

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
15 February 2007 (15.02.2007)

(10) International Publication Number
WO 2007/019444 A2

(51) International Patent Classification:
C12N 15/09 (2006.01)

(74) **Agents: KILYK, JR, John** et al.; Leydig, Voit & Mayer, Ltd, 180 North Stetson Avenue, Two Prudential Plaza, Suite 4900, Chicago, Illinois 60601-6780 (US).

(21) International Application Number:

PCT/US2006/030721

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

(22) International Filing Date: 4 August 2006 (04.08.2006)

(25) Filing Language: English

(26) **Publication Language:** English

(30) Priority Data:
60/705,964 5 August 2005 (05.08.2005) US

(71) Applicant (for all designated States except US): EUCLID DIAGNOSTICS LLC [US/US]; Chicago Technology Park, 2201 West Campbell Park Drive, Chicago, Illinois 60612 (US).

(72) Inventors; and

(75) **Inventors/Applicants (for US only): FREIJE, Wadiha** [LB/US]; 117 Elgin Avenue, Forest Park, IL 60130 (US). **BRIKUN, Igor** [US/US]; 117 Elgin Avenue, Forest Park, Illinois 60130 (US).

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

[Continued on next page]

(54) Title: SUBTRACTIVE SEPARATION AND AMPLIFICATION OF NON-RIBOSOMAL TRANSCRIBED RNA (NRRNA)

28S rRNA sequence:

[illegible]

(57) Abstract: The invention provides a method of separating non-ribosomal transcribed RNA (nrRNA) fragments from ribosomal RNA (rRNA) and rRNA fragments. The method comprises (i) providing a sample comprising rRNA, rRNA fragments, and nrRNA fragments, and (ii) providing a plurality of probes. The probes hybridize to RNA targeting sequences of at least 50% of the contiguous regions of the rRNA and to rRNA fragments comprising the rRNA targeting sequences. The method further comprises (iii) adding the plurality of probes to the sample, (iv) hybridizing the probes to the rRNA and rRNA fragments to form rRNA-probe complexes and rRNA fragment-probe complexes, and (v) separating the rRNA-probe complexes and rRNA fragment-probe complexes. The invention also provides a method of amplifying an nrRNA fragment, a method of analyzing nrRNA expression, a method of determining the level of nrRNA in a sample, and a kit and system useful in any of the foregoing methods.



Published:

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

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

SUBTRACTIVE SEPARATION AND AMPLIFICATION OF NON-RIBOSOMAL TRANSCRIBED RNA (nrRNA)

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This patent application claims the benefit of U.S. Provisional Patent Application No. 60/705,964, filed August 5, 2005.

TECHNICAL FIELD OF THE INVENTION

[0002] The invention relates to methods comprising separating messenger RNA (mRNA) and/or non-coding RNA (ncRNA), referred to collectively herein as non-ribosomal transcribed RNA (nrRNA), from a sample, such as a sample of total RNA. The invention also relates to kits and systems for use in such methods.

BACKGROUND OF THE INVENTION

[0003] Most mRNA isolation kits rely on the oligo dT-mediated purification of mRNA. This method is disadvantageous when mRNA is degraded and, hence, fragmented. The oligo dT only binds to the polyA tail of mRNA and, therefore, only purifies intact mRNA and molecules of mRNA that contain polyA tails. Molecules of mRNA that do not contain polyA tails are not purified. For this reason, when performing expression analysis using RNA isolated from fixed tissues, scientists have relied on the reverse transcription of total RNA using random hexamers even though over 95% of the RNA present in the sample is ribosomal (Godfrey et al., J. Molec. Diagnostics 2(2): 84-91 (2000)). This approach is disadvantageous because it is not sensitive and does not enable optimal detection of mRNA present in low copy numbers.

[0004] To overcome this limitation, scientists have used oligonucleotide primers designed to anneal to a plurality of genes for the reverse transcription reaction (see, for example, U.S. Patent Application Publication Nos. 2004/0259105 and 2005/0095634). Using this approach, only the expression level of a limited number of genes can be determined. Furthermore, only small portions of those genes, which are upstream of the primer site, can be analyzed, since the majority of the RNA fragments are less than 500 base pairs in length. In addition, if the fragments are very short, such as less than 50 base pairs, it may be extremely difficult to design an assay to determine expression levels, whether by polymerase chain reaction, ligation-mediated amplification, or microarray analysis, for example.

[0005] Murphy and Whitley (U.S. Patent Application Publication No. 2003/0175709 and International Patent Application Publication No. WO 03/054162) have proposed the use of a bridging nucleic acid, which includes at least one targeting region and at least one bridging region, and a capture nucleic acid, which includes a capture region and a non-reacting structure. The targeting region is complementary to a targeted region of a targeted nucleic acid to be removed. The bridging region is complementary to the capture region. The nonreacting structure is a compound that does not react with a nucleic acid. The bridging nucleic acid can include up to 10 or more targeting regions that are complementary to different, non-overlapping targeted regions from the same or different targeted nucleic acids. This method is disadvantageous in that the use of multiple targeting regions on a single bridging nucleic acid can decrease the efficiency of hybridization between the targeting regions and the complementary targeted regions. In addition, the use of multiple targeting regions on a single bridging nucleic acid can interfere with hybridization between the bridging region of the bridging nucleic acid and the capture region of the capture nucleic acid. These effects also can decrease efficiency of recovery of non-ribosomal RNA. The methods disclosed by Murphy and Whitley also result in increased background hybridization due to the length of the oligonucleotides used.

[0006] The invention seeks to overcome the disadvantages of the currently available isolation kits by providing a method that enables the purification of nrRNA fragments that do not contain polyA tails as well as intact mRNA and/or ncRNA. This and other objects and advantages of the invention, as well as additional inventive features, will become apparent from the detailed description provided herein.

BRIEF SUMMARY OF THE INVENTION

[0007] The invention provides a method of separating non-ribosomal transcribed RNA (nrRNA) fragments from ribosomal RNA (rRNA) and rRNA fragments. The method comprises (i) providing a sample comprising rRNA, rRNA fragments, and nrRNA fragments, wherein the rRNA comprises multiple contiguous regions of about 100 base pairs. Each contiguous region comprises an rRNA targeting sequence, and the rRNA fragments comprise one or more of the rRNA targeting sequences. The method further comprises (ii) providing a plurality of probes. The probes hybridize to (a) different RNA targeting sequences of at least 50% of the contiguous regions of the rRNA and (b) rRNA

fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the rRNA. The method further comprises (iii) adding the plurality of probes to the sample, (iv) hybridizing the probes to the rRNA and rRNA fragments to form rRNA-probe complexes and rRNA fragment-probe complexes, and (v) separating the rRNA-probe complexes and rRNA fragment-probe complexes from the sample, thereby separating nrRNA fragments from rRNA and rRNA fragments.

[0008] The invention further provides a method of amplifying a cDNA complementary to an nrRNA fragment. The method comprises (i) separating an nrRNA fragment from rRNA and rRNA fragments according to the inventive method, wherein the nrRNA fragment comprises a 5' end. The method also comprises (ii) ligating a first oligonucleotide to the 5' end of the nrRNA fragment, (iii) hybridizing to the nrRNA fragment a second oligonucleotide, (iv) extending a DNA strand from the second oligonucleotide to generate a first cDNA, and (v) amplifying the first cDNA, thereby amplifying a cDNA complementary to an nrRNA fragment.

[0009] In addition, the invention provides a method of amplifying nrRNA fragments. The method comprises (i) separating nrRNA fragments from rRNA and rRNA fragments according to the inventive method. The method further comprises (ii) amplifying the nrRNA fragments. Alternatively, the method further comprises (ii) ligating a first oligonucleotide to the 5' end of the nrRNA fragment; (iii) hybridizing to the nrRNA fragment a second oligonucleotide; and (iv) extending a DNA strand from the second oligonucleotide to generate a first cDNA having a 5' end and a 3' end. The method then comprises (v) degrading the nrRNA fragment and, optionally, removing any free oligonucleotides and/or free RNA fragments; (vi) synthesizing a complementary strand to the first cDNA; (vii) optionally purifying the double-stranded DNA copy; and (viii) transcribing RNA from the double-stranded DNA copy. Synthesizing a complementary strand to the first cDNA in step (vi) comprises (a) contacting the first cDNA with a third oligonucleotide such that the third oligonucleotide hybridizes to the 3' end of the first cDNA and (b) extending the complementary strand from the third oligonucleotide, whereupon a double-stranded DNA copy of the nrRNA fragment is generated. The third oligonucleotide hybridizes to at least a portion of the first oligonucleotide and optionally contains an RNA polymerase promoter sequence.

[0010] Still further provided by the invention is a method of analyzing nrRNA expression. The method comprises (i) separating nrRNA fragments from rRNA and rRNA fragments according to the inventive method and (ii) analyzing nrRNA expression. Additionally, the invention provides a method of determining the level of nrRNA in a sample. The method comprises (i) separating nrRNA fragments from rRNA and rRNA fragments according to the inventive method and (ii) labeling the separated nrRNA fragments with a detectable label. The method of determining the level of nrRNA in a sample further comprises (iii) providing an array comprising a collection of fixed DNAs that hybridize to a plurality of target sequences located within an expression sequence of a gene of interest. The fixed DNAs comprise from about 15 nucleotides to about 750 nucleotides, and at least one target sequence is interspersed approximately every 500 base pairs of the expression sequence. The method additionally comprises (iv) applying the labeled nrRNA fragments to the array under hybridization conditions; (v) optionally removing unbound and nonspecifically bound nrRNA fragments from the array; (vi) detecting the labeled nrRNA fragments that have specifically annealed to the fixed DNAs in the array; and (vii) quantitating the level of labeled nrRNA fragments detected in step (vi) to determine the level of nrRNA in the sample.

[0011] The invention also provides a kit for separating nrRNA fragments from a sample comprising rRNA and rRNA fragments. The kit comprises a plurality of probes that hybridize to (a) different rRNA target sequences of at least 50% of contiguous regions of about 100 base pairs of the rRNA, which contiguous regions comprise the rRNA targeting sequences and (b) rRNA fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the rRNA. Preferably, the kit also comprises instructions.

[0012] Likewise, the invention provides a system for separating nrRNA fragments from a sample comprising rRNA and rRNA fragments. The system comprises a device adapted for the separation of nucleic acids by hybridization. The system also comprises a plurality of probes that hybridize to (a) different rRNA target sequences of at least 50% of contiguous regions of about 100 base pairs of the rRNA, which contiguous regions comprise the rRNA targeting sequences, and (b) rRNA fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the rRNA. The

plurality of probes is positioned within the device to separate the rRNA and rRNA fragments from nrRNA.

BRIEF DESCRIPTION OF THE FIGURES

[0013] Figures 1A - 1D show the nucleotide sequences of the 28S (SEQ ID NO: 1), 18S (SEQ ID NO: 2), 5.8S (SEQ ID NO: 3), 5S (SEQ ID NO: 4), 16S (SEQ ID NO: 5), and 12S (SEQ ID NO: 6) rRNAs. Sequences are presented from left to right, top to bottom, in accordance with convention.

[0014] Figures 2A - 2D show a set of oligonucleotides (SEQ ID NOs: 7-110) that can hybridize to rRNA and are suitable for use in the method. The "Start" positions are given relative to the sequences of Figures 1A - 1D.

[0015] Figures 3A - 3C show a set of forward and reverse primers (SEQ ID NOs: 111-165 and 191-245) that can amplify fragments corresponding to six rRNA species. The "Left" and "Right" designations indicate forward and reverse primers, respectively.

[0016] Figures 4A - 4M compare 28S rRNA sequences and provide the consensus sequence (SEQ ID NOs: 166-184).

[0017] Figure 5 is a diagram illustrating the preparation of double-stranded cDNA for amplification.

DETAILED DESCRIPTION OF THE INVENTION

[0018] The invention is predicated on the discovery that the use of an exhaustive set of probes enables the efficient removal of ribosomal RNA (rRNA) (i.e., intact rRNA and/or rRNA fragments) from a sample of total RNA. In doing so, the invention enables the separation and/or purification and/or enrichment of non-ribosomal transcribed RNA (nrRNA) fragments that do not contain polyA tails, such as are commonly found in degraded RNA. The invention does not suffer from the disadvantages attendant the use of currently available methods.

[0019] The invention provides a method of separating nrRNA, such as messenger RNA (mRNA) and non-coding RNA (ncRNA), from a sample, such as a sample comprising total RNA. The inventive method can be performed on any sample suitable for separation of RNA species. Indeed, any suitable sample of, for example, total RNA can be used. Preferably, degradation of RNA in the sample has been minimized. However, the invention is advantageous in that it can be used on degraded RNA, such as RNA recovered from

preserved tissue samples, e.g., paraffin-embedded tissue samples, biopsies, surgical specimens, and any other source containing degraded RNA. Such RNA is usually less than 500 base pairs in size and may not represent the entire population of mRNA species present in original tissue samples. The invention also permits the recovery of ncRNA species that are present in a total RNA sample, such as small interfering RNA (siRNA) and small nuclear RNA (snRNA). Researchers have great difficulty in efficiently isolating such small RNA species using currently known methods.

[0020] The method comprises (i) removing rRNA from the sample of total RNA by hybridizing the rRNA with a plurality of probes comprising at least one probe that can hybridize to a complementary sequence within about 100 contiguous base pairs of the rRNA to be removed such that at least one probe hybridizes to a complementary sequence within at least about every 100 contiguous base pairs of the rRNA to be removed, and (ii) purifying the nrRNA.

[0021] Alternatively, the invention provides a method of separating nrRNA fragments from rRNA and rRNA fragments. The method comprises: (i) providing a sample comprising rRNA, rRNA fragments, and nrRNA fragments; (ii) providing a plurality of probes; (iii) adding the plurality of probes to the sample; (iv) hybridizing the probes to the rRNA and rRNA fragments to form rRNA-probe complexes and rRNA fragment-probe complexes, and (v) separating the rRNA-probe complexes and rRNA fragment-probe complexes from the sample (i.e., separating the hybridized RNA from the sample of RNA). Optionally, simultaneously with or subsequently to removing rRNA from the sample (e.g., the sample of total RNA), the method can further comprise removing any other RNA that is not of interest, such as unwanted mRNA species that are abundantly expressed or tRNA(s). It will be understood that "separating" the rRNA-probe complexes and/or rRNA fragment-probe complexes (or any other RNA) from the sample does not require complete elimination of rRNA or rRNA fragments from the sample. If desired, additional rounds of hybridization in accordance with the inventive method can be performed to further remove unwanted RNA. For example, the method can be performed on a sample of total RNA to achieve a sample comprising no more than 20% rRNA. Preferably, the method yields a sample or composition comprising no more than 10% rRNA, and more preferably the resulting composition comprises no more than 5% rRNA. Most preferably, the composition resulting from the inventive method comprises no more than 1% rRNA. The inventive

method preferably further comprises (vi) recovering and purifying the nrRNA fragments. In view of the above, the method is particularly suited for enriching a composition for nrRNA and fragments thereof by removing unwanted RNAs, such as rRNA and rRNA fragments.

[0022] The invention comprises employing a plurality of probes to saturate unwanted rRNA in a sample, thereby providing multiple "handles" for removing the unwanted material using any suitable separation technique. In this regard, rRNA and rRNA fragments can be conceptually divided into multiple contiguous regions of about 100 base pairs. For example, an rRNA or rRNA fragment that comprises 1000 base pairs comprises ten contiguous regions of about 100 base pairs. Each contiguous region comprises one or more RNA targeting sequences. Preferably, two or more of the multiple contiguous regions comprise different rRNA targeting sequences. An "rRNA targeting sequence" refers to a nucleotide sequence within the rRNA or rRNA fragment that is complementary to, and thereby hybridizes with, one or more probes of the plurality. The rRNA fragments of the sample desirably comprise one or more of the rRNA targeting sequences.

[0023] A "plurality of probes" is a collection of molecules, such as oligonucleotides, that bind rRNA target sequences within the rRNA or rRNA fragments of the sample. Preferably, the probes hybridize to (a) different RNA targeting sequences of at least 50% of the contiguous regions of the rRNA and (b) rRNA fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the rRNA. In other words, if an rRNA comprises ten contiguous regions of about 100 base pairs, the probes hybridize to RNA targeting sequences in at least five of the ten contiguous regions. By "different" RNA targeting sequences is meant separate, individual RNA targeting sequences, which may or may not be unique. Desirably, multiple probes hybridize over the length of the unwanted RNA to ensure efficient removal. For example, the probes of the inventive method can hybridize to rRNA targeting sequences of at least 60% of the contiguous regions (e.g., the probes hybridize to rRNA targeting sequences of at least 70% of the contiguous regions). Even more preferably, the probes hybridize to rRNA targeting sequences of at least 80% of the contiguous regions (e.g., the probes hybridize to rRNA targeting sequences of at least 90% of the contiguous regions). Still more preferably, the probes hybridize to rRNA targeting sequences of at least 95% of the contiguous regions. Even more preferably, the probes hybridize to rRNA targeting sequences of all of the

contiguous regions (i.e., rRNA targeting sequences located in each and every contiguous region).

[0024] In this regard, rRNA and fragments thereof and, optionally, any other RNA that is not of interest, can be removed from a sample, e.g., a sample of total RNA, by hybridizing the RNA to a plurality of probes comprising, for example, at least one tagged oligonucleotide (i.e., DNA or RNA, including modified base analogs), tagged single-stranded nucleic acid molecule (i.e., DNA or RNA), or tagged DNA mimic, such as a locked nucleic acid (LNA; Koshkin et al., *Tetrahedron* 54: 3607-3630 (1998); Koshkin et al., *JACS* 120: 13252-13253 (1998); and Wahlestedt et al., *PNAS* 97: 5633-5638 (2000)) and/or a DNA analog with a non-phosphodiester backbone (such as a peptide nucleic acid (PNA; Nielson, *Molec. Biotech.* 26: 233-248 (2004); and Egholm et al., *Nature* 365(6446): 566-568 (1993)) or morpholino derivatives (Summerton, *BioChim. Biophys. Acta* 1489:141-158 (1999)), which is complementary to the RNA that is to be removed.

[0025] The rRNA and the rRNA fragments in the sample can comprise eukaryotic rRNA and eukaryotic rRNA fragments, although a sample also can comprise bacterial RNA species. In addition, the probes, including any tagged oligonucleotides, tagged single-stranded nucleic acid molecules, and/or DNA mimics, can be complementary to 18S, 28S, 5S, 5.8S, 12S, or 16S rRNA (see Figures 1A - 1D and 2A - 2D). The sequences of different rRNA species are highly conserved; however, the probes (e.g., oligonucleotides or molecules complementary to the rRNA species) can contain sequences complementary to polymorphic regions within the rRNA sequences to maximize removal of rRNA species from the sample (see, e.g., Figures 4A - 4M for examples of such polymorphic regions). Such probes (e.g., oligonucleotides and/or molecules complementary to the rRNA species) can be used in various combinations. The rRNA and rRNA fragments can comprise at least two rRNAs selected from the group consisting of 28S, 18S, 5.8S, 5S, 12S, and 16S rRNAs. Thus, the plurality of probes can comprise one that is complementary to 18S rRNA and another that is complementary to 28S rRNA. Under certain circumstances, such as when an nrRNA of interest is present in low copy number, it can be desirable to remove unwanted RNA fragments that are abundant in the sample. It may be desirable to use an additional oligonucleotide and/or single-stranded nucleic acid molecule probe in addition to probes that are complementary to 18S and 28S, such as one or more or even all of 12S, 16S, 5.8S, and 5S rRNAs, alone or in further combination with probes (such as an oligonucleotide, a

single-stranded nucleic molecule and/or a DNA mimic) that are complementary to tRNA and/or that are complementary to any mRNA that may be abundantly expressed in the RNA sample under investigation and that is not of interest.

[0026] Probes (e.g., oligonucleotide probes) can be synthesized using any *in vitro* chemical synthesis known in the art. Oligonucleotide probes can be as short as around 12 nucleotides or as long as around 150 nucleotides, such as around 15 or 17 nucleotides up to around 100 nucleotides, whereas a single-stranded nucleic acid molecule probe can be as short as 75 nucleotides up to the full-length of the RNA molecule that it complements. Oligonucleotides can be selected using a Primer selection program (e.g., Vector NTI (Invitrogen, Carlsbad CA) or Primer 3 (Rozen et al., "Primer 3 on the WWW for general users and for biologist programmers." In: Krawetz and Misener, eds. Bioinformatics Methods and Protocols, Methods in Molecular Biology. Humana Press, Totowa, NJ (2000), pp. 365-386)) using suitable parameters for primer length, base composition, and melting temperatures. It is preferred that the oligonucleotide probe have a GC content of about 35-80%, and a melting temperature of about 50-75°C. The optimal oligonucleotide length and melting temperature for the probes can be determined by the composition of the buffers used for the hybridization. Primer design programs select oligonucleotides based on the user's input of relevant parameters, including minimizing secondary structures, palindromic sequences, and hairpin loops. The oligonucleotides can be further analyzed by sequence homology search programs (e.g., BLAST) to minimize cross-hybridization to other genomic sequences, especially homology to expressed transcripts. It is preferred that the set of oligonucleotides used as probes for the hybridization contain at least one oligonucleotide complementary to each fragment (i.e., contiguous region) of the rRNA, wherein each fragment (i.e., contiguous region) is from about 100 base pairs to about 250 base pairs.

