



US 20250250631A1

(19) United States

(12) Patent Application Publication

Praveen et al.

(10) Pub. No.: US 2025/0250631 A1

(43) Pub. Date: Aug. 7, 2025

(54) KELCH DOMAIN CONTAINING 7B  
(KLHDC7B) VARIANTS AND USES  
THEREOF(71) Applicant: Regeneron Pharmaceuticals, Inc.,  
Tarrytown, NY (US)(72) Inventors: **Kavita Praveen**, Chester Springs, PA  
(US); **Giovanni Coppola**, Tarrytown,  
NY (US); **Manuel Allen Revez**  
**Ferreira**, Tarrytown, NY (US); **Lauren**  
**Gurski**, Tarrytown, NY (US); **Aris**  
**Baras**, Tarrytown, NY (US); **Meghan**  
**Drummond Samuelson**, Tarrytown,  
NY (US); **Goncalo Abecasis**,  
Tarrytown, NY (US)

(21) Appl. No.: 19/060,224

(22) Filed: Feb. 21, 2025

**Related U.S. Application Data**(63) Continuation of application No. 18/303,193, filed on  
Apr. 19, 2023, now Pat. No. 12,258,627, which is a  
continuation of application No. 17/307,653, filed on  
May 4, 2021, now Pat. No. 11,674,177.(60) Provisional application No. 63/087,245, filed on Oct.  
4, 2020, provisional application No. 63/020,746, filed  
on May 6, 2020.**Publication Classification**(51) **Int. Cl.**

<i>C12Q 1/6874</i>	(2018.01)
<i>A61K 38/00</i>	(2006.01)
<i>C12Q 1/6851</i>	(2018.01)
<i>C12Q 1/6876</i>	(2018.01)
<i>G01N 33/68</i>	(2006.01)

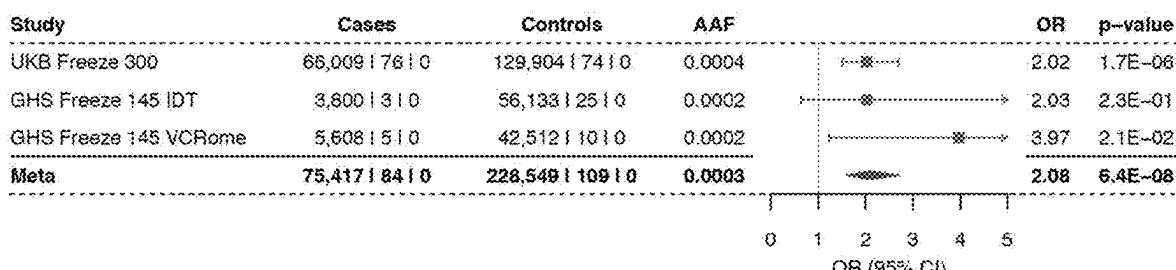
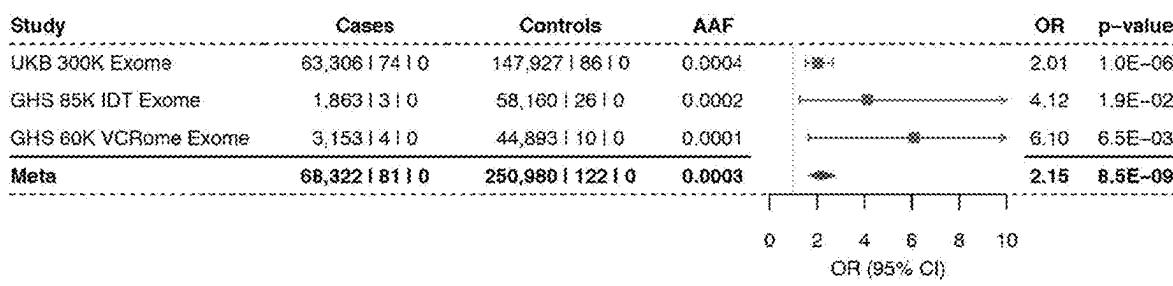
(52) **U.S. Cl.**

CPC .....	<i>C12Q 1/6874</i> (2013.01); <i>C12Q 1/6851</i> (2013.01); <i>C12Q 1/6876</i> (2013.01); <i>G01N</i> <i>33/6818</i> (2013.01); <i>A61K 38/00</i> (2013.01); <i>C12Q 2600/118</i> (2013.01); <i>C12Q 2600/158</i> (2013.01)
-----------	---

(57)

**ABSTRACT**

The present disclosure provides methods of treating subjects having hearing loss, methods of identifying subjects having an increased risk of developing hearing loss, and methods of detecting Kelch Domain Containing 7B (KLHDC7B) variant nucleic acid molecules and variant polypeptides.

**Specification includes a Sequence Listing.****A) Position: 22:50548704:CG:C; cDNA: c.540delG; Amino acid: Lys181fs (truncates at aa 245)**

## A) Position: 22:50548704:CGC; cDNA: c.540delG; Amino acid: Lys181fs (truncates at aa 245)

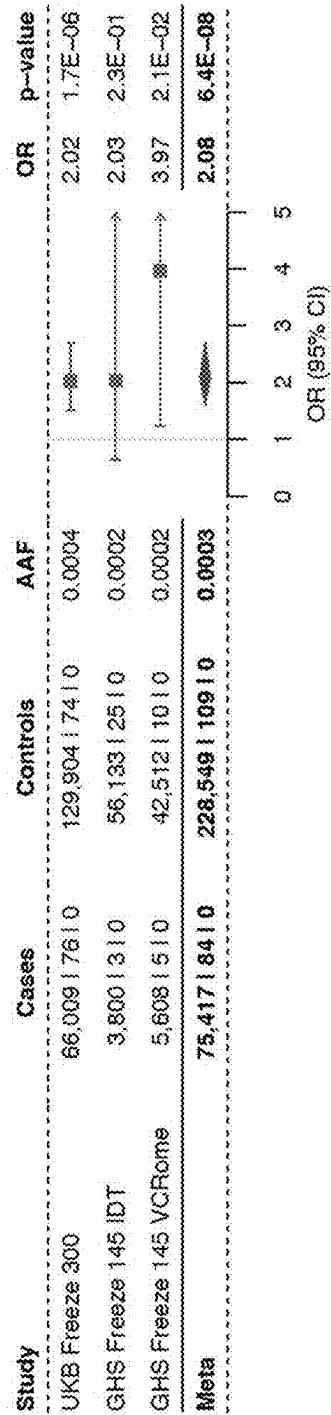
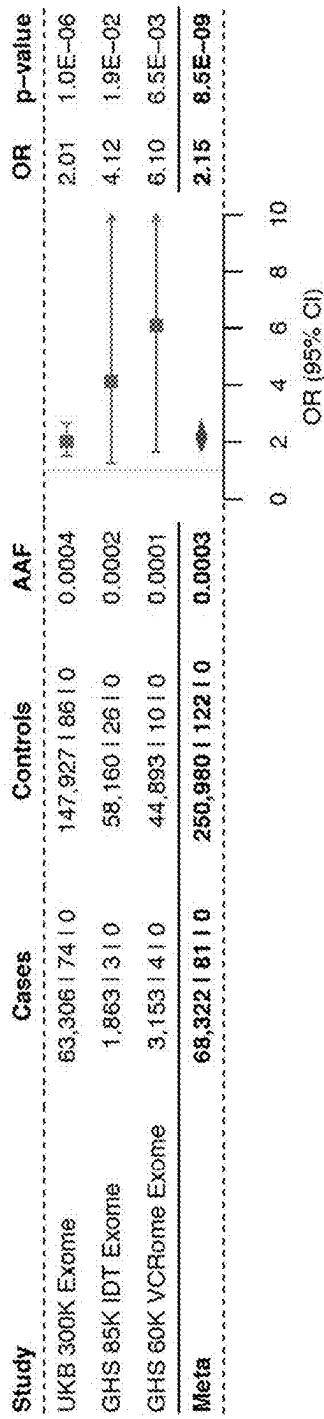


Figure 1

B) Position: 22:50549067:AG:A; cDNA: c.905delG; Amino acid: Gly302fs (truncates at aa 335)

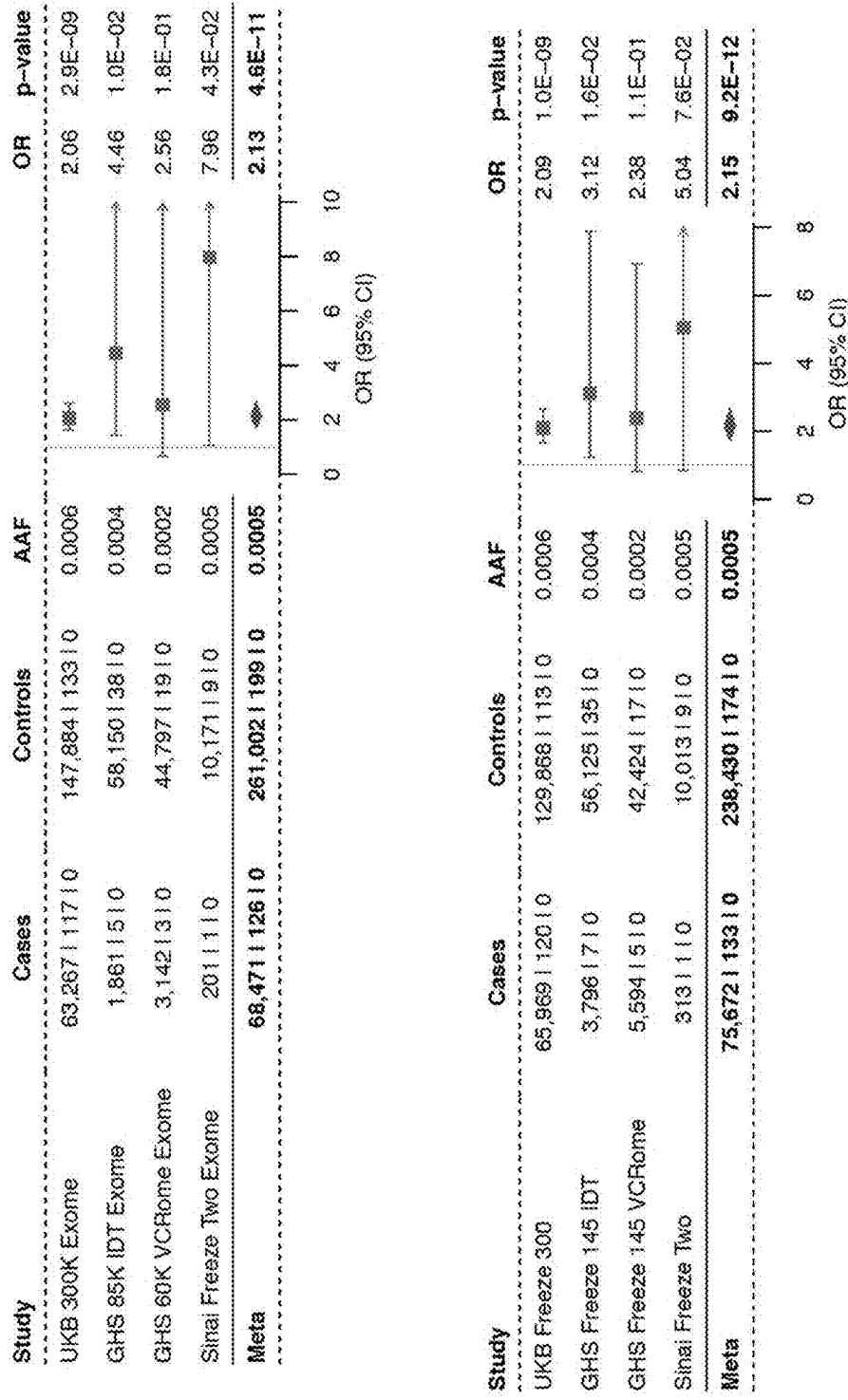


Figure 1 (cont.)

## C) Position: 22:50549676:G&gt;A; cDNA: c.1510G&gt;A; Amino acid: Val504Met

Study	Cases	Controls	AAF		OR	p-value
UKS 30K Imputed	92,009   9,036   283	221,108   19,348   371	0.0434	*	1.14	1.3E-27
CGPS 85K IMI Imputed	1,711   156   2	53,828   4,324   101	0.0410	**	1.11	1.2E-01
CGPS 63K VCFome Imputed	2,915   238   12	41,580   3,370   62	0.0391	**	1.07	1.1E-01
Mammo Freeze Two Imputed	305   28   1	25,624   1,813   37	0.0358	**	1.00	7.1E-01
Sinai Freeze Two Imputed	189   13   0	3,504   667   10	0.0385	**	0.99	4.9E-01
<b>Meta</b>	<b>97,123   9,466   298</b>	<b>251,644   29,522   581</b>	<b>0.0417</b>	<b>*</b>	<b>1.17</b>	<b>3.3E-28</b>

OR (95% CI)

Figure 1 (cont.)

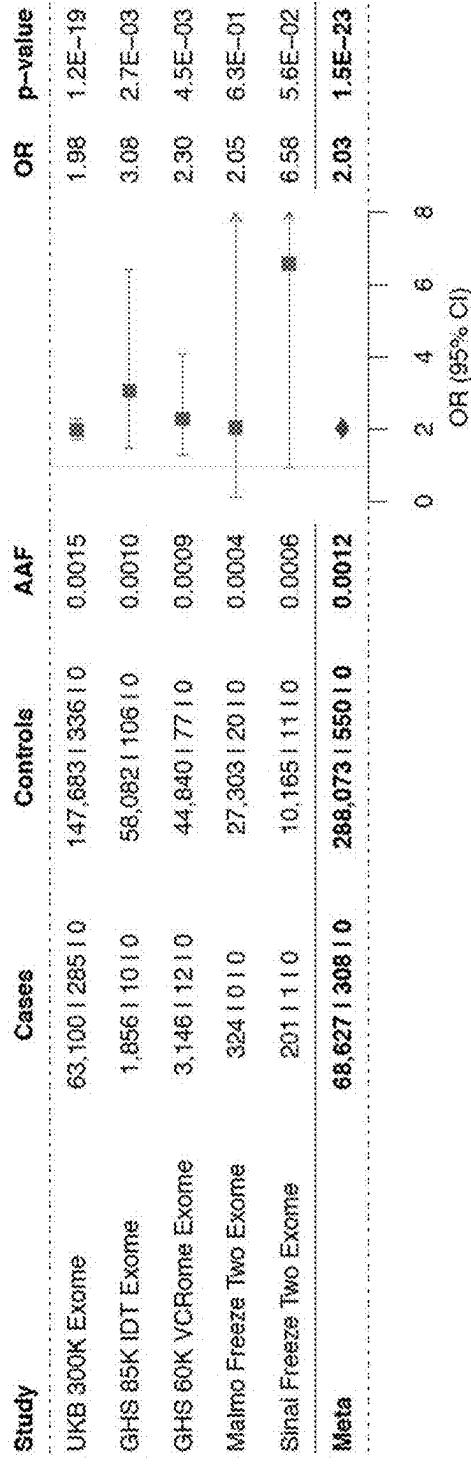
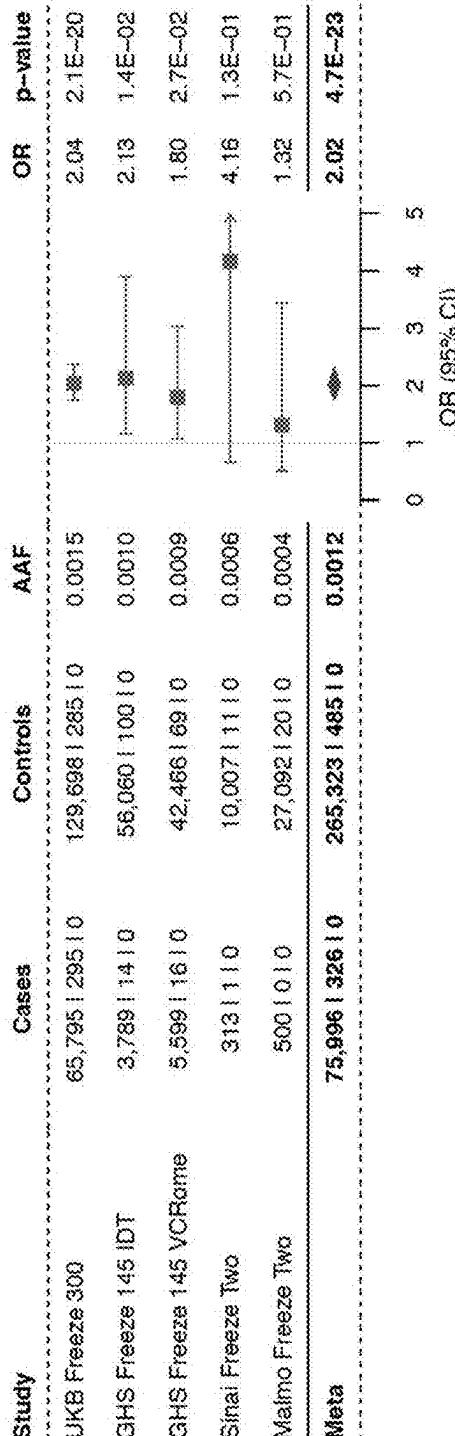
***KLHDC7B M1 (MAF<1%; aggregate of pLOF variants only)******KLHDC7B Burden (MAF<1%; pLOF variants only)***

Figure 2

**KELCH DOMAIN CONTAINING 7B  
(KLHDC7B) VARIANTS AND USES  
THEREOF**

REFERENCE TO SEQUENCE LISTING

[0001] This application includes a Sequence Listing submitted electronically as an XML file named 381204358SEQ, created on Feb. 20, 2025, with a size of 190,219 bytes. The Sequence Listing is incorporated herein by reference.

FIELD

[0002] The present disclosure relates generally to the treatment of subjects having hearing loss, methods of identifying subjects having an increased risk of developing hearing loss, and methods of detecting KLHDC7B variant nucleic acid molecules and variant polypeptides.

BACKGROUND

[0003] Auditory dysfunction in humans is an ongoing problem in the medical fields of otology and audiology. About 300 million people worldwide currently suffer from moderate to severe hearing loss, and this number is expected to increase to 700 million by the year 2015. Auditory dysfunction is a common consequence of aging in Western societies. Approximately 17 percent of Americans have hearing loss and half of that number are under the age of 65. It is predicted that the number of Americans with hearing loss will exceed 70 million by the year 2030.

[0004] Auditory dysfunctions typically arise from both acute and chronic exposures to loud sounds, ototoxic chemicals, and aging. Hearing impairments can be attributed to a wide variety of causes, including infections (e.g., otitis media), genetic predisposition, mechanical injury, tumors, loud sounds or prolonged exposure to noise, aging, and chemical-induced ototoxicity (e.g., antibiotics or platin drugs) that damages neurons and/or hair cells of the peripheral auditory system. This can be caused by acute noise or can be progressive over time. Sounds exceeding 85 decibels can cause hearing loss and is generated by sound sources such as, gun shots, exploding bombs, jet engines, power tools, and musical concerts. Other common everyday activities and products also give rise to high intensity noise such as use of hair dryers, MP3 players, lawn mowers, and blenders. Military personnel are particularly at risk for noise induced hearing loss due to typical military noise exposures. Side effects of noise-induced hearing loss include tinnitus (ringing in the ears), diminished speech understanding, hyperacusis, and various types of auditory processing impairments. Exposures to commonly used medications may also induce auditory dysfunctions. For instance, subjects treated with anticancer therapies, antibiotics, and other medications often develop hearing loss as a side effect. Furthermore, exposure to industrial chemicals and gasses may induce auditory impairments.

[0005] The prevalence of hearing loss after damage to the mammalian cochlea has been thought to be due to a lack of spontaneous regeneration of hair cells and/or neurons, the primary components to detect sound. Humans are born with about 15,000 inner ear hair cells and hair cells do not regenerate after birth. Supporting cells, which surround hair cells in the normal cochlear epithelium, have potential to differentiate into new hair cells in the neonatal mouse following ototoxic damage. Using lineage tracing, the new

hair cells, predominantly outer hair cells, have been shown to arise from Lgr5-expressing inner pillar and third Deiters cells, and new hair cell generation has been shown to incrementally be increased by pharmacological inhibition of Notch.

[0006] Permanent damage to the hair cells of the inner ear results in sensorineural hearing loss, leading to communication difficulties in a large percentage of the population. Hair cells are the receptor cells that transduce the acoustic stimulus. Regeneration of damaged hair cells provide an avenue for the treatment of a condition that currently has no therapies other than prosthetic devices. Although hair cells do not regenerate in the mammalian cochlea, new hair cells in lower vertebrates are generated from epithelial cells, called supporting cells, that surround hair cells.

[0007] Currently, very few cases of hearing loss can actually be cured. Audiological devices such as hearing aids have limitations including the inability to improve speech intelligibility. Of those impacted by hearing impairments, less than 20 percent presently use hearing instruments. In cases of age-related, noise- or drug-induced auditory dysfunctions, often the only effective way to currently "treat" the disorder or reduce its severity is prevention, such as by avoiding excessive noise and using ear protectors, practicing a healthy lifestyle, and avoiding exposure to ototoxic drugs and substances if possible.

[0008] Thus, there remains a long felt need to protect auditory cells before injury and preserve/promote the function of existing cells after injury.

[0009] Kelch Domain Containing 7B (KLHDC7B) is a protein member of the Kelch superfamily, proteins involved in cellular processes such as cytoskeletal rearrangement and protein degradation, and also have roles in extracellular communication, cell morphology, gene expression and actin binding. In addition, members of this superfamily can be co-opted by a virus after an infection. Alterations in this protein superfamily have been associated with various types of cancer, including leukemia, lung, prostate, brain, and Hodgkin's disease. KLHDC7B was identified as being hypermethylated, yet upregulated, in breast cancer cells. Moderate levels of KLHDC7B expression were observed in hair cells of the ear, while outer hair cells seem to show slightly higher expression (gEAR portal).

SUMMARY

[0010] The present disclosure provides methods of identifying a subject having an increased risk for developing hearing loss, wherein the methods comprise: determining or having determined the presence or absence of a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide in a biological sample obtained from the subject; wherein: i) when the subject is KLHDC7B reference, then the subject does not have an increased risk for developing hearing loss; and ii) when the subject is heterozygous or homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide, then the subject has an increased risk for developing hearing loss.

[0011] The present disclosure also provides methods of treating a subject with a therapeutic agent that treats or inhibits hearing loss, wherein the subject has hearing loss, the methods comprising the steps of: determining whether the subject has a KLHDC7B missense variant nucleic acid

molecule encoding a KLHDC7B predicted loss-of-function polypeptide by: i) obtaining or having obtained a biological sample from the subject; and ii) performing or having performed a sequence analysis on the biological sample to determine if the subject has a genotype comprising the KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide; and administering or continuing to administer the therapeutic agent that treats or inhibits hearing loss in a standard dosage amount to a subject that is KLHDC7B reference; and administering or continuing to administer the therapeutic agent that treats or inhibits hearing loss in an amount that is the same as or greater than a standard dosage amount to a subject that is heterozygous or homozygous for the KLHDC7B missense variant nucleic acid molecule; wherein the presence of a genotype having the a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide indicates the subject has an increased risk of developing hearing loss.

[0012] The present disclosure also provides methods of detecting a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide in a subject comprising assaying a sample obtained from the subject to determine whether a nucleic acid molecule in the sample is: i) a genomic nucleic acid molecule comprising a nucleotide sequence: comprising an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof; lacking a guanine at a position corresponding to position 2,807 according to SEQ ID NO:1, or the complement thereof; or lacking a guanine at a position corresponding to position 3,170 according to SEQ ID NO:1, or the complement thereof; ii) an mRNA molecule comprising a nucleotide sequence: comprising an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; comprising an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or the complement thereof; comprising an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or the complement thereof; comprising an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or the complement thereof; lacking a guanine at a position corresponding to position 2,807 according to SEQ ID NO:3, or the complement thereof; lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:4, or the complement thereof; lacking a guanine at a position corresponding to position 2,503 according to SEQ ID NO:5, or the complement thereof; lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:6, or the complement thereof; lacking a guanine at a position corresponding to position 3,170 according to SEQ ID NO:3, or the complement thereof; lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:4, or the complement thereof; lacking a guanine at a position corresponding to position 2,866 according to SEQ ID NO:5, or the complement thereof; or lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:6, or the complement thereof; or iii) a cDNA molecule comprising a nucleotide sequence: comprising an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; comprising an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or the complement thereof; comprising an

adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, or the complement thereof; comprising an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, or the complement thereof; lacking a guanine at a position corresponding to position 2,807 according to SEQ ID NO:11, or the complement thereof; lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:12, or the complement thereof; lacking a guanine at a position corresponding to position 2,503 according to SEQ ID NO:13, or the complement thereof; lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:14, or the complement thereof; lacking a guanine at a position corresponding to position 3,170 according to SEQ ID NO:11, or the complement thereof; lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:12, or the complement thereof; lacking a guanine at a position corresponding to position 2,866 according to SEQ ID NO:13, or the complement thereof; or lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:14, or the complement thereof.

[0013] The present disclosure also provides methods of detecting the presence of a KLHDC7B V1145M, V504M, V405M, K822fs, K181fs, K82fs, G943fs, G302fs, or G203fs variant polypeptide, comprising performing an assay on a sample obtained from a subject to determine whether a KLHDC7B protein in the sample: comprises a methionine at a position corresponding to position 1,145 according to SEQ ID NO:22, comprises a methionine at a position corresponding to position 504 according to SEQ ID NO:23, comprises a methionine at a position corresponding to position 405 according to SEQ ID NO:24, terminates at a position corresponding to position 885 according to SEQ ID NO:43 and lacks amino acids at positions corresponding to positions 886 to 1,235 of SEQ ID NO:19, terminates at a position corresponding to position 244 according to SEQ ID NO:44 and lacks amino acids at positions corresponding to positions 245 to 594 of SEQ ID NO:20, terminates at a position corresponding to position 145 according to SEQ ID NO:45 and lacks amino acids at positions corresponding to positions 146 to 495 of SEQ ID NO:21, terminates at a position corresponding to position 975 according to SEQ ID NO:46 and lacks amino acids at positions corresponding to positions 976 to 1,235 of SEQ ID NO:19, terminates at a position corresponding to position 334 according to SEQ ID NO:47 and lacks amino acids at positions corresponding to positions 335 to 594 of SEQ ID NO:20, and terminates at a position corresponding to position 235 according to SEQ ID NO:48 and lacks amino acids at positions corresponding to positions 236 to 495 of SEQ ID NO:21.

[0014] The present disclosure also provides therapeutic agents that treat or inhibit hearing loss for use in the treatment of hearing loss in a subject having: i) a genomic nucleic acid molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence: comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof; lacks a guanine at a position corresponding to position 2,807 according to SEQ ID NO:1, or the complement thereof; or lacks a guanine at a position corresponding to position 3,170 according to SEQ ID NO:1, or the complement thereof; ii) an mRNA molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises: an adenine at a position

corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or the complement thereof; an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or the complement thereof; an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or the complement thereof; an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof; or iii) a cDNA molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or the complement thereof; an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, or the complement thereof; an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or the complement thereof; an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0015]** FIG. 1 (Panels A, B, and C) shows that a common missense variant, and rare, predicted loss-of-function (pLOF) variants in KLHDC7B are associated with increased risk for hearing loss. The association with the pLOF variants suggest that the missense is likely to be loss of or reduced function, and that reduction in KLHDC7B confers an increase in the risk for hearing loss.

**[0016]** FIG. 2 shows an aggregate of rare (minor allele frequency <1%), pLOF variants in KLHDC7B is associated

with hearing loss. This suggests that there are additional loss of function variants that increase the risk for hearing loss in carriers.

#### DESCRIPTION

**[0017]** Various terms relating to aspects of the present disclosure are used throughout the specification and claims. Such terms are to be given their ordinary meaning in the art, unless otherwise indicated. Other specifically defined terms are to be construed in a manner consistent with the definitions provided herein.

**[0018]** Unless otherwise expressly stated, it is in no way intended that any method or aspect set forth herein be construed as requiring that its steps be performed in a specific order. Accordingly, where a method claim does not specifically state in the claims or descriptions that the steps are to be limited to a specific order, it is in no way intended that an order be inferred, in any respect. This holds for any possible non-expressed basis for interpretation, including matters of logic with respect to arrangement of steps or operational flow, plain meaning derived from grammatical organization or punctuation, or the number or type of aspects described in the specification.

**[0019]** As used herein, the singular forms “a,” “an” and “the” include plural referents unless the context clearly dictates otherwise.

**[0020]** As used herein, the term “about” means that the recited numerical value is approximate and small variations would not significantly affect the practice of the disclosed embodiments. Where a numerical value is used, unless indicated otherwise by the context, the term “about” means the numerical value can vary by +10% and remain within the scope of the disclosed embodiments.

**[0021]** As used herein, the term “comprising” may be replaced with “consisting” or “consisting essentially of” in particular embodiments as desired.

**[0022]** As used herein, the term “isolated”, in regard to a nucleic acid molecule or a polypeptide, means that the nucleic acid molecule or polypeptide is in a condition other than its native environment, such as apart from blood and/or animal tissue. In some embodiments, an isolated nucleic acid molecule or polypeptide is substantially free of other nucleic acid molecules or other polypeptides, particularly other nucleic acid molecules or polypeptides of animal origin. In some embodiments, the nucleic acid molecule or polypeptide can be in a highly purified form, i.e., greater than 95% pure or greater than 99% pure. When used in this context, the term “isolated” does not exclude the presence of the same nucleic acid molecule or polypeptide in alternative physical forms, such as dimers or alternatively phosphorylated or derivatized forms.

**[0023]** As used herein, the terms “nucleic acid”, “nucleic acid molecule”, “nucleic acid sequence”, “polynucleotide”, or “oligonucleotide” can comprise a polymeric form of nucleotides of any length, can comprise DNA and/or RNA, and can be single-stranded, double-stranded, or multiple stranded. One strand of a nucleic acid also refers to its complement.

**[0024]** As used herein, the term “subject” includes any animal, including mammals. Mammals include, but are not limited to, farm animals (such as, for example, horse, cow, pig), companion animals (such as, for example, dog, cat), laboratory animals (such as, for example, mouse, rat, rabbits), and non-human primates (such as, for example, apes

and monkeys). In some embodiments, the subject is a human. In some embodiments, the subject is a patient under the care of a physician.

[0025] A rare variant in the KLHDC7B gene associated with an increased risk of developing hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss, in humans has been identified in accordance with the present disclosure. For example, a genetic alteration that changes the guanine nucleotide of position 3,778 in the human KLHDC7B reference (see, SEQ ID NO:1) to adenine has been observed to indicate that the human having such an alteration may have an increased risk of developing hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss. It is believed that no rare (minor allele frequency <1%) and predicted loss-of-function variants of the KLHDC7B gene or protein have any known association with hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss. Altogether, the genetic analyses described herein surprisingly indicate that the KLHDC7B gene and, in particular, a variant in the KLHDC7B gene, associates with an increased risk of developing hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss. Therefore, subjects that have a KLHDC7B variant nucleic acid molecule or polypeptide that associates with an increased risk of developing hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss, may be treated such that hearing loss is prevented, the symptoms thereof are reduced, and/or development of symptoms is repressed. Accordingly, the present disclosure provides methods of leveraging the identification of such variants in subjects to identify or stratify risk in such subjects of developing hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss, or to diagnose subjects as having an increased risk of developing hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss, such that subjects at risk or subjects with active disease may be treated accordingly.

[0026] For purposes of the present disclosure, any particular subject can be categorized as having one of three KLHDC7B genotypes: i) KLHDC7B reference; ii) heterozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide; or iii) homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide. A subject is KLHDC7B reference when the subject does not have a copy of a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide. A subject is heterozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide when the subject has a single copy of a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide. A KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide is any KLHDC7B nucleic acid molecule (such as, a genomic nucleic acid molecule, an mRNA molecule, or a cDNA molecule) encoding a KLHDC7B polypeptide having a partial loss-of-function, a complete loss-of-function, a predicted partial loss-of-function, or a predicted complete loss-of-function. A subject who has a KLHDC7B polypeptide having a partial loss-of-function (or predicted

partial loss-of-function) is hypomorphic for KLHDC7B. The KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide can be any nucleic acid molecule encoding KLHDC7B V1145M, V504M, V405M, K822fs, K181fs, K82fs, G943fs, G302fs, or G203fs. A subject is homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide when the subject has two copies of a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide.

[0027] For subjects that are genotyped or determined to be heterozygous or homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide, such subjects have an increased risk of developing hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss. For subjects that are genotyped or determined to be heterozygous or homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide, such subjects can be treated with an agent effective to treat hearing loss, such as conductive hearing loss, sensorineural hearing loss, or neural hearing loss.

[0028] In any of the embodiments described herein, the KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide can be any KLHDC7B nucleic acid molecule (such as, for example, genomic nucleic acid molecule, mRNA molecule, or cDNA molecule) encoding a KLHDC7B polypeptide having a partial loss-of-function, a complete loss-of-function, a predicted partial loss-of-function, or a predicted complete loss-of-function. For example, the KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide can be any nucleic acid molecule encoding KLHDC7B V1145M, V504M, V405M, K822fs, K181fs, K82fs, G943fs, G302fs, or G203fs. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B V1145M. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B V504M. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B V405M. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B K822fs. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B K181fs. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B K82fs. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B G943fs. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B G302fs. In some embodiments, the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B G203fs.

[0029] In any of the embodiments described herein, the KLHDC7B predicted loss-of-function polypeptide can be any KLHDC7B polypeptide having a partial loss-of-function, a complete loss-of-function, a predicted partial loss-of-function, or a predicted complete loss-of-function. In any of the embodiments described herein, the KLHDC7B predicted loss-of-function polypeptide can be any of the KLHDC7B polypeptides described herein including, for example, KLHDC7B V1145M, V504M, V405M, K822fs, K181fs, K82fs, G943fs, G302fs, or G203fs. In some

embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B V1145M. In some embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B V504M. In some embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B V405M. In some embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B K822fs. In some embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B K181fs. In some embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B K82fs. In some embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B G943fs. In some embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B G302fs. In some embodiments, the KLHDC7B predicted loss-of-function polypeptide is KLHDC7B G203fs.

[0030] In any of the embodiments described herein, hearing loss is conductive hearing loss, sensorineural hearing loss, or neural hearing loss. In any of the embodiments described herein, hearing loss is conductive hearing loss. In any of the embodiments described herein, hearing loss is sensorineural hearing loss. In any of the embodiments described herein, hearing loss is neural hearing loss.

[0031] Symptoms of hearing loss include, but are not limited to, hearing problem (muffling of speech and other sounds, difficulty understanding words, especially against background noise or in a crowd, or trouble hearing consonants), ringing in the ears, sensitivity to sound, and speech delay in a child.

[0032] The present disclosure also provides methods of treating a subject with a therapeutic agent that treats or inhibits hearing loss, wherein the subject has hearing loss. In some embodiments, the methods comprise determining whether the subject has a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide by obtaining or having obtained a biological sample from the subject, and performing or having performed a sequence analysis on the biological sample to determine if the subject has a genotype comprising the KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide. The methods comprise administering or continuing to administer the therapeutic agent that treats or inhibits hearing loss in a standard dosage amount to a subject that is KLHDC7B reference. The methods further comprise administering or continuing to administer the therapeutic agent that treats or inhibits hearing loss in an amount that is the same as or greater than a standard dosage amount to a subject that is heterozygous or homozygous for the KLHDC7B missense variant nucleic acid molecule. The presence of a genotype having the KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide indicates the subject has an increased risk of developing hearing loss. In some embodiments, the subject is KLHDC7B reference. In some embodiments, the subject is heterozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide. In some embodiments, the subject is homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide.

[0033] In some embodiments, the methods of treatment further comprise detecting the presence or absence of a KLHDC7B missense variant nucleic acid molecule encod-

ing a KLHDC7B predicted loss-of-function polypeptide in a biological sample from the subject. As used throughout the present disclosure, a "KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide" is any KLHDC7B nucleic acid molecule (such as, for example, genomic nucleic acid molecule, mRNA molecule, or cDNA molecule) encoding a KLHDC7B polypeptide having a partial loss-of-function, a complete loss-of-function, a predicted partial loss-of-function, or a predicted complete loss-of-function.

[0034] Detecting the presence or absence of a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide in a biological sample from a subject and/or determining whether a subject has a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide can be carried out by any of the methods described herein. In some embodiments, these methods can be carried out in vitro. In some embodiments, these methods can be carried out in situ. In some embodiments, these methods can be carried out in vivo. In any of these embodiments, the nucleic acid molecule can be present within a cell obtained from the subject.

[0035] The present disclosure also provides methods of treating a subject with a therapeutic agent that treats or inhibits hearing loss, wherein the subject has hearing loss. In some embodiments, the method comprises determining whether the subject has a KLHDC7B predicted loss-of-function polypeptide by obtaining or having obtained a biological sample from the subject, and performing or having performed an assay on the biological sample to determine if the subject has a KLHDC7B predicted loss-of-function polypeptide. The methods comprise administering or continuing to administer the therapeutic agent that treats or inhibits hearing loss in a standard dosage amount to a subject that is KLHDC7B reference. The methods further comprise administering or continuing to administer the therapeutic agent that treats or inhibits hearing loss in an amount that is the same as or greater than a standard dosage amount to a subject that has a KLHDC7B predicted loss-of-function polypeptide. The presence of a KLHDC7B predicted loss-of-function polypeptide indicates the subject has an increased risk of developing hearing loss. In some embodiments, the subject has a KLHDC7B predicted loss-of-function polypeptide. In some embodiments, the subject does not have a KLHDC7B predicted loss-of-function polypeptide.

[0036] Detecting the presence or absence of a KLHDC7B predicted loss-of-function polypeptide in a biological sample from a subject and/or determining whether a subject has a KLHDC7B predicted loss-of-function polypeptide can be carried out by any of the methods described herein. In some embodiments, these methods can be carried out in vitro. In some embodiments, these methods can be carried out in situ. In some embodiments, these methods can be carried out in vivo. In any of these embodiments, the polypeptide can be present within a cell obtained from the subject.

[0037] Examples of therapeutic agents that treat or inhibit hearing loss include, but are not limited to: antioxidants, calcium-channel blockers, anti-inflammatory drugs (such as steroids), apoptosis inhibitors, D-methionine, ebselen, N-acetylcysteine, lipoic acid, combination of ebselen and allopurinol, resveratrol, neurotrophic factors (such as

T-817MA), caspase inhibitors (such as z-DEVD-fmk), copper transport inhibitors (such as cimetidine and copper sulphate), and micronutrients with antioxidant vitamins.

**[0038]** In some embodiments, the dose of the therapeutic agents that treat or inhibit hearing loss can be increased by about 10%, by about 20%, by about 30%, by about 40%, by about 50%, by about 60%, by about 70%, by about 80%, or by about 90% for subjects that are heterozygous or homozygous for a KLHDC7B predicted loss-of-function variant (i.e., a greater amount than the standard dosage amount) compared to subjects that are KLHDC7B reference (who may receive a standard dosage amount). In some embodiments, the dose of the therapeutic agents that treat or inhibit hearing loss can be increased by about 10%, by about 20%, by about 30%, by about 40%, or by about 50%. In addition, the dose of therapeutic agents that treat or inhibit hearing loss in subjects that are heterozygous or homozygous for a KLHDC7B predicted loss-of-function variant can be administered more frequently compared to subjects that are KLHDC7B reference.

**[0039]** In some embodiments, the dose of the therapeutic agents that treat or inhibit hearing loss can be increased by about 10%, by about 20%, by about 30%, by about 40%, by about 50%, by about 60%, by about 70%, by about 80%, or by about 90% for subjects that are homozygous for a KLHDC7B predicted loss-of-function variant compared to subjects that are heterozygous for a KLHDC7B predicted loss-of-function variant. In some embodiments, the dose of the therapeutic agents that treat or inhibit hearing loss can be increased by about 10%, by about 20%, by about 30%, by about 40%, or by about 50%. In addition, the dose of therapeutic agents that treat or inhibit hearing loss in subjects that are homozygous for a KLHDC7B predicted loss-of-function variant can be administered more frequently compared to subjects that are heterozygous for a KLHDC7B predicted loss-of-function variant.

**[0040]** Administration of the therapeutic agents that treat or inhibit hearing loss can be repeated, for example, after one day, two days, three days, five days, one week, two weeks, three weeks, one month, five weeks, six weeks, seven weeks, eight weeks, two months, or three months. The repeated administration can be at the same dose or at a different dose. The administration can be repeated once, twice, three times, four times, five times, six times, seven times, eight times, nine times, ten times, or more. For example, according to certain dosage regimens a subject can receive therapy for a prolonged period of time such as, for example, 6 months, 1 year, or more.

**[0041]** Administration of the therapeutic agents that treat or inhibit hearing loss can occur by any suitable route including, but not limited to, parenteral, intravenous, oral, subcutaneous, intra-arterial, intracranial, intrathecal, intraperitoneal, topical, intranasal, or intramuscular. Pharmaceutical compositions for administration are desirably sterile and substantially isotonic and manufactured under GMP conditions. Pharmaceutical compositions can be provided in unit dosage form (i.e., the dosage for a single administration). Pharmaceutical compositions can be formulated using one or more physiologically and pharmaceutically acceptable carriers, diluents, excipients or auxiliaries. The formulation depends on the route of administration chosen. The term "pharmaceutically acceptable" means that the carrier,

diluent, excipient, or auxiliary is compatible with the other ingredients of the formulation and not substantially deleterious to the recipient thereof.

**[0042]** The terms "treat", "treating", and "treatment" and "prevent", "preventing", and "prevention" as used herein, refer to eliciting the desired biological response, such as a therapeutic and prophylactic effect, respectively. In some embodiments, a therapeutic effect comprises one or more of a decrease/reduction in hearing loss, a decrease/reduction in the severity of hearing loss (such as, for example, a reduction or inhibition of development or hearing loss), a decrease/reduction in symptoms and hearing loss-related effects, delaying the onset of symptoms and hearing loss-related effects, reducing the severity of symptoms of hearing loss-related effects, reducing the severity of an acute episode, reducing the number of symptoms and hearing loss-related effects, reducing the latency of symptoms and hearing loss-related effects, an amelioration of symptoms and hearing loss-related effects, reducing secondary symptoms, preventing relapse to hearing loss, decreasing the number or frequency of relapse episodes, increasing latency between symptomatic episodes, increasing time to sustained progression, speeding recovery, and/or increasing efficacy of or decreasing resistance to alternative therapeutics, following administration of the agent or composition comprising the agent. A prophylactic effect may comprise a complete or partial avoidance/inhibition or a delay of hearing loss development/progression (such as, for example, a complete or partial avoidance/inhibition or a delay) following administration of a therapeutic protocol. Treatment of hearing loss encompasses the treatment of subjects already diagnosed as having any form of hearing loss at any clinical stage or manifestation, the delay of the onset or evolution or aggravation or deterioration of the symptoms or signs of hearing loss, and/or preventing and/or reducing the severity of hearing loss.

**[0043]** The present disclosure also provides methods of identifying a subject having an increased risk for developing hearing loss. In some embodiments, the method comprises determining or having determined in a biological sample obtained from the subject the presence or absence of a KLHDC7B missense variant nucleic acid molecule (such as a genomic nucleic acid molecule, mRNA molecule, and/or cDNA molecule) encoding a KLHDC7B predicted loss-of-function polypeptide. When the subject lacks a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide (i.e., the subject is genotypically categorized as a KLHDC7B reference), then the subject does not have an increased risk for developing hearing loss. When the subject has a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide (i.e., the subject is heterozygous or homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide), then the subject has an increased risk for developing hearing loss.

**[0044]** Determining whether a subject has a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide in a biological sample from a subject and/or determining whether a subject has a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide can be carried out by any of the methods described herein. In some embodiments, these methods can

be carried out in vitro. In some embodiments, these methods can be carried out in situ. In some embodiments, these methods can be carried out in vivo. In any of these embodiments, the nucleic acid molecule can be present within a cell obtained from the subject.

[0045] In some embodiments, when a subject is identified as having an increased risk of developing hearing loss, the subject is further treated with a therapeutic agent that treats or inhibits hearing loss, as described herein. In some embodiments, when the subject is heterozygous or homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide, the subject is administered the therapeutic agent that treats or inhibits hearing loss in a dosage amount that is the same as or greater than a standard dosage amount. In some embodiments, when the subject is homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide, the subject is administered the therapeutic agent that treats or inhibits hearing loss in a dosage amount that is the same as or greater than the dosage amount administered to a subject that is heterozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide. In some embodiments, the subject is KLHDC7B reference. In some embodiments, the subject is heterozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide. In some embodiments, the subject is homozygous for a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide.

[0046] The present disclosure also provides methods of detecting the presence or absence of a KLHDC7B missense variant genomic nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide in a biological sample from a subject, and/or a KLHDC7B missense variant mRNA molecule encoding a KLHDC7B predicted loss-of-function polypeptide in a biological sample from a subject, and/or a KLHDC7B missense variant cDNA molecule encoding a KLHDC7B predicted loss-of-function polypeptide produced from an mRNA molecule in a biological sample from a subject. It is understood that gene sequences within a population and mRNA molecules encoded by such genes can vary due to polymorphisms such as single-nucleotide polymorphisms. The sequences provided herein for the KLHDC7B variant genomic nucleic acid molecule, KLHDC7B variant mRNA molecule, and KLHDC7B variant cDNA molecule are only exemplary sequences. Other sequences for the KLHDC7B variant genomic nucleic acid molecule, variant mRNA molecule, and variant cDNA molecule are also possible.

[0047] The biological sample can be derived from any cell, tissue, or biological fluid from the subject. The sample may comprise any clinically relevant tissue, such as a bone marrow sample, a tumor biopsy, a fine needle aspirate, or a sample of bodily fluid, such as blood, gingival crevicular fluid, plasma, serum, lymph, ascitic fluid, cystic fluid, or urine. In some cases, the sample comprises a buccal swab. The sample used in the methods disclosed herein will vary based on the assay format, nature of the detection method, and the tissues, cells, or extracts that are used as the sample. A biological sample can be processed differently depending on the assay being employed. For example, when detecting any KLHDC7B variant nucleic acid molecule, preliminary

processing designed to isolate or enrich the sample for the genomic DNA can be employed. A variety of techniques may be used for this purpose. When detecting the level of any KLHDC7B variant mRNA, different techniques can be used to enrich the biological sample with mRNA. Various methods to detect the presence or level of an mRNA or the presence of a particular variant genomic DNA locus can be used.

[0048] In some embodiments, detecting a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide in a subject comprises assaying or genotyping a biological sample obtained from the subject to determine whether a KLHDC7B genomic nucleic acid molecule in the biological sample, and/or a KLHDC7B mRNA molecule in the biological sample, and/or a KLHDC7B cDNA molecule produced from an mRNA molecule in the biological sample, comprises one or more variations that cause a loss-of-function (partial or complete) or are predicted to cause a loss-of-function (partial or complete).

[0049] In some embodiments, the methods of detecting the presence or absence of a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide (such as, for example, a genomic nucleic acid molecule, an mRNA molecule, and/or a cDNA molecule produced from an mRNA molecule) in a subject, comprise performing an assay on a biological sample obtained from the subject. The assay determines whether a nucleic acid molecule in the biological sample comprises a particular nucleotide sequence.

[0050] In some embodiments, the nucleotide sequence comprises: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2 (for genomic nucleic acid molecules), an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7 (for mRNA molecules), or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15 (for cDNA molecules obtained from mRNA molecules).

[0051] In some embodiments, the nucleotide sequence comprises: an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8 (for mRNA molecules), or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16 (for cDNA molecules obtained from mRNA molecules).

[0052] In some embodiments, the nucleotide sequence comprises: an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9 (for mRNA molecules), or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17 (for cDNA molecules obtained from mRNA molecules).

[0053] In some embodiments, the nucleotide sequence comprises: an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10 (for mRNA molecules), or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18 (for cDNA molecules obtained from mRNA molecules).

[0054] In some embodiments, the nucleotide sequence: lacks a guanine at a position corresponding to position 2,807 according to SEQ ID NO:1 (for genomic nucleic acid molecules), lacks a guanine at a position corresponding to position 2,807 according to SEQ ID NO:3 (for mRNA molecules), or lacks a guanine at a position corresponding to position 2,807 according to SEQ ID NO:11 (for cDNA molecules obtained from mRNA molecules).

[0055] In some embodiments, the nucleotide sequence: lacks a guanine at a position corresponding to position 673 according to SEQ ID NO:4 (for mRNA molecules), or lacks a guanine at a position corresponding to position 673 according to SEQ ID NO:12 (for cDNA molecules obtained from mRNA molecules).

[0056] In some embodiments, the nucleotide sequence: lacks a guanine at a position corresponding to position 2,503 according to SEQ ID NO:5 (for mRNA molecules), or lacks a guanine at a position corresponding to position 2,503 according to SEQ ID NO:13 (for cDNA molecules obtained from mRNA molecules).

[0057] In some embodiments, the nucleotide sequence: lacks a guanine at a position corresponding to position 673 according to SEQ ID NO:6 (for mRNA molecules), or lacks a guanine at a position corresponding to position 673 according to SEQ ID NO:14 (for cDNA molecules obtained from mRNA molecules).

[0058] In some embodiments, the nucleotide sequence: lacks a guanine at a position corresponding to position 3,170 according to SEQ ID NO:1 (for genomic nucleic acid molecules), lacks a guanine at a position corresponding to position 3,170 according to SEQ ID NO:3 (for mRNA molecules), or lacks a guanine at a position corresponding to position 3,170 according to SEQ ID NO:11 (for cDNA molecules obtained from mRNA molecules).

[0059] In some embodiments, the nucleotide sequence: lacks a guanine at a position corresponding to position 1,036 according to SEQ ID NO:4 (for mRNA molecules), or lacks a guanine at a position corresponding to position 1,036 according to SEQ ID NO:12 (for cDNA molecules obtained from mRNA molecules).

[0060] In some embodiments, the nucleotide sequence: lacks a guanine at a position corresponding to position 2,866 according to SEQ ID NO:5 (for mRNA molecules), or lacks a guanine at a position corresponding to position 2,866 according to SEQ ID NO:13 (for cDNA molecules obtained from mRNA molecules).

[0061] In some embodiments, the nucleotide sequence: lacks a guanine at a position corresponding to position 1,036 according to SEQ ID NO:6 (for mRNA molecules), or lacks a guanine at a position corresponding to position 1,036 according to SEQ ID NO:14 (for cDNA molecules obtained from mRNA molecules).

