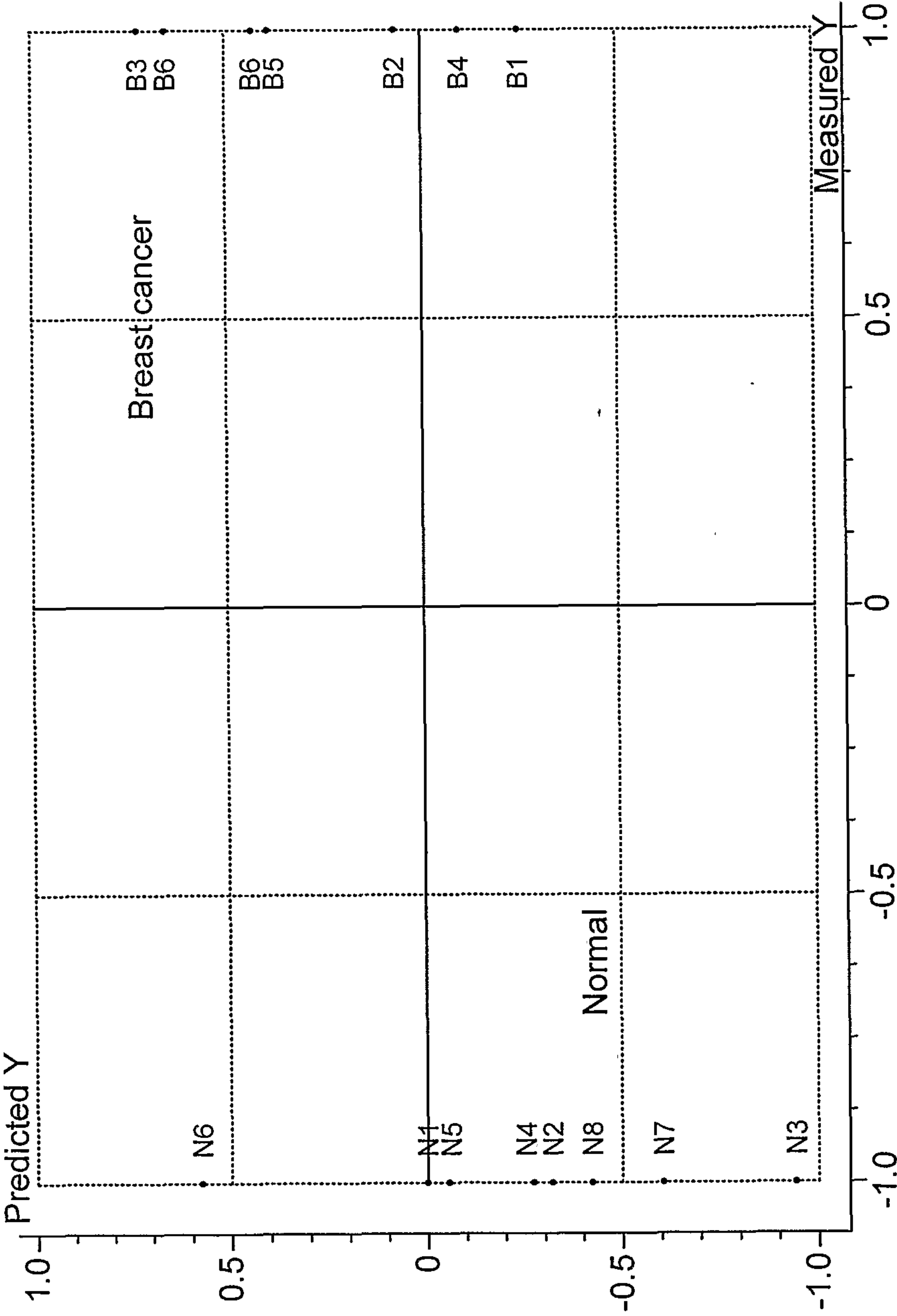


FIG. 6

Projection of normal (women with suspected mammogram but no breast cancer) and breast cancer samples onto a classification model generated by PLSR-DA using the data of 111 cDNAs already on file

