

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

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
31 August 2006 (31.08.2006)

PCT

(10) International Publication Number  
**WO 2006/091579 A2**

(51) International Patent Classification:  
*CI2Q I/68* (2006.01)

(74) Agent: PROVENCE, David, L.; MUETING, RAASCH & GEBHARDT, P.A., P.O. Box 581415, Minneapolis, Minnesota 55458-1415 (US).

(21) International Application Number:

PCT/US2006/006078

(22) International Filing Date:

22 February 2006 (22.02.2006)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/655,172 22 February 2005 (22.02.2005) US

(71) Applicant (for all designated States except US): REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; 450 McNamara Alumni Center, Minneapolis, Minnesota 55455 (US).

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

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

Published:

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

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

WO 2006/091579 A2

(54) Title: IDENTIFICATION OF A GENE ASSOCIATED WITH SPINOCEREBELLAR ATAXIA TYPE 5 AND METHODS OF USE

(57) Abstract: The present invention provides methods that include analyzing an SCA5 polynucleotide, and determining whether the SCA5 polynucleotide includes a mutation. The methods may be used to identify a subject that is at risk or not at risk for developing spinocerebellar ataxia type 5. The present invention also provides isolated polynucleotides having a mutation present in an SCA5 polynucleotide.

5 IDENTIFICATION OF A GENE ASSOCIATED WITH SPINOCEREBELLAR  
ATAXIA TYPE 5 AND METHODS OF USE

CONTINUING APPLICATION DATA

This application claims the benefit of U.S. Provisional Application Serial No. 60/655,172 filed February 22, 2005, which is incorporated by reference herein.

10

GOVERNMENT FUNDING

The present invention was made with government support under Grant No. NS33958 and PO1 NS33718, awarded by the National Institutes of Health. The Government may have certain rights in this invention.

15

BACKGROUND

The dominant spinocerebellar ataxias (SCAs) are a heterogeneous group of neurodegenerative disorders characterized by incoordination of gait, limb, and eye movements, slurred speech and swallowing difficulties. Nine of the 11 known SCA mutations are microsatellite repeat expansions (Schols et al., Lancet Neurol 3, 291-304 (2004)). In 1994, SCA5 was mapped to 11q13, a centromeric region with suppressed recombination (Ranum et al., Nature Genetics 8, 280-284 (1994)). MRI and autopsy findings show cerebellar cortical atrophy, Purkinje cell loss and thinning of the molecular layer (Liquori et al., Spinocerebellar ataxia type 5 (SCA5) in Cerebellar Ataxias ed. M. Pandolfo, Cambridge University Press pp 445-450. in The Cerebellum and its Disorders (eds. Manto, M.U. & Pandolfo, M.) 445-450 (Cambridge University Press, Cambridge, 2002)). Additional SCA5 families from France and Germany were reported with similar clinical and neuroradiological findings (Stevanin et al., Neurology 53, 1355-1357 (1999), and Burk et al., Neurology 62, 327-329 (2004)).

The significance of identifying ataxia genes provides an improved method for diagnosis of individuals with the disease and allows the possibility of prenatal/presymptomatic diagnosis for better classification of ataxias.

## SUMMARY OF THE INVENTION

The invention relates to the newly discovered correlation between mutations in the protein β-III spectrin (encoded by the *SPTBN2* gene) and the disease spinocerebellar ataxia type 5 (SCA5). It has been discovered that β-III spectrin mutations cause SCA5 in an 11-generation American kindred descended from President Lincoln's grandparents, and two additional families. β-III spectrin is highly expressed in Purkinje cells and has been shown to stabilize the glutamate transporter EAAT4 at the surface of the plasma membrane. Dramatic differences in EAAT4 and GluRδ2 were found by Western and cell fractionation in SCA5 autopsy tissue. Cell culture studies demonstrated that wildtype but not mutant β-III spectrin stabilizes EAAT4 at the plasma membrane. Spectrin mutations are a novel cause of ataxia and neurodegenerative disease that affect membrane proteins involved in glutamate signaling.

In one aspect, the present invention provides methods that include analyzing an *SCA5* polynucleotide, and determining whether the *SCA5* polynucleotide includes a mutation. The *SCA5* polynucleotide can be obtained from a subject, where a subject at risk of having SCA5 has a mutation in an *SCA5* polynucleotide, or a subject not at risk of having SCA5 does not have a mutation in an *SCA5* polynucleotide. The subject may or may not display at least one symptom of ataxia. The *SCA5* polynucleotide can be a genomic *SCA5* polynucleotide or a processed *SCA5* polynucleotide. The analyzing can include amplification of the *SCA5* polynucleotide, hybridization of the *SCA5* polynucleotide to a second polynucleotide, sequencing a portion of the *SCA5* polynucleotide, or a combination thereof. The *SCA5* polynucleotide may contain a mutation, and the mutation may be present in an exon. A mutation in an exon may result in an *SCA5* polypeptide having an amino acid sequence different than SEQ ID NO:2. The type of mutation may be, for instance, a mutation in a nucleotide corresponding to exon 7 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 12 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 14 of the *SCA5* polynucleotide, or a combination thereof.

The present invention also provides a method for identifying a subject not at risk for developing spinocerebellar ataxia type 5. The method includes analyzing nucleotides of *SCA5* polynucleotide, and determining if the polynucleotide includes a mutation, wherein a subject not at risk of having SCA5 does not have a mutation in an *SCA5* polynucleotide.