[0027] Single-stranded nucleic acid molecules can be synthesized *in vitro* (i.e., PCR or transcription-coupled amplification) or *in vivo* (i.e., replicated in living cells using recombinant DNA vectors). Such molecules can be synthesized as fragments or full-length molecules, and can be employed as probes in the context of the invention.

[0028] The probes of the invention can be tagged. By "tagged" is meant adding a molecule (a "tag") to the probe (e.g., oligonucleotide, single-stranded nucleic acid molecule, and/or DNA mimic) that will permit downstream manipulations, such as removal. For example, the tag can be biotin, which, upon binding to streptavidin, such as streptavidin

attached to a support (e.g., a magnetic bead or plastic surface), results in the removal of rRNA and any other RNA that is not of interest from the sample. Alternatively, the tag can be digoxigenin, which, upon binding to an immobilized anti-digoxigenin antibody, such as an antibody attached to a substrate (e.g., a magnetic bead or a microtiter plate), results in the removal of rRNA and any other RNA that is not of interest from the sample. Alternatively, the tag can be an amine linker that can be used to immobilize the oligonucleotide on a solid support.

[0029] The method of the invention comprises hybridizing the probes to the rRNA and rRNA fragments to form rRNA-probe complexes and rRNA fragment-probe complexes, which are then separated from the sample. Hybridization refers to the annealing of single stranded oligonucleotides to form double stranded oligonucleotides in an environment below the melting temperature of the double stranded oligonucleotide. Hybridization conditions are well known in the art and can be adapted to accommodate specific probes, rRNA sequences, and desired level of specificity of unwanted RNA removal. In this regard, "stringency" of hybridization reactions is determined by probe length, homology between the sequences, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. In addition, the higher the degree of homology between the probe and rRNA or rRNA fragment, the higher the relative temperature that can be used. Hybridization reactions are further described in, for example, Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, (1995), and U.S. Patent No. 7,081,340.

[0030] The hybridization conditions can include, but are not limited to, "stringent conditions" or "high stringency conditions" which typically: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ, during hybridization, a denaturing agent, such as formamide (e.g., 50% (v/v) formamide), 50 mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5X SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, optionally followed by a high-stringency wash consisting of 0.1X SSC containing EDTA at 55°C.

[0031] "Moderately stringent conditions" are described in, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Press, 1989.

Moderately stringent conditions can be achieved using a washing solution and hybridization conditions (e.g., temperature and ionic strength) less stringent than those described above. For example, moderately stringent conditions can comprise an overnight incubation at 37°C in a solution comprising: 20% formamide, 5X SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6). The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like. DNA recovered in the remaining wash buffer can be added to the original hybridization mix, and hybridization can be repeated to further reduce the amount of rRNA in the sample. If desired, the rRNA of the sample is discarded, and the nrRNA is recovered in the hybridization buffer. In this regard, the hybridization step of the inventive method can be repeated to reduce the concentration of unwanted RNA, such as unwanted rRNA or rRNA fragments, in the sample. The hybridization step can be repeated by administering fresh probes to the reaction to form additional rRNA-probe complexes and rRNA fragment-probe complexes with the residual rRNA and rRNA fragments that may remain in the sample. Further reducing the concentration of rRNA and rRNA fragments further enriches the composition with desired nrRNA.

[0032] In view of the above, the invention further provides a composition comprising mRNA, which has been separated from a sample of total RNA and which includes N-terminal and/or internal molecules of mRNA. "N-terminal molecule of mRNA" means a molecule of mRNA that includes the transcription start site. "Internal molecule of mRNA" means a molecule of mRNA that does not include the transcription start site or the polyA tail. The mRNA can be amplified.

[0033] Also provided by the invention is an improved method of amplifying nrRNA fragments (such as mRNA or fragments thereof). The method comprises (i) separating nrRNA fragments from rRNA and rRNA fragments according to the inventive method described herein, and (ii) amplifying the nrRNA fragments. The improvement provided by the invention comprises using either of (i) nrRNA that has been separated from a sample of total RNA in accordance with the above-described method or (ii) the above-described composition. Although the inventive method enriches a sample for desired nrRNA, the composition can be further enriched with desired nrRNA by amplifying the nrRNA fragments in the resulting composition. The amplification of RNA, and oligonucleotides in general, has been the subject of numerous patents (see, e.g., U.S. Patent Nos. 5,545,522;

5,716,785; 6,291,170; 6,642,034; 6,593,086; and 6,794,141, and U.S. Patent Application Publication No. 2005/0009101). Examples of methods that can be used to amplify mRNA include, but are not limited to, concatenation of RNA fragments, reverse transcription to generate the first cDNA strand, and strand displacement amplification; and attachment of an RNA polymerase promoter to the first or second strand cDNA followed by a transcription reaction to generate a multitude of copies of the starting cDNAs. A promoter is a sequence that can be recognized and bound by an RNA polymerase with subsequent transcription of the DNA fragment that is located downstream of the 3' end of the promoter. Examples of such promoters include, but are not limited to, the T3, T4, T7, SP6, and Q beta replicase promoters.

[0034] The invention further provides an improved method of performing expression analysis of nrRNA. The method of analyzing nrRNA expression comprises (i) separating nrRNA fragments from rRNA and rRNA fragments according to the inventive method described herein and (ii) analyzing nrRNA expression. Examples of methods that can be used to perform expression analysis include, but are not limited to, PCR, quantitative PCR, and microarrays (PCR protocols, Bartlett et al., eds. Humana Press, Totowa, NJ (2003); RT-PCR protocols, O'Connell, ed. Humana Press (2002); Gene cloning and analysis by RT-PCR, Siebert et al., eds. Eaton Publishing, Natick, MA (1998); A-Z quantitative PCR, Bustin, ed. IUL press, LaJolla, CA (2004); DNA microarrays: Gene expression applications, Jordan, ed. Springer-Verlag, New York, NY (2001); Microarrays Methods and Applications: Nuts and Bolts, Hardiman, ed. DNA Press, Eagleville, PA (2003); and Guide to analysis of microarray data, 2nd ed., Knudsen, ed. Wiley & Sons, Indianapolis, IN (2004)). The improvement provided by the invention comprises using either of (i) mRNA that has been separated from a sample of total RNA in accordance with the above-described method or (ii) the above-described composition.

[0035] A kit is also provided by the invention. The kit comprises instructions and at least one reagent for separating nrRNA from a sample of total RNA by (i) removing rRNA from the sample of total RNA, and, optionally, simultaneously or subsequently removing any other RNA that is not of interest, and (ii) purifying the nrRNA. The at least one reagent can comprise at least one oligonucleotide, single-stranded nucleic acid molecule, or DNA mimic that is complementary to (i) an rRNA selected from the group consisting of 18S, 28S, 12S, 16S, 5.8S and 5S and/or (ii) a tRNA.

[0036] In addition, the invention provides a kit for separating nrRNA fragments from a sample comprising rRNA and rRNA fragments (e.g., eukaryotic rRNA and eukaryotic rRNA fragments). Preferably, the rRNA and rRNA fragments comprise 28S, 18S, 5.8S, 5S, 16S, and 12S rRNAs or fragments thereof. The inventive kit comprises a plurality of probes that hybridize to different rRNA target sequences of at least 50% of contiguous regions of about 100 base pairs of the rRNA, which contiguous regions comprise the rRNA targeting sequences. The plurality of probes also hybridize to rRNA fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the rRNA. The kit also, optionally, further comprises instructions.

[0037] To increase specificity and efficiency of removal of unwanted rRNA from the sample, the probes of the kit preferably hybridize to rRNA targeting sequences of at least 60% of the contiguous regions (more preferably at least 70% of the contiguous regions). More preferably, the probes hybridize to rRNA targeting sequences of at least 80% of the contiguous regions (e.g., at least 90% of the contiguous regions). Even more preferably, the probes hybridize to rRNA targeting sequences of at least 95% of the contiguous regions (e.g., the probes hybridize to rRNA targeting sequences of all of the contiguous regions).

[0038] Also provided is a method of amplifying nrRNA fragments, such as those present in degraded RNA. The method comprises (i) separating an nrRNA fragment from rRNA and rRNA fragments according to the inventive method. The method further comprises (ii) ligating a first oligonucleotide to the 5' end of nrRNA fragments, which have been separated from a sample (e.g., a sample of total RNA) in accordance with the above method. The first oligonucleotide (a) can comprise a first unique sequence, (b) preferably is of sufficient length to hybridize to a complementary oligonucleotide under annealing (i.e., hybridization) conditions as are known in the art, and (c) optionally contains an RNA polymerase promoter sequence. The first oligonucleotides serves as an anchor sequence at the 5' end of nrRNA fragments, thereby facilitating the subsequent analysis of the fragments, in particular short fragments, such as those of less than about 50 base pairs. Desirably, the first oligonucleotide does not contain a sequence that is present in the genome of the species from which the nrRNA sample is derived. Preferably, the first oligonucleotide is at least about 10 nucleotides in length, more preferably, at least about 25 nucleotides in length, and most preferably, at least about 35 nucleotides in length. Examples of RNA polymerase promoter sequences include, but are not limited to, T3, T7,

SP6, T4, and Q β replicase. If desired, non-ligated nucleotides can be removed from the reaction.

[0039] The method further comprises (iii) annealing (i.e., hybridizing) to the nrRNA fragment a second oligonucleotide. The second oligonucleotide preferably comprises a second unique sequence and a random sequence, wherein the random sequence is of sufficient length to hybridize to the nrRNA under annealing (i.e. hybridization) conditions as are understood in the art. Desirably, the unique sequence of the second oligonucleotide does not contain a sequence that is present in the genome of the species from which the nrRNA sample is derived. Optionally, the unique sequence of the second oligonucleotide comprises an RNA polymerase promoter. As such, it is possible to obtain a nucleotide chain comprising RNA polymerase promoters capable of directing transcription in opposite directions. The random sequence can be from about 6 to about 35 nucleotides in length, preferably from about 6 to about 12 nucleotides in length. The random sequence must anneal to the nrRNA to be amplified. One of ordinary skill in the art will appreciate that step (iii) involves heating the sample and then cooling the sample to allow the second oligonucleotide to anneal to the nrRNA (e.g., mRNA) and incubating the sample at the annealing temperature in accordance with methods known in the art.

[0040] The method still further comprises (iv) extending a DNA strand from the second oligonucleotide to generate a first cDNA having a 5' and a 3' end. The DNA strand can be extended using any suitable method such as, for example, reverse transcription. Several methods of reverse-transcription for obtaining cDNA are known in the art. If the sample of nrRNA is small, it may be desirable to amplify the cDNA. Therefore, if desired, the cDNA may be amplified using any suitable method, such as PCR. Then, the method comprises (v) degrading the nrRNA (i.e., the nrRNA fragment). The nrRNA can be degraded using one or more RNA-degrading enzymes, such as RNases. The enzymes will need to be removed prior to transcription, such as by phenol-chloroform extraction in accordance with methods known in the art. The method optionally further comprises removing any free oligonucleotides and/or RNA fragments in accordance with methods known in the art, such as size-exclusion column purification, preferential precipitation, and the like.

[0041] Afterwards, the method comprises (vi) synthesizing the complementary strand to the first cDNA by annealing a third oligonucleotide to the 3' of the first cDNA and extending a strand from the third oligonucleotide. In other words, a complementary strand

to the first cDNA is synthesized by (a) contacting the first cDNA with a third oligonucleotide such that the third oligonucleotide hybridizes to the 3' end of the first cDNA, and (b) extending the complementary strand from the third oligonucleotide, whereupon a double-stranded DNA copy of the nrRNA fragment is generated. The third oligonucleotide is of sufficient length to hybridize to at least a portion of the first oligonucleotide under annealing conditions and optionally contains an RNA polymerase promoter sequence as previously described. The third oligonucleotide can comprise a unique sequence that is not present in the genome of the species from which the nrRNA sample was extracted. If the free oligonucleotides and/or RNA fragments are not removed in step (v), then step (vi) should be conducted at an elevated annealing temperature to prevent the second oligonucleotide and any RNA fragments from annealing to the first cDNA, in which case a thermostable DNA polymerase, preferably with proofreading capabilities, is used for synthesis of the complementary strand to the first cDNA.

[0042] The method optionally comprises (vii) purifying the double-stranded DNA copy of the nrRNA (i.e., the nrRNA fragment) in accordance with methods known in the art (e.g., binding to glass milk or ion-exchange resin, proteinase K digestion and phenol-chloroform extraction, etc.). In addition, the double-stranded DNA copy can be amplified using, for instance, PCR. The method further comprises (viii) transcribing RNA from the double-stranded DNA in accordance with methods known in the art, such as with commercially available kits (e.g., such as those available from Ambion (Austin, TX) or Epicentre (Madison, WI)). The purified double-stranded copy of the nrRNA carries two RNA polymerase promoters that can be used to generate RNA copies of the coding or the noncoding strands.

[0043] In addition, the invention provides a method of amplifying a cDNA complementary to an nrRNA fragment. The method comprises (i) separating an nrRNA fragment from rRNA and rRNA fragments according to the invention. The method further comprises (ii) ligating a first oligonucleotide, such as the first oligonucleotide described above, to the 5' end of the nrRNA fragment. The method then comprises (iii) hybridizing to the nrRNA fragment a second oligonucleotide and (iv) extending a DNA strand from the second oligonucleotide to generate a first cDNA. The first cDNA is then (v) amplified, thereby amplifying a cDNA complementary to an nrRNA fragment.

[0044] In view of the above, the invention provides a kit comprising instructions and at least one reagent for generating a double-stranded DNA copy of mRNA in accordance with the above-described kit. The kit can further comprise the kit described above, in which case the at least one reagent comprises at least one tagged oligonucleotide, single-stranded nucleic acid molecule, or DNA mimic that is complementary to (i) an rRNA selected from the group consisting of 18S, 28S, 12S, 16S, 5.8S and 5S and/or (ii) a tRNA.

[0045] The invention further provides a method of determining the level of nrRNA in a sample of nrRNA, which has been separated from total RNA as described above. The method comprises (i) separating nrRNA fragments from rRNA and rRNA fragments according to the inventive method, and (ii) labeling the separated nrRNA with a detectable label. The nrRNA can be labeled by any suitable method as is known in the art. Examples of suitable labels include biotin, fluorescent tags, radioactive labels, and the like. The method further comprises (iii) providing an array comprising a collection of fixed DNAs that hybridize to a plurality of target sequences located within an expression sequence of a gene of interest. The fixed DNAs (e.g., DNA fragments) comprise from about 15 nucleotides to about 750 nucleotides, such as from about 30 nucleotides to about 750 nucleotides. At least one target sequence is interspersed approximately every 500 base pairs of the expression sequence. In other words, preferably, a fixed DNA hybridizes to a target sequence that is within 500 base pairs of another target sequence. More preferably, a target sequence recognized by a fixed DNA is located less than 500 base pairs from other target sequences recognized by fixed DNAs. Also preferably, the collection of fixed DNAs comprises a sufficient number of fixed DNAs to hybridize to target sequences of about 50% or more of the expression sequence. In other words, the collection preferably comprises DNA fragments that correspond to at least about 50% of each gene of interest and, preferably, the collection comprises at least one unique or discrete DNA fragment for every 500 contiguous base pairs of each gene of interest. The method further comprises (iv) applying the labeled nrRNA fragments to the array under hybridization conditions, i.e., under conditions that allow annealing of the labeled nrRNA with complementary fragments in the array of DNA fragments, and (v) optionally removing unbound and nonspecifically bound nrRNA fragments from the array. The method then comprises (vi) detecting the labeled nrRNA fragments that have specifically annealed to the fixed DNAs in the array (i.e., the nrRNA that has annealed to complementary fragments in the array of DNA

fragments); and (vii) quantitating the level of labeled nrRNA fragments detected in (iv) to thereby determine the level of nrRNA in the sample. The method can comprise (viii) correlating the level of nrRNA in the sample to a level of expression of a gene of interest. For example, the amount of nrRNA in the sample indicates whether the gene of interest is highly or poorly expressed.

[0046] An array for use in the inventive method can include any one-dimensional, two-dimensional or substantially two-dimensional (as well as a three-dimensional) arrangement of addressable regions bearing particular probes associated with that region. The arrays for use in the invention generally are arrays of probes as described above, and can include nucleic acids, including oligonucleotides, polynucleotides, cDNAs, RNAs, synthetic mimetics thereof, and the like. Oligonucleotide probes can be attached to the array substrate at any point along the nucleic acid chain, but are generally attached at one of their termini (e.g., the 3' or 5' terminus).

[0047] Any suitable array substrate can be used in the context of the invention including, for example, siliceous materials, e.g., glass, fused silica, ceramics and the like; metals such as stainless steel, aluminum, and alloys thereof; polymers, e.g., plastics and other polymeric materials such as polysulfone, poly(vinylidene fluoride), poly(ethyleneterephthalate), polyurethane, e.g., nonporous polyurethane, fluoropolymers such as polytetrafluoroethylene (e.g., TeflonTM) or the like, polypropylene, polystyrene, polycarbonate, PVC, nylon, and blends thereof. With arrays that read using fluorescence, the substrate may be of a material that emits low fluorescence upon illumination with the excitation light. Additionally, the substrate may be relatively transparent to reduce the absorption of the incident illuminating laser light and subsequent heating if the focused laser beam travels too slowly over a region. Any given substrate may carry one, two, four or more arrays disposed on a front surface of the substrate. Depending upon the use, multiple arrays can be used which are the same or different from one another. Each array may cover an area of less than 100 cm², or even less than 50 cm², 10 cm² or 1 cm². The substrate carrying one or more arrays can be any shape, such as a rectangular solid (although other shapes are possible). Arrays are further described in, for example, U.S. Patent No. 7,022,157.

[0048] Arrays can be fabricated in a number of ways, including, for example, drop deposition, direct synthesis, light directed fabrication, or by printing. Methods of

generating arrays are further described in, for example, U.S. Patent Nos. 6,171,797; 6,180,351; 6,232,072; 6,242,266; and 6,323,043, as well as U.S. Patent Application No. 09/302,898, filed Apr. 30, 1999 by Caren et al. (a counterpart of which has been published as U.S. Patent Application No. 2004/0203138 A1), and the references cited therein, which are hereby incorporated by reference. For example, the labeled nrRNA can be incubated with an array as described herein in a commercially available buffer (e.g., such as those available from Sigma Chemical Co., St. Louis, MO), for example.

[0049] Unbound and nonspecifically bound nrRNA can be removed by washing in a salt-containing buffer at various temperatures/stringencies. The method employed in detecting the labeled nrRNA will be determined largely by the label employed in accordance with well-known methods. The level of detected labeled nrRNA can be quantitated after subtracting the background. If desired, the nrRNA can be initially amplified in accordance with the above-described methods.

[0050] Accordingly, in view of the above, the invention also provides a microarray for analysis of nrRNA fragments for genes of interest. The microarray comprises a collection of DNA fragments, wherein each fragment ranges in size from about 15 nucleotides to about 750 nucleotides (e.g., from about 30 base pairs to about 750 base pairs), such as from about 50 base pairs to about 600 base pairs or from about 60 base pairs to about 500 base pairs. The collection preferably comprises at least one unique DNA fragment (or discrete DNA fragment) for every 500 contiguous base pairs of each gene of interest. In other words, the collection preferably comprises a DNA fragment (i.e., probe) that hybridizes to a target DNA (e.g., gene) of interest within 500 base pairs of another DNA fragment of the collection. Preferably, the collection comprises DNA fragments that correspond to at least about 50% of each gene of interest, as described further herein.

[0051] Thus, further provided is a kit comprising instructions and at least one reagent for separating nrRNA from a sample of total RNA by (i) removing rRNA from the sample of total RNA, and, optionally, simultaneously or subsequently removing any other RNA that is not of interest, and (ii) purifying the nrRNA, and the above-described microarray. The at least one reagent comprises at least one tagged oligonucleotide, single-stranded nucleic acid molecule, or DNA mimic that is complementary to (i) an rRNA selected from the group consisting of 18S, 28S, 12S, 16S, 5.8S and 5S and/or (ii) a tRNA.

[0052] In addition, the invention provides a system for separating nrRNA fragments from a sample comprising rRNA and rRNA fragments, such as eukaryotic rRNA (e.g., comprise 28S, 18S, 5.8S, 5S, 16S, and 12S rRNAs or fragments thereof). The system comprises (i) a device adapted for the separation of nucleic acids by hybridization; and (ii) a plurality of probes positioned within or upon the device to separate the rRNA and rRNA fragments from nrRNA. The probes hybridize to (a) different rRNA target sequences of at least 50% of contiguous regions of about 100 base pairs of the rRNA, which contiguous regions comprise the rRNA targeting sequences. Preferably, the probes of the plurality hybridize to rRNA targeting sequences of at least 60% of the contiguous regions (more preferably at least 70% of the contiguous regions) to further increase the efficiency of separation. More preferably, the probes hybridize to rRNA targeting sequences of at least 80% of the contiguous regions (e.g., at least 90% of the contiguous regions). Even more preferably, the probes hybridize to rRNA targeting sequences of at least 95% of the contiguous regions (e.g., the probes hybridize to rRNA targeting sequences of all of the contiguous regions). The probes also hybridize to (b) rRNA fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the rRNA.

[0053] Any device that employs hybridization to separate of nucleic acids can be used in the context of the inventive method. For example, the device can be an array, such as the array described above, where probes are attached to a substrate to facilitate separating target oligonucleotides from a sample. The device can be a filtration apparatus wherein probes are attached to filters to allow oligonucleotide separation following centrifugation. Likewise, the device can be a microfluidic device. Microfluidic devices can comprise biochannels or microchannels comprising arrays of probes to capture target RNA in samples.