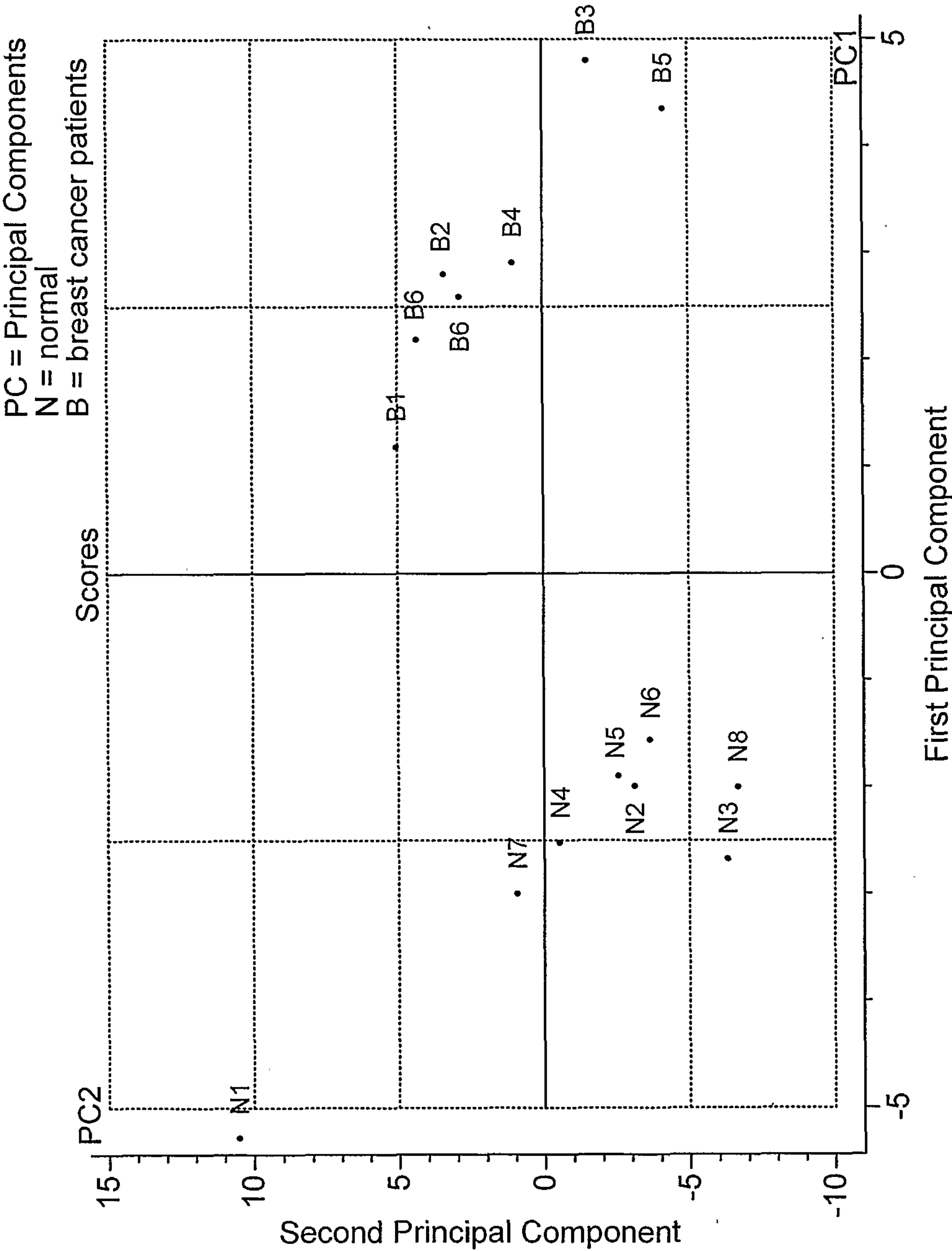


FIG. 7
Prediction of normal (women with suspected mammogram but no breast cancer) and breast cancer samples based on 4 principal components using the data of 111 cDNAs already on file