The *SCA5* polynucleotide can be a genomic *SCA5* polynucleotide or a processed *SCA5* polynucleotide. The analyzing can include amplification of the *SCA5* polynucleotide, hybridization of the *SCA5* polynucleotide to a second polynucleotide, sequencing a portion of the *SCA5* polynucleotide, or a combination thereof. The *SCA5* polynucleotide 5 may contain a mutation, and the mutation may be present in an exon. A mutation in an exon may result in an *SCA5* polypeptide having an amino acid sequence different than SEQ ID NO:2. The type of mutation may be, for instance, a mutation in a nucleotide corresponding to exon 7 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 12 of the *SCA5* polynucleotide, a mutation in a nucleotide 10 corresponding to exon 14 of the *SCA5* polynucleotide, or a combination thereof.

The present invention further provides a method for identifying a subject at risk for developing spinocerebellar ataxia type 5. The method includes analyzing nucleotides of *SCA5* polynucleotide, and determining if the polynucleotide includes a mutation, wherein a subject at risk of having *SCA5* has a mutation in an *SCA5* polynucleotide. The 15 subject may or may not display at least one symptom of ataxia. The *SCA5* polynucleotide can be a genomic *SCA5* polynucleotide or a processed *SCA5* polynucleotide. The analyzing can include amplification of the *SCA5* polynucleotide, hybridization of the *SCA5* polynucleotide to a second polynucleotide, sequencing a portion of the *SCA5* polynucleotide, or a combination thereof. The *SCA5* polynucleotide may contain a 20 mutation, and the mutation may be present in an exon. A mutation in an exon may result in an *SCA5* polypeptide having an amino acid sequence different than SEQ ID NO:2. The type of mutation may be, for instance, a mutation in a nucleotide corresponding to exon 7 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 12 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 14 of the *SCA5* 25 polynucleotide, or a combination thereof.

The present invention provides a method for determining whether a subject has spinocerebellar ataxia type 5 (*SCA5*). The method includes analyzing an *SCA5* polynucleotide for a mutation, and determining whether the subject displays a symptom of *SCA5*, wherein having a mutation in an *SCA5* polynucleotide and having a symptom of 30 *SCA5* indicates the subject has *SCA5*. The *SCA5* polynucleotide can be a genomic *SCA5* polynucleotide or a processed *SCA5* polynucleotide. The analyzing can include amplification of the *SCA5* polynucleotide, hybridization of the *SCA5* polynucleotide to a second polynucleotide, sequencing a portion of the *SCA5* polynucleotide, or a combination thereof. The *SCA5* polynucleotide may contain a mutation, and the mutation

may be present in an exon. A mutation in an exon may result in an *SCA5* polypeptide having an amino acid sequence different than SEQ ID NO:2. The type of mutation may be, for instance, a mutation in a nucleotide corresponding to exon 7 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 12 of the *SCA5* 5 polynucleotide, a mutation in a nucleotide corresponding to exon 14 of the *SCA5* polynucleotide, or a combination thereof.

Also included in the present invention is a kit for detecting an *SCA5* polynucleotide, including a primer pair that will amplify a portion of an *SCA5* polynucleotide. The present invention also provides an isolated polynucleotide including 10 a mutant of SEQ ID NO:1 or a portion thereof. The mutation present in the polynucleotide may be a mutation in a nucleotide corresponding to exon 7 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 12 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 14 of the *SCA5* polynucleotide, or a combination thereof. The isolated polynucleotide may be 15 to 500 15 nucleotides. Also included is a vector including an isolated polynucleotide of the present invention, and a cell including the vector.

The terms "comprises" and variations thereof do not have a limiting meaning where these terms appear in the description and claims. Unless otherwise specified, "a," "an," "the," and "at least one" are used interchangeably and mean one or more than one.

20

#### BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Pedigree of the Lincoln *SCA5* family. (a) An 11-generation *SCA5* kindred descended from paternal grandparents of President Abraham Lincoln. Squares 25 and circles represent males and females, respectively, shaded symbols represent affected individuals, symbols with a dot indicate obligate mutation carriers, and diagonal lines denote individuals who are deceased. The asterisks beneath the symbols indicate individuals whose blood samples were obtained for analysis. (b) Enlargement of a portion of the pedigree showing the common ancestry of the two branches and their relationship to President Lincoln. The pedigrees of the two branches are abbreviated, and the genders of Josiah and Mary and individuals in generations III, IV, V are masked in (a) 30 to preserve confidentiality.

Figure 2. Mapping and cloning of the *SCA5* mutations. (a) Critical regions defined by recombination events in the three *SCA5* families are indicated by black arrows. The

boundaries of the French critical region are not defined because no recombination events were found among affected family members. Markers defining recombination events, along with other published markers are shown. (b) BAC map spanning the SCA5 region. A panel of 445 novel di-, tri-, tetra-, and penta-nucleotide repeat markers were used to refine

5 the SCA5 region and search for haplotype conservation between the families.