[0062] In some embodiments, the nucleotide sequence: comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof; lacks a guanine at a position corresponding to position 2,807 according to SEQ ID NO:1, or the complement thereof; or lacks a guanine at a position corresponding to position 3,170 according to SEQ ID NO:1, or the complement thereof.

[0063] In some embodiments, the nucleotide sequence: comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; comprises an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or the complement thereof; comprises an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or the complement thereof; comprises an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or the complement thereof; lacks a guanine at a position corresponding to position 2,807 according to SEQ ID NO:3, or the complement thereof; lacks a guanine at a

position corresponding to position 673 according to SEQ ID NO:4, or the complement thereof; lacks a guanine at a position corresponding to position 2,503 according to SEQ ID NO:5, or the complement thereof; lacks a guanine at a position corresponding to position 673 according to SEQ ID NO:6, or the complement thereof; lacks a guanine at a position corresponding to position 3,170 according to SEQ ID NO:3, or the complement thereof; lacks a guanine at a position corresponding to position 1,036 according to SEQ ID NO:4, or the complement thereof; lacks a guanine at a position corresponding to position 2,866 according to SEQ ID NO:5, or the complement thereof; or lacks a guanine at a position corresponding to position 1,036 according to SEQ ID NO:6, or the complement thereof.

[0064] In some embodiments, the nucleotide sequence: comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; comprises an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or the complement thereof; comprises an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, or the complement thereof; comprises an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, or the complement thereof; lacks a guanine at a position corresponding to position 2,807 according to SEQ ID NO:11, or the complement thereof; lacks a guanine at a position corresponding to position 673 according to SEQ ID NO:12, or the complement thereof; lacks a guanine at a position corresponding to position 2,503 according to SEQ ID NO:13, or the complement thereof; lacks a guanine at a position corresponding to position 673 according to SEQ ID NO:14, or the complement thereof; lacks a guanine at a position corresponding to position 3,170 according to SEQ ID NO:11, or the complement thereof; lacks a guanine at a position corresponding to position 1,036 according to SEQ ID NO:12, or the complement thereof; lacks a guanine at a position corresponding to position 2,866 according to SEQ ID NO:13, or the complement thereof; or lacks a guanine at a position corresponding to position 1,036 according to SEQ ID NO:14, or the complement thereof.

[0065] In some embodiments, the biological sample comprises a cell or cell lysate. Such methods can further comprise, for example, obtaining a biological sample from the subject comprising a KLHDC7B genomic nucleic acid molecule or mRNA molecule, and if mRNA, optionally reverse transcribing the mRNA into cDNA. Such assays can comprise, for example determining the identity of these positions of the particular KLHDC7B nucleic acid molecule. In some embodiments, the method is an *in vitro* method.

[0066] In some embodiments, the determining step, detecting step, or sequence analysis comprises sequencing at least a portion of the nucleotide sequence of the KLHDC7B genomic nucleic acid molecule, the KLHDC7B mRNA molecule, or the KLHDC7B cDNA molecule in the biological sample, wherein the sequenced portion comprises one or more variations that cause a loss-of-function (partial or complete) or are predicted to cause a loss-of-function (partial or complete).

[0067] In some embodiments, the determining step, detecting step, or sequence analysis comprises sequencing at least a portion of: the nucleotide sequence of the KLHDC7B genomic nucleic acid molecule in the biological sample, wherein the sequenced portion comprises a position corresponding to position 3,778 according to SEQ ID NO:2, or





genomic nucleic acid molecule in the biological sample, wherein the sequenced portion comprises: a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof; positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, or the complement thereof; or positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, or the complement thereof. When the sequenced portion of the KLHDC7B nucleic acid molecule in the biological sample comprises: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, then the KLHDC7B nucleic acid molecule in the biological sample is a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide.

[0080] In some embodiments, the determining step, detecting step, or sequence analysis comprises sequencing at least a portion of the nucleotide sequence of the KLHDC7B mRNA molecule in the biological sample, wherein the sequenced portion comprises: a position corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; a position corresponding to position 1,644 according to SEQ ID NO:8, or the complement thereof; a position corresponding to position 3,474 according to SEQ ID NO:9, or the complement thereof; a position corresponding to position 1,644 according to SEQ ID NO:10, or the complement thereof; positions corresponding to positions 2,806-2,807 according to SEQ ID NO:27, or the complement thereof; positions corresponding to positions 672-673 according to SEQ ID NO:28, or the complement thereof; positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or the complement thereof; positions corresponding to positions 672-673 according to SEQ ID NO:30, or the complement thereof; positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or the complement thereof; positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or the complement thereof; positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or the complement thereof; or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof. When the sequenced portion of the KLHDC7B nucleic acid molecule in the biological sample comprises: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or an AG dinucleotide at positions corre-

sponding to positions 1,035-1,036 according to SEQ ID NO:34, then the KLHDC7B nucleic acid molecule in the biological sample is a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide.

[0081] In some embodiments, the determining step, detecting step, or sequence analysis comprises sequencing at least a portion of the nucleotide sequence of the KLHDC7B cDNA molecule in the biological sample, wherein the sequenced portion comprises: a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; a position corresponding to position 1,644 according to SEQ ID NO:16, or the complement thereof; a position corresponding to position 3,474 according to SEQ ID NO:17, or the complement thereof; a position corresponding to position 1,644 according to SEQ ID NO:18, or the complement thereof; positions corresponding to positions 2,806-2,807 according to SEQ ID NO:35, or the complement thereof; positions corresponding to positions 672-673 according to SEQ ID NO:36, or the complement thereof; positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37, or the complement thereof; positions corresponding to positions 672-673 according to SEQ ID NO:38, or the complement thereof; positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, or the complement thereof; positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or the complement thereof; positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or the complement thereof; or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof. When the sequenced portion of the KLHDC7B nucleic acid molecule in the biological sample comprises: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, then the KLHDC7B nucleic acid molecule in the biological sample is a KLHDC7B missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide.

[0082] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B: genomic nucleic acid molecule that is proximate to a position corresponding to position 3,778 according to SEQ ID NO:2; mRNA molecule that is proximate to a position corresponding to position 3,778 according to SEQ ID NO:7; and/or

cDNA molecule that is proximate to a position corresponding to position 3,778 according to SEQ ID NO:15; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B: genomic nucleic acid molecule corresponding to position 3,778 according to SEQ ID NO:2; mRNA molecule corresponding to position 3,778 according to SEQ ID NO:7; and/or cDNA molecule corresponding to position 3,778 according to SEQ ID NO:15; and c) determining whether the extension product of the primer comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, and/or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15.

[0083] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B mRNA molecule that is proximate to a position corresponding to position 1,644 according to SEQ ID NO:8, and/or cDNA molecule that is proximate to a position corresponding to position 1,644 according to SEQ ID NO:16; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B mRNA molecule corresponding to position 1,644 according to SEQ ID NO:8, and/or cDNA molecule corresponding to position 1,644 according to SEQ ID NO:16; and c) determining whether the extension product of the primer comprises an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, and/or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16.

[0084] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B mRNA molecule that is proximate to a position corresponding to position 3,474 according to SEQ ID NO:9, and/or cDNA molecule that is proximate to a position corresponding to position 3,474 according to SEQ ID NO:17; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B mRNA molecule corresponding to position 3,474 according to SEQ ID NO:9, and/or cDNA molecule corresponding to position 3,474 according to SEQ ID NO:17; and c) determining whether the extension product of the primer comprises an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, and/or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17.

[0085] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B mRNA molecule that is proximate to a position corresponding to position 1,644 according to SEQ ID NO:10, and/or cDNA molecule that is proximate to a position corresponding to position 1,644 according to SEQ ID NO:18; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B mRNA molecule corresponding to position 1,644 according to SEQ ID NO:10, and/or cDNA molecule corresponding to position 1,644 according to SEQ ID NO:18; and c) determining whether the extension product of the primer comprises an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10,

and/or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18.

[0086] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B: genomic nucleic acid molecule that is proximate to positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25; mRNA molecule that is proximate to positions corresponding to positions 2,806-2,807 according to SEQ ID NO:27; and/or cDNA molecule that is proximate to positions corresponding to positions 2,806-2,807 according to SEQ ID NO:35; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B: genomic nucleic acid molecule that is proximate to positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25; mRNA molecule that is proximate to positions corresponding to positions 2,806-2,807 according to SEQ ID NO:27; and/or cDNA molecule that is proximate to positions corresponding to positions 2,806-2,807 according to SEQ ID NO:35; and c) determining whether the extension product of the primer comprises a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35.

[0087] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B: mRNA molecule that is proximate to positions corresponding to positions 672-673 according to SEQ ID NO:28; and/or cDNA molecule that is proximate to positions corresponding to positions 672-673 according to SEQ ID NO:36; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B: mRNA molecule that is proximate to positions corresponding to positions 672-673 according to SEQ ID NO:28; and/or cDNA molecule that is proximate to positions corresponding to positions 672-673 according to SEQ ID NO:36; and c) determining whether the extension product of the primer comprises a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36.

[0088] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B: mRNA molecule that is proximate to positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29; and/or cDNA molecule that is proximate to positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B: mRNA molecule that is proximate to positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29; and/or cDNA molecule that is proximate to positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37; and c) determining whether the extension product of the primer comprises a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or a CG



of the nucleotide sequence of the KLHDC7B mRNA molecule that is proximate to: a position corresponding to position 3,778 according to SEQ ID NO:7, a position corresponding to position 1,644 according to SEQ ID NO:8, a position corresponding to position 3,474 according to SEQ ID NO:9, a position corresponding to position 1,644 according to SEQ ID NO:10, positions corresponding to positions 2,806-2,807 according to SEQ ID NO:27, positions corresponding to positions 672-673 according to SEQ ID NO:28, positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, positions corresponding to positions 672-673 according to SEQ ID NO:30, positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B mRNA molecule corresponding to: position 3,778 according to SEQ ID NO:7, position 1,644 according to SEQ ID NO:8, position 3,474 according to SEQ ID NO:9, position 1,644 according to SEQ ID NO:10, positions 2,806-2,807 according to SEQ ID NO:27, positions 672-673 according to SEQ ID NO:28, positions 2,502-2,503 according to SEQ ID NO:29, positions 672-673 according to SEQ ID NO:30, positions 3,169-3,170 according to SEQ ID NO:31, positions 1,035-1,036 according to SEQ ID NO:32, positions 2,865-2,866 according to SEQ ID NO:33, or positions 1,035-1,036 according to SEQ ID NO:34; and c) determining whether the extension product of the primer comprises: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34.

[0096] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B cDNA molecule that is proximate to: a position corresponding to position 3,778 according to SEQ ID NO:15, a position corresponding to position 1,644 according to SEQ ID NO:16, a position corresponding to position 3,474 according to SEQ ID NO:17, a position corresponding to position 1,644 according to SEQ ID NO:18, positions corresponding to positions 2,806-2,807 according to SEQ ID NO:35, positions corresponding to positions 672-673 according to SEQ ID NO:36, positions corresponding to positions 2,502-

2,503 according to SEQ ID NO:37, positions corresponding to positions 672-673 according to SEQ ID NO:38, positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42; b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B cDNA molecule corresponding to: position 3,778 according to SEQ ID NO:15, position 1,644 according to SEQ ID NO:16, position 3,474 according to SEQ ID NO:17, position 1,644 according to SEQ ID NO:18, positions 2,806-2,807 according to SEQ ID NO:35, positions 672-673 according to SEQ ID NO:36, positions 2,502-2,503 according to SEQ ID NO:37, positions 672-673 according to SEQ ID NO:38, positions 3,169-3,170 according to SEQ ID NO:39, positions 1,035-1,036 according to SEQ ID NO:40, positions 2,865-2,866 according to SEQ ID NO:41, or positions 1,035-1,036 according to SEQ ID NO:42; and c) determining whether the extension product of the primer comprises: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42.

[0097] In some embodiments, the assay comprises sequencing the entire nucleic acid molecule. In some embodiments, only a KLHDC7B genomic nucleic acid molecule is analyzed. In some embodiments, only a KLHDC7B mRNA is analyzed. In some embodiments, only a KLHDC7B cDNA obtained from KLHDC7B mRNA is analyzed.

[0098] In some embodiments, the determining step, detecting step, or sequence analysis comprises: a) amplifying at least a portion of the nucleic acid molecule that encodes the KLHDC7B polypeptide, wherein the amplified portion comprises: i) an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof; ii) an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; and/or iii) an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; b) labeling the amplified nucleic acid molecule with a detectable label; c) contacting the labeled nucleic acid molecule with a support comprising an alteration-specific probe, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleic acid sequence of the







positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof; and d) detecting the detectable label.

[0113] In some embodiments, the nucleic acid molecule is mRNA and the determining step further comprises reverse-transcribing the mRNA into a cDNA prior to the amplifying step.

[0114] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof; an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; and/or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; and detecting the detectable label.

[0115] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or the complement thereof, and/or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or the complement thereof; and detecting the detectable label.

[0116] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or the complement thereof, and/or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, or the complement thereof; and detecting the detectable label.

[0117] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or the complement thereof, and/or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, or the complement thereof; and detecting the detectable label.

[0118] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: a CG dinucleotide at positions corresponding to

positions 2,806-2,807 according to SEQ ID NO:25, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or the complement thereof; or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, or the complement thereof; and detecting the detectable label.

[0119] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or the complement thereof; or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, or the complement thereof; and detecting the detectable label.

[0120] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or the complement thereof; or a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37, or the complement thereof; and detecting the detectable label.

[0121] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or the complement thereof; or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, or the complement thereof; and detecting the detectable label.

[0122] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, or the complement thereof; an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, or the complement thereof; and detecting the detectable label.

[0123] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an

alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or the complement thereof; and detecting the detectable label.

[0124] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or the complement thereof; and detecting the detectable label.

[0125] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof; and detecting the detectable label.

[0126] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, or the complement thereof; and detecting the detectable label.

[0127] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or the complement thereof; an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or the complement

thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or the complement thereof; an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or the complement thereof; an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof; and detecting the detectable label.

[0128] In some embodiments, the determining step, detecting step, or sequence analysis comprises: contacting the nucleic acid molecule in the biological sample with an alteration-specific probe comprising a detectable label, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleotide sequence of the amplified nucleic acid molecule comprising: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or the complement thereof; an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, or the complement thereof; an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or the complement thereof; an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof; and detecting the detectable label.

[0129] Alteration-specific polymerase chain reaction techniques can be used to detect mutations such as SNPs in a nucleic acid sequence. Alteration-specific primers can be used because the DNA polymerase will not extend when a mismatch with the template is present.

[0130] In some embodiments, the nucleic acid molecule in the sample is mRNA and the mRNA is reverse-transcribed

into a cDNA prior to the amplifying step. In some embodiments, the nucleic acid molecule is present within a cell obtained from the subject.

[0131] In some embodiments, the assay comprises contacting the biological sample with a primer or probe, such as an alteration-specific primer or alteration-specific probe, that specifically hybridizes to a KLHDC7B variant genomic sequence, variant mRNA sequence, or variant cDNA sequence and not the corresponding KLHDC7B reference sequence under stringent conditions, and determining whether hybridization has occurred.

[0132] In some embodiments, the assay comprises RNA sequencing (RNA-Seq). In some embodiments, the assays also comprise reverse transcribing mRNA into cDNA, such as by the reverse transcriptase polymerase chain reaction (RT-PCR).

[0133] In some embodiments, the methods utilize probes and primers of sufficient nucleotide length to bind to the target nucleotide sequence and specifically detect and/or identify a polynucleotide comprising a KLHDC7B variant genomic nucleic acid molecule, variant mRNA molecule, or variant cDNA molecule. The hybridization conditions or reaction conditions can be determined by the operator to achieve this result. The nucleotide length may be any length that is sufficient for use in a detection method of choice, including any assay described or exemplified herein. Such probes and primers can hybridize specifically to a target nucleotide sequence under high stringency hybridization conditions. Probes and primers may have complete nucleotide sequence identity of contiguous nucleotides within the target nucleotide sequence, although probes differing from the target nucleotide sequence and that retain the ability to specifically detect and/or identify a target nucleotide sequence may be designed by conventional methods. Probes and primers can have about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or 100% sequence identity or complementarity with the nucleotide sequence of the target nucleic acid molecule.

[0134] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2 (genomic nucleic acid molecule), an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7 (mRNA molecule), or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, and a second primer derived from the 3' flanking sequence adjacent to an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15. In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15.

[0135] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8 (mRNA molecule), or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, and a second primer derived from the 3' flanking sequence adjacent to an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16 to produce an amplicon that is indicative of the presence of the SNP at positions encoding an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16.

[0136] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16. In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16.

nucleotide sequence comprising: an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9 (mRNA molecule), or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, and a second primer derived from the 3' flanking sequence adjacent to an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17 to produce an amplicon that is indicative of the presence of the SNP at positions encoding an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17.

[0137] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10 (mRNA molecule), or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, and a second primer derived from the 3' flanking sequence adjacent to an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18 to produce an amplicon that is indicative of the presence of the SNP at positions encoding an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including posi-

tions comprising an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18.

[0138] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25 (genomic nucleic acid molecule), a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27 (mRNA molecule), or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, and a second primer derived from the 3' flanking sequence adjacent to a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35 to produce an amplicon that is indicative of the presence of the SNP at positions encoding a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35.

[0139] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28 (mRNA molecule), or a CG dinucleotide at

positions corresponding to positions 672-673 according to SEQ ID NO:36 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36 and a second primer derived from the 3' flanking sequence adjacent to a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36 to produce an amplicon that is indicative of the presence of the SNP at positions encoding a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36.

[0140] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29 (mRNA molecule), or a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37 and a second primer derived from the 3' flanking sequence adjacent to a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37 to produce an amplicon that is indicative of the presence of the SNP at positions encoding a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the

primer pair flanks a region including positions comprising a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37.

[0141] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30 (mRNA molecule), or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38 and a second primer derived from the 3' flanking sequence adjacent to a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38 to produce an amplicon that is indicative of the presence of the SNP at positions encoding a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38.

[0142] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26 (genomic nucleic acid molecule), an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31 (mRNA molecule), or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to an AG

dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, and a second primer derived from the 3' flanking sequence adjacent to an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39 to produce an amplicon that is indicative of the presence of the SNP at positions encoding an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26, an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39.

[0143] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32 (mRNA molecule), or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40 and a second primer derived from the 3' flanking sequence adjacent to an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40 to produce an amplicon that is indicative of the presence of the SNP at positions encoding an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40. In

some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40.

[0144] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33 (mRNA molecule), or an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41 and a second primer derived from the 3' flanking sequence adjacent to an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41 to produce an amplicon that is indicative of the presence of the SNP at positions encoding an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41.

[0145] In some embodiments, to determine whether a KLHDC7B nucleic acid molecule (genomic nucleic acid molecule, mRNA molecule, or cDNA molecule), or complement thereof, within a biological sample comprises a nucleotide sequence comprising: an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34 (mRNA molecule), or an AG dinucleotide

at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42 (cDNA molecule), the biological sample can be subjected to an amplification method using a primer pair that includes a first primer derived from the 5' flanking sequence adjacent to an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42 and a second primer derived from the 3' flanking sequence adjacent to an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42 to produce an amplicon that is indicative of the presence of the SNP at positions encoding an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42. In some embodiments, the amplicon may range in length from the combined length of the primer pairs plus one nucleotide base pair to any length of amplicon producible by a DNA amplification protocol. This distance can range from one nucleotide base pair up to the limits of the amplification reaction, or about twenty thousand nucleotide base pairs. Optionally, the primer pair flanks a region including positions comprising an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nucleotides on each side of positions comprising an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42.

[0146] Similar amplicons can be generated from the mRNA and/or cDNA sequences. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose, such as the PCR primer analysis tool in Vector NTI version 10 (Informax Inc., Bethesda Md.); PrimerSelect (DNASTAR Inc., Madison, Wis.); and Primer3 (Version 0.4.0.COPYRGT., 1991, Whitehead Institute for Biomedical Research, Cambridge, Mass.). Additionally, the sequence can be visually scanned and primers manually identified using known guidelines.

[0147] Illustrative examples of nucleic acid sequencing techniques include, but are not limited to, chain terminator (Sanger) sequencing and dye terminator sequencing. Other methods involve nucleic acid hybridization methods other than sequencing, including using labeled primers or probes directed against purified DNA, amplified DNA, and fixed cell preparations (fluorescence in situ hybridization (FISH)). In some methods, a target nucleic acid molecule may be amplified prior to or simultaneous with detection. Illustrative examples of nucleic acid amplification techniques include, but are not limited to, polymerase chain reaction (PCR), ligase chain reaction (LCR), strand displacement amplification (SDA), and nucleic acid sequence based amplification (NASBA). Other methods include, but are not limited to, ligase chain reaction, strand displacement amplification, and thermophilic SDA (tSDA).

[0148] In hybridization techniques, stringent conditions can be employed such that a probe or primer will specifically hybridize to its target. In some embodiments, a polynucle-

otide primer or probe under stringent conditions will hybridize to its target sequence to a detectably greater degree than to other non-target sequences, such as, at least 2-fold, at least 3-fold, at least 4-fold, or more over background, including over 10-fold over background. In some embodiments, a polynucleotide primer or probe under stringent conditions will hybridize to its target nucleotide sequence to a detectably greater degree than to other nucleotide sequences by at least 2-fold. In some embodiments, a polynucleotide primer or probe under stringent conditions will hybridize to its target nucleotide sequence to a detectably greater degree than to other nucleotide sequences by at least 3-fold. In some embodiments, a polynucleotide primer or probe under stringent conditions will hybridize to its target nucleotide sequence to a detectably greater degree than to other nucleotide sequences by at least 4-fold. In some embodiments, a polynucleotide primer or probe under stringent conditions will hybridize to its target nucleotide sequence to a detectably greater degree than to other nucleotide sequences by over 10-fold over background. Stringent conditions are sequence-dependent and will be different in different circumstances.

[0149] Appropriate stringency conditions which promote DNA hybridization, for example, 6× sodium chloride/sodium citrate (SSC) at about 45° C., followed by a wash of 2×SSC at 50° C., are known or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Typically, stringent conditions for hybridization and detection will be those in which the salt concentration is less than about 1.5 M Na<sup>+</sup> ion, typically about 0.01 to 1.0 M Na<sup>+</sup> ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C. for short probes (such as, for example, 10 to 50 nucleotides) and at least about 60° C. for longer probes (such as, for example, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Optionally, wash buffers may comprise about 0.1% to about 1% SDS. Duration of hybridization is generally less than about 24 hours, usually about 4 to about 12 hours. The duration of the wash time will be at least a length of time sufficient to reach equilibrium.

[0150] The present disclosure also provides methods of detecting the presence of a KLHDC7B predicted loss-of-function polypeptide comprising performing an assay on a sample obtained from a subject to determine whether a KLHDC7B polypeptide in the subject contains one or more variations that causes the polypeptide to have a loss-of-function (partial or complete) or predicted loss-of-function (partial or complete). The KLHDC7B predicted loss-of-function polypeptide can be any of the KLHDC7B truncated variant polypeptides described herein. In some embodiments, the methods detect the presence of KLHDC7B V1145M, V504M, V405M, K822fs, K181fs, K82fs, G943fs, G302fs, or G203fs. In some embodiments, the methods detect the presence of KLHDC7B V1145M. In some embodiments, the methods detect the presence of KLHDC7B V504M. In some embodiments, the methods detect the presence of KLHDC7B V405M. In some embodiments, the methods detect the presence of KLHDC7B K822fs. In some embodiments, the methods detect the presence of KLHDC7B K181fs. In some embodiments, the methods detect the presence of KLHDC7B K82fs. In some embodiments, the methods detect the presence of KLHDC7B G943fs. In some embodiments, the methods

detect the presence of KLHDC7B G302fs. In some embodiments, the methods detect the presence of KLHDC7B G203fs.

[0151] In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether a KLHDC7B polypeptide in the sample comprises a methionine at a position corresponding to position 1,145 according to SEQ ID NO:22. In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether a KLHDC7B polypeptide in the sample comprises a methionine at a position corresponding to position 504 according to SEQ ID NO:23. In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether a KLHDC7B polypeptide in the sample comprises a methionine at a position corresponding to position 405 according to SEQ ID NO:24.

[0152] In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether the sample contains a KLHDC7B truncated variant polypeptide terminating at a position corresponding to position 885 according to SEQ ID NO:43. In some embodiments, the KLHDC7B truncated variant polypeptide lacks amino acids at positions corresponding to positions 886 to 1,235 of SEQ ID NO:19. In some embodiments, the KLHDC7B truncated variant polypeptide comprises or consists of SEQ ID NO:43.

[0153] In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether the sample contains a KLHDC7B truncated variant polypeptide terminating at a position corresponding to position 244 according to SEQ ID NO:44. In some embodiments, the KLHDC7B truncated variant polypeptide lacks amino acids at positions corresponding to positions 245 to 594 of SEQ ID NO:20. In some embodiments, the KLHDC7B truncated variant polypeptide comprises or consists of SEQ ID NO:44.

[0154] In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether the sample contains a KLHDC7B truncated variant polypeptide terminating at a position corresponding to position 145 according to SEQ ID NO:45. In some embodiments, the KLHDC7B truncated variant polypeptide lacks amino acids at positions corresponding to positions 146 to 495 of SEQ ID NO:21. In some embodiments, the KLHDC7B truncated variant polypeptide comprises or consists of SEQ ID NO:45.

[0155] In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether the sample contains a KLHDC7B truncated variant polypeptide terminating at a position corresponding to position 975 according to SEQ ID NO:46. In some embodiments, the KLHDC7B truncated variant polypeptide lacks amino acids at positions corresponding to positions 976 to 1,235 of SEQ ID NO:19. In some embodiments, the KLHDC7B truncated variant polypeptide comprises or consists of SEQ ID NO:46.

[0156] In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether the sample contains a KLHDC7B truncated variant polypeptide terminating at a position corresponding to position 334 according to SEQ ID NO:47. In some embodiments, the KLHDC7B truncated variant polypeptide lacks amino acids at positions corresponding to

positions 335 to 594 of SEQ ID NO:20. In some embodiments, the KLHDC7B truncated variant polypeptide comprises or consists of SEQ ID NO:47.

[0157] In some embodiments, the methods comprise performing an assay on a sample obtained from a subject to determine whether the sample contains a KLHDC7B truncated variant polypeptide terminating at a position corresponding to position 235 according to SEQ ID NO:48. In some embodiments, the KLHDC7B truncated variant polypeptide lacks amino acids at positions corresponding to positions 236 to 495 of SEQ ID NO:21. In some embodiments, the KLHDC7B truncated variant polypeptide comprises or consists of SEQ ID NO:48.

[0158] In some embodiments, the detecting step comprises sequencing at least a portion of the polypeptide that comprises a position corresponding to position 1,145 according to SEQ ID NO:22 or SEQ ID NO:19. In some embodiments, the detecting step comprises sequencing at least a portion of the polypeptide that comprises a position corresponding to position 504 according to SEQ ID NO:23 or SEQ ID NO:20. In some embodiments, the detecting step comprises sequencing at least a portion of the polypeptide that comprises a position corresponding to position 405 according to SEQ ID NO:24 or SEQ ID NO:21.

[0159] In some embodiments, the detecting step comprises sequencing at least a portion of a KLHDC7B polypeptide that may comprise positions corresponding to any positions that are C-terminal to position 885 according to SEQ ID NO:43. If amino acids are detected in the KLHDC7B polypeptide at positions corresponding to positions 886 to 1,235 according to SEQ ID NO:19, then such KLHDC7B polypeptide is a KLHDC7B reference polypeptide. An absence of positions 886 to 1,235 according to SEQ ID NO:19 in the KLHDC7B polypeptide indicates that the KLHDC7B polypeptide terminates at position 885 according to SEQ ID NO:43 and is a KLHDC7B predicted loss-of-function polypeptide.

[0160] In some embodiments, the detecting step comprises sequencing at least a portion of a KLHDC7B polypeptide that may comprise positions corresponding to any positions that are C-terminal to position 244 according to SEQ ID NO:44. If amino acids are detected in the KLHDC7B polypeptide at positions corresponding to positions 245 to 594 according to SEQ ID NO:20, then such KLHDC7B polypeptide is a KLHDC7B reference polypeptide. An absence of positions 245 to 594 according to SEQ ID NO:20 in the KLHDC7B polypeptide indicates that the KLHDC7B polypeptide terminates at position 244 according to SEQ ID NO:44 and is a KLHDC7B predicted loss-of-function polypeptide.

[0161] In some embodiments, the detecting step comprises sequencing at least a portion of a KLHDC7B polypeptide that may comprise positions corresponding to any positions that are C-terminal to position 145 according to SEQ ID NO:45. If amino acids are detected in the KLHDC7B polypeptide at positions corresponding to positions 146 to 495 according to SEQ ID NO:21, then such KLHDC7B polypeptide is a KLHDC7B reference polypeptide. An absence of positions 146 to 495 according to SEQ ID NO:21 in the KLHDC7B polypeptide indicates that the KLHDC7B polypeptide terminates at position 145 according to SEQ ID NO:45 and is a KLHDC7B predicted loss-of-function polypeptide.



KLHDC7B polypeptide at positions corresponding to positions 335 to 594 according to SEQ ID NO:20, then such KLHDC7B polypeptide is a KLHDC7B reference polypeptide. A lack of detection of positions 335 to 594 according to SEQ ID NO:20 in the KLHDC7B polypeptide indicates that the KLHDC7B polypeptide terminates at position 334 according to SEQ ID NO:47 and is a KLHDC7B predicted loss-of-function polypeptide.

[0171] In some embodiments, the detecting step comprises an immunoassay for detecting the presence of a KLHDC7B polypeptide that comprises or consists of SEQ ID NO:48. In some embodiments, the KLHDC7B polypeptide consists of SEQ ID NO:48. In some embodiments, the detecting step comprises detecting at least a portion of a KLHDC7B polypeptide that may comprise positions corresponding to any positions that are C-terminal to position 235 according to SEQ ID NO:48. If amino acids are detected in the KLHDC7B polypeptide at positions corresponding to positions 236 to 495 according to SEQ ID NO:21, then such KLHDC7B polypeptide is a KLHDC7B reference polypeptide. A lack of detection of positions 236 to 495 according to SEQ ID NO:21 in the KLHDC7B polypeptide indicates that the KLHDC7B polypeptide terminates at position 235 according to SEQ ID NO:48 and is a KLHDC7B predicted loss-of-function polypeptide.

[0172] In some embodiments, when the subject does not have a KLHDC7B predicted loss-of-function polypeptide, then the subject does not have an increased risk for developing hearing loss or any of conductive hearing loss, sensorineural hearing loss, or neural hearing loss. In some embodiments, when the subject has a KLHDC7B predicted loss-of-function polypeptide, then the subject has an increased risk for developing hearing loss or any of conductive hearing loss, sensorineural hearing loss, or neural hearing loss.

[0173] The present disclosure also provides isolated nucleic acid molecules that hybridize to KLHDC7B variant genomic nucleic acid molecules, KLHDC7B variant mRNA molecules, and/or KLHDC7B variant cDNA molecules (such as any of the genomic variant nucleic acid molecules, mRNA variant molecules, and cDNA variant molecules disclosed herein). In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes a position corresponding to position 3,778 according to SEQ ID NO:2, position 3,778 according to SEQ ID NO:7, or position 3,778 according to SEQ ID NO:15. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes a position corresponding to position 1,644 according to SEQ ID NO:8, or position 1,644 according to SEQ ID NO:16. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes a position corresponding to position 3,474 according to SEQ ID NO:9, or position 3,474 according to SEQ ID NO:17. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes a position corresponding to position 1,644 according to SEQ ID NO:10, or position 1,644 according to SEQ ID NO:18. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, positions corresponding to positions 2,806-2,807

according to SEQ ID NO:27, or positions corresponding to positions 2,806-2,807 according to SEQ ID NO:35. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes positions corresponding to positions 672-673 according to SEQ ID NO:28, or positions corresponding to positions 672-673 according to SEQ ID NO:36. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes positions corresponding to positions 672-673 according to SEQ ID NO:30, or positions corresponding to positions 672-673 according to SEQ ID NO:38. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes positions corresponding to positions 3,169-3,170 according to SEQ ID NO:26; positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41. In some embodiments, the isolated nucleic acid molecules hybridize to a portion of the KLHDC7B nucleic acid molecule that includes positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42.

[0174] In some embodiments, such isolated nucleic acid molecules comprise or consist of at least about 5, at least about 8, at least about 10, at least about 11, at least about 12, at least about 13, at least about 14, at least about 15, at least about 16, at least about 17, at least about 18, at least about 19, at least about 20, at least about 21, at least about 22, at least about 23, at least about 24, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, at least about 200, at least about 300, at least about 400, at least about 500, at least about 600, at least about 700, at least about 800, at least about 900, at least about 1000, at least about 2000, at least about 3000, at least about 4000, or at least about 5000 nucleotides. In some embodiments, such isolated nucleic acid molecules comprise or consist of at least about 5, at least about 8, at least about 10, at least about 11, at least about 12, at least about 13, at least about 14, at least about 15, at least about 16, at least about 17, at least about 18, at least about 19, at least about 20, at least about 21, at least about 22, at least about 23, at least about 24, or at least about 25 nucleotides. In some embodiments, the isolated nucleic acid molecules comprise or consist of at least about 18 nucleotides. In some

embodiments, the isolated nucleic acid molecules comprise or consists of at least about 15 nucleotides. In some embodiments, the isolated nucleic acid molecules consist of or comprise from about 10 to about 35, from about 10 to about 30, from about 10 to about 25, from about 12 to about 30, from about 12 to about 28, from about 12 to about 24, from about 15 to about 30, from about 15 to about 25, from about 18 to about 30, from about 18 to about 25, from about 18 to about 24, or from about 18 to about 22 nucleotides. In some embodiments, the isolated nucleic acid molecules consist of or comprise from about 18 to about 30 nucleotides. In some embodiments, the isolated nucleic acid molecules comprise or consist of at least about 15 nucleotides to at least about 35 nucleotides.

[0175] In some embodiments, such isolated nucleic acid molecules hybridize to KLHDC7B variant nucleic acid molecules (such as genomic nucleic acid molecules, mRNA molecules, and/or cDNA molecules) under stringent conditions. Such nucleic acid molecules can be used, for example, as probes, primers, alteration-specific probes, or alteration-specific primers as described or exemplified herein, and include, without limitation primers, probes, antisense RNAs, shRNAs, and siRNAs, each of which is described in more detail elsewhere herein, and can be used in any of the methods described herein.

[0176] In some embodiments, the isolated nucleic acid molecules hybridize to at least about 15 contiguous nucleotides of a nucleic acid molecule that is at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% identical to KLHDC7B variant genomic nucleic acid molecules, KLHDC7B variant mRNA molecules, and/or KLHDC7B variant cDNA molecules. In some embodiments, the isolated nucleic acid molecules consist of or comprise from about 15 to about 100 nucleotides, or from about 15 to about 35 nucleotides. In some embodiments, the isolated nucleic acid molecules consist of or comprise from about 15 to about 100 nucleotides. In some embodiments, the isolated nucleic acid molecules consist of or comprise from about 15 to about 35 nucleotides.

[0177] In some embodiments, the isolated alteration-specific probes or alteration-specific primers comprise at least about 15 nucleotides, wherein the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the portion comprises a position corresponding to: position 3,778 according to SEQ ID NO:2, or the complement thereof; position 3,778 according to SEQ ID NO:7, or the complement thereof; or position 3,778 according to SEQ ID NO:15, or the complement thereof. In some embodiments, the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence comprising positions corresponding to: positions 3,778-3,780 according to SEQ ID NO:2, or the complement thereof; positions 3,778-3,780 according to SEQ ID NO:7, or the complement thereof; and/or positions 3,778-3,780 according to SEQ ID NO:15, or the complement thereof.

[0178] In some embodiments, the isolated alteration-specific probes or alteration-specific primers comprise at least about 15 nucleotides, wherein the alteration-specific probe or alteration-specific primer comprises a nucleotide

sequence which is complementary to a portion of a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the portion comprises a position corresponding to: position 1,644 according to SEQ ID NO:8, or the complement thereof; or position 1,644 according to SEQ ID NO:16, or the complement thereof. In some embodiments, the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence comprising positions corresponding to: positions 1,644-1,646 according to SEQ ID NO:8, or the complement thereof and/or positions 1,644-1,646 according to SEQ ID NO:16, or the complement thereof.

[0179] In some embodiments, the isolated alteration-specific probes or alteration-specific primers comprise at least about 15 nucleotides, wherein the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the portion comprises a position corresponding to: position 3,474 according to SEQ ID NO:9, or the complement thereof; or position 3,474 according to SEQ ID NO:17, or the complement thereof. In some embodiments, the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence comprising positions corresponding to: positions 3,474-3,476 according to SEQ ID NO:9, or the complement thereof; and/or positions 3,474-3,476 according to SEQ ID NO:17, or the complement thereof.

[0180] In some embodiments, the isolated alteration-specific probes or alteration-specific primers comprise at least about 15 nucleotides, wherein the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the portion comprises a position corresponding to: position 1,644 according to SEQ ID NO:10, or the complement thereof; or position 1,644 according to SEQ ID NO:18, or the complement thereof. In some embodiments, the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence comprising positions corresponding to: positions 1,644-1,646 according to SEQ ID NO:10, or the complement thereof; and/or positions 1,644-1,646 according to SEQ ID NO:18, or the complement thereof.

[0181] In some embodiments, the isolated alteration-specific probes or alteration-specific primers comprise at least about 15 nucleotides, wherein the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the portion comprises a position corresponding to: a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:27, or the complement thereof; or a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, or the complement thereof. In some embodiments, the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence comprising positions corresponding to: a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:25, or the complement thereof; a CG dinucleotide at positions corresponding to



corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or the complement thereof.

[0188] In some embodiments, the isolated alteration-specific probes or alteration-specific primers comprise at least about 15 nucleotides, wherein the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the portion comprises a position corresponding to: an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof. In some embodiments, the alteration-specific probe or alteration-specific primer comprises a nucleotide sequence which is complementary to a portion of a nucleotide sequence comprising positions corresponding to: an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof.

[0189] In some embodiments, the alteration-specific probes and alteration-specific primers comprise DNA. In some embodiments, the alteration-specific probes and alteration-specific primers comprise RNA.

[0190] In some embodiments, the probes and primers described herein (including alteration-specific probes and alteration-specific primers) have a nucleotide sequence that specifically hybridizes to any of the nucleic acid molecules disclosed herein, or the complement thereof. In some embodiments, the probes and primers specifically hybridize to any of the nucleic acid molecules disclosed herein under stringent conditions.

[0191] In some embodiments, the primers, including alteration-specific primers, can be used in second generation sequencing or high throughput sequencing. In some instances, the primers, including alteration-specific primers, can be modified. In particular, the primers can comprise various modifications that are used at different steps of, for example, Massive Parallel Signature Sequencing (MPSS), Polony sequencing, and 454 Pyrosequencing. Modified primers can be used at several steps of the process, including biotinylated primers in the cloning step and fluorescently labeled primers used at the bead loading step and detection step. Polony sequencing is generally performed using a paired-end tags library wherein each molecule of DNA template is about 135 bp in length. Biotinylated primers are used at the bead loading step and emulsion PCR. Fluorescently labeled degenerate nonamer oligonucleotides are used at the detection step. An adaptor can contain a 5'-biotin tag for immobilization of the DNA library onto streptavidin-coated beads.

[0192] The probes and primers described herein can be used to detect a nucleotide variation within any of the KLHDC7B variant genomic nucleic acid molecules, KLHDC7B variant mRNA molecules, and/or KLHDC7B variant cDNA molecules disclosed herein. The primers described herein can be used to amplify KLHDC7B variant genomic nucleic acid molecules, KLHDC7B variant mRNA molecules, or KLHDC7B variant cDNA molecules, or a fragment thereof.

[0193] The present disclosure also provides pairs of primers comprising any of the primers described above. For

example, if one of the primers' 3'-ends hybridizes to a guanine at a position corresponding to position 3,778 according to SEQ ID NO:1 (rather than adenine) in a particular KLHDC7B nucleic acid molecule, then the presence of the amplified fragment would indicate the presence of a KLHDC7B reference genomic nucleic acid molecule. Conversely, if one of the primers' 3'-ends hybridizes to an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2 (rather than guanine) in a particular KLHDC7B nucleic acid molecule, then the presence of the amplified fragment would indicate the presence of the KLHDC7B variant genomic nucleic acid molecule. In some embodiments, the nucleotide of the primer complementary to the adenine at a position corresponding to position 3,778 according to SEQ ID NO:2 can be at the 3' end of the primer. In addition, if one of the primers' 3'-ends hybridizes to a guanine at a position corresponding to position 3,778 according to SEQ ID NO:3 (rather than adenine) in a particular KLHDC7B nucleic acid molecule, then the presence of the amplified fragment would indicate the presence of a KLHDC7B reference mRNA molecule. Conversely, if one of the primers' 3'-ends hybridizes to an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7 (rather than guanine) in a particular KLHDC7B mRNA molecule, then the presence of the amplified fragment would indicate the presence of the KLHDC7B variant mRNA molecule. In some embodiments, the nucleotide of the primer complementary to the adenine at a position corresponding to position 3,778 according to SEQ ID NO:7 can be at the 3' end of the primer. In addition, if one of the primers' 3'-ends hybridizes to a guanine at a position corresponding to position 3,778 according to SEQ ID NO:11 (rather than adenine) in a particular KLHDC7B nucleic acid molecule, then the presence of the amplified fragment would indicate the presence of a KLHDC7B reference cDNA molecule. Conversely, if one of the primers' 3'-ends hybridizes to an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15 (rather than guanine) in a particular KLHDC7B cDNA molecule, then the presence of the amplified fragment would indicate the presence of the KLHDC7B variant cDNA molecule. In some embodiments, the nucleotide of the primer complementary to the adenine at a position corresponding to position 3,778 according to SEQ ID NO:15 can be at the 3' end of the primer.

[0194] If, for example, one of the primers' 3'-ends hybridizes to a guanine at a position corresponding to position 1,644 according to SEQ ID NO:4 (rather than adenine) in a particular KLHDC7B nucleic acid molecule, then the presence of the amplified fragment would indicate the presence of a KLHDC7B reference mRNA molecule. Conversely, if one of the primers' 3'-ends hybridizes to an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8 (rather than guanine) in a particular KLHDC7B mRNA molecule, then the presence of the amplified fragment would indicate the presence of the KLHDC7B variant mRNA molecule. In some embodiments, the nucleotide of the primer complementary to the adenine at a position corresponding to position 1,644 according to SEQ ID NO:8 can be at the 3' end of the primer. In addition, if one of the primers' 3'-ends hybridizes to a guanine at a position corresponding to position 1,644 according to SEQ ID NO:12 (rather than adenine) in a particular KLHDC7B nucleic acid molecule, then the presence of the amplified fragment would







primer complementary to the AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34 can be at the 3' end of the primer. In addition, if one of the primers' 3'-ends hybridizes to an AGG trinucleotide at positions corresponding to positions 1,035-1,037 according to SEQ ID NO:14 (rather than an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42) in a particular KLHDC7B nucleic acid molecule, then the presence of the amplified fragment would indicate the presence of a KLHDC7B reference cDNA molecule. Conversely, if one of the primers' 3'-ends hybridizes to an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42 (rather than AGG trinucleotide at positions corresponding to positions 1,035-1,037 according to SEQ ID NO:14) in a particular KLHDC7B cDNA molecule, then the presence of the amplified fragment would indicate the presence of the KLHDC7B variant cDNA molecule. In some embodiments, the nucleotide of the primer complementary to the AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42 can be at the 3' end of the primer.

[0205] In the context of the disclosure "specifically hybridizes" means that the probe or primer (such as, for example, the alteration-specific probe or alteration-specific primer) does not hybridize to a nucleic acid sequence encoding a KLHDC7B reference genomic nucleic acid molecule, a KLHDC7B reference mRNA molecule, and/or a KLHDC7B reference cDNA molecule.

[0206] In some embodiments, the probes (such as, for example, an alteration-specific probe) comprise a label. In some embodiments, the label is a fluorescent label, a radiolabel, or biotin.

[0207] The present disclosure also provides supports comprising a substrate to which any one or more of the probes disclosed herein is attached. Solid supports are solid-state substrates or supports with which molecules, such as any of the probes disclosed herein, can be associated. A form of solid support is an array. Another form of solid support is an array detector. An array detector is a solid support to which multiple different probes have been coupled in an array, grid, or other organized pattern. A form for a solid-state substrate is a microtiter dish, such as a standard 96-well type. In some embodiments, a multiwell glass slide can be employed that normally contains one array per well.

[0208] The present disclosure also provides molecular complexes comprising or consisting of any of the KLHDC7B nucleic acid molecules (genomic nucleic acid molecules, mRNA molecules, or cDNA molecules), or complement thereof, described herein and any of the alteration-specific primers or alteration-specific probes described herein. In some embodiments, the KLHDC7B nucleic acid molecules (genomic nucleic acid molecules, mRNA molecules, or cDNA molecules), or complement thereof, in the molecular complexes are single-stranded. In some embodiments, the KLHDC7B nucleic acid molecule is any of the genomic nucleic acid molecules described herein. In some embodiments, the KLHDC7B nucleic acid molecule is any of the mRNA molecules described herein. In some embodiments, the KLHDC7B nucleic acid molecule is any of the cDNA molecules described herein. In some embodiments, the molecular complex comprises or consists of any of the KLHDC7B nucleic acid molecules (genomic nucleic acid molecules, mRNA molecules, or cDNA molecules), or

complement thereof, described herein and any of the alteration-specific primers described herein. In some embodiments, the molecular complex comprises or consists of any of the KLHDC7B nucleic acid molecules (genomic nucleic acid molecules, mRNA molecules, or cDNA molecules), or complement thereof, described herein and any of the alteration-specific probes described herein.

[0209] In some embodiments, the molecular complex comprises or consists of an alteration-specific primer or an alteration-specific probe hybridized to a genomic nucleic acid molecule comprising a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the alteration-specific primer or the alteration-specific probe is hybridized to an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof.

[0210] In some embodiments, the molecular complex comprises or consists of an alteration-specific primer or an alteration-specific probe that is hybridized to an ATG codon at positions corresponding to positions 3,778-3,780 according to SEQ ID NO:2.

[0211] In some embodiments, the molecular complex comprises or consists of a genomic nucleic acid molecule that comprises SEQ ID NO:2.

[0212] In some embodiments, the molecular complex comprises or consists of an alteration-specific primer or an alteration-specific probe hybridized to an mRNA molecule comprising a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the alteration-specific primer or the alteration-specific probe is hybridized to: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or the complement thereof; an adenine at a position corresponding to position 3,474 according to SEQ ID NO:9, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:10, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 according to SEQ ID NO:27, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:29, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or the complement thereof; an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:31, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or the complement thereof; an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:33, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof.

[0213] In some embodiments, the molecular complex comprises or consists of an alteration-specific primer or an alteration-specific probe that is hybridized to: an AUG codon at positions corresponding to positions 3,778-3,780 according to SEQ ID NO:7, an AUG codon at positions corresponding to positions 1,644-1,646 according to SEQ ID NO:8, an AUG codon at positions corresponding to positions 3,474-3,476 according to SEQ ID NO:9, or an

AUG codon at positions corresponding to positions 1,644-1,646 according to SEQ ID NO:10.

[0214] In some embodiments, the molecular complex comprises or consists of an mRNA molecule that comprises SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, or SEQ ID NO:34.

[0215] In some embodiments, the molecular complex comprises or consists of an alteration-specific primer or an alteration-specific probe hybridized to a cDNA molecule comprising a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the alteration-specific primer or the alteration-specific probe is hybridized to: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or the complement thereof; an adenine at a position corresponding to position 3,474 according to SEQ ID NO:17, or the complement thereof; an adenine at a position corresponding to position 1,644 according to SEQ ID NO:18, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,806-2,807 to SEQ ID NO:35, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, or the complement thereof; a CG dinucleotide at positions corresponding to positions 2,502-2,503 according to SEQ ID NO:37, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, or the complement thereof; an AG dinucleotide at positions corresponding to positions 3,169-3,170 according to SEQ ID NO:39, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or the complement thereof; an AG dinucleotide at positions corresponding to positions 2,865-2,866 according to SEQ ID NO:41, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof.

[0216] In some embodiments, the molecular complex comprises or consists of an alteration-specific primer or an alteration-specific probe that is hybridized to: an ATG codon at positions corresponding to positions 3,778-3,780 according to SEQ ID NO:15, an ATG codon at positions corresponding to positions 1,644-1,646 according to SEQ ID NO:16, an ATG codon at positions corresponding to positions 3,474-3,476 according to SEQ ID NO:17, or an ATG codon at positions corresponding to positions 1,644-1,646 according to SEQ ID NO:18.

[0217] In some embodiments, the molecular complex comprises or consists of a cDNA molecule that comprises SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, or SEQ ID NO:42.

[0218] In some embodiments, the molecular complex comprises an alteration-specific probe or an alteration-specific primer comprising a label. In some embodiments, the label is a fluorescent label, a radiolabel, or biotin. In some embodiments, the molecular complex further comprises a non-human polymerase.

[0219] The nucleotide sequence of a KLHDC7B reference genomic nucleic acid molecule (hg38 chr22: 50,545,899-

50,551,023; ENST00000648057.3) is set forth in SEQ ID NO:1. Referring to SEQ ID NO:1, position 3,778 is a guanine.

[0220] A variant genomic nucleic acid molecule of KLHDC7B exists, wherein the guanine at position 3,778 (referring to SEQ ID NO:1) is replaced with an adenine. The nucleotide sequence of this KLHDC7B variant genomic nucleic acid molecule is set forth in SEQ ID NO:2 (rs36062310).

[0221] Another variant genomic nucleic acid molecule of KLHDC7B exists, wherein a guanine at position 2,807 (referring to SEQ ID NO:1) is deleted. The nucleotide sequence of this KLHDC7B variant genomic nucleic acid molecule is set forth in SEQ ID NO:25 (rs746113253).

[0222] Another variant genomic nucleic acid molecule of KLHDC7B exists, wherein a guanine at position 3,170 (referring to SEQ ID NO:1) is deleted. The nucleotide sequence of this KLHDC7B variant genomic nucleic acid molecule is set forth in SEQ ID NO:26 (rs749405486).

[0223] The nucleotide sequence of a KLHDC7B reference mRNA molecule is set forth in SEQ ID NO:3 (ENST00000648057.3). Referring to SEQ ID NO:3, position 3,778 is a guanine. The nucleotide sequence of another KLHDC7B reference mRNA molecule is set forth in SEQ ID NO:4 (ENST00000395676.4). Referring to SEQ ID NO:4, position 1,644 is a guanine.