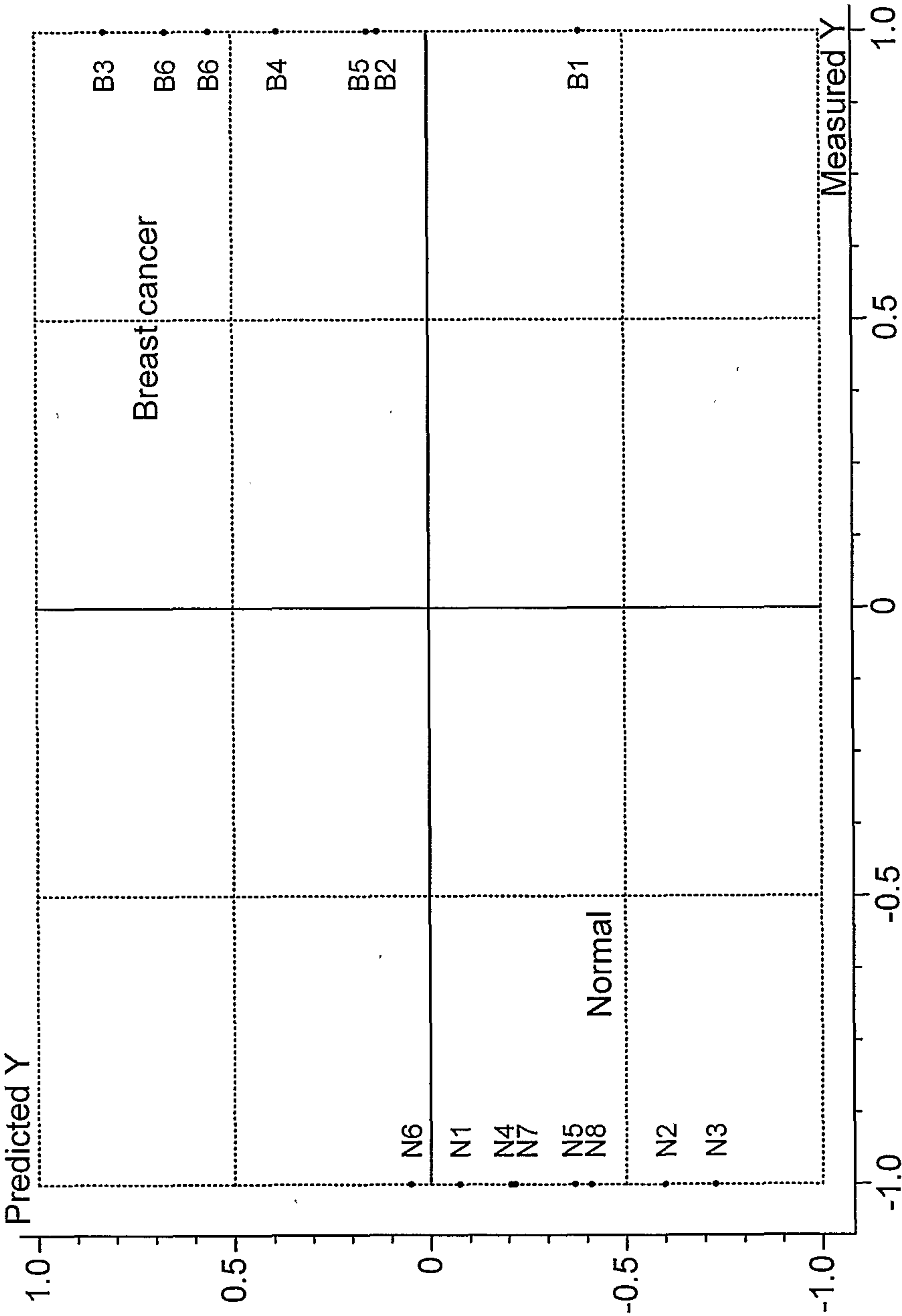


FIG. 8

Projection of normal (women with suspected mammogram but no breast cancer) and breast cancer samples onto a classification model generated by PLSR-DA using the data of 345 sequenced cDNAs for breast cancer