Chromosome-separated cell lines haploid for the affected or the normal chromosome 11 were generated from an affected American family member and used in this screen to directly and unambiguously define the affected haplotype. The enlarged BACs, highlighted in gray, span a 255 kb region of haplotype conservation between the American and French

10 families, containing 11 novel polymorphic STR markers and 8 SNPs (size and NCBI

accession number noted). The three BACs generated from the affected SCA5 haploid cell line are depicted in black along with their relative position and size. The approximate sizes and locations of genes present on the SCA5 specific BAC clones are illustrated by black blocks. The block shaded in gray represents the gene *SPTBN2*. (c) Illustration of *SPTBN2*

15 gene (top) and protein structure (bottom). The relative size and location of the 3' / 5'-UTR and exons are represented by clear and solid squares, respectively. Locations of the three mutations are indicated by arrows on the gene and protein diagrams.  $\beta$ -III spectrin is a 2,390 amino acid protein that is highly homologous to the four other human  $\beta$ -spectrin proteins. Known domains in the protein are specified along with the seventeen spectrin

20 repeats. The calponin-homology (CH)/actin binding domain (ABD), ankyrin binding domain (ANK), and pleckstrin-homology domain (PH) are shaded in gray. The functional unit of spectrin is typically a non-covalently-joined tetrameric complex consisting of two alpha and two beta spectrin subunits. An asterisk (\*) indicates that the direction of *SPTBN2* transcription relative to chromosome 11q is reversed.

25 Figure 3. The three SCA5 mutations and  $\beta$ -III spectrin expression. PCR analysis and the corresponding genotype for the three SCA5 families are illustrated for each mutation. Sequence electropherograms and the corresponding amino acid sequence are also shown. (a) American SCA5 mutation. The PCR analysis generated a 222 bp normal allele and a 183 bp deleted allele. The sequence of SCA5 BAC DNA is shown with the

30 deletion mutation relative to control. The two arrows indicate the two possible deletion sites, and the corresponding 39-base deletions including one of the two flanking TGGA tetranucleotides is underlined. The two TGGA tetranucleotides flanking the American deletion are reminiscent of the deletions caused by slipped-mispairing (Krawczak et al., Hum Genet 86, 425-441 (1991). (b) French SCA5 mutation. The  $[\gamma\text{-}^{33}\text{P}]$  ATP-labeled

PCR generated a 105 bp normal allele and a 90 bp deleted allele. Sequence of the heterozygous and deletion specific PCR product is shown. Arrows indicate the site of the mutation and the 15-base deletion is underlined. (c) German SCA5 mutation. The T to C base change, which converts a leucine to a proline, is depicted. The allele-specific PCR 5 produced a 177 bp normal allele and a 158 bp mutation allele. Amino acid sequence comparisons, of a region containing the German SCA5 mutation (L253P), of five human beta spectrins and beta spectrins from other species are shown. The leucine residue (marked with arrow) which is mutated in the German family is conserved in all five of human beta spectrin proteins and evolutionarily conserved in multiple species. Amino 10 acid alignments were performed with Clustal W (available online through the World Wide Web at, for instance, the Kyoto University Bioinformatics Center). While previously reported polymorphisms were also found in each family, these mutations were the only unreported differences, and the only changes that would alter the corresponding protein. (d) RT-PCR analysis of American SCA5 and control cerebellar tissues. The 15 normal *SPTBN2* amplified product is 227 bp and the product containing the deletion is 188 bp. There was no amplification in the RT- or no RNA control lanes. SCA5-cbl RT+, cerebellum from SCA5 autopsy with reverse transcriptase; SCA5-cbl RT-, cerebellum from SCA5 autopsy without reverse transcriptase (control and should not see product); Cont-cbl RT+, cerebellum from Normal autopsy with reverse transcriptase; CONT-cbl 20 RT-, Cerebellum from normal autopsy without reverse transcriptase (control and should not see product). (e) Immunohistochemistry of control and American SCA5 cerebellar tissues. Sections were stained with an antibody raised against the N-terminal portion of the  $\beta$ -III spectrin (Santa Cruz Biotechnology, Santa Cruz, CA), and visualized at 200X magnification. Enlarged images of the Purkinje cells are also shown (630X). Purkinje 25 cell loss, dendritic atrophy and significant thinning of the molecular layer are seen in SCA5 compared to control.

Figure 4. Western, immunohistochemistry and TIRF microscopy: effects of mutant  $\beta$ -III spectrin on EAAT4. EAAT4 immunoblot comparisons of lysates extracted with RIPA buffer (a, c) or 8M urea and 4% SDS (b, d). EAAT4 and calbindin are both 30 highly expressed in Purkinje cells, with little or no expression in other cells within the cerebellar cortex. When possible, samples were normalized for Purkinje cell loss with calbindin. Markedly less EAAT4 relative to the calbindin control was extracted from human SCA5 cerebella compared to control tissue in the RIPA extracts (a) but similar levels of EAAT4 were found in the harsher 8M urea, 4% SDS buffer (b). As a control,

we examined murine extracts from homozygous 12 week old SCA1 B05 mice but did not observe similar increases in EAAT4 in the urea vs. the RIPA extracts (c, d). EAAT4 immunohistochemistry of American SCA5 (e), murine SCA1 (f) and corresponding human and murine controls. Sections were stained with EAAT4 antibody, and visualized at 200X magnification. Enlarged images of the Purkinje cells are also depicted (630X). Darker EAAT4 staining was observed in the SCA5 Purkinje cell bodies (representative sample) but not in Purkinje cells from SCA1 transgenic mice or controls. (g-i) EAAT4 fast lateral trafficking is modulated by  $\beta$ -III spectrin interaction. (g) A superimposed image shows the total lateral movement of EAAT4 when expressed with an empty vector in HEK293 cells (arrows). (h) EAAT4 was co-transfected with wildtype  $\beta$ -III spectrin and no lateral fast movement was seen. (i) EAAT4 was co-transfected with mutant  $\beta$ -III spectrin containing the 39bp SCA5 deletion and fast movement was seen again (arrows).