[0224] The nucleotide sequence of another KLHDC7B reference mRNA molecule is set forth in SEQ ID NO:5 (NM\_138433.4). Referring to SEQ ID NO:5, position 3,474 is a guanine.

[0225] The nucleotide sequence of another KLHDC7B reference mRNA molecule is set forth in SEQ ID NO:6 (BC009980). Referring to SEQ ID NO:6, position 1,644 is a guanine.

[0226] A variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 3,778 (referring to SEQ ID NO:3) is replaced with an adenine. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:7 (ENST00000648057.3).

[0227] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 1,644 (referring to SEQ ID NO:4) is replaced with an adenine. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:8 (ENST00000395676.4).

[0228] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 3,474 (referring to SEQ ID NO:5) is replaced with an adenine. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:9 (NM\_138433.4).

[0229] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 1,644 (referring to SEQ ID NO:6) is replaced with an adenine. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:10 (BC009980).

[0230] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 2,807 (referring to SEQ ID NO:3) is deleted. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:27 (ENST00000648057.3).

[0231] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 673 (referring to SEQ ID NO:4) is deleted. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:28 (ENST00000395676.4).

[0232] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 2,503 (referring to SEQ ID NO:5) is deleted. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:29 (NM\_138433.4).

[0233] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 673 (referring to SEQ ID NO:6) is deleted. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:30 (BC009980).

[0234] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 3,170 (referring to SEQ ID NO:3) is deleted. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:31 (ENST00000648057.3).

[0235] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 1,036 (referring to SEQ ID NO:4) is deleted. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:32 (ENST00000395676.4).

[0236] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 2,866 (referring to SEQ ID NO:5) is deleted. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:33 (NM\_138433.4).

[0237] Another variant mRNA molecule of KLHDC7B exists, wherein the guanine at position 1,036 (referring to SEQ ID NO:6) is deleted. The nucleotide sequence of this KLHDC7B variant mRNA molecule is set forth in SEQ ID NO:34 (BC009980).

[0238] The nucleotide sequence of a KLHDC7B reference cDNA molecule is set forth in SEQ ID NO:11 (ENST00000648057.3). Referring to SEQ ID NO:11, position 3,778 is a guanine.

[0239] The nucleotide sequence of another KLHDC7B reference cDNA molecule is set forth in SEQ ID NO:12 (ENST00000395676.4). Referring to SEQ ID NO:12, position 1,644 is a guanine.

[0240] The nucleotide sequence of another KLHDC7B reference cDNA molecule is set forth in SEQ ID NO:13 (NM\_138433.4). Referring to SEQ ID NO:13, position 3,474 is a guanine.

[0241] The nucleotide sequence of another KLHDC7B reference cDNA molecule is set forth in SEQ ID NO:14 (BC009980). Referring to SEQ ID NO:14, position 1,644 is a guanine.

[0242] A variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 3,778 (referring to SEQ ID NO:11) is replaced with an adenine. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:15 (ENST00000648057.3).

[0243] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 1,644 (referring to SEQ ID NO:12) is replaced with an adenine. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:16 (ENST00000395676.4).

[0244] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 3,474 (referring to SEQ ID NO:13) is replaced with an adenine. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:17 (NM\_138433.4).

[0245] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 1,644 (referring to SEQ ID NO:14) is replaced with an adenine. The nucleotide

sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:18 (BC009980).

[0246] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 2,807 (referring to SEQ ID NO:11) is deleted. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:35 (ENST00000648057.3).

[0247] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 673 (referring to SEQ ID NO:12) is deleted. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:36 (ENST00000395676.4).

[0248] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 2,503 (referring to SEQ ID NO:13) is deleted. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:37 (NM\_138433.4).

[0249] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 673 (referring to SEQ ID NO:14) is deleted. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:38 (BC009980).

[0250] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 3,170 (referring to SEQ ID NO:11) is deleted. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:39 (ENST00000648057.3).

[0251] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 1,036 (referring to SEQ ID NO:12) is deleted. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:40 (ENST00000395676.4).

[0252] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 2,866 (referring to SEQ ID NO:13) is deleted. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:41 (NM\_138433.4).

[0253] Another variant cDNA molecule of KLHDC7B exists, wherein the guanine at position 1,036 (referring to SEQ ID NO:14) is deleted. The nucleotide sequence of this KLHDC7B variant cDNA molecule is set forth in SEQ ID NO:42 (BC009980).

[0254] The genomic nucleic acid molecules, mRNA molecules, and cDNA molecules can be from any organism. For example, the genomic nucleic acid molecules, mRNA molecules, and cDNA molecules can be human or an ortholog from another organism, such as a non-human mammal, a rodent, a mouse, or a rat. It is understood that gene sequences within a population can vary due to polymorphisms such as single-nucleotide polymorphisms. The examples provided herein are only exemplary sequences. Other sequences are also possible.

[0255] Also provided herein are functional polynucleotides that can interact with the disclosed nucleic acid molecules. Examples of functional polynucleotides include, but are not limited to, antisense molecules, aptamers, ribozymes, triplex forming molecules, and external guide sequences. The functional polynucleotides can act as effectors, inhibitors, modulators, and stimulators of a specific activity possessed by a target molecule, or the functional polynucleotides can possess a de novo activity independent of any other molecules.

[0256] The isolated nucleic acid molecules disclosed herein can comprise RNA, DNA, or both RNA and DNA.

The isolated nucleic acid molecules can also be linked or fused to a heterologous nucleic acid sequence, such as in a vector, or a heterologous label. For example, the isolated nucleic acid molecules disclosed herein can be within a vector or as an exogenous donor sequence comprising the isolated nucleic acid molecule and a heterologous nucleic acid sequence. The isolated nucleic acid molecules can also be linked or fused to a heterologous label. The label can be directly detectable (such as, for example, fluorophore) or indirectly detectable (such as, for example, hapten, enzyme, or fluorophore quencher). Such labels can be detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. Such labels include, for example, radiolabels, pigments, dyes, chromogens, spin labels, and fluorescent labels. The label can also be, for example, a chemiluminescent substance; a metal-containing substance; or an enzyme, where there occurs an enzyme-dependent secondary generation of signal. The term "label" can also refer to a "tag" or hapten that can bind selectively to a conjugated molecule such that the conjugated molecule, when added subsequently along with a substrate, is used to generate a detectable signal. For example, biotin can be used as a tag along with an avidin or streptavidin conjugate of horseradish peroxidase (HRP) to bind to the tag, and examined using a calorimetric substrate (such as, for example, tetramethylbenzidine (TMB)) or a fluorogenic substrate to detect the presence of HRP. Exemplary labels that can be used as tags to facilitate purification include, but are not limited to, myc, HA, FLAG or 3×FLAG, 6×His or polyhistidine, glutathione-S-transferase (GST), maltose binding protein, an epitope tag, or the Fc portion of immunoglobulin. Numerous labels include, for example, particles, fluorophores, haptens, enzymes and their calorimetric, fluorogenic and chemiluminescent substrates and other labels.

**[0257]** The disclosed nucleic acid molecules can comprise, for example, nucleotides or non-natural or modified nucleotides, such as nucleotide analogs or nucleotide substitutes. Such nucleotides include a nucleotide that contains a modified base, sugar, or phosphate group, or that incorporates a non-natural moiety in its structure. Examples of non-natural nucleotides include, but are not limited to, dideoxynucleotides, biotinylated, aminated, deaminated, alkylated, benzylated, and fluorophor-labeled nucleotides.

**[0258]** The nucleic acid molecules disclosed herein can also comprise one or more nucleotide analogs or substitutions. A nucleotide analog is a nucleotide which contains a modification to either the base, sugar, or phosphate moieties. Modifications to the base moiety include, but are not limited to, natural and synthetic modifications of A, C, G, and T/U, as well as different purine or pyrimidine bases such as, for example, pseudouridine, uracil-5-yl, hypoxanthin-9-yl (I), and 2-aminoadenin-9-yl. Modified bases include, but are not limited to, 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl uracil and cytosine, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo (such as, for example, 5-bromo), 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine, 7-methylad-

enine, 8-azaguanine, 8-azaadenine, 7-deazaguanine, 7-deazaadenine, 3-deazaguanine, and 3-deazaadenine.

**[0259]** Nucleotide analogs can also include modifications of the sugar moiety. Modifications to the sugar moiety include, but are not limited to, natural modifications of the ribose and deoxy ribose as well as synthetic modifications. Sugar modifications include, but are not limited to, the following modifications at the 2' position: OH; F; O-, S-, or N-alkyl; O-, S-, or N-alkenyl; O-, S- or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl, alkenyl, and alkynyl may be substituted or unsubstituted C<sub>1-10</sub>alkyl or C<sub>2-10</sub>alkenyl, and C<sub>2-10</sub>alkynyl. Exemplary 2' sugar modifications also include, but are not limited to, —O[(CH<sub>2</sub>)<sub>n</sub>O]<sub>m</sub>CH<sub>3</sub>, —O(CH<sub>2</sub>)<sub>n</sub>OCH<sub>3</sub>, —O(CH<sub>2</sub>)<sub>n</sub>NH<sub>2</sub>, —O(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>, —O(CH<sub>2</sub>)<sub>n</sub>—ONH<sub>2</sub>, and —O(CH<sub>2</sub>)<sub>n</sub>ON[(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>]<sub>2</sub>, where n and m are from 1 to about 10. Other modifications at the 2' position include, but are not limited to, C<sub>1-10</sub>alkyl, substituted lower alkyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH<sub>3</sub>, OCN, Cl, Br, CN, CF<sub>3</sub>, OCF<sub>3</sub>, SOCH<sub>3</sub>, SO<sub>2</sub>CH<sub>3</sub>, ONO<sub>2</sub>, NO<sub>2</sub>, N<sub>3</sub>, NH<sub>2</sub>, heterocycloalkyl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. Similar modifications may also be made at other positions on the sugar, particularly the 3' position of the sugar on the 3' terminal nucleotide or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Modified sugars can also include those that contain modifications at the bridging ring oxygen, such as CH<sub>2</sub> and S. Nucleotide sugar analogs can also have sugar mimetics, such as cyclobutyl moieties in place of the pentofuranosyl sugar.

**[0260]** Nucleotide analogs can also be modified at the phosphate moiety. Modified phosphate moieties include, but are not limited to, those that can be modified so that the linkage between two nucleotides contains a phosphorothioate, chiral phosphorothioate, phosphorodithioate, phosphotriester, aminoalkylphosphotriester, methyl and other alkyl phosphonates including 3'-alkylene phosphonate and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates. These phosphate or modified phosphate linkage between two nucleotides can be through a 3'-5' linkage or a 2'-5' linkage, and the linkage can contain inverted polarity such as 3'-5' to 5'-3' or 2'-5' to 5'-2'. Various salts, mixed salts, and free acid forms are also included. Nucleotide substitutes also include peptide nucleic acids (PNAs).

**[0261]** The present disclosure also provides vectors comprising any one or more of the nucleic acid molecules disclosed herein. In some embodiments, the vectors comprise any one or more of the nucleic acid molecules disclosed herein and a heterologous nucleic acid. The vectors can be viral or nonviral vectors capable of transporting a nucleic acid molecule. In some embodiments, the vector is a plasmid or cosmid (such as, for example, a circular double-stranded DNA into which additional DNA segments can be ligated). In some embodiments, the vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Expression vectors include, but are not limited to, plasmids, cosmids, retroviruses, adenoviruses,

adeno-associated viruses (AAV), plant viruses such as cauliflower mosaic virus and tobacco mosaic virus, yeast artificial chromosomes (YACs), Epstein-Barr (EBV)-derived episomes, and other expression vectors known in the art.

[0262] Desired regulatory sequences for mammalian host cell expression can include, for example, viral elements that direct high levels of polypeptide expression in mammalian cells, such as promoters and/or enhancers derived from retroviral LTRs, cytomegalovirus (CMV) (such as, for example, CMV promoter/enhancer), Simian Virus 40 (SV40) (such as, for example, SV40 promoter/enhancer), adenovirus, (such as, for example, the adenovirus major late promoter (AdMLP)), polyoma and strong mammalian promoters such as native immunoglobulin and actin promoters. Methods of expressing polypeptides in bacterial cells or fungal cells (such as, for example, yeast cells) are also well known. A promoter can be, for example, a constitutively active promoter, a conditional promoter, an inducible promoter, a temporally restricted promoter (such as, for example, a developmentally regulated promoter), or a spatially restricted promoter (such as, for example, a cell-specific or tissue-specific promoter).

[0263] Percent identity (or percent complementarity) between particular stretches of nucleotide sequences within nucleic acid molecules or amino acid sequences within polypeptides can be determined routinely using BLAST programs (basic local alignment search tools) and PowerBLAST programs (Altschul et al., J. Mol. Biol., 1990, 215, 403-410; Zhang and Madden, Genome Res., 1997, 7, 649-656) or by using the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison Wis.), using default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489). Herein, if reference is made to percent sequence identity, the higher percentages of sequence identity are preferred over the lower ones.

[0264] The present disclosure also provides compositions comprising any one or more of the isolated nucleic acid molecules, genomic nucleic acid molecules, mRNA molecules, and/or cDNA molecules disclosed herein. In some embodiments, the composition is a pharmaceutical composition. In some embodiments, the compositions comprise a carrier and/or excipient. Examples of carriers include, but are not limited to, poly (lactic acid) (PLA) microspheres, poly (D,L-lactic-coglycolic-acid) (PLGA) microspheres, liposomes, micelles, inverse micelles, lipid cochleates, and lipid microtubules. A carrier may comprise a buffered salt solution such as PBS, HBSS, etc.

[0265] As used herein, the phrase "corresponding to" or grammatical variations thereof when used in the context of the numbering of a particular nucleotide or nucleotide sequence or position refers to the numbering of a specified reference sequence when the particular nucleotide or nucleotide sequence is compared to a reference sequence (such as, for example, SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:11). In other words, the residue (such as, for example, nucleotide or amino acid) number or residue (such as, for example, nucleotide or amino acid) position of a particular polymer is designated with respect to the reference sequence rather than by the actual numerical position of the residue within the particular nucleotide or nucleotide sequence. For example, a particular nucleotide sequence can be aligned to a reference sequence by introducing gaps to optimize resi-

due matches between the two sequences. In these cases, although the gaps are present, the numbering of the residue in the particular nucleotide or nucleotide sequence is made with respect to the reference sequence to which it has been aligned.

[0266] For example, a nucleic acid molecule comprising a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2 means that if the nucleotide sequence of the KLHDC7B genomic nucleic acid molecule is aligned to the sequence of SEQ ID NO:2, the KLHDC7B sequence has an adenine residue at the position that corresponds to position 3,778 of SEQ ID NO:2. The same applies for mRNA molecules comprising a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, and cDNA molecules comprising a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15. In other words, these phrases refer to a nucleic acid molecule encoding a KLHDC7B polypeptide, wherein the genomic nucleic acid molecule has a nucleotide sequence that comprises an adenine residue that is homologous to the adenine residue at position 3,778 of SEQ ID NO:2 (or wherein the mRNA molecule has a nucleotide sequence that comprises an adenine residue that is homologous to the adenine residue at position 3,778 of SEQ ID NO:7, or wherein the cDNA molecule has a nucleotide sequence that comprises an adenine residue that is homologous to the adenine residue at position 3,778 of SEQ ID NO:15).

[0267] As described herein, a position within a KLHDC7B genomic nucleic acid molecule that corresponds to position 3,778 according to SEQ ID NO:2, for example, can be identified by performing a sequence alignment between the nucleotide sequence of a particular KLHDC7B nucleic acid molecule and the nucleotide sequence of SEQ ID NO:2. A variety of computational algorithms exist that can be used for performing a sequence alignment to identify a nucleotide position that corresponds to, for example, position 3,778 in SEQ ID NO:2. For example, by using the NCBI BLAST algorithm (Altschul et al., Nucleic Acids Res., 1997, 25, 3389-3402) or CLUSTALW software (Sievers and Higgins, Methods Mol. Biol., 2014, 1079, 105-116) sequence alignments may be performed. However, sequences can also be aligned manually.

[0268] The amino acid sequence of a KLHDC7B reference polypeptide is set forth in SEQ ID NO:19. Referring to SEQ ID NO:19, the KLHDC7B reference polypeptide is 1,235 amino acids in length. Referring to SEQ ID NO:19, position 1,145 is valine, position 822 is lysine, and position 943 is glycine.

[0269] The amino acid sequence of another KLHDC7B reference polypeptide is set forth in SEQ ID NO:20. Referring to SEQ ID NO:20, the KLHDC7B reference polypeptide is 594 amino acids in length. Referring to SEQ ID NO:20, position 504 is valine, position 181 is lysine, and position 302 is glycine.

[0270] The amino acid sequence of another KLHDC7B reference polypeptide is set forth in SEQ ID NO:21. Referring to SEQ ID NO:21, the KLHDC7B reference polypep-

tide is 495 amino acids in length. Referring to SEQ ID NO:21, position 405 is valine, position 82 is lysine, and position 203 is glycine.

[0271] A KLHDC7B variant polypeptide exists (V1145M or Val1145Met), the amino acid sequence of which is set forth in SEQ ID NO:22. Referring to SEQ ID NO:22, the KLHDC7B variant polypeptide is 1,235 amino acids in length. Referring to SEQ ID NO:22, position 1,145 is methionine.

[0272] Another KLHDC7B variant polypeptide exists (V504M or Val504Met), the amino acid sequence of which is set forth in SEQ ID NO:23. Referring to SEQ ID NO:23, the KLHDC7B variant polypeptide is 594 amino acids in length. Referring to SEQ ID NO:23, position 504 is methionine.

[0273] Another KLHDC7B variant polypeptide exists (V405M or Val405Met), the amino acid sequence of which is set forth in SEQ ID NO:24. Referring to SEQ ID NO:24, the KLHDC7B variant polypeptide is 495 amino acids in length. Referring to SEQ ID NO:24, position 405 is methionine.

[0274] A KLHDC7B truncated variant polypeptide exists (K822fs or Lys822fs), the amino acid sequence of which is set forth in SEQ ID NO:43. Referring to SEQ ID NO:43, the KLHDC7B variant polypeptide is 885 amino acids in length. Referring to SEQ ID NO:43, the KLHDC7B variant polypeptide is truncated at position 885 and does not contain amino acids at positions corresponding to positions 886 to 1,235 of SEQ ID NO:19. Referring to SEQ ID NO:43, position 822 is serine.

[0275] Another KLHDC7B truncated variant polypeptide exists (K181fs or Lys181fs), the amino acid sequence of which is set forth in SEQ ID NO:44. Referring to SEQ ID NO:44, the KLHDC7B variant polypeptide is 244 amino acids in length. Referring to SEQ ID NO:44, the KLHDC7B variant polypeptide is truncated at position 244 and does not contain amino acids at positions corresponding to positions 245 to 594 of SEQ ID NO:20. Referring to SEQ ID NO:44, position 181 is serine.

[0276] Another KLHDC7B truncated variant polypeptide exists (K82fs or Lys82fs), the amino acid sequence of which is set forth in SEQ ID NO:45. Referring to SEQ ID NO:45, the KLHDC7B variant polypeptide is 145 amino acids in length. Referring to SEQ ID NO:45, the KLHDC7B variant polypeptide is truncated at position 145 and does not contain amino acids at positions corresponding to positions 146 to 495 of SEQ ID NO:21. Referring to SEQ ID NO:45, position 82 is serine.

[0277] Another KLHDC7B truncated variant polypeptide exists (G943fs or Gly943fs), the amino acid sequence of which is set forth in SEQ ID NO:46. Referring to SEQ ID NO:46, the KLHDC7B variant polypeptide is 975 amino acids in length. Referring to SEQ ID NO:46, the KLHDC7B variant polypeptide is truncated at position 975 and does not contain amino acids at positions corresponding to positions 976 to 1,235 of SEQ ID NO:19. Referring to SEQ ID NO:46, position 943 is arginine.

[0278] Another KLHDC7B truncated variant polypeptide exists (G302fs or Gly302fs), the amino acid sequence of which is set forth in SEQ ID NO:47. Referring to SEQ ID NO:47, the KLHDC7B variant polypeptide is 334 amino acids in length. Referring to SEQ ID NO:47, the KLHDC7B variant polypeptide is truncated at position 334 and does not

contain amino acids at positions corresponding to positions 335 to 594 of SEQ ID NO:20. Referring to SEQ ID NO:47, position 302 is arginine.

[0279] Another KLHDC7B truncated variant polypeptide exists (G203fs or Gly203fs), the amino acid sequence of which is set forth in SEQ ID NO:48. Referring to SEQ ID NO:48, the KLHDC7B variant polypeptide is 235 amino acids in length. Referring to SEQ ID NO:48, the KLHDC7B variant polypeptide is truncated at position 235 and does not contain amino acids at positions corresponding to positions 236 to 495 of SEQ ID NO:21. Referring to SEQ ID NO:48, position 203 is arginine.

[0280] The nucleotide and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three-letter code for amino acids. The nucleotide sequences follow the standard convention of beginning at the 5' end of the sequence and proceeding forward (i.e., from left to right in each line) to the 3' end. Only one strand of each nucleotide sequence is shown, but the complementary strand is understood to be included by any reference to the displayed strand. The amino acid sequence follows the standard convention of beginning at the amino terminus of the sequence and proceeding forward (i.e., from left to right in each line) to the carboxy terminus.

[0281] The present disclosure also provides therapeutic agents that treat or inhibit hearing loss for use in the treatment of hearing loss (or for use in the preparation of a medicament for treating hearing loss) in a subject, wherein the subject has any of the genomic nucleic acid molecules, mRNA molecules, and/or cDNA molecules encoding a KLHDC7B polypeptide described herein. The therapeutic agents that treat or inhibit hearing loss can be any of the therapeutic agents that treat or inhibit hearing loss described herein.

[0282] In some embodiments, the subject comprises: a genomic nucleic acid molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises: an adenine at a position corresponding to position 3,778 according to SEQ ID NO:2, or the complement thereof; an mRNA molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:7, or the complement thereof; a cDNA molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an adenine at a position corresponding to position 3,778 according to SEQ ID NO:15, or the complement thereof; or a KLHDC7B polypeptide that comprises a methionine at a position corresponding to position 1,145 according to SEQ ID NO:22.

[0283] In some embodiments, the subject comprises: an mRNA molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an adenine at a position corresponding to position 1,644 according to SEQ ID NO:8, or the complement thereof; a cDNA molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an adenine at a position corresponding to position 1,644 according to SEQ ID NO:16, or the complement thereof; or a KLHDC7B polypeptide that comprises a methionine at a position corresponding to position 504 according to SEQ ID NO:23.



complement thereof; or a KLHDC7B polypeptide that comprises an arginine at a position corresponding to position 203 according to SEQ ID NO:48.

[0293] In some embodiments, the subject comprises: a genomic nucleic acid molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises: an mRNA molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof; or a cDNA molecule having a nucleotide sequence encoding a KLHDC7B polypeptide, wherein the nucleotide sequence comprises an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof.

[0294] All patent documents, websites, other publications, accession numbers and the like cited above or below are incorporated by reference in their entirety for all purposes to the same extent as if each individual item were specifically and individually indicated to be so incorporated by reference. If different versions of a sequence are associated with an accession number at different times, the version associated with the accession number at the effective filing date of this application is meant. The effective filing date means the earlier of the actual filing date or filing date of a priority application referring to the accession number if applicable. Likewise, if different versions of a publication, website or the like are published at different times, the version most recently published at the effective filing date of the application is meant unless otherwise indicated. Any feature, step, element, embodiment, or aspect of the present disclosure can be used in combination with any other feature, step, element, embodiment, or aspect unless specifically indicated otherwise. Although the present disclosure has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims.

[0295] The following examples are provided to describe the embodiments in greater detail. They are intended to illustrate, not to limit, the claimed embodiments. The following examples provide those of ordinary skill in the art

with a disclosure and description of how the compounds, compositions, articles, devices and/or methods described herein are made and evaluated, and are intended to be purely exemplary and are not intended to limit the scope of any claims. Efforts have been made to ensure accuracy with respect to numbers (such as, for example, amounts, temperature, etc.), but some errors and deviations may be accounted for. Unless indicated otherwise, parts are parts by weight, temperature is in ° C. or is at ambient temperature, and pressure is at or near atmospheric.

## EXAMPLES

### Example 1: A Missense Variant and Predicted Loss-of-Function Variants in KLHDC7B are Associated with Increased Risk for Hearing Loss

[0296] A genome-wide and exome-wide analysis of self-reported and ICD code based hearing loss was carried out in UK Biobank, Geisinger (GHS) and other datasets. A common missense variant (FIG. 1, Panel C) and two rare, predicted loss-of-function (pLOF) variants (FIG. 1, Panels A and B) in KLHDC7B were associated with increased risk for hearing loss in meta-analysis of UK Biobank and 3 other cohorts. In addition, an aggregate of rare (minor allele frequency of less than 1%), pLOF variants in KLHDC7B also show an association with increased risk for hearing loss in the meta-analysis (FIG. 2) suggesting that KLHDC7B loss of function variants in addition to the two described in FIG. 1 confer an increased risk for hearing loss in carriers. The association with loss of function variants further suggests that reduced function of KLHDC7B is detrimental to hearing ability.

[0297] Various modifications of the described subject matter, in addition to those described herein, will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims. Each reference (including, but not limited to, journal articles, U.S. and non-U.S. patents, patent application publications, international patent application publications, gene bank accession numbers, and the like) cited in the present application is incorporated herein by reference in its entirety and for all purposes.

---

## SEQUENCE LISTING

```

Sequence total quantity: 48
SEQ ID NO: 1      moltype = DNA length = 5125
FEATURE          Location/Qualifiers
source           1..5125
mol_type = genomic DNA
organism = Homo sapiens

SEQUENCE: 1
agaaggcaggat tggctgtat gacagcacag agctcaaggaa cgctgcctga ggacacctggg 60
gcctacgagg aggagaagag ggcaggagat tttgggggtgc ttgcagagac cctggggctcc 120
tatccctgcca taaggctcgcc tatcttccctga tatctgcacg caggccctac tgacaccccc 180
aggcctgagt gcaaggcaggac acccccacat tccaggcccc tggaggactg gtccaccccta 240
actgggcaggc ccttggggca ggcgctggcc ggtgcctcagg cccaggccctc tggctctgc 300
atgcactgccc agccctccat cttgcacccctt gaaccatgtt ccaggccacc 360
ttggagcccg atggccccctt ctggggctgg gactggggaca gtgacaatgtt ctggggatgt 420
gctgtgttggcc cccttcctggc gttggctgtt gttggctggca cttttttttt cttttttttt 480
tttgggttggcc ggcacgatca agaggccggca gaaccgggtt ccacggccctt cggggctcaa 540
cctcatcagg caggaggagc tggatgtggcc ctgcaacccga agtctaaatgtt cttttttttt 600
agcgaggggcc agagcccaagg gcaggggaaa ccagaggcccc caggacggcc ggccaggcc 660
cctgtccctt ctgcaggccccc gggcgggggcc ctggccggca tggccgggtt tccactcaag 720
acggctgtcg aggaggccccc cagagggca ttggacacggc aacggggccag tgccaccccc 780

```

-continued

gcggcccccc	gagcggaaagg	aaaggagct	cccaggcca	gaactgcct	cctggcagg	840
agcgaagcag	ggggatgtc	cggggccctc	ctgatccact	tcactctcg	gagcccttgc	900
agcgaagcgg	aggcggagac	agggtgtc	aggggctcct	ctcgccaggc	cgcaggcccc	960
gcggggcaac	aggacactgg	ccccctggag	gcggggcgcc	ggccctcggg	ctcgatgggg	1020
agaggccggg	gcccggggcg	gccccatggac	gctggctcg	gagacagagc	ccggccgccc	1080
cggaaactgg	accgcgtccg	cctggggcgcc	gcggggagcc	tgtgggacgc	gttggacggg	1140
gcgcgcggcc	tggacgcccc	cgcgcgcgg	ctccccacag	gaccccaact	cgcccaggag	1200
cccgactcc	cggcgtgc	cgccccccgc	gcctgcac	ctgggtttca	gacggaaaggc	1260
tctggggcca	agggtgtcg	gagcaggag	gcctcggggg	tcctgcccc	cgaggaggc	1320
tggcccttggg	ttagcgagg	gttccccggg	acccggagct	ttggcccaagc	cccaagactcc	1380
acgcgcggcc	ggcttagag	tccgcctcaa	ggtgccca	tctcgatccca	aggggccgggt	1440
gcacagggg	cttacatgc	cgggagggc	ggggctgaca	gtctccgaga	taacagtcc	1500
gcgcgtacc	tggggccac	cgggcccccg	gagcaagcaa	agccggctgc	agccggccac	1560
agccgcgcgc	cctccggag	ccgtgagcc	cgcccggcc	ccgcctcc	gccecgca	1620
cccgccccgg	ggttccacc	tgaagccctg	actctccccc	tccttcaga	cttttgccc	1680
cttgagggtt	ccaggatcc	tccctgggc	gaaaatcta	gacggccccc	agccccaa	1740
tcagcctcg	ccaaagtctt	aacttcag	ccagcctcg	tcctagcccc	ageccctggct	1800
tcatcccca	gttcgcacc	aaccttcag	accacccat	ccttcatccc	cacccatgc	1860
ccagcccccc	cttccaaatcc	agctccat	tcacccccc	cccccagccc	aagtccagct	1920
gcagcccoaa	ctcccaaaaa	agccccatgc	ccagttccaa	ccctcacacc	cccatcccc	1980
gccttaacc	cagtcccaac	cccagccct	agccccatgc	caactccagc	cctaacc	2040
gcgcataccc	cagccctaa	cccagtc	accccaagccc	taagcccc	tccaaactcc	2100
gcgcgcaccc	cagccgcac	ccctgc	gccccaccc	cagccccaa	cccaacc	2160
gcgcgcaccc	ctggccccc	tgcgggtc	aaggctcagg	agagtgtgc	tctcccccagg	2220
cgttaccagg	aggggcagg	ctcagcc	tggggaaacc	ttattgc	gttcttaga	2280
agccacccct	tcccaagg	agacaggcc	caaggatgt	tcccggggc	gttccccggg	2340
agcccccgtt	gtcccgac	ttccacac	tctgaggaca	gacacggccc	cttcttc	2400
gtggggcag	tcatagg	aggatcagg	ggctgttgc	aggctgggg	tcagccac	2460
ccaaagaact	ccgagacc	cgatcgcc	agccccagg	ctcccccagg	cctaagagga	2520
gagggaacca	gggagaaaag	tctagacc	ctggccca	ccgcgtatgc	cagggcccc	2580
gcacacccc	ccgcgc	ggccgtcc	ccggccgc	cctccctgc	gagggcgtca	2640
cagccgttac	cccaatgc	ggaaacg	agggtgc	tgcggccag	ctcgagcag	2700
gaggctcagg	ccggccctc	ggggaccc	caagggagg	cgccggggg	ggggggcagc	2760
cctggccccc	gcagcggg	gctcacgg	aagcaggagg	aggccccgg	gctcatgt	2820
tttctcaga	ggccccgggg	ttgggggtt	gtgggggg	cccgaaagg	cagctcccg	2880
gccttggac	ccgcaccc	ggcagcc	ggcagcc	tggacccgg	cagttcgt	2940
gacgtgttgg	cttttgc	gcagaccc	gagccggcc	tggcggagg	gacccatcg	3000
ctgtatgc	acaacatgc	gctgatgc	ggagaccc	gcctctacc	ccggctg	3060
gcgcgcaccc	gogggcgt	cctccatgc	ccggccggcc	ggggccggc	gtgtgtggc	3120
gtctctcgatc	tgc	ccatgc	ctaccatgg	ggccctc	ggccctc	3180
ggcgaggagc	ctccctggc	ggccctgt	tcctgc	tacccatgc	cctgc	3240
ttcaaccc	gggagaac	ctggggcc	ctgacccagg	tgcggagg	ggcccgc	3300
cgccgcgtcg	gtctctgc	catgcacaa	tacttgtt	ttgcgggggg	catcggttgc	3360
tccgggtc	aggccgtcg	cttcaacag	tgcttgc	acaacccct	gaccaacatc	3420
tggaggcagg	tccggccat	cgacagg	cgacccagg	tcaatgt	ggccctgg	3480
ggggtgtct	atgcacatgg	tggc	atgc	tggatgt	cgacccgc	3540
acagacgc	ggacccac	cgcgc	ccgcagg	ccttccctgt	ggcccacag	3600
getgtggc	ggccgtgg	catctacgt	acccgggg	accccttca	ccgcgtc	3660
aggtacaggc	ccgtgaa	tgcttgg	gatgtcc	acatgtc	ccacccgc	3720
tccagcgaca	tcgtgg	acttgc	gggggtt	ctgtacc	tcgacc	3780
ggcgcccg	tgatgc	caacacat	acccgtc	ggagcagg	tgccctc	3840
ccctgc	ccccccccc	atgc	ccaccc	tggacactt	ttactgc	3900
aacccca	tcaatgc	cttacgg	tctgggggg	ctgcccagg	ccagccaa	3960
gagctgc	cctcc	ggggaccc	ggggctc	gtccatt	cctgact	4020
ccccctgg	acccgc	gacccat	tgatgt	gca	gacccat	4080
cgtgtctc	cagg	ggaccc	ctctgg	ggccctg	ggccgggg	4140
gtctggatcg	gaacttctg	cttgg	tggacactt	tcccttc	ctttaaa	4200
tgtcgattat	tttgc	agactcc	acccctt	tgc	ccacacca	4260
gactgttcc	tgc	acttca	ccgttac	ttac	gacacccc	4320
cctgtctgt	ggacttcc	tcc	ccagg	act	gtgttcc	4380
acttgc	ctgg	atgc	ccagg	atgt	ccacccat	4440
cctatgtct	gtc	acccat	ggagg	ggat	tttccat	4500
ggactttctg	ctgtctaa	ggccccc	ggtacca	ccgt	atgttgc	4560
ttggcactgt	ggatcc	ggccct	ggcc	actgt	tttgc	4620
agccctgt	tctgc	atgc	cccg	cc	ccaccc	4680
gggctca	ccctctg	ccttgg	cttccc	ccct	ccaccc	4740
cttctcgcca	accccgatc	tctaaat	ggct	ccct	ccat	4800
tttagtctgt	agacttgg	gagacat	tttcc	tttgc	tttgc	4860
gacagtgg	gtctgtc	caagag	gtgg	gac	tttgc	4920
tcgcacgg	tccaggat	acag	ccat	gtat	ccat	4980
tga	aaaa	ccat	ccat	ccat	ccat	5040
gatgtccct	tctggccct	tttttgc	tttgc	tttgc	tttgc	5100
aatacattt	aattcacc	aagca				5125

-continued

-continued

---

cctgtctgtg	ggactcccta	ttcccttagag	ccagggactg	atgcgttccc	acagacaagg	4380
acttggctcg	ctggagctc	gctgagccga	gagaggaggg	ggtagaaaaac	attcacactt	4440
cctatgtct	gtcagcggaa	caggagca	aaaagtcccc	aggcaacggc	ctcgctctg	4500
ggactttctg	cctgtctaa	ggccctccca	ggtagccaacc	ccgtagctat	ctgggtctgt	4560
ttggcactgt	ggatttcaa	gggccttagaa	cccttgcctc	tgaaacttgt	ccgctggtgc	4620
agccctgtcg	tctgcagctc	ctgeccatac	ccccagccca	caccaggcca	ggcccaactc	4680
gggttcacca	ccctctgcag	ccttggtggg	ctctcccccage	ccctccggaa	ggccacccca	4740
cttctcgoca	accccccgtc	tctaatgag	gcttgagctgt	ccctcttagtt	ctgccccctt	4800
tttagtctgt	agacttggac	gagacatgg	acttccctt	tccttgcctt	ataaaatgt	4860
gacagtggac	gtctgtcacc	caagaggtt	gtgggagaca	agatcacagc	tatgacacc	4920
tcgcacgggt	tccaggatgc	acagcaaat	ccatgtatgc	tttttccccc	ttacgcact	4980
tgaacccat	gttagaaaaag	tgaataacatc	tgactgtgt	ccactccaaac	ctccagctg	5040
gatgtccctg	tctggccct	ttttctgttt	tttattctat	gttcagcacc	actggcacca	5100
aatacatttt	aattcaccga	aagca				5125

SEQ ID NO: 3            moltype = RNA    length = 5125

FEATURE                          Location/Qualifiers  
source                          1..5125  
mol\_type = mRNA  
organism = Homo sapiens

SEQUENCE: 3

agaaggcagg	tggctgtat	gacagcacag	agtcagggaa	cgctgcctga	ggaccctggg	60
gcctacggg	aggagaagag	ggcaggaggt	ggtaggggtt	ttgcagagac	cctgggtctcc	120
tatccctgcca	taagcctcgc	tgttctctga	tatctgcagc	caggccctac	tgacacccccc	180
aggcctgagt	gcaagcagag	acccaccaat	teccaggcccc	tggaggactg	gtccacccctt	240
actggggcagc	ccttggggca	ggcgttggcc	ggtgtccctag	cccaggccctc	tgtgtctgc	300
atgcactgcc	agcctggccat	caggctcta	ttgcagccct	gaaccatgt	ccaggggccac	360
tttggagccag	atggcttccc	ctggggctgg	gactggggca	gtgacaatga	ctgggatagt	420
gtctgtgtgg	ccctctctggc	gctggctgt	gtggctgc	cagcgtggc	cttacactgg	480
ttttggctccg	ggcacatca	agaggccggca	gaaccgggtt	ccacagccct	cggggctcaa	540
cctcatcagg	caggaggagc	tgatggcc	ctgcaacccat	agtctaagg	cagtgtatggc	600
aggcgaggggc	aggcgccagg	gcaggggaaa	ccagagcccc	caggacgggg	ccagcagagc	660
cctgtccctg	ctgcacgcgc	gggggggggg	ctggccgc	tggcccggt	tccactcaag	720
acggcgtgtcg	aggaggcccc	cagagaggca	tttaggacagc	aacggggcag	tgccacccccc	780
ggggggccccc	gagcggggagg	aaaggaggct	cccaaggccgg	gcaactgcct	cctggggcagg	840
agcgaagcag	ggggggatgtc	cgccccccctc	ctgatccact	tcactctcg	gagccctggc	900
agcgaagggg	aggcgagagc	agggtgtgtc	aggcggtct	ctgcgcaggc	cgcaggcccc	960
gccccggcaac	aggacactgg	ccccctggcg	gccccctgggg	ctcgatgggg	1020	
agaggcgccgg	gccccggccgg	gccccgggg	gccccgggg	gccccgggg	1080	
cgggaaactgg	ccccggccgg	ccccgggggg	ccccgggggg	ccccgggggg	1140	
gccccggccccc	ttggacggccca	ccccgggggg	ccccgggggg	ccccgggggg	1200	
ccccgactcc	ccccggccgg	ccccgggggg	ccccgggggg	ccccgggggg	1260	
tctggggccca	agggtggctg	gaggcgagg	gccccgggg	tccctggccccc	cgaggaggcc	1320
tggccctggg	tccagcgagg	ggtccccggg	acccggaggt	ttggccca	ccccagactcc	1380
acggcgccccc	ggctagagag	tccgcctcaa	ggtgcggcc	tctcgccca	aggggccgggt	1440
gccccgggg	cttacatgc	cgccggggcc	ggggctgaca	gtccccgaga	taacagtccct	1500
gccccggccccc	ttggacggccca	ccccgggggg	ccccgggggg	ccccgggggg	1560	
ccccgactcc	ccccggccgg	ccccgggggg	ccccgggggg	ccccgggggg	1620	
tccggggccca	gggtccccc	tgggggggg	ccccgggggg	ccccgggggg	1680	
ctggagggtt	ccccaggatcc	ttccgtggg	gaaaatctca	gagccggcc	agccccaaat	1740
tcagccttcag	cccaactgtt	aacttcgt	ccagcctcag	tcctagccccc	agccctgtct	1800
tcateccccca	gtcagcggacc	accatcagec	accatccca	cctccatcccc	cacccatcgcc	1860
ccagcccccc	ctccaaacctc	agtcacta	tcaacccca	ccccagcccc	aatcccaat	1920
ccagccgcaat	ctccagcccc	agccccagtc	ccagtc	cccccccccc	ccccccggat	1980
ccctactccccc	ccatccca	ccccggccca	agccccatct	caactccagc	cttaacccca	2040
cccgccatccc	ccatccca	ccccggccca	ccccggccca	taagccggcc	tccaaactcca	2100
cccccaaccc	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2160
cccgccatccc	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2220
ccgttccatgg	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2280
ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2340
ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2400
gtggggacag	tcatagggac	aggatcaggg	ggctgggtt	aggctgggg	tcagccacag	2460
ccaaagaatct	ccgagacccaa	ccgatcgccc	agcccagacc	ctccccca	cctaaagagga	2520
ggggggacca	ggggaaaaaa	tctagaccc	ctggccca	ccgcgtatgg	ccggggcccc	2580
gcacagcccc	cccgccggcc	ggccgttccc	cccgccggcc	cctctcttc	ggggccgtca	2640
ccatccca	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2700
gagggtcgcc	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2760
cctggccccc	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2820
ttttctgtgg	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2880
ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	2940
gacgtgtgg	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	3000
ctatgtgtcg	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	3060
ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	3120
gtccctcgatc	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	3180
ggcgaggaggc	ccatccca	ccccggccca	ccccggccca	ccccggccca	ccccggccca	3240
ttcaacccccc	ggggaaacac	ctggccggcc	ctgacc	ccccggccca	ccccggccca	3300

-continued

---

cggggctcg	gtctctgcac	catgcacaad	tacctgtttc	tgggggggg	catccgtggc	3360
tccgggtcoca	aggccgtctg	cttcaacag	gtcttcgtct	acaacccct	gaccaacatc	3420
tggagccagg	tteggccat	gcagcaggcc	cgagcccgac	tcaagctgt	ggccctggac	3480
gggctgtct	atgccccatgg	tggcataatgc	ctgtacacga	tggagtgtca	cgaccggcga	3540
acagacgcct	ggaccccacg	cgcgcactc	ccgcaggca	ccttccctgt	ggccacacag	3600
getgtggcct	gocgtgggaa	catctacgtc	acccgggggt	acctcttcta	ccgctgtctc	3660
aggtacacgc	cctgtgaagga	tgcctgggac	gagtggccat	acatgtccag	ccacccggct	3720
tccagcaca	tcgtggact	ggggggctc	ctgtacccgt	tgcacccgt	ggggggctg	3780
ggccgcgcg	tgtatgcgcta	caacacagt	acccggctct	ggagcaggcc	tgcctccctg	3840
cccccgtccg	ccccccccc	actgcactgc	accacccct	gcaacacat	ttactgcctc	3900
aaccccaagg	tcaactggcac	cttacacgg	tctgggggg	ctgcacccat	ccaggccaa	3960
gagctgcagg	ccttccctt	gggggacacc	gggggtctca	gtccattcat	cctgactctg	4020
ccccctgagg	accggctgca	gacctactc	tgagtggca	gcagagaacc	aaagctgctt	4080
cgtgtcttc	caggagacac	ctctgggat	ggggctgtac	ggccggggct	cagggaagg	4140
gctgggatcg	gaacttctgt	ctcttgg	tggacaact	tcccttctg	ctttaaagg	4200
tgtcgattat	tttgaagccc	agacttctc	agctcttctc	tgcctccac	tccacaccca	4260
gactgtttcc	tgactcaatt	ccgtactac	ttacagaccc	tctcagctg	ctgacacccc	4320
cctgtctgt	ggacttctca	tttccttag	ccaggact	atgcgttcc	acagacaagg	4380
acttggctcg	ctggagctct	gctgagccg	gagaggagg	ggtagaaaac	attcacactt	4440
cctatgtct	gtcagcagg	caggagacaa	aaacgtcccc	aggcaacggc	ctgcctctg	4500
ggactttctg	ctgtctcaa	ggccctccca	ggtaccaacc	ccgtagctat	ctgggtctgt	4560
ttggacttgt	ggatttctaa	ggggcttaga	cccttgcctc	tgaaactgtt	ccgctgggtc	4620
agccctgtcg	tctgcagctc	ctgccccat	ccccagccca	caccaggcca	ggcccaactcc	4680
ggggctcacca	ccctctgcag	ccttgggg	ccctccaggaa	ggccacccca	4740	
cttctcgcca	accccccate	tctaaatg	gcctgagcgt	caccctagtt	ctggcccttt	4800
ttagctgtgt	agacttggac	gagacattt	acttccctt	tcccttgc	ataaaatgt	4860
gacagtgcac	gtctgtcacc	caagagat	gtggagac	agatcacagc	tatgagcacc	4920
tgcacccgt	tccaggatgc	acagccaa	ccatgatgc	ttttctccc	ttacgcactt	4980
tgaacccat	gttagaaaaa	tgaatacato	tgactgtgt	ccactccaa	ctccagcctg	5040
gatgtccctg	tctggccct	ttttctgtt	tttattctat	gttcagcacc	actggcacca	5100
aatacatttt	aattcaccga	aagca				5125

---

SEQ ID NO: 4	moltype = RNA	length = 2990				
FEATURE	Location/Qualifiers					
source	1..2990					
	mol_type = mRNA					
	organism = Homo sapiens					
SEQUENCE: 4						
ccacccatc	cccaacccca	accccagccg	catccccctc	cccagctgac	gggtcaaaac	60
ctcaggag	tgtggcttc	cccaggcg	accaggagg	gcaggctca	gccagctgg	120
gaaacccat	tccatgttt	cttagaagcc	accccttccc	caggcaagac	aggcccaag	180
ggagtgtccc	gagggcggtt	ccgggagcc	ccgtgggtc	cagcacttcc	acacactctg	240
aggacagacca	ccggccctc	tcttcgtgg	ggacagtcat	agggacaggt	acagggggcc	300
tggttgaggg	tggaggatcg	ccacagccaa	gaagctccga	gaccaacgg	tcgcccagcc	360
caagccctcc	ccaggccct	agaggagg	gaaccaggaa	aaaagtcta	gaccggctgc	420
cccaageccg	gatgccagg	ggcccccgc	agccccccgc	gcagaggccg	cctggcccg	480
ccggcccttc	ctctcgagg	cgctcacac	cggttacccaa	gtacggaaa	cgcagcagg	540
gcgaaatccg	ccccagatcg	gacggagg	tcaggccgg	cgccctgggg	gaccctcaag	600
ggggggccgg	ggggggagggg	ggcgcctc	ccggcccg	egggggctc	acggaaaac	660
aggaggaggc	ccggaaatctc	atgggtttc	tgcaagagg	cgggggttt	gggggtgtt	720
aggggggcccg	gaagggccacg	tccggggcc	tggagcccg	cacggggcga	gccttgcggc	780
ggcgccgtgg	ctggggcgt	tgcctgg	tgctggcctt	tgcccgac	cacggagac	840
ccggccgtgc	ccaggagacc	tacgcgtg	tgacgcacaa	ctgtgcga	gtctgggg	900
accctgtcc	ctaccgcgg	ctgagcgcgg	ccgacccgca	gchgatcc	agcctgcgga	960
ccggccgggg	ccggccgtgg	ctggggctcc	tcgtactgc	cagccttac	cagggggcc	1020
getcagggt	ccccaggcc	cetcgtgg	aggacgcctc	tgccggggcc	cctgtgttcc	1080
tgcctctacc	tgcacccgt	tatgttca	acccctggaa	gaacaccc	cgcccccctg	1140
cccaagggtcc	ccaggaggcc	ccgttccgg	gtcgccgt	ctgcacccat	cacaactacc	1200
tgtttctgtc	ggggggcata	ctgtggccgt	gtggccaa	cgctctgtcc	aacgaggat	1260
tctgtctacc	ccctctgtac	aatcttgcg	ggcaggatcg	ccggccgg	1320	
cccaaggctca	gtgttgcgc	tgcacccgt	tgcgtatgc	catcggtgc	gaatgcctgt	1380
acagcatgg	gtgtacacgc	ccgcgaacag	acgcctggac	ccacacgc	ccactcccg	1440
caggcaccc	ccctgtggcc	cacggagg	tggctccgg	tggggacatc	tacgttacc	1500
gggggtcacct	cttctaccgc	ctgtggccgt	acagccccc	gaaggatgt	tggggacat	1560
ccccatcac	cccgccac	ccggccgtca	gcgcacatcg	ggcactgggg	ggcttccctg	1620
acccgttccg	ctgtgtccgg	ggcgtgggg	ccggccgtgt	gcgcacaa	acatgtcc	1680
gctccctgg	caagggtggcc	tccctggcc	tgcctggccc	cgccccact	cactgcacca	1740
ccctggccaa	caccatttac	tgcctcaacc	cccaagg	tgcacccatc	acggctctg	1800
ggggggactgc	ccagttccag	gcacaggag	tgcaaggccct	cccttgggg	agcacccgggg	1860
tcctcgtatcc	attcatctcg	actctgc	ctgaggaccc	tcaactctg	tgtttctgt	1920
tggcaggcc	agaacccaa	ctgttcgt	gtctccagg	gagaccc	tgggtatggc	1980
ctgagaggcc	gggggtcagg	gaaggggctg	ggatcgaa	ttctctgt	tgtttctgt	2040
caacttcc	cttctgttt	aaagggttgc	gattat	tttgcac	tccctcagcc	2100
tctttctgtc	cctacttcca	ccccagact	gtttccctgac	tcaatccgt	acctacttac	2160
agacccttc	agcttgcgt	cacccctgt	tctgtggac	tcccttattcc	ctagagcc	2220
ggactgtatgc	gtctccacag	acaaggactt	ggctcgctg	agctctgt	agccgagaga	2280

-continued

---

ggagggggta	aaaaacattc	acacttcata	tgctctgtca	gcaggacagg	gagcaaaaac	2340
gtccccccgc	aacgcctcg	cctctggaa	tttctgcctt	tcctaaggcc	tcccaggta	2400
ccaaccccggt	agctatctgg	gtctgtttgg	cactgtggat	tctcaaggcc	ctagaaccct	2460
tgcctctgaa	actggccgc	tggtcagcc	ctgtgttttgc	cagctctgc	ccataccccc	2520
agcccacacc	aggccaggcc	cactccgggc	tcaccacccct	ctgcagccct	gtggggctct	2580
cccaagccccc	ccagaaggccc	accccaacttc	tcgccaaccc	ccgatcttca	aatggggct	2640
gagcgttacc	ctagttctgc	cccttttttag	ctgtgttagac	ttggacgaga	catttgactt	2700
cccttttcc	ttgtctataa	aatgtggaca	gtgacgtctt	gtcacccaaag	agagtgtgg	2760
gagacaagat	cacagctatg	agcacctcg	acggtgttca	ggatgeacag	cacaatccat	2820
atgcgtttt	ctcccttac	gcacccgtaa	acccatgtca	aaaaagtggaa	tacatctgac	2880
tgtgtccac	tccaaacctc	agcttggat	tcctgttcc	ggccctttt	ctgtttttta	2940
tttatatgttc	agcaccactg	gcaccaaata	catttaat	caccgaaagc		2990