Alternatively, a microfluidic reaction apparatus can be configured to have one or more individual reaction chambers in direct communication with a microarray of oligonucleotide probes to perform multiple, parallel, thermally controlled hybridization reactions.

Microfluidic-style devices have been developed for oligonucleotide amplification, detection, and/or hybridization assays, as described in International Patent Applications WO 96/15450; WO 96/15576; WO 97/27324; WO 96/39252; WO 96/39260; WO 97/16561; WO 97/16835; WO 97/37755; WO 97/43629; and WO 98/13683; U.S. Patent Nos. 5,061,336; 5,071,531; 5,110,745; 5,126,022; 5,135,627; 5,147,607; 5,296,375; 5,304,487;

5,486,335; 5,498,392; 5,569,364; 5,585,069; 5,587,128; 5,593,838; 5,603,351; 5,631,337; 5,632,876; 5,637,469; 5,643,738; 5,681,484; 5,726,026; 5,747,169; 5,750,015; 5,755,942; 5,770,029; 5,843,767; and U.S. Patent Application Publication No. 2005/0009101 A1.

[0054] The following examples serve to illustrate the invention but are not intended to limit its scope in any way.

EXAMPLE 1

[0055] This example describes a method of recovering mRNA from paraffin-embedded tissues by subtraction with oligonucleotides complementary to rRNA.

[0056] Five 10 micron sections of paraffin-embedded, benign prostatic hyperplasia tissues are extracted 3X with xylene to remove the paraffin and 3X with ethanol. After air drying, the tissue is incubated in 100 μ l of proteinase K buffer at 50°C for four days, with additional aliquots of proteinase K buffer added every 12 hours. The proteinase K buffer comprises 10 millimolar Tris (pH 8), 5 millimolar EDTA, 100 millimolar NaCl, 0.1% SDS, and 2 μ g/mL proteinase K. The RNA is then isolated using total RNA isolation methods, such as lysis in guanidium thiocyanate followed by phenol-chloroform extraction and ethanol precipitation. A number of commercial kits are suitable for this purpose, such as Totally RNA (Ambion), Perfect RNA Eukaryotic Kit (Eppendorf, Westbury, NY), and RNeasy kit (Qiagen, Valencia, CA). The RNA is extracted according to manufacturer's directions. Following extraction and precipitation, the RNA is subjected to DNase treatment, phenol chloroform extraction, and precipitation. The RNA resuspended in H₂O, and stored at -80°C.

[0057] To remove the rRNA, one microgram of the total RNA is incubated with biotinylated oligonucleotides in a hybridization buffer to promote the annealing of the oligonucleotide to the rRNA fragments. A sample of suitable oligonucleotides is shown in Figures 2a - 2d. Any suitable buffer can be used for this step, including buffers containing tetramethylammonium chloride (TMACl; see, e.g., U.S. Patent No. 5,633,134 in regard to the use of quaternary ammonium salt). One such buffer is composed of 50 mM Tris HCl, pH 8.0, 150 mM NaCl, 10 mM EDTA, and 0.1% SDS. The oligonucleotide mix includes an equimolar amount of all oligonucleotides that are complementary to 28S, 18S, 5.8S, 5S, 12S, and 16S rRNA. One microgram of RNA and 2 micrograms of oligonucleotides are

resuspended in 10 μ l of H₂O and incubated at 75°C for 10 minutes, followed by immersion in an ice bath for 10 minutes. Ten microliters of 10X buffer and 80 μ l of H₂O are added, and the solution is incubated at 65°C for between 10 minutes and several hours. Following the incubation step, a 50 μ l aliquot of streptavidin-coated magnetic beads (Dynal, Oslo, Norway) is added to the hybridization solution, which is then mixed gently to allow for binding of the biotinylated oligonucleotides to the streptavidin-coated beads. The beads are then precipitated using a magnet, and the supernatant is transferred to a fresh tube. If desired, the hybridization step can be repeated one or more times by adding a fresh aliquot of the oligonucleotides. The supernatant is extracted with phenol-chloroform, and the RNA is precipitated with the addition of 2 volumes of ethanol. The RNA is resuspended in H₂O, quantitated using the picogreen DNA quantitation kit (Invitrogen, Carlsbad, CA), and stored at -80°C. The RNA is then ready for reverse transcription.

[0058] This example illustrates a method of separating nrRNA from rRNA and rRNA fragments using the inventive method.

EXAMPLE 2

[0059] This example demonstrates a method of preparing probes for use in the invention.

[0060] To generate single-stranded DNA fragments complementary to rRNA, PCR primers are designed to amplify segments of the rRNA species that range in size from around about 100 base pairs up to the full-length rRNA template. The primers listed in Figures 3A - 3C are designed to amplify fragments corresponding to all six rRNA species. The amplification reaction is performed using MasterTaq kit from Eppendorf according to the manufacturer's direction. The buffer is supplemented with 10% DMSO or 5% formamide when needed to improve the amplification of GC rich sequences. The amplification is performed for a total of 35 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds, and extension for 1 minute at 72 °C in the presence of 25 pmoles of the forward and reverse primers. The amplified rRNA fragments are used as templates to generate the single-stranded DNA complementary to the rRNA using asymmetric PCR reactions. This is achieved by reducing the concentration of the forward primer to 2.5 pmoles during the PCR amplification, which leads to an excess of the desired strand. Following asymmetric PCR, the DNA is purified using phenol-chloroform

extraction and ethanol precipitation. The DNA fragments are quantitated and mixed in equimolar amounts.

[0061] To remove rRNA from a sample comprising total RNA, 1 μ g of the total RNA is incubated with 2 μ g of the single-stranded DNA fragments complementary to rRNA. The hybridization and rRNA removal are performed as described above.

[0062] This example illustrates a method of preparing probes for use in the inventive method, kit, and system.

EXAMPLE 3

[0063] This example describes a method of analyzing nrRNA expression. In particular, this example describes reverse transcription and detection of p53 mRNA.

[0064] Reverse transcription is performed using the cMaster RT_{plus} PCR kit (Eppendorf) using random primers or a combination of random primers and oligo dT primers according to the supplier's directions. Twenty nanograms of selected RNA (e.g., nrRNA isolated according to the method of Example 1) are used for each reaction. The reaction is stopped by heating at 70°C for 15 minutes, followed by the addition of 1 unit of Rnase H and incubation at 37°C for 20 minutes to degrade the RNA. The cDNA is now ready for RT-PCR. The transcription reaction is stored at -20°C.

[0065] To detect the p53 transcript, 1 μ l of the cDNA is amplified using primers designed to amplify a portion of the p53 transcript. Examples of suitable primer pairs are:
F1: ctgtccgtccaagcaatggatg (SEQ ID NO: 185); R1: ggagcttcacatggacctgggtc (SEQ ID NO: 246) (89 base pairs),
F2: caataggtgtgcgtcagaagcacc (SEQ ID NO: 186); R2: caaacaccagtcaggccaacttg (SEQ ID NO: 247) (86 base pairs), and
F3: ggtctcacagtgtgcccaggctg (SEQ ID NO: 187); R3: ttgtaatcccagcactctgggag (SEQ ID NO: 248) (86 base pairs).

One microliter of the cDNA is amplified using the primer pairs shown above for 40 cycles using the MasterTaq PCR amplification kit (Eppendorf) in a 15 μ l reaction containing 25 pmoles of each primer. The annealing temperature is 60°C. The reactions are separated on a 7.5% acrylamide gel and stained with ethidium bromide to detect the amplification products.

[0066] This example illustrates a method of analyzing nrRNA expression from a sample of nrRNA separated from rRNA in accordance with the invention.

EXAMPLE 4

[0067] This example describes amplification of mRNA fragments using transcription-coupled amplification.

[0068] One hundred nanograms of selected RNA fragments (e.g., nrRNA fragments isolated according to the method of Example 1) are ligated to 1 picomole of RNA oligonucleotide (GGAUGAACGUAGGAAGCUUG (SEQ ID NO: 188)) using 2 units of T4 RNA ligase in a 20 μ l-reaction (NEB, Beverly, MA) according to supplier's recommendations. Following the ligation reaction, the RNA is purified by adding 80 μ l of TE8 to the ligation reaction and then gently pipetting the entire volume to the top of a Sephadex G25 column (Pharmacia, Peapack, NJ). The reaction is spun at 4,000 rpm for 1 minute, and the recovered RNA is precipitated by adding 0.1 volumes of 5.2 M sodium acetate and 2 volumes of ethanol. The RNA is resuspended in 20 μ l of TE8. The reverse transcription reaction is performed as described in the Superscript Choice System for cDNA synthesis (Invitrogen), except that the primers used to initiate the priming of the first strand are composed of a T7 promoter followed by around 6 to 12 random nucleotides (GGAATTAATACGACTCACTATAGGGN9 (SEQ ID NO: 189)). The reaction is treated with RNase H at 30°C for 30 minutes. The second strand synthesis is initiated using the following primer, which includes a T3 promoter (GCGCGAAATTAACCCTCACTAAA GGGATGAACGTAGGAAGCTTG (SEQ ID NO: 190)). Following the second strand synthesis, the cDNA is extracted with phenol-chloroform, ethanol precipitated, and resuspended in TE8. The cDNA (Figure 5) is ready for the amplification step using the MEGAshortscript high yield transcription kit (Ambion). The transcription reaction is performed according to the manufacturer's recommendations. Depending on the choice of promoter used for the transcription reaction, the sense or the antisense strand can be generated. The amplified RNA can serve as a template for reverse transcription and RT-PCR as described in Example 3 or as a probe for use in microarray hybridization.

[0069] This example illustrates a method of amplifying nrRNA separated from rRNA according to the method of the invention.

[0070] All references, including publications, patent applications, and patents, cited herein are hereby incorporated by reference to the same extent as if each reference were individually and specifically indicated to be incorporated by reference and were set forth in its entirety herein.

[0071] The use of the terms “a” and “an” and “the” and similar referents in the context of describing the invention (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms “comprising,” “having,” “including,” and “containing” are to be construed as open-ended terms (i.e., meaning “including, but not limited to,”) unless otherwise noted. Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

[0072] Preferred embodiments of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Variations of those preferred embodiments may become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventors expect skilled artisans to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

WHAT IS CLAIMED IS:

1. A method of separating non-ribosomal transcribed RNA (nrRNA) fragments from ribosomal RNA (rRNA) and rRNA fragments, wherein the method comprises:

(i) providing a sample comprising rRNA, rRNA fragments, and nrRNA fragments, wherein the rRNA comprises multiple contiguous regions of about 100 base pairs, each contiguous region comprising an rRNA targeting sequence, and wherein the rRNA fragments comprise one or more of the rRNA targeting sequences;

(ii) providing a plurality of probes, wherein the probes hybridize to (a) different RNA targeting sequences of at least 50% of the contiguous regions of the rRNA and (b) rRNA fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the rRNA;

(iii) adding the plurality of probes to the sample,

(iv) hybridizing the probes to the rRNA and rRNA fragments to form rRNA-probe complexes and rRNA fragment-probe complexes, and

(v) separating the rRNA-probe complexes and rRNA fragment-probe complexes from the sample,

thereby separating nrRNA fragments from rRNA and rRNA fragments.

2. The method of claim 1, wherein the method further comprises (vi) recovering and purifying the nrRNA fragments.

3. The method of claim 1 or claim 2, wherein the sample comprises total RNA.

4. The method of any of claims 1-3, wherein the probes hybridize to rRNA targeting sequences of at least 60% of the contiguous regions.

5. The method of claim 4, wherein the probes hybridize to rRNA targeting sequences of at least 70% of the contiguous regions.

6. The method of claim 5, wherein the probes hybridize to rRNA targeting sequences of at least 80% of the contiguous regions.

7. The method of claim 6, wherein the probes hybridize to rRNA targeting sequences of at least 90% of the contiguous regions.
8. The method of claim 7, wherein the probes hybridize to rRNA targeting sequences of at least 95% of the contiguous regions.
9. The method of claim 8, wherein the probes hybridize to rRNA targeting sequences of all of the contiguous regions.
10. The method of any of claims 1-9, wherein the rRNA and the rRNA fragments comprise eukaryotic rRNA and eukaryotic rRNA fragments.
11. The method of any of claims 1-10, wherein the rRNA and rRNA fragments comprise at least two rRNAs selected from the group consisting of 28S, 18S, 5.8S, 5S, 12S and 16S rRNAs.
12. A method of amplifying nrRNA fragments, wherein the method comprises:
 - (i) separating nrRNA fragments from rRNA and rRNA fragments according to the method of any of claims 1-11, and
 - (ii) amplifying the nrRNA fragments.
13. A method of analyzing nrRNA expression, wherein the method comprises:
 - (i) separating nrRNA fragments from rRNA and rRNA fragments according to the method of any of claims 1-11, and
 - (ii) analyzing nrRNA expression.
14. A method of amplifying a cDNA complementary to an nrRNA fragment, wherein the method comprises:
 - (i) separating an nrRNA fragment from rRNA and rRNA fragments according to the method of any of claims 1-11, wherein the nrRNA fragment comprises a 5' end;
 - (ii) ligating a first oligonucleotide to the 5' end of the nrRNA fragment;

(iii) hybridizing to the nrRNA fragment a second oligonucleotide;
(iv) extending a DNA strand from the second oligonucleotide to generate a first cDNA;
(v) amplifying the first cDNA,
thereby amplifying a cDNA complementary to an nrRNA fragment.

15. A method of amplifying an nrRNA fragment, wherein the method comprises:
(i) separating an nrRNA fragment from rRNA and rRNA fragments according to the method of any of claims 1-11, wherein the nrRNA fragment comprises a 5' end and a 3' end;
(ii) ligating a first oligonucleotide to the 5' end of the nrRNA fragment;
(iii) hybridizing to the nrRNA fragment a second oligonucleotide;
(iv) extending a DNA strand from the second oligonucleotide to generate a first cDNA having a 5' end and a 3' end;
(v) degrading the nrRNA fragment and, optionally, removing any free oligonucleotides and/or free RNA fragments;
(vi) synthesizing a complementary strand to the first cDNA by
(a) contacting the first cDNA with a third oligonucleotide such that the third oligonucleotide hybridizes to the 3' end of the first cDNA, wherein the third oligonucleotide hybridizes to at least a portion of the first oligonucleotide and wherein the third oligonucleotide optionally contains an RNA polymerase promoter sequence, and
(b) extending the complementary strand from the third oligonucleotide, whereupon a double-stranded DNA copy of the nrRNA fragment is generated;
(vii) optionally purifying the double-stranded DNA copy; and
(viii) transcribing RNA from the double-stranded DNA copy,
thereby amplifying the nrRNA fragment.

16. The method of claim 15, wherein the second oligonucleotide comprises a unique sequence, a random sequence, and, optionally, an RNA polymerase promoter sequences, wherein the random sequence is of sufficient length to hybridize to the nrRNA.

17. A method of determining the level of nrRNA in a sample, wherein the method comprises:

- (i) separating nrRNA fragments from rRNA and rRNA fragments according to the method of any of claims 1-11;
- (ii) labeling the separated nrRNA fragments with a detectable label;
- (iii) providing an array comprising a collection of fixed DNAs that hybridize to a plurality of target sequences located within an expression sequence of a gene of interest, wherein the fixed DNAs comprise from about 15 nucleotides to about 750 nucleotides, wherein at least one target sequence is interspersed approximately every 500 base pairs of the expression sequence;
- (iv) applying the labeled nrRNA fragments to the array under hybridization conditions;
- (v) optionally removing unbound and nonspecifically bound nrRNA fragments from the array;
- (vi) detecting the labeled nrRNA fragments that have specifically annealed to the fixed DNAs in the array; and
- (vii) quantitating the level of labeled nrRNA fragments detected in step (vi) to thereby determine the level of nrRNA in the sample.

18. The method of claim 17, wherein the method further comprises (viii) correlating the level of nrRNA to a level of expression of the gene of interest.

19. A kit for separating non-ribosomal transcribed RNA (nrRNA) fragments from a sample comprising ribosomal RNA (rRNA) and rRNA fragments, wherein the kit comprises a plurality of probes that hybridize to (a) different rRNA target sequences of at least 50% of contiguous regions of about 100 base pairs of the rRNA, which contiguous regions comprise the rRNA targeting sequences and (b) rRNA fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the rRNA.

20. The kit of claim 19, wherein the probes hybridize to rRNA targeting sequences of at least 60% of the contiguous regions.

21. The kit of claim 20, wherein the probes hybridize to rRNA targeting sequences of at least 70% of the contiguous regions.
22. The kit of claim 21, wherein the probes hybridize to rRNA targeting sequences of at least 80% of the contiguous regions.
23. The kit of claim 22, wherein the probes hybridize to rRNA targeting sequences of at least 90% of the contiguous regions.
24. The kit of claim 23, wherein the probes hybridize to rRNA targeting sequences of at least 95% of the contiguous regions.
25. The kit of claim 24, wherein the probes hybridize to rRNA targeting sequences of all of the contiguous regions.
26. The kit of any of claims 19-25, wherein the rRNA and rRNA fragments comprise eukaryotic rRNA and eukaryotic rRNA fragments.
27. The kit of any of claims 19-26, wherein the rRNA and rRNA fragments comprise 28S, 18S, 5.8S, 5S, 16S, and 12S rRNAs or fragments thereof.
28. The kit of any of claims 19-27 further comprising instructions.
29. A system for separating non-ribosomal transcribed RNA (nrRNA) fragments from a sample comprising ribosomal RNA (rRNA) and rRNA fragments, the system comprising:
- (i) a device adapted for the separation of nucleic acids by hybridization; and
 - (ii) a plurality of probes that hybridize to (a) different rRNA target sequences of at least 50% of contiguous regions of about 100 base pairs of the rRNA, which contiguous regions comprise the rRNA targeting sequences, and (b) rRNA fragments comprising the rRNA targeting sequences hybridized to by the probes in the contiguous regions of the

rRNA, wherein the plurality of probes is positioned within the device to separate the rRNA and rRNA fragments from nrRNA.

30. The system of claim 29, wherein the probes hybridize to rRNA targeting sequences of at least 60% of the contiguous regions.

31. The system of claim 30, wherein the probes hybridize to rRNA targeting sequences of at least 70% of the contiguous regions.

32. The system of claim 31, wherein the probes hybridize to rRNA targeting sequences of at least 80% of the contiguous regions.

33. The system of claim 32, wherein the probes hybridize to rRNA targeting sequences of at least 90% of the contiguous regions.

34. The system of claim 33, wherein the probes hybridize to rRNA targeting sequences of at least 95% of the contiguous regions.

35. The system of claim 34, wherein the probes hybridize to rRNA targeting sequences of all of the contiguous regions.

36. The system of any of claims 29-35, wherein the rRNA and rRNA fragments comprise 28S, 18S, 5.8S, 5S, 16S, and 12S rRNAs or fragments thereof.