Figure 5. Subcellular distribution of EAAT4 and GluR $\delta$ 2. Subcellular fractionation of cerebellar homogenates from human SCA5 and control autopsy tissue was analyzed by Western blots with EAAT4 and GluR $\delta$ 2, and as a control, clathrin light chain antibodies. P1 nuclear pellet; S1 postnuclear supernatant; P2 crude synaptosomal fraction; S2 supernatant of the crude synaptosomal fraction; LP1 pellet obtained after lysis of synaptosomes.

Figure 6. Nucleotide sequence of a genomic *SPTBN2* gene and amino acid sequence of SPTBN2 polypeptide. Exons are shown in capital letters, and introns are shown in small letters. The locations of single nucleotide polymorphisms (SNP) are underlined and the dbSNP rs# cluster id is shown above each SNP. rs5792396, presence or absence of a C; rs10702473, presence or absence of AAA; rs5792395, presence or absence of a G immediately before the underlined C; rs11286358, presence or absence an A. The sequence listing reflects the different nucleotides that can be present at each of the remaining SNPs.

## DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS OF THE INVENTION

30

### *Compositions*

The present invention includes polynucleotides associated with SCA5, polypeptides encoded by the polynucleotides, and methods for identifying such

polynucleotides and polypeptides. As used herein, the term "polynucleotide" refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxynucleotides.

A polynucleotide may include nucleotide sequences having different functions, including for instance coding sequences such as exons, and non-coding sequences such as introns,

5 regulatory sequences, and the like. A polynucleotide can be obtained directly from a natural source, or can be prepared with the aid of recombinant, enzymatic, or chemical techniques. A polynucleotide can be linear or circular in topology, and can be, for example, a portion of a vector, such as an expression or cloning vector, or a fragment.

Polynucleotides can be single-stranded or double-stranded, and the sequence of the

10 second, complementary strand is dictated by the sequence of the first strand. The term "polynucleotide" is therefore to be broadly interpreted as encompassing a single stranded nucleic acid polymer, its complement, and the duplex formed thereby.

"Complementarity" of polynucleotides refers to the ability of two single-stranded polynucleotides to base pair with each other, in which an adenine on one polynucleotide

15 will base pair with a thymidine (or uracil, in the case of RNA) on the other, and a cytidine on one polynucleotide will base pair with a guanine on the other. Two polynucleotides are complementary to each other when a nucleotide sequence in one polynucleotide can base pair with a nucleotide sequence in a second polynucleotide. For instance, 5'-ATGC and 5'-GCAT are fully complementary, as are 5'-GCTA and 5'-TAGC.

20 As used herein, the term "polypeptide" refers broadly to a polymer of two or more amino acids joined together by peptide bonds. The term "polypeptide" also includes molecules which contain more than one polypeptide joined by a disulfide bond, or complexes of polypeptides that are joined together, covalently or noncovalently, as multimers (e.g., dimers, tetramers). Thus, the terms peptide, oligopeptide, and protein are 25 all included within the definition of polypeptide and these terms are used interchangeably. It should be understood that these terms do not connote a specific length of a polymer of amino acids, nor are they intended to imply or distinguish whether the polypeptide is produced using recombinant techniques, chemical or enzymatic synthesis, or is naturally occurring.

30 A polynucleotide or polypeptide may be isolated. An "isolated" polypeptide or polynucleotide means a polypeptide or polynucleotide that has been removed from its natural environment. A polypeptide or polynucleotide may be purified, i.e., essentially free from any other polypeptide or polynucleotide and associated cellular products or other impurities. A "purified" polypeptide or polynucleotide is one that is at least 60%

free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated. Polypeptides and nucleotides that are produced outside the organism in which they naturally occur, e.g., through chemical or recombinant means, are considered to be isolated and purified by definition, since they were never present in a natural environment.

A polynucleotide of the present invention, referred to herein interchangeably as an *SCA5* polynucleotide and an *SPTBN2* polynucleotide, is a polynucleotide originating from the long arm of human chromosome 11 (11q13), between the microsatellite markers KADSCA5-184 and D11S970. An *SCA5* polynucleotide may be genomic or processed.

5 A genomic *SCA5* polynucleotide includes a polynucleotide that encodes an unprocessed preRNA (i.e., an RNA molecule that includes both exons and introns), and the preRNA. When placed under the control of appropriate regulatory sequences, a genomic *SCA5* polynucleotide produces an mRNA. The boundaries of a genomic *SCA5* polynucleotide are generally determined by a transcription initiation site at its 5' end and a transcription 10 terminator at its 3' end. A genomic *SCA5* polynucleotide typically includes introns and exons. A regulatory sequence is a polynucleotide that regulates expression of a genomic sequence to which it is operably linked. A non-limiting example of a regulatory sequence includes promoters. "Operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A 15 regulatory sequence is "operably linked" to a genomic sequence when it is joined in such a way that expression of the genomic sequence is achieved under conditions compatible with the regulatory sequence.

20

An example of a genomic *SCA5* polynucleotide is shown in Figure 6 (SEQ ID NO:1). Other examples are disclosed at Genbank accession number NM\_006946 and AB008567. A genomic *SCA5* polynucleotide typically includes 37 exons, and encodes a 25 polypeptide of 2,390 amino acids (Stankewich et al., Proc. Nat. Acad. Sci. USA, 95:14158-14163 (1998)), an example being SEQ ID NO:2. This polypeptide is often referred to in the art as  $\beta$ III spectrin. A processed *SCA5* polynucleotide is the mRNA originating from transcription of the genomic *SCA5* polynucleotide followed by removal 30 of the nucleotides corresponding to the introns. An example of a processed *SCA5* polynucleotide is SEQ ID NO:1 without the nucleotides corresponding to the introns. A processed *SCA5* polynucleotide also includes a DNA polynucleotide derived from the

mRNA, for instance a cDNA. Furthermore, an *SCA5* polynucleotide of the present invention also includes the complements of a genomic or processed *SCA5* polynucleotide.