---

SEQ ID NO: 5                    moltype = RNA   length = 4837  
 FEATURE                        Location/Qualifiers  
 source                        1..4837  
 mol\_type = mRNA  
 organism = Homo sapiens

SEQUENCE: 5  
 actgccaccc tgccatcagg cctctattgc agccctgaac catgatccag ggcacccctgg 60  
 agccagatgg tccccctctgg ggctggact gggacagtga caatgaactgg gatagtgtcg 120  
 tgtgtggccct ctggcgctg gctgtgggtg ctggccacagc gctggccctta cactgtgttg 180  
 gctccggggca cgatcaagag gccgcacggaa cgggtgtccac aaccctctggg gctcaacactc 240  
 atcaggcagg aggagcttag cttggccctgc aaccgcggatc taaggctagt gatggcagcg 300  
 agggggcagag cccaggccag gggaaaccag agccccccagg acgcggccagc cagagccctg 360  
 tcocctgtgc agcgcggggc gggggccctgg cggccatggc cccgcttcca ctcaagacgg 420  
 ctgtcgagga ggcccccggaga gagggatattt gacagcaacgc gggcgtgtcc acccccgccgg 480  
 ccccccggagc ggaaggaaag gagcttccca ggcaggccac tgccctctgg ggcaggagcg 540  
 aagcagggggg gatgtccgc cccctcttgc tccacttcac tccctggagc cctggcagcg 600  
 aagcggggggc gggacggatg ggtgtcaggg cgttctctcg ccaggccgc gggcccccggg 660  
 gcaacacgga cactggcccc tggcaggccg ggcggggccctc ctcgggctcg atggggagag 720  
 gccggggggcc gggccggccg atggacgtcg gtcggggaga cagagccgc cggcccccgg 780  
 aactggaccc gtcggccctg ggcgcgcggg ggacgtgtg ggacgcgggtg gacggggccg 840  
 ccccccgttga cgcccccgcg cggggccctc ccacaggacc cccactcgcc caggagcccg 900  
 cactccggccg gtcggccctg ccccccggcc tgcacgttcc gtcctcaacgc gaaaggctcg 960  
 gggcccaagggg tggctggagc aaggaggccct cgggggttcc tggccctggga ggaggctggc 1020  
 cctgggtttag caggaggatc cccggccaccc ggacgttggc cccagcccca gactccacgc 1080  
 gcccctggct agagagtccg cctcaaggtc gcccactctc gtcccaagggg cccgggtggca 1140  
 caggggccctt ctagatccgc gggggccggg ctgcacgttc cccgatataac agtcttgcgg 1200  
 ctgcacccgtt gtcggccctg ccccccggcc tgcacgttcc gtcctcaacgc gaaaggctcg 1260  
 gggccgcctc cccggccctg cggcgttcc ccccccggcc gcaactcccg 1320  
 gcccgggggtt cccacctgaa gccctgactc tccctcttc ttcaagactt ttggccctgg 1380  
 aggttaccca gtagcccttc gtggccggaa atctcaacgc ggcgcggcc ccaagttcag 1440  
 cctcagccca agtcttaact tcagtcacg ctcacgttc atggccggcc ctgggttcat 1500  
 ccccccggatc agcaccaccc tcagccacca cttcaacccct atccccccac tcagccccc 1560  
 ccccaactctc aacccatcgat ccaacttcaa ccccaactcc agccccaactg ccactgcag 1620  
 ccgcaactcc acggccggcc cccatcccgat tcccaacccct cacaccccca tccccaggcc 1680  
 taaccccaatg cccaaaccca gccctaagcc cagtcaccaatccacccatccca accccccgg 1740  
 catccccccgc cttaaacccca gtcccaaccc cccagcttca actccaggccc 1800  
 caaccccaatc cgcataccctt gccccagccc ccacccatccca accccccgg 1860  
 catccccctgc cccatcgatc gggtaaaagg ctcaggagag tgcgttccctc cccaggccgt 1920  
 accaggagggg cggaggctca gccagcttca gaaacccatgg tgcctatgtt cttagaaagcc 1980  
 acccccttccc caggcaagac agggcccaag qgagtgttcc gaggcgggtt cccggggagcc 2040  
 ccgtgggtcc cagcaacttc acacactctg aggacagaca cggccctct ttttcaatgg 2100  
 ggapcggatcat agggcggatg acggggggcc tgggttggcc tggagggttcc ccacagccaa 2160  
 gaaggtccga gaccaacggg tggccggccatc cccaggccca accccccgg 2220  
 gaaccagggg gaaaggatcta gacccggctg cccaaaggcc gatgcggccgg gggcccccac 2280  
 ageccccccgc gcagaggccg cttggccccc cggccctctc ctctcgagg cgttcacacgc 2340  
 cggtaacccca gtagccggaa cgcacggatc gggaaatccg cccgagctcg gagcaggagg 2400  
 tcaggccggc cggccctgggg gacccatcgat tggggccggcc ggggggggggg ggcggccctg 2460  
 cccggccgcag cggggccgtc acggccaaagg aggaggccgg cccggacgtc atgggtttc 2520  
 tgcacggggc cgggggttgg ggggtgttgg agggggccccc gaagcccaagc tcccccggcc 2580  
 tggagccggc cacggccggca gccctggccg ggcggcttggaa cttttttttt tttttttttt 2640  
 tgcgtggccct tggcccgccatc cccggccatggc cccggccatggc tttttttttt tttttttttt 2700  
 tgagcgcacaa cttgttgcgtc gtgtgtttttt tttttttttt tttttttttt tttttttttt 2760  
 ccgcaccgcga ggcgcacccatc cccggccatggc cccggccatggc tttttttttt tttttttttt 2820  
 tgcgtacttc cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2880  
 aggagccctcc tgcggccggcc cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2940  
 acccccccggg gaaacccatcgat cccggccatggc cccggccatggc tttttttttt tttttttttt 3000  
 gctgtgggttcc cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 3060  
 gtgccaaggcc cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 3120  
 gccagggttcg gcccatcgat cccggccatggc cccggccatggc tttttttttt tttttttttt 3180  
 tgcgtctatgc cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 3240  
 acggccctggac cccacccgcgc cccactccccc cccggccatggc cccggccatggc tttttttttt 3300  
 tggccctggcc tggggccatcgat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 3360  
 acaggccccgtt gaaaggatgtc tggggccatcgat tttttttttt tttttttttt tttttttttt tttttttttt 3420

-continued

ggcacatcg	ggcactgggg	ggcttccgt	accgcttcga	cctgcgtggg	ggcggtggcg	3480
ccggccgtat	gcgcgtacaac	acagtggcc	gctccctggag	cagggtctgcc	tccctgcccc	3540
tgcggccccc	cgcccccaact	actgcacca	ccctggggca	caccatttc	tgcgtcaacc	3600
cccaaggatca	tgccaccctt	acggttttcc	ggggggactgc	ccaagtccag	gcacaaggagc	3660
tgcagccctt	ccccctgggg	agcacccgggg	tcctcaagtcc	attcatccgt	actctggccc	3720
ctgaggaccg	gctgcagacc	tcactctgg	tggcaggcag	agaaccaaa	ctgttctcgct	3780
gtcttcggcc	gagacccctt	ttggatgggc	ctgagggcc	gggggtcagg	gaaggggggc	3840
ggatcgaaac	ttcttcgtct	tggttctgg	caacttcc	cttcgttta	aaagggtgtc	3900
gattatttt	aageccagac	tccctcaogcc	tctttctggc	cctcaactcca	caccaggact	3960
gtttctgtac	tcaattccgt	acctacttac	agaccccttc	agttgtgtca	cccccccttg	4020
tctgttgggg	tccttatttc	tcagacggcc	ggactgtgc	gttcacacag	acaaggactt	4080
ggctctgtgg	agctctgtct	agccggagaga	ggggggggta	aaaaacattc	acacttccta	4140
tgctctgtca	gcaggacagg	gagcaaaaac	gtccccaggc	aacgccttc	cctctgggac	4200
tttctgtctt	tcctaaaggcc	tccccaggt	ccaaacccctg	agttatctgg	gtctgtttgg	4260
cacttgtgtt	tctcaagggc	tcagaacgg	tgcctctgaa	actggccgc	tgggtcagcc	4320
ctgtctgtct	cagtcgtctc	ccataccccc	agccccaccc	ggccggcc	cactccggcc	4380
tcaccaccc	ctgcagccct	gtggggctct	ccagccct	ccagaagccc	accccaacttc	4440
tcggcccaaa	ccggatctca	aataggggct	gagggctcacc	ctagtctcgc	cccttttttg	4500
ctgtgttagac	ttggacgaga	cattttgg	ccctttctcc	ttgtctataa	aatgtggaca	4560
tgggacgtct	gtcaccacca	agaggtttgg	gagacaagat	cacagctatg	agcaccttc	4620
acggtgttcca	ggatgcacag	cacaatccat	gatgggttt	ctcccttac	gcactttgaa	4680
accctatgtca	gaaaatgtaa	tacatctgc	tgtgtccac	tccaaaccc	agactggatg	4740
tccctgtctg	ggccctttt	ctgtttttt	ttctatgttc	agcacca	gcaccaaata	4800
catttttaatt	caccggaaagc	aaaaaaaaaa	aaaaaaaaaa			4837

```
SEQ ID NO: 6          moltype = RNA  length = 2099
FEATURE           Location/Qualifiers
source            1..2099
                  mol_type = mRNA
                  organism = Homo sapiens
```

```
SEQ ID NO: 7          moltype = RNA    length = 5125
FEATURE              Location/Qualifiers
source               1..5125
                     mol_type = mRNA
                     organism = Homo sapiens
```

SEQUENCE: 7  
organism = HOMO sapiens  
agaagcagg tggctgtat gacagcacag agtcctggaa cgctgcctga ggaccctggg 60  
ataccatggg aggagaaggg gcggcaggat gttgggggtc ttgcagagac cctgggttcc 120  
tatccctcca taaggcttcg ttatccctta tatcttcacgg cttggccctac tgaccccccc 180

-continued

---

ggccctgagt	gcaaggagag	accccacat	tcccaggccc	tggaggactg	gtccacetta	240
actggcagc	ccttggggc	ggccgtggcc	gttgcctca	cccgaggc	tgtgtctgc	300
atgcactgcc	agectggccat	caggcctcta	ttgcagccct	gaaccatgtat	ccaggcacc	360
ttggagccag	atggtccccct	ctggggctgg	gactgggaca	tgacaaatga	ctggatagt	420
gctgtgtgg	ccctcttgcc	gtggctgtg	gtggctccca	cagcgtggc	cttacactgg	480
tttggctccg	ggcacatca	agaggccggca	gaaccgggt	ccacagocct	cggggctcaa	540
cctcatcagg	caggaggagc	tgagctggcc	ctgcaacccga	agtctaaatgt	cagtgtatggc	600
agcgaggggc	agagcccgag	gcaggggaaa	ccagagcccc	caggacgggg	ccagctagagc	660
cctgtccctg	ctgcagcgcc	gggggggggc	ctggccggca	tggccggct	tccactcaag	720
acggctgtcg	aggggccccc	cagaggaggca	tttaggacagc	aacggggccag	tgccacccccc	780
gcccggccca	gagcggaaagg	aaaggagcc	ccaaaggccag	gcaactgcct	cctggccagg	840
agcgaagcag	ggggatgtc	cgcggggcc	ctgatccact	tcaactctcg	gagccctggc	900
agcgaagcgg	aggcgagac	agggtgtc	agggcgct	ctcgccagc	cgcaggcccc	960
gcggggcaac	aggacactgg	ccccctggag	gccccggcgg	ggccctcg	ctcgatgggg	1020
agaggccggg	gccccggcgc	ggggatggcc	gtggctcg	gagacagagc	ccggccccc	1080
cggaaactgg	acccgtctcg	cctggggcc	qccgggaaqc	tgtggacgc	gttggacggg	1140
gcccggccca	ttggacggcca	cggcgccgg	ctccccacag	gaccccaact	cgcggcaggag	1200
cccgactcc	cgggcgtggc	cgctccccc	gcctgcagc	ctgggttcca	gacggaaaggc	1260
tctggggcca	agggtgtcg	gagcggggag	gctctggggg	tcctgcccc	cgggaggagc	1320
tggcccttgg	ttagcaggga	gttccgggg	acccggggat	tggcccaagc	cccagactcc	1380
aegcgccccc	ggctagagag	tccgcctaa	gttgcggcc	tctgttcca	aggggccgggt	1440
gccccagggg	cctacgtat	cgggcgaggc	ggggctgaca	gttcccgaga	taacagtcc	1500
gcggctgacc	ttggggccac	ccggccggc	gagaacgca	agccgggtgc	agccggccac	1560
agccggcgc	cctccggag	cctggagcc	ccggccgcgt	ccggccccc	gcccccaagc	1620
ccggccccc	gttccccacc	tgaagccctg	actctccct	ctcccttca	cttttggcc	1680
ctggagggtt	cccaggatcc	ttccgtggg	gaaaatctca	gagcggccgc	agccccaaat	1740
tcagcctcag	ccaaatgttt	aactctatgt	ccagcctca	tcctagoccc	agccctgtgt	1800
tcatccccca	gttcagcacc	acccatccaa	ccttcattccc	cacccatggcc	1860	
ccagccccag	ctcccaacc	agctccaa	tcaacccca	ccccagcccc	aagtccagct	1920
gcagccgca	ctccagcccc	agccccagtc	ccagttccaa	ccctcaaccc	ccccatcccc	1980
gcctcaacc	cagtcccaac	ccggccctca	agccccatgc	caactccagc	cctaacc	2040
ggccatcc	cggccatcc	ccccatccaa	ccggccggcc	tttttttttt	tttttttttt	2100
gcccccaacc	cagccgcata	ccctggccca	gcccccaact	cagcccaac	ccccatcccc	2160
gcgcgcatecc	ctggccccc	tgacgggtca	aaacccatgg	agagtgtggc	tctcccccagg	2220
cgttaccagg	aggggccagg	ctcaggcc	ttggggaaacc	ttatgtcat	gttctttaga	2280
agccaccctt	tcccccggca	agacggggcc	caagggatgt	tcccgaggcc	gttcccccggg	2340
agccccgtgg	gtcccgagcc	ttccacacac	tctgaggaca	gacacggccc	cttcttctca	2400
gtggggcag	tcatagggac	aggtacagg	ggccctgggt	aggctgggg	tcagccacag	2460
ccaaagact	cggagccaa	cggatcgcc	agccccagcc	ctcccccagg	cctaagagga	2520
gagggaatcca	ggggggaaag	ttagaccc	ctggcccaaa	ccgcgtatggc	caggggcccc	2580
gcacagcccc	ccgcgcagag	ggccctctgg	ccccggggct	cttcctctgc	gaggggctca	2640
cagccgtac	cccaatc	gaaacgcgc	aggtgcgaa	tcgccccgg	ctcgaggccag	2700
gaggctggc	ggccggccctc	ggggggacc	caagggggagg	cgccgggggg	ggggggccgc	2760
cctgcggg	cgccggggc	gttcacggaa	aacccgggg	aggccgggg	gttcatgg	2820
tttctcgaga	ggccgggggg	ttgggggtgt	gttgcgggg	cccgaaaggc	cagtcctccgg	2880
gccttgaggc	ccggccatggc	ggcagccctg	ccggccggcc	tggacccgg	cagttgcgt	2940
gacgtgtgg	ccttgc	gcacacgg	gagccgggg	tggcggagg	gacccatcg	3000
ctgtatggc	aaacactgt	cgatgtgt	ggagacccgt	gtcttaccc	ccggctgtgc	3060
ggggccgacc	gogagccat	cctcagctg	ccggacccgg	ggggccggcc	gttgcggc	3120
gtctctgtac	tgccca	ctaccagg	ggcgcgtc	gttcccccag	ggggccctcg	3180
ggcgccggc	ctcc	ggccctgtgt	tcctgc	tacccgc	cttgcgttg	3240
ttcaaccccc	gggagaacac	ctggggccc	ctgacccgg	tggccgggg	ggcccccgt	3300
cgccggcgt	gtctctgcac	catgcacaa	tacccgttcc	ttgggggggg	cattccgtgc	3360
tccgggtcc	aggccgtctg	cttcaacgg	gttccatgt	acaacccct	gaccaacatc	3420
tggagccagg	ttccggccat	gcacccgg	cgacccgg	tcaactgt	ggccctggac	3480
ggggctgtct	atgcctatgg	tggcgtat	tcgtacag	tggagtgtca	cgacccggc	3540
acagacgcct	ggacccca	ccgcgcact	cccgccagg	cattccctgt	ggccca	3600
gctgtggct	ggcggtgg	catctacgt	accgggg	accccttct	ccgcctgtc	3660
aggtacagg	ccgtaa	tggtgggg	gagtgc	acagtgc	ccacccgcgt	3720
tccagcgaca	tcgtggccat	ggggggctt	ctgtaccc	tcgactgt	ggccggat	3780
ggggctgtct	atgcctatgg	tggcgtat	tcgtacag	tggagtgtca	cgacccggc	3840
ccctgcggcc	cccccccccc	actgcact	accacccctgg	gcaacaccat	ttactgc	3900
aaccccccagg	tcactgcac	tttacccgt	tctgggggg	ctgcccagg	ccaggccaa	3960
gagctgcggc	ccttccctt	ggggggatcc	gttccat	ccctgtact	ccgcctgtc	4020
ccccctgg	acccgtgtca	gacccat	tgatgtgg	ccacccgt	aaaggtgt	4080
cgtgtcttc	caggagacc	ctccctggat	ggcccttgg	ggacccgg	tgcctcc	4140
gctggatcg	gaacttctg	ctcttgc	tggaca	actccctt	ctttaaaatgt	4200
tgtcgttat	tttgc	ggatcc	agcccttcc	tgccc	tccacacca	4260
gactgtttcc	tgactca	ccgtaccc	ttacagat	tctcag	ccgacaccc	4320
cctgtctgt	ggactcc	ttccct	ttcccttag	ccaggact	atgcgttcc	4380
acttggctcg	ctggagct	tgtac	gagagg	ggtagaaa	attccacact	4440
cctatgtct	gtcagc	ccggag	aaacgtcccc	aggcaac	ctcgctct	4500
ggactttctg	cctgtctaa	ggccccc	gttacca	ccgtat	ctgggtctgt	4560
ttggactgt	ggattctca	ggccct	tgaaactgt	ccgt	ttgggt	4620
agccctgtct	tctgc	ccgc	ccccagcc	caccagg	ggcccac	4680
gggtctacca	ccctctgc	cttgc	gggggg	ccctcc	ccccacccca	4740

-continued

cttctcgcca	accccccgate	tctaaatgag	gcctgagcgt	caccctagtt	ctgccccctt	4800
tttagctgtgt	agacttggac	gagaggattt	acttccctt	ctccctgtct	ataaaatgtg	4860
gacagtggac	gtctgtcacc	caagagatgt	gtgggagaca	agatcacagc	tatgacacc	4920
tcgcacgggt	tccaggatgc	acagcacaat	ccatgatgcg	ttttctcccc	ttacgcactt	4980
tgaaacccat	gctagaaaaag	tgaataacatc	tgactgtgt	ccactccaaac	ctccagcctg	5040
atgtgtccctg	tctggccctt	ttttctgttt	tttatttctat	gttcagcacc	actggcacca	5100
aatacatttt	aattcaccga	aagca				5125

SEQ ID NO: 8                    moltype = RNA   length = 2990  
 FEATURE                        Location/Qualifiers  
 source                        1..2990  
 mol\_type = mRNA  
 organism = Homo sapiens

SEQUENCE: 8

ccacctcagc	ccccaaaaaa	accccaagccg	catccccctgc	cccagctgac	gggtcaaaagc	60
ctcaggagag	tgtggctctc	cccaggcgt	accaggaggg	gcagggtctca	qccagctgg	120
gaaaaccttat	tgccatggtt	cttagaagcc	accccttccc	caggcaagac	aggccccaaag	180
ggagtgtccc	gagggccgggt	ccggggagcc	ccgtgggtcc	cagcacttcc	acacactctg	240
aggacagaca	ccggccctct	tcttcagtgg	ggacagtcat	aggggacaggt	acagggggcc	300
tggttgaggc	tggagggtcag	ccacagccaa	gaagactccga	gaccaacgga	tcgccccagcc	360
cagaccctcc	cccaggccta	agaggagagg	gaaccaggga	gaaaagtcta	gaccggctgc	420
cccaaggccg	gtatggccagg	ggcccccggd	agccccccgc	gcagaggccg	cctggccccc	480
cggccctctc	ctctggcagg	cgcttacacago	cggttacccca	gtatggccagg	cgccggccgg	540
gcgaaatcgc	cccgaggctcg	gaggcagggg	tcaggccggc	ccgcctgggg	gaccctcaag	600
gggagggggc	gggggggggg	ggcagccctg	ccggccggcg	cgggggcgctc	acggaaaagc	660
aggaggaggc	ccggaaatgc	atggttttc	tgcaagaggcc	cggggggttgg	gggggtgggtgg	720
aggggggcccg	gaaggccccagg	tccggggcc	tggagccccg	cacggccggca	gccttgcggc	780
ggcggcgttga	cttggggcagt	tgcctggacy	tgcctggctt	tgcctcggcag	cacggagagc	840
ccggcctggc	gcaggagacc	tacgcgctga	tgagcgacaa	cctgctgcga	gtgtggggag	900
acccgtgtct	ctaccggccgg	ctgagcgcgg	ccgaccgcga	gcgcacatctc	agcctgcggg	960
ccggccgggg	ccgggggggt	ctggggctgt	tcgtactgc	cagcctctac	cagggggggcc	1020
gtcaggggct	ccccaggggc	cctctggcgc	aggagccctc	tgcggccggc	cctgtgtccc	1080
tgcctctacc	tgcgcacctg	catgtgtca	accccccggg	gaacacccgg	cgccccctga	1140
cccaagggtgc	cgaggaggcc	ccgcttcggg	gctgcggct	ctgcaccaatg	cacaactacc	1200
tgtttctggc	ggggggggcata	ctgttcggcc	gtggccaaaggc	cgctctgtcc	aacgagggtct	1260
tctgtctacaa	ccctctgtacc	aatatcttga	gccagggttc	gcccattgcg	caggcccgag	1320
cccaaggctca	gtgttgtggc	ctggacgggg	tgcctatgc	catcggtggc	gaatgcctgt	1380
acagcatgga	gtgtctacgc	ccgcgaacag	acgcctggac	cccacgcgcg	ccactccccg	1440
caggcactt	ccctgtggcc	cacggggcc	tggggccgtt	ttggggacatc	taatgcaccc	1500
gggggtcacct	tttcttacccg	tcgttcgggt	acagccccgt	gaaggatgtt	ttggggacgagt	1560
gcocccatag	tgcccaagcc	ccgcgttca	gcgcacatcg	ggcacttggg	ggcttcctgt	1620
accgcttgcg	cttgcgtgg	ggcatgggg	ccgcgcgtat	gcgcataaac	acagtgcaccc	1680
gtccctgggg	cgagggtggc	tccctggccc	tgcggccccc	cgccccactg	cactgcacca	1740
ccctggggca	caccattac	tgcctcaacc	cccaagggtca	tgcacccctc	acgggtctctg	1800
ggggggactcaa	ccatgttccag	gccaaggago	tgcagccctt	cccccttgggg	agcaccgggg	1860
tccttcgtcc	attcatctcg	actctggccc	ctgaggaccg	gtgcagacc	tcactctgag	1920
tggcaggccag	agaacccaaat	ctgttcgtcg	gtcttcagg	gagacccccc	ttggatgggc	1980
ctgagggggcc	gggggttcagg	gaaggggctg	ggatcgaaat	ttctctgtct	tgtttcttgg	2040
caactttccc	tttctgtttt	aaagggtgtc	gattattttt	aagcccgatc	tccctcagcc	2100
tctttctggc	cctcaactca	cacccagact	gttctctgtc	tcatttcgtt	acctactttac	2160
agacccttc	actgttcgtc	caccccccgt	tctgtgggac	tcccttattcc	ctagagccag	2220
ggactgtatgc	gtctccacag	acaaggactt	ggctctgtcg	agccgagaga	2280	
ggaggggggta	gaaacacattc	acacttccca	tgcctctgtca	gcaggacagg	gacccaaaaac	2340
gtccccaggc	aacggccctcg	cctctggac	tttctgtctt	tccttaaggcc	tccccaggta	2400
caaaacccgt	actatcttgc	gtctgtttgg	cactgtgtat	tctcaaggcc	ctagaacccct	2460
tgcctctgtt	actgttcgtc	tgggtggcc	ctgttcgtcg	cagtcgttc	ccataccccc	2520
agccccaccc	aggccggcc	cactccggcc	tcaccaccc	ctgcagcc	gtggggctct	2580
cccaaggccct	ccagaagcccc	accccaactt	tcgccaaccc	ccgatctta	aatggggct	2640
gagcgttccatc	ctatgttcgtc	ccctttttat	ctgtgtatgc	ttggacgaga	cattttact	2700
ccctttcttc	ttgttctttaa	aatgtggaca	gtggacgtct	gtcaccacaa	agagtgttgg	2760
gagacaatag	cacagctatgc	acgcacatcg	acggtgttca	ggatgcacag	cacaatccat	2820
gatgcgtttt	ctcccccattac	gcactttgaa	accatgtctt	gaaaagtggaa	tacatctgac	2880
tgtgtcttcc	tccaaacctcc	agcctggat	tccctgtctg	ggcccttttt	ctgtttttta	2940
ttctatgttc	agcaccacttgc	gcaccaaata	cattttaaat	caccgaaagc		2990

SEQ ID NO: 9                    moltype = RNA   length = 4837  
 FEATURE                        Location/Qualifiers  
 source                        1..4837  
 mol\_type = mRNA  
 organism = Homo sapiens

SEQUENCE: 9

actggccaggc	tgccatcagg	cctctattgc	agccctgaac	catgatccag	ggcaccttgg	60
agccagatgg	tcccctctgg	ggctggact	ggggacagtga	caatgactgg	gatagtgttgc	120
tgcctggccct	cctggggctg	gtgtgggtgg	ctggccacagc	gctggccctt	cactgggttg	180
gtccctggggca	cgatcaagag	gcccagaac	cggtgtccac	agccctcggt	gctcaaccc	240
atcaggcagg	aggagcttgc	ctggccctgc	aacccggatc	taagggtcgt	gtggcagcgc	300

-continued

---

gggggcagag	cccaggcag	gggaaaccag	agcccccaagg	acgcggccag	cagagccctg	360
tcctctgc	acgcggccgg	gggggcctgg	ccgocatggc	ccggcttcca	ctcaagacgg	420
ctgtcgagga	ggccgcaga	gaggcattag	gacagcaacg	ggcagtgcc	accccccgcgg	480
ccccccgagc	ggaaggaaag	gagccctcca	ggccaggcac	tgcctctcg	ggcaggagcg	540
aagcaggggg	gatgtccgcc	cccctctga	tccacttac	tcctcgagc	cctggcagcg	600
aagcggaggc	ggagacaggt	ggtgtcaggg	cgtcctctcg	ccaggeccca	ggcccccgcgg	660
ggcaacagga	cactggccc	ttggcaggcgg	ggcgggggcc	ctcggtctcg	atggggagag	720
ggcggggccc	ggggggcgg	atggacgtg	gtctgggaga	cagageccgc	cgccccggg	780
aactggacc	gtccgctg	ggcgcgcgg	ggagcgtgt	ggacgegggt	gacggggccg	840
ccgcccgtga	cgcccaacgc	cgggcctcc	ccacaggacc	cccaactcgcc	caggagccg	900
cactccggc	gtgtccgg	ccccggccc	tgaagctcg	gtctcaagac	gaaggctctg	960
gggccaaggg	ttggctggagc	agggaggct	cgggggctcc	tgcctccgg	ggaggctggc	1020
cctgggtcag	cagggggtc	ccgggcaccc	ggagcttgg	cccageccca	gactccacgc	1080
gcccctggct	agagagtct	cctcaaggg	gcccacttc	gtcccaaggg	ccgggttggca	1140
caggggccta	cgatgcggc	gaggggggg	ctgacagtc	ccgagataac	agtctgtccg	1200
ctgacccctgg	ggccacccgg	ccccccggg	aaggcaaggc	ggctgcaggc	ggccacacgc	1260
gegcgcctc	ccggagccgt	gagectcgcc	cgcgctccg	ctccccgccc	gcagetcccg	1320
gcccgggggtt	ccccactgta	gcctgtactc	tcccctctcc	ttcagactt	ttgcccctgg	1380
aggttaccca	ggatcttcc	ttggggaaa	atctcagagc	ggcgccggcc	ccaaagtccag	1440
cctcagccca	agtcttaact	tcagtcagg	cctcagtc	agccccagcc	ctgggttcat	1500
cccccaagtc	agcacaacc	tcagccacca	cctcaaccc	atccccacc	tcagecccg	1560
ccccagtc	aacctcagtc	cccaacttca	ccccagcccc	agccccaaat	ccagctgcag	1620
ccgcaactcc	agccccagcc	ccagttcccg	tcccaaccc	cacccccca	ttcccaaggcc	1680
taaccccaagt	cccaaaaaaa	gccttaagcc	caagtcacac	tccagcccta	acccccaggcg	1740
catccccagc	cttaaaaaaa	gtcccaaccc	cagccctaag	ccccagetca	actccagcccc	1800
caaccccaagc	ccgcacccct	gccccagcc	ccacccagcc	ccccacccca	acccccaggcg	1860
catccccccg	cccgatgtac	gggttaaagg	cttagagag	tgtggcttc	ccccaggcg	1920
accaggaggg	ggaggctctc	ggccatgtt	gaaaaccttat	ttccatgttt	tttagaaagcc	1980
accccttccc	ccaggcaagac	aggccccaa	ggagtgtccc	gagggggggtt	cccgggagcc	2040
ccgtgggtcc	cagcaacttc	acacactctg	aggacagaca	cgccccctct	tcttcagttgg	2100
ggacagtcat	aggggacagg	acaggggggg	ttggtagggc	ttggggatcg	ccacagccaa	2160
gaagctccg	gaccaacgg	tcggccaggc	caagccctcc	ccccggctca	agaggagagg	2220
gaaccaggga	gaaaaagtcta	gaccggctc	cccaagccgc	gatgccccagg	ggcccccgcac	2280
agccccccgc	gcagaggccg	cctggccccc	cgcccttc	ctctgcgggg	cgctcacagc	2340
cggttaccc	gtacggaaa	cgccggcggt	gcgaaatcgc	cccgagctcg	gaggcaggagg	2400
tcaggccggc	cgccctgggg	gacccttaag	gggggggg	gggggggggg	ggcagccctg	2460
ccggcccgag	cgggggcgctc	acggaaaago	aggaggaggc	ccggaaagctc	atgggttttc	2520
tgcagaggcc	cggggggttgg	gggggtgtgg	agggggccccc	gaagccccgc	tccggggccc	2580
tggagccccc	cacggccgg	gcctggcgcc	ggggggctgg	cctggggact	tgcctggacg	2640
tgtgtggctt	tgcctggagc	cacggagac	ccggcctggc	gcagggagacc	tacgcgtcga	2700
tgagcgtacaa	ctctgtcgca	gtgtgggg	acccgtgtct	ctaccgcgg	ctgagcgcgg	2760
ccgaccgcgca	gcgcacccctc	agcctcgccg	ccggccgggg	ccgggggggt	ctgggggtcc	2820
tgcgtactcc	cagccctctc	caggggggcc	gctcaggggct	ccccaggggc	cctcgtggcg	2880
aggagcttcc	tgcggggcc	cctgtctacc	tgcgcaccc	ctatgttca	2940	
accccccggg	gaacaccc	ccggccctga	ccccagggtc	cgaggaggcc	ccggttccgg	3000
gtgtgggtct	ctgcacccat	cacaactacc	tgtttctggc	ggggggcatc	cgtggctccg	3060
gtgtccaaaggc	cgctgtctcc	aacgggttct	tctgtctacc	ccctctgacc	aacatctgg	3120
cccgagggtcg	ccccatcgac	caggcccgag	ccccgtccaa	gtgtggggcc	ctggacgggg	3180
tgtctatgtc	catcggtggc	gaatgcctgt	acacgcattg	gtgtctacc	ccgcgcacag	3240
acgcctggac	cccacgcgc	ccactccccc	caggcacctt	ccctgtggcc	cacggaggct	3300
tggcctggcc	tggggacatc	tacgctccgg	gggggttacc	cttctaccgc	ctgttcagg	3360
acagccccgg	taaggatgtc	ttggggatgt	gcccatacag	tgccagccac	ccggcgttcca	3420
ccgacatcg	ggcactgggg	ggcttctgt	acccgttccg	cctgtcgccg	ggcatgggg	3480
ccgcctgtat	gcgcataaac	acgtgtacc	gtcccttgg	cagggtgtcc	tccctgc	3540
tgcctccccc	cgccccactc	cactgcacca	ccctgggca	caccattac	tgcctcaacc	3600
cccaagggtac	tgcacccatc	acggctctcg	ggggggactc	ccagttcccg	gccaaggagg	3660
tgcagccctt	ccccctgggg	agcaccgggg	tccctagtc	attcatctc	actctgc	3720
ctgaggaccg	gctcgacacc	tcactctgag	tggcaggcg	agaaccaaag	ctgttcgtct	3780
gtctcttccagg	gagcccttc	ttggatgggg	ctgagaggcc	ggggcttcc	gaaggggtct	3840
ggatcggaac	ttctctgtct	ttgttctgg	caactttcc	cttctgtctt	aaagggttgc	3900
gattattttgc	aaggcccgac	tccctccgg	tctttctgc	cctctacttca	caccccgact	3960
gtttctgtac	tcaattccgt	acctacttac	agacccttc	agctgtctga	caccccccgt	4020
tctgtgggac	tcccttatcc	ctagagccag	ggactgtatgc	gtctccacag	acaaggactt	4080
ggctctcggtt	agctctgtct	agcggagagg	gggggggttgc	aaaaacatc	acacttctca	4140
tgtctgtac	tgccaggcagg	gagcaaaaa	gtccccccgg	aacggccctcg	cctctgggac	4200
tttctgtctg	tcttaaggcc	tcccccaggta	ccaaacccctgt	agctatctgg	gtctgtttgg	4260
cactgtggat	tctcaaggcc	ctagaaccct	tgcctctgaa	actggtccgc	ttgtgcagcc	4320
ctgtctgtctg	cagccctctc	ccatcccccc	agcccacacc	aggccaggcc	cactccggcc	4380
tcaccaccc	ctgcacccct	gtggggctc	cccaaggccc	ccagaacccc	accccaccc	4440
tcgccaaccc	ccgatctctc	aatggggct	gagcgtcacc	ctatgttctc	cccttttttag	4500
ctgtgtac	ttggacgaga	cattgtactt	ccctttctcc	ttgtctat	aatgtggaca	4560
gtggacgtct	gtcaccacag	agatgtgtgg	gagacaagat	cacagctatg	agcacccctc	4620
acgggtgtcc	ggatgcacag	cacaatccat	gatgcgtttt	ctccccctac	gcactttgaa	4680
acccatgtca	aaaaagtgtaa	tacatctgac	tgtgtcccac	tccaaaccc	agcctggatg	4740
tccctgtctg	ggcccttttt	ctgttttttta	ttctatgttca	agcaccactg	gcaccaaaata	4800
catttaattt	caccgaaagc	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	4837

---

-continued

---

```

SEQ ID NO: 10      moltype = RNA  length = 2099
FEATURE          Location/Qualifiers
source           1..2099
                 mol_type = mRNA
                 organism = Homo sapiens
SEQUENCE: 10
ccacctcagg cccaaaccca accccagcg catccccgtc cccagctgac gggtaaaagc 60
ctcaggagag tgggttctc cccaggcgct accaggaggg gcaggtctca gccagctggg 120
gaaacccttat tgccatgtt cttaaagccc acccccttc caggcaagac aggccccaaag 180
ggagtgttcc gggggcggtt cccgggagcc cctgtgggtcc cagacttcc acacactctg 240
aggacagaca cggccccctt ctccatgtt ggacagtcat agggacaggt acagggggcc 300
tgggtgaggc tggaggatcc ccacagccaa gaagctccga gaccaacgga tcgcccagcc 360
cagaccctcc cccaggccca agaggaggg gaaccaggaa gaaaagtcta gacccgcgtc 420
ccaaagegcg cgtcccgagg ggccccgcac ageccccccgc geagaggccg cctggcccg 480
cggcccttcctt ctctcgagg cgtcacacg cggtaaaaaa gtcacggaaa cgcaacgggt 540
gcgaaatcgc cccgagctcg gagcaggagg teagggcgcc cgccctcgaaa gaccctcaag 600
gggaggcgcc gggggggggg ggccggccctg cccggcccgag cggggcgccctc acggaaaagc 660
aggaggaggc cgggaatcgc atgggtttc tgcagaggcc cgggggttgg ggggtgggtgg 720
aggggggcccg aaagcccccggc tccggggccca tggaggccgc cacggccgca gcccctcgcc 780
gggggcttggc cttggggcgtt tgcctggacg tgcgtggctt tgcctcggcag cacggagggc 840
ccggcccttgc gcaaggaggacc tacggcgcttgc tgacgcacaa cctgtcgca gtgtggggag 900
accgggttcccttccatggcccttgc ctggccggcccttgc cccggccggccatcc acggcttgcgg 960
ccggccgggg cccggccgttgc tccggccgttgc tgcgtacttgc cagcccttac cagggggggcc 1020
getcagggtt ccccaaggggc cctcgatggcc agggcccttc tgcggggggcc cctgtgttccc 1080
tgccttacc tgcgttccatgc catgttca accccccggaa gaacacccggggccccccttgc 1140
cccaagggtcc cggaggaggcc cccgttccggg gtcgggggttctc ctgcacccatgc cacaacttacc 1200
tgggttcttgc gggggggccatc ctggccgttccggg tggccaaaggcc cgtctgttcc aacgggggtt 1260
tctgttccatcc cccctctgacc aacatcttgc gccagggttgc gcccattgcag caggggccgg 1320
cccaagcttccatccatggcccttgc ctggccggggc tgcgttccatgc catcggttgcg gaatgttccgt 1380
acagcttgcg gtcgttccatgc ccggccaaacg acggccgttgc accacccggcc ccaactcccg 1440
caggccatcc cccctgtggcc caggccgttgc tggccctggcc tggggacatc tacgttccatcc 1500
gggggttccatccatccatggcccttgc ctggccgttccggg acggggcccttgc tggggatgtt 1560
gcccatacatc tgccagccac cggcggttccaa ggcacatcgat ggcactgggg ggcttccgtt 1620
accggccatccatggcccttgc ctggccgttccggg ggcacatggcc cccggccgttgc ggcacatccatccatggcc 1680
gctccctggatccatccatggcccttgc tccctggggcc tggccggccccc cggccggccatccatggcc 1740
ccctggggccatccatggcccttgc tggccatccatccatggcccttgc acggccatccatggcccttgc 1800
ggggggacttgc ccagggttccatccatggcccttgc gccaaggagcc tgcagcccttgc ccccttgggggg 1860
tccctcgttccatccatggcccttgc actcttgcggcc ctggaggacatccatggcccttgc tcaacttgcgg 1920
tggccaggccatccatggcccttgc ctgttccatggcccttgc gtcgttccatggcccttgc tggggatggggc 1980
ctggaggccatccatggcccttgc gggggccatccatggcccttgc gatccatccatggcccttgc tggggatggggc 2040
caactttccatccatggcccttgc aaagggttgc gattatatttggaaaaaaa 2099

SEQ ID NO: 11      moltype = DNA  length = 5125
FEATURE          Location/Qualifiers
source           1..5125
                 mol_type = other DNA
                 organism = Homo sapiens
SEQUENCE: 11
agaaggcagggt tggctgttgc gacagcacag agtcaggaa cgtgttgcggaa ggacccttggg 60
gcctacaggc aggaaagaaatggccaggatgttgggggttgc ttgcaggagac cttggggctcc 120
tatccctgcata taaggccttc tgcgttgcgttgc tatcttgcgtc caggcccttac tgacccccc 180
aggcccttgcgttgc gcaaggaggatccatccatggcccttgc tcccaaggccc tggaggacttgc gtccaccc 240
actggccaggcc ctttggggccatccatggcccttgc ggcgttgcggcc ggtgttgcgttgc cccaggcccttgc tgcgttgcgttgc 300
atgcacttgcg acccttgcgttgc caggcccttac tgcgttgcgttgc gacccatgttgc cccaggcccttgc 360
ttggccaggccatccatggcccttgc atggggccatccatggcccttgc ctgggggttgg gactggggccatccatggcccttgc gtcgttgcgttgc 420
gtctgttgcgttgc ccccttgcgttgc gtcgttgcgttgc tggccatccatggcccttgc cccaggcccttgc 480
tttggccatccatggcccttgc ggcacatccatggcccttgc agggccggccatccatggcccttgc gacccaggcccttgc 540
cttccatccatggcccttgc caggccggccatccatggcccttgc tgccatccatggcccttgc agtcttgcgttgc cccaggcccttgc 600
ageggccggccatccatggcccttgc agggccggccatccatggcccttgc gtcgttgcgttgc cccaggcccttgc 660
cctgttccatccatggcccttgc ctggccatccatggcccttgc tggccggccatccatggcccttgc tccacttgc 720
acggccatccatggcccttgc caggccggccatccatggcccttgc tggccatccatggcccttgc tggccatccatggcccttgc 780
ggggcccccggccatccatggcccttgc gaggccggccatccatggcccttgc aaaggccatccatggcccttgc cccaggcccttgc 840
agccggccatccatggcccttgc gggggatgttgc cccccccttgc ctgttccatccatggcccttgc tcaacttgcgttgc gacccaggcccttgc 900
agccggccatccatggcccttgc aggggttgcgttgc agggccatccatggcccttgc tcccttgcgttgc cccaggcccttgc 960
gcggggccatccatggcccttgc ccccttgcgttgc gtcgttgcgttgc tggccggccatccatggcccttgc tccacttgc 1020
agaggccggccatccatggcccttgc gggggatgttgc tggccatccatggcccttgc gacccaggcccttgc 1080
cgaaaacttgc acccccttgcgttgc cccccccttgc ctgttccatccatggcccttgc gggggatgttgc tggccatccatggcccttgc 1140
gcggccggccatccatggcccttgc tggccatccatggcccttgc cccccccttgc gtcgttgcgttgc tggccatccatggcccttgc 1200
cccgccatccatggcccttgc cccccccttgc ctgttccatccatggcccttgc tggccatccatggcccttgc tggccatccatggcccttgc 1260
tctggggccatccatggcccttgc aggggttgcgttgc gtcgttgcgttgc tggccatccatggcccttgc cccaggcccttgc 1320
tggccatccatggcccttgc tggccatccatggcccttgc gtcgttgcgttgc tggccatccatggcccttgc tggccatccatggcccttgc 1380
acggccatccatggcccttgc gtcgttgcgttgc tggccatccatggcccttgc gtcgttgcgttgc tggccatccatggcccttgc 1440
gccacaggccatccatggcccttgc cccccccttgc ctgttccatccatggcccttgc tggccatccatggcccttgc tggccatccatggcccttgc 1500
gcggccatccatggcccttgc tggccatccatggcccttgc gtcgttgcgttgc tggccatccatggcccttgc tggccatccatggcccttgc 1560
agccggccatccatggcccttgc cccccccttgc ctgttccatccatggcccttgc tggccatccatggcccttgc tggccatccatggcccttgc 1620

```

-continued

---

cccgccccgg	ggttcccacc	tgaagccctg	actctccct	ctccctcaga	cttttgccc	1680
cttgagggtta	cccaaggatcc	ttccgtgggg	gaaaatctca	gagcggggcc	agccccaaagt	1740
tcagecctcag	cccaaggatcc	aacctcagct	ccagecctcag	tectagcccc	agccctgtgt	1800
tcatccccca	gtcagcaccc	aacctcagcc	accacctcaa	cctcatcccc	cacccatcgcc	1860
ccagccccag	ctccaaacctc	agctccaaact	tcaaccccaag	ccccagcccc	aagtccagct	1920
gcagccgcaa	ctccagcccc	agccccagtc	ccagtcggcc	ccctcacacc	ccccatcccc	1980
gcctctaacc	cagtcggcc	cccagccctaa	agccccagtc	caactccagc	cttaacccca	2040
gecgcatccc	cagccctaac	cccagccctaa	acccccagcc	taagccccagc	tccaaactcca	2100
gcuccaacc	cagccgcate	ccctgcccc	gccccccact	cagccccaaac	ccccacccca	2160
geccgcattcc	ctgccccagc	tgacgggtca	aaqccteagg	agagtgtggc	tctcccccagg	2220
cgttacccagg	ggggggagggt	ctcagccagc	tggggaaacc	ttatgtccat	ggtttttaga	2280
agccacccct	tcccccggca	agacaggccc	caagggagt	tcccggggc	ggttttccggg	2340
agccccgtgg	gtccccggac	ttccacacac	tctgaggaca	gacacggcc	cttcttctca	2400
gtggggacatg	tcatatggcc	aggatcaggg	ggcttgggtt	aggctggagg	ttagccacag	2460
ccaagaagct	cggagaccaa	cggatcggcc	agccccagcc	ctcccccaag	cctaagagga	2520
gagggaacca	ggggaaaaag	tctatggcc	ctggcccaag	ctggcatgcc	caggggcccc	2580
gcacagcccc	ccgcgcagag	gcccctggc	ccgcggccct	ctccctctgc	gagggcgtca	2640
cacccgggtac	cccaactcag	gaaacgcgc	aggtgtcgaa	tcgccccggag	ctcgaggcag	2700
gagggtcggc	ccgcggccctc	gggggaccct	caaggggggg	cgccccgggg	ggggggcagc	2760
cctgcggccc	gcagggggcc	gctcggccaa	aaagcaggagg	aggccccggaa	gctcatgtg	2820
tttctgcaga	ggccgggggg	ttgggggggt	gtggaggggc	ccggaaagcc	cagctcccg	2880
gcocctggcgc	ccgcgcacggc	ggcagccctgg	ccggggccgc	tggacctgg	cagttgcctg	2940
gacgtgtcg	cccttcggca	gacacggcc	gagccccggcc	tggccggcgg	gacccatcg	3000
ctgtatgcgc	acaacccctgt	gctgatgtcg	ggagacccct	gcctctaccc	ccggctgagc	3060
geggccgacc	cgccgcgcac	cctcagccgt	cgccacggcc	ggggccgggc	gggtgtggc	3120
gtctctgtac	tgcccccgcct	ctaccagggg	ggccgcgtca	ggctccccag	ggccctctcg	3180
ggcgaggagc	ctccgcggcc	ggcccccgtg	ccctgcgtca	ccctgcgtg	3240	
ttaaaccccc	ggggaaaaac	ctggggggcc	ctgaccctgg	tggccggagg	ggccccgtt	3300
cggggctgcg	gtctctgcac	catgcacaa	tacccgtttc	tggggggggg	catccgtggc	3360
tccgggtcgc	aggccgtctg	ctccaacgg	gtctctgtc	acaacccctt	gacccaaatc	3420
tggaggccagg	tccggccat	gcaacgggg	cgagccccgg	tcaagctgg	ggccctggac	3480
gggctgtct	atggccatgg	tggcaatgc	ctgtacagca	tggagtgct	cgaccggcga	3540
acagacgcct	ggaccccaac	cgccgcac	cccgccaggca	ccttccctgt	ggccacccag	3600
gctgtggcct	ggcgtgggg	catctacgt	accgggggtc	accccttcta	ccgcctgtc	3660
aggtacaccc	ccgttgcgg	tgccgtgggg	gagtgcccat	acagtgcac	ccacccgggt	3720
tccagcgcaca	tcgtggca	gggggggtc	ctgttccgt	tcgacccgt	ggggggcgt	3780
ggcgccgcgc	tgtatgccta	caacacgtg	accggctcc	ggagcaggcc	tgccctccgt	3840
ccccctggcc	cccccccccc	actgcactg	accaccctgg	gcaacaccat	ttactgcctc	3900
aaccccccagg	tcaatccatc	cttcgcggcc	tctgggggg	ctgcccagg	ccaggccaa	3960
gagetgcgc	cctttccctt	ggggggacc	gggggtctca	gtccattcat	cctgactctg	4020
ccccctggagg	accgggtgc	gacccactc	tgatggccag	cgacagaaacc	aaagctgtt	4080
cgctgcgtct	cgaggagacc	ctccctggat	ggccctgaga	ggccggggct	cagggaaagg	4140
gtctggatcg	gaaacttcctg	ctctgtttc	tggacaactt	tcccttcctc	ttttaatgg	4200
tgtcgattat	tttgaagccc	agacttcctt	agecttcctt	tgecccttcac	tccacaccca	4260
gactgtttcc	tgtactcaatt	ccgtacccat	ttacagacc	tctcgttgc	ctgacccccc	4320
cctgtctgt	ggactcccta	ttcccttagag	ccagggactg	atgcgtctcc	acagacaagg	4380
acttggctcg	tcggagatcg	gtctggcc	gagaggagg	ggtagaaaaac	attcacactt	4440
cctatgtct	tcctgcggca	ggggggccaa	aaaatgtccc	aggcaacggc	ctgcctctg	4500
ggactttctg	ctgtctctaa	ggccctccca	ggtaccaacc	ccgtagotat	ctgggtctgt	4560
ttggcactgt	ggatttcata	ggcccttagaa	ccctgtcctc	tgaaactgg	ccgctgtgtc	4620
agccctgtcg	tctgcgtc	ctgcctatc	ccccagccca	caccaggcc	ggccctactcc	4680
gggctccatca	cccttcgtcg	cctgtgggg	ctctccccc	ccctccggaa	gcccacccca	4740
cttctcgccca	accccccata	tctaatq	gcttgcgcgt	caccctagg	ctgccccctt	4800
tttagtgcgt	agacttggac	gagacattt	acttccctt	tccttgcgt	ataaaatgt	4860
gacagtggac	gtctgtcc	caagagatgt	gtggagacaa	agatcacac	tatgacacc	4920
tcgcacgggt	tccaggatgc	acagacaaat	ccatgtgc	ttttctcccc	ttacgcactt	4980
tgaacccat	gttagaaaaag	tataatcata	tgactgtcg	ccactccaaac	ctccagctg	5040
gatgtccctg	tctggccct	ttttctgttt	tttattctat	gttcagcacc	actggcacca	5100
aatacatttt	aattcaccga	aagca				5125