1/26

28S rRNA sequence:

cgcgacctcagatcagacgtggcgacccgctgaatttaagcatattagtcagcggaggaaaagaaactaaccagg
attccctcagtaacggcgagtgaaacaggaagagcccagcgcgaatccccgccccgcggggcgcgggacatgt
ggcgtagcgaagacccgctccccggcgccgctcgtggggggcccaagtccttctgatcagggcccagcccgtgga
cgggtgtgaggccggtagcggccggcgcgcccggtcttcccgagtcgggttgcttggaatgcagcccaaag
cgggtggtaaactccatctaaggctaaataccggcacgagaccgatagtaacaagtaccgtaagggaagtgga
aaagaactttgaagagagagttcaagagggcgtaaaccggttaagaggtaaacgggtgggtccgcgcagtcgg
cccgaggattcaacccggcgggcggtccggccgtgtcggcgcccgcggtatcttcccgccccccgttccctccg
acccctccacccgcccctccctccccgcgcgccctcctcctcctccccggaggggggcggtccggcggggtgcgg
gggtggggcgggcgggggccgggggtggggtcggcgggggaccgtccccggaccggcgaccggccgcccgg
gcgcatttccaggcgggtgcgcgcgaccggctccgggacggctgggaaggccggcggggaaggtggctcggg
gggccccgtccgtccgtccgtcctcctccccgtctccgcccccgccccgcgtcctccctcgggagggcgcg
gggtcggggcgggcgggcgggcgggcggtggcgggcgggcgggggcgggcgggaccgaaaccccccccgagt
gttacagccccccggcagcagcactcgccgaatcccgggccgagggagcgagaccgctcgccgcgtctccc
ccctccggcgcccacccccgcgggaatccccgcgaggggggtctccccggcgcgggcgccggcgtcctcgtg
ggggggcggggcccacccctcccacggcgcgaccgctctcccacccctcctccccgcgcccccgccccggcgacg
gggggggtgccgcgcgcgggtcgggggcgggggcggaactgtcccagtcgccccggcggggtcgcgccgtcg
ggccccggggaggttctcggggccacgcgcgcgtccccgaagagggggacggcgagcgagcgacggg
gtcggcgggcgacgtcggtacccacccgaccgctctgaaacacggaccaaggagttaaacacgtgcgcgagtcg
ggggctcgcacgaaagccgcccgtggcgcaatgaaggtgaaggccggcgcgctcgccggccgaggtgggatccc
gaggcctctccagtccgcccagggggcaccacccggcccgtctcgcccgccgccccggggaggtggagcacgagc
gcacgtgttaggacccgaaagatggtgaactatgcctgggcagggcgaaagccagaggaaactctggtggaggtcc
gtacgggtcctgacgtgaaatcggtcgtccgacctgggtatagggcgaaagactaatcgaaccatctagtagctg
gtccctccgaagttccctcaggatagctggcgctctcgcagaccgcacccccgccacgcagtttatccggtga
aagcgaatgattagaggtcttggggccgaaacgatctcaacctattctcaaaactttaaatgggtaagaagcccggtc
gctggcggtggagccgggggtggaatgcgagtgccctagtggggccacttttgtaagcagaactggcgctcgggatga
accgaacgcccgggttaaggcgcccgatgccgacgctcatcagaccccagaaaagggtgttggtgatatagacagc
aggacggtggccatggaagtcggaatccgctaaggagtggtgaacaactcacctgccgaatcaactagccctgaaa
atggatggcgctggagcgtcggggccatacccgccgctcgccggcagtcgagagtggaaggagcgggcggggg
cgggcgcgcgcgcgcgctgtggtgtgcgtcgagggcgggcgggcgggcgggcggggggtgtgggtccttc
ccccgccccccccccacgcctcctccctcctccccgcacgccccgctccccgccccggagccccgcggagc
tacgccgcgacgagtaggagggccgctcggtgagcctgaagcctagggcgcgggccccgggtggaggccgcc
gcaggtgcagatcttggtggtagtagcaaataattcaaacgagaactttgaaggccgaagtggagaagggttccatgt
gaacagcagttgaacatgggtcagtcggtcctgagagatgggcgagcgccgttccgaagggaaggcgatggcct
ccgttgccctcggccgatcgaaggagtcgggttcagatccccgaatccggagtgggcgagatgggcgcccgcga
ggcgtccagtgcggtaacgcgaccgatcccgagaaagccggcgggagccccggggagagttctctttctttgtgaa
gggcagggcgccctggaatgggttcgccccgagagagggggcccgctgccttgaaagcgtcgcggttccggcggc

FIG. 1A

2/26

gtccggtgagctctcgctggcccttgaaaaatccgggggagaggggtgtaaactctcgcccgggccggtacccatataccg
cagcaggttccaaggtgaacagcctctggcatgttggaacaatgtaggtaaggggaagtcgggaagccggatccgt
aacttcgggataaggattggctctaagggtcgggtcgggtcggggcgcggaagcggggctgggcgcgcgccg
cggctggacgaggcgcgcgccccccacgcccggggcacccccctcgcggccctcccccgccccacccgcg
gcgcccgtcgtccctccccaccccgcgccctctctctctctctctccccgctccccgtctccccctccccggggga
gcgcccgcgtgggggcgcggcgggggggagaagggtcggggcggcaggggccgcgcggcggccgcccggggcg
gccggcgggggcaggtccccgcgagggggggccccggggacccggggggccggcggcggcgaggactctgga
cgcgagccggggcccttcccggtggatcgccccagctcgggcgggcgctcgcgcccgccccggggagcccggcg
cggcgcggcgcgccccccacccccacccacgctcggtcgcgcgcgcgctccgctgggggaggagcggtcggg
cggcggcggtcgcgggcggcggggcggggcgggttcgtcccccgccctaccccccgccccgtccgcccc
gtccccctctctctcggcgcgcggcggcggcggcggcaggcggcggaggggcccgcgggcccgggtccccccgc
cgggtccgcccccggggccgcggttcgcgcgcgccctgcctcgccgggcgcctagcagccgacttagaactggt
gcggaccaggggaatccgactgtttaattaaacaaagcatcggaaggcccgcggcgggtgttgacgcgatgtg
atttctgccagtgtctgaatgtcaaagtgaagaaattcaatgaagcgcggttaaaccggcgggagtaactatgactc
tcttaaggtagccaaatgcctcgtcatctaattagtacgcgcatgaatggatgaacgagattccactgtccctaccta
ctatccagcgaaaccacagccaagggaacgggcttggcggaatcagcggggaaagaagacctgttgagcttga
ctctagtctggcacggtgaagagacatgagagggttagaataagtgggaggccccggcgcccccccggtgtcc
cgcgaggggcccggggcggggtccgcggccctcggggcccgggtgaaataccactactctgatcgtttttactg
acccgggtgaggcgggggggcgagcccaggggctctcgcttctggcgccaagcgcccggccgggcccgcgga
cccgtccggggacagtgccaggtggggagtttgactggggcggtacacctgtcaaaccggtaacgcaggtgtccta
aggcgagctcagggaggacagaaacctcccgtggagcagaagggcaaaagctcgcttgatcttgatttcagtacg
aatacagaccgtgaaagcggggcctcacgatccttctgaccttttggtttaagcaggagggtgtcagaaaagttacc
acagggataactggcttgtggcgggccaagcggtcatagcgacgtcgcttttgatccttcgatgtcggtcttcctatcatt
gtgaagcagaattcgccaagcgttggattgttcacccactaatagggaaacgtgagctgggttagaccgtcgtgagac
aggtagttttacctactgatgatgtgtgttgccatggtaacctgtcagtagagaggaaccgcaggttcagacattt
gggtgatgtgcttggctgaggagccaatggggcgaagctaccatctgtgggattatgactgaacgcctctaagtcaga
atcccggccaggcgaacgatacggcagcgccgcggagcctcggttggcctcggaatagccgggtcccccgctgtcc
ccgcccggggccgccccccccctccacgcgccccgcggcgggagggcgcggtccccgcgcgcgccgggacc
gggggtccgggtcgagtgcccttcgtcctgggaaacggggcgcgccgggaaaggcgggccgccccctcgccgctc
acgcaccgcacgttcgtggggaacctggcgctaaccattcgtagacgacctgttctgggtcgggggttcgtacgta
gcagagcagctccctcgctgcgatctattgaaagtcagccctcgacacaaggggttgc [SEQ ID NO: 1]

18S rRNA sequence:

cgtgctcctcccgtcgccgtccgggcccgtccgtccgtccgtccgtcctcctcgctnnnccggggcgccggggcc
cgctctcacnggcccccgnnnnngtccnggccgtcggggcctcgccgcgctctacctacctggtgatcctg
ccagtagcatatgcttgtctcaaagattaagccatgcatgtctaagtacgcacggccgggtacagtgaactgcgaatg
gtcattaaatcagttatggttcctttggtcgctcgctcctctcctacttgataactgtggaattctagagctaatacatgc

FIG. 1B

3/26

[illegible]

5.8S rRNA sequence:

cgactcttagcgggtggatcactcggctcgtgcgtcgatgaagaacgcagctagctgcgagaattaatgtgaattgcag
gacacattgatcatcgacacttcgaacgcacttgcggccccgggttctcccggggctacgcctgtctgagcgtcgct
[SEQ ID NO: 3]

5S rRNA sequence:

gtctacggccataccaccctgaacgcgcccgatctcgctgatctcggaagctaagcagggctcgggcctggttagtac
ttggatgggagaccgcctgggaataccgggtgctgtaggcttt [SEQ ID NO: 4]

FIG. 1C

4/26

16S rRNA sequence:

sequence:gctaaacctagccccaaccccactccaccttactaccagacaaccttagccaaaccattacccaaat
aaagtataggcgatagaaattgaaacctggcgcaatagatatagtaccgcaagggaaagatgaaaaattataacc
aagcataatatagcaaggactaaccctataccttctgcataatgaattaaactagaaataactttgcaaggagagcca
aagctaagacccccgaaaccagacgagctacctaagaacagctaaaagagcacaccgctctatgtagcaaaata
gtgggaagatttataggttagaggcgacaaacctaccgagcctgggtgatagctgggtgtccaagatagaatcttagttc
aactttaaatgtgccacagaacccctctaaatcccctgtaaatttaactgttagtccaaagaggaacagctcttggaca
ctaggaaaaaacctgttagagagagtaaaaaatttaaccccatagtaggcctaaaagcagccaccaattaagaa
agcgttcaagctcaacacccactacctaataatcccaaacatataactgaactcctcacaccaattggaccaatc
tatccctatagaagaactaatgttagtataagtaacatgaaaacattctcctccgcataagcctgcgtcagattaaa
acactgaactgacaattaacagcccaatatctacaatcaaccaacaagtcattattacccctactgtcaacccaacac
aggcatgctcataaggaaagggttaaaaaagtaaaagggaactcggcaaatcttaccgctgtttaccaaaaaca
tcacctctagcatcaccagttattagaggcaccgctgccagtgacacatgttaacggccgcggtaccctaaccgtg
caaaggtagcataatcactgttccctaaatagggacctgtatgaatggctccacgaggggtcagctgtctctactttta
ccagtgaaattgacctgcccgtaagaggcgggcataacacagcaagacgagaagaccctatggagctttaattta
ttaatgcaaacagtacctaacaacccacaggtcctaactaccaaacctgcattaaaaatttcggttggggcgacct
cggagcagaacccaacctccgagcagtagctgaagacttcaccagtc aaagcgaactactatactcaattgatcc
aataacttgaccaacggaacaagttaccctagggataacagcgcaatcctattctagagtccatatcaacaataggg
ttacgacctcgatgttggatcaggacatcccgatgggtgcagccgctattaaagggtcgttgttcaacgattaaagtcct
acgtgatctgagttcagaccggagtaatccaggtcgggttctatctaccttcaaattcctccctgtacgaaaggacaaga
gaaataaggcctacttcacaaagcgccctccccgtaaatgatcatctcaacttagtattatacccacacccacca
agaacagggtt [SEQ ID NO: 5]

12S rRNA sequence:

aatagggttggcttagcctttctattagctcttagtaagattacacatgcaagcatccccgttccagtgaagttcacctcta
aatcaccacgatcaaaaggaacaagcatcaagcacgcagcaatgcagctcaaaacgcttagcctagccacaccc
ccacgggaaacagcagtgattaaccttttagcaataaacgaaagttaactaagctataactaaccacagggttggta
atttcgtgccagccaccgcggtcacacgattaacccaagtcaatagaagccggcgtaaagagtgttttagatcaccc
cctccccaataaagctaaaactcacctgagttgtaaaaaactccagttgacacaaaatagactacgaaagtggcttta
acatatctgaacacacaatagctaagacccaaactgggattagatacccccactatgcttagccctaaacctcaacag
ttaaatcaacaaaactgctcgccagaacactacgagccacagctaaaactcaaaggacctggcggtgcttcatac
cctctagaggagcctgttctgtaatcgataaaccccgatcaacctcaccacctcttgcagcctatataccgccatcttc
agcaaacctgatgaaggctacaaagtaagcgcaagtacccacgtaaagacgttaggtcaaggtgtagcccatga
gggtggcaagaaatgggctacattttctacccagaaaaactacgatagcccttatgaaacttaagggctgaaggtggat
ttagcagtaaaactaagagtagagtgttagttgaacagggccctgaagcggtacacaccgcccgtcacccctctca
agtatacttcaaaggacatttaactaaaacccctacgcatttatatagaggagacaagtcgtaacatggtaagtgtact
ggaaagtgcacttggacgaa [SEQ ID NO: 6]

FIG. 1D

SUBSTITUTE SHEET (RULE 26)

5/26

rRNA species	Sequence	Start	Length	SEQ ID NO:
18SRNA	ACGACGGACGGACGGACGGAC	50	21	SEQ ID NO. 7
18SrRNA	TAGGTAAGGTAGAGCGCGGCGA	142	22	SEQ ID NO. 8
18SrRNA	TTCAGTGTACCGGCCGTGCGTA	224	22	SEQ ID NO. 9
18SrRNA	GCCCGTCGGCATGTATTAGCTC	326	22	SEQ ID NO. 10
18SrRNA	TCTAGAGTCACCAAAGCCGCCG	439	22	SEQ ID NO. 11
18SrRNA	TGGTCACCATGGTAGGCACGG	541	21	SEQ ID NO. 12
18SrRNA	GGTCGGGAGTGGGTAATTTGCG	642	22	SEQ ID NO. 13
18SrRNA	GCCCTCCAATGGATCCTCGTTA	743	22	SEQ ID NO. 14
18SrRNA	CCGCTCCAAGATCCAACCTACG	837	22	SEQ ID NO. 15
18SrRNA	GCGGGACACTCAGCTAAGAGCA	925	22	SEQ ID NO. 16
18SrRNA	TTCTAGCTGCGGTATCCAGGC	1002	22	SEQ ID NO. 17
18SrRNA	CCGGTCCAAGGAATTTACCTC	1117	22	SEQ ID NO. 18
18SrRNA	TCTGATCGTCTTCGAACCTCCG	1201	22	SEQ ID NO. 19
18SrRNA	GACTTTGGTTTCCCGGAAGCTG	1294	22	SEQ ID NO. 20
18SrRNA	TTAAGCCGCAGGCTCCACTCCT	1382	22	SEQ ID NO. 21
18SrRNA	CTAAGAACGGCCATGCACCACC	1486	22	SEQ ID NO. 22
18SrRNA	TCTAAGAAGTTGGGGGACGCCG	1589	22	SEQ ID NO. 23
18SrRNA	TAGGCACACGCTGAGCCAGTCA	1697	22	SEQ ID NO. 24
18SrRNA	CTGGGAATTCCTCGTTCATGGG	1789	22	SEQ ID NO. 25
18SrRNA	CCGATCCGAGGGCCTCACTAAA	1890	22	SEQ ID NO. 26
18SrRNA	GCTCCGTTAATGATCCTTCCGC	2017	22	SEQ ID NO. 27
18SrRNA	CGAACGAACGAGCGAGCGAAC	2133	21	SEQ ID NO. 28
28SrRNA	GTCGCCACGTCTGATCTGAG	26	20	SEQ ID NO. 29
28SrRNA	CTCTTCCCTGTTCACTCGCCGT	109	22	SEQ ID NO. 30
28SrRNA	CCTCGATCAGAAGGACTTGGGC	211	22	SEQ ID NO. 31
28SrRNA	TTTACCACCCGCTTTGGGCTG	310	21	SEQ ID NO. 32
28SrRNA	ACGGTTTCACGCCCTCTTGAAC	415	22	SEQ ID NO. 33
28SrRNA	CGCCGGGTTGAATCCTCC	469	18	SEQ ID NO. 34
28SrRNA	CCTCCGGGGAGGAGGAGGAGG	579	21	SEQ ID NO. 35
28SrRNA	GCACCGCCTGGAAATGCGCC	692	20	SEQ ID NO. 36

FIG. 2A

6/26

rRNA species	Sequence	Start	Length	SEQ ID NO:
28SrRNA	GATTCGGCGAGTGCTGCT	930	18	SEQ ID NO. 37
28SrRNA	AGAGCGGTCGCGCCGTGGGAG	1082	21	SEQ ID NO. 38
28SrRNA	CGTGGCCCCGAGAGAACCTC	1220	20	SEQ ID NO. 39
28SrRNA	GTCCGTGTTTCAAGACGGGTCG	1311	22	SEQ ID NO. 40
28SrRNA	CGCCGGCCTTCACCTTCATT	1386	20	SEQ ID NO. 41
28SrRNA	TCGGGTCCTAACACGTGCGCT	1502	21	SEQ ID NO. 42
28SrRNA	TTTCGCCCCCTATACCCAGGTCG	1611	22	SEQ ID NO. 43
28SrRNA	TCGGCCCCAAGACCTCTAATCA	1744	22	SEQ ID NO. 44
28SrRNA	ACTAGGCACTCGCATTCCACCC	1830	22	SEQ ID NO. 45
28SrRNA	GGGTCTGATGAGCGTCGGCAT	1915	21	SEQ ID NO. 46
28SrRNA	GGGCTAGTTGATTCGGCAGGTG	2015	22	SEQ ID NO. 47
28SrRNA	CCCTCCGACGCACACCACAC	2130	20	SEQ ID NO. 48
28SrRNA	AGGCTTCAAGGCTCACCGCAG	2288	21	SEQ ID NO. 49
28SrRNA	CCCTTCTCCACTTCGGCCTTCA	2384	22	SEQ ID NO. 50
28SrRNA	TCGCCCATCTCTCAGGACCGA	2437	21	SEQ ID NO. 51
28SrRNA	TTCGGGGATCTGAACCCGACT	2514	21	SEQ ID NO. 52
28SrRNA	CGCCCTGCCCTTCACAAAGAAA	2629	22	SEQ ID NO. 53
28SrRNA	TTTTCAAGGGCCAGCGAGAGC	2723	21	SEQ ID NO. 54
28SrRNA	TGTTCCAACATGCCAGAGGCTG	2812	22	SEQ ID NO. 55
28SrRNA	CGACCCAGCCCTTAGAGCCAA	2883	21	SEQ ID NO. 56
28SrRNA	GTGGGGAGGGAGCGAGCGG	3016	19	SEQ ID NO. 57
28SrRNA	CTGCCGCCCGACCCTTCTC	3119	20	SEQ ID NO. 58
28SrRNA	CGGCTCGCGTCCAGAGTCC	3224	19	SEQ ID NO. 59
28SrRNA	CCGAGGAGGAGGGGGGAAC	3455	19	SEQ ID NO. 60
28SrRNA	CGCACCAGTTCTAAGTCGGCTG	3588	22	SEQ ID NO. 61
28SrRNA	TTGACATTCAGAGCACTGGGCA	3687	22	SEQ ID NO. 62
28SrRNA	CGTTCATCCATTCATGCGCGTC	3796	22	SEQ ID NO. 63
28SrRNA	TGATTCCGCCAAGCCCGTTC	3863	20	SEQ ID NO. 64
28SrRNA	CGGGGGCCTCCCATTATTCTA	3949	22	SEQ ID NO. 65
28SrRNA	CGCCTCACCGGGTCAGTGAAA	4057	21	SEQ ID NO. 66

FIG. 2B

7/26

rRNA species	Sequence	Start	Length	SEQ ID NO:
28SrRNA	TCAAACCTCCCCACCTGGCACTG	4152	22	SEQ ID NO. 67
28SrRNA	CCTTCTGCTCCACGGGAGGTTT	4236	22	SEQ ID NO. 68
28SrRNA	ATCGTGAGGCCCCGCTTTTAC	4300	21	SEQ ID NO. 69
28SrRNA	CGACGTCGCTATGAACGCTTGG	4393	22	SEQ ID NO. 70
28SrRNA	GTGAACAATCCAACGCTTGGCG	4464	22	SEQ ID NO. 71
28SrRNA	TCTGAACCTGCGGTTCTCTCG	4584	22	SEQ ID NO. 72
28SrRNA	GCCTGGGCGGGATTCTGACTTA	4680	22	SEQ ID NO. 73
28SrRNA	GTTTCCCAGGACGAAGGGCA	4850	20	SEQ ID NO. 74
28SrRNA	CGACCCAGAAGCAGGTCGTCTA	4952	22	SEQ ID NO. 75
28SrRNA	CAAACCCTTGTGTCGAGGGCTG	5023	22	SEQ ID NO. 76
5.8S rRNA	TCATCGACGCACGAGCCGAG	40	20	SEQ ID NO. 77
5.8S rRNA	GCAAGTGCGTTCGAAGTGTCTG	113	21	SEQ ID NO. 78
5.8S rRNA	AGCGACGCTCAGACAGGCGTA	156	21	SEQ ID NO. 79
5SrRNA	CTCTTAGCGGTGGATCACTCGG	25	22	SEQ ID NO. 80
5SrRNA	GTCGATGATCAATGTGTCCTGCAA	96	24	SEQ ID NO. 81
5SrRNA	AGCGACGCTCAGACAGGCGTAG	156	22	SEQ ID NO. 82
12SrRNA	ACTCACTGGAACGGGGATGCTTG	73	23	SEQ ID NO. 83
12SrRNA	TGCGTGCTTGATGCTTGTTCT	121	22	SEQ ID NO. 84
12SrRNA	GGCACGAAATTGACCAACCCTG	244	22	SEQ ID NO. 85
12SrRNA	ACGCCGGCTTCTATTGACTTGG	290	22	SEQ ID NO. 86
12SrRNA	GGTCTTAGCTATTGTGTGTTTCAG	418	23	SEQ ID NO. 87
12SrRNA	TGTGGCTCGTAGTGTTCTGGCG	507	22	SEQ ID NO. 88
12SrRNA	AGAGGGATATGAAGCACCGCCA	548	22	SEQ ID NO. 89
12SrRNA	AGAGGTGGTGAGGTTGATCGGG	598	22	SEQ ID NO. 90
12SrRNA	TCAGGGTTTGCTGAAGATGGCG	636	22	SEQ ID NO. 91
12SrRNA	TTGCCACCTCATGGGCTACACC	707	22	SEQ ID NO. 92
12SrRNA	GGCGGTGTGTACGCGCTTCAGG	840	22	SEQ ID NO. 93
12SrRNA	CCATGTTACGACTTGTCTCCTCT	923	23	SEQ ID NO. 94
16SrRNA	GTCTGGTAGTAAGGTGGAGTGG	41	22	SEQ ID NO. 95
16SrRNA	TTCATCTTCCCTTGCGGTACT	132	22	SEQ ID NO. 96