An *SCA5* polynucleotide of the present invention includes one or more mutations. The *SCA5* polynucleotide depicted at SEQ ID NO:1 is an example of a normal non-mutated genomic *SCA5* polynucleotide, also referred to herein as a wildtype genomic *SCA5* polynucleotide. Likewise, a wildtype processed *SCA5* polynucleotide has a nucleotide sequence as depicted in SEQ ID NO:1 without the sequences corresponding to the introns. Several single nucleotide polymorphisms (SNPs) have been identified in normal *SCA5* polynucleotides. The locations of these SNPs are shown in Figure 6. The presence of a SNP in an *SCA5* polynucleotide is not considered to be a mutation. One skilled in the art will understand that additional SNPs are likely to be discovered on an ongoing basis, and at an increasing rate, especially in view of the recent sequencing of the human genome. Any change in nucleotide sequence of an *SCA5* polynucleotide when compared to SEQ ID NO:1 is considered to be a mutation.

A mutation may exist in nucleotides corresponding to a 5' upstream region, a 5' untranslated region (UTR), an exon, an intron, a 3' UTR, or a 3' downstream region. A mutation may influence the amount of polypeptide produced by an *SCA5* polynucleotide, for instance by altering transcription of the genomic *SCA5* polynucleotide, or altering (such as destabilizing) translation of the mRNA. A mutation may alter the amino acid sequence of an *SCA5* polypeptide. Examples of mutations include, for instance, deletions, insertions, duplications, and point mutations. A mutation may result in, for instance, a frameshift, an amino acid substitution, insertion or deletion of amino acids, and/or premature termination of translation through the presence of a stop codon.

In some aspects, a mutation in an *SCA5* polynucleotide is in an exon, and results in an amino acid sequence that is altered when compared to the amino acid sequence depicted at SEQ ID NO:2. As is well known, many mutations to a nucleotide sequence encoding a polypeptide can result in a silent mutation, i.e., the nucleotide mutation has no effect on the amino acid sequence. Due to the third-base degeneracy present in a codon, the base in the third position of a codon is often not significant, and a change in the third nucleotide of a codon often does not result in a different amino acid. For instance, the codon AAA and AAG both encode the amino acid lysine. A mutation in an *SCA5* polynucleotide can result in an *SCA5* polypeptide having altered activity. Altered activities include decreased stabilization of EAAT4 at the plasma membrane of Purkinje cells.

It is expected that mutations can be present in essentially any location of an *SCA5* polynucleotide, preferably a location that corresponds to an exon. For instance, a mutation can be in a region encoding the amino terminal domain of an *SCA5* polypeptide, such as the actin-binding domain, or in one of the spectrin repeat domains including, but 5 not limited to, the third spectrin repeat domain. Without intending to be limiting, several mutations have been detected in *SCA5* polynucleotides. For instance, mutations have been detected in the sequences corresponding to exon 7, including the sequences spanning nucleotides 7654-7769, and specifically a T to C (A to G on the non-coding strand) at nucleotide 7755. Mutations have been detected in the sequences corresponding 10 to exon 12, including the sequences spanning 13582-13884, and specifically a deletion of nucleotides 13,823-13,861 or 13,827-13,865. Mutations have also been detected in the sequences corresponding to exon 14, including the sequences spanning nucleotides 15932-16802, and specifically a deletion of nucleotides 16010-16024.

The present invention also includes shorter polynucleotides that correspond to a 15 portion of a genomic or processed *SCA5* polynucleotide. In some aspects the shorter polynucleotides are referred to herein as primers and probes. A polynucleotide of this aspect of the invention has a nucleotide sequence that is complementary to a nucleotide sequence of a genomic *SCA5* polynucleotide, or the complement thereof. In some embodiments, a polynucleotide of this aspect of the invention includes consecutive 20 nucleotides selected from nucleotides 1 - 159, 160 - 316, 317-5418, 5419 - 5570, 5571 - 6004, 6005 - 6178, 6179 - 6992, 6993 - 7084, 7085 - 7228, 7229 - 7309, 7310 - 7653, 7654 - 7769, 7770 - 10370, 10371 - 10483, 10484 - 10630, 10631 - 10818, 10819 - 12380, 12381 - 12498, 12499 - 13100, 13101 - 13259, 13260 - 13581, 13582 - 13884, 13885 - 15562, 15563 - 15716, 15717- 15931, 15932 - 16802, 16803 - 19678, 19679 - 19816, 19817 - 20117, 20118 - 20874, 20875 - 21791, 21792 - 21994, 21995 - 22317, 22318 - 22408, 22409 - 22614, 22615 - 22761, 22762 - 24859, 24860 - 25123, 25124 - 27036, 27037 - 27261, 27262 - 27538, 27539 - 27628, 27629 - 27953, 27954 - 28214, 28215 - 28399, 28300 - 28430, 28431 - 28659, 28660 - 28864, 28865 - 29741, 29742 - 30116, 30117 - 31115, 31116 - 31360, 31361 - 31455, 31456 - 31594, 31595 - 32218, 32219 - 32303, 32304 - 32549, 32550 - 32746, 32747 - 33087, 33088 - 33230, 33231 - 33319, 33320 - 33395, 33396 - 33480, 33481 - 33531, 33532 - 33751, 33752 - 33972, 33973 - 34231, 34232 - 34405, 34406 - 34958, 34959 - 35001, 35002 - 35294, 35295 - 35525, 35526 - 36147 of SEQ ID NO:1, or the complements thereof. Also included are portions of these polynucleotides, wherein the portion is at least 100 consecutive