SEQ ID NO: 12 moltype = DNA length = 2990  
 FEATURE Location/Qualifiers  
 source 1..2990  
 mol\_type = other DNA  
 organism = Homo sapiens

SEQUENCE: 12  
 ccacctcagc cccaaaccca accccagccg catccccctgc cccagctgac gggtcaaagc 60  
 ctcaggagag tggcgtctc cccaggcgc accaggaggg gcaggtctca gccagctggg 120  
 gaaacccat tatccatgtttt cttagaagcc acccccttccc caggcaagac aggccccaa 180  
 ggagtgttcc ggggggggtt cccggggagcc cctgtgggtcc cagcacttcc acacactctg 240  
 aggacacaca cggccccctt tcttcgttgc ggacagtcat agggacaggt acagggggcc 300  
 tggttgaggc tggagggtcag ccacagccaa gaagctccga gaccaacggg tcgccccagc 360  
 cagaccctcc cccaggccctaa agaggagagg gaaccaggaa gaaaagtcta gaccgcgtc 420  
 cccaaaggccgc gatgcccagg gggcccccgc acggcccccgc cctggccccc 480  
 cggccctctc ctctgcgagg cgctcacaac cggtacccca gctacggaaa cgcagcagg 540  
 gggaaatcgcc ccccgatctcg gaggcaggagg tcaggccggc cgcctcgccc gaccctcaag 600

-continued

---

gggagggccc	ggggggaggggg	ggcagccctg	ccggccgcga	cgggggcgtc	acggaaaagc	660
aggaggaggc	ccggagactc	atgggttcc	tgcaagaggcc	cggggggtgg	gggggtgggtgg	720
agggggccccg	gaagccacgc	tccggggccc	tggagccgc	cacggggcga	gcccctgcggc	780
ggccggctgg	ctggggcagt	tgccctggacg	tgctggctt	tgcccacag	cacggagagc	840
ccggcctggc	gcaggagacc	tacgctgtga	tgagcgacaa	cctgctgcga	gtgctggag	900
acccgtgcct	ctaceggccgg	ctgagcgcgg	ccgacccgcga	ggcgcatttc	agccctgcgg	960
ccggccgggg	ccggggcggtg	ctggggctcc	tegtactgc	cagectctac	caggggggcc	1020
getcaggggc	ccccaggggg	cctctgggg	aggagctcc	tgccggggcc	cctgtgtccc	1080
tgcctctacc	tgcgeacctg	catgtttca	accccccggg	gaacacctgg	cggccctgta	1140
cccaagggtcc	cgaggagggc	ccgcttcggg	gtgtgggtct	ctgcacccat	cacaactacc	1200
tgttttggc	ggggggggatc	ctgtggctcc	gtgccaaggc	cgtgttgc	aacgagggtt	1260
tctgctacaa	ccctctgacc	aacatctgga	gccaggatcg	gcccattcg	caggcccgag	1320
cccaagctcaa	gctgggtggcc	ctggacgggc	tgtcttatgc	categggtgc	gaatgcctgt	1380
acacgatgta	gtgtctacgc	ccgcgaacag	acgcctggac	ccacacgcgc	ccactccccc	1440
caggcacctt	ccctgtggcc	cacgaggctt	tgccctgcgg	tggggacata	tagtcaccg	1500
ggggtcaacct	cttctaccgc	ctgtctgggt	acagccccgt	gaaggatgt	tggacgagt	1560
gccccatcac	tgccagccac	cggggttca	gegacatcg	ggcactgggg	ggcttcctgt	1620
accgcttcga	ctgtctgggg	ggcgtgggg	ccgcgtgtat	ggctcataac	acagtgcac	1680
gtctcttgag	cagggtgtcc	tcctctcccd	tgccttc	ccggccccc	cactgcacca	1740
ccctggggaa	caccatttac	tgcctcaacc	cccaaggatc	tgccaccc	acggtctctg	1800
gggggactgc	ccagttccag	gccaaggago	tgcaagccct	ccccctgggg	agcacccggg	1860
tctctcgtcc	attcatactgc	actctggcc	ctgaggaccg	gtctcgaccc	tcactctgag	1920
tggcaggcag	gaaaaaaaag	ctgtttcg	gtcttcagg	gagacccctc	tgggatgggc	1980
ctgagaggccc	gggggttcagg	taaagggtgt	gattatttt	aagcccagac	tccctcagcc	2100
caactttccc	cttctgtttt	aaagggtgt	gatcttgc	tttttttttt	acctacttac	2160
tctttctggc	cttcactc	cacccagat	gttctctgac	tcaattccgt	tttttttttt	2220
agaccccttc	agtttgcgt	caccccccgt	tctgtgggg	tcccttattcc	ctagagccag	2280
ggactgtgc	gtctccacag	aaacggactt	ggctcgtgt	agctctgt	agccgagaga	2340
ggaggggggt	aaaaacattc	acacccctta	tgtctgtca	gcaggacagg	gagaaaaaac	2400
gtccccccaggc	aacgcctc	cctctgggg	tttttgcgtt	ttcttaaggcc	tcccccaggta	2460
caaccccccgt	actatctgg	gtctgtgg	taagggtgt	tctcaaggcc	ctagaaccct	2520
tgctctgaa	atgttgcgg	tggtcagcc	ctgtgttgc	cagtcctgc	ccatcccc	2580
agccccacacc	aggccaggcc	cactccggc	tcaccaccc	ctgcagoc	tttgggctct	2640
cccagccct	ccagaagccc	accccaactt	tcgccaacc	ccgatctca	aatggggct	2700
gagcgttcc	ctagtctgc	cccttttttag	ctgtgttagac	ttggacggaga	catttgactt	2760
cccttttcc	ttgtctataa	aatgtggaca	gtggacgt	gtcacc	agagttgtgg	2820
gagacaagat	cacagctatg	agcacctcg	acggtgttca	ggatgcac	cacaatccat	2880
gatgcgtttt	ctccccttac	gcactttga	accatgtca	aaaaagtggaa	tacatctgac	2940
ttctatgttc	agcaccact	gccaataata	cattttaaat	caccgaaagc	2990	

SEQ ID NO: 13	moltype = DNA	length = 4837				
FEATURE	Location/Qualifiers					
source	1..4837					
	mol_type = other DNA					
	organism = Homo sapiens					
SEQUENCE: 13						
actggccagcc	tgccatcagg	cctctattgc	agccctgaac	catgtatccag	ggcaccttgg	60
agccagatgg	tcccctctgg	ggctgggact	gggacagtga	caatgtactt	gatagtgtgt	120
tgctggccct	cctggcgtg	gctgtgttgg	ctggccacage	gtgtgttca	cactgtttt	180
gtcccccggc	cgatcaagag	ggcccgagac	cggtgttcac	agccctcggg	gtcaacactc	240
atcaggcagg	aggagatgt	ctggccctgc	aaccgaatgc	taagggtgt	gtggcagcgt	300
aggggcagag	cccaaggcg	ggggaaaccag	agccccccagg	acgcggccag	cagacccctg	360
tcctctgtgc	agcgcggggc	ggggggctgg	ccggccatggc	ccgggttcca	ctcaagacgg	420
ctgtcgagga	ggcccccggc	gaggccattt	gacagcaacg	gggcgtgtcc	accccccggg	480
ccccccggc	ggaaggaaag	gaccccttca	ggccaggccac	tgcctccctg	ggcaggacgc	540
aacgcaggggg	gtatgtccgc	ccctcttcttgc	tccacttcac	tcctcggtac	cctggcagcg	600
aacgcggaggc	ggagacagggt	ggtgtcaggg	ctgtctctcg	ccaggccgca	ggcccccggg	660
ggcaacagga	cactgtccgc	tggcgggggg	ggcgccggcc	ctcggttctcg	atggggagag	720
ggccggggccg	ggggggccgg	atggacgt	gtctgggg	cagacccgc	cgccccccgg	780
aactggaccc	gttccgtctg	ggccgggggg	ggacgtgtgt	ggacgtgtgt	gacggggccg	840
ccggccctgg	cccccacgc	ccggccctcc	ccacaggacc	cccactcgcc	caggagcccg	900
cactccccc	gttccgtctg	ccccggccc	tgccagccctg	gtctcgac	gaagggtctg	960
ggggccagg	ttggcggtgg	aggggggcc	cggggggtcc	tgccccccgg	ggagggttggc	1020
cctgggttc	cagggtggc	ccggccaccc	ggacgttttg	cccaagccca	gactccacgc	1080
ggccctgg	agagacttgc	cctcaagttc	gtcccaacgg	ccgggttcca	tttggccctgg	1140
caggggccta	cgatgtccgc	gaggccgggg	ctgacagtc	ccgagataac	agtctgtccg	1200
ctgacactgg	ggccacccgg	ccccggaga	aagcaaaac	ggctgcagcc	ggccacagcc	1260
ggccgcctc	ccggagccgt	gagccctgc	cgccgtccgc	ctccccccccc	gcagctcccg	1320
ggccgggggtt	cccaactgtaa	ggccctgtact	tcctctctcc	ttcagactt	tttggccctgg	1380
aggttaccca	ggatcttcc	gtggggc	aaatctc	ggcgttcc	ccaaatcc	1440
cctcagecc	agtcttaact	tcagtc	cctcagtc	agccccagcc	ctggttccat	1500
cccccaagtc	agcaccac	tcagccacca	cctcaaccc	atcccccc	tcagccccag	1560
ccccagctc	aacctcag	ccaaatcc	cccaaccc	tcagccac	ccccccagcc	1620
ccgcaactcc	agccccagcc	ccagttcc	tcccaaccc	cacacccca	tcccagcc	1680
taacccca	cccaacccca	ggccctaa	tcagccctt	accccccc	tttggccctgg	1740

-continued

cateccccagc	cctaacccca	gtcccaaccc	cagccctaag	cccagctca	actccagccc	1800
caacccccagc	cgccatccct	gcccagccccc	ccacccctcgc	cccaacccca	acccccagccg	1860
cateccccctgc	cccagctgc	gggtcaaaggc	ctcaggagaa	tgtggcttc	cccagggcgt	1920
accaggagggg	gcaggcttca	gccagctggg	gaaaccttat	tgccatggtt	cttagaagcc	1980
accccttccc	caggcaagac	agggcccaag	ggagtgtccc	gagggegggtt	cccgggagcc	2040
cctgtgggtcc	cagcacatcc	acacactctg	aggacacaga	eggccccctt	tcttcagtgg	2100
ggacagtcat	agggacaggt	acagggggcc	tgggtgagge	tggaggtcag	ccacagccaa	2160
gaagctccga	gaccacggta	tcgeccagec	cajacccctcc	cccaggocca	agaggagagg	2220
gaaccaggga	gaaaagtcta	gacecgtgc	ccaaageccgc	gatgcccaag	ggccccgcac	2280
acgccccccgc	cgagggccgc	cctggccccc	cggcccttc	ctctgcegggg	cgctcacacgc	2340
cgttacccca	gctacggaaa	cgcagcagggt	cgaaatcgc	cccggatctg	gagcaggagg	2400
tcagccggc	cgccctgggg	gaccctcaag	ggggagggccc	gggggaggggg	ggcagccctg	2460
cggccgcag	cgggggcgtc	acggaaaagc	aggaggaggc	ccggaaagctc	atggtgttcc	2520
tgcagaggcc	cggggggtgg	gggggtgtgg	agggggccccc	gaagcccaagc	tcccccggcc	2580
tggagccgc	cacggccgca	gcccgtggc	gggggtgggg	cctggggcgt	tgcgtggacg	2640
tgcgtggctt	tgcccaaggc	cacggagac	ccggccctgc	cgaggaaaccc	tafcgtctga	2700
tgagcgacaa	cctgtgtcga	gtgtgtggag	accctgtccct	ctaccggccgg	ctgagcgggg	2760
cgaccccgoga	gogccatctc	agccctggga	ceggccgggg	ceggccgggt	ctgggggttc	2820
tgcgtactcc	cagectctac	cagggggccc	getcgggtcc	ccccaggggc	cctctgtggcg	2880
aggagcttcc	tcggccggcc	cctgtgtccc	tgccttaccc	tgcgcaccc	catgtgtca	2940
accccccggg	gaacacactgg	cgggccctga	cccaggtgc	cgaggaggcc	ccgcttcggg	3000
gtgtgggtct	ctgcaccatc	cacaactacc	tgtttctggc	ggggggccatc	cgtggctccg	3060
gtgtccaaaggc	cgctgtgtcc	aacgggttcc	tctgtctaca	coctctgacc	aacatctgg	3120
gcccagggtc	gcccattgcag	caggccccag	cccagctca	gtgtgtggcc	ctggacgggc	3180
tgcctatgtc	cacgggtggc	gaatgtctgt	acagcatgg	gtgtacggac	ccgcegaacag	3240
acgcctggac	ccccacggcg	ccactccccc	caggcacctt	ccctgtggcc	cacgaggctg	3300
tggcctggcg	tggggacate	taacttacccg	gggggtcacct	tttctaccgg	ctgttcagg	3360
acagccccgt	gggggtgtgt	ttggggcagg	ggccatatacg	tgccagccac	ccgcgtttca	3420
ggcacatcg	ggcaactgggg	ggcttccctgt	accgttctga	cctgtgtggg	ggcgtgggg	3480
ccgcccgtat	ggcgatcacaac	acagtggcc	gtccctggag	cagggtgtcc	tccctgcccc	3540
tgcctccggc	cgccccccgt	cactgtggcc	ccctggggca	caccatttac	tgcctcaacc	3600
cccaagggtac	tggccaccc	acggctctgt	ggggggactgc	ccagttccag	gccaaggagc	3660
tgcagccctt	ccccctgggg	agcacccgggg	tccctcgttcc	attcatctg	actctgcggc	3720
ctgaggaccg	gtgtcgacacc	tcactctgt	ttggcaggcag	agaaccaaag	ctgcttcgt	3780
gtctctccagg	gagaccctcc	ttggatgggg	cttggagggcc	gggggtcagg	gaagggggctg	3840
ggatcggaac	ttccctgtgt	tggttttggg	caactttccc	ttctgtgttt	aaaggttgc	3900
gattattttt	aaggccagac	tccctcgttcc	tcttttctgg	ccttcacttca	cacccagact	3960
gtttcctgtac	tcaattccgt	acctacttac	agacccttc	agettgtgt	caccccccgt	4020
tctgtggggc	tccctatcc	ctagtgccgt	ggactgtatgc	gtctccacag	acaaggact	4080
ggctcgctgg	agctctgtgt	agccctgtgt	ggggggggaa	gaaaacatcc	acacttctca	4140
tgcgtgtca	ggaggacagg	gagcaaaaa	gtccccggcc	aaccccccgt	cctctggac	4200
tttctgtgtt	tccttaaggcc	tccccaggta	ccaaccccccgt	agttatctgg	gtctgtttgg	4260
caactgtgtat	tctcaaggcc	ctgaaaccctt	tgctctgtgg	actgggtccgc	ttgtgtcagcc	4320
ctgtgtgtgt	cagcttcgtc	ccatataccc	tgcccaacacc	aggccaggcc	cactccgggc	4380
tcaccaccct	ctgcacccctt	gtggggcttc	cccaagccccc	ccagaagccc	accccaactt	4440
tcgcacccccc	ccgatctcta	aatggggctt	gagcgttacc	ctagttctgc	cccttttttg	4500
ctgtgtgtac	ttggggatgg	catttttttt	ttgtctataa	aatgtggaca	4560	
gtggaaatgt	gtcacccaa	aggggtgtgg	gagacaagat	cacatgtat	agcacccctgc	4620
acgggtgtca	gtatgtacag	caataccat	gtatgtttt	ctcccttac	gcacttttgg	4680
acccatgtca	gaaaagtgaa	tacatctgt	tgtgtccac	tccaaacctcc	agcctggat	4740
tcctctgtgt	ggccctttttt	ctgttttttta	tccatgtttc	agcaccactg	gcaccaaata	4800
catttttaattt	caccgaaagc	aaaaaaaaaa	aaaaaaaaaa			4837

SEQ ID NO: 14                  moltype = DNA length = 2099  
 FEATURE  
 source                  Location/Qualifiers  
 1..2099  
 mol\_type = other DNA  
 organism = Homo sapiens

SEQUENCE: 14  
 ccacccctcgc cccaaacccca accccagccg catccccctgc cccagctgc gggtcaaaggc 60  
 ctcaaggagag tgggtgtctc cccaggccgt accaggagggg gcagggttca gccagctggg 120  
 gaaaccccttat tgccatggtt ctttagaaggcc acccccttccc caggcaagac aggccccaaag 180  
 ggagtgttccc gggggccgggtt cccggggagcc ccgtgggtcc cagcacatctg 240  
 aggacagacca cggccctctt tcttcagtgg ggacagtcat agggacaggtt acagggggcc 300  
 tgggtggggc tggagggtcag ccacagccaa gaagctccga gaccaacggga tcgcccagcc 360  
 cagaccctccccc cccaggccctt aagggagggg gaaacccatca gaccggcttc 420  
 cccaaageccgc gatgcccaagg ggcggccgcac ageccccccgcg ccagggccgg cctggccccc 480  
 cggcccttc ctcgtgtccgg cgctcacago cggtacccca gctacggaaa cgcagccagg 540  
 gggaaatccgc cccgaaatccgc gaggcaggagg tcaggccggc ccctccgggg gaccctcaag 600  
 ggggggggggggg ggcaggccctg cccggccgcac ccggggggccgtt acggaaaagc 660  
 aggaggaggcc cccggaaatccgc atgggttttgc tgcagaggcc ccgggggttgg ggggtgggtgg 720  
 agggggccccc gaaaggccaggc tcccccggcc tggggccggc eacggccggca gcccctggcc 780  
 gggggccgttca cctggggcgtt tgccctggacg tgctggctt tgcccagccag cacggagac 840  
 cccggccctgc gcaggagacc tacggcgttca tgagcgtacaa cctgtgtccga gtgtgtgg 900  
 acccggtgtccctt ataccggccgg ctgagcgtccgg cccaccggcga ggcctccatcc 960  
 cccggccgggg cccggccgttcc ctggggcttc tgcgtacttgc caggcccttac 1020

-continued

---

gctcagggt	ccccaggggc	cctcggtggcg	aggagccctcc	tgccggggcc	cctgttccc	1080
tgccttacc	tgcgcacccgt	catgtttca	accccccggga	gaacacctgg	cgccccctga	1140
cccagggtcc	cgaggaggcc	cgcgttcggg	gctgcccgtct	ctgcacatcg	cacaactacc	1200
tgtttctggc	ggggggccatc	cgtggctccg	gtgcacaaggc	cgctctgtcc	aacgaggct	1260
tctgtctaa	ccctctgacc	aacatctgga	gccaggcttcg	gccccatgcag	caggccccag	1320
cccaaggctaa	gtgtggggcc	ctggacgggg	tgtctatgc	catcggtggc	gaatgcctgt	1380
acagcatggc	gtgtcacgac	ccgcgaacacg	acgcctggac	cccacgcggg	ccactccccg	1440
caggcaccc	ccctgtggcc	cacaggctg	tggccctggcc	tggggacatc	tacgtcaccg	1500
ggggtcaccc	cttctaccgc	ctgtcagg	acagccccgt	gaaggatgtct	tgggacgagt	1560
gcccatacag	tgcccaagcc	cgcgcttcca	cgacatcgct	ggcactgggg	ggcttccctgt	1620
accgcttcg	cctgtgggg	ggcgctgggg	cgccctgtgt	gctgatcacaa	acagtgcacc	1680
gctcctggag	cagggtgtcc	tccctggccc	tgcccccccc	cgccccactg	cactgcacca	1740
ccctgggca	caccatttac	tgcctcaacc	cccaggctac	tgccacccctc	acggctctgt	1800
gggggactgc	ccaggcttcag	gccaagggg	tgcagccct	ccccctgggg	agcacccggg	1860
tccctcgtcc	attcatctgt	acttgc	ctgaggaccg	gtcgacgacc	tcaactctgag	1920
tggcaggcag	agaaccaaag	ctgttcgt	gtcttcagg	gagacccctc	tgggatggc	1980
ctgagaggcc	ggggctcagg	gaaggggctg	ggatcgaa	ttcctgtct	tgtttctgga	2040
caactttccc	cttctgttcc	aaagggttgc	gattatttt	aaaaaaaaa	aaaaaaaaa	2099

SEQ_ID NO: 15	moltype = DNA	length = 5125				
FEATURE	Location/Qualifiers					
source	1..5125					
	mol_type = other DNA					
	organism = Homo sapiens					
SEQUENCE: 15						
agaaggcagg	tggctgtgtat	gacgcacag	agctcaggaa	cgctgcctga	ggacccttggg	60
gcctacgagg	aggagaagaag	ggcaggaggt	gttgggggtgc	tttcagagac	cctgggtctcc	120
tatccctggca	taaggctctgc	tatcttcgt	tatctgcacg	caggccctac	tgacaccccc	180
aggcctgagt	gcaaggcagag	accccccatt	tcccaggccc	tggaggactg	gtccacccctt	240
actggggcgc	cottttggca	ggcgtgtggcc	gggtgcctca	cccaggccctc	tgtgtctgc	300
atgcactgc	acgccttcgt	caggctcta	tttcagccct	gaaccatgtat	ccaggccacc	360
ttggagcctag	atggccccct	ctggggctgg	gactgggaca	gtgacaatga	ctgggatagt	420
gctgtgttgg	ccctcttggc	gctggctgt	gttggctcca	cagcgctggc	cttacactgg	480
tttggcttcg	ggcacgtca	agaggcggca	gaaccgggt	ccacagccct	cggggctcaa	540
cctcatcagg	caggaggcgc	tgagttggcc	ctgcaacccg	agtctaaagg	cagtgtatggc	600
agcggaggcc	agagccagg	gcaggggaaa	ccagagcccc	caggacgggg	ccagcagagc	660
cctgtccctg	ctgcacgc	gggcgggggg	ctggccgc	tggcccggt	tccactcaag	720
acggctgtcg	aggaggccc	cagagaggca	tttaggacagc	aacggggcag	tgccaccccc	780
ggggcccccc	gagcggggcc	aaaggaggcc	cccaaggccgt	gcactgcct	cttggggcagg	840
agcgaagcag	gggggtatgc	cgccccccctc	ctgatccact	tcactctcg	gagcccttggc	900
agcgaagcgg	aggcgagac	agggtgttc	agggcgtct	ctcgccaggc	cgcaggcccc	960
gcggggcaac	aggacactgg	cccctggcag	gcggggcgccg	ggccctcggt	ctcgatgggg	1020
agaggccggg	gcggggccgg	gcggatggcc	gttggctegg	gagacagacg	ccggccgc	1080
cgaaaaactgg	acccgttcg	cctggggccc	gccccggggcc	gttgggacgc	gttggacggg	1140
ggccggccccc	ttggacccca	cgccggccgg	ctccccacac	gaccggccact	cgccggagg	1200
cccgcaetcc	cgccgtgc	cgctccccgg	gcctgcacg	ctgggtctca	gacggaaaggc	1260
tctggggcoca	agggttgtgtc	gaggcagggg	gttgggggg	tccctgc	ccggaggaggc	1320
ttggccctgg	ttcaggggg	gttccccggg	accggggat	ttggccca	cccgacatcc	1380
acgcggccct	gtctagagag	tccctctca	gttgc	ccccccat	aggggccgg	1440
gcacacgggg	cttacgtgc	cgccggggcc	ggggctgaca	gttcccggaa	taacagtct	1500
ggccgttacc	ttggggccac	ccggggccccc	gaggcaagca	agccgggtgc	agccggccac	1560
agccggccgc	cttccgggg	cggtggccgt	cgccgtccccc	ccggccgt	ccccggccgt	1620
ccccggccgg	gttccccacc	tgaaggccctg	acttccct	tcccttcaga	ctttttggcc	1680
ctggagggtt	cccaggatcc	ttccgtggcc	gaaaatctca	gagccggccc	agccccaaat	1740
tcagccttcg	cccaaggatcc	aacttcgt	ccagcctc	ttctagcccc	agccctgtct	1800
tcateccccca	gtcagccacc	aacctcagec	accacccatc	cetcatecccc	cacccatcgcc	1860
ccagcccccc	cttccaaacc	agctcact	taacccctt	ccccagcccc	aagtccatgt	1920
gcagccgca	cttccagcccc	agccccatgc	ccagtc	cccttcacacc	ccccatcccc	1980
gccttaacc	cagtcacca	cccccaccc	agcccaatgc	ccactcc	ccctaaacc	2040
gcgcgcaccc	cagcccaac	ccccatcca	accccaatgc	taagccccc	tcccaacttca	2100
gcucccaacc	cagccgcac	ccctggccca	gccccccatc	cagcccaatcc	ccccaaacc	2160
gcgcgcaccc	ctgc	ggccgttca	aaggctcagg	agagtgtgg	tctcccccagg	2220
cgttaccagg	agggggcagg	ctcagccago	tggggaaacc	ttatttgc	gttccat	2280
agccacccct	tcccccggca	agacaggcc	caaggggat	tcccggggcc	ggttcccccgg	2340
agccccgtgg	gttccggcc	ttccacaccc	tctggggaca	gacacggccc	cttcttctca	2400
gttggggacac	tcataggac	aggatcagg	ggccctgggtt	aggctgggg	tccggccacag	2460
ccaaagaagct	cogagacc	cgatcgccc	agcccaatgc	ctccccccagg	cctaaagagga	2520
gaggggacca	ggggagaaaag	tctagacc	ctggccca	ccgcgtatgc	cagggcccc	2580
gcacagcccc	ccgcgtac	ggccgtctgg	cccgccggcc	cctcccttc	ccggccgtca	2640
cagccggat	cccaaggatcc	gaaacgcgc	agggtgc	ccggccggcc	ctcgagcc	2700
gagggtcagg	ccggccgc	ggggggaccc	caaggggagg	ccgggggggg	ggggggccgc	2760
cctggccccc	ccgcggggcc	gctc	aaaggcgggg	aggccgggg	gtctatgtgt	2820
tttctgtcaga	ggcccccgggg	ttgggggggt	gttggggggc	cccgga	agccctcccg	2880
gcctctgggc	ccgcacccggc	ggcagccctg	ccggccgggg	tggacctgg	cagttgcctg	2940
gacgtgtctgg	catttgc	ccgcacccgg	gagccggcc	ttggccgggg	gacccatcg	3000
ctgtatgacgc	acaacactgt	cgcgatgt	ggagccccc	cccttc	ccggccgt	3060

-continued

ggggccgacc	gcgagcgcac	cctcagccctg	cggaacggcc	ggggccgggc	gggtgtggc	3120
gtctctgtac	tgcggccgtc	ctaccagggg	ggccgcgtcag	ggctcccccag	gggcctcg	3180
ggcgaggcgc	cctctggcgc	ggcccccgtg	tcctgtcetc	tacctgcgca	cctgcatgtg	3240
ttaaaccccc	gggagaacac	ctggggggcc	ctgacccagg	tgcccgagg	ggcccccgtt	3300
cggggctgcg	gtctctgcac	catgcacaac	tacctgtttc	tggegggggg	catccgtggc	3360
tccgggtgoca	aggccgtctg	ctccaaacgag	gtcttctgtc	acaacccct	gaccaacatc	3420
tggagccagg	tteggcccat	gcagcaggcc	cgagccccag	tcaagctgt	ggccctggac	3480
gggctgtct	atgccccatgg	tggggatgc	ctgtacacga	tggagtgtca	cgaccggcga	3540
acagacgcct	ggacccacg	cgcgcactc	ccgcaggca	ccttccctgt	ggccacacag	3600
gtctgtggct	gcccgtgggg	catctacgtc	acccgggggtc	accttcttc	ccgcctgtc	3660
aggtacacgc	cctgtggaaa	tgttggggac	gagtggccat	acagtggccag	ccacccgggt	3720
tccagcgcaca	tcgtggca	ggggggctc	ctgttccgc	tcgacccgt	ggggggcatg	3780
ggcgccgcgc	tgatgcgcta	caacacagt	accggcttc	ggagcagg	tgccctccgt	3840
cccccgtccgc	ccccggcccc	actgcactgc	accacccctgg	gcaacacat	ttactgcctc	3900
aaccccccagg	tcaactgcac	cttcacggc	tctggggggg	ctggccagg	ccaggccaag	3960
gagctgcage	ccttcccc	ggggggacc	gggggtctca	gttcattcat	cctgactctg	4020
ccccctgagg	accgggtgca	gacccactc	tgagtggca	geagagaacc	aaagctgtt	4080
cgtgtcttc	caggagac	ctccgtggat	gggcctgtat	ggccgggggt	cagggaagg	4140
gctgggatcg	gaacttctg	ctcttggtt	tggacaactt	tcccccttcg	ctttaaaaggt	4200
tgtcgattat	tttgaagccc	agacttccct	agcccttttc	tgcccccac	tccacaccca	4260
gactgtttcc	tgactcaatt	ccgtacactac	ttacagacc	tcttagtgc	ctgacacccc	4320
cctgtctgt	ggactccct	ttcccttag	ccaggactgt	atgcgttcc	acagacaagg	4380
acttggcttc	cttggatct	gctggccgt	gagaggagg	ggttagaaaac	attcacactt	4440
cctatgtct	gtcagcagga	caggaggaa	aaacgtcccc	aggcaacggc	ctcgctctg	4500
ggactttctg	ctgttctaa	ggccctccca	ggtaccaacc	ccgttagctat	ctgggtctgt	4560
ttggcactgt	ggatttctaa	gggccttagaa	cccttgcctc	tgaaactgtt	ccgctgggtc	4620
agccctgtgt	tctgtgtc	ctgcacatac	ccccagccca	caccaggca	ggcccaactc	4680
gggcttcacca	ccctctgtcg	ccttgggg	ccctccccc	ggcccccacca	4740	
cttctcgoc	accccccgtc	tctaaatgag	gcctgagcgt	caccctagtt	ctggcccttt	4800
ttagctgtgt	agacttggac	gagacattt	acttccctt	tcccttgc	ataaaaatgt	4860
gacagtgac	gtctgtcacc	caagagat	gtggggagaca	agatcacagc	tatgacacc	4920
tccacccgt	tccaggatgc	caagacaaat	ccatgtatgc	ttttctccc	ttacgca	4980
tgaacccat	gttagaaaaag	tgaataacatc	tgactgtgc	ccactccaa	ctccagcctg	5040
gatgtccctg	tctggggcc	ttttctgtt	tttattctat	gttcagcacc	actggcacca	5100
aatacattt	aattccacca	aagca				5125

SEQ ID NO: 16                  moltype = DNA length = 2990  
 FEATURE                  Location/Qualifiers  
 source                  1..2990  
 mol\_type = other DNA  
 organism = Homo sapiens

SEQUENCE: 16						
ccacccatgc	cccaacccca	accccaagg	cattccctgc	cccaagctgac	gggtcaaaac	60
ctcaggagag	tgtggcttc	ccccaggcg	accaggagg	ggagggttc	ccagctgg	120
gaaacccctat	tgccatggtt	cttagaagcc	acccttccc	caggcaagac	aggccccaag	180
ggagtggtcc	gagggcggtt	cccgggagcc	cctgggttc	cagcaacttc	acacactctg	240
aggacagac	ccggcccttc	tcttcgtgg	ggacgttc	atggagacgtt	acagggggcc	300
tgttgtgggg	tggaggatcg	ccacacggaa	gaagctccga	gaccaacgg	tcgcccac	360
ccagccctcc	ccccaggctca	agaggagg	gaaccagg	gaaaagtcta	gaccggctc	420
cccaagegcgc	gtatgcac	ggcccccgcac	agccccccgc	gcagaggcc	cctggcccg	480
ccggcccttc	ctctgtcgagg	cgctcacac	cgttacccca	gtatcgaaa	cgccagcagg	540
gcgaaatcgc	cccgagatcg	gagcaggagg	tcaggccggc	ccctcggggg	gaccctcaag	600
gggaggccgg	ggggggagggg	ggcagccctg	ccggcccgac	cgggggcgctc	acggaaaac	660
aggaggaggc	ccggaaagctc	atgggtttc	tgcaaggagg	cggggggttg	gggggtgtgg	720
ggggggccccc	gaagccccc	tccggggccc	tggagccgc	cacggccgca	gccctgcggc	780
ggcggtcgat	ctctggatgg	tgcctggac	tgcgttgc	tgcctggat	caoggagagc	840
ccggccgtgc	gcaggaggacc	tacgtcgta	tgacgcacaa	cctgtcgca	gtgtgggg	900
accctgtct	ctaccgcgg	ctgagcgcgg	ccgaccgcga	gacgcaccc	agcctgcgga	960
ccggccgggg	ccggccgtgt	tgcggcgcc	tgcgtactgc	cagccttcac	cagggggg	1020
gctcagggtc	ccccaggccc	cctctgtgg	aggagcttc	tgccggggcc	cctgttccc	1080
tgcctctacc	tgcgcacatc	tatgttca	ccatgtgg	ccatgtgg	ccatgtgg	1140
cccaagggtcc	cgaggaggcc	ccgttccgg	gtcgccgtt	ctgcacat	cacaactacc	1200
tgtttctggc	ggggggccatc	ctgttgcctc	gtgccaagg	ctgttgc	aatgggtt	1260
tctgttacaa	ccctctgtacc	aaatctgtt	gccaagggtt	ccgcacatgc	caggcccg	1320
cccaaggctaa	gtctggggcc	ctggacgggg	tgcgttatgc	ccatgtgg	gaatgcctt	1380
acacgtatgc	gtgtacac	ccgcgcac	acgcctgtt	ccacgcgc	ccactcccc	1440
caggcacctt	ccctgtggcc	cacgaggctg	tggctgcgg	tggggacatc	tacgtcacc	1500
gggggtccatc	cttctaccgc	ctgttgcagg	acagccccgt	gaaggatgt	tgggacgat	1560
gcccatacg	tgccacccac	ccgcgttca	ccgcacatcg	ggcacttgg	ggcttccct	1620
accgcgttca	cctgtgttgc	ggcatgggg	ccgcgtgtat	ccatgtgg	gaatgcct	1680
gctccctgtgg	cagggtgtcc	tccctggccc	tgcctggccc	ccgcggccat	cactgcacca	1740
ccctggggca	eaccatttac	tgcctcaacc	cccaagggtac	tgccaccc	acggcttc	1800
ggggggactgc	ccagggtccag	gccaaggago	tgcagccctt	cccccttgggg	agcacccgggg	1860
tccctcgtcc	attcatctct	actctggccc	ctgaggacc	tcaactct	tgcgttgc	1920
tggcaggccag	agaaccaaag	ctgttgcgt	gtctgcacatc	tgggatggc	1980	
ctgagaggcc	gggggtccagg	gaaggggtcg	ggatcgaaac	ttctctgtt	tgtttctgg	2040

-continued

caactttccc	cttctgtctt	aaagggttgc	gattatttg	aageccagac	tccctcagcc	2100
tctttctgcc	cctcactcca	caccaggact	gttccctgac	tcaatccgt	acctaacttc	2160
agaccatctc	agcttgtgt	ccccccctcg	tctgtgggac	tcccttatcc	ctgagcccg	2220
ggactgtatc	gtctccacac	acaaggactt	ggctcgtctgg	atgttcgtgt	aggccagaga	2280
ggagggggta	gaaaacatc	acacttctca	tgctctgtca	gcaggacagg	gagoaaaaac	2340
gtccccaggc	aacgccttcg	cctctggac	tttctgcctg	tccataaggcc	tcccccaggta	2400
ccaaacccgt	agctatcttcg	tgctgttgg	cactgtggat	tctcaaggcc	ctagaacccct	2460
tgctctgtaa	actggcccg	tggtgtcagcc	ctgtgtctgt	cactgtctgc	ccatcccccc	2520
agcccacacc	aggccaggcc	cactccggc	tcaccaccc	ctgcagcctt	gtggggctct	2580
ccggacccct	ccagaacccc	acccccata	tcggccaccc	ccgtatctca	aatggagett	2640
gagcgtcacc	ctatgttcg	cccttttttag	ctgtgtagac	ttggacgaga	catttgactt	2700
cccttttccc	ttgtctataa	aatgtggaca	tgggacgtct	gtcaccaag	agagtgttgt	2760
gagacaagat	cacagctatg	agcacctcg	acgggtgtcca	ggatgcacag	cacaatccat	2820
gtatgcgttt	ctcccccattac	gcactttgtaa	accatcatgt	gaaaatgtaa	tacatctgac	2880
tgtgtccca	tccaaacccct	agccctgtat	tccctgtctg	ggcccttttt	ctgtttttta	2940
ttctatgttc	agcaccacat	gcaccaaata	cattttaaat	cacccaaagc		2990

```
SEQ ID NO: 17          moltype = DNA  length = 4837
FEATURE                Location/Qualifiers
source                 1..4837
                      mol_type = other DNA
                      organism = Homo sapiens
```

-continued

---

tgtcttatgc	catacggtggc	gaatgcctgt	acagcatgga	gtgtctacgac	ccgcgaacag	3240
acgcctggac	cccacgcgcg	ccactccccc	caggcacctt	ccctgtggcc	cacgaggctg	3300
tggectgcgg	tggggacatc	tacgtcaccg	ggggtcaccc	cttctaccgc	ctgtcagggt	3360
acagccccgt	gaaggatgtct	tggacagttt	gccccatcacg	tgccagccac	ccgcgttcca	3420
gegacatcg	ggcaactgggg	ggcttctgt	accgcttgcg	cctgctgccc	ggcatggcg	3480
ccggccgtat	gogctacaac	acagtgcacc	gtccctggag	cagggtgc	tccctgc	3540
tgccegeccc	cccccaactg	cactgcacca	ccctggggca	caccatcac	tgcctcaacc	3600
cccaaggtaac	tgcacatcc	acggctctgt	gggggactgc	ccagttccag	gccaaaggag	3660
tgcagccctt	ccccttgggg	agcacccggg	tcctcagtcc	attcatctcg	actctgc	3720
ctgaggacgc	gtgcacaggc	tcactctgt	ttgcaggcag	agaaccaaag	ctgtttcg	3780
getcttcagg	gagacccatcc	ttggatggcc	ctgagaggcc	ggggctcagg	gaaggggctg	3840
ggatcgaaac	ttccctgtct	tggttttgc	caacttccc	tttctgtt	aaagggttgc	3900
gattattttg	aagccagac	tcctctcage	tctttctgc	cctcaactca	cacccagact	3960
gttctctgc	tcaatttcgt	acctactac	agacccttc	agcttgcgt	cacccctcg	4020
tctgtggac	tccctatcc	ctagacccg	ggactgtatc	gttccacag	acaaggact	4080
ggctcgctgg	agctctgtcg	agccqagaga	ggagggggta	aaaaacatc	ttcttccta	4140
tgtctgtca	gcaggacagg	gagcaaaaac	gtccccaggc	aacgcctcg	cctctggac	4200
tttctgcgtt	tcctaaaggcc	tcccccaggta	ccaaacccctgt	agctatctgg	gtctgtttgg	4260
cactgtggat	tctcaaggcc	ctagaaccct	tgcctctgaa	actggtccgc	ttgtgcagcc	4320
ctgctgttcg	cagcttcgtc	ccatccccc	agccccacacc	aggccaggcc	cactccgggc	4380
tcaccaccc	ctgcagcc	gtggggctct	cccagccccc	ccagaagccc	accccaactc	4440
tcgccaaccc	ccgatctcta	aatggggct	gagcgtcacc	ctagtttc	cccttttttag	4500
ctgtgttagac	ttggacgaga	catttgcatt	ccctttctcc	ttgtctataa	aatgtggaca	4560
gtggacgtct	gtcacccaa	agatgttgc	gagacaagat	cacagctatg	agcacctcgc	4620
acggtgtcca	ggatgcacag	cacaatccat	gatgcgtttt	ctcccttac	gcactttgaa	4680
acccatgtca	aaaaagtgaa	tacatctgc	tgtgtccac	tccaaacctcc	agcttggat	4740
tcctgtctg	ggccctttt	ctgtttttt	ttctatgttc	agcaccactg	gcaccaaata	4800
caatttaattt	caccgaaagc	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	4837

---

SEQ ID NO: 18                    moltype = DNA   length = 2099  
 FEATURE                        Location/Qualifiers  
 source                        1 .. 2099  
 mol\_type = other DNA  
 organism = Homo sapiens

SEQUENCE: 18

ccacctcagc	ccccaaaaaa	accccagccg	catccccctgc	cccagctgac	gggtcaaaagc	60
ctcaggagag	tgtggctctc	cccagcgt	accaggaggg	gcaggtctca	gccagctgg	120
gaaacacctat	tgccatggtt	cttagaagcc	accccttccc	caggcaagac	aggcccaag	180
ggaggtgtccc	gaggggcggt	ccggggggcc	ccgtgggtcc	cagcaactcc	acacactctg	240
aggacagaca	ccggccctct	tcttcgtgg	ggacagtc	aggggacagg	acagggggcc	300
tgttgagggc	tggagggtcag	ccacagccaa	gaaagcttgc	gaccaaggaa	tcgcccagcc	360
cagaccctcc	cccaggctca	agaggagagg	gaaccaggga	aaaaagtcta	gaccgcgtc	420
cccaaggccg	gatgcggcagg	ggccccccgc	agccccccgc	gcagaggccg	cctggccccc	480
cggectetc	cttcgtcagg	cgcteacago	cggtacccca	getacggaaa	cgcagcagg	540
gcaaaatcgc	cccgagctcg	gaggcagg	tcaggccggc	cgccctgggg	gaccctcaag	600
ggggagggccc	gggggggggg	ggcagccctg	ccggccggag	cgggggcg	acggaaaagc	660
aggaggaggc	ccggaaatgc	atgggtttc	tgcaaggagg	cggggggttgg	gggggtgggtgg	720
aggggccccg	gaaaggccagc	tccggggccc	tgggagggcc	cacggggcc	gccttcgggc	780
ggggcgttgc	cttggggcgt	tgcttggacg	tgcttgcctt	tgcccagcc	cacggagagc	840
ccggcgttgc	gcaggagacc	tacgcgtca	tgagcgtca	cctgtgtcc	gtgctgggg	900
acccgtgtct	ctaccggccgg	ctgagcgtgg	ccgaccgcga	gcccacatc	agccgtccgg	960
ccggccgggg	ccggccgtgt	ctgggggttc	tgtactgc	cagecttcac	caggggggcc	1020
gtccagggtt	cccccaaggc	cctgtgtgg	aggagcttc	tgccggggcc	cctgtgtccc	1080
tgccttctacc	tgcgtaccc	catgtttca	accccccgg	gaacacctgg	cgccccctga	1140
cccaaggatgc	cgaggaggcc	ccgttctgg	gtgtgggtt	ctgcacat	cacaactacc	1200
tgtttcttgc	ggggggccat	ctgtggctcc	gtgccaagg	cgctgttcc	aacgggtt	1260
tctgtctaca	ccctctgtcc	aaatctgtt	gcccgttcc	gcccacat	caggcccgag	1320
cccaagctca	gtgtgtggcc	ctggacgggc	tgtctatgc	categggttgc	gaatgcctgt	1380
acacgtatgt	gtgtgtacgc	ccgcgtacag	aeccctggac	cccaacgcgc	ccactccccc	1440
caggcacatc	ccctgtggcc	cacggtgtt	ttggcttgc	ttggggacatc	taegtcaccg	1500
ggggtcaccc	cttctaccc	ctgtgtgtt	acagccccgt	aaaggatgtt	ttggacgagt	1560
gccccatcac	tgccagccac	ccgggttcc	gacacatcg	ggcactgggg	ggcttctgt	1620
accgcttcga	ctgtgtcgcc	ggcatgggg	ccggcgtgt	gctgttcc	acagtgtacc	1680
gtccctgggg	cagggtgtcc	tccctggccc	cgccccactg	cactgcacca	1740	
ccctggggca	caccatcc	tgcctcaacc	cccaagggtc	tgccacatc	acgggttctg	1800
ggggggactgc	ccagttccag	gccaaggagc	tgccgttcc	cccttgggg	agcaccgggg	1860
tcctcgttcc	attcatctcg	actctggccc	ctgaggaccc	gtgtcgtacc	tcactctgt	1920
tggcaggcag	agaaccaaag	ctgtttcg	gtcttccagg	gagaccctcc	ttgggtgggc	1980
ctgagaggcc	gggggtcagg	gaaggggctg	ggatcgaaac	ttccctgtct	tgtttcttgc	2040
caactttccc	tttctgtttt	aaagggttgc	tattttttt	aaaaaaa	aaaaaaa	2099

SEQ ID NO: 19                    moltype = AA   length = 1235  
 FEATURE                        Location/Qualifiers  
 source                        1 .. 1235  
 mol\_type = protein  
 organism = Homo sapiens

-continued

---

SEQUENCE: 19

MIGQTLEPDG	PLWGWDWDSD	NDWDASVLAL	LALAVVAATA	LALHWFGSGH	DQEAAEPVST	60
ALGAQPHQAG	GAEALALQPKS	KVSDGSEGQS	PGQGKPEPPG	RQQSQPVPA	APGGGLAAMA	120
RLPLKTAVEE	ARREALGQQR	GSATPAAPRA	EGKEPPRPGT	ALLGRSEAGG	MSAPLLIHFT	180
PRSPGSEAEA	ETGGVRASSR	QAAGPAGQD	TGPWQAGAGP	SGSMGRGRGR	RRRMDAGSGD	240
RARRPRKLDP	LRLGAAGSVW	DAVGAAALD	AHARGLPTGP	PLAQEPALPA	LPAPRALQPG	300
SQTEGSGAKG	GWSREASGVP	APGGGWPWVS	REVPGTRSFQ	PAPDSTRPWL	ESPPQGRPLS	360
SQGPGATGAY	DAGEAGADSS	RDNSPAADLG	PTRPPEQAKP	AAAGHSRAPS	RSREPRPRSA	420
SPPAAPGPGF	PPEALTLPS	SDFLPLEVTQ	DPSVGENLRA	APAPSSASAQ	VLT SAPASVL	480
APALASSPSS	APTSATTSTS	SPT SAPAPAP	TSAPTSTPAP	APSPAAAATP	APAPVVPVTL	540
TPPSPALTPV	PTPALSPAPT	PALTPAASP	LTPVPTPALS	PAPTPAPTPA	ASPA PAPTS	600
PTPTPAASPA	PADGSKPQES	VALPRRYQEC	QVSASWGNLI	AMVLRSHFPF	RQDRPQGSVP	660
RAVPGSPVGP	STSTHSEDRH	GPSSSVGTVI	GTGTGGLVEA	GGQPQPRSSE	TNGSPSPDPP	720
PGLRGEGTRE	KSLDPLPQAA	MPRGPAQPPA	QRPGPQAASS	SARRSQVPVQ	LRKRSRCEIA	780
PSSEQEVRP	ASGDPQGEAP	GE GGSPAGRS	GALTEKQEEA	RKLMVFLQRP	GGWGVVEGPR	840
KPSSRALEPA	TAALRRLLD	LGSCLDVLA	AQOHGEPCLA	QETYALMSDN	LLRVLGDPC	900
YRRLSAADRE	RILSLRTGRG	RAVLGVLVLP	SLYQGGRSGL	PRGPRGEPP	AAAPVSLPLP	960
AHLHVFNPRE	NTWRPLTQVP	EEAPLRCGCL	CTMHNYLFLA	GGIRGSGAKA	VCSNEVFCYN	1020
PLTNIWSQVR	PMQQARAQLK	LVALDGLLYA	IGGECLYSME	CYDPRTDAWT	PRAPLPGTF	1080
PVAHEAVACR	GDIYVTGGHL	FYRLRYSVP	KDAWDECPYS	ASHR RSSDIV	ALGGFLYRF	1140
LLRGVGAAVM	RYNTVTGSWS	RAASLPLPAP	APLHCTTLGN	TIYCLNPQVT	ATFTVSGGTA	1200
QFQAKELQPF	PLGSTGVLS	FILTLPPEDR	LQTS			1235

SEQ ID NO: 20            moltype = AA    length = 594

FEATURE                    Location/Qualifiers

source                    1..594

mol\_type = protein

organism = Homo sapiens

SEQUENCE: 20

MVLRSHPFPR	QDRPQGSVPR	AVPGSPVGPS	TSTHSEDRHG	PSSSVGTIVG	TGTGGLVEAG	60
GQPQPRSSSET	NGSPSPDPPP	GLRGEGTREK	SLDPLPQAAM	PRGPAQPPA	RPPGPAASSS	120
ARRSQPVQQL	RKRSRCEIA	SSEQFVRPA	SGDPQGEAPG	EGGSPAGRS	ALTEKQEAR	180
KLMVFLQRPG	GWGVVEGPRK	PSSRALEPAT	AAALRRRLD	GSCLDVLAFA	QOHGEPCLAQ	240
ETYALMSDNL	LRVLGDPCLY	RRLSAADRE	IILSLRTGRGR	AVLGVLVLP	LYQGGRSGLP	300
RGPRGEPPA	AAPVSLPLP	HLHVFNPREN	TWRPLTQVPE	EAPLRCGCLC	TMHNYLFLAG	360
GIRGSGAKAV	CSNEVPCYNP	LTNIWSQVRP	MQQARAQLK	VALDGLLYAI	GGECLYSMEC	420
YDPRTDAWT	RAPLPGTFF	VAHEAVACRG	DIYVTGGHLF	YRLLRYSVPK	DAWDECPYS	480
SHRSSDIVA	LLGGFLYRF	LRGVGAAMR	YNTVTGSWSR	AASLPLPAPA	PLHCTTLGN	540
IYCLNPQVTA	TFTVSGGTAQ	FQAKELQPF	LGSTGVLS	FILTLPPEDR	QTS	594

SEQ ID NO: 21            moltype = AA    length = 495

FEATURE                    Location/Qualifiers

source                    1..495

mol\_type = protein

organism = Homo sapiens

SEQUENCE: 21

MRGRGPAQPA	QRPPGPAASS	SARRSQVPQ	LRKRSRCEIA	PSSEQEVRP	ASGDPQGEAP	60
GEGGSPAGRS	GALTEKQEEA	RKLMVFLQRP	GGWGVVEGPR	KPSSRALEPA	TAALRRLD	120
LGSCLDVLA	AQOHGEPCLA	QETYALMSDN	LLRVLGDPC	YRRLSAADRE	RILSLRTGRG	180
RAVLGVLVLP	SLYQGGRSGL	PRGPRGEPP	AAAPVSLPLP	AHLHVFNPRE	NTWRPLTQVP	240
EEAPLRCGCL	CTMHNYLFLA	GGIRGSGAKA	VCSNEVFCYN	PLTNIWSQVR	PMQQARAQLK	300
LVALDGLLYA	IGGECLYSME	CYDPRTDAWT	PRAPLPGTFF	PVAHEAVACR	GDIYVTGGHL	360
FYRLRYSVP	KDAWDECPYS	ASHR RSSDIV	ALGGFLYRF	LLRGVGAAM	RYNTVTGSWS	420
RAASLPLPAP	APLHCTTLGN	TIYCLNPQVT	ATFTVSGGTA	QFQAKELQPF	PLGSTGVLS	480
FILTLPPEDR	LQTS					495