FIG. 2C

8/26

rRNA species	Sequence	Start	Length	SEQ ID NO:
16SrRNA	CTGTTCTTAGGTAGCTCGTCTGG	262	23	SEQ ID NO. 97
16SrRNA	AACCAGCTATCACCAGGCTCGG	351	22	SEQ ID NO. 98
16SrRNA	CCTAGTGTCCAAAGAGCTGTTCC	459	23	SEQ ID NO. 99
16SrRNA	TGGGTGTTGAGCTTGAACGCTT	550	22	SEQ ID NO. 100
16SrRNA	TCTGACGCAGGCTTATGCGGAG	680	22	SEQ ID NO. 101
16SrRNA	TGAGCATGCCTGTGTTGGGTTG	774	22	SEQ ID NO. 102
16SrRNA	CCCTCGTGGAGCCATTCATACA	977	22	SEQ ID NO. 103
16SrRNA	TCGTCTTGCTGTGTTATGCCCG	1050	22	SEQ ID NO. 104
16SrRNA	ACTGCTCGGAGGTTGGGTTCTG	1180	22	SEQ ID NO. 105
16SrRNA	AGGGTAACTTGTTCCGTTGGTC	1260	22	SEQ ID NO. 106
16SrRNA	TTAATAGCGGCTGCACCATCGG	1359	22	SEQ ID NO. 107
16SrRNA	CTCCGGTCTGAACTCAGATCAC	1412	22	SEQ ID NO. 108
16SrRNA	GGGGAAGGCGCTTTGTGAAGTA	1501	22	SEQ ID NO. 109
16SrRNA	ACCCTGTTCTTGGGTGGGTGTG	1557	22	SEQ ID NO. 110

FIG. 2D

9/26

rRNA	Left Primer	SEQ ID NO:	Right primer	SEQ ID NO:	Product size	Left	Right
28SrRNA	GACCTCAGATCAGACGTGGCGA	SEQ ID NO. 111	AGCAACCCGACTCCGGGAAGA	SEQ ID NO. 191	278	4,22	281,21
28SrRNA	GTCGGGTGCTTGGGAATGC	SEQ ID NO. 112	GACCCGCGCGCGGTTGAAT	SEQ ID NO. 192	206	271,20	476,20
28SrRNA	TCCGCCCCGAGGATTCAA	SEQ ID NO. 113	CGCACCCGCTGGAAATGC	SEQ ID NO. 193	249	445,18	693,18
28SrRNA	GCATTTCCAGGCGGTGCG	SEQ ID NO. 114	GATTCGGCGAGTGCTGCT	SEQ ID NO. 194	255	676,18	930,18
28SrRNA	AGCAGCACTCGCCGAATC	SEQ ID NO. 115	TGGCCCCGAGAGAACCCTC	SEQ ID NO. 195	306	913,18	1218,18
28SrRNA	ACCCCTCCACACGGCGCGA	SEQ ID NO. 116	CGTGTTCAGACGCGGTCGG	SEQ ID NO. 196	251	1058,18	1308,20
28SrRNA	CACGGACCAAGGAGTCTAACA	SEQ ID NO. 117	CCAGGCATAGTTCACCATCTTT	SEQ ID NO. 197	219	1305,21	1523,22
28SrRNA	CGAAAGATGGTGAACATATGCCT	SEQ ID NO. 118	GCCCCAAGACCTCTAATCATTC	SEQ ID NO. 198	242	1500,22	1741,22
28SrRNA	CCGAAACGATCTCAACCTATTTC	SEQ ID NO. 119	TGTTACACACTCCTTAGCGGATT	SEQ ID NO. 199	250	1741,22	1990,23
28SrRNA	CACCTGCCGAATCAACTAGCCC	SEQ ID NO. 120	AGGCTTCAAGGCTCACCGCA	SEQ ID NO. 201	295	1994,22	2288,20
28SrRNA	CCAGCCTCCTCCCTCCTCCT	SEQ ID NO. 121	CAGGACCGACTGACCCATGTTTC	SEQ ID NO. 201	244	2182,19	2425,22
28SrRNA	GAACAGCAGTTGAACATGGGT	SEQ ID NO. 122	CTGCCCTTCACAAAGAAAAGAG	SEQ ID NO. 202	233	2393,21	2625,22
28SrRNA	CGGGGAGAGTTCTCTTTTCTTT	SEQ ID NO. 123	GCCGACTTCCCTTACCTACATT	SEQ ID NO. 203	241	2593,22	2833,22
28SrRNA	CAATGTAGGTAAGGGAAGTCGG	SEQ ID NO. 124	GAGAGAGAGAGAGAGAGGGCG	SEQ ID NO. 204	231	2811,22	3041,21
28SrRNA	CTCTCTCTCTCCCCCGCTCC	SEQ ID NO. 125	AGCTGGGGCGATCCACGGGA	SEQ ID NO. 205	220	3031,20	3250,20
28SrRNA	TCCCGTGGATCGCCCCAGCT	SEQ ID NO. 126	CCGAGGAGGAGGGGGGAACG	SEQ ID NO. 206	225	3231,20	3455,20
28SrRNA	GGCGGGCGGGCGGTTGCT	SEQ ID NO. 127	CGCACCAGTTCTAAGTCGGCTG	SEQ ID NO. 207	208	3381,20	3588,22

FIG. 3A

10/26

rRNA	Left Primer	SEQ ID NO:	Right primer	SEQ ID NO:	Product size	Left	Right
28SrRNA	CCTAGCAGCCGACTTAGAACTG	SEQ ID NO. 128	GACAGTGGGAATCTCGTTTCATC	SEQ ID NO. 208	249	3562,22	3810,22
28SrRNA	TCTCTTAAGGTAGCCAAATGCC	SEQ ID NO. 129	GGCCTCCCACTTATTCTACA	SEQ ID NO. 209	210	3737,22	3946,21
28SrRNA	CTACTATCCAGCGAAACCACAG	SEQ ID NO. 130	AGAGTAGTGGTATTTACCCGGC	SEQ ID NO. 210	213	3816,22	4028,22
28SrRNA	CGGGAAAGAAGACCCTGTTGA	SEQ ID NO. 131	CTCGCCTTAGGACACCTGCGT	SEQ ID NO. 211	336	3865,22	4200,21
28SrRNA	CTCAGGGAGGACAGAAACCTC	SEQ ID NO. 132	AATGATAGGAAGAGCCGACATC	SEQ ID NO. 212	229	4201,21	4429,22
28SrRNA	GATGTCGGCTCTTCCTATCAIT	SEQ ID NO. 133	CGTTCAGTCATAATCCCACAGA	SEQ ID NO. 213	247	4408,22	4654,22
28SrRNA	ACTAATAGGGAACGTGAGCTGG	SEQ ID NO. 134	GGATTCTGACTTAGAGGCGTTC	SEQ ID NO. 214	205	4467,22	4671,22
28SrRNA	CTAAGTCAGAATCCCGCCAGG	SEQ ID NO. 135	GTTAGCGCCAGGTTCCCCAC	SEQ ID NO. 215	267	4658,22	4924,21
28SrRNA	TCGCCGCGCTCTACCTTACCTA	SEQ ID NO. 136	GCCCCGTCGGCATGTATTAGCTC	SEQ ID NO. 216	206	121,22	326,22
28SrRNA	GAGCTAATACATGCCGACGGGC	SEQ ID NO. 137	GAATCGAACCCCTGATCCCCCGT	SEQ ID NO. 217	265	305,22	569,22
28SrRNA	AACGTCTGCCCTATCAACTTTC	SEQ ID NO. 138	TGGACTCATTTCCAATTACAGGG	SEQ ID NO. 218	222	488,22	709,22
28SrRNA	GGCCCTGTAATTGGAATGAGT	SEQ ID NO. 139	GGGACACTCAGCTAAGAGCATC	SEQ ID NO. 219	238	686,21	923,22
28SrRNA	TGCTCTTAGCTGAGTGTCGCCGC	SEQ ID NO. 140	GCCGGTCCAAGGAATTTACACC	SEQ ID NO. 220	215	904,22	1118,21
28SrRNA	TAGAGGTGAAATTCCTTGACC	SEQ ID NO. 141	GTTTCAGCTTTGCAACCATACTC	SEQ ID NO. 221	239	1094,22	1332,23
28SrRNA	TGAAACTTAAAGGAATTGACGGA	SEQ ID NO. 142	AGTTAGCATGCCAGAGTCTCGT	SEQ ID NO. 222	219	1327,23	1545,22
28SrRNA	CACGGACAGGATTGACAGATT	SEQ ID NO. 143	CATCTAAGGGCATCACAGACCT	SEQ ID NO. 223	234	1418,21	1651,22
28SrRNA	CCCAACTTCTTAGAGGGACAAGT	SEQ ID NO. 144	AGCTTATGACCCCGCACTTACTG	SEQ ID NO. 224	233	1576,23	1808,22

FIG. 3B

11/26

rRNA	Left Primer	SEQ ID NO:	Right primer	SEQ ID NO:	Product size	Left	Right
18S rRNA	CTGTGATGCCCTTAGATGTC	SEQ ID NO. 145	CTCACTAAACCATCCAATCGGT	SEQ ID NO. 225	244	1634,21	1877,22
18S rRNA	ATGAACGAGGAATTCCAGTAAG	SEQ ID NO. 146	CCTACGGAAACCTTGTTACGAC	SEQ ID NO. 226	218	1771,23	1988,22
18S rRNA	ACCTGCGGAAGGATCATTAAC	SEQ ID NO. 147	ACGCCCTTCTTTCTCTCTCTCT	SEQ ID NO. 227	242	1992,21	2233,22
16S rRNA	CCCACCTCCACCTTACTACCAGA	SEQ ID NO. 148	GGTCTTAGCTTTGGCTCTCCTT	SEQ ID NO. 228	214	19,22	232,22
16S rRNA	CAAGGAGAGCCAAAGCTAAGAC	SEQ ID NO. 149	AGTGTCCAAAGAGCTGTTCCCTC	SEQ ID NO. 229	247	210,22	456,22
16S rRNA	AGAACCCCTCTAAATCCCCTTGT	SEQ ID NO. 150	ATTGGGTGTGAGGAGTTTCAGTT	SEQ ID NO. 230	205	392,22	596,22
16S rRNA	CAAAGAGGAACAGCTCTTTGGA	SEQ ID NO. 151	CGGAGGAGAAATGTTTTCATGTT	SEQ ID NO. 231	233	431,22	663,22
16S rRNA	TAACTGAACCTCCTCACACCCAA	SEQ ID NO. 152	CTTTCCTTATGAGCATGCCTGT	SEQ ID NO. 232	210	574,22	783,22
16S rRNA	ACAGGCATGCTCATAAAGGAAAG	SEQ ID NO. 153	TAAGAGACAGCTGAACCCCTCGT	SEQ ID NO. 233	231	762,22	992,22
16S rRNA	ACCGTGCAAAAGGTAGCATAATC	SEQ ID NO. 154	GCAGGTTTGGTAGTTTAGGACC	SEQ ID NO. 234	216	912,22	1127,22
16S rRNA	AGCAAGACGAGAAGACCCCTATG	SEQ ID NO. 155	CTAGGGTAACCTTGTTCCGTTGG	SEQ ID NO. 235	222	1041,22	1262,22
16S rRNA	TTCACCAGTCAAAGCGAACTAC	SEQ ID NO. 156	GGTAGATAGAAAACCGACCTGGA	SEQ ID NO. 236	245	1193,22	1437,22
16S rRNA	GATCCAATAACTTGACCAACGG	SEQ ID NO. 157	CTCTTGTCCTTTTCGTACAGGGA	SEQ ID NO. 237	244	1226,22	1469,22
12S rRNA	AAGATTACACATGCAAGCATCC	SEQ ID NO. 158	TGACTTGGGTTAATCGTGTGAC	SEQ ID NO. 238	239	38,22	276,22
12S rRNA	GTCACACGATTAAACCCCAAGTCA	SEQ ID NO. 159	GTAGTGTTCTGGCGAGCAGTT	SEQ ID NO. 239	245	255,22	499,21
12S rRNA	GATACCCCACTATGCTTAGCCC	SEQ ID NO. 160	CTTGCGCTTACTTTGTAGCCCTT	SEQ ID NO. 240	229	432,22	660,22
12S rRNA	TTCTGTAAATCGATAAACCCCGCA	SEQ ID NO. 161	ATCCACCTTCGACCCCTTAAGTT	SEQ ID NO. 241	218	560,22	777,22

FIG. 3C

12/26

rRNA	Left Primer	SEQ ID NO:	Right primer	SEQ ID NO:	Product size	Left	Right
12SrRNA	GATCAACCTCACCACCTCTTG	SEQ ID NO. 162	GCCCTGTTCAACTAAGCACTCT	SEQ ID NO. 242	239	580,21	818,22
12SrRNA	GGCTACATTTTCTACCCCGAGAA	SEQ ID NO. 163	TGCACCTTCCAGTACACTTACCA	SEQ ID NO. 243	230	714,22	943,23
5_8SrRNA	GACTCTTAGCGGTGGATCACTC	SEQ ID NO. 164	TTCGAAAGTGTGCGATGATCAATGT	SEQ ID NO. 244	103	2,22	104,23
5SrRNA	GTCTACGGCCATACCACCCT	SEQ ID NO. 165	CGGTCTCCCATCCAAGTACTAA	SEQ ID NO. 245	93	1,20	93,22

FIG. 3D

13/26

		Section 44									
		(2796)	2796	2810	2820	2830	2840	2850	2860		
gi_36013_emb_X69341.1_HSRGE5	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36014_emb_X69357.1_HSRGE6	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36015_emb_X69358.1_HSRGE9	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36021_emb_X69348.1_HSRHFEF11	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36022_emb_X69347.1_HSRHFEF3	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36023_emb_X69356.1_HSRHFEF5	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36024_emb_X69349.1_HSRHFEF6	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36025_emb_X69350.1_HSRHFEF7	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36026_emb_X69351.1_HSRHFEF8	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36039_emb_X69354.1_HSRHT29	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36040_emb_X69352.1_HSRHT2918	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36041_emb_X69359.1_HSRHT294	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36066_emb_X69361.1_HSRLAN52	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36067_emb_X69353.1_HSRLAN56	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36068_emb_X69360.1_HSRLAN58	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36071_emb_X69364.1_HSRMRC58	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
gi_36072_emb_X69362.1_HSRMRC59	(1)	---	---	---	---	---	---	---	---	GGATCCGTAAC	TTCCGGGATAAG
28SrRNA (2796)		TCTGGCATGTTGGAACAATGTAGGTAAGGGAAGTCGGCAAGCCCGATCCGTAAC	TTCCGGGATAAG								
Consensus (2796)										GGATCCGTAAC	TTCCGGGATAAG

FIG. 4A

	(2861)	2861	2870	2880	2890	2900	2910	2925
								Section 45
gi_36013_emb_X69341.1_HSRGE5	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						2925
gi_36014_emb_X69357.1_HSRGE6	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36015_emb_X69358.1_HSRGE9	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36021_emb_X69348.1_HSRHFEF11	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36022_emb_X69347.1_HSRHFEF3	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36023_emb_X69356.1_HSRHFEF5	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36024_emb_X69349.1_HSRHFEF6	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36025_emb_X69350.1_HSRHFEF7	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36026_emb_X69351.1_HSRHFEF8	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36039_emb_X69354.1_HSRHT29	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36040_emb_X69352.1_HSRHT2918	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36041_emb_X69359.1_HSRHT294	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36066_emb_X69361.1_HSRLAN52	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36067_emb_X69353.1_HSRLAN56	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36068_emb_X69360.1_HSRLAN58	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36071_emb_X69364.1_HSRMRC58	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
gi_36072_emb_X69362.1_HSRMRC59	(23)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
28SrRNA	(2861)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						
Consensus	(2861)	GATTGGCTCTAAGGGCTGGGTCTGGTCTGGGCTGGGGCGCGAAGCGGGGCTGGGCGCGCGCGGGC						

FIG. 4B

[illegible]

FIG. 4C

	(3056)	3056	3070	3080	3090	3100	3110	3120	Section 48
gi_36013_emb_X69341.1_HSRGE5	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36014_emb_X69357.1_HSRGE6	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36015_emb_X69358.1_HSRGE9	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36021_emb_X69348.1_HSRHFEF11	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36022_emb_X69347.1_HSRHFEF3	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36023_emb_X69356.1_HSRHFEF5	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36024_emb_X69349.1_HSRHFEF6	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36025_emb_X69350.1_HSRHFEF7	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36026_emb_X69351.1_HSRHFEF8	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36039_emb_X69354.1_HSRHT29	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36040_emb_X69352.1_HSRHT2918	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36041_emb_X69359.1_HSRHT294	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36066_emb_X69361.1_HSRLAN52	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36067_emb_X69353.1_HSRLAN56	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36068_emb_X69360.1_HSRLAN58	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36071_emb_X69364.1_HSRMRC58	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
gi_36072_emb_X69362.1_HSRMRC59	(218)	GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
28SrRNA (3053)		GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							
Consensus (3056)		GTCTTCCCCCTCCCCGGGGAGCGCCGCTGGGGCGGCGCGGGGGGAGAAAGGTCGGGGCGG							

FIG. 4E

	(3121)	3121	3130	3140	3150	3160	3170	3185
gi_36013_emb_X69341.1_HSRGE5	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36014_emb_X69357.1_HSRGE6	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36015_emb_X69358.1_HSRGE9	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36021_emb_X69348.1_HSRHFEF11	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36022_emb_X69347.1_HSRHFEF3	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36023_emb_X69356.1_HSRHFEF5	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36024_emb_X69349.1_HSRHFEF6	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36025_emb_X69350.1_HSRHFEF7	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36026_emb_X69351.1_HSRHFEF8	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36039_emb_X69354.1_HSRHT29	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36040_emb_X69352.1_HSRHT2918	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36041_emb_X69359.1_HSRHT294	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36066_emb_X69361.1_HSRLAN52	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36067_emb_X69353.1_HSRLAN56	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36068_emb_X69360.1_HSRLAN58	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36071_emb_X69364.1_HSRMRC58	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
gi_36072_emb_X69362.1_HSRMRC59	(283)	CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
28SrRNA (3117)		CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG
Consensus (3121)		CAGGGGGCGG	GCGGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG	CGCGCGGCGG

FIG. 4F

	(3251)	3251	3260	3270	3280	3290	3300	3315
gi_36013_emb_X69341.1_HSRGE5	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36014_emb_X69357.1_HSRGE6	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36015_emb_X69358.1_HSRGE9	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36021_emb_X69348.1_HSRHFEF11	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36022_emb_X69347.1_HSRHFEF3	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36023_emb_X69356.1_HSRHFEF5	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36024_emb_X69349.1_HSRHFEF6	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36025_emb_X69350.1_HSRHFEF7	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36026_emb_X69351.1_HSRHFEF8	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36039_emb_X69354.1_HSRHT29	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36040_emb_X69352.1_HSRHT2918	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36041_emb_X69359.1_HSRHT294	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36066_emb_X69361.1_HSRLAN52	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36067_emb_X69353.1_HSRLAN56	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36068_emb_X69360.1_HSRLAN58	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36071_emb_X69364.1_HSRMRC58	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
gi_36072_emb_X69362.1_HSRMRC59	(283)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
28SrRNA	(3243)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG
Consensus	(3251)	CCCCAGCTGCGGCGGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG	CGTTCGCGGCGGCGG

FIG. 4H

21/26

		Section 52						
		3316	3330	3340	3350	3360	3370	3380
gi_36013_emb_X69341.1_HSRGE5	(470)	AC	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36014_emb_X69357.1_HSRGE6	(471)	CACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36015_emb_X69358.1_HSRGE9	(470)	ACCCGCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36021_emb_X69348.1_HSRHFEF11	(470)	C	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36022_emb_X69347.1_HSRHFEF3	(470)	ACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36023_emb_X69356.1_HSRHFEF5	(471)	CAGGCGCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36024_emb_X69349.1_HSRHFEF6	(471)	CACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36025_emb_X69350.1_HSRHFEF7	(470)	ACGCGCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36026_emb_X69351.1_HSRHFEF8	(470)	ACCGCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36039_emb_X69354.1_HSRHT29	(470)	ACCCGCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36040_emb_X69352.1_HSRHT2918	(471)	ACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36041_emb_X69359.1_HSRHT294	(470)	ACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36066_emb_X69361.1_HSRLAN52	(471)	ACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36067_emb_X69353.1_HSRLAN56	(470)	C	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36068_emb_X69360.1_HSRLAN58	(470)	C	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36071_emb_X69364.1_HSRMRC58	(470)	C	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
gi_36072_emb_X69362.1_HSRMRC59	(471)	ACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
28S rRNA	(3307)	ACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				
Consensus	(3316)	ACCCCA	CCCCACGTCTCG	TCGCGCGCGCGTCCGCTGGGGCGGGGAGCGGTCTGGGGCGGGC				

FIG. 4I

22/26

[illegible]

FIG. 4J

	(3446)	3446	3460	3470	3480	3490	3500	3510	Section 54
gi_36013_emb_X69341.1_HSRGE5	(593)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36014_emb_X69357.1_HSRGE6	(593)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36015_emb_X69358.1_HSRGE9	(600)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36021_emb_X69348.1_HSRHFEF11	(598)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36022_emb_X69347.1_HSRHFEF3	(594)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36023_emb_X69356.1_HSRHFEF5	(598)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36024_emb_X69349.1_HSRHFEF6	(600)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36025_emb_X69350.1_HSRHFEF7	(600)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36026_emb_X69351.1_HSRHFEF8	(598)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36039_emb_X69354.1_HSRHT29	(598)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36040_emb_X69352.1_HSRHT2918	(599)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36041_emb_X69359.1_HSRHT294	(598)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36066_emb_X69361.1_HSRLAN52	(599)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36067_emb_X69353.1_HSRLAN56	(594)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36068_emb_X69360.1_HSRLAN58	(593)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36071_emb_X69364.1_HSRMRC58	(593)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
gi_36072_emb_X69362.1_HSRMRC59	(595)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
28SrRNA	(3435)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	
Consensus	(3446)	CC GTTCCCCCTCCTCTCGGCGCGGGCGGGCGGC						GGCGGCGAGAGGGCGCG	