nucleotides, at least 200 consecutive nucleotides, at least 300 consecutive nucleotides, at least 400 consecutive nucleotides, or at least 500 consecutive nucleotides. Other polynucleotides of this aspect of the invention include the polynucleotides depicted at Table 1 (SEQ ID NOs:3 to 77), or the complements thereof. In some embodiments, a 5 polynucleotide of this aspect of the invention includes a mutation. For instance, a polynucleotide can include nucleotides 13773-13822 and 13867- 13917 (i.e., reflects one of the American mutations described in detail herein), include nucleotides 13777-13826 and 13867- 13920 (i.e., reflects one of the American mutations described in detail herein), include nucleotides 15879-15929 and 16803-16853 (i.e., reflects the French mutation 10 described in detail herein), or the complements thereof. Typically, a polynucleotide of this aspect of the invention has at least about 95% sequence identity, preferably at least about 97% sequence identity, most preferably, about 100% sequence identity with the target sequence to which the primer hybridizes.

Also included in the present invention are primer pairs. As used herein, the term 15 "primer pair" means two oligonucleotides designed to flank a region of a polynucleotide to be amplified. The polynucleotide to be amplified can be referred to as the template polynucleotide. In some aspects, the template polynucleotide is a genomic *SCA5* polynucleotide. Methods for amplifying a polynucleotide are discussed herein. One primer is complementary to nucleotides present on one strand at one end of a template 20 polynucleotide and another primer is complementary to nucleotides present on the other strand at the other end of the template polynucleotide. For example, in some aspects the primers of a primer pair may be used to amplify nucleotides corresponding to one or more exons, or nucleotides corresponding to a portion of an exon. When the template polynucleotide is obtained from genomic DNA, one or both of the primers of the primer 25 pair may be complementary to nucleotides corresponding to an intron. Examples of primer pairs are disclosed at Table 1, and those skilled in the art will recognize that other primer pairs can be easily made using the sequence present at SEQ ID NO:1 and routine methods. A polynucleotide of this aspect of the invention includes, in increasing order of preference, at least 15 consecutive nucleotides, at least 18 consecutive nucleotides, at 30 least 20 consecutive nucleotides, at least 24 consecutive nucleotides, or at least 27 consecutive nucleotides. Typically, a polynucleotide of this aspect of the invention has at least about 95% sequence identity, preferably at least about 97% sequence identity, most preferably, about 100% sequence identity with the target sequence to which the primer hybridizes.

A polynucleotide of the invention can be inserted in a vector. A vector is a replicating polynucleotide, such as a plasmid, phage, or cosmid, to which another polynucleotide may be attached so as to bring about the replication of the attached polynucleotide. Construction of vectors containing a polynucleotide of the invention 5 employs standard ligation techniques known in the art. See, e.g., Sambrook et al, *Molecular Cloning: A Laboratory Manual.*, Cold Spring Harbor Laboratory Press (1989). A vector can provide for further cloning (amplification of the polynucleotide), i.e., a cloning vector, or for expression of the polypeptide encoded by the coding region, i.e., an expression vector. The term vector includes, but is not limited to, plasmid vectors, viral 10 vectors, cosmid vectors, or artificial chromosome vectors. Typically, a vector is capable of replication in a bacterial host, for instance *E. coli*. Preferably the vector is a plasmid. Selection of a vector depends upon a variety of desired characteristics in the resulting construct, such as a selection marker, vector replication rate, and the like. Suitable host 15 cells for cloning or expressing the vectors herein are prokaryote or eukaryotic cells. The vector may contain an entire *SCA5* polynucleotide, or a portion thereof, for instance, a region of nucleotides corresponding to an exon, an intron, or a combination thereof. The present invention also includes cells containing a polynucleotide of the invention inserted in a vector, and cells containing a polypeptide encoded by a polynucleotide of the invention inserted in a vector.

20 The present invention also includes *SCA5* polypeptides containing one or more mutations. The *SCA5* polypeptide depicted at SEQ ID NO:2 is an example of a normal non-mutated genomic *SCA5* polypeptide, also referred to herein as a wildtype *SCA5* polypeptide. Several single nucleotide polymorphisms (SNPs) have been identified in 25 normal *SCA5* polypeptides that result in a different amino acid sequence. The locations of these SNPs are shown in Figure 6. The presence of an altered amino acid sequence due to a SNP in an *SCA5* polypeptide is not considered to be a mutation. Any change in amino acid sequence of an *SCA5* polypeptide when compared to SEQ ID NO:2 is considered to be a mutation, and is a polynucleotide of the present invention. A mutation 30 can result in an *SCA5* polypeptide containing a mutation in the amino terminal domain, such as the actin-binding domain, in one of the spectrin repeat domains including, but not limited to, the third spectrin repeat domain, or in the carboxy terminal domain.