PC = Principal Components
N = normal
B = breast cancer patients

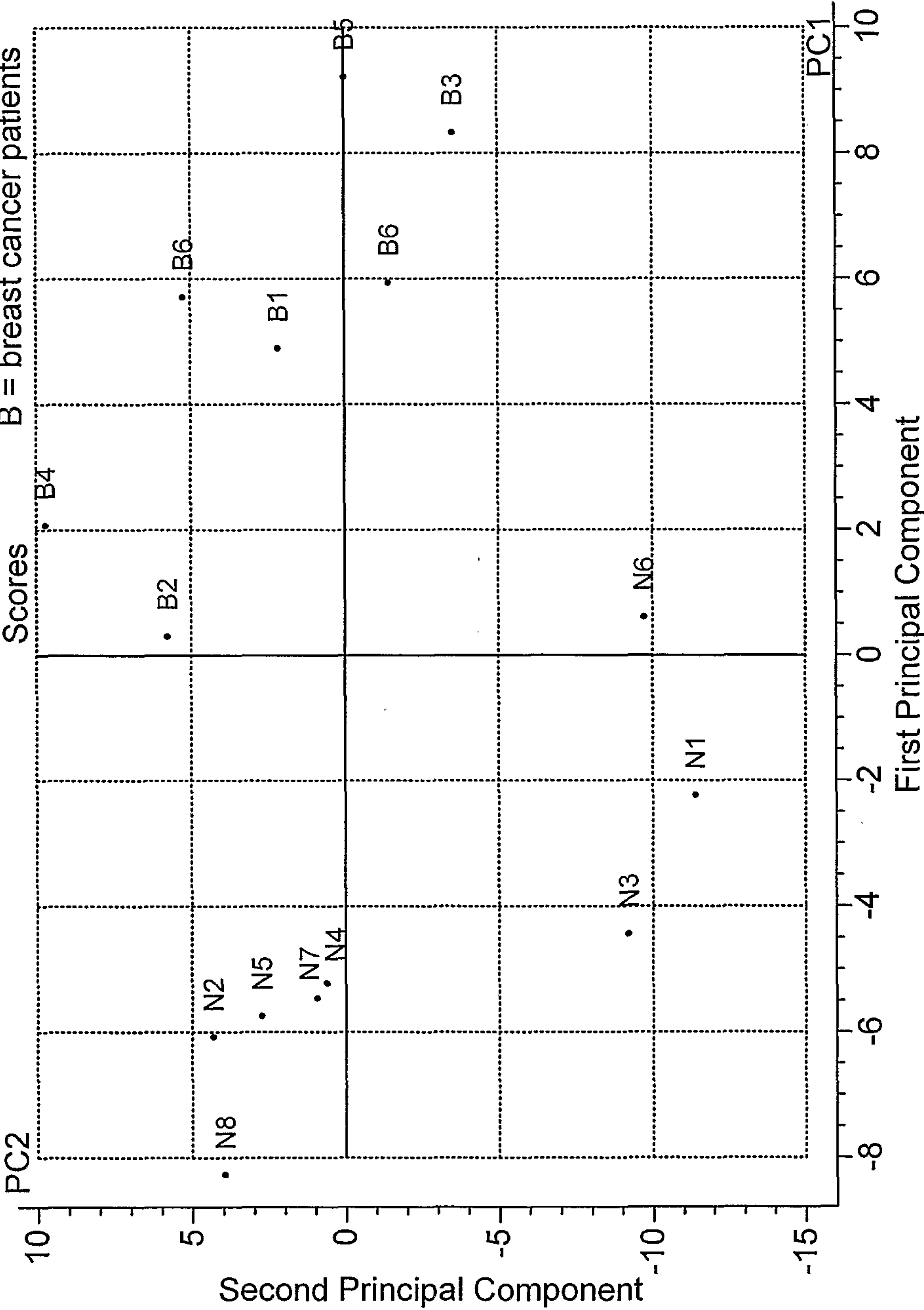


FIG. 9
Prediction of normal (women with suspected mammogram but no breast cancer) and breast cancer samples based on 3 principal components using the data of 345 sequenced cDNAs