SEQ ID NO: 22            moltype = AA    length = 1235

FEATURE                    Location/Qualifiers

source                    1..1235

mol\_type = protein

organism = Homo sapiens

SEQUENCE: 22

MIGQTLEPDG	PLWGWDWDSD	NDWDASVLAL	LALAVVAATA	LALHWFGSGH	DQEAAEPVST	60
ALGAQPHQAG	GAEALALQPKS	KVSDGSEGQS	PGQGKPEPPG	RQQSQPVPA	APGGGLAAMA	120
RLPLKTAVEE	ARREALGQQR	GSATPAAPRA	EGKEPPRPGT	ALLGRSEAGG	MSAPLLIHFT	180
PRSPGSEAEA	ETGGVRASSR	QAAGPAGQD	TGPWQAGAGP	SGSMGRGRGR	RRRMDAGSGD	240
RARRPRKLDP	LRLGAAGSVW	DAVGAAALD	AHARGLPTGP	PLAQEPALPA	LPAPRALQPG	300
SQTEGSGAKG	GWSREASGVP	APGGGWPWVS	REVPGTRSFQ	PAPDSTRPWL	ESPPQGRPLS	360
SQGPGATGAY	DAGEAGADSS	RDNSPAADLG	PTRPPEQAKP	AAAGHSRAPS	RSREPRPRSA	420
SPPAAPGPGF	PPEALTLPS	SDFLPLEVTQ	DPSVGENLRA	APAPSSASAQ	VLT SAPASVL	480
APALASSPSS	APTSATTSTS	SPT SAPAPAP	TSAPTSTPAP	APSPAAAATP	APAPVVPVTL	540
TPPSPALTPV	PTPALSPAPT	PALTPAASP	LTPVPTPALS	PAPTPAPTPA	ASPA PAPTS	600
PTPTPAASPA	PADGSKPQES	VALPRRYQEC	QVSASWGNLI	AMVLRSHFPF	RQDRPQGSVP	660
RAVPGSPVGP	STSTHSEDRH	GPSSSVGTVI	GTGTGGLVEA	GGQPQPRSSE	TNGSPSPDPP	720
PGLRGEGTRE	KSLDPLPQAA	MPRGPAQPPA	QRPGPQAASS	SARRSQVPVQ	LRKRSRCEIA	780
PSSEQEVRP	ASGDPQGEAP	GE GGSPAGRS	GALTEKQEEA	RKLMVFLQRP	GGWGVVEGPR	840

-continued

---

KPSSRALEPA TAAALRRRLD LGSCLDVLAQ AQQHGEPLA QETYALMSDN LLRVLGDPCL	900
YRRLSAADRE RILSLRTGRF RAVLGVLVLP SLYQGGRSGL PRGPRGEPP AAAPVSLPLP	960
AHLHVFNPRE NTWRPLTQVP EEAPLRGCGL CTMHNYLFIA GGIRGSGAKA VCSNEVFCYN	1020
PLTNIWSQVR PMQQARAQLK LVALDGLLYA IGGECLYSME CYDPRTDWAT PRALPLAGTF	1080
PVAHEAVACR GDIYVTGGHL FYRLLRYSPV KDAWDECPSY ASHRRSSDIV ALGGFLYRFD	1140
LIRGMGAAVM RYNTVTGSWS RAASLPLPAP APLHCTTLGN TIYCLNPQVT ATFTVSGGTA	1200
QFQAKELQPF PLGSTGVLSP FILTLPEDR LQTS	1235

SEQ ID NO: 23	moltype = AA length = 594
FEATURE	Location/Qualifiers
source	1..594
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 23	
MVLRSHPFPR QDRPQGSVPR AVPGSPVGPS TSTHSEDRHG PSSSVGTIVG TGTGGLVEAG	60
GQPQPRSSSET NGSPSPDPPP GLRGEGTREK SLGPLQAM PRGPAOPPAQ RPPGPAASSS	120
ARRSQPVQPL RKRSRCEIAP SSEQEVRPAA SGDPOGEAPG EGGSPAGRSG ALTEKQEEAR	180
KLMVFLQRPG CGWVVEGPRI PSSRALEPAT AAALRRRLD GSCLDVLAFA QOHGEPLAQ	240
ETYALMSDNL ILSLRTGRGR AVLGVLVLPs LYQGGRSGLP	300
RGRGEPEPPA AAPVSLPLPA HLHVNPNPRE TWRLPTQVPE EAPLRGCGLC TMHNYLFAG	360
GIRGSGAKAV CSNEVFCYNP LTNIWSQVRP MQQARAQLKL VALDGLLYAI GGECLYSMEC	420
YDPRTDWATP RAPLPGATFP VAHEAVACRG DIVVTGGHLF YRLLRYSPVK DAWDECPSA	480
SHR RSSDIVA LGGFLYRFDL LRGMGAAVM RYNTVTGSWSR AASLPLPAPA PLHCTTLGN	540
TYCLNPQVTA TFTVSGGTAQ FQAKELQPF PLGSTGVLSP FILTLPEDR LQTS	594

SEQ ID NO: 24	moltype = AA length = 495
FEATURE	Location/Qualifiers
source	1..495
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 24	
MPRGPAOPA QRPPGPAASS SARRSQPVQ LRKRSRCEIA PSSEQEVRPA ASGDPQGEAP	60
GEGGSPAGRS GALTEKQEEA RKLMLVFLQRP GGWGVVEGPRI KPSSRALEPA TAAALRRRLD	120
LGSCLDVLAQ AQQHGEPLA QTYALMSDN LLRVLGDPCL YRRLSAADRE RILSLRTGRG	180
RAVLGVLVLP SLYQGGRSGL PRGPRGEPEPP AAAPVSLPLP AHLHVNPNPRE NTWRPLTQVP	240
EEAPLRGCGL CTMHNYLFIA GGIRGSGAKA VCSNEVFCYNP PLTNIWSQVR PMQQARAQLK	300
LVALDGLLYA IGGECLYSME CYDPRTDWAT PRALPLAGTF PVAHEAVACR GDIYVTGGHL	360
FYRLLRYSPV KDAWDECPSY ASHRRSSDIV ALGGFLYRFDL LRGMGAAVM RYNTVTGSWS	420
RAASLPLPAP APLHCTTLGN TIYCLNPQVT ATFTVSGGTA QFQAKELQPF PLGSTGVLSP	480
FILTLPEDR LQTS	495

SEQ ID NO: 25	moltype = DNA length = 5124
FEATURE	Location/Qualifiers
source	1..5124
	mol_type = genomic DNA
	organism = Homo sapiens
SEQUENCE: 25	
agaaggcagg tggctgtat gacagcacag agtcaggaa gcgtgcctga ggaccctggg	60
gcttacgagg aggagaagag ggcaggagct ggtgggggtc ttgcagagac cctgggtctcc	120
tatccctgc taagcctcg tgcgtctgtatctgc caggcctac tgacaccccc	180
aggcctgat gcaaggcagg accccacat tcccaggccc tggaggactg gtccaccccta	240
actggcgcgc cttggggca ggcgtgtggc ggtgcctcag cccaggcctc tgcgtctgc	300
atgcactgc accctgcatt caggcctata ttgcagccct gaaccatgtat ccaggcacc	360
ttggagccag atggccctct ctggggctgg gactgggaca gtgacaatga ctggatagt	420
gtctgtctgg ccctctggc gtcgtctgtg gtggctgcca cagcgtggc cttacactgg	480
tttggctccg ggcacgtatca agaggcggca gaaccgggtt ccacaggctt cggggctcaa	540
cctccatcagg caggaggagc tgatgcctc ctgcacccga agtctaaaggc tgcgtatggc	600
agcggaggccc agaggccagg gcaggggaaa ccagagcccc caggacggg ccaggcagac	660
cctgtccctg ctgcacgcgg gggcgggccca tggcccggtt tccactcaag	720
acggcgtcg aggaggccc cagagaggca ttagagacgc aacggggcag tggccacccc	780
ggggcccccc gageggaggc aaaggaggct cccaggccgg cgaactccctt cttggccagg	840
agcgaagcag gggggatgtc cgcggccctc ctgatccact tcactccctg gagccctggc	900
agcgaagcgg aggccggagac aggtggtgctc agggcgtctc ctgcggcaggc cgcaggcccc	960
ggggggcaac aggacactgg cccctggccg gggggccgggg ggcctccggg ctgcgtgggg	1020
agaggccggg gccggccggc gggatggac gctggctggc gagacagacgc cccggccccc	1080
cggaaacttgg acccgctccg cctggggccca gggggaggcc tggggacgc ggtggacccgg	1140
ggccggccccc tggacgcccc cgcggcgccg ctcccccacag gaccccccact cgcccaaggag	1200
cccgcaactcc cggcgctgca cgcggccctc gcccctgcac ctgggtctca gacggaaaggc	1260
tctggggcaca agggtggtcg gaggcaggagc gcctcgccccc cggaggaggc	1320
tggcccttggg tcagcaggaa ggtcccgccg acccgagact tggccccagg cccagactcc	1380
acgcgcgcctt ggctagagat tccgcctcaaa ggtggccccc ttcgtccca agggccgggt	1440
gcacacagggg cttacgatgc cggcgaggcc ggggtgacca gtcggcgaga taacagtcct	1500
ggcgctgacc tggggccac cggggccccg gagcaagcaa agccggctgc agccggccac	1560
agccgcgcgc cttccggagc cctgtgagct cggccctcccc gcccggcagct	1620
cccgccccccg gtttccacc tgaaggccctg actctccctt ctccttcaga cttttgc	1680
ctggagggtta cccaggatcc ttccgtggc gaaaatctca gagccggccccc agcccaagt	1740

-continued

tcagcctcg cccaagtctt	aacttcagct ccagcctcg	tccttagcccc agccctggct	1800
tcatcccca gtcagcacc	aacccatcgcc accacacctca	cctcatcccc cacccatcgcc	1860
ccagccccag ctccaaaccc	agctccaaact tcaaccccaag	cccagcccc aagtccagct	1920
gcagccgcaa ctccagcccc	agccccagtc ccagtcctaa	ccctcacacc cccatcccc	1980
gccctaacc cagtcctaaac	cccagccata agccagctc	caactccagc cctaaccctta	2040
geogcatccc cagccctaaac	cccagtcctaa accccagcccc	taagccccagc tccaaactcca	2100
gcccccaaccc cagccgcatc	ccctgccccca gccccccaccc	cagecccaac cccaaacccca	2160
geogcatccc ctgccccagc	tgacgggtca aagctcagg	agagttgtggc ttctcccaagg	2220
cgctaccagg aggggcagggt	ctcagccago tggggaaacc	ttattgcctt ggttctttaga	2280
agccacccct tccccaggca	agacaggccc caaggaggat	tcccgaggcc ggttcccccgg	2340
agccctgtgg tcccgagcac	tccacacac tttggggaca	gacacggccc cttttttca	2400
gtggggacag tcatagggac	aggtacagggg ggcttgggtt	aggctggagg tcagccacag	2460
ccaagaagct ccgagaccaa	cggtatcgccc agccagac	ctcccccaagg ctaagagga	2520
gagggaaacca gggggaaaaag	tctacatcccc ctgcccccaag	ccgcgtatggc cagggggcccc	2580
gcacagcccc ccgcgcatgc	gcccgttgc cccgcggcc	cctctctgc gaggegetca	2640
cacccgtatc cccagctac	gaaaacggcgc aggtgcgaaa	tgccccccgg ctcggagcag	2700
gagggtcaggc ggccgcgtcc	gggggaccct caagggggagg	cgccgggggg gggggggcagc	2760
cctgcggccc gcaagggggc	gctcacggaa aagcaggagg	aggccccaaag ctcatgggt	2820
ttctgcagag cccgggggtt	tgggggggtt tgaggggggc	ccggaggccg agctccccgg	2880
ccttggggcc cggccacccgg	gcagccctgc ggcggggcgt	ggacactggc agttgcctgg	2940
acgtgtggc ctttgcctag	cagcacggag agccggccct	ggcgccggag acctacgcgc	3000
tgtatgagoga caacctgtg	cgatgtgtgg gagacccgt	cctctaccgc cggctgagcg	3060
cggccgcgtcc cgagccgtatc	ctcagccgtc ggacccggcc	ggggggggcc gtgttggggc	3120
tcctctgtact gcccaccc	tccagggggg gcccgttgc	gtccccccagg ggccctctgt	3180
gcgaggagcc tcctgtggcg	gcccctgtgt ccctgcctt	acctgegcac ctgcattgtgt	3240
taaacccccc ggagaaaccc	tggggggccc tgacccagg	gcccggagg gcccccttc	3300
ggggctgggg tttctgcacc	atgcacaactt acctgtttt	ggggggggggc atccgtggct	3360
ccgggtccaa ggccgtctgc	tccacgggg tttctctgtca	aaaccccttc accaaacatct	3420
ggagccagggt tcggccatcg	cagcaggccc gagcccaagct	caagctgggt gcccctggacg	3480
gggtgtctta tgccatcggt	ggcgaatggc tgcacagcat	ggagtgtctac gaccccgccaa	3540
cacacgcgtt gacccaccc	ggcccaactcc cccgcaggac	cttccctgtg gcccacgggg	3600
ctgtggctt cgggtggggc	attcacgtca cgggggggtca	cctcttcac cgcctgtca	3660
gttacagccg cgtgaaggat	gcttggggacg agtgccttata	cagtgcggcc caccggcggt	3720
ccagcgacat cgtggactt	gggggcttcc tgcacccgtt	cgacccgtc cggggcggt	3780
ggccgcgggt gatgcgttac	aacacaggta ccggcttccgt	gaggccgggt gcctccctgc	3840
ccctggccgc ccccgcccc	ctgcacgtca ccacccctgg	caacacatt tactgcctca	3900
accccaagggt cactgcacc	ttcacgggtca ctggggggac	tgcccaagttt caggccaaagg	3960
agctgcgtcc ctccccctt	ggggggcccg gggcctctcg	ccatttcac ctgactctgc	4020
ccccctgggg cccggctgtc	ggccgttactt gatgtggcagg	catagaaacca aagctgtttc	4080
gctgtcttcc agggagaccc	tccttggatg ggcgttggag	ggccggggcc agggaaagggg	4140
cttggatcttgc aacttcttc	tctttttttt ggacaactttt	cccccttcgc tttaaagggtt	4200
gtcgattatt ttgaagccca	gactccctca gcctttttt	gcctcttact ccacacccag	4260
actgttcttgc gactcaatc	cgatctactt tacagaccc	ctcagcttgc tgacacccccc	4320
ctgttctgtgg gactccat	tcctctatggc cggggactga	tgcgttccca cagacaaggaa	4380
cttggctctgc tggatctgt	ctgagccagc agaggagggg	gtagaaaaca ttcacactt	4440
ctatgtctgc tgcacggac	agggggccaa aacgtcccc	ggcaacgcgc tgcctctgg	4500
gacttcttcg ctgttcataa	ggccctccca gtcaccaacc	cgatgttatac tgggtctgtt	4560
tggcaacttgc gatgttcag	ggccatggaaat ctttgcctt	gaaatggtc cgctggtca	4620
ggccctgtgt ctgcgttcc	tgcctatacc cccagccccc	accaggccag gcccactccg	4680
ggtcacccac cctctgcgtc	cttggggggc tctcccgac	cctccagaag cccacccac	4740
ttctcgccaa ccccccgtat	cttaatgggg cctggatgttc	accctgttgc tgccctttt	4800
tagctgtta gacttggacg	agacatgttgc ttccctttt	tccttgcata taaaatgtgg	4860
acagtggacat ctgttcaccc	aaagagatgtt tggggagacaa	gatcacaatgt atqacaccc	4920
cgcacggtgt ccaggatgca	cagcacaatc catgtgtgt	tttctccctt ttcgcacttt	4980
gaaaacccatg ctagaaaaatg	gaatcatct gactgtgttc	cactccaaacc tccagccctgg	5040
atgttcttgt ctggccctt	tttctgtttt ttatctatg	ttcagccacca ctggcaccaa	5100
atacatttta attcaccgaa	agca		5124

SEQ ID NO: 26 moltype = DNA length = 5124

FEATURE Location/Qualifiers

source 1..5124

mol\_type = genomic DNA

organism = Homo sapiens

SEQUENCE: 26

agaaggcagggt tggctgtat	gacagcacag agctcaggaa	cgctgcctga ggaccctggg	60
gcctacaggagg aggagaagag	ggcaggagct ggtggggtgc	ttgcagagac cctggggctcc	120
tatcctgcata taagcctcgc	tgtctctgtatctgc	caggccctac tgacaccccc	180
aggcctgaggat gcaaggcagag	accccaccaat tcccaggcccc	tggaggactg gtccacccctta	240
actggggcaggc ctttggggca	ggcgttggcc ggtgcctcag	cccaggccctc tgcgtctgc	300
atgcactgtcc agccctggccat	caggccctata ttgcagccccc	gaaccatgt ccaggccacc	360
ttggagccagg atggccccct	ctggggctgg gactgggaca	gtgacaatgt ctggatgtt	420
gctgtgttgg ccctcttggc	gctggctgtg gtggcttca	cagcgttggc cttacactgg	480
tttggctccgc ggcacgtatca	agaggccgtt ccacagccct	cggggctcaa	540
cctcatcagg caggaggaggc	tgagctggcc ctgcaaccga	agtctaagggt cagtgttggc	600
agcgaggggc agagcccaagg	gcagggggaa ccagaccc	caggacgggg ccacgcagagc	660
cctgtccctgc ctgcagcgcc	ggggggggcc ctggccggca	tggcccggtc tccactcaag	720

-continued

acggctgtcg	aggaggccc	cagagaggca	ttaggacagc	aacggggcag	tgccacccc	780
ggggcccccc	gagcggaaagg	aaaggagcc	cccaggccag	gaactgcct	cctgggcagg	840
agcgaagcag	ggggatgtc	cgccccctc	ctgatccact	tcactctcg	gagccctggc	900
agcgaagcgg	aggcgagac	aggtgtgtc	aggcgctct	ctcgccagc	cgcaggcccc	960
gcggggcaac	aggacactgg	cccctggcag	gcggggcgccg	ggccctcggg	ctcgatgggg	1020
agaggccgg	gcggggcgccg	gcggatggac	gttggctcg	gagacagagc	ccgcggcccc	1080
cgaaaaactgg	accgcgtccg	cetgggegc	cgggggagcg	tgtggacgc	ggttggacgg	1140
gecgccgccc	ttggacgcccc	cgegcggcc	ctccccacag	gaccggccact	cgcggaggag	1200
cccgcaactcc	ggcgctgcc	cgctccccgc	gcctgcage	ctgggtctca	gacggaaaggc	1260
tctggggcoca	agggtgtgg	gaggcggag	gcctggggg	tccctgcucc	cgggaggagc	1320
tggccctggg	tcagcgggg	gttcccggg	acccggagat	ttggccca	cccgactcc	1380
acgcgcgcct	ggctagagag	tccgcctcaa	gggtgc	ccac	ttccgtccca	1440
gecacagggg	cttacgatgc	cgcgagggc	ggggctgaca	gtccccgaga	taacagtct	1500
gecgctgacc	tggggccccc	ccggggccccc	gaggcaagca	agccggctgc	agccggccac	1560
agecgcgccc	cctccggag	cggtgagcc	cgcgcgcgtc	cgccccccccc	gcccgcagct	1620
cccgcccccgg	gttcccccac	tgaaggccctg	acttccctc	ctccctcaga	ctttttgcc	1680
ctggaggtta	cccaggatcc	tccgtggc	gaaaatcta	gagcggcgcc	agccccca	1740
tcagcctcag	cccaagtctt	aacttcgt	ccagcctcag	tcctagcccc	agccctgtgt	1800
tcatccatcc	gtcagcacc	aacctcage	accacccat	cctcatcccc	cacccatc	1860
ccagcccccc	ctccaaaccc	agtcacta	taacccccc	ccccagcccc	aagtccat	1920
gcagccgca	ctccagcccc	agccccagtc	ccagtc	ccctc	acaccatcccc	1980
gcocctaacc	cagtc	cccaagccca	agccccagtc	caactcc	ccatacccca	2040
gcgcgcatacc	cagccctaa	cccagtc	accccgcc	taagcc	ccatccactcca	2100
gcccccaacc	cagccgcate	ccctggccca	gccccccac	cagccca	cccccaacc	2160
gcgcgcatacc	ctgccccage	tgacgggtca	aagcctcagg	agagtgtgg	tctcccccagg	2220
cgttaccagg	aggggcagg	ctcagccag	tggggaaacc	ttattgc	catggttttaga	2280
agccacccct	tcggccagg	agacggcc	caaggagat	tcccgagg	gcgttcccccgg	2340
agcccccgtgg	gtcccagcac	ttccacac	tctgaggaca	gacacggcc	ctcttctca	2400
gtggggacag	tcataggac	aggtacaggg	ggctctgg	aggctggagg	tcagcacag	2460
ccaaagaacct	cggagacc	cggtacgccc	agcccagacc	ctccccccagg	cctaagagga	2520
gaggggaaacc	ggggaa	tctagacc	ctggccca	cgcgatggc	caggggcccc	2580
gacagcccc	cggcgag	ggccctgtgg	cccgccgg	cttccttc	gaggcgctca	2640
cagccgtatc	cccagtc	gaaacgc	agg	tcgcccc	ctcgagg	2700
gagggtcaggc	ggggccctc	gggggac	caagggagg	cgccgggg	ggggggc	2760
cctggccggc	gaagccgg	gtcacgg	aa	gaggccggaa	gtctatgt	2820
tttctgcaga	ggccgggggg	ttgggggtgt	tgggagg	cccgga	agccccc	2880
gcctggagc	ccggccac	ggcagcc	ccggccgg	tgac	ctggc	2940
gacgtgtgg	ccttgc	gcagcac	gagccggc	tg	ggc	3000
ctgtatgc	acaac	ccgtgtgt	ggagaccc	gc	cttc	3060
gcccggccat	ccgtgtgt	cc	ggagccgg	cc	ccctgg	3120
ggatccat	tcggccat	cagcagg	gagcc	cc	ccctcg	3180
ggatccat	tcggccat	ggc	aat	cc	ccctcg	3240
ccatccat	tcggccat	ggc	aat	cc	ccctcg	3300
ccatccat	tcggccat	atc	ea	cc	ccctcg	3360
ccatccat	tcggccat	atc	ea	cc	ccctcg	3420
ccatccat	tcggccat	atc	ea	cc	ccctcg	3480
ccatccat	tcggccat	atc	ea	cc	ccctcg	3540
ccatccat	tcggccat	atc	ea	cc	ccctcg	3600
ccatccat	tcggccat	atc	ea	cc	ccctcg	3660
ccatccat	tcggccat	atc	ea	cc	ccctcg	3720
ccatccat	tcggccat	atc	ea	cc	ccctcg	3780
ccatccat	tcggccat	atc	ea	cc	ccctcg	3840
ccatccat	tcggccat	atc	ea	cc	ccctcg	3900
ccatccat	tcggccat	atc	ea	cc	ccctcg	3960
ccatccat	tcggccat	atc	ea	cc	ccctcg	4020
ccatccat	tcggccat	atc	ea	cc	ccctcg	4080
ccatccat	tcggccat	atc	ea	cc	ccctcg	4140
ccatccat	tcggccat	atc	ea	cc	ccctcg	4200
ccatccat	tcggccat	atc	ea	cc	ccctcg	4260
ccatccat	tcggccat	atc	ea	cc	ccctcg	4320
ccatccat	tcggccat	atc	ea	cc	ccctcg	4380
ccatccat	tcggccat	atc	ea	cc	ccctcg	4440
ccatccat	tcggccat	atc	ea	cc	ccctcg	4500
ccatccat	tcggccat	atc	ea	cc	ccctcg	4560
ccatccat	tcggccat	atc	ea	cc	ccctcg	4620
ccatccat	tcggccat	atc	ea	cc	ccctcg	4680
ccatccat	tcggccat	atc	ea	cc	ccctcg	4740
ccatccat	tcggccat	atc	ea	cc	ccctcg	4800
ccatccat	tcggccat	atc	ea	cc	ccctcg	4860
ccatccat	tcggccat	atc	ea	cc	ccctcg	4920
ccatccat	tcggccat	atc	ea	cc	ccctcg	4980
ccatccat	tcggccat	atc	ea	cc	ccctcg	5040
ccatccat	tcggccat	atc	ea	cc	ccctcg	5100
ccatccat	tcggccat	atc	ea	cc	ccctcg	5124

-continued

-continued

---

actgtttct	gactcaattc	cgtacactact	taacagacccct	ctcagettgc	tgacaccccc	4320
ctgtctgtgg	gactctctat	tcccttagago	caggactga	tgcgttcca	cagacaagga	4380
cttggctcgc	ttggagctctg	ctgagcccgag	agaggagggg	gtagaaaaaca	ttcacacettc	4440
ctatgtctg	ttagcaggac	agggagcaaa	aacgtcccc	ggcaacgccc	tcgcctctgg	4500
gactttctg	ctgtcttaag	gectccccag	gtaccaaacc	cgtagttatac	tgggtctgtt	4560
tggcacttgt	gatttctcaag	ggcttagaaad	ccttgctctt	gaaactggtc	cgctggtgca	4620
gcctctgtgt	ctgcagctcc	tgccccatacc	cccaagccccc	accaggccag	gcccactccg	4680
ggtcacccac	ctctgtcage	cttggggggc	tctcccaagcc	cctccagaag	cccacccac	4740
ttctegeccaa	cccccgatct	cttaatgagg	cctgaggttc	accctagttc	tgccttcttt	4800
tagctgtta	gactttggacg	agacatttg	cttcccttgc	ttaaatgtgg	4860	
acagtggacg	tctgttgcaccc	aagagatgtt	tgggagacaa	gatcacatgt	atgagcacct	4920
cgcacgggt	ccaggatgca	cggcacaatc	catgtgcgt	tttctccctt	taegcacttt	4980
gaaaacccatg	ctagaaaagt	gaatacatct	gactgtgctc	cactccaacc	tccagctgg	5040
atgtccctgt	ctggggccctt	tttctgtttt	ttattctatg	ttcagcacca	ctggcaccaa	5100
atacatttta	attcaccgaa	agca				5124

SEQ ID NO: 28	moltype = RNA	length = 2989				
FEATURE	Location/Qualifiers					
source	1..2989					
	mol_type = mRNA					
	organism = Homo sapiens					
SEQUENCE: 28						
ccacacctc	cccaaaaa	accccaagccg	catccccctgc	cccagctgac	gggtcaaa	60
ctcaggagag	tgtggctc	cccaggcgt	accaggagg	gcagggttca	gccagctgg	120
gaaaccttat	tgccatggtt	cttagaaged	accccttccc	caggcaagac	agggcccaag	180
ggagtgtccc	gaggggcggtt	cccgggagcc	ccgtgggtcc	cagcacttcc	acacactctg	240
aggacagacaa	ccggccctct	tcttcgttg	ggacagtcat	aggggacagg	acagggggcc	300
tgggttggac	tggagggttc	ccacagccaa	gaagctccga	gaccaacgga	tcgcccagcc	360
caagaccctcc	cccaggccta	agaggaggagg	gaaccaggga	aaaaagtcta	gaccggctgc	420
ccaaaggccgc	gtatggccagg	ggcccccgc	agccccccgc	gcagaggccg	cctggcccg	480
ccgccttc	ctctgtggac	cgctcacatc	cggtacggaa	ctacggaaa	cgccgcagg	540
gcgaaatccg	cccgagtcg	gaggaggagg	tcaggccgc	cgccctcggg	gacccctaag	600
gggaggggccc	gggggggggg	ggcagccctg	ccggccgcag	cggggggc	acggaaa	660
aggaggaggc	ccgaagctca	tgggtttct	gcagaggccc	gggggttgg	gggttgttgg	720
ggggggccccc	aaggccagat	cccgggccct	ggagccccc	acggccgcag	cccttgccgc	780
gcccgttggac	ctggggcgtt	gcctggacgt	gtggccctt	gcccagcagc	acggagagcc	840
ccggccttggc	cgaggacac	acgcgtgtat	gagcgcacaa	ctgctgogag	tgctgggaga	900
cccggttgc	taccggccgc	tgagcgcggc	cgaccgcgag	cgcatctca	gcctgcggac	960
ccggccggggc	ccggccgtgc	tggggcttgc	cgtaactggcc	agccttacc	agggggggcc	1020
ctcagggttc	cccaaggggcc	ctcgtggcga	ggagcctctt	ggggggccc	ctgtgtccct	1080
gcctctactt	gogcacactc	atgtttca	ccccggggaa	aacacctggc	ggccctgtac	1140
ccagggttgc	gaggaggccc	cgcttcgggg	ctgggttctc	tgcaccatgc	acaactacat	1200
gtttctggc	ggggggatcc	gttggatccgg	tgccaa	agggttgc	gtcggatcc	1260
ctgtetacaac	ctctgttgc	acatctgg	ccagggttccg	ccatgttgc	aggccgcagc	1320
ccagctcaag	ctgggtggcc	tggacgggt	gtctatgc	atcggtggc	aatgcctgt	1380
cagcatggag	tgtctacgacc	cgcgaacaga	cgcttggacc	ccacgcgcgc	cactccccgc	1440
aggcacttc	ctgtgtggcc	acgggtgt	ggcgttgcgt	ggggacatct	acgtcacccg	1500
gggttacactc	ttcttacggcc	tgtctacggta	caggccctgt	aaggatgtt	gggacgttg	1560
ccatatacg	gcacgcaccc	ggcgttccag	cgacatcg	cgactgggg	gttctctgt	1620
ccgcgttgcac	ctgtgtggcc	ggtggggc	cgccgtatgt	cgctacaaca	cagtgcaccc	1680
ctctctggac	agggtgttgc	ccctggccct	gcccccccc	ccccactgc	actgcacac	1740
cctggggcaac	accattttact	gctcaaccc	ccagggttact	gcccattca	cggttctcg	1800
ggggactgc	caggccagg	ccaaaggat	qacgccttc	cccttggga	gcacgggggt	1860
cetcagttca	ttcatectga	ctctggcccc	tgaggaccgg	ctgcagact	cactctgagt	1920
ggcaggcaga	gaaccaa	tggtctcg	ctctccagg	agacccctct	gggatggcc	1980
tgagggccgt	gggttcagg	aagggggtgg	gtatggaa	acttctgtt	gtttctggac	2040
aactttcccc	ttctgttta	agggtgtcg	atttttga	acggcaact	ccctcaggct	2100
ctttctgtcc	ctcaactccac	accaggact	tttctgtact	caatccgta	cctacttaca	2160
gacccttca	gtttgttgc	accccttctgt	ctgtggact	ccctattccc	tagagccagg	2220
gactgtgtcg	tctccacaga	caaggact	gtctgtcg	gtctgtcg	gcccggagag	2280
gagggggttag	aaaacatca	cacttccat	gtctgtcg	caggacagg	agcaaaaacg	2340
tccccaggca	acgccttcgc	ctctggact	ttctgtct	cctaaggct	ccccaggat	2400
caacccccgt	gtatcttgg	tctgtttgg	actgtggatt	ctcaaggccc	tagaacctt	2460
gcctctgttca	ttgggtcgct	gggtcagccc	tgttgtctgc	agctcttgc	cataccccc	2520
gcccacacca	ggccaggccc	actccgggtt	caccaccctt	tgccgttgc	ttgggtcttc	2580
ccagcccttc	cagaagccca	ccccacttc	cgccaacccc	cgatcttta	atgaggct	2640
aggtgttaccc	tagtttgc	ccttttttgc	tgtgttagact	tggacgagac	atttgact	2700
cctttctctt	tgtctataaa	atgtggacag	tggacgtctg	tcacccaaga	gagttgtgg	2760
agacaagatc	acagctatga	gcacccgc	cggtgtccag	gtatgcacagc	acaatccat	2820
atgcgttttc	tccccttacg	cactttgaaa	cccatgtctag	aaaagtgaat	acatctgact	2880
gtgtcttact	ccaaacccatca	gcctggatgt	ccctgtctgg	gcccctttc	tgtttttat	2940
tctatgttca	gcaccaactgg	caccaatac	attttaattc	accgaaagc		2989

SEQ ID NO: 29	moltype = RNA	length = 4836
FEATURE	Location/Qualifiers	
source	1..4836	

-continued

mol\_type = mRNA  
 organism = Homo sapiens  
 SEQUENCE: 29  
 actggccagc tgccatcagg cctctattgc agccctgaac catgatccag ggacacccagg 60  
 agccagatgg tccccctctgg ggctggact gggacagtga caatgactgg gatagtgtgg 120  
 tgcgtggccct cctggcgctg gctgtgggtt ctgcacacagc gctggcccta caactgggtt 180  
 gtcggggggca cgataaaggag gcccggaaac cgggtttccac agccctccgg gctcaaccct 240  
 atcaggcagg aggacgtcg atggcccttc aaccgaagtc taaggctcgt gatggcagcg 300  
 aggggcagag cccaggccgg gggaaacccgg agccccccagg acgcggccag cagacccctg 360  
 tccctgtgc agcgccgggc gggggccctgg ccgcattggc cgggttcca ctaaagacgg 420  
 ctgtcgagga ggcggccaga gaggcattag gacacaaacg gggcgtgtcc acccccgccgg 480  
 cccccccggc ggaaggaaag ggcctccca ggcggccac gtcctctctg ggcaggagcg 540  
 aagcgggggg gatgtccggc cccctcttgc tccacttcac tcctggcgc cttggcagcg 600  
 aagcgggggg ggagacagggt ggtgttcagg gtcctctcg ccaggccgca ggcggccgg 660  
 ggcacacagga cactggccccc tggcaggccg ggcggggccct ctcggctctg atggggagag 720  
 gggggggggc gggggggggc atggacgtcg gtccggaga cagaccccgcc gggccgggg 780  
 aacttggacc cgtccggctg ggcggccggg ggacgtgtg ggacgcgggt gacggggggcc 840  
 cccggccctgg cgcggccacgg cgccggctcc ccacaggacc cccactcgcc caggagccgg 900  
 cactccggc gtcggccctg ccccgccggc tgacgttcgg gtctcagacg gaaggctctg 960  
 gggccaaaggc tggctggggc aggaggacccgt cgggggtccc tgccggggca gggggctggc 1020  
 ctgggggttcg caggagggtc cggggccaccgg gggatcttgg cccaggccccca gactccacgg 1080  
 gccccctggct agagacttcg cctcaaggtc gcccactctc gtcccaaggd cccgggtgcca 1140  
 cagggggccta cgatggccgc gaggccgggg ctgacagtc ccgagataac agtctctggcc 1200  
 ctgacccctgg gcccacccgg ccccccggcgg aaccaaaaggc ggtgcagcc ggcacacggc 1260  
 ggcggccctc cggggccggc tgacgttcg cccggccccc ctcggccccc gcaactccgg 1320  
 gccccgggtt cccactgtaa gcccgtactc tccctcttc tcaagactt ttggcccttg 1380  
 aggttaccca ggatcttcc gtggggggaaa atctcaagac ggcgcacggcc ccaagttcag 1440  
 ctcctggccca agtcttaact tcacgttcgg cctcaacttc tcccaacttc agcccccggcc 1500  
 cccccactgc agccaccaacc tcacccacca cccacttc acccccccacc tcaagccccc 1560  
 ccccaacttc aacctcaactt ccaacttcac ccccaaggccc acggccaaatg ccacgttcag 1620  
 cggcaacttc agccccccaggcc ccaggccatcc tcccaacccct cacacccccc tcccaacggcc 1680  
 taaccccaactt cccaaacccca gcccataagcc cagtcacactc tccaggccctt accccaggccg 1740  
 catccccccggc ecttaacccca gtcggccaccc cgcctcaactt cccacttca acctccaggcc 1800  
 caaccccaaggc cgcatccctt gcccggggcc cccacttcac cccaaacccca accccaggcc 1860  
 catcccttcg cccagttcac gggteaaaggc ctcaggagag tgggtctctc cccaggccgt 1920  
 accaggaggg gcagggttcg gccagcttgg gaaaccttat tgccatggg ttttagaaaggcc 1980  
 acccccttcggc caggcaagac aggcccccaag ggagtgtccc ggggggggtt cccggggggcc 2040  
 cctgggggttc cagcacccctt acacacttc acggacagaca cggccctctt tcttcgttg 2100  
 ggacactcat acggggagggt acaggggggcc tgggtggcgg tggaggctcgg ccacagccaa 2160  
 gaagctccga gaccacacgg tggccctggc cagacccctcc cccaggcccta agaggagagg 2220  
 gaaccaggagg gaaaaactgtc gaccggctgc cccacccggc gatggccggc gggccggcc 2280  
 agccccccggc gcaaggccggc cttggccccc cggccctctc tctggcgagg cgctccacagg 2340  
 eggtaacccca gtcacggaaa cgcacggatc ggcggatcc cccggactccg gaggaggagg 2400  
 tcaggccggc cgccctgggg gaccctcaag gggaggccgc ggggggggg ggcacccctg 2460  
 cggccggccag cggggccgttc acggaaaaggc acggggggcc cggacactca tgggttttt 2520  
 geagggccggc ggggggttggg ggggttggg gggggccccc aacccggccat cccggccctt 2580  
 ggaggccggc acggggccggc cccctggccgg cggccgttggc ctggccgtt gcttggccgt 2640  
 gtcggccctt gcccggccgg acggggaggcc cggccctggcg caggagaccc acggcgctgt 2700  
 gagcgacaaac ctgtcgccgg tgctggggaga cccgttgcctc taccggccgc tgacggccggc 2760  
 cgacggccggc cgcacacttc gcttcggccac cggccggggcc cgggggggtgc tggggctgt 2820  
 ctgtactggcc acggcttaccc acggggggcc ctcagggttc cccaggggcc ctgtggccga 2880  
 ggagccctctt gggggggccctt ctgttccttgc gcttcaactt ggcacccctc atgtgttcaa 2940  
 ccccccgggg aacacctggc gggcccttgc acgggttccccc gaggaggccc cgcttccgggg 3000  
 ctgggggttc tgcacccatgc acaactactt gtttccggc gggggggccat ctgtggccctgg 3060  
 tgcccaaggcc gtctgtcaca acgggttctt ctgtcataac cctctggccca acatctggag 3120  
 ccagggttccg cccatgtcaca acggggggccac cccacttcac gtcggccccc tggacggggct 3180  
 gtcctatgc atcggtggcg aatgccttgc acggatggag tgctacgacc cgcaacacaga 3240  
 cgcctggacc ccacggccgc cactcccccggc aggacacttc cctgtggggcc acggggctgt 3300  
 ggcctggccgt gggggacatc acgttcacccgg gggtcaccc ttcactccggc tgctcaggta 3360  
 cggccctgggtt aagggttcgtt gggggagggtt cccatccatgc ggcacccccc ggggttccgg 3420  
 cgacatcggtt gcaactggggg gtttcgttgc acgggttccggc ctgttcggcc gctggggccg 3480  
 cccctgtatgc cgctacacaca cagtgtacccg ctcctggggc acgggggtcc cccggccctt 3540  
 gccccccccc gccggccacttc acgtgcaccc acgggttccccc accatataact gcttcaacc 3600  
 ccagggttccat gccacccatca cgggttccggc gggggggcc cagggttccggc ccaaggaggat 3660  
 gcaaccccttc cccctggggg gcaacccgggg cccacttcac ttcatcttgc ctctggccccc 3720  
 tgaggacccgg ctgcagacact cacttcgtgtt ggcaggccaga gacccaaaggc tgcttcgtt 3780  
 ctctccagggg agacccctctt gggatggggcc tgagggggccg gggctcaggg aaggggctgg 3840  
 gatggccatcc ttcgtgttcc gtttccgttgc aactttccccc ttctgtttaa aagggttggc 3900  
 attatttgtt gccccccggact cccttcggact ctgttcgtccct ctcactccac accggactgt 3960  
 ttcttcgtact caatttcgtt cttacttaca gacccttcac gtcgtgtccac accccccctgt 4020  
 ctgtggggact cccttattccc tagagccagg gactgtatgcg ttcacccacaga caaggacttg 4080  
 gtcgtgttgc gtcgtgttgc gccggagagag gagggggttag aaaacattca cacttcctat 4140  
 gtcgtgttgc caggacaggc agccaaaaacg tcccccaggca acggccctcgcc ctctggggcc 4200  
 ttcttcgtact cttacttcgtt cccctggggcc acggggggccat gtcgtgtccac accccccctgt 4260  
 atgtgttggatt ctcaaggggcc tggaaaccttgc gtcgtgttgc gtcgtgtccgt ggtgcagcc 4320  
 tgctgttgc gtcgtgttgc cataccccca gcccacccca gggccaggccc actccggccctt 4380

-continued

---

caccacccctc	tgcagccctg	tggggctc	ccagccccctc	cagaagccca	ccccacttct	4440
cgcacaacccc	cgatctcaa	atgaggccctg	agcgtcaccc	tagttctgcc	ccttttttagc	4500
tgtgttagact	tggacgagac	atttgacttc	cctttctctt	tgtctataaa	atgtggacag	4560
tggacgtctg	tcacccaaga	gagtttgggg	agacaagatc	acagctatga	gcacctcgca	4620
cggtgtccag	gatgcacagc	acaatccatg	atgcgttttc	tcccttaa	cactttgaaa	4680
cccatgttag	aaaagtgaat	acatctgact	gtgctccact	ccaaacctca	gcctggatgt	4740
ccctgtctgg	gccccttttc	tgttattat	tctatgttca	gcaccactgg	caccaaatac	4800
attttaattc	accgaaagca	aaaaaaaaaa	aaaaaaaaaa			4836

SEQ ID NO: 30	moltype = RNA	length = 2098				
FEATURE	Location/Qualifiers					
source	1..2098					
	mol_type = mRNA					
	organism = Homo sapiens					
SEQUENCE: 30						
cacacccatc	cccaacccca	accccagccg	catccccctgc	cccaagctgac	gggtcaaaagc	60
ctcaggagag	tgtggcttc	cccaggcgct	accaggagggg	gcagggtctca	gccagctggg	120
gaaacccat	ttggccatgtt	cttagaaagcc	accccttccc	caggcaagac	aggccccaaag	180
ggagtgtccc	gagggggggtt	ccccggggcc	cegtgggttcc	cggactttcc	acacactctg	240
aggacagaca	cgccccctct	tcttcgttgg	ggacagtcat	aggaggaggt	acagggggcc	300
tggttgaggc	tggaggctag	ccacagccaa	gaagctccga	gaccaacgga	tcgcccagcc	360
caaaaaatccc	ccccaggctt	agaggaggg	gaaccaggga	aaaaagtcta	gaccggctgc	420
ccaaagecgc	gatggccagg	ggcccccgcac	ageccccccgc	gcaggaggccg	cctggccccc	480
cggcccttc	ctctcgagg	cgctcacago	cggttacccca	gtacggaaa	cgcagcagg	540
gegaaatcgc	cccgagctcg	gacggaggagg	tcaggccggc	cgccctgggg	gaccctcaag	600
gggaggggccc	gggggggggg	ggcagccctg	cgccggccag	cgggggccctc	acggaaaagc	660
aggaggaggc	ccggaaatc	tggttttct	gcagaggccc	gggggttggg	gggttgttgg	720
ggggccccccgg	aaaggccat	ccccccccc	ggagcccccgc	acggccggcag	ccctgcggcg	780
ggggctgtgg	ctggggcgtt	gccttgacgt	gtggccctt	gcccagca	acggagagcc	840
cggcctggcg	caggagacct	acgcgtctat	gacgcacaa	ctgctggcg	tgctgggaga	900
cccgcttc	taccggccgg	tgaggccggc	cgacggccgg	cgacccatcca	gccttgccgac	960
cgccggggcc	cgccggggcc	tggggctct	cgtaactggcc	accccttacc	aggggggccc	1020
ctcaggggctc	cccaggggcc	ctctggcg	ggagcccttc	gcggcgcccc	ctgtgtccct	1080
gcctctactt	gcgcaccc	atgtgttca	ccccccggag	aacaccatggc	ggccctgtac	1140
cacggctggcc	gaggggggcc	cgcttcgggg	ctgggggttcc	tgccacatgc	acaactacc	1200
gtttctggcc	ggggggatcc	gtgggttcc	tgccaaggcc	gtctgttca	acggagttt	1260
ctgctacaac	cctctgacca	acatctggag	ccaggttccg	ccatgcagc	aggccggagc	1320
ccagctcaag	ctggggccc	tggacgggt	gtcttatgc	atcggtggc	aatgcctgt	1380
caacatgggg	tgtctacgg	cgccacca	cgccgtggac	ccacggccgc	cactccccc	1440
aggcaccc	cctgtggccc	acggggctgt	ggccggccgt	ggggacatct	acgtcacccg	1500
gggtcaccc	ttctaccggc	tgctcaggta	caagccccgt	aaggatgtt	gggacggatg	1560
cccatatacg	gccagccacc	ggcgttccag	cgacatctgt	gcactgggg	gtttctgt	1620
cgccttcgac	ctgtcgccgg	gctgtggcc	cgccgtgtat	cgctacaaca	cagtgtaccc	1680
ctctgggg	agggggttct	ccctgggg	gcccggcccc	gcccactgtc	actgcaccc	1740
cctggggcaac	accatctact	gcctcaacc	ccaggtca	gcacaccc	cggttctgt	1800
ggggactggc	cgatccagg	ccaaggagct	gcagcccttc	ccctgggg	gcacccgggt	1860
cctcagttca	ttcatctgt	ctctggcc	tgaggaccc	ctgcagac	cactctgt	1920
ggcaggcaga	aaacccaaagc	tgtttcgctg	cttccaggag	agacccctct	gggatggg	1980
ttagaggccg	ggggctcagg	aaggggctgg	gatcggaa	tcctgttctt	gttctggac	2040
aactttcccc	ttctgttta	aagggtgtcg	attatgttga	aaaaaaaaaa	aaaaaaaaaa	2098

SEQ ID NO: 31	moltype = RNA	length = 5124				
FEATURE	Location/Qualifiers					
source	1..5124					
	mol_type = mRNA					
	organism = Homo sapiens					
SEQUENCE: 31						
agaagcagg	tggctgtat	gacagcacag	agtcaggaa	cgctgcctga	ggaccctggg	60
gcctacagg	aggagaagag	ggcaggaggct	gggtgggtgc	ttgcagagac	cctgggtctcc	120
tatcctgc	taaggcctc	tgttctgt	tatctgc	caggccctac	tgacaccc	180
aggcctgtgt	gcaagcagag	acccccat	teccaggccc	tggaggactg	gtccaccc	240
actgggcage	ccttggggca	ggcgtggcc	ggtgcctca	cccaggcc	tgtgtctgc	300
atgcactgc	agccgtccat	caggccct	ttgcagcc	gaaccatgtat	ccagggcacc	360
ttggaggcc	atgggtcc	ctggggctgg	gactggggac	gtgacaatga	ctgggatagt	420
gtctgtgtgg	ccctcttgc	gtgggtgt	gtgggtgc	cagcgctggc	cttacactgg	480
tttggcttc	ggcagcatca	agaggggcc	gaacccgtgt	ccacaggcc	cggggctcaa	540
cctcatcagg	caggaggagg	tgatgtggc	ctgcaacc	agtctaa	cagtgtatgg	600
agcgaggggc	agagccccc	gcagggaaa	ccagagccc	caggacgg	ccagcagac	660
cctgtccctg	ctgcacgc	gggggggg	ctggccgc	tggcccggt	tccactca	720
acggctgtcg	aggggccccc	cagagaggca	ttaggacac	aacgggg	tgccaccc	780
gcggccccc	gagcggaa	aaaggagct	cccaggcc	gactgc	cctggccagg	840
agcgaagcag	gggggtatgtc	cgcccc	ctgatcc	tcactctcg	gacccctgg	900
agcgaagcgg	aggcgagac	agggtgtc	agggcgtct	ctgc	ccgcaggcccc	960
gcggggcaac	aggacactgg	cccc	ggggcc	ctcgatgggg		1020
agaggccgg	ggccggccgg	gcccgtgg	ggggcc	gagacagac	ccgcggcccc	1080
cgaaaaactgg	acccgc	cctggggcc	gggggg	gttggac	gggtggac	1140