*Methods of Use*

The identification of a genomic sequence that is associated with a disease allows for improved diagnosis of the disease. The present invention discloses that a mutation in an *SCA5* polynucleotide is associated with the disease spinocerebellar ataxia type 5  
5 (SCA5). The present invention includes methods for detecting a polynucleotide of the present invention, such as an *SCA5* polynucleotide including a mutation in an *SCA5* polynucleotide, methods for identifying a subject not at risk for developing SCA5, and methods for identifying a subject that has or is at risk for developing SCA5. The methods of the present invention typically include analyzing an *SCA5* polynucleotide, generally a  
10 portion of an *SCA5* polynucleotide, and determining whether the *SCA5* polynucleotide comprises a mutation.

As used herein, "at risk" describes a subject having an *SCA5* polynucleotide that contains a mutation. Preferably, the mutation is present in a nucleotide corresponding to an exon. Preferably, the mutation results in an *SCA5* polypeptide having an amino acid sequence that is different than the amino acid sequence disclosed at SEQ ID NO:2. More than one mutation may be present. An at risk subject includes an individual who may be manifesting at least one symptom of SCA5, as well as a subject who may develop at least one symptom of SCA5 in the future. Symptoms of SCA5 include incoordination of gait, limb, and eye movements, slurred speech and swallowing difficulties. The evaluation of  
15 such symptoms is routine and easily accomplished by a person of ordinary skill. A subject that does not have an *SCA5* polynucleotide containing a mutation as described herein is expected to not display symptoms of SCA5 during his or her lifetime, and is considered to be "not at risk."

The methods of the present invention include analyzing an *SCA5* polynucleotide,  
25 and determining whether the *SCA5* polynucleotide includes a mutation. The source of polynucleotides is typically a biological sample that includes genomic DNA and/or processed RNA. As used herein, a "biological sample" refers to a sample of material (solid or fluid) obtained from an individual, including but not limited to, for example, blood, plasma, serum, or tissue. A biological sample may be treated to obtain  
30 polynucleotides, for instance, DNA or RNA. A subject can be a rat, mouse, human, chimpanzee, or gorilla, preferably human. The *SCA5* polynucleotide that is analyzed may be an entire *SCA5* polynucleotide, and is typically a portion of an *SCA5* polynucleotide.

The present invention provides methods for analyzing an *SCA5* polynucleotide, including at least a portion of an *SCA5* polynucleotide. In one aspect, the method

includes amplifying nucleotides of an *SCA5* polynucleotide of a subject to form amplified polynucleotides, preferably including amplified nucleotides that correspond to an exon, and detecting the amplified polynucleotides. Preferably, nucleotides are amplified by PCR. In PCR, a molar excess of a primer pair is added to a biological sample that

5 includes polynucleotides, preferably genomic DNA. The primers are extended to form complementary primer extension products which act as template for synthesizing the desired amplified polynucleotides. The conditions for amplifying a polynucleotide by PCR vary depending on the nucleotide sequence of primers used, and methods for determining such conditions are routine in the art.

10 Various types of amplification techniques are known and used routinely, such as allele-specific PCR, cold PCR, hot PCR, reverse-transcriptase PCR, and the like. These and other amplification techniques are known in the art and are used routinely. In view of the disclosure of SEQ ID NO:1, the skilled person can easily adapt an amplification technique to be used in identifying mutations in an *SCA5* polynucleotide. Examples of  
15 primers that can be used in the methods of the present invention include those depicted in Table 1 (SEQ ID NOs:3-77).

After amplification, the sizes of the amplified polynucleotides may be determined, for instance by gel electrophoresis, and compared. The amplified polynucleotides can be visualized by staining (e.g., with ethidium bromide) or labeling with a suitable label  
20 known to those skilled in the art, including radioactive and nonradioactive labels. Typical radioactive labels include  $^{33}\text{P}$ . Nonradioactive labels include, for example, ligands such as biotin or digoxigenin as well as enzymes such as phosphatase or peroxidases, or the various chemiluminescers such as luciferin, or fluorescent compounds like fluorescein and its derivatives. Optionally, the nucleotide sequence of an amplified polynucleotide  
25 can be determined.

In another aspect of the methods for analyzing an *SCA5* polynucleotide containing a mutation, polynucleotide probes are used that hybridize to a polynucleotide. As used herein, "hybridizes," "hybridizing," and "hybridization" means that a probe forms a noncovalent interaction with a target polynucleotide under standard conditions. Standard  
30 hybridizing conditions are those conditions that allow a probe to hybridize to a target polynucleotide. Such conditions are readily determined for a probe and the target polynucleotide using techniques well known to the art, for example see Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: New York (1989). Preferred probes useful in the present invention hybridize to a target

polynucleotide by using prehybridization in a hybridization buffer, preferably RAPID-HYB buffer (Amersham, Piscataway, NJ), at 60° for 1 hour, and hybridization overnight at 60°C. Preferably, at least 4 x 10<sup>7</sup> counts per minute (cpm) total of the labeled probe is used in the hybridization. When the probe used is at least about 200 nucleotides, the  
5 wash conditions used are: 2 washes for 5 minutes each at room temperature in a solution containing 2X SSC (one liter of 20X SSC contains 175.3 grams NaCl and 88.2 grams sodium citrate, pH 7.0) and 0.05% sodium dodecyl sulfate (SDS), followed by 2 to 3 washes for 30 minutes each at 52° in a solution containing 0.15X SSC and 0.1% SDS. Other hybridization conditions for use when the probe is at least about 200 nucleotides  
10 use the same prehybridization and hybridization conditions as described above, but the wash conditions used are: 2 washes for 5 minutes each at room temperature in a solution containing 2X SSC and 0.05% SDS, followed by 1 wash for 15 minutes at 50°C in a solution containing 0.15X SSC and 0.1% SDS, followed by 1 wash for 10 minutes at 50°C in a solution containing 0.15X SSC and 0.1% SDS. When the probe used is about  
15 20 to about 22 nucleotides, the same prehybridization and hybridization conditions described above are used, but the wash conditions used are: two 15 minute washes at 45°C in 2xSSC and 0.1% SDS. The nucleotide sequence of a target DNA molecule is generally a sequence complementary to the probe. The hybridizing probe may contain 1 to 10 nonhybridizing nucleotides, preferably no greater than 5, more preferably no greater  
20 than 2 nonhybridizing nucleotides, that do not interfere with forming the noncovalent interaction. The nonhybridizing nucleotides of a probe may be located at an end or within the hybridizing probe. Thus, a probe does not have to be complementary to all the nucleotides of the target DNA sequence as long as there is hybridization under standard hybridization conditions.