FIG. 4K

	(3511)	3511	3520	3530	3540	3550	3560	3575
Section 55								
gi_36013_emb_X69341.1_HSRGE5	(648)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36014_emb_X69357.1_HSRGE6	(664)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36015_emb_X69358.1_HSRGE9	(662)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36021_emb_X69348.1_HSRHFEF11	(649)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36022_emb_X69347.1_HSRHFEF3	(656)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36023_emb_X69356.1_HSRHFEF5	(664)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36024_emb_X69349.1_HSRHFEF6	(658)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36025_emb_X69350.1_HSRHFEF7	(653)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36026_emb_X69351.1_HSRHFEF8	(662)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36039_emb_X69354.1_HSRHT29	(659)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36040_emb_X69352.1_HSRHT2918	(663)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36041_emb_X69359.1_HSRHT294	(659)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36066_emb_X69361.1_HSRLAN52	(658)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36067_emb_X69353.1_HSRLAN56	(650)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36068_emb_X69360.1_HSRLAN58	(648)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36071_emb_X69364.1_HSRMRC58	(648)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
gi_36072_emb_X69362.1_HSRMRC59	650)	CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
28SrRNA (3493)		CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						
Consensus (3511)		CGGGCCGGTCCCCCGCGGGTCCGCGGGCCCGCGGTTCGCGGGCCCGCGCTCGCCCTCGGCC						

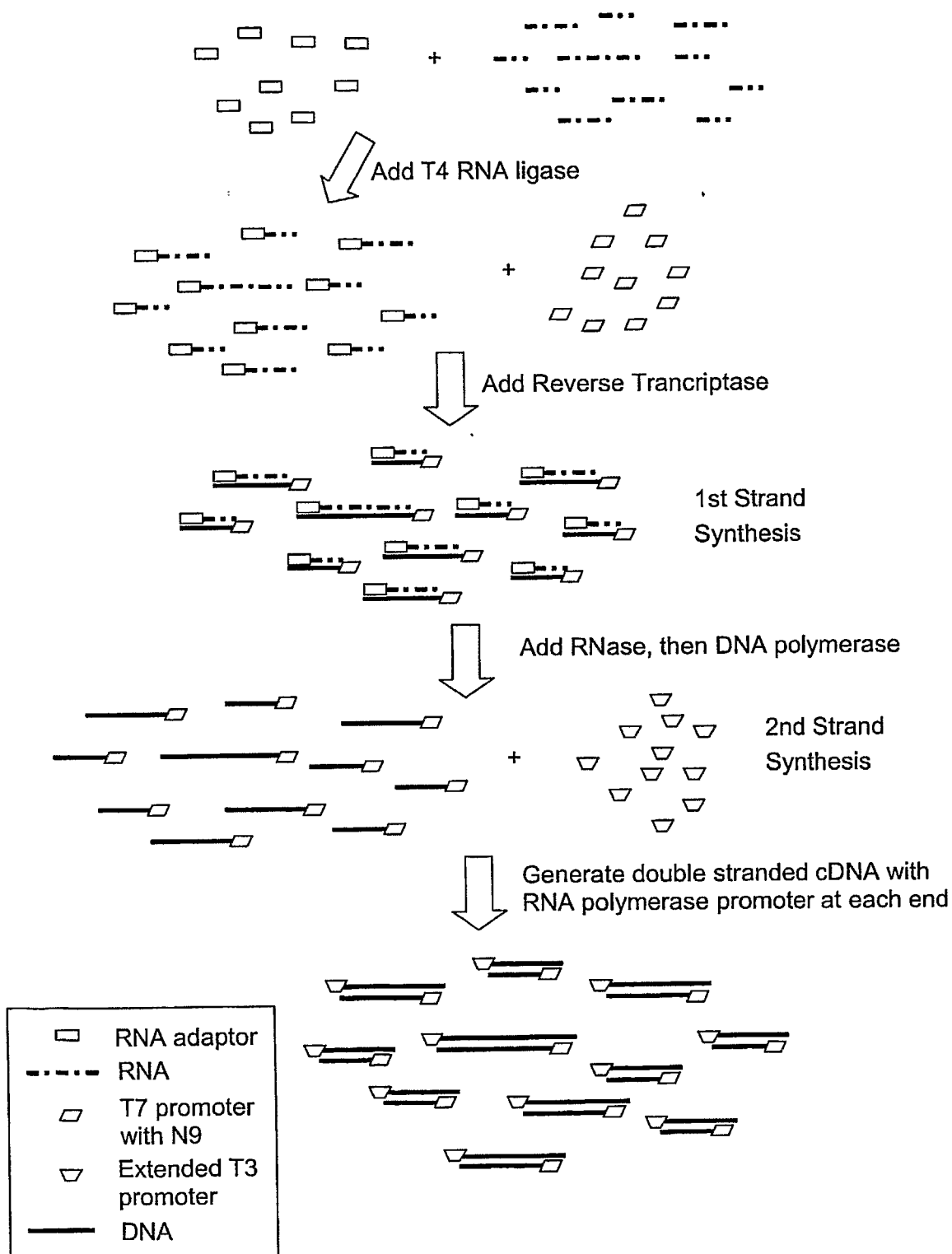
FIG. 4L

Section 56		3630	3620	3610	3600	3590	3576	3576	3576
gi_36013_emb_X69341.1_HSRGE5	SEQ ID NO:166								
gi_36014_emb_X69357.1_HSRGE6	SEQ ID NO:167								
gi_36015_emb_X69358.1_HSRGE9	SEQ ID NO:168								
gi_36021_emb_X69348.1_HSRHFEF11	SEQ ID NO:169								
gi_36022_emb_X69347.1_HSRHFEF3	SEQ ID NO:170								
gi_36023_emb_X69356.1_HSRHFEF5	SEQ ID NO:171								
gi_36024_emb_X69349.1_HSRHFEF6	SEQ ID NO:172								
gi_36025_emb_X69350.1_HSRHFEF7	SEQ ID NO:173								
gi_36026_emb_X69351.1_HSRHFEF8	SEQ ID NO:174								
gi_36039_emb_X69354.1_HSRHT29	SEQ ID NO:175								
gi_36040_emb_X69352.1_HSRHT2918	SEQ ID NO:176								
gi_36041_emb_X69359.1_HSRHT294	SEQ ID NO:177								
gi_36066_emb_X69361.1_HSRLAN52	SEQ ID NO:178								
gi_36067_emb_X69353.1_HSRLAN56	SEQ ID NO:179								
gi_36068_emb_X69360.1_HSRLAN58	SEQ ID NO:180								
gi_36071_emb_X69364.1_HSRMRC58	SEQ ID NO:181								
gi_36072_emb_X69362.1_HSRMRC59	SEQ ID NO:182								
28S rRNA	SEQ ID NO:183								
Consensus	SEQ ID NO:184								

FIG. 4M

26/26

FIG. 5



SEQUENCE LISTING

<110> FREIJE, WADIHA
BRIKUN, IGOR

<120> SUBTRACTIVE SEPARATION AND AMPLIFICATION OF NON-RIBOSOMAL
TRANSCRIBED RNA (nrRNA)

<130> 701046

<150> US 60/705,964

<151> 2005-08-05

<160> 248

<170> PatentIn version 3.3

<210> 1

<211> 5025

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 1

```

cgcgacctca gatcagacgt ggcgacccgc tgaatttaag catattagtc agcggaggaa      60
aagaaactaa ccaggattcc ctcatgaacg gcgagtgaac aggggaagagc ccagcgccga      120
atccccgccc cgcgggggcgc gggacatgtg gcgtacggaa gacccgctcc ccggcgccgc      180
tcgtgggggg ccgaagtcct tctgatcgag gccagcccg tggacggtgt gaggccggta      240
gcggccggcg cgcgcccggg tcttccccga gtcgggttgc ttgggaatgc agcccaaagc      300
gggtggtaaa ctccatctaa ggctaaatac cggcacgaga ccgatagtca acaagtaccg      360
taagggaaaag ttgaaaagaa ctttgaagag agagttcaag agggcgtgaa accgttaaga      420
ggtaaacggg tgggtccgc gcagtccgcc cggaggattc aaccggcgcg cgggtccggc      480
cgtgtcggcg gcccgcgga tcttccccgc cccccgttcc tcccgacccc tccacccgcc      540
ctcccttccc ccgcccggcc tcctcctcct ccccgagggg ggcgggctcc ggcgggtgcg      600
ggggtgggcg ggcggggccg ggggtggggt cggcggggga ccgtcccccg gaccggcgac      660
cggccgcccgc cgggcgcatt tccaggcggt gcgccgcgac cggctccggg acggctggga      720
aggcccggcg gggaaggtag ctcggggggc cccgtccgtc cgtccgtcct cctcctcccc      780
cgtctccgcc ccccgggccc gcgtcctccc tcgggagggc gcgcgggtcg gggcggcggc      840
ggcggcgggc gtggcgggcg cggcgggggc ggcgggaccg aaaccccccc cgagtgttac      900
agcccccccg gcagcagcac tcgccgaatc ccggggccga gggagcgaga cccgtcgccg      960
cgctctcccc cctcccggcg cccacccccg cgggaatccc cgcgaggggg gtctcccccg     1020
gcgcggcgcc ggcgtctcct cgtggggggg ccggggccacc cctccacagg cgcgaccgct     1080
ctccaccccc tcctccccgc gcccccgccc cggcgacggg ggggggtgccg cgcgcgggtc     1140

```

ggggggcggg	gcggaactgtc	cccagtgcg	cccgggcggg	tcgcgccgtc	gggcccgggg	1200
gaggttctct	cggggccacg	cgcggtccc	ccgaagaggg	ggacggcgga	gcgagcgcac	1260
ggggtcggcg	gcgacgtcgg	ctaccacccc	gacccgtctt	gaaacacgga	ccaaggagtc	1320
taacacgtgc	gcgagtcggg	ggctcgcacg	aaagccgccc	tggcgcaatg	aaggtgaagg	1380
ccggcgcgct	cgccggccga	ggtgggatcc	cgaggcctct	ccagtccgcc	gaggggcacc	1440
accggcccgt	ctcgcccgcc	gcgccgggga	ggtggagcac	gagcgcacgt	gttaggaccc	1500
gaaagatggt	gaactatgcc	tgggcagggc	gaagccagag	gaaactctgg	tggaggtccg	1560
tagcggtcct	gacgtgcaaa	tcggtcgtcc	gacctgggta	taggggcgaa	agactaatcg	1620
aaccatctag	tagctggttc	cctccgaagt	ttccctcagg	atagctggcg	ctctcgcaga	1680
cccgaacgat	ctcaacctat	tctcaaactt	taaatgggta	agaagcccgg	ctcgctggcg	1740
tggaagccgg	gtggaatg	agtgcctagt	gggccacttt	tggtaagcag	aactggcgct	1800
gcggggatgaa	ccgaacgccg	ggttaaggcg	ccgatgccg	acgctcatca	gacccagaaa	1860
aaggtgttgg	ttgatataga	cagcaggacg	gtggccatgg	aagtcggaat	ccgctaagga	1920
gtgtgtaaca	actcacctgc	cgaatcaact	agccctgaaa	atggatggcg	ctggagcgtc	1980
gggcccatac	ccggccgtcg	ccggcagtcg	agagtggacg	ggagcggcgg	gggcggcggc	2040
gcgcgcgcgc	gtgtggtgtg	cgtcggaggg	cgccggcgcc	ggcggcggcg	ggggtgtggg	2100
gtccttcccc	cgccccccc	cccacgcctc	ctccctcctt	cccggccacg	ccccgctccc	2160
cgcccccgga	gccccgcgga	gctacgccgc	gacgagtagg	agggccgctg	cggtgagcct	2220
tgaagcctag	ggcgcgggcc	cgggtggagg	ccgccgcagg	tgcagatctt	ggtggtagta	2280
gcaaataattc	aaacgagaac	tttgaaggcc	gaagtggaga	agggttccat	gtgaacagca	2340
gttgaacatg	ggtcagtcgg	tcctgagaga	tgggcgagcg	ccgttccgaa	gggacgggcg	2400
atggcctccg	ttgccctcgg	ccgatcgaaa	gggagtcggg	ttcagatccc	cgaatccgga	2460
gtggcgagga	tgggcgccc	gagggctcca	gtgcggtaac	gcgaccgatc	ccggagaagc	2520
cgccggggagc	cccgggggaga	gttctctttt	ctttgtgaag	ggcagggcg	cctggaatgg	2580
gttcgccccg	agagaggggc	ccgtgccttg	gaaagcgctc	cggttccggc	ggcgctccggt	2640
gagctctcgc	tggcccttga	aaatccgggg	gagaggggtg	aaatctcgcg	ccgggcccgt	2700
cccatatccg	cagcaggtct	ccaaggtgaa	cagcctctgg	catgttgga	caatgtaggt	2760
aagggaaagtc	ggcaagccgg	atccgtaact	tcgggataag	gattggctct	aagggctggg	2820
tcggctcggg	tggggcgcg	agcggggctg	ggcgcgcgcc	gcggctggac	gagggcgcg	2880
ccccccccac	gccccgggga	ccccctcgc	ggccctcccc	cgccccaccc	gcgcgcgcg	2940
						3000

ctcgctccct	ccccaccccg	cgccctctct	ctctctctct	cccccgctcc	ccgtcctccc	3060
ccctccccgg	gggagcgccg	cgtggggggc	cggcgggggg	agaagggtcg	gggcggcagg	3120
ggccgcgcgg	ggccgcgcgg	ggcggccggc	gggggcaggt	ccccgcgagg	ggggccccgg	3180
ggacccgggg	ggccggcggc	ggcgcggact	ctggacgcga	gccggggcct	tcccgtggat	3240
cgccccagct	gcggcggggc	tcgcggccgc	ccccggggag	cccggcggcg	gcgcggcgcg	3300
ccccccaccc	ccaccccacg	tctcggtcgc	gcgcgcgtcc	gctggggggc	ggagcggtcg	3360
ggcggcggcg	gtcggcgggc	ggcggggcgg	ggcggttcgt	ccccccgccc	tacccccccg	3420
gccccgtccg	ccccccgttc	ccccctctct	ctcggcgcgc	ggcggcggcg	gcggcaggcg	3480
gcggaggggc	cgcgggccgg	tccccccgc	cggttcgcgc	cccggggccg	cggttcgcgc	3540
cgcgctctgc	ctcggccggc	gcctagcagc	cgacttagaa	ctggtgcgga	ccaggggaaat	3600
ccgactgttt	aattaaaaca	aagcatcgcg	aaggcccgcg	gcgggtgttg	acgcgatgtg	3660
atctctgccc	agtgtcttga	atgtcaaagt	gaagaaattc	aatgaagcgc	gggtaaacgg	3720
cgggagtaac	tatgactctc	ttaaggtagc	caaatgcctc	gtcatctaata	tagtgacgcg	3780
catgaatgga	tgaacgagat	tcccactgtc	cctacctact	atccagcgaa	accacagcca	3840
agggaaacggg	cttggcggaa	tcagcgggga	aagaagaccc	tggttagactt	gactctagtc	3900
tggcacggtg	aagagacatg	agaggtgtag	aataagtggg	aggcccccg	cgcccccccg	3960
gtgtccccgc	gagggggccc	gggcggggtc	cgcgccctg	cgggccgccc	gtgaaatacc	4020
actactctga	tcgttttttc	actgaccggg	tgaggcgggg	gggcgagccc	gaggggctct	4080
cgcttctggc	gccaaagcgc	cgcccggccg	ggcgcgaccc	gctccgggga	cagtgccagg	4140
tggggagttt	gactggggcg	gtacacctgt	caaacggtaa	cgcaggtgtc	ctaaggcgag	4200
ctcagggagg	acagaaacct	cccgtggagc	agaagggcaa	aagctcgctt	gatcttgatt	4260
ttcagtacga	atacagaccg	tgaaagcggg	gcctcacgat	ccttctgacc	ttttgggttt	4320
taagcaggag	gtgtcagaaa	agttaccaca	gggataactg	gcttgtggcg	gccaaagcgtt	4380
catagcgacg	tcgctttttg	atccttcgat	gtcggctctt	cctatcattg	tgaagcagaa	4440
ttcgccaagc	gttggtttgt	tcaccacta	atagggaaac	tgagctgggt	ttagaccgtc	4500
gtgagacagg	ttagttttac	cctactgatg	atgtgttggt	gccatggtaa	tcctgctcag	4560
tacgagagga	accgcagggt	cagacatttg	gtgtatgtgc	ttggctgagg	agccaatggg	4620
gcgaagctac	catctgtggg	attatgactg	aacgcctcta	agtcagaatc	ccgcccaggc	4680
gaacgatacg	gcagcgccgc	ggagcctcgg	ttggcctcgg	atagccggtc	ccccgcctgt	4740
ccccgcgggc	gggcccggcc	ccccctccac	cgccccgccc	cgggagggcg	cgtgccccgc	4800
cgcgcgccgg	gaccgggggtc	cggtgcggag	tgcccttcgt	cctgggaaac	ggggcgcggc	4860
cggaaaggcg	gccgccccct	cgcccgtcac	gcaccgcacg	ttcgtgggga	acctggcgct	4920

aaaccattcg tagacgacct gcttctgggt cgggggtttcg tacgtagcag agcagctccc 4980
 tcgctgcgat ctattgaaag tcagccctcg acacaagggt ttgtc 5025

<210> 2
 <211> 2235
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<220>
 <221> misc_feature
 <222> (61)..(64)
 <223> "n" may be any nucleotide

<220>
 <221> misc_feature
 <222> (88)..(88)
 <223> "n" may be any nucleotide

<220>
 <221> misc_feature
 <222> (97)..(101)
 <223> "n" may be any nucleotide

<220>
 <221> misc_feature
 <222> (106)..(106)
 <223> "n" may be any nucleotide

<400> 2
 cgctgctcct cccgtcgccg tccgggcccc tccgtccgtc cgtccgtcgt cctcctcgct 60
 nnnncggggc gccgggcccc tctcacngg ccccgnnnn ngtcnnggcc cgtcggggcc 120
 tcgccgcgct ctaccttacc tacctggttg atcctgccag tagcatatgc ttgtctcaaa 180
 gattaagcca tgcattgtcta agtacgcacg gccggtacag tgaaactgcg aatggctcat 240
 taaatcagtt atgggttcctt tggtcgctcg ctctctcct acttgataa ctgtggtaat 300
 tctagagcta atacatgccg acgggcgctg acccccttcg cgggggggat gcgtgcattt 360
 atcagatcaa aaccaaccg gtcagcccct ctccggcccc ggccgggggg cgggcgccgg 420
 cggctttggt gactctagat aacctcgggc cgatcgcacg cccccgtgg cggcgacgac 480
 ccattcgaac gtctgcccta tcaactttcg atggtagtcg ccgtgcctac catggtgacc 540
 acgggtgacg gggaatcagg gttcgattcc ggagagggag cctgagaaac ggctaccaca 600
 tccaaggaag gcagcaggcg cgcaaattac cactccccga cccggggagg tagtgacgaa 660
 aaataacaat acaggactct ttcgaggccc tgtaattgga atgagtccac tttaaactct 720
 ttaacgagga tccattggag ggcaagtctg gtgccagcag ccgcggtaat tccagctcca 780
 atagcgtata ttaaagttgc tgcagttaaa aagctcgtag ttggatcttg ggagcgggcg 840

```

ggcgggtccgc cgcgaggcga gccaccgccc gtccccgccc ctgacctctc ggcgccccct 900
cgatgctctt agctgagtgt cccgcggggc ccgaagcgtt tactttgaaa aaattagagt 960
gttcaaagca ggcccagacc gcctggatac cgagctagg aataatggaa taggaccgcg 1020
gttctatttt gttggttttc ggaactgagg ccatgattaa gagggacggc cgggggcatt 1080
cgtattgcgc cgctagaggt gaaattcctt ggaccggcgc aagacggacc agagcgaaag 1140
catttgccaa gaatgttttc attaatacaag aacgaaagtc ggagggttcga agacgatcag 1200
ataccgtcgt agttccgacc ataaacgatg ccgaccggcg atgcggcggc gttattccca 1260
tgaccgcgcg ggcagcttcc gggaaaccaa agtctttggg ttccgggggg agtatggttg 1320
caaagctgaa acttaaagga attgacggaa gggcaccacc aggagtggag cctgcggtt 1380
aatttgactc aacacgggaa acctcaccg gcccgacac ggacaggatt gacagattga 1440
tagctctttc tcgattccgt ggggtggtgg gcatggccgt tcttagttgg tggagcgatt 1500
tgtctggtta attccgataa cgaacgagac tctggcatgc taactagtta cgcgaccccc 1560
gagcggtcg cgcccccaa cttcttagag ggacaagtgg cgttcagcca cccgagattg 1620
agcaataaca ggtctgtgat gcccttagat gtccggggct gcacgcgcgc tacactgact 1680
ggctcagcgt gtgcctaccc tacgccgga ggcgcgggta acccgttgaa cccattcgt 1740
gatggggatc ggggattgca attattcccc atgaacgagg aattcccagt aagtgcgggt 1800
cataagcttg cgttgattaa gtccctgccc tttgtacaca ccgcccgtcg ctactaccga 1860
ttggatggtt tagtgaggcc ctcggatcgg ccccgccggg gtcggccac ggccctggcg 1920
gagcgtgag aagacggtcg aacttgacta tctagaggaa gtaaaagtcg taacaagggt 1980
tccgtaggtg aacctgcgga aggatcatta acggagcccc gacggcggcc cgggcgggc 2040
ccgcgccgcg cttccctccg cacaccacc ccccaccgc gacggcgcgt gcgggcgggg 2100
ccgtgcccgt tcgttcgctc gctcgttcgt tcgccgccc gcccggccgc gagagccgag 2160
aactcgggag ggagacgggg gagagagaga gagagagaga gagagagaga gagagagaga 2220
gaaagaaggg cgtgt 2235