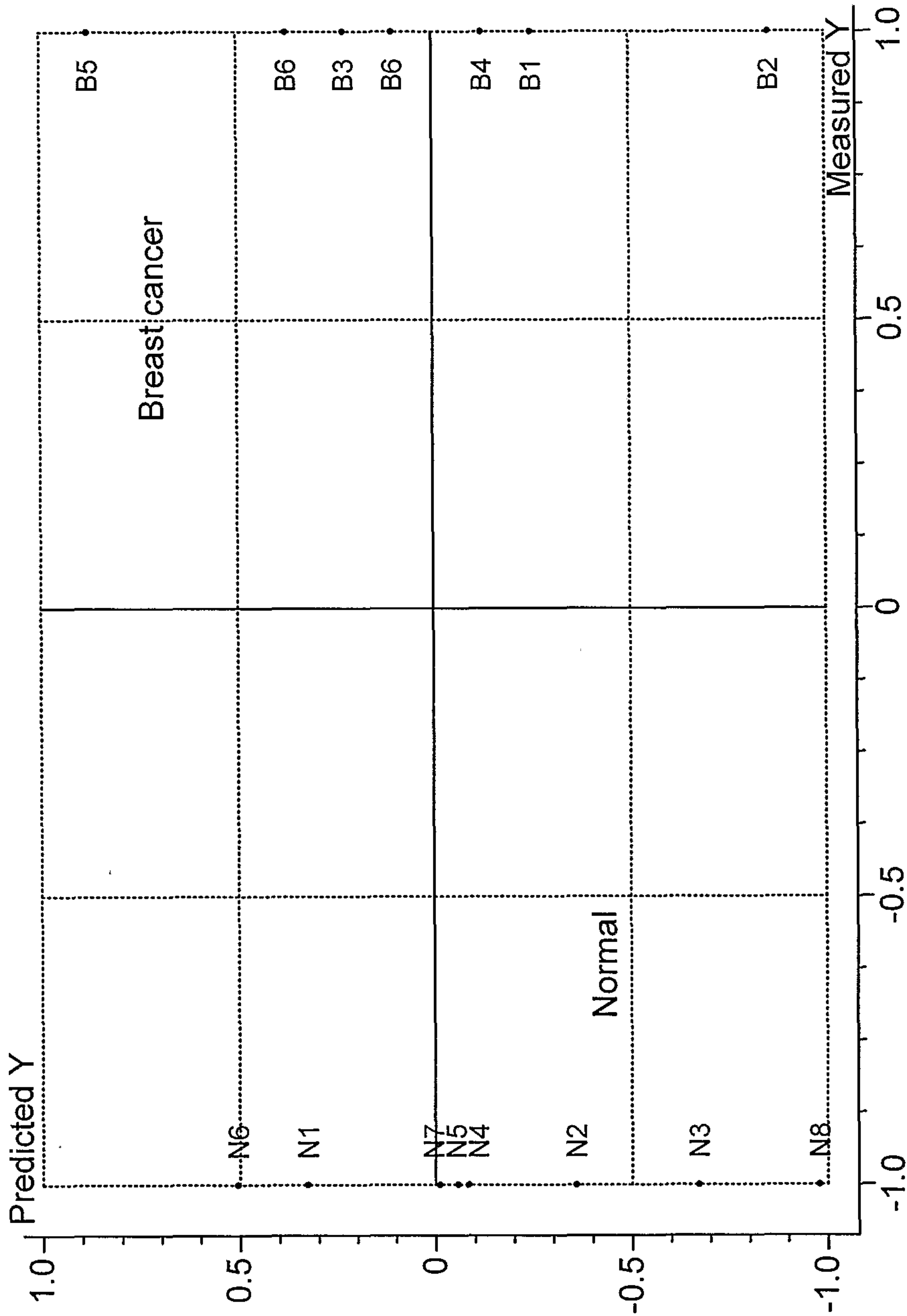


FIG. 10

Projection of non-alzheimer and alzheimer samples onto a classification model generated by PLSR-DA using the data of 520 sequenced cDNAs