-continued

---

ccgcgcgcgc	tggacgccc	cgcgcgeggc	ctccccacag	gaccggccact	cggccaggag	1200
cccgcatccc	ccgcgtgcc	cgccccccgc	gcctcgacgc	ctgggttcca	gacggaaaggc	1260
tctggggcga	agggtggcgt	gagcaggggg	gcctcggggg	tccctgcgcc	cgaggaggagc	1320
tggccctggg	ttagcaggga	ggtccccggc	accggagact	ttggcccaagc	cccgactcc	1380
acgcgcgcct	ggctagagag	tcgcctca	ggtcgcccac	tctegtcaca	agggccgggt	1440
gcacacgggg	cttacatgc	cggegaggcc	ggggctgtaca	gctcccgaga	taacagtcc	1500
gcgcgtgacc	tggggccac	ceggcccccg	gagcaagcc	agccggctgc	agccggccac	1560
agccgcgcgc	ctccccggag	ccgtgagct	cgccccggct	ccgcctcccc	gccccgagct	1620
cccgccccc	ggtccccacc	tgaagccctg	actctccct	tccttcaga	cttttgccc	1680
ctggaggta	cccgaggcc	tccctgggg	gaaaatctca	gagcggggcc	agccccaaqt	1740
tcagccttag	ccaaatgttt	aacttcagot	ccagccttag	tcctatggcc	agccctgtgt	1800
tcatccccca	gtcagcacc	aacctccgc	accacccctaa	cctctatccc	cacccatggcc	1860
ccagccccag	ctccaaatcc	agtcctcaact	tcaacccctcg	ccccagcccc	aagtccatgt	1920
gcagccgcaa	ctccagcccc	agccccggct	ccagtcctca	ccctcacacc	cccatccccca	1980
gccttaaccc	cagtcctaaac	ccccggctta	ageccagctc	caactccagc	ctaaccctca	2040
gcccgcattcc	cagccctaa	ccccggctta	accggagcc	taagcccaagc	tccaaatccca	2100
gcccccaacc	cageggcatc	ccctggccca	gccccccact	cagccccaa	ccccaaaccct	2160
gcccgcattcc	ctggcccgagc	tgacgggtca	aagctctagg	agagtgtggc	tctcccccagg	2220
cgctaccagg	agggggcagggt	ctcagccago	tggggaaacc	ttattggcat	ggttctttaga	2280
agccaccacct	tcccccggca	agacggggcc	caaggagtg	tcccgagggc	ggttcccccgg	2340
agccccgtgg	gtccccggac	ttccacacac	tctgaggaca	gacacggccc	cttcttctca	2400
gtggggacag	tcatagggg	aggttacagg	ggctctgttt	aggctgtgggg	tcagccacag	2460
ccaagaagct	cggagccaa	cggtatggcc	ageccagacc	ctccccccagg	cttaagaggg	2520
gagggaacca	ggggaaaaag	tctagatccgg	tcgccccca	ccgcgtatgg	caggggcccc	2580
geacagecccc	cggcgccagag	gcccgcctgg	ccccggccct	ctccctctgc	gaggccgtca	2640
caaggcggtac	cccaatgtac	gaaaacggc	aggtgtcgaa	tccggccggag	ctcgaggcag	2700
gaggctggc	cgccggcccte	ggggggaccc	caagggggg	cgccgggggg	ggggggccgc	2760
cctggccggcc	ggcgccggcc	gtctcggca	aaggcagg	aggccgggg	gctcatgttg	2820
tttctcaga	ggccccgggg	ttgggggggt	gttgaggggc	cccggaaggc	cagctcccg	2880
gccttggagc	cggccacccgc	ggcagccctg	cgccggggcc	tggacccctgg	cagttgcctg	2940
gacgtgtgg	ccttgcggca	gcgacacgg	gaggccggcc	tggcgeagga	gacctacggc	3000
ctgtatggcg	acaacccgt	gctgtgtgt	ggagacccgt	gctctatcc	ccggctgtgc	3060
gcccgcacc	gogagggcat	cctcagccgt	cggtacccggc	ggggccgggc	gggtgtggc	3120
gtctctgtac	tgccccggct	ctaccagggg	ggccgtctcg	ggctcccccag	ggccctctgt	3180
gcaaggagcc	tcctggggcg	gccccctgt	ccctgcctct	acctggccac	ctgcatgtgt	3240
tcaacccccc	ggagaacacc	tgggggcccc	tgacccagg	gcccggagg	gcccccttcc	3300
ggggctgggg	tctctgcacc	atgcacaact	acccgtttct	gggggggggc	atccgtgtgt	3360
ccgggtgcac	ggccgtctgc	tccaaacggg	tcttctgtca	caaccctctg	accaacatct	3420
ggagccagg	tccggccat	cagcaggccc	gagccccagg	caagctgtgt	ggccctggacg	3480
ggctgtctca	tgcctatcgat	ggcgaatggc	tgtacagat	ggagtgtgtac	gaccggcgaa	3540
caacacgtc	gacccacccgc	gcccactcc	ccgcaggcc	cttccctgtg	gccccacggg	3600
ctgtggctg	cggtggggac	atctacgtca	ccgggggtca	cctttctac	cgccctgtca	3660
gttacagcc	ctgtggggat	gtctggggac	agtgccatca	cagtggccac	caccggcggt	3720
ccagegacat	ctgtggactgt	gggggggtcc	tgttaccgtt	cgactctgt	cggggggtgg	3780
gcccgcgggt	gtatggctac	aacacatgt	ccggctctgg	gagcagggt	gcctccctgc	3840
ccctggccgc	ccccggccca	ctgcaatgt	ccaccctgg	caacaccatt	tactgcctca	3900
accccccagg	cacttcgtcc	ttcaatgtt	ctggggggac	tgccctgtc	caggccaaagg	3960
agtcgaccc	tttccccctt	ggggacccgg	gggttccat	tccattatc	ctgactctgc	4020
ccctctggag	ccggctgtcc	acccctactt	gtgtggcagg	cagagaaatca	aagtcgttcc	4080
gtctgtctcc	agggagaccc	tcctgggtat	ggctctgtat	ggccgggttc	agggaaagggg	4140
ctggatgtgg	aacttctgtc	tcttgcattt	ggacaactt	cccttttgc	tttaaagggt	4200
gtcgattatt	tttggatggca	gactccctca	gctcttttgc	cccccctact	ccacaccagg	4260
actgtttctt	gactcaatcc	cgttactact	taacggccct	ctcagttgc	tgacaccccc	4320
ctgtctgtgg	gactccctat	tcctctagat	caggactgt	tgctgttcca	cagacaagga	4380
cttggctgtc	ttggatgtctg	ctgagccgc	agaggagggg	gtagaaataa	ttcacacttc	4440
ctatgtctg	tcagcggac	aggggagaaa	aacgtcccc	ggcaacggcc	tcgcctctgg	4500
gactttctgc	ctgtcttaag	gcctccccc	gtaccaatcc	cgtagtata	tgggtctgtt	4560
tggcaactgt	gatttcata	ggcctagaac	ccttgcctct	gaaactggtc	cgctgggtca	4620
gccctgtgt	ctgcagctcc	tgccatata	cccaagccac	accaggccag	gcccactccg	4680
ggttcaccc	ccttcgtcc	cttgggggg	tctcccaat	cctcccaaa	cccacccac	4740
tttctgcacca	ccccccatct	ctaatatgg	cctggatgtc	accctatgtc	tgcccccttt	4800
tagctgtgt	gacttggacg	agacatttga	tttcccttcc	tccttgcata	taaaatgtgg	4860
acagtggacg	tctgtcaccc	aagagatgt	tgggagacaa	gatcacatgt	atgagcacct	4920
cgacacgggt	ccaggatgtc	cagcacaatc	catgtatgt	tttctccctt	taacgttc	4980
gaaacccatgt	ctgaaaatgt	gaatacatct	gactgtgtctc	cactccaaacc	tccagccctgg	5040
atgtccctgt	ctggggccctt	tttctgtttt	ttattctatg	ttagccacca	ctggcaccaa	5100
atacattta	attcaccgaa	agca				5124

SEQ ID NO: 32 moltype = RNA length = 2989

FEATURE Location/Qualifiers

source 1..2989

mol\_type = mRNA

organism = Homo sapiens

SEQUENCE: 32

ccacctcaggcc ccaacccatcc accccagccg catccccctgc cccagctgac gggtcaaagc 60  
ctcaggagag tttgggtcttc cccaggccgt accaggagg ggccagctgg 120

-continued

---

gaaaccttat	tgccatggtt	cttagaagcc	acccttccc	caggcaagac	aggccccaaag	180
ggagtgtccc	gaggcgcggtt	cccgggagcc	ccgtgggtcc	cagcaactcc	acacactctg	240
aggacagaca	ccggccctct	tettcaagtgg	ggcacagtcat	agggcagagt	acagggggcc	300
tgttgtggc	tggagggtca	ccacagccaa	gaagctccga	gaccaacgga	tcgcccagcc	360
cagacccctcc	cccaggccata	agaggagagg	gaaccaggga	aaaaagtcta	gaccgcgtgc	420
ccaaagccgc	gatgcccgagg	ggcccccgcac	ageccccccgc	gcagaggccg	cctggccccc	480
cgcccteetc	ctctgcgagg	cgctcacago	cggtacccca	gtacggaaa	ccgcacaggt	540
gegaaatcgc	cccgacgtcg	gagcaggagg	tcaggccggc	ccgcctcggg	gaccctcaag	600
gggaggggccc	gggggggggg	ggcagccctg	ccggccgca	eggggegctc	acggaaaagc	660
aggaggaggc	ccggaggactc	atgggttcc	tgcaagaggcc	cgggggttgg	gggtgtgg	720
aggggccccg	aaagccca	tccggggccc	ttggagccgc	eacggggca	gcctgcggc	780
ggccggctgga	cttggggcgt	tgctggacg	tgctggctt	tgcccaagcag	cacggagac	840
ccggcctggc	gcaggagacc	tacgcgtca	tgagcgaaca	cctgctgca	gtgctgggg	900
accctgtcct	ctaccggccgg	ctgacgcgg	ccgacccgcga	gcccattctc	agccctgcgg	960
ccggccgggg	ccgggggggt	ctgggggtcc	tegtactgc	cagectctac	caggggggcc	1020
gtctcaggcgt	ccccaggccc	ctehtggca	ggagccctct	gcccggccccc	ctgtgtccct	1080
gcctctacct	gcgcacactc	atgtgttca	ccccggggag	aacacctggc	ggccctgtac	1140
ccaggtgcgc	gaggggggcc	cgcttcgggg	ctgggggtcc	tgccacattgc	acaactacct	1200
gtttctggcg	gggggatccc	gtgggtccgg	tgccaaggcc	gtctgtccca	acgggttctt	1260
ctgctacaa	cctctgacca	acatctggag	ccaggttcgg	cccatcagc	aggcccggac	1320
ccagctcaag	ctggggccc	tggaagggtt	gtctatgc	atcgggtggc	aatgcgtgt	1380
cagcatggag	tgctacgacc	cgccaaacaga	cgccctggacc	ccacgcgcgc	cactcccccgc	1440
aggcacccctc	ctggggggcc	acgagggtgt	ggcgtccgt	ggggacatct	acgtcacccgg	1500
gggtcaccc	ttctaccggc	tgctcaggta	caagccccgt	aaggatgtt	gggacgagt	1560
ccatatacgt	ggcagccacc	ggcgttccag	cgacategtg	gactggggg	gttctctgt	1620
ccgcttcgcac	ctgtcgccgg	gctgtggggd	cgccgtgtat	cgctacaaca	cagtgacccgg	1680
ctctggggc	agggtgtccct	ccctggccct	ggccggccccc	gccccactgc	actgcaccc	1740
cctggggcaac	accattact	gcctcaacc	ccaggtca	gccacccatca	cggtctctgg	1800
ggggactgcc	catgtccagg	ccaaggagct	gcagcccttc	cccttgggg	gcacccgggt	1860
cctcagtc	ttcatcttgc	ctctggcccc	tgaggaccgg	ctgcagact	cactctgtgt	1920
ggcagggaga	aaaccaaaagc	tgctcgtcg	cttccaggag	agaccctct	gggatgggg	1980
tgagggggcc	gggctcagg	aaaggggctgg	gatcggaaat	tcctgtctt	gttctggac	2040
aactttcccc	ttctgttta	aagggtgtcg	attatttga	agcccaagact	ccctcagcc	2100
ctttctgtcc	ctcactccac	acccagactg	tttctctact	caattccgt	cctacttaca	2160
gacccttc	gtttgtcgac	accccccctg	ctgtgggact	cccttattcc	tagagccagg	2220
gactgtatcg	tctccatcaga	caaggactgt	gctcgtcg	gtctgtcg	gccgagagag	2280
gagggggtag	aaaacatca	cactccat	gtctgtcg	caggacagg	agcaaaaacg	2340
tcccccaggca	acgcctcgc	ctctggact	ttctgtctgt	cctaaggct	ccccaggat	2400
caaccccgta	gotatctgg	tctgtttgg	actgtggatt	ctcaaggccc	tagaaccctt	2460
gcctctgaaa	ctgggtcgat	gggtcagccc	tgctgtcg	agetcctgc	cataccccc	2520
gcccacacca	ggccagggcc	actccgggt	caccaccc	tgccgttcc	tggggctc	2580
ccagccccctc	cagaagccca	ccccactct	cgccaaaccc	cgatctctaa	atgaggctg	2640
adggtcaacc	tagtctcg	cctttagtgc	tggtgtact	tgacggagac	atttgactc	2700
ccttcttc	tgtcttaaaa	atgtggacag	ttggacgtctg	tcacccaaga	gatgtgtgg	2760
agacaagatc	acagatcatg	gcaccccgca	cggtgtccag	gatgcacagc	acaatccatg	2820
atgcgttttc	tccccttacg	cactttgaaa	cccatgtcg	aaaagtgaat	acatctgtact	2880
gtgtcttact	ccaaacctca	gctctggatgt	ccctgtctgg	gcccttttc	tgtttttat	2940
ctatgttca	gcaaccactgg	caccaaatac	atttaattc	accgaaaggc		2989

SEQ ID NO: 33                  moltype = RNA    length = 4836

FEATURE                  Location/Qualifiers

source

1..4836

mol\_type = mRNA

organism = Homo sapiens

SEQUENCE: 33

actgcacagcc	tgccatcagg	cctctattgc	agccctgaac	catgatccag	ggcaccttgg	60
agccagatgg	tcccctctgg	ggctggact	ggcacagtga	caatgaactgg	gatagtgtg	120
tgttgtggct	cttgcgtcg	gctgtgttgg	ctgccccagg	gctggcctta	cactgttttgc	180
gtcccgccca	cgatcaagag	ggcccgagaa	cggtgtccac	agccctcggg	gctcaacactc	240
atcaggcagg	aggaggtcg	ctggggctgc	aaccgaatgc	taagggtcgat	gtggcagcg	300
agggggcagag	cccaggcgag	ggggaaaccag	agccccccagg	acgcggccag	cagggccctg	360
tcctctgtgc	agcggccgggc	ggggggctgg	ccggccatggc	ccgggttcca	ctcaagacgg	420
ctgtcgagga	ggcccgccaga	gaggcattag	gacagcaacg	gggcagtggc	accccccgg	480
ccccccggcc	ggaaggaaag	gaggccatcca	ggccaggccac	tgcccttc	ggcaggagcg	540
aaggcagggg	gtatccgg	cccccttcga	tccacttc	tcctcgaggc	cctggcagcg	600
aaggcggaggc	ggagcagagg	gggtcagg	ctgtcttc	ccaggccca	ggcccccgg	660
ggcaacagaga	cactggcccc	ttggcaggcg	ggccggggcc	ctcggtctcg	atggggagag	720
ggccggggcc	ggggccggcc	atggacgtcg	gtctgggaga	cagagccgc	cgccccccgg	780
aactggaccc	gtcccgctc	ggcccgccgg	ggagcgtgt	ggacggcggt	gacggggcc	840
ccgcctctgg	cccccacccg	cgccgttcc	ccacaggacc	cccaactcg	caggagcccg	900
cactccccc	gtgtccctcg	ccccccccc	tgccgttcc	gtctcaagacg	gaaggctcg	960
ggggccaaagg	tggcttgg	agggggctc	cgggggtccc	tgcccccgga	ggaggcttgc	1020
cctgggttct	caggggatgt	ccggggaccc	ggagcttttg	cccaagccca	gactccacgc	1080
gcccctggct	agagagatcc	cctcaagg	gcccactctc	gtcccaagg	ccgggttcc	1140
caggggccat	cgatgcgcgc	gaggccgggg	ctgacagatc	ccgagataac	agtctgc	1200
ctgacactgg	gccccccccc	ccccccgg	aaggcaaaagc	ggctcgagcc	ggccacagcc	1260

-continued

gcgcggccctc	cggagccgt	gagccctcgcc	cgcgctccgc	ctccccgccc	gcagctcccg	1320
gcgcgggggtt	cccccttgaa	gccctgactc	tcccctctcc	ttcagactt	ttgccccctgg	1380
aggtaaccca	ggatcttc	gtggggaaa	atctcagagc	ggcgccagcc	ccaagttcag	1440
cctcagccca	agtcttaact	tcagtcagg	cctcagtcct	agccccagcc	ctgggttcat	1500
cccccaagtc	agcaccaacc	tcagccacca	cctcaacctc	atcccccaacc	tcagccccag	1560
cccccaagtc	aacctcagct	ccaaactcaa	ccccagccccc	agccccaaagt	ccagtcgcag	1620
ccgcaactcc	ageccccagcc	ccagtccccag	teccaaacccct	cacaccccca	ccccagccccc	1680
taaccccaagt	cccaaaaaaa	gccctaagcc	cagtcacaa	tccageccct	acccccagccg	1740
catccccccg	ctaaacccca	gtcccaaccc	cagccctaa	ccccagtcctca	actccagccc	1800
caaccccaagc	ccgcactccct	gccccagcc	ccacacttcag	ccccacccca	acccccagccg	1860
catccccctgc	ccccatgc	gggtccaa	ctcaggagaa	tgtggcttc	ccccaggcgct	1920
accaggaggg	gcaggcttc	gccagctggg	gaaaaccttat	tgccatgggt	cttagaaagcc	1980
accccttccc	caggcaagac	aggecccaag	ggagtgtccc	gagggegggt	cccgggagcc	2040
ccgtgggtcc	cagcactcc	acacactct	aggacagaca	ccggccctct	tcttcagttgg	2100
ggacagtcat	aggggacaggt	acagggggcc	tgggtggag	tgggtggat	ccacagccaa	2160
gaagctccga	gaccacccgg	tcgccccgg	caagacccccc	ccccaggctca	agaggagagg	2220
gaaccaggga	gaaaagtcta	gaccgctgc	cccaagccgc	gatgcccagg	ggccccgcac	2280
agccccccgc	gcagaggccg	cctggccccc	cggtcttc	ctctgcgggg	cgctcacagc	2340
cggttacccc	gctacggaaa	cgcagcagg	cgcaaatcgc	ccccagctcg	gagcaggagg	2400
tcaggccggc	ccgcctccggg	gaccctaa	ggggaggccg	ggggggagggg	ggcagccctg	2460
ccggccgcag	ggggggcgtc	acggaaaago	aggaggaggc	ccggaaagtc	atggtgttcc	2520
tgcaagggcc	cggggggttgg	gggggtgtgg	agggggccccc	gaagccca	tccccggggcc	2580
tggagccgc	cacggccgg	gcccgtggc	gggggtgtgg	cctggggact	tgccctggacg	2640
tgttgtccctt	tgccccccgg	cacggagag	ccggccgtgc	ccaggagacc	tacgctgtga	2700
tgagcgacaa	cttgtgtgcg	gtgtgggg	accctgtact	ctaccgcgg	ctgagcggcg	2760
ccgaccgcga	gwgacatcc	acgtcgccg	ccggccgggg	ccggccgggt	ctgggggtcc	2820
tcgtactggc	cagctctac	caggggggcc	gtcaggggct	ccccaggggc	ctgtggccg	2880
ggagcccttc	ggggggcc	ctgtgtccct	gcctctac	gcgcac	atgtttca	2940
ccccccgggg	aacaccttgc	ggccctgtac	ccaggtgccc	gaggaggccc	cgcttcgggg	3000
ctgcgggttc	tcacccatgc	acaactact	gttctgtgg	gggggcattcc	tgggctccgg	3060
tgccaaagg	gtctgtcc	acgggttct	ctgttacaa	cctctgtacca	acatctggag	3120
ccagggttggc	ccatcgacg	ggggccgg	ccagctca	ctgggtggcc	tggacgggct	3180
gtctatgc	atcggtggcg	aatgcgtt	ca	gatcatgg	tgctacgacc	3240
cgctggacc	ccacgcgcgc	cactccccgc	aggcac	cc	ctgtggccc	3300
ggctgtccgt	ggggggatct	acgtgttcc	gggttccat	ttcttacccgg	tgcttgcgtt	3360
cageccccgt	aaggatgtt	ggggacgtt	ccatacag	gcacccccc	ggcggtccag	3420
cgacatcggt	gacttgggg	gttctctgt	ccgttcc	ctgtgtgg	gcgttggcgc	3480
cgccgtgtat	cgcttacaca	cagtgtccgg	ctcttgg	agggtgtct	ccctgtccct	3540
ggccggcccc	gccccacttc	actgttcc	cctggcaac	accattttact	gccttcaaccc	3600
ccaggtaact	ccacatc	cggttctgg	ggggactgtc	cgatgttcc	ccaaaggagct	3660
gcagcccttc	cccttgggg	gcacgggg	cctcagtc	ttcatctgt	ctctgtccccc	3720
tgaggacggg	tgccagac	cactctg	ggcaggcaga	gaaccaaa	tgcttcgt	3780
cttcacagg	agaccccttc	gggttggcc	tgagggccgg	gggttccgg	aaggggctgg	3840
gatcggaat	tcctgtt	gttttcc	aacttttcc	ttctgtt	aaagggtgtcg	3900
attattttga	agcccaact	cccttgcgtt	ctttctgtcc	ctcactcc	accccaact	3960
tttctgtact	caatccgt	cttacttaca	gacccttc	gtttgttgc	accccccctgt	4020
ctgttgttgc	ccatcttcc	tagatccgg	gactgttgc	tccatccaga	caaggacttgc	4080
getcgctgt	gtctgtgt	gcccggag	gagggggtag	aaaatccat	cacttctat	4140
getctgttag	caggacagg	ggggggat	ccccggcc	acgccttc	ctcttggact	4200
ttctgtctgt	cttcttgc	ccccgggt	caacccgt	gtatatgtt	tctgtttgg	4260
actgtgttgc	tcacggggc	tagatcc	gttctgtt	gtgttccgt	ggtgcac	4320
tgctgtgtc	agcttcttc	catatccca	gcccacacca	ggccaggccc	actccgggt	4380
caccaccc	tcacgttcc	tgggtctc	ccagccccc	cagaagccca	ccccactt	4440
cgccaaaccc	cgatcttca	atggggct	agcgatcacc	tagtttgc	cttttttag	4500
tgtgtact	tggttgc	atggactt	ccttttctt	tgttataaa	atgtggac	4560
tggacgttgc	tcacccaa	gagttgtgg	agacaagat	acatgtat	gcacccgtca	4620
cggtgtccag	gtgcac	acatccat	atgcgtttt	tcccttac	cacttgc	4680
ccatgttag	aaaatgt	acatctgact	gtgttcc	acttccat	ccaaacccca	4740
cctgttgc	gccc	tgtttt	tctatgt	tcacccat	gcctggatgt	4800
atttaat	accgaaag	aaaaaaa	aaaaaaa	aaaaaaa	aaaaaaa	4836

SEQ ID NO: 34 moltype = RNA length = 2098  
 FEATURE Location/Qualifiers  
 source 1..2098  
 mol\_type = mRNA  
 organism = Homo sapiens

SEQUENCE: 34  
 ccacctcagc cccaaacccca acccccagg catccccgtc cccagctgac gggtcaagc 60  
 ctcaaggagat tgggttgc cccaaggccgtt | accaggaggg | gcagggttcc | gccagctggg | 120 || gaaacccctat | tttgcgtt | cttaga | acccttcc | caggcaagac | aggccccaa | 180 |
ggagtgttcc	gagggcggtt	ccggggggcc	ccgtgggtcc	cagcacttcc	acacactctg	240
aggacagaca	ccggccctct	tcttcgttgg	ggacagtcat	agggacaggt	acagggggcc	300
tgggttggcc	tgggttgc	ccacagccaa	gaagcttcc	gaccaacgg	tcgcccagcc	360
cacaccc	ccacggcc	agaggaggg	gaaaatgt	gacccgttcc	420	
cccaagecgc	gtgcac	ggccccc	acccccccgc	gcagaggccg	cctggccccc	480
cggcccttc	ctctgtcg	cgctc	ccatccat	cgatcgat	540	

-continued

---

gcaaaatcgc	cccgagctcg	gaggcaggagg	tcaggccggc	cgectcgaaa	gaccctcaag	600
gggagggccc	ggggggagggg	ggcagccccgt	ccggccgcag	cgggggcgctc	acggaaaagc	660
aggaggagggc	cgggaaggtc	atgggtttc	tgcagggcc	cgggggttgg	gggggtgggtgg	720
aggggccccg	gaagccccgc	tccggggccc	tggagccccg	cacggggcga	gccctgcggc	780
ggcggtctgg	cttggggcagt	tgctggacg	tgctggctt	tgccccagcag	cacggagagc	840
ceggccctggc	ggaggagacc	tacgctgta	tgagcgcaca	cctgtcgca	gtgttgggag	900
accctgtgcct	ctaccggccg	ctgagcggg	ccgaccggca	gecacttc	agccgtcgga	960
ceggccgggg	ccggggcggt	ctggggctcc	tcgtactgc	cagcccttac	cagggggccc	1020
getcagggt	ccccaggccc	ctcgtggcga	ggagcccttc	ggggcgccc	ctgtgtccct	1080
getcttaccc	ggcgactctg	atgtgttcaa	ccccgggggaa	aacacctggc	ggccctgtac	1140
caagggtggcc	ggaggggccc	cgcttggggg	ctgagggtctc	tgacccatgc	acaactaccc	1200
gtttctggcc	gggggcatcc	tggttccgg	tgccaaaggcc	gtctgtcca	acgagggttt	1260
ctgtctacaac	cctctgacca	acatctggag	ccaggttccg	ccatgcagc	aggcccgagc	1320
caagctcaag	ctgggtggcc	tggacgggt	gtctatgc	atcgggtggc	aatgcctgt	1380
cagcatggag	tgtctaccc	cgcgaacaga	cgcgtggac	ceacggcgcc	cactccccgc	1440
aggcacccct	cttggggccc	acgaggctgt	ggcttgcgt	ggggacatct	acgtcacccg	1500
gggtcacctc	tttacccgcc	tgctcaggta	cagccccgt	aaggatgtt	gggacgagt	1560
cccatacatgt	ggccatccccc	ggcgttccag	cgacategt	gcactgggg	gtttctgtta	1620
ccgcttcgac	ctgtcgccgg	gctgtgtat	cgctacaaca	cagtgacccg	1680	
cttctggagc	agggtgtctt	cccttgcctt	ggccccc	gccccactgc	actgcaccc	1740
cctggggcaac	accattttact	gcctcaaccc	ccaggtca	gcacaccca	cggttcttgc	1800
ggggactgtcc	cagggtccagg	ccaaagggtt	gcagcccttc	cccttgggg	gcacccgggt	1860
cctcagtc	tttacccgtt	ctctggccccc	tgaggacccg	ctgcagact	cactctgt	1920
ggcaggcaga	gaaccaaaagc	tgcttcgtgt	ctctccagg	agaccccttct	ggggatggggc	1980
tgagaggccc	gggctcaggg	aaggggctgg	gatcggact	tcctgtctt	gtttctggac	2040
aacttcccc	ttctgttta	aagggttgtcg	attatttta	aaaaaaaaa	aaaaaaaaa	2098

---

SEQ ID NO: 35	moltype = DNA	length = 5124				
FEATURE	Location/Qualifiers					
source	1..5124					
	mol_type = other DNA					
	organism = Homo sapiens					
SEQUENCE: 35						
agaagcagggt	tggctgtat	gacagcacag	agctcaggaa	cgctgcctga	ggacccttggg	60
gcctacagg	aggagaagag	ggcaggaggt	gttgggggtgc	tttcagagac	cctggggctcc	120
tatccctgc	taagccctgc	tgtctctgt	tatctgc	caggccctac	tgacaccccc	180
aggccctgagt	gcaaggcagag	accccccatt	tcccaggccc	tggaggactg	gtccaccc	240
actgggcagc	ccttggggca	ggcgtgtggc	gggtgcctcg	cccaggcc	tgtgtctgc	300
atgcactgc	acgcctccat	caggccctta	tttcagccct	gaaccatgt	ccaggccacc	360
ttggagggcag	atgggttccct	ctggggctgg	gactggggca	tgacaaatga	ctggggatagt	420
gtctgtgtgg	ccctcttggc	gtctgtgt	gttgctgc	cagcgtggc	cttactctgg	480
tttggctccg	ggcacatcg	agaggcggca	gaaccgggt	ccacagcc	cggggctcaa	540
cctccatcagg	caggaggcgc	tgtagtggcc	ctgcaacccg	agtctaagg	cagtgtggc	600
agcgaggggc	agagcccgagg	gcaggggaaa	ccagagcc	caggacggg	ccagcagagc	660
cctgtccctg	ctgcacgc	gggggggg	ctggccgc	ttggcccgct	tccactcaag	720
acggctgtcg	aggaggcccc	cagagaggca	tttaggac	aacggggcag	tgccaccccc	780
ggggcccccc	gagcggggagg	aaaggaggat	cccaaggcc	gactgcct	cttggggcagg	840
acgcgaaggcag	ggggggatgtc	cgccccctc	ctgttccact	taactctcg	gagcccttggc	900
acgcgaaggcgg	aggcgagac	agggtgtc	agggcgtct	ctgcacggc	cgccaggccc	960
gcggggcaac	aggacactgg	cccctggcag	gcggggcg	ggccctcg	ctcgatgggg	1020
agaggccggg	ggccgggggg	ggggatggac	gttggctcg	gagacagac	ccggccggccc	1080
cgaaaaactgg	acccgcctcg	cctggggcgc	ggggggagg	tgtgggacgc	gggtggacgg	1140
ggccggccccc	ttggacggccca	cgccggcgc	ctccccacag	gaccccaact	cgccaggag	1200
ccgcactcc	cgccgtgc	cgccccccg	gcctgc	ctgggtctca	gacggaaagg	1260
tctggggcoca	agggtggct	gaggcaggag	gcctcg	ccccccccc	cgaggaggc	1320
tggcccttgg	tcagcaggga	ggteccgggg	acccgggat	ttggccca	cccagactcc	1380
acgcgcctct	ggcttagag	tccctcaaa	ggtgc	ccatcc	agggccgggt	1440
gcccacaggg	cctacatgc	cgccgaggcc	ggggctgaca	gtccccaga	taacatctt	1500
ggcgctgacc	ttggggccac	ccggcccccc	gagcaagca	agccggctgc	agccggccac	1560
ageccgcgc	cctccgggg	ccgtgaccc	cgccgcgc	cgccctcc	gcccgcacgt	1620
ccggcccccgg	ggttccca	tgaaggccctg	actctccctt	ctccctca	ctttttggcc	1680
ctggaggtta	cccaggatcc	ttccgtggcc	gaaaatctca	gagcggcc	agccccaa	1740
tcagcctcag	cacaaggctt	aacttcag	ccagcctcg	tcctagcccc	agccctgt	1800
tcatccca	gttcacggcc	aaacccatcc	cctcatcccc	caccc	tcaccc	1860
ccagcccccc	cctccaa	agctccaa	tcaacccca	ccccagccc	aagtcc	1920
gcagccgca	ctccagcccc	agccccatgc	ccagtc	ccctcacacc	ccccatcccc	1980
gcccatacc	cagtc	cccaaggcc	agccca	caactcc	ccataccca	2040
gcccgcattcc	cagccctaa	cccagtc	acccca	taagcc	tcacactca	2100
gcccccaacc	cagccgc	ccctggccc	gccccca	cggcc	cccaacccca	2160
ggcgatccc	ctggccc	tgacgggtca	aaggctc	aggatgtggc	tctccca	2220
cgatccatgg	aggggcagg	ctcagcc	ttggggaa	ttatgtccat	gttctttaga	2280
agccacccct	tcccccagg	caagacggcc	caagg	ggatgtggc	gggtcccccgg	2340
agccccgtgg	gtccccagg	ttccacac	tctgaggaca	gacacggccc	cttcttctca	2400
gtggggacag	tcatagg	ggtacagg	aggctgggg	aggctggagg	tcagccacag	2460
ccaagaagct	ccgagacca	cgatcgccc	agccca	ctccccca	cctaagagga	2520
gaggggacca	ggggaaa	tctagaccc	ctgccccaa	cgccgtatgc	cgccggccc	2580

-continued

gcacagcccc	ccgcgcagag	gccgcctggc	cccgcggccct	cctcccttcgc	gaggcgctca	2640
cagccgttac	cccgatcgtacg	gaaacgcgcg	agggtgcgaaaa	tgcggccggag	ctcgaggacag	2700
gagggtcaggc	ccggccgcctc	gggggaccct	caaggggggg	cgccgggggg	ggggggcagc	2760
cctgcggccc	gaagcgggggc	gctcacggaa	aaggcaggagg	aggcccgaaag	ctcatgggt	2820
ttctcgacag	gccccgggggt	tgggggggtgg	tggagggggc	ccggaaagccc	agctccccgg	2880
ccttgaggcc	cccccacggcg	gcagccctgc	gggggggggt	ggacccctggc	agtggccctgg	2940
acgtgtcgcc	ctttggccag	cagcacggag	ageccgggg	ggcgcaggag	acccatcgcc	3000
tgtatgacgca	caacccgtctg	cgatgtctgg	gagacccctgt	cctctaccgc	cggtgtggcg	3060
cggcccgaccg	cgagcgcatc	ctcagccctgc	ggacccggcc	ggccgggggg	gtgtggcg	3120
tcctctgtact	gccccggctc	taccaggggg	ggccgtcaagg	gtcccccaagg	ggccctctgt	3180
gcgaggaggcc	tcttgcggcg	gcccctgtg	cctgtgtctt	acccgtcgac	ctgcattgt	3240
tcaacccccc	ggagaaacacc	tgggggggcc	tgacccagggt	gcccgggggg	gcccccttc	3300
gggggtcgccg	tctctgcacc	atgeacaact	acctgtttct	ggccgggggg	atccgtggct	3360
cccggtccaa	ggccgtctgc	tccaaacggc	tcttctgtca	caacccctctg	accaacatct	3420
ggagccagggt	tccggccatg	cagcaggccc	gagcccaagg	caagatgggt	ggccctggacg	3480
ggctgtctta	tgccatcggt	ggcgaatggc	tgtacagcat	gagatgtctac	gaccggcgaa	3540
cagacgcctg	gacccacgc	gcgcactcc	ccgcaggcc	cttccctgtg	gccccacgagg	3600
cttgtggctt	ccgtggggat	atctacgtca	ccgggggtca	cctctttcac	cgccgtctca	3660
ggttacagcc	cgtagggat	gtttggggacg	agtggccata	cagtggccagc	cacccggctt	3720
ccagcgtacat	cgccgtactg	ggggggcttc	tgtacccgtt	cgacccgtcg	cggggggtgg	3780
gcgcgcgggt	gtatgcgtac	aacacagtga	ccggctctgg	gagcagggt	gcccccttc	3840
ccttgcggcc	ccccccccca	ctgcactgca	ccacccctggg	caacacattt	tactgcctca	3900
accccccagg	cactgtccacc	ttcacggct	ctggggggac	tgccggggat	cggccaaagg	3960
agtcgcggcc	cttcccccgt	ggggccacgg	gggtccctca	tccatccatc	ctgactctgc	4020
ccctgtggaa	ccggctgcag	acccactct	gagtggcagg	cagagaacca	aagctgtttc	4080
gtgtgtctcc	agggagaccc	tccttgggtg	ggccctgagag	gccccggctc	agggaaagggg	4140
cttggatcggt	aacttgcgtc	tcttttttgc	ggacaactt	ccccctgtgc	tttaaagggtt	4200
gtcgatttatt	ttaaaggccca	gactccctca	gcctctttct	gccccctact	ccacacccag	4260
actgttttct	gactcaattc	cgtaccaact	tacagaccc	ctcagettgc	tgacacccccc	4320
ctgtctgtgg	gactccctat	tccctagago	cagggactga	tgcgcttc	cagacaagga	4380
cttggctcgc	tggagctcg	ctgagccgg	agaggagggg	tgtagaaaaaca	ttcacacttc	4440
ctatgtctgt	ttagcaggac	aggaggaaa	aacgtccccca	ggcaacggcc	tcgcctctgg	4500
gactttctgc	ctgtcttaag	gcctcccaag	gtaccaaccc	cgtagctata	tgggtctgtt	4560
tggcaactgt	gattctcaag	ggcctagaaac	ccttgcctct	gaaactggc	cgctgggtca	4620
gcctctgtgt	ctgcacgtcc	tgccatccatc	ccccggccac	accaggccag	gccccactcc	4680
gggttcaccc	cctctcgacg	cttgcgggg	tctcccgcc	cctccagaag	ccccccccc	4740
ttctcgccaa	cccccgatct	ctaaatgggg	cctgaggcgtc	accctgtttc	tgccctttt	4800
tagctgtgt	gacttggacg	agacatttga	cttcccttcc	tccttgcata	taaaatgtgg	4860
acagttggacg	tctgtccaccc	aaggaggat	tggggagacaa	gatcacatgt	atggcacct	4920
cgcacgggt	ccaggatgtca	cagcacaatc	catgatgcgt	tttccctcc	tacgcacattt	4980
gaaacccatcg	ctagaaaaatg	gatacatctt	gactgtgtctc	cactccaa	tccagccctgg	5040
atgtccctgt	ctggggccctt	tttctgtttt	ttattctatg	ttcagcacca	ctggcaccaa	5100
atacattttt	attcaccggaa	agca				5124

SEQ ID NO: 36	moltype = DNA	length = 2989				
FEATURE	Location/Qualifiers					
source	1..2989					
	mol_type = other DNA					
	organism = Homo sapiens					
SEQUENCE: 36						
ccacacctacgc	cccaacccca	accccgccg	catcccttcgc	cccagctgac	gggtcaaaagc	60
ctcaggagag	tgtggcttc	ccaggcgct	accaggagggg	gcaggctc	gccagctgg	120
gaaacccattat	tgccatcggtt	cttagaaagcc	accccttc	caggcaagac	aggccccaag	180
ggagggtgtccc	gaggggcggtt	cccgggaged	ccgtgggtcc	cagcacatcc	acacactctg	240
aggacagacaca	ccggccctct	tcttcgtgt	ggacagtcgt	agggaggat	acagggggcc	300
tggtttgaggc	tggagggtcg	ccacagccaa	gaagctccga	gaccaacgg	tcgcccagcc	360
caagacccctcc	cccaaggccct	agaggagagg	gaacccaggga	aaaatgtcta	gaccggctgc	420
cccaageccgc	gatgcccgagg	ggcccccgcac	ageccccccgc	gcagaggccg	cctggccccc	480
ccggccttc	ctctgcgggg	cgctcacago	cggtacccca	gtaacggaaa	cgccaggccgt	540
gcgaaatcg	cccgacgtcg	gaggcagggg	tcaggccggc	cgccctgggg	gaccctcaag	600
ggggaggccccc	gggggggggg	ggcagccctcg	ccggcccgac	cgggggccgtc	acggaaaagc	660
aggaggaggcc	ccgaagctca	tggttttct	gcagaggccc	gggggttggg	gggtgggtgg	720
ggggccccccgg	aagcccaagct	cccgggccct	ggagccggcc	acggccggcc	ccctgcggcg	780
ggggctggac	ctggggcggt	gcctggacgt	gtggccctt	gcccaggccg	acggagagcc	840
cggccttc	cgaggacatct	acgcgtcgat	gaggcgacaa	ctgtcgccgg	tgctggggaga	900
cccggtcc	taccggccgc	tgaggcgccgc	cgaccgcgg	cgcatctca	gcctgcggac	960
cgggccggggc	cgggcggtgc	tggggctct	cgtactgc	accccttacc	agggggggcc	1020
ctcagggtc	cccaggggcc	ctcgtggcg	ggagccct	gcggccccc	ctgtgtccct	1080
gcctctactt	gogcacactgc	atgtgttca	cccccggggg	aacaccttgc	ggccctctgc	1140
ccagggtgc	gaggaggccc	cgcttcgggg	ctgcgggtctc	tgccatgtc	acaactactt	1200
gtttctggcc	gggggcattcc	gtggctccgg	tgccaaggcc	gtctgttc	acgagggttt	1260
ctgttacaa	cctctgtacca	acatctggag	ccagggtcg	ccatgtc	aggcccgagc	1320
ccagctcaag	ctgggtggccc	tggacgggct	gctctatgc	atcggtggcg	aatgcctgt	1380
cacatgtggag	tgctacgacc	cgcaaacaga	cgccctggacc	ccacgcgcgc	cactcccgcc	1440
aggcaccttc	cctgtggccc	acgaggctgt	ggectggccgt	ggggacatct	acgtcaccgg	1500
gggtcaccc	ttctaccggcc	tgctcaggta	cagccccgtg	aggatgttt	gggacagatg	1560

-continued

---

ccatacaggt	gccagccacc	ggcggtccag	cgacatcg	gcaactgggg	gcttcctgta	1620
cgcgttgcac	ctgtcgccgg	gctgtggggc	cgccgtgat	cgctacaaca	cagtgacccg	1680
ctcttggagc	agggtgcgt	ccctggccc	gcccggcc	ccccactgc	actgcacac	1740
cctggggac	accattact	gcctcaacc	ccaggtca	gcaccccca	cggtctcg	1800
ggggactgc	catgtccagg	ccaaggagct	gcagccctc	cccttgggg	gcaccgggt	1860
cctcagttca	tccatctga	ctctggccc	tgaggacgg	ctgcagact	cactcgat	1920
ggcaggcaga	gaaccaaagc	tgcttcgt	ctctccagg	agaccctct	gggatgggc	1980
tgagaggcgc	gggctcagg	aaggggctgg	gatcggaa	tcctgtctt	gttctggac	2040
aacttcccc	ttctgttta	aaggtgtcg	attatttga	agcccaact	ccctcagct	2100
cttctgcoc	tcactccac	accggactg	tttcctgact	caatccgt	cctacttaca	2160
gaccctctca	gttgcgtac	acccctcg	ctgtggggact	cccttattcc	tagagccagg	2220
gactgtcg	tccacacaga	caaggactg	gtcgcgtga	gctctgtga	gccgagagag	2280
gaggggtag	aaaacattca	cacttctat	gtctgtca	caggacagg	agcaaaaac	2340
tcccccggca	acggccctgc	ctctgggg	ttctgcctt	cctaaggct	ccccaggat	2400
caacccccgg	gtatctgg	tctgtgggc	actgtggatt	ctcaagggg	tagaaccct	2460
gcctctgaaa	ctggctcggt	ggtcagcc	tgtgtctgc	agtcgttgc	cataacccca	2520
gccccaccca	ggccaggccc	actccgggt	caccaccc	tgca	ccgttgc	2580
ccagccctc	cagaagccca	ccccactct	cgccaa	ccatctaa	atgaggct	2640
agegtcaccc	tagttctgc	ccttgc	tgtgttagact	tggacgagac	atttgactt	2700
cctttctct	tgtctttaaa	atgtggacag	tggacgtctg	tcacccaga	gagtttggg	2760
agacaagatc	acagctatga	gcacccgtca	cggtgtccag	gatgcacagc	acaatccat	2820
atgcgtttc	tcccttta	cacttggaa	cccatgtca	aaaagtgtat	acatctgact	2880
gtgtccact	ccaacccctca	gccttggatgt	ccctgtctgg	gcctttttc	tgttttttat	2940
tctatgttca	gcacccactgg	caccaatac	attttaatc	accgaaagc		2989

---

SEQ ID NO: 37 moltype = DNA length = 4836  
 FEATURE Location/Qualifiers  
 source 1..4836  
 mol\_type = other DNA  
 organism = Homo sapiens

SEQUENCE: 37  
 actgccagcc tgccatcagg ccttattgc agccctgaac catgatccag ggcaccttgg 60  
 agccagatgg tcccccttgg ggctggact gggacagtgta caatgttcg gatagtgtc 120  
 tgctggccct cttggcgctg gctgtgggt ggccacacg gctggccta cactggttt 180  
 gtcggggggca cgatcaagag gccggcagaad cgggtgtccac agccctcg 240  
 atcaggcagg aggagctgat ctggccctgc aaccgaatgc taaggctgt gatggcagcg 300  
 agggggcagag ccacggccag gggaaaccag agccccccagg acgcggccag cagacccctg 360  
 tccctgtgc agcgcggggc gggggctgg ccgcgcattgc cccgcgttca ctcagaacagg 420  
 ctgtcgagga gggcccgaga gaggcattag gacagcaac gggcgtgtcc acccccgccg 480  
 ccccccggc ggaaggaaag gaccccttca ggcaggcaca tgccttcctg ggcaggagcg 540  
 aacgcagggg gatgtccgc ccccttctga tccacttcac tcctcgaggc cttggcagcg 600  
 aagcggggc gggagacaggt ggtgtcaggg cgtccctcgc ccaggccgca gggcccgccg 660  
 gcaacacgaa cactggccccc tggcaggcgg ggcggggggcc ctcggggctcg atggggagag 720  
 gcccggggccg gggggggccg atggacgtcg gtcggggaga cagacccgc cggcccccgg 780  
 aactggaccg gtccgcgtc ggccgcgggg ggacgtgtcg ggacgcgggt gacggggccg 840  
 ccgcgccttgc gccccacgcg cgcgcgtcc ccacaggacc cccactcgcc caggaccccg 900  
 cactccggc gtcggccgtc cccggccggc tgacgcgtcg gtctcagacg gaaggctcg 960  
 gggccaaagg gttgtggaggc agggggcgt cgggggttcc tgccccggg ggggttggc 1020  
 ctgggttgcg caggagggtc ccggggccccc ggagcttgg cccagccca gactccacgc 1080  
 gccccttgcg agagagtccg cctcaagggtc gcccactctc gtcccaaggg ccgggttgc 1140  
 caggggccca cgtatccggc gaggccgggg ctgacagctc cccgatatac agtctcg 1200  
 ctgacctggg gcccacccgg ccccccggc aacggccggc ggtcggcc ggcacacgac 1260  
 ggcgcgcctc ccggaggccgt gacccgtccgc cgcgcgtcc ccccccggcc gcaactcccg 1320  
 gcccggggcc cccacctgaa gcctgactc tcccttctcc ttcaactt ttggccctgg 1380  
 aggttaccca ggatecttcc gtggggcgaa atctcagacg ggcgcggccc ccaagttcag 1440  
 cctcagccca agtttactt acgttcact tcacgttcg ctcgttccat agcccgccg 1500  
 ccccccagtc aaccaaccc tcacccca cctcaacccct atccccccac tcagccccc 1560  
 ccccaagtc aacccatgat ccaacttcaa cccacggccc agcccaactt ccacgtcg 1620  
 ccgcacactcc acggcccgcc ccacgtccac tcccaacccca tcccccggcc 1680  
 taacccccatg ccaacccca gcccataa cagcttcaac tccacgtccca accccgccg 1740  
 catccccccgc ctaacccca gtcaccaaccc cggccatccaa cccactccaa actccggcc 1800  
 caacccccccgc cgcataccct gccccagccc ccacccatgc cccaaacccca accccgccg 1860  
 catccccctgc ccacgttgac gggtcaaago tcacggagag tgggttgc tccaggccgt 1920  
 accaggaggcc gcaaggctca gccagctgg gaaacccat tggccatgtt cttaaagcc 1980  
 accccttccc caggcaagac agggcccaag ggatgttcc gaggccgggt cccggggagcc 2040  
 ccgtgggtcc cggcacttcc acacactcg aggacagaca cggccctct tcttcgttg 2100  
 ggacagtcat agggacagggt acaggggggc tgggttggggc tggagggtcg ccacacgca 2160  
 gaagctccga gaccaacggc tcgcccagcc cagaccctcc cccaggccca agaggagagg 2220  
 gaaccaggaa gaaaaatcta gacccgtcg cccaaaggccg gatgcggccg gggcccgcc 2280  
 agccccccgc gcaaggccgc cttggccccc cggccctcc tcctcgttcc cgttcacacg 2340  
 cggtacccca gtcacggaa cgcaggccgt gcaaaatcg cccgactcg gaggccgg 2400  
 tcaggccggc ccgcctgggg gacccatcaag gggaggccg gggggggggg ggcggccctg 2460  
 cccggccgcag cggggccgtc acggaaaago aggaggccg cggaaactca tgggtttct 2520  
 gcaaggccgc gggggtttggg ggggttggaa gggggccccc aagcccaactt cccggggccct 2580  
 ggagccgcgc acggccggcag ccctgcggcc gccggctggac ctggccgtt gcttggacgt 2640  
 gctggccctt gcccacgcg acggagagcc cggccatggcg caggagacact acgcgcgtat 2700

-continued

---

gagcgcacaac	ctgctgcgag	tgctggaga	cccggtgcctc	taccggggc	tgagcgggc	2760
cggaccggcg	ccgcattctca	gcctcgggac	cgggccgggc	cgggcggtgc	tgggcggtct	2820
cgtactgcgc	agectctacc	aggggggcgg	ctcaggggc	cccaggggcc	ctcggtggcga	2880
ggagccctct	ggggggggcc	ctgtgtccct	gcttctactc	gcccaccc	atgtgttcaa	2940
ccccccgggg	aacacctggc	ggccctgtac	ccaggtgccc	gaggaggccc	cgcttcgggg	3000
ctgcgggttc	tgcaccatgc	acaactactt	gtttctggcg	ggggggcatcc	gtgggtccgg	3060
tgccaaaggcc	gttgtgttca	acgggggttt	ctgttacaa	ccttgcacca	acatctggag	3120
ccagggttgg	cccatgcg	aggcccgago	ccagctca	ctgggtggccc	tggaegggct	3180
gctctatgcc	atcggtggcg	aatgcctgt	cacatggag	tgctacgacc	cgccaaacaga	3240
cgcctggacc	ccacgcgcgc	cactccccc	aggcaccc	cctgtggccc	acgaggctgt	3300
ggccgtccgt	ggggacatct	acgtcggcc	gggtcaccc	ttcttaccc	tgttcaggta	3360
cagccccctg	aaggatgtt	ggggcggatgt	ccccatacagt	gcccggcc	ggcggttcag	3420
cgacatctgt	gcactgggg	gttctctgt	cegcgttc	ctgtcgcc	gcgtggggc	3480
cgccctgtat	cgcttacaca	cagtgcaccc	ctctggggc	agggtgtct	ccctggccct	3540
gccccccatc	gccccatc	actgcaccc	cetggggc	accatattaa	gcctcaaccc	3600
ccaggtaact	gcacccatca	cggttctgg	ggggactgc	cagtccagg	ccaaggaaqct	3660
gcagcccttc	cccttggga	gcacccgggt	cctcagtcc	tttcatctga	ctctggccccc	3720
tgaggacggc	ctgcagac	cactctgt	ggcaggcaga	gaaccaa	tggttcgt	3780
ctctccatct	ggggatgggc	tgagggccgg	gggggtcg	tccatccaga	caaggactt	3840
gatccggaa	ctctgttctt	tttttggc	aacttcccc	ttctgttta	aagggtgt	3900
attattttga	agcccgact	ccctcagct	ttttctggc	ctcactccac	acccagactg	3960
tttcctgact	caatttgc	cttacttca	gccccttc	gtttgttgc	acccccc	4020
ctgtggact	cccttattcc	tagagggc	gactgtatgc	tccatccaga	caaggactt	4080
getcgtgg	gtctgttca	gcccggat	gagggggt	aaaatctca	cacttccat	4140
getctgttca	caggacagg	agaaaaaa	ttcccaagg	acgcctc	ctctggact	4200
tttgcgtt	cttaaggcc	ccccagg	caacccgt	gtatctgg	tctgttggc	4260
actgtggatt	ctcaaggcc	tagaaccc	gttctgt	ctgggtcg	ggtgcaagg	4320
tgctgttgc	agctcttc	catatccca	gcccaccca	ggccaggccc	actccggct	4380
caccaccc	tgccagctt	ttgggtct	ccagccctc	cagaagocca	ccccactt	4440
cgcccaaccc	cgatcttca	atggggcc	agcgcttac	tagtctgg	cttgcgtt	4500
tgtgtact	tggaacggac	attttgc	ccttttct	tgtctataa	atgtggacag	4560
ttggacgttgc	tcaccaaga	gatgttgg	ggacatgt	ggggacat	gcaactc	4620
cgggtgtcc	gatgcac	acaatccat	atcggttt	ttcccttac	cacttggaa	4680
cccatgttag	aaaagtgt	atctgtact	gtgttcc	ccaacccat	gcctggatgt	4740
ccctgttgc	gcccattt	ttttttt	tctatgttca	gcaccact	ggcaccactt	4800
atttttaatc	accgaaagca	aaaaaaaa	aaaaaaa	aaaaaaa	aaaaaaa	4836