```

<210> 3
<211> 156
<212> DNA
<213> Artificial

<220>
<223> Synthetic

```

<400> 3
cgactcttag cgggtggatca ctcggctcgt gcgtcgatga agaacgcagc tagctgcgag 60
aattaatgtg aattgcagga cacattgatc atcgacactt cgaacgcact tgcggccccg 120

```

ggttcctccc ggggctacgc ctgtctgagc gtcgct

156

<210> 4
<211> 121
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 4
gtctacggcc ataccaccct gaacgcgccc gatctcgtct gatctcggaa gctaagcagg 60
gtcgggcctg gttagtactt ggatgggaga ccgcctggga ataccgggtg ctgtaggcctt 120
t 121

<210> 5
<211> 1558
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 5
gctaaaccta gcccacaaacc cactccacct tactaccaga caaccttagc caaaccattt 60
accacaaataa agtataggcg atagaaattg aaacctggcg caatagatat agtaccgcaa 120
gggaaagatg aaaaattata accaagcata atatagcaag gactaaccctt tataccttct 180
gcataatgaa ttaactagaa ataactttgc aaggagagcc aaagctaaga cccccgaaac 240
cagacgagct acctaagaac agctaaaaga gcacaccgt ctatgtagca aaatagtggg 300
aagatttata ggtagaggcg acaaacctac cgagcctggg gatagctggg tgtccaagat 360
agaatcttag ttcaacttta aatttgccca cagaaccctc taaatcccct tgtaaattta 420
actgttagtc caaagaggaa cagctctttg gacactagga aaaaaccttg tagagagagt 480
aaaaaattta acacccatag taggcctaaa agcagccacc aattaagaaa gcgttcaagc 540
tcaacacca ctacctaaaa aatcccaaac atataactga actcctcaca cccaattgga 600
ccaatctatc accctataga agaactaatg ttagtataag taacatgaaa acattctcct 660
ccgcataagc ctgcgtcaga ttaaaacact gaactgacaa ttaacagccc aatatctaca 720
atcaaccaac aagtcattat taccctcact gtcaacccaa cacaggcatg ctcataagga 780
aagggttaaaa aaagtaaaag gaactcggca aatcttaccg cgctgttta ccaaaaacat 840
cacctctagc atcaccagta ttagaggcac cgcttgccca gtgacacatg tttaacggcc 900
gcggtaccct aaccgtgcaa aggtagcata atcacttggt ccttaaataag ggacctgtat 960
gaatggctcc acgagggttc agctgtctct tacttttaac cagtgaaatt gacctgcccg 1020
tgaagaggcg ggcataacac agcaagacga gaagacccta tggagcttta atttattaat 1080

gcaaacagta cctaacaaac ccacaggtcc taaactacca aacctgcatt aaaaatttcg 1140
 gttggggcga cctcggagca gaaccaacc tccgagcagt acatgctaag acttcaccag 1200
 tcaaagcgaa ctactatact caattgatcc aataacttga ccaacggaac aagttaccct 1260
 agggataaca gcgcaatcct attctagagt ccatatcaac aatagggttt acgacctcga 1320
 tgttgatca ggacatcccg atggtgcagc cgctattaaa gggtcgtttg ttcaacgatt 1380
 aaagtcctac gtgatctgag ttcagaccgg agtaatccag gtcggtttct atctaccttc 1440
 aaattcctcc ctgtacgaaa ggacaagaga aataaggcct acttcacaaa gcgccttccc 1500
 ccgtaaataga tatcatctca acttagtatt ataccacac ccaccaaga acagggtt 1558

<210> 6
 <211> 953
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 6
 aatagggtttg gtcctagcct ttctattagc tcttagtaag attacacatg caagcatccc 60
 cgttccagtg agttcacctt ctaaatacacc acgatcaaaa ggaacaagca tcaagcacgc 120
 agcaatgcag ctcaaaacgc ttagcctagc cacaccccca cgggaaacag cagtgattaa 180
 ccttttagcaa taaacgaaag ttttaactaag ctataactaac cccaggggtg gtcaatttcg 240
 tgccagccac cgcggtcaca cgattaaccc aagtcaatag aagccggcgt aaagagtgtt 300
 ttagatcacc ccctcccaa taaagctaaa actcacctga gttgtaaaaa actccagtgt 360
 acacaaaata gactacgaaa gtggccttta catatctgaa cacacaatag ctaagaccca 420
 aactgggatt agatacccca ctatgcttag ccctaaacct caacagttaa atcaacaaaa 480
 ctgctcgcca gaacactacg agccacagct taaaactcaa aggacctggc ggtgcttcat 540
 atccctctag aggagcctgt tctgtaatcg ataaaccccg atcaacctca ccacctcttg 600
 ctgagcctat ataccgccat cttcagcaaa ccctgatgaa ggctacaaag taagcgcaag 660
 taccacgta aagacgttag gtcaagggtg agcccatgag gtggcaagaa atgggctaca 720
 ttttctaccc cagaaaacta cgatagccct tatgaaactt aagggtcgaa ggtggattta 780
 gcagtaaact aagagtagag tgcttagttg aacagggccc tgaagcgcggt acacaccgcc 840
 cgtcaccctc ctcaagtata cttcaaagga catttaacta aaacccttac gcatttatat 900
 agaggagaca agtcgtaaca tggtaagtgt actggaaagt gcacttggac gaa 953

<210> 7
 <211> 21
 <212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 7

acgacggacg gacggacgga c

21

<210> 8

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 8

taggtaaggt agagcgcggc ga

22

<210> 9

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 9

ttcactgtac cggccgtgcg ta

22

<210> 10

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 10

gcccgtcggc atgtattagc tc

22

<210> 11

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 11

tctagagtca ccaaagccgc cg

22

<210> 12

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 12

tggtcaccat ggtaggcacg g

21

<210> 13

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 13

ggtcgggagt gggtaatttg cg

22

<210> 14

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 14

gccctccaat ggatcctcgt ta

22

<210> 15

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 15

ccgctcccaa gatccaacta cg

22

<210> 16

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 16

gcgggacact cagctaagag ca

22

<210> 17

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 17

ttcctagctg cggtatccag gc 22

<210> 18
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 18
ccggtccaag gaatttcacc tc 22

<210> 19
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 19
tctgatcgtc ttcgaacctc cg 22

<210> 20
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 20
gactttgggt tcccggaagc tg 22

<210> 21
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 21
ttaagccgca ggctccactc ct 22

<210> 22
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 22
ctaagaacgg ccatgcacca cc 22

<210> 23
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 23
tctaagaagt tgggggacgc cg

22

<210> 24
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 24
taggcacacg ctgagccagt ca

22

<210> 25
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 25
ctgggaattc ctcgttcatg gg

22

<210> 26
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 26
ccgatccgag ggcctcacta aa

22

<210> 27
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 27
gctccgttaa tgatccttcc gc

22

<210> 28
<211> 21
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 28

cgaacgaacg agcgagcgaa c

21

<210> 29

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 29

gtcgccacgt ctgatctgag

20

<210> 30

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 30

ctcttcctg ttcactcgcc gt

22

<210> 31

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 31

cctcgatcag aaggacttgg gc

22

<210> 32

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 32

tttaccaccc gctttgggct g

21

<210> 33

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 33
acggtttcac gccctcttga ac

22

<210> 34
<211> 18
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 34
cgccgggttg aatcctcc

18

<210> 35
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 35
cctccgggga ggaggaggag g

21

<210> 36
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 36
gcaccgcctg gaaatgcgcc

20

<210> 37
<211> 18
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 37
gattcggcga gtgctgct

18

<210> 38
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 38

agagcgggtcg cgccgtggga g

21

<210> 39
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 39
cgtggcccg agagaacctc

20

<210> 40
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 40
gtccgtgttt caagacgggt cg

22

<210> 41
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 41
cgccggcctt caccttcatt

20

<210> 42
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 42
tcgggtccta acacgtgcgc t

21

<210> 43
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 43
tttcgccct ataccaggt cg

22

<210> 44
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 44
tcggcccca gacctcta ca

22

<210> 45
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 45
actaggcact cgcattccac cc

22

<210> 46
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 46
gggtctgatg agcgtcggca t

21

<210> 47
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 47
gggctagttg attcggcagg tg

22

<210> 48
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 48
ccctccgacg cacaccacac

20

<210> 49
<211> 21
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 49

aggcttcaag gctcaccgca g

21

<210> 50

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 50

cccttctcca cttcggcctt ca

22

<210> 51

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 51

tcgcccattct ctcaggaccg a

21

<210> 52

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 52

ttcgggggatc tgaacccgac t

21

<210> 53

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 53

cgccctgccc ttcacaaaga aa

22

<210> 54

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 54

ttttcaaggg ccagcgagag c

21

<210> 55

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 55

tgttccaaca tgccagaggc tg

22

<210> 56

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 56

cgacccagcc cttagagcca a

21

<210> 57

<211> 19

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 57

gtggggaggg agcgagcgg

19

<210> 58

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 58

ctgccgccc gacccttctc

20

<210> 59

<211> 19

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 59

cggtctgcgt ccagagtcc

19

<210> 60
<211> 19
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 60
ccgaggagga ggggggaac

19

<210> 61
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 61
cgcaccagtt ctaagtcggc tg

22

<210> 62
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 62
ttgacattca gagcactggg ca

22

<210> 63
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 63
cgttcatcca ttcattgcgcg tc

22

<210> 64
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 64
tgattccgcc aagcccgttc

20

<210> 65
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 65
cgggggcctc ccacttattc ta

22

<210> 66
<211> 21
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 66
cgccctcaccg ggtcagtgaa a

21

<210> 67
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 67
tcaaactccc cacctggcac tg

22

<210> 68
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 68
ccttctgctc cacgggaggt tt

22

<210> 69
<211> 21
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 69
atcgtgaggc cccgctttca c

21

<210> 70
<211> 22
<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 70

cgacgtcgct atgaacgctt gg

22

<210> 71

<211> 22

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 71

gtgaacaatc caacgcttgg cg

22

<210> 72

<211> 22

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 72

tctgaacctg cggttcctct cg

22

<210> 73

<211> 22

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 73

gcctgggcgg gattctgact ta

22

<210> 74

<211> 20

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 74

gtttcccagg acgaagggca

20

<210> 75

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 75
cgacccagaa gcaggtcgtc ta

22

<210> 76
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 76
caaacccttg tgtcgagggc tg

22

<210> 77
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 77
tcatcgacgc acgagccgag

20

<210> 78
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 78
gcaagtgcgt tcgaagtgtc g

21

<210> 79
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 79
agcgacgctc agacaggcgt a

21

<210> 80
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 80

ctcttagcgg tggatcactc gc 22

<210> 81
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 81
gtcgatgatc aatgtgtcct gcaa 24

<210> 82
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 82
agcgacgctc agacaggcgt ag 22

<210> 83
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 83
actcactgga acggggatgc ttg 23

<210> 84
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 84
tgcgtgcttg atgcttggtc ct 22

<210> 85
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 85
ggcacgaaat tgaccaaccc tg 22

<210> 86
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 86
acgccggctt ctattgactt gg

22

<210> 87
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 87
ggctcttagct attgtgtgtt cag

23

<210> 88
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 88
tgtggctcgt agtgttctgg cg

22

<210> 89
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 89
agagggatat gaagcaccgc ca

22

<210> 90
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 90
agaggtggtg aggttgatcg gg

22

<210> 91
<211> 22
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 91

tcagggtttg ctgaagatgg cg

22

<210> 92

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 92

ttgccacctc atgggctaca cc

22

<210> 93

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 93

ggcgggtgtgt acgcgcttca gg

22

<210> 94

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 94

ccatgttacg acttgtctcc tct

23

<210> 95

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 95

gtctggtagt aaggtggagt gg

22

<210> 96

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 96
ttcatctttc ccttgcggtg ct

22

<210> 97
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 97
ctgttcttag gtagctcgtc tgg

23

<210> 98
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 98
aaccagctat caccaggctc gg

22

<210> 99
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 99
cctagtgtcc aaagagctgt tcc

23

<210> 100
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 100
tgggtgttga gcttgaacgc tt

22

<210> 101
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 101

tctgacgcag gcttatgcgg ag

22

<210> 102
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 102
tgagcatgcc tgtgttggtg tg

22

<210> 103
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 103
ccctcgtgga gccattcata ca

22

<210> 104
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 104
tcgtcttgct gtgttatgcc cg

22

<210> 105
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 105
actgctcgga ggttggttc tg

22

<210> 106
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 106
agggtaactt gttccgttgg tc

22

<210> 107
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 107
ttaatagcgg ctgcaccatc gg

22

<210> 108
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 108
ctccggtctg aactcagatc ac

22

<210> 109
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 109
ggggaaggcg ctttgtgaag ta

22

<210> 110
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 110
accctgttct tgggtgggtg tg

22

<210> 111
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 111
gacctcagat cagacgtggc ga

22

<210> 112
<211> 20
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 112

gtcgggttgc ttgggaatgc

20

<210> 113

<211> 18

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 113

tccgcccgga ggattcaa

18

<210> 114

<211> 18

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 114

gcatttccag gcggtgcg

18

<210> 115

<211> 18

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 115

agcagcactc gccgaatc

18

<210> 116

<211> 18

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 116

accctccca cggcgca

18

<210> 117

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 117

cacggaccaa ggagtctaac a

21

<210> 118

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 118

cgaaagatgg tgaactatgc ct

22

<210> 119

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 119

ccgaaacgat ctcaacctat tc

22

<210> 120

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 120

cacctgccga atcaactagc cc

22

<210> 121

<211> 19

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 121

ccacgcctcc tcccctcct

19

<210> 122

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 122

gaacagcagt tgaacatggg t 21

<210> 123
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 123
cggggagagt tctcttttct tt 22

<210> 124
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 124
caatgtaggt aagggaagtc gg 22

<210> 125
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 125
ctctctctct cccccgctcc 20

<210> 126
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 126
tcccgtaggt cgccccagct 20

<210> 127
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 127
ggcggggcgg ggcgggttcgt 20

<210> 128
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 128
cctagcagcc gacttagaac tg

22

<210> 129
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 129
tctcttaagg tagccaaatg cc

22

<210> 130
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 130
ctactatcca gcgaaaccac ag

22

<210> 131
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 131
cggggaaaga agaccctggt ga

22

<210> 132
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 132
ctcagggagg acagaaacct c

21

<210> 133
<211> 22
<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 133

gatgtcggct cttcctatca tt

22

<210> 134

<211> 22

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 134

actaataggg aacgtgagct gg

22

<210> 135

<211> 22

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 135

ctaagtcaga atcccgccca gg

22

<210> 136

<211> 22

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 136

tcgccgcgct ctaccttacc ta

22

<210> 137

<211> 22

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 137

gagctaatac atgccgacgg gc

22

<210> 138

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 138

aacgtctgcc ctatcaactt tc

22

<210> 139

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 139

ggccctgtaa ttggaatgag t

21

<210> 140

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 140

tgctcttagc tgagtgtccc gc

22

<210> 141

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 141

tagaggtgaa attccttgga cc

22

<210> 142

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 142

tgaaacttaa aggaattgac gga

23

<210> 143

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 143

cacggacagg attgacagat t

21

<210> 144
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 144
cccaacttct tagagggaca agt

23

<210> 145
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 145
ctgtgatgcc cttagatgtc c

21

<210> 146
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 146
atgaacgagg aattcccagt aag

23

<210> 147
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 147
acctgcggaa ggatcattaa c

21

<210> 148
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 148
cccactccac cttactacca ga

22

<210> 149
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 149
caaggagagc caaagctaag ac

22

<210> 150
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 150
agaaccctct aaatcccctt gt

22

<210> 151
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 151
caaagaggaa cagctctttg ga

22

<210> 152
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 152
taactgaact cctcacaccc aa

22

<210> 153
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 153
acaggcatgc tcataaggaa ag

22

<210> 154
<211> 22
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 154

accgtgcaaa ggtagcataa tc

22

<210> 155

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 155

agcaagacga gaagacccta tg

22

<210> 156

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 156

ttcaccagtc aaagcgaact ac

22

<210> 157

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 157

gatccaataa cttgaccaac gg

22

<210> 158

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 158

aagattacac atgcaagcat cc

22

<210> 159

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 159

gtcacacgat taacccaagt ca

22

<210> 160

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 160

gatacccccac tatgcttagc cc

22

<210> 161

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 161

ttctgtaatc gataaacccc ga

22

<210> 162

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 162

gatcaacctc accacctctt g

21

<210> 163

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 163

ggctacattt tctaccccag aa

22

<210> 164

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 164

gactcttagc ggtggatcac tc

22

<210> 165
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 165
 gtctacggcc ataccaccct

20

<210> 166
 <211> 759
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 166
 ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
 gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg 120
 ggcaccccc tgcgggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca 180
 ccccgcgccc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
 cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccaggggccc ggcggcgggc 300
 cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccgggggggc 360
 ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
 cggcgtcgcg gccgcccccg gggagccccg cgcccggcgc gccccccca cccccacgtc 480
 tcgtcgcgcg cgcgtccgct gggggcgggg agcggtcggg cggcggcggt cggcgggcgg 540
 cggggcgggg cggttcgtcc ccccgcccta ccccccggc cccgtccgcc ccccttccc 600
 cctcctcct cggcgcgcg cggcggcggc gggcggcgga ggggccgcgg gccggtcccc 660
 cccgccgggt ccgcccccg ggcgcgggt ccgcgcgcgc ctgcctcgg ccggcgcccta 720
 gcagccgact tagaactggg gcgaccagg ggaatccga 759

<210> 167
 <211> 775
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 167
 ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
 gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg 120

```

ggcaccccc tcgcggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca 180
ccccgcgccc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgtgg gggcgggcgc ggggggagaa gggtcggggc gccaggggccc ggcggcgggc 300
cgccgggggc cccggcggcg gggcacggtc ccccgagagg gcccgggcac ccggggggcc 360
ggcgggcgcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
cggcgtcgcg gccgcccccg gggagccccg cggccggcgc gccccccccc cccccccacc 480
ccacgtctcg tcgcgcgcgc gtccgctggg ggcggggagc ggtcgggcgg cggcggtcgg 540
cgggcggcgg ggcggggcgg ttcgtcccc cgccctaccc ccccggcccc gtccgcccc 600
ccttcccccc tcctcctcgg cgcgcggcgg cggcgggcgc ggcggcaggc ggcggagggg 660
ccgcgggccc gtcccccccg ccgggtccgc ccccggggcc gcggttccgc gcgcgcctcg 720
cctcgggcgg cgcctagcag ccgacttaga actggtgcgg accaggggaa tccga 775

```

<210> 168
<211> 773
<212> DNA
<213> Artificial

<220>
<223> Synthetic

```

<400> 168
ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg 120
ggcaccccc tcgcggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca 180
ccccgcgccc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgtgg gggcgggcgc ggggggagaa gggtcggggc gccaggggccc ggcggcgggc 300
cgccgggggc cccggcggcg gggcacggtc ccccgagagg gcccgggcac ccggggggcc 360
ggcgggcgcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
cggcgtcgcg gccgcccccg gggagccccg cggccggcgc gcccccccca ccccccccc 480
acgtctcgtc gcgcgcgcgt ccgctggggg cggggagcgg tcgggcggcg gcggtcggcg 540
ggcgggcggg cggggcggtt cgtccccccg ccctaccccc ccggccccgt ccgccccccc 600
ttccccctc ctctcggcg cgcgggcgcg gcggcggcgg cggcaggcgg cggagggggc 660
gcgggcccgt cccccccgc ggggtccgcc ccggggccgc ggttccgcgc gcgcctcgcc 720
tcggccggcg cctagcagcc gacttagaac tggtgcggac caggggaatc cga 773

```

<210> 169
<211> 760
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 169

```

ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc      60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg      120
ggcaccccc tcgcggccct ccccgcccc accccacgcg cgccgctcgc tccctccccg      180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag      240
cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccaggggcc ggcggcggcc      300
cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccgggggggc      360
ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg      420
cggcgtcgcg gccgcccccg gggagcccgg cgccggcgcg gcccccccc caccacacgt      480
ctcgtcgcgc gcgcgtccgc tgggggcggg gagcggtcgg gcggcggcgg tcggcgggcg      540
gcggggcggg gcggttcgtc ccccgccct acccccccgg ccccgctcgc ccccccttc      600
ccccctctcc tcggcgcgcg gcggcggcgg cgggcggcgg aggggccgcg ggccggtccc      660
ccccgcggg tccgcccccg gggccgcggt tccgcgcgcg cctcgcctcg gccggcgcct      720
agcagccgac ttagaactgg tgcggaccag gggaatccga      760

```

<210> 170

<211> 767

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 170

```

ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc      60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg      120
ggcaccccc tcgcggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca      180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag      240
cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccaggggcc ggcggcggcc      300
cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccgggggggc      360
ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg      420
cggcgtcgcg gccgcccccg gggagcccgg cgccggcgcg gccccccca cccccacccc      480
acgtctcgtc gcgcgcgcgt ccgctggggg cggggagcgg tcgggcggcg gcggtcggcg      540
ggcggcgggg cggggcggtt cgtccccccg ccctaccccc ccggccccgt ccgccccccc      600
ttccccctc ctctcggcg gcggcggcg gcggcggcag gcggcggagg ggccgcgggc      660