PC = Principal Components
A= non-alzheimer
AD = alzheimer patients

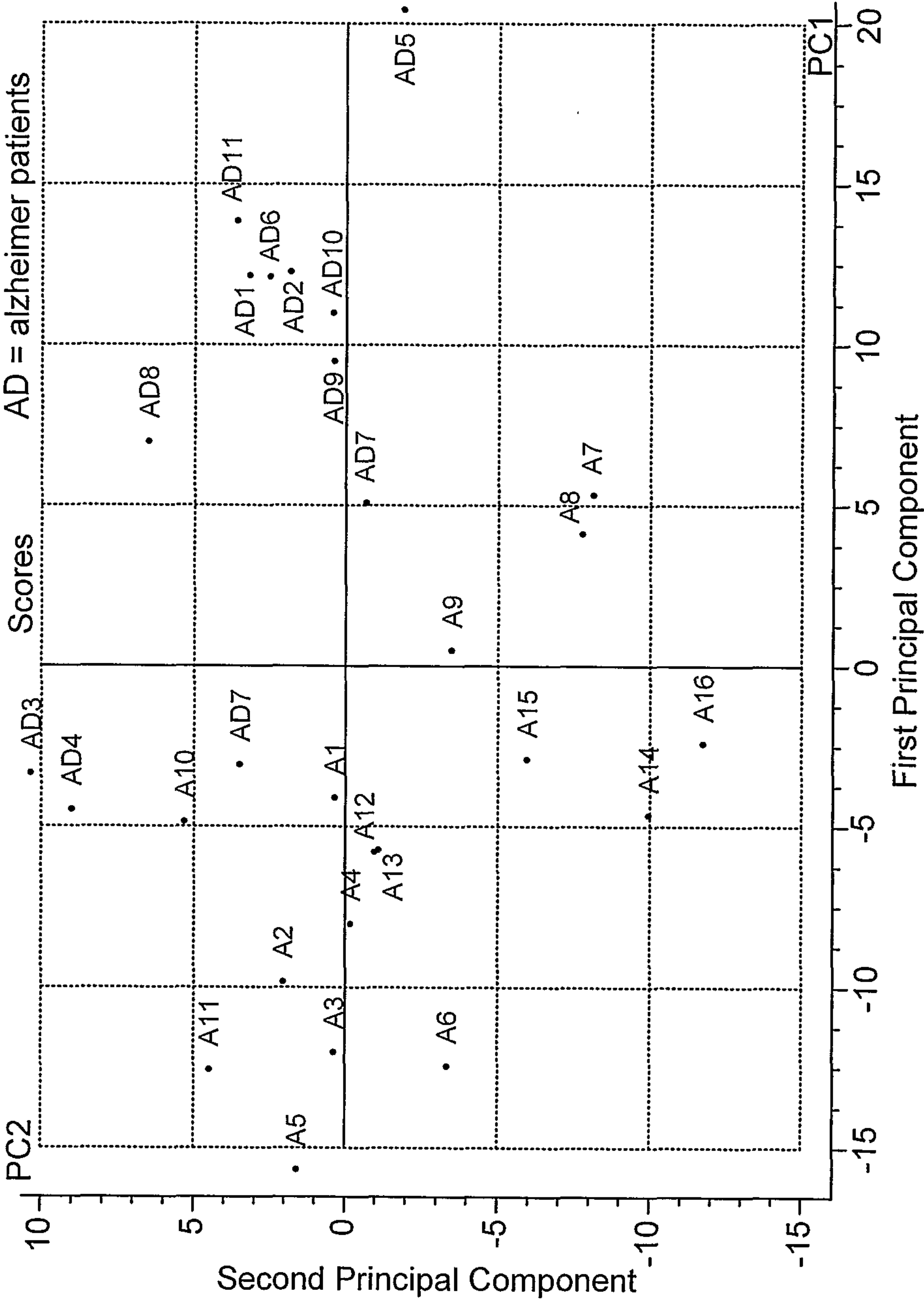


FIG. 11
Prediction of non-alzheimer and alzheimer samples based on 4
principal components using the data of sequenced cDNAs