---

SEQ ID NO: 38	moltype = DNA	length = 2098			
FEATURE	Location/Qualifiers				
source	1..2098				
	mol_type = other DNA				
	organism = Homo sapiens				
SEQUENCE: 38					
ccacccatc	cccaaaaa	acccca	cccgatgc	gggtcaaa	60
ctcaggag	tggtgttct	ccca	ccaggcgt	accaggagg	120
gaaacctt	tgccatgtt	ttt	tttgcgtt	tttgcgtt	180
ggagtgtcc	gggggggg	cccg	ccgtgggt	cacacttcc	240
aggacagaca	ggggccct	tttt	tttgcgtt	ggacatgt	300
ttgttgagg	ttggagg	ggacatgt	ggggacat	acaggggcc	360
cagaccctcc	cccaggc	ccac	ccatcc	ggacatgt	420
ccaaaggcc	gtatcc	ggccccc	ccatcc	ggacatgt	480
cggccttc	ctctgc	ccgtt	ccgttac	ggatcc	540
gcaaaatc	cccg	ccgtt	ccgttac	ggatcc	600
ggggggcc	gggggggg	ccgc	ccggcc	ccgttac	660
aggaggaggc	ccga	ccgtt	ccatcc	ggatcc	720
ggggccccc	aa	cccg	ccatcc	ggatcc	780
ggcgctgt	gttgg	ccgtt	ccatcc	ggatcc	840
ccggcttgc	gggg	ccgtt	ccatcc	ggatcc	900
cccggtcc	tacc	ccgtt	ccatcc	ggatcc	960
ccggccggc	ccgg	ccgtt	ccatcc	ggatcc	1020
ctcagggtc	ccca	ccgtt	ccatcc	ggatcc	1080
gctctactt	gtgttca	cccc	ccatcc	ggatcc	1140
ccagggtcc	gaggagg	ccgtt	ccatcc	ggatcc	1200
ttttctggc	gggggat	ccgtt	ccatcc	ggatcc	1260
ctgttacaa	cctctg	ccatcc	ccatcc	ggatcc	1320
ccagcttca	ctgg	ccatcc	ccatcc	ggatcc	1380
cagcatgg	tgctac	ccgtt	ccatcc	ggatcc	1440
aggcaccc	ccttgg	ccatcc	ccatcc	ggatcc	1500
gggtcaccc	tttac	ccatcc	ccatcc	ggatcc	1560
cccatatc	ccatcc	ccatcc	ccatcc	ggatcc	1620
ccgccttc	ccatcc	ccatcc	ccatcc	ggatcc	1680
ctctggag	agggtgt	ccatcc	ccatcc	ggatcc	1740
cctgggac	accatt	ccatcc	ccatcc	ggatcc	1800
ggggactg	cagg	ccatcc	ccatcc	ggatcc	1860
cctcagtcc	tttac	ccatcc	ccatcc	ggatcc	1920
ggcaggc	gaacaa	ccatcc	ccatcc	ggatcc	1980

-continued

---

tgagaggccg	gggctcaggg	aaggggctgg	gatcggaaact	tcctgtctt	gtttctggac	2040
aacttcccc	ttctgttta	aagtttgtcg	attattttga	aaaaaaaaaa	aaaaaaaaaa	2098
<b>SEQ ID NO: 39</b>		<b>moltype = DNA</b>	<b>length = 5124</b>			
<b>FEATURE</b>		<b>Location/Qualifiers</b>				
<b>source</b>		<b>1..5124</b>				
		<b>mol_type = other DNA</b>				
		<b>organism = Homo sapiens</b>				
<b>SEQUENCE: 39</b>						
agaaggcagg	tggctgtat	gacagcacag	agctcaggaa	cgtgcctga	ggaccctggg	60
gcctacgagg	aggagaagag	ggcaggagct	ggtgggggtgc	ttgcagagac	cctgggtctc	120
tatcctgcga	taagcctgc	tgtctcctga	tatctgcacg	caggccctac	tgacacccc	180
aggcctgagt	gcaagcagag	accccacca	tcccaggccc	tggaggactg	gtccaccta	240
actggcagc	ccttggggc	ggcgcgtggc	ggtgcctca	cccaggcc	tgtgtctgc	300
atgcactgc	agectgcacat	cagcgtccat	ttgcagccct	gaaccatgt	ccagggcacc	360
ttggagccag	atggccccct	ctggggctgg	gactgggaca	tgacaaatga	ctggatagt	420
gtctgtctgg	ccctctggc	gctggctgt	gtggctccca	cagcgtggc	cttacactgg	480
tttggctccg	ggcaatcata	agaggccgca	gaaccgggt	ccacgcct	cggggctcaa	540
cctcatcagg	caggaggagc	tgagctggc	ctgcaacc	agtctaagg	cagtgtatggc	600
agcgaggggc	agagcccaag	gcaggggaaa	ccagagcccc	caggacggg	ccagcagagc	660
cctgtccctg	ctgcagcggc	gggggggggc	ctggccgcca	tggcccgct	tccactcaag	720
acggctgtcg	aggaggcccc	cagagaggca	tttagacagc	aacggggcag	tgccacccc	780
gccccgggg	gagcggggag	aaaggagcc	cccaggcc	gactgcct	cttggggcagg	840
agcgaagcag	ggggatgtc	ccccccctc	ctgatccact	tcaactctcg	gagccctggc	900
agcgaagcgg	aggcggagac	agggtgtc	agggcgtct	ctgcccaggc	cgcaggcccc	960
gccccggcaac	aggacactgg	ccccctggag	gccccggcgg	ggccctcg	ctcgatggg	1020
agaggcgggg	gccccgggg	gccccgggg	gtggatggd	gtggctcg	gagacagagc	1080
cgaaaaactgg	acccgcgtcg	cctggggcgc	gccccgggg	tgtggggcgc	ggtggacggg	1140
gccccggccc	ttggacggcc	cgccgcggc	cttccccac	gacccccc	cgccccaggag	1200
cccgactcc	cgccgcgtgc	cgccccccgc	gccccggc	ctgggttc	gacggaaaggc	1260
tctggggcc	agggtggctg	gagcgggg	gccccgggg	tccctgccc	cgggaggagc	1320
tggccctggc	tcagcggg	ggtccccgg	acccggag	ttggcccg	ccccagactc	1380
acccgcacct	ggctagagag	tccgcctaa	ggtcggcc	tctcgatcca	agggccgggt	1440
gccacagggg	cctacgatgc	cgccgaggcc	ggggctgaca	gtcccccaga	taacagtct	1500
gccccgtacc	tggggccac	ccggggcc	ggggatggc	agccgggtc	agccggccac	1560
agccgcgcgc	cctccggag	cggtgacgc	accacccatc	cctccatccc	cacccatgc	1620
ccccggccgg	gttcccccac	tgaaggcc	acttccct	tcccttcaga	ctttttggcc	1680
ctggaggtta	cccaggatcc	ttccgtggg	gaaaatctc	gagcggcc	agccccaa	1740
tcagcgttc	cccaactgtt	aacttcgt	ccagcctc	tcctagoccc	agccctgtct	1800
tcatccccca	gtcagcacc	aacctcaga	accacccatc	cctccatccc	cacccatgc	1860
ccagccccag	ctcccaactc	agctcaca	tcaacccatc	ccccagcccc	aagtcacat	1920
gcagccgca	ctccagcccc	agccccatc	ccagtc	ccctcacacc	ccccatcccc	1980
gccccatccc	cagtcacca	ccccatcc	agccca	caactccagc	ccctcacacc	2040
gccccatccc	cagccatcc	ccccatcc	ccccatcc	ccccatcc	ccccatcc	2100
gccccatccc	cagccatcc	ccccatcc	ccccatcc	ccccatcc	ccccatcc	2160
gccccatccc	ctggccccc	tgacgggt	aaacctcagg	agactgtgg	tctcccccagg	2220
cgttacccagg	agggggagg	ctcggcc	tggggaaagg	ttatgtccat	gttctttaga	2280
agccacccct	tcccccggca	agacgggg	ttccgggg	ttccgggg	gtttcccccgg	2340
agccccgtgg	gtcccccggc	ttccacac	tctgaggaca	gacacggccc	cttcttc	2400
gtggggacag	tcataggg	aggtcaggg	ggctgggt	aggctgggg	tcagccacag	2460
caaaagact	ccggacccaa	cggtcgc	agccca	ctcccccagg	cctaagggaga	2520
gagggaacaa	gggagaaaag	tctagacc	ccggatgc	ccggggcc	ccggggcc	2580
gcacagcccc	ccgcgcagag	ggccgcgt	ccccggcc	ctcccttc	gaggcgctca	2640
cagccgtac	cccaatc	gaaacgc	aggcgc	tcgcccc	tcggaggcag	2700
gagggtcagg	ccggccgc	ggggggac	caaggggagg	ccgggggg	ggggggcagc	2760
cctgcggcc	ccggccggc	gtcgcggc	gtcgcggc	agccca	ccggccggc	2820
tttctgcaga	ggccgggggg	ttgggggt	gtgggggg	ccgggggg	cagccccccgg	2880
gccccatccc	ccggccacggc	ggcagcc	ccggccggc	ttggac	cagttccctgc	2940
gacgtgtctgg	cttgcggcc	gcacacgg	gagccgg	ttggcggc	gacccatcg	3000
ctgtatgcgc	acaacatgt	cggtgtct	ggagacccgt	gtctctac	ccggctgtgc	3060
gccccatccc	cgacgcgc	cctcagc	ccggccggc	ggggccggc	gggtgtggc	3120
gtccctgtac	tcggccatc	ctaccagg	ggccgtc	ggctccccc	ggccctctgc	3180
gcgaggaggcc	tcctgcggc	gccccatgt	ccctgcctc	acctgc	ctgcgtgt	3240
tcacccatccc	ggagaaaccc	tggccggcc	tgacccatgt	gcccggag	cccccgttc	3300
ggggctgggg	tctctgc	atgcaca	acctgttt	gggggggg	atccgtgtct	3360
cccggtccaa	ggccgtctc	tccaaac	ccctctc	accacat	cccccgttc	3420
ggagccagg	tccggccat	cagcggcc	gagccat	caagctgt	cccccgttc	3480
gggtgtctca	tgccatcg	ggcgaatgc	tgtacag	ggagtgt	cccccgttc	3540
cajacgcctc	gacccacgc	ggccca	ccgcaggc	ttccctgt	gccccacgg	3600
ctgtggcctg	ccgtggggac	atctacgt	ccgggggt	cctcttac	cccccgttc	3660
ggtagacccc	cgtgtacgg	gttggggac	agtgccata	cagtgc	cccccgttc	3720
ccagcgacat	cgtggact	gggggttcc	tgtaccgtt	ccacccgt	cccccgttc	3780
gccccatccc	gtatgcgtac	aacacatgt	ccggccctc	gagcagggt	cccccgttc	3840
ccctgcggcc	ccccatccc	ctgcact	ccaccat	tactgc	cccccgttc	3900
acccatccc	cactgcac	ttcactgt	ctgggggg	tgcccgat	cccccgttc	3960
agctgcagcc	cttccccc	ggggcacc	gggtcctc	tccattc	ctgactctc	4020

-continued

---

cccccgtgagga	ccggctgcag	acctcaactct	gagtggcagg	cagagaacca	aagctgtttc	4080
gtctgttc	aggggagaccc	tcctggatgg	ggctgtgagag	gccggggctc	agggaaagggg	4140
ctggggatcg	aacttctgc	tcttggttct	ggacaacttt	cccccttgc	tttaaagggtt	4200
gtcgattt	ttaaagccca	gactccctca	gcctcttct	ccccctact	ccacaccagg	4260
actgtttct	gactcaattc	cgtaccaact	tacagaccc	ctcagettgc	tgacacccccc	4320
ctgtctgtgg	gactccctat	tccttagago	caggactga	tgcgcttcca	cagacaaggaa	4380
cttgggtcgt	tggagctctg	ctgagccgag	agaggagggg	gtaaaaaacaa	ttcacacttc	4440
ctatgtctg	ttagcaggac	agggagcaaa	aacgtcccc	ggcaacggccc	tcgctctgg	4500
gactttctgc	ctgtcttaag	gcctccccag	gtaccaacc	cgtatetac	tgggtctgtt	4560
tggcactgtg	gatttcataa	ggcttagaga	ccttgcctt	gaaactgtgc	cgtgtgtgca	4620
gcccgtgt	gttcacatcc	tccttatacc	cccaaggccac	accaggccag	gcccactccg	4680
ggctcaccac	ctctgcage	cttgggggc	tctcccgac	cctccagaag	cccacccac	4740
ttctcgeccaa	cccccgatct	cttaatgagg	cctgagctgc	accctagttc	tgccccttt	4800
tagctgtta	gactgttgacg	agacatttg	cttccctt	tccttgccta	taaaatgtgg	4860
acagtggacg	tctgtcacc	aagagatgt	tgggagacaa	gateacagct	atgacacact	4920
cgcacggtgt	ccaggatgca	cgcacaaatc	catgtatgc	tttctccct	tafcacactt	4980
gaaaaccatg	ctagaaaagt	gaatacatct	gactgtgctc	cactccaacc	tcagccctgg	5040
atgtccctgt	ctggccctt	tttctgtttt	ttattctatg	ttcagcacca	ctggcaccaa	5100
atacatttta	attcaccgaa	agca				5124

---

SEQ ID NO: 40	moltype = DNA	length = 2989				
FEATURE	Location/Qualifiers					
source	1..2989					
	mol_type = other DNA					
	organism = Homo sapiens					
SEQUENCE: 40						
ccacacctc	cccaacccca	accccaagccg	catccccctgc	ccccagctgac	gggtcaaaagc	60
ctcaggagag	tgtggctc	cccaggcgct	accaggagggg	gcagggtctca	gccagctggg	120
gaaaccttat	tgccatgg	cttagaaged	accccttccc	caggcaagac	aggccccaag	180
ggagtggtcc	gagggcggtt	cccgggagcc	ccgtgggtcc	cagcacttcc	acacactctg	240
aggacagacca	ccggccctct	tcttcagtt	ggacagtc	atggagacagg	acagggggcc	300
tgttgtggc	tggaggatcg	ccacagccaa	gaagctccga	gaccaacgga	tcgcccagcc	360
caagacccccc	cccaggctca	agaggagagg	gaaccaggga	aaaaagtctt	gaccggctgc	420
cccaaggccg	gatgcccagg	ggcccccgcac	agccccccgc	gcagaggccg	cctggcccg	480
ccgcctctc	ctctcgagg	cgctcacago	cggtatccccc	gtacggaaa	cgccgcagg	540
gcgaaatcgc	cccgagatcg	gagcaggagg	tcaggccgc	cgccctgggg	gaccctcaag	600
gggaggcgcc	ggggggagggg	ggcagccctg	ccggccgcag	cgggggcgctc	acggaaaagc	660
aggaggaggc	ccggaaatctc	atgggtttt	tgcaagggcc	cgggggttgg	gggggtgtgg	720
aggggccccc	gaagggccagg	tccggggccc	tggagccccc	cacggggcga	gcctcgccgc	780
gggggtctgg	ctctgggg	tgcctgg	tgctggcc	tgeccagcag	cacggagac	840
ccggcctgc	gcaggagacc	tacgcgtca	tgacgcacaa	cctgctgcga	gtgtgggg	900
accctgtct	ctaccggcc	ctgagcgcgg	ccgaccgcga	gcbcattc	agcctgcgga	960
ccggccgggg	ccgggggtgt	ctgggcgtcc	tcgtactgc	caggggggcc	1020	
gtcaggaggt	ccccaggccc	ctcggtggcc	ggagectct	ggggccccc	ttgtgtccct	1080
gcctctactt	gogcaactgc	atgtttca	ccccccggag	aacacctggc	ggccctctgac	1140
ccagggtccc	gaggaggccc	cgcttcgggg	ctgggtctc	tgcaaccatgc	acaactacct	1200
gtttctggcc	ggggggatct	gtgggtccgg	tgccaaaggcc	gtctgtccca	acgaggctt	1260
ctgttacaa	cctctgacca	acatctggag	ccagggtccg	ccatgcacgc	aggccgcagc	1320
ccagctcaag	ctgggtggcc	tggagggt	gtctatgc	atcggtggc	aatgcctgt	1380
cagcatggag	tgtctacgacc	cgcgaacaga	cgcttgacc	ccacgcgcgc	cactccccgc	1440
aggcacttc	ctctgtggcc	acggggctgt	ggctgtccgt	ggggacatct	acgttacccgg	1500
gggttaccc	ttcttacggcc	tgctcaggta	cgccccgg	aaggatgtt	gggacagatg	1560
cccatatacg	gcacccccc	ggcgttccag	cgacatcg	gactgggg	gttctctgt	1620
ccgttgcac	ctgtcgccgg	gctggggcgc	cgccgtatgc	cgatgaccgg	1680	
ctctctggac	agggggtct	ccctggccc	gccccccc	gccccactgc	actgcacac	1740
cctggccac	accattttact	gcctcaacc	ccagggtact	gcccattca	cggtctctgg	1800
ggggactggc	caggccagg	ccaaaggatct	gcagcccttc	cccttgggg	gcacgggggt	1860
cctcagttca	ttcatacttca	ctctggcccc	tgaggaccgg	ctgcagac	cactctgagt	1920
ggcaggcaga	gaaccaaagg	tgctccagg	agacccct	gggatggcc	1980	
tgaggggttag	gggtctcagg	aagggggg	gatcggaaact	tcctgtctt	gtttctggac	2040
aactttcccc	ttctgtttta	aaagggtgt	attatttga	ggcccaact	ccctcagcc	2100
ctttctggcc	ctcaactccac	accaggactg	tttctgtact	caattcogta	cctacttaca	2160
gaccctctca	gtttctgtac	acccccc	ctgtgggact	cccttattccc	tagaggcagg	2220
gactgtatcg	tctccacaca	caaggactg	gtctgtgtca	gccggagagag	2280	
gagggggtag	aaaacattca	cacttctt	gtctgtca	caggacagg	agcaaaaacg	2340
tcccccaggca	acgccttc	ctctggact	ttctgtctgt	cctaaggct	ccccaggatc	2400
caacccctgt	gtatatggg	tctgtttgg	actgtggatt	ctcaaggccc	tagaacctt	2460
gcctctgaaa	ctgggtcgct	ggtgagccc	tgctgtctgc	agctctgtcc	cataacccca	2520
gcccacacca	ggccacggccc	actccgggt	caccaccc	tgccagccctg	tggggctctc	2580
ccagccccc	cagaaggccca	ccccacttct	cgccaaacccc	cgatcttca	atgaggctg	2640
agcgtaacc	tagttctgc	ccttttttgc	tgtgttagact	tggaacggac	atttgactt	2700
cctttctct	tgtctataaa	atgtggacag	tggacgtctg	tcacccaaga	gagttgtggg	2760
agacaagatc	acagctatga	gcacccgc	cggtgtccag	gatgcacac	acaatccatg	2820
atgcgttttc	tcccttta	cactttgaaa	cccatgttag	aaaagtgaat	acatctgact	2880
gtgtccact	ccaacccca	gcctggatgt	ccctgtctgg	gccctttttc	tgttttttat	2940
tctatgtca	gcacccactgg	caccaatac	attttaat	accgaaagc		2989

---

-continued

---

```

SEQ ID NO: 41      moltype = DNA  length = 4836
FEATURE          Location/Qualifiers
source           1..4836
                 mol_type = other DNA
                 organism = Homo sapiens
SEQUENCE: 41
actgccagcc tgcccatcagg ccttattgc ageccctgaac catgatccag ggcacccctgg 60
agccagatgg tccccctctgg ggctgggact gggacagtga caatgactgg gatagtgtcg 120
tgtgtggccct ctggcgctg gctgtggctg ctggccacagc gctggcccta cactggttg 180
getccgggca cgcataaagag gccggaaagaad cgggtgtccac agccctcggg gctcaacetc 240
atcaggcagg aggagctgag ctggccctgc aaccgaagt tcaggtaagt gatggcageg 300
agggggcagag cccaggcag gggaaaccag agccccccagg acgcggccag cagagccctg 360
tcctgtcgc agcgcggggc gggggccctgg ccggccatggc ccggccctcca ctcaagacgg 420
ctgtcgaggg gggcccgaga gaggcattag gacagcaacag gggcagtggc accccccggg 480
ccccccgagc ggaaggaaag gagccctcca ggcggggcac tcggccctctg ggcaggagcg 540
aagcaggggg gatgtccggc cccctcttga tccacttcac tccttcggagc cctggcagcg 600
aagcggaggc ggagacagggt ggtgtcaggg cgttctctcg ccagggccga ggcggggccgg 660
ggcaacacgga cactggcccc tggcggccgg ggcggggccctgg ctcgggtctcg atggggagag 720
gcccggggccg gggggggccgg atggacgtg gtcggggaga gagagccccc cgccccccgg 780
aactggaccc gtcggccctg ggcgcgggg ggagcgtgtg ggacgggggtg gacggggccg 840
ccgccttgga cgcggccacgg cggggccctc ccacaggacc cccactccggc caggagcccg 900
caactccggc gtcggccctg ccccccggcgg tcaagccctgg gtcttcagacg gaaggctctg 960
ggggcaagggg tggctggagc gggggggctt cgggggttccc tggggggggg ggaggctggc 1020
cctgggttag cagggggggtc ccggggccccc ggagctttgg ecacgeccca gactccacgc 1080
gcccctggct agagagtccg ctcctaaaggtd gcccacttc ttcccaagggg ccgggttggca 1140
caggggccctt gatgtccggc gggggggggg ctgacagtcg cccagatatac agtcttcgg 1200
ctgacccctggg gggcccccggg ccccccggago aagcaaaaggc ggcttcggacc ggccacagcc 1260
ggcgccctc ccggggccgt gagectcgcc cgggctccgc ctcccccggcc gcacgtcccg 1320
gcccgggggtt cccaccttggaa gcccctactc tcccttccttc ttcaagactt ttggccctgg 1380
agtttccatc ggttccatcc gttggggggaaa atcttcagatc ggcggccaggg ccaagttcg 1440
cctcagccca agtcttaact tcagttccag ctccttcgttcc acggccccc ctgggttcat 1500
ccccccagtc agcacaacc tcagccatcca cttcaacccctt atccccccacc tcagccccc 1560
ccccagtc aacccatcgat ccaacttcaa ccccaagcccc acggcccaatg ccacgtcgag 1620
ccggcaacttcc acggcccccgg ccgttccatcc tcccaacccca tccccccggcc 1680
taaccccccgtt cccaaaccctt ggccttccaaacg cagttccaaac tccagcccttta accccggcc 1740
catccccccggc cttaaaccctt gttcccaaccc cagcccttcaag cccagttccca actccagccc 1800
caaccccccggc cgcataccctt gccccagccc ccacccctcgc cccaaacccca accccggccg 1860
catccccccggc cccatcgatc gggcttccaaaggc ctccaggatggg tggggcccttc cccaggccgt 1920
accaggaggcc gggatgttca gccaggtttttt gaaacccctt ggccttccat tttagaaagcc 1980
accccttccggc caggcaagac acggcccaag ggagtgttcc gagggggggtt ccggggggcc 2040
ccgtgggtcc cagccacttcc acacactctg aggacagaca cggcccccctt ttttcgttgg 2100
ggacagtccat aggccgggggg tgggttgggg tggggggggcc ccacagccaa 2160
gaaggttccg gaccacccgg tggggccccc cccaggccca agaggaggagg 2220
gaaccagggg gaaaaggctta gaccggctgc cccaaaggccc gatggccccc ggccggccac 2280
agccccccggc gcaaggccgg cctggccccc cggcccttcgc ctcttcgggg cgttcacagc 2340
cggttccatcc gatgtccggaa cgcggccatgg cccggatccgg gaggaggagg 2400
tcaggccggc cccctccgggg gacccttcaag ggggggggggg gggcccccctg 2460
ccggcccgccg cggggggccgt acggaaaaggc aggaggaggcc cgggaagctc atgggtttcc 2520
tgcagaggcc cgggggggttgg ggggttgggg agggggccccc gaagccccc tccggggccc 2580
tggagggccg caccggggca gcccctgggg ggcggccgttgc cctggggact tgcctggacg 2640
tgctggccctt tggccatccggc cacggagacg cccggccctggcc gcaggagacc taegcgctga 2700
tgagcacaatc cttgttccggc gtgttgggg acggccgttcc taccggccgg ctgagccgg 2760
ccgaccggca ggcgttccatc acggccatcc acggccatcc acggccatcc acggccatcc 2820
tcgtactggc caccgttccatc cggggggggcc gtcagggttcc cccaggggcc ctggccatcc 2880
ggagcccttcc cggggggccgg ctgttccatcc ggcgttccatc ggcgttccatc atgtgttccaa 2940
ccccccgggg aacacccttgg gggcccttgc ccagggtccc gaggaggccc cgttccgggg 3000
ctggccgttcc tgcaccatgc acaactaccc ttttctggcc gggggccatcc gtggccatcc 3060
tgccaaaggcc gtctgttccaa acgggttccatc ctgttccatcc acatctggag 3120
ccagggttccg cccatccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3180
gtcttatggc atccgttccggc aatgttccatc cggatgttccatc gtcgttccatc ggcgttccatc 3240
ccgttccatcc cccatccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3300
ggccctggccgt gggggccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3360
caggcccttcc acggccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3420
cgacatcggttccg cccatccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3480
ccggccgttccatcc cccatccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3540
gcccggccccc gccccccatcc acgttccatcc acggccatcc acggccatcc acggccatcc 3600
ccagggttccatcc acggccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3660
gcaggcccttcc acggccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3720
tgaggacccgg ctggccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3780
cttccatccggc acacccttcc gggatggggcc tggatggggcc gggccatcc acggccatcc 3840
gatccggaaact ttcgttccatcc gtttctggac aatttccccc ttcttccatcc acggccatcc 3900
attatccatcc acggccatcc acggccatcc acggccatcc acggccatcc acggccatcc 3960
tttccatccact caattccatcc acggccatcc acggccatcc acggccatcc acggccatcc 4020
ctgtggggact cccttccatcc acggccatcc acggccatcc acggccatcc acggccatcc 4080
gtctcgatcc gtcgttccatcc acggccatcc acggccatcc acggccatcc acggccatcc 4140

```

-continued

---

gctctgtcag	caggacaggg	agcaaaaacg	tccccaggca	acgcctcgc	ctctggact	4200
ttctgcgtt	cctaaggcc	ccccaggta	caaccccgta	gttatctgg	tctgtttggc	4260
actgtggatt	ctaaggccc	tagaacccctt	gcctctgaaa	ctggtcggct	ggtgccagccc	4320
tgtgtctgc	actctctgcc	cataccccc	gccccaccca	ggccaggccc	actccggct	4380
caccacccctc	tgcagcctt	tggggctc	ccagccctc	cagaagccca	ccccacttct	4440
cgcacaacccc	cgatctctaa	atgaggccctg	agggtcaccc	tagttctgg	ccttttttagc	4500
tgtgttagact	tggacgagac	atttgactt	cttttctct	tgtctataaa	atgtggacag	4560
tggacgttctg	tcacccaaga	gagtttgtgg	agacaagatc	acagctatga	gcacctcgca	4620
cggtgtccag	gatgcacagc	acaatccatg	atgcgtttc	tccctta	cactttgaaa	4680
cccatcttag	aaaatgtaat	acatctact	gtgctccact	ccaaacctca	gccttgatgt	4740
cctgtctgg	gcccctttt	tgtttttt	tctatgttca	gcaccactgg	caccaaatac	4800
attttaattc	accgaaagca	aaaaaaaaaa	aaaaaaaaaa			4836

SEQ ID NO: 42            moltype = DNA    length = 2098  
 FEATURE                Location/Qualifiers  
 source                1..2098  
 mol\_type = other DNA  
 organism = Homo sapiens

SEQUENCE: 42

ccacctcagc	cccaaaaa	accccagccg	atccccctgc	cccaagctgac	gggtcaaa	60
ctcaggagag	tgtggcttc	cccaggcgt	accaggagg	gcaggcttca	gccagctgg	120
gaaacccat	tttgcgtt	ctttagaagcc	accccttccc	caggcaagac	aggccccaa	180
ggagtgtccc	gagggggg	ccccgggg	ccccgggg	ccgtgggtcc	cggacttcc	240
aggacagaca	ccccccctc	tcttc	ggacactt	ggggaggat	acagggggcc	300
tgttgtggc	ttgggg	ccacagccaa	gaagctccg	gaccaacgg	tcgcccagcc	360
caaaaa	ccccccccc	agaggaggg	gaaccaggaa	aaaaatctt	gaccggctgc	420
cccaagccgc	gtgtccgg	ggccccggc	agccccccgc	gcagggccg	ctggccccc	480
cggcccttc	ctctgcgg	cgctcacago	cggtacccca	gctacggaaa	cgcagcagg	540
ggaaaatcgc	ccccggctc	gagcaggagg	tcaggccgg	cgccctgggg	gaccctcaag	600
ggggaggcc	gggggggg	ggcagccct	ccggccgg	cgggggcc	acggaaa	660
aggaggaggc	ccggaggc	atgggtttt	tgcaggagg	cggggggtt	gggggtttgg	720
aggggccccg	aaagccccc	tccggggcc	tggagccgc	cacggccgca	gcccctggcc	780
ggggggcttga	cttggggcgt	tgcctggac	tgcctggctt	tgccca	cacggagac	840
ccggccttgc	gcaggagacc	tacgcgtt	tgagcgtt	tgctgttgg	900	
accctgtctt	ctaccggcc	ctggcgttgc	ccgacccgg	gcccatttcc	acccgttgg	960
ccggccgggg	ccgggggg	ctgggggtt	tgcgtactt	cagcccttac	cagggggg	1020
gctcagggtt	ccccaggcc	ctctggcc	ggagccctt	gcccggggcc	ctgtgttcc	1080
gcctctactt	gcgcaccc	atgtttt	ccccgggg	aacaccc	ggcccttgc	1140
ccaggtgtcc	gagggggg	cgcttgggg	ctgggttcc	tgcacccat	acaacttac	1200
gtttctggc	ggggggatcc	gtgggttcc	tgccaa	gtctgttca	acggatgtt	1260
ctgttacaa	cctctgtt	acatctgg	ccaggttcc	ccatgttca	aggccgg	1320
ccagctcaag	ctggggccc	tggacggg	gtcttat	tgcgttgg	aatgttctt	1380
caacatgg	gtctgttcc	cgccacac	cgccctgg	ccacggccgc	cactccccc	1440
aggcaccc	cctgtggcc	acggggctgt	ggcttggcc	ggggacat	acgttacccg	1500
gggttaccc	tttctggcc	tgcctgg	caagccctg	aaggatgtt	gggacgggt	1560
ccctatacgt	cccaaggcc	ggcgttcc	cgacatgt	gactgggg	gtttctgt	1620
ccgcttgcac	ctgttgcgg	gtgttgcgg	cgccgttgc	cgatcaca	tgttgcgg	1680
ctctgggg	agggttgc	ccctggcc	gcccggcc	gcccacttgc	actgcacac	1740
cctgggcaac	accatttt	gcctcaacc	ccaggttact	gcaccc	cggttcttgc	1800
ggggacttgc	cgatcc	ccaaaggact	gcagccctt	cccttgggg	gcacccgg	1860
cctcagtca	tccatctt	ctctggccc	tgaggaccc	ctgcagact	cacttgcgt	1920
ggcaggcaga	gaaccaaa	tgcgttgc	ctctccagg	agacccctt	ggggatggcc	1980
tgagaggccg	gggctcagg	aagggttgc	gtatggaa	tcttgcgtt	gttttgcgac	2040
aacttcccc	ttctgttta	aagggttgc	attat	tttgcgtt	aaaaaaaaaa	2098

SEQ ID NO: 43            moltype = AA    length = 885  
 FEATURE                Location/Qualifiers  
 source                1..885  
 mol\_type = protein  
 organism = Homo sapiens

SEQUENCE: 43

MIQGTLEPDG	PLWGWDWSD	NDWDSAVLAL	LALAVVAATA	LALHWFGSH	DQEAAEPVST	60
ALGAQPHQAG	GAEELALQPKS	KVSDGSEGQS	PGGGKPEPPG	RQQSPVPAA	APGGGLAAMA	120
RLPLKTAVEE	AREALLGQQR	GSATPAAPRA	EGKEPPRPGT	ALLGRSEAGG	MSAPLLIHT	180
PRSPGSEAEA	ETGGVRASSR	QAAGPGQQD	TGPWQAGAGP	SGMSGRGRGR	RRRMDAGSGD	240
RARRPRKLDP	LRLGAAGSVW	DAVDGAAALD	AHARGLPTGP	PLAQEPALPA	LPAPRALQPG	300
SQTEGSGAKG	GWSREASGPV	APGGGWPWV	REVPGTRSRFG	PAPDSTRPW	ESPPQGRPLS	360
SQGPGATGAY	DAGEAGADSS	RDNSPAADLG	PTRPPEQAKP	AAAGHSRAPS	RSREPRPRS	420
SPPAAPGPGF	PPEALTLPS	SDFLPLEVTC	DPSVGENLRA	APAPSSASAQ	VLTSPASVL	480
APALASSPSS	APTSATTSTS	SPTSAPAP	TSAPTSTPAP	APSPAAAATP	APAPPVPTL	540
TPPSPALTPV	PTPALSPAP	PALTPAASP	LTPVPTPALS	PAPTPAPTA	ASPAAPAPTS	600
PTPTPAASPA	PADGSKPQES	VALPRRYQEG	QVSASWGNLI	AMVLRSHPFP	RQDRPQGSVP	660
RAVPGSPVGP	STSTHSEDRH	GPSSSVGTVI	GTGTGGLV	GGQPQPRSE	TNGSPSPDPP	720
PGLRGEGTRE	KSLDPLPQAA	MPRGPAQPPA	QRPPGPAASS	SARRSQVVPQ	LRKRSRCEIA	780
PSSEQEVRPA	ASGDPQGEAP	GECCSPAGRS	GALTEKQEEA	RSSWCFCRGP	GVGGWWRGPG	840
SPAPGPWSPPP	RRQPCGGGWT	WAWAWTCWPL	PSSTESPAWR	RRPRT		885

---

-continued

---

```

SEQ ID NO: 44      moltype = AA  length = 244
FEATURE          Location/Qualifiers
source           1..244
mol_type = protein
organism = Homo sapiens
SEQUENCE: 44
MVLRSHPFPR QDRPQGSVPR AVPGSPVGPS TSTHSEDRHG PSSSVGTIVIG TGTGGLVEAG 60
GQPQPRSET NGSPSPDPPP GLRGEGTREK SLGPLPQAAM PRGPAQPAQ RPPGPAASSS 120
ARRSQPVQPL RRKSRCIEIAP SSEQEVRPAA SGDPQGEAPG EGGSPAGRSG ALTEKQEEAR 180
SSWCFCRGPG VGGWWRGPGS PAPGPWSPPR RQPCGGGWTW AWAUTCWPLP SSTESPAWR 240
RPTR                                         244

SEQ ID NO: 45      moltype = AA  length = 145
FEATURE          Location/Qualifiers
source           1..145
mol_type = protein
organism = Homo sapiens
SEQUENCE: 45
MPPGPAQPA QRPPGPAASS SARRSQPVQ LRKRSRCIEIA PSSEQEVRPA ASGDPQGEAP 60
GEGGSPAGRS GALTEKQEEA RSSWCFCRG GVGGWWRGPG SPAPGPWSPP RRQPCGGGWT 120
WAUTCWPLP PSSTESPAWR RRPTR                                         145

SEQ ID NO: 46      moltype = AA  length = 975
FEATURE          Location/Qualifiers
source           1..975
mol_type = protein
organism = Homo sapiens
SEQUENCE: 46
MIQGTLEPDG PLWGWDWDSD NDWDASVLAL LALAVVAATA LALHWFGSGH DQEAAEPVST 60
ALGAQPHQAG GAEALALQPKS KVSDGSEGQS PGQGKPEPPG RGQQSPVPAA APGGGLAAMA 120
RLPLKTAVEE ARREALGQQR GSATPAPRA EGKEPPRPGT ALLGRSEAGG MSAPLILHFT 180
PRSPGSEAEA ETGGVRASSR QAAGPAGQOD TGPWQAGAGP SGSMGRGRGR RRRMDAGSGD 240
RARRPRKLDP LRLGAAGSVW DAVDGAAALD AHARGLPTGP PLAQEPALPA LPAPRALQPG 300
SQTEGSGAKG GWSREASGVP APGGGWPWVS REVPGTRSFQ PAPDSTRPWL ESPPOGRPLS 360
SQGPGATGAY DAGEAGADSS RDNSPAADLG PTRPPEQAKP AAAGHSRAPS RSREPRPRSA 420
SPPAAPPGPF PPEALTLPPSP SDFLPLEVTQ DPSVGENLRA APAPSSASAQ VLTSAPASVL 480
APALASPSS APTSATSTS SPTSPAPAP TSAPTSTPAP APSPAAAATP APAPVPVPTL 540
TPPPSPALTPV PTPLSPAPT PALTTPAASP LTPVPTPAPL PAPTPAPTPA ASPAPAPTS 600
PTPTPAASPA PADGSKPQES VALPRRYQEG QVSASWGNLI AMVLRSHFPF RQDRPOGSVP 660
RAVPGSPVGP STSTHSEDRH GPSSSVGTVI GTGTGGLVEA GGQPQPRSSE TNQSPSPDPP 720
PGLRGEGTRE KSLDPLPQAA MPPGPAQPPA QRPPGPAASS SARRSQPVQ LRKRSRCIEIA 780
PSSEQEVRPA ASGDPQGEAP GEGGSPAGRS GALTEKQEEA RKLMVFLQRP GGWGVVEGPR 840
KPSSRALEPA TAAALRRRLD LGSCLDVLAF AQHQHGEPLA QETYALMSDN LLRVLGDPCL 900
YRRLSAADRE RILSLRTGRG RAVLGVLVLP SLYQGGRSGL PRALVARSLL RRPLCPCLYL 960
RTCMCSTPGR TPGGP                                         975

SEQ ID NO: 47      moltype = AA  length = 334
FEATURE          Location/Qualifiers
source           1..334
mol_type = protein
organism = Homo sapiens
SEQUENCE: 47
MVLRSHPFPR QDRPQGSVPR AVPGSPVGPS TSTHSEDRHG PSSSVGTIVIG TGTGGLVEAG 60
GQPQPRSET NGSPSPDPPP GLRGEGTREK SLGPLPQAAM PRGPAQPAQ RPPGPAASSS 120
ARRSQPVQPL RRKSRCIEIAP SSEQEVRPAA SGDPQGEAPG EGGSPAGRSG ALTEKQEEAR 180
KLMVFLQRPQ GEGGVVVEGPRK PSSEALEPAT AAALRRRLDL GSCLDVLAFA QHQHGEPLAQ 240
ETYALMSDNLL RVLVGDPCLY RRLSAADRE ILSLRTGRGR AVLGVLVLP LYQGGRSGLP 300
RALVARSLLRPLCPCLYL TCMCSTPGR PGPP                                         334

SEQ ID NO: 48      moltype = AA  length = 235
FEATURE          Location/Qualifiers
source           1..235
mol_type = protein
organism = Homo sapiens
SEQUENCE: 48
MPPGPAQPA QRPPGPAASS SARRSQPVQ LRKRSRCIEIA PSSEQEVRPA ASGDPQGEAP 60
GEGGSPAGRS GALTEKQEEA RKLMVFLQRP GGWGVVEGPR KPSSRALEPA TAAALRRRLD 120
LGSCLDVLAF AQHQHGEPLA QETYALMSDN LLRVLGDPCL YRRLSAADRE RILSLRTGRG 180
RAVLGVLVLP SLYQGGRSGL PRALVARSLL RRPLCPCLYL RTMCSTPGR TPGGP                                         235

```

---

1. A method of treating a subject with a therapeutic agent that treats or inhibits hearing loss, wherein the subject has hearing loss, the method comprising the steps of:
  - determining whether the subject has a Kelch Domain Containing 7B (KLHDC7B) missense variant nucleic acid molecule encoding a KLHDC7B predicted loss-of-function polypeptide by:
    - obtaining or having obtained a biological sample from the subject; and
    - performing or having performed a sequence analysis on the biological sample to determine if the subject has a genotype comprising the KLHDC7B missense variant nucleic acid molecule encoding the KLHDC7B predicted loss-of-function polypeptide; and
    - administering or continuing to administer the therapeutic agent that treats or inhibits hearing loss in a standard dosage amount to a subject that is KLHDC7B reference; and
    - administering or continuing to administer the therapeutic agent that treats or inhibits hearing loss in an amount that is the same as or greater than a standard dosage amount to a subject that is heterozygous or homozygous for the KLHDC7B missense variant nucleic acid molecule;
  - wherein the presence of a genotype having the KLHDC7B missense variant nucleic acid molecule encoding the KLHDC7B predicted loss-of-function polypeptide indicates the subject has an increased risk of developing hearing loss,
  - wherein the KLHDC7B missense variant nucleic acid molecule encodes KLHDC7B K181fs or KLHDC7B G302fs, and
  - wherein the therapeutic agent comprises an antioxidant, a calcium-channel blocker, an anti-inflammatory drug, an apoptosis inhibitor, D-methionine, ebselen, N-acetylcysteine, lipoic acid, a combination of ebselen and allopurinol, resveratrol, a neurotrophic factor, a caspase inhibitor, a copper transport inhibitor, or a micronutrients with an antioxidant vitamin.
- 2-3. (canceled)
4. The method according to claim 1, wherein the KLHDC7B missense variant nucleic acid molecule encoding the KLHDC7B predicted loss-of-function polypeptide is:
  - a genomic nucleic acid molecule having a nucleotide sequence: lacking a guanine at a position corresponding to position 2,807 according to SEQ ID NO:1; or lacking a guanine at a position corresponding to position 3,170 according to SEQ ID NO:1; or
  - an mRNA molecule having a nucleotide sequence: lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:4; lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:6; lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:4; or lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:6; or
  - a cDNA molecule produced from an mRNA molecule, wherein the cDNA molecule has a nucleotide sequence: lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:12; lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:14; lacking a guanine at a position

corresponding to position 1,036 according to SEQ ID NO:12; or lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:14.

5. (canceled)
6. The method according to claim 1, wherein the sequence analysis comprises sequencing at least a portion of the nucleotide sequence of the KLHDC7B mRNA molecule in the biological sample, wherein the sequenced portion comprises: positions corresponding to positions 672-673 according to SEQ ID NO:28, or the complement thereof; positions corresponding to positions 672-673 according to SEQ ID NO:30, or the complement thereof; positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or the complement thereof; or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof;
 

wherein when the sequenced portion of the KLHDC7B mRNA molecule in the biological sample comprises: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, then the KLHDC7B mRNA molecule in the biological sample is a KLHDC7B missense variant mRNA molecule encoding a KLHDC7B predicted loss-of-function polypeptide.
7. The method according to claim 1, wherein the sequence analysis comprises sequencing at least a portion of the nucleotide sequence of the KLHDC7B cDNA molecule produced from the mRNA molecule in the biological sample, wherein the sequenced portion comprises: positions corresponding to positions 672-673 according to SEQ ID NO:36, or the complement thereof; positions corresponding to positions 672-673 according to SEQ ID NO:38, or the complement thereof; positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or the complement thereof; or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof;
 

wherein when the sequenced portion of the KLHDC7B cDNA molecule in the biological sample comprises: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, then the KLHDC7B cDNA molecule is a KLHDC7B missense variant cDNA molecule encoding a KLHDC7B predicted loss-of-function polypeptide.
8. (canceled)
9. The method according to claim 1, wherein the sequence analysis comprises:
  - a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B mRNA molecule that is proximate to: positions corresponding to positions 672-673 according to SEQ ID NO:28, positions corresponding to positions

672-673 according to SEQ ID NO:30, positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34;

- b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B mRNA molecule corresponding to: positions 672-673 according to SEQ ID NO:28, positions 672-673 according to SEQ ID NO:30, positions 1,035-1,036 according to SEQ ID NO:32, or positions 1,035-1,036 according to SEQ ID NO:34; and
- c) determining whether the extension product of the primer comprises: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34.

**10.** The method according to claim 1, wherein the sequence analysis comprises:

- a) contacting the biological sample with a primer hybridizing to a portion of the nucleotide sequence of the KLHDC7B cDNA molecule that is proximate to: positions corresponding to positions 2,806-2,807 according to SEQ ID NO:35, positions corresponding to positions 672-673 according to SEQ ID NO:36, positions corresponding to positions 672-673 according to SEQ ID NO:38, positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42;
- b) extending the primer at least through the position of the nucleotide sequence of the KLHDC7B cDNA molecule corresponding to: positions 672-673 according to SEQ ID NO:36, positions 672-673 according to SEQ ID NO:38, positions 1,035-1,036 according to SEQ ID NO:40, or positions 1,035-1,036 according to SEQ ID NO:42; and
- c) determining whether the extension product of the primer comprises: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42.

**11.** The method according to claim 1, wherein the sequence analysis comprises sequencing the entire nucleic acid molecule.

**12.** (canceled)

**13.** The method according to claim 1, wherein the sequence analysis comprises:

- a) amplifying at least a portion of the nucleic acid molecule that encodes the KLHDC7B polypeptide, wherein the portion comprises: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42 according to SEQ ID NO:34;

NO:32, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof;

- b) labeling the amplified nucleic acid molecule with a detectable label;
- c) contacting the labeled nucleic acid molecule with a support comprising an alteration-specific probe, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleic acid sequence of the amplified nucleic acid molecule comprising: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:28, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:30, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:32, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:34, or the complement thereof; and
- d) detecting the detectable label.

**14.** The method according to claim 1, wherein the sequence analysis comprises:

- a) amplifying at least a portion of the nucleic acid molecule that encodes the KLHDC7B polypeptide, wherein the portion comprises: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof;
- b) labeling the amplified nucleic acid molecule with a detectable label;

- c) contacting the labeled nucleic acid molecule with a support comprising an alteration-specific probe, wherein the alteration-specific probe comprises a nucleotide sequence which hybridizes under stringent conditions to the nucleic acid sequence of the amplified nucleic acid molecule comprising: a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:36, or the complement thereof; a CG dinucleotide at positions corresponding to positions 672-673 according to SEQ ID NO:38, or the complement thereof; an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:40, or the complement thereof; or an AG dinucleotide at positions corresponding to positions 1,035-1,036 according to SEQ ID NO:42, or the complement thereof; and
- d) detecting the detectable label.

**15.** The method according to claim 14, wherein the nucleic acid molecule in the sample is mRNA and the mRNA is reverse-transcribed into cDNA prior to the amplifying step.

**16-60.** (canceled)

**61.** The method according to claim 1, wherein: the anti-inflammatory drug comprises a steroid, the neurotrophic factor comprises T-817MA, the caspase inhibitor comprises z-DEVD-fmk, and the copper transport inhibitor comprises cimetidine or copper sulphate.

**62.** A method of treating a subject having hearing loss with a therapeutic agent that treats or inhibits hearing loss, wherein the subject has been determined to have a Kelch Domain Containing 7B (KLHDC7B) missense variant nucleic acid molecule encoding KLHDC7B K181fs or KLHDC7B G302fs, and wherein the therapeutic agent comprises an antioxidant, a calcium-channel blocker, an anti-inflammatory drug, an apoptosis inhibitor, D-methionine, ebselen, N-acetylcysteine, lipoic acid, a combination of ebselen and allopurinol, resveratrol, a neurotrophic factor, a caspase inhibitor, a copper transport inhibitor, or a micro-nutrients with an antioxidant vitamin.

**63.** The method according to claim 62, wherein the KLHDC7B missense variant nucleic acid molecule is: a genomic nucleic acid molecule having a nucleotide sequence: lacking a guanine at a position corresponding to position 2,807 according to SEQ ID NO:1; or lacking a guanine at a position corresponding to position 3,170 according to SEQ ID NO:1; or an mRNA molecule having a nucleotide sequence: lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:4; lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:6; lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:4, or lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:6; or a cDNA molecule produced from an mRNA molecule, wherein the cDNA molecule has a nucleotide sequence: lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:12; lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:14; lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:12, or lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:14.

**64.** The method of claim 62, wherein the subject has conductive hearing loss, sensorineural hearing loss, or neural hearing loss.

**65.** The method of claim 62, wherein the subject has been determined to be heterozygous for the KLHDC7B missense variant nucleic acid molecule encoding KLHDC7B K181fs.

**66.** The method of claim 62, wherein the subject has been determined to be heterozygous for the KLHDC7B missense variant nucleic acid molecule encoding KLHDC7B G302fs.

**67.** The method of claim 62, wherein the KLHDC7B missense variant nucleic acid molecule is a genomic nucleic acid molecule having a nucleotide sequence lacking a guanine at a position corresponding to position 2,807 or corresponding to position 3,170 according to SEQ ID NO:1.

**68.** The method of claim 62, wherein the KLHDC7B missense variant nucleic acid molecule is an mRNA molecule having a nucleotide sequence lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:4 or corresponding to position 673 according to SEQ ID NO:6.

**69.** The method of claim 62, wherein the KLHDC7B missense variant nucleic acid molecule is an mRNA molecule having a nucleotide sequence lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:4 or corresponding to position 1,036 according to SEQ ID NO:6.

**70.** The method of claim 62, wherein the KLHDC7B missense variant nucleic acid molecule is a cDNA molecule produced from an mRNA molecule, wherein the cDNA molecule has a nucleotide sequence lacking a guanine at a position corresponding to position 673 according to SEQ ID NO:12 or corresponding to position 673 according to SEQ ID NO:14.

**71.** The method of claim 62, wherein the KLHDC7B missense variant nucleic acid molecule is a cDNA molecule produced from an mRNA molecule, wherein the cDNA molecule has a nucleotide sequence lacking a guanine at a position corresponding to position 1,036 according to SEQ ID NO:12 or corresponding to position 1,036 according to SEQ ID NO:14.

**72.** The method according to claim 62, wherein: the anti-inflammatory drug comprises a steroid, the neurotrophic factor comprises T-817MA, the caspase inhibitor comprises z-DEVD-fmk, and the copper transport inhibitor comprises cimetidine or copper sulphate.

\* \* \* \* \*