```

cggtcccccc cgccgggtcc gccccgggg ccgcggttcc gcgcgcgcct cgcctcggcc 720
ggcgcctagc agccgactta gaactggtgc ggaccagggg aatccga 767

<210> 171
<211> 775
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 171
ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
gaagcggggc tgggcgcgcg ccgcggttg acgaggcgcc gccgcccccc ccacgcccgg 120
ggcaccccc tcgcggccct cccccgcccc accccacgcg cgccgctcgc tccctcccca 180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgttg gggcggcggc ggggggagaa gggtcggggc gccagggggc ggcggcggcc 300
cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccggggggcc 360
ggcggcggcg gcgactcttg acgcgagccg ggcccttccc gtggatcgcc ccagctcgcg 420
cggcgtcgcg gccgcccccg gggagcccgg cggccggcgc gccccccccc cacccccacc 480
ccacgtctcg tcgcgcgcgc gtccgctggg ggcggggagc ggtcgggcgg cggcggtcgg 540
cgggcggcgg ggcggggcgg ttcgtcccc cgccctaccc ccccggcccc gtccgcccc 600
ccttcccccc tcctcctcgg cgcgcggcgg cggcggcggc ggcggcaggc ggcggagggg 660
ccgcggggcg gtcccccccg ccgggtccgc cccgggggcc gcggttccgc gcgcgcctcg 720
cctcggcccg cgcctagcag ccgacttaga actggtgcgg accaggggaa tccga 775

<210> 172
<211> 769
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 172
ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
gaagcggggc tgggcgcgcg ccgcggttg acgaggcgcc gccgcccccc ccacgcccgg 120
ggcaccccc tcgcggccct cccccgcccc accccacgcg cgccgctcgc tccctcccca 180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgttg gggcggcggc ggggggagaa gggtcggggc gccagggggc ggcggcggcc 300
cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccggggggcc 360

ggcggcgggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
 cggcgctcgcg gccgcccccg gggagcccgg cggccggcgc gccccccccc caccaccacc 480
 ccacgtctcg tcgcgcgcgc gtccgctggg ggcggggagc ggtcgggcgg cggcggtcgg 540
 cgggcggcgg ggcggggcgg ttcgtcccc cgccctaccc ccccgggccc gtccgcccc 600
 ccttcccccc tccctcctcg cgcgcgggcg cggcggcggc aggcggcgga ggggccgcgg 660
 gccgggtcccc cccgccgggt ccgcccccg ggcgcgggtt ccgcgcgcgc ctgcctcgg 720
 ccggcgccta gcagccgact tagaactggg gcggaccagg ggaatccga 769

<210> 173
 <211> 764
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 173
 ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
 gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccc ccacgcccgg 120
 ggcaccccc tcgcggccct ccccgcccc accccacgcg cgccgctcgc tccctccccg 180
 ccccgcgccc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
 cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccaggggccc ggcggcggcc 300
 cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccgggggggc 360
 ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
 cggcgctcgcg gccgcccccg gggagcccgg cggccggcgc gccccccca cccaccacc 480
 acgtctctgc gcgcgcgcgt ccgctggggg cggggagcgg tcgggcggcg gcggtcggcg 540
 ggcggcgggg cggggcggtt cgtcccccg ccctaccccc ccggccccgt ccgccccccc 600
 tccccccctc ctctcggcg cgcggcggcg gcggcggggc gcggaggggc cgcgggccgg 660
 tccccccgc cgggtccgcc cccggggccg cggttccgcg cggcctcgc ctcggccggc 720
 gcctagcagc cgacttagaa ctggtgcgga ccaggggaat ccga 764

<210> 174
 <211> 773
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 174
 ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
 gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccc ccacgcccgg 120

```

ggcaccccc tcgcgccct ccccgcccc acccacgcg cgccgctcg tccctcccc 180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgtgg gggcgggcg ggggggagaa gggtcggggc gccagggggc ggcggcggcc 300
cgccgggggc cccggcgggc gggcacggtc ccccgaggag gcccgggcac ccggggggcc 360
ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
cggcgtcgcg gccgccccg gggagcccgg cggccggcgc gccccccca cccccaccc 480
acgtctctgc gcgcgcgcgt ccgctggggg cggggagcgg tcgggcggcg gcggtcggcg 540
ggcggcgggg cggggcggtt cgtcccccg ccctaccccc ccggccccgt ccgccccccc 600
ttccccctc ctctcggcg cgcggcggcg gcggcggcgg cggcaggcgg cggaggggcc 660
gcgggcccgt ccccccgcc gggtcgcgcc ccggggccgc ggttcgcgc gcgcctcgcc 720
tcggccggcg cctagcagcc gacttagaac tggtgcggac caggggaatc cga 773

```

<210> 175
<211> 770
<212> DNA
<213> Artificial

<220>
<223> Synthetic

```

<400> 175
ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgc gccgcccccc ccacgcccgg 120
ggcaccccc tcgcgccct ccccgcccc acccacgcg cgccgctcg tccctcccca 180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgtgg gggcgggcg ggggggagaa gggtcggggc gccagggggc ggcggcggcc 300
cgccgggggc cccggcgggc gggcacggtc ccccgaggag gcccgggcac ccggggggcc 360
ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
cggcgtcgcg gccgccccg gggagcccgg cggccggcgc gccccccca cccccaccc 480
acgtctctgc gcgcgcgcgt ccgctggggg cggggagcgg tcgggcggcg gcggtcggcg 540
ggcggcgggg cggggcggtt cgtcccccg ccctaccccc ccggccccgt ccgccccccc 600
ttccccctc ctctcggcg cgcggcggcg gcggcggcgg cgggcggcgg aggggcccgc 660
ggccggtccc ccccgccggg tccgccccg gggccgcggt tccgcgcgc cctcgccctg 720
gccggcgcct agcagccgac ttagaactgg tgcggaccag gggaatccga 770

```

<210> 176
<211> 774
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 176

```

ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc      60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg      120
ggcaccccc tgcgggccct ccccgcccc accccacgcg cgccgctcgc tccctccccg      180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag      240
cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccaggggcc ggcggcggcc      300
cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccggggggcc      360
ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg      420
cggcgctcgc gccgcccccg gggagccccg cggccggcgc gccccccccc acccccaccc      480
cacgtctcgt cgcgcgcgcg tccgctgggg gcggggagcg gtcggggcgg ggcggtcggc      540
gggcggcggg gcggggcggt tcgtcccccc gccctacccc cccggccccg tccgcccccc      600
cttccccct cctcctcggc gcgcggcggc ggcggcggcg gcggcaggcg gcggaggggc      660
cgcgggcccg tccccccgc cgggtccgcc cccggggccg cggttccgcg cgcgcctcgc      720
ctcggccggc gcctagcagc cgacttagaa ctggtgcgga ccaggggaat ccga          774

```

<210> 177

<211> 770

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 177

```

ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc      60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg      120
ggcaccccc tgcgggccct ccccgcccc accccacgcg cgccgctcgc tccctccccg      180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag      240
cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccaggggcc ggcggcggcc      300
cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccggggggcc      360
ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg      420
cggcgctcgc gccgcccccg gggagccccg cggccggcgc gcccccccca cccccacccc      480
acgtctcgtc gcgcgcgcgt ccgctggggg cggggagcgg tcgggcggcg gcggtcggcg      540
ggcggcgggg cggggcggtt cgtccccccg ccctaccccc cgggccccgt ccgccccccc      600
ttccccctc ctccctcggc gcgcggcggc gcggcgggcg cgggcggcgg aggggcccgcg      660

```

ggccggtccc ccccgccggg tccgcccccg gggccgcggt tccgcgcgcg ctcgcctcg 720
 gccggcgcct agcagccgac ttagaactgg tgcggaccag gggaatccga 770

<210> 178
 <211> 769
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 178
 ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
 gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg 120
 ggcaccccc tgcgggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca 180
 cccgcgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
 cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccagggggc ggcggcggcc 300
 cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccggggggcc 360
 ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
 cggcgtcgcg gccgcccccg gggagccccg cggccggcgc gccccccccc acccccaccc 480
 cacgtctcgt cgcgcgcgcg tccgctgggg gcggggagcg gtcgggcggc ggcggtcggc 540
 gggcggcggg gcggggcggg tcgtcccccc gccctacccc cccggccccg tccgcccccc 600
 cttccccct ctcctcggc gcgcggcggc ggcggcgggc aggcggcgga ggggccgcgg 660
 gccggtcccc cccgccgggt ccgcccccg ggcgcgggtt ccgcgcgcgc ctcgcctcgg 720
 ccggcgccta gcagccgact tagaactggt gcggaccagg ggaatccga 769

<210> 179
 <211> 761
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 179
 ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
 gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg 120
 ggcaccccc tgcgggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca 180
 cccgcgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
 cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccagggggc ggcggcggcc 300
 cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccggggggcc 360

```

ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
cggcgtcgcg gccgcccccg gggagcccgg cggccggcgc gcccccccc caccacacgt 480
ctcgctcgcg gcgctccgc tgggggcggg gagcggtcgg gcggcggcgg tcggcgggcg 540
gcggggcggg gcggttcgtc cccccgccct acccccccg cccgctccgc ccccccttc 600
ccccctcctc ctcggcgcg gcggcggcgg gcggcggcgg gaggggccc gcggcggtcc 660
ccccgcgggt gtccgcccc ggggcccggg ttccgcgcgc gcctgcctc ggccggcgcc 720
tagcagccga cttagaactg gtgcggacca ggggaatccg a 761

```

<210> 180
 <211> 759
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

```

<400> 180
ggatccgtaa cttcgggata aggattggct ctaagggtg ggtcggtcgg gctggggcgc 60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg 120
ggcacccccc tcgcggccct ccccgcccc acccacgcg cgcgctcgc tccctcccca 180
ccccgcgccc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgtgg gggcggcggc ggggggagaa gggtcggggc gccagggggc gcgggcggcc 300
cgccgggggc cccggcggcg gggcacggtc ccccgaggg gcccgggcac ccggggggcc 360
ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
cggcgtcgcg gccgcccccg gggagcccgg cggccggcgc gcccccccc acccacgtc 480
tcgtcgcgcg cgcgtccgtt gggggcgggg agcggtcggg cggcggcggg cggcgggcg 540
cggggcgggg cggttcgtcc ccccccccta ccccccgcc cccgtccgcc ccccttccc 600
ccctcctcct cggcgcgcg cggcggcggc gggcggcgga ggggcccgg gccggtcccc 660
cccgccgggt ccgcccccg ggcgcgggtt ccgcgcgcgc ctcgcctcgg ccggcgcccta 720
gcagccgact tagaactggt gcggaccagg ggaatccga 759

```

<210> 181
 <211> 759
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

```

<400> 181
ggatccgtaa cttcgggata aggattggct ctaagggtg ggtcggtcgg gctggggcgc 60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg 120

```

```

ggcaccccc tcgcggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca 180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgtgg gggcgggcgc ggggggagaa gggtcggggc gccaggggccc ggcggcgggc 300
cgccgggggc cccggcgggc gggcacggtc ccccgaggag gcccgggcac ccgggggggc 360
ggcgggcgcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
cggcgtcgcg gccgcccccg gggagccccg cggccggcgc gcccccccc accccacgtc 480
tcgtcgcgcg cgcgtccgct gggggcgggg agcggtcggg cggcggcggt cggcgggcgg 540
cgggggcgggg cggttcgtcc ccccgcccta ccccccggc cccgtccgcc ccccttccc 600
ccctctctct cggcgcgcgcg cggcgggcgc gggcggcgga ggggcccgcg gccggtcccc 660
cccgccgggt ccgcccccg ggcgcgggtt ccgcgcgcgc ctgcctcgg ccggcgcccta 720
gcagccgact tagaactggt gcggaccagg ggaatccga 759

```

<210> 182
<211> 761
<212> DNA
<213> Artificial

<220>
<223> synthetic

```

<400> 182
ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc 60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg 120
ggcaccccc tcgcggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca 180
ccccgcgcc tctctctctc tctctcccc gctccccgtc ctccccctc cccgggggag 240
cgccgcgtgg gggcgggcgc ggggggagaa gggtcggggc gccaggggccc ggcggcgggc 300
cgccgggggc cccggcgggc gggcacggtc ccccgaggag gcccgggcac ccgggggggc 360
ggcgggcgcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctgcgg 420
cggcgtcgcg gccgcccccg gggagccccg cggccggcgc gcccccccc ccacccacg 480
tctcgtcgcg cgcgcgtccg ctgggggcgg ggagcggtcg ggcggcgggc gtcggcgggc 540
ggcggggcgg ggcggttcgt cccccgccc tcccccccc gcccggtccg ccccccttc 600
ccccctctc ctcggcgcgc ggcggcgggc gcaggcgggc gaggggccgc gggccggtcc 660
ccccgcggg gtccgcccc ggggcccgcg ttccgcgcgc gcctgcctc ggccggcgcc 720
tagcagccga cttagaactg gtgcggacca ggggaatccg a 761

```

<210> 183
<211> 827
<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 183

tctggcatgt tggaacaatg taggtaaggg aagtcggcaa gccggatccg taacttcggg	60
ataaggattg gctctaaggg ctgggtcggg cgggctgggg cgcaagcgg ggctgggagc	120
gcgccgcggc tggacgaggc gcgcgcccc cccacgccc gggcaccccc ctgcggggcc	180
tccccgcgcc caccgcgcg cgccgctcgc tccctcccca cccgcgccc tctctctctc	240
tctctccccc gctccccgtc ctccccctc cccgggggag cgccgctgg gggcgcggcg	300
gggggagaag ggtcggggcg ccaggggccc cgcgggcgcc gccggggcgg ccggcggggg	360
caggtccccg cgaggggggg cccgggcacc cggggggccc gcggcggcgc ggactctgga	420
cgcgagccgg gcccttccc tggatcgccc cagctcgggc gggcgtcgc gccgccccg	480
gggagcccg cgcgggcgcg gcgcgcccc caccgccacc ccacgtctc gtcgcgcgcg	540
cgtccgctgg gggcgggagc ggtcggggcg cgcggtcgg cgggcggcgg ggcggggcg	600
ttcgtcccc cgccctacc cccgggccc gtccgcccc cttcccccc tcctcctcgg	660
cgcgggcgcg cgcgggcggc aggcggcgga ggggcccgg gccgggtccc cccgccgggt	720
ccgcccccg ggccgcggtt ccgcgcgcgc ctgcctcgg ccggcgcta gcagccgact	780
tagaactggg gcggaccagg ggaatccgac tgtttaatta aaacaaa	827

<210> 184

<211> 771

<212> DNA

<213> Artificial

<220>

<223> Synthetic

<400> 184

ggatccgtaa cttcgggata aggattggct ctaagggctg ggtcggtcgg gctggggcgc	60
gaagcggggc tgggcgcgcg ccgcggctgg acgaggcgcc gccgcccccc ccacgcccgg	120
ggcaccccc tcgcggccct ccccgcccc accccacgcg cgccgctcgc tccctcccca	180
ccccgcgcc tctctctctc tctctccccc gctccccgtc ctccccctc cccgggggag	240
cgccgctgg gggcggcggc ggggggagaa gggtcggggc gccaggggccc ggcggcgccc	300
cgccgggggc cccggcggcg gggcacggtc ccccgagggg gcccgggcac ccggggggcc	360
ggcggcggcg gcgactctgg acgcgagccg ggcccttccc gtggatcgcc ccagctcgg	420
cggcgtcgc gccgccccg gggagcccc cgccggcgcc gcccccccc acccccccacc	480
ccacgtctc tcgcgcgcgc gtccgctggg ggcggggagc ggtcgggcgg cgcggtcgg	540
cgggcggcgg ggcggggcg ttcgtcccc cgccctacc cccgggccc gtccgcccc	600

ccttcccccc tcctcctcgg cgcgcggcgg cggcggcggg cgaggcggcg gaggggcccgc 660
 gggccggtcc cccccgccgg gtccgcccc ggggccgcgg ttccgcgcgc gcctcgcctc 720
 ggccggcgcc tagcagccga cttagaactg gtgcggacca ggggaatccg a 771

<210> 185
 <211> 23
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 185
 cttgccgtcc caagcaatgg atg 23

<210> 186
 <211> 24
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 186
 caataggtgt gcgtcagaag cacc 24

<210> 187
 <211> 24
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 187
 ggtctcacag tgttgcccag gctg 24

<210> 188
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 188
 ggaugaacgu aggaagcuug 20

<210> 189
 <211> 26
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<220>
 <221> misc_feature
 <222> (26)..(26)
 <223> "n" may be between 6 and 12 of any nucleotide

<400> 189
 ggaattaata cgactcacta tagggn 26

<210> 190
 <211> 44
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 190
 gcgcgaaatt aaccctcact aaagggatga acgtaggaag cttg 44

<210> 191
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 191
 agcaaccgga ctccgggaag a 21

<210> 192
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 192
 gacccgccgc cgggttgaat 20

<210> 193
 <211> 18
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 193
 cgcaccgcct ggaaatgc 18

<210> 194
 <211> 18
 <212> DNA
 <213> Artificial

<220>
<223> Synthetic

<400> 194
gattcggcga gtgctgct 18

<210> 195
<211> 18
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 195
tggccccgag agaacctc 18

<210> 196
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 196
cgtgtttcaa gacgggtcgg 20

<210> 197
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 197
ccaggcatag ttcaccatct tt 22

<210> 198
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 198
gcccccaagac ctctaatacat tc 22

<210> 199
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 199
tgttacacac tccttagcgg att

23

<210> 200
<211> 20
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 200
aggcttcaag gctcaccgca

20

<210> 201
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 201
caggaccgac tgacccatgt tc

22

<210> 202
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 202
ctgcccttca caaagaaaag ag

22

<210> 203
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 203
gccgacttcc cttacctaca tt

22

<210> 204
<211> 21
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 204
gagagagaga gagagagggc g

21

<210> 205
<211> 20
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 205
agctggggcg atccacggga

20

<210> 206
<211> 20
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 206
ccgaggagga ggggggaacg

20

<210> 207
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 207
cgcaccagtt ctaagtcggc tg

22

<210> 208
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 208
gacagtggga atctcgttca tc

22

<210> 209
<211> 21
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 209
gggcctccca cttattctac a

21

<210> 210

<211> 22
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 210
 agagtagtgg tatttcaccg gc 22

<210> 211
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 211
 ctcgccttag gacacctgcg t 21

<210> 212
 <211> 22
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 212
 aatgatagga agagccgaca tc 22

<210> 213
 <211> 22
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 213
 cgttcagtca taatcccaca ga 22

<210> 214
 <211> 22
 <212> DNA
 <213> Artificial

<220>
 <223> Synthetic

<400> 214
 ggattctgac ttagaggcgt tc 22

<210> 215
 <211> 21
 <212> DNA
 <213> Artificial

<220>
<223> Synthetic

<400> 215
gttttagcgcc aggttcccca c 21

<210> 216
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 216
gcccgtcggc atgtattagc tc 22

<210> 217
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 217
gaatcgaacc ctgattcccc gt 22

<210> 218
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 218
tggactcatt ccaattacag gg 22

<210> 219
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 219
gggacactca gctaagagca tc 22

<210> 220
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 220
gccggtccaa ggaatttcac c 21

<210> 221
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 221
gtttcagctt tgcaaccata ctc 23

<210> 222
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 222
agtttagcatg ccagagtctc gt 22

<210> 223
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 223
catctaaggg catcacagac ct 22

<210> 224
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 224
agcttatgac ccgcacttac tg 22

<210> 225
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 225
ctcactaaac catccaatcg gt 22

<210> 226
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 226
cctacggaaa ccttggttacg ac

22

<210> 227
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 227
acgcccttct ttctctctct ct

22

<210> 228
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 228
ggtcttagct ttggctctcc tt

22

<210> 229
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 229
agtgtccaaa gagctgttcc tc

22

<210> 230
<211> 22
<212> DNA
<213> Artificial

<220>
<223> synthetic

<400> 230
attgggtgtg aggagttcag tt

22

<210> 231

<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 231
cggaggagaa tgttttcatg tt

22

<210> 232
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 232
ctttccttat gagcatgcct gt

22

<210> 233
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 233
taagagacag ctgaaccctc gt

22

<210> 234
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 234
gcaggtttgg tagtttagga cc

22

<210> 235
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 235
ctagggtaac ttgttccggt gg

22

<210> 236
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 236
ggtagataga aaccgacctg ga 22

<210> 237
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 237
ctcttgcct ttcgtacagg ga 22

<210> 238
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 238
tgacttgggt taatcgtgtg ac 22

<210> 239
<211> 21
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 239
gtagtgttct ggcgagcagt t 21

<210> 240
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 240
cttgcgctta cttttagacc tt 22

<210> 241
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 241
atccaccttc gacccttaag tt 22

<210> 242
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 242
gccctgttca actaagcact ct 22

<210> 243
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 243
tgcactttcc agtacactta cca 23

<210> 244
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 244
ttcgaagtgt cgatgatcaa tgt 23

<210> 245
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 245
cggtctccca tccaagtact aa 22

<210> 246
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 246
ggagcttcat ctggacctgg gtc 23

<210> 247
<211> 25
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 247
caaaacacca gtgcaggcca acttg

25

<210> 248
<211> 23
<212> DNA
<213> Artificial

<220>
<223> Synthetic

<400> 248
ttgtaatccc agcactctgg gag

23