25 In one embodiment of this aspect of the invention, the methods include digesting genomic DNA of a subject with a restriction endonuclease to obtain polynucleotides, and probing the polynucleotides under hybridizing conditions with a detectably labeled probe. The digestion of genomic DNA with endonucleases is routine in the art, and numerous endonucleases are known. Typically, the polynucleotides resulting from digestion are  
30 fractionated, for instance by gel electrophoresis, denatured to yield single stranded polynucleotides, and then exposed to the probe under hybridizing conditions. The probe that has hybridized to the polynucleotide is then detected, and the size of the hybridized polynucleotide may then be determined. The presence or absence of the mutation can be inferred by the approximate molecular weight of the detected polynucleotide. The

presence of a mutation indicates the person has or is at risk, and the absence of a mutation indicates the person is not at risk.

Other methods can be used to analyze an *SCA5* polynucleotide. Examples include, but are not limited to, ligase-mediated detection techniques (Landegren, U.S. Pat. No. 4,988,617), fluorescent in situ hybridization (Stokke, U.S. Pat. No. 5,633,365 and Pinkel, U.S. Pat. No. 5,665,549), direct DNA sequencing, PFGE analysis, Southern or Northern blotting, single-stranded conformation analysis (SSCA), RNase protection assay, allele-specific oligonucleotide (Wallace, U.S. Pat. No. 5,639,611), dot blot analysis, denaturing gradient gel electrophoresis (Borresen, U.S. Pat. No. 5,190,856), RFLP (Helentjaris, U.S. Pat. No. 5,324,631) and PCR-SSCP. Methods for detecting and quantifying gene sequences, such as mutated genes and oncogenes, in for example biological fluids are described in Sorenson (U.S. Pat. No. 5,496,699).

The present invention also provides a kit for identifying whether a subject is at risk or not at risk for developing *SCA5*. The kit includes the primers and/or probes discussed above in a suitable packaging material in an amount sufficient for at least one assay. Optionally, other reagents such as buffers and solutions needed to practice the invention are also included. Optionally, other reagents such as buffers and solutions needed to practice the invention are also included. As used herein, the phrase "packaging material" refers to one or more physical structures used to house the contents of the kit. The packaging material is constructed by well known methods, generally to provide a sterile, contaminant-free environment. The packaging material may have a label which indicates that the polynucleotides can be used for identifying whether a subject is at risk or not at risk for developing *SCA5*. In addition, the packaging material contains instructions indicating how the materials within the kit are employed. As used herein, the term package or container refers to a receptacle such as glass, plastic, paper, foil, and the like, capable of holding within fixed limits the primers and/or probes. Thus, for example, a package can be a plastic vial used to contain milligram quantities of a primer pair. "Instructions for use" typically include a tangible expression describing the reagent concentration or at least one assay method parameter, such as the relative amounts of reagent and sample to be admixed, maintenance time periods for reagent/sample admixtures, temperature, buffer conditions, and the like.

The present invention is illustrated by the following examples. It is to be understood that the particular examples, materials, amounts, and procedures are to be

interpreted broadly in accordance with the scope and spirit of the invention as set forth herein.

## EXAMPLES

5

### Materials and Methods

*Human subjects.* All participating subjects and control individuals referred to this study signed an informed consent form as approved by the Human Subjects Committee at the University of Minnesota or by the participating institutions. Unrelated control DNA samples were obtained from the CEPH panel and from healthy North Americans (n=500).  
10 DNA was extracted from peripheral venous blood using the Puregene kit (Gentra Systems, Plymouth, MN).

*Generation of chromosome-separated cell lines.* Mouse/human hybrid cell lines haploid for the affected or normal copy of chromosome 11 were generated at GMP Genetics (Waltham, MA) by fusing mouse E2 cells with human lymphoblastoid cells from an affected American family member, as previously described (Papadopoulos et al., Nat Genet 11, 99-102 (1995)). In brief, lymphoblast cells from an affected individual were electrofused to mouse E2 cells and HAT plus geneticin was used to select against unfused E2 and lymphoblast cells, respectively. The surviving colonies were expanded  
15 and clones containing only a single copy of the affected or normal chromosome 11 were selected by typing microsatellite markers that spanned the SCA5 region.  
20

*Screening of microsatellite repeat markers in the SCA5 region.* Microsatellite repeat markers were amplified by PCR using a [ $\gamma$ -<sup>33</sup>P] ATP tagged primer. Products were separated on 4% denaturing polyacrylamide gels and visualized by autoradiography.  
25 Genotyping of the single affected chromosome allowed for the exclusion of repeat-expansion mutations in non-polymorphic markers. The E2 mouse DNA was used as a negative control to confirm the amplified product was specific to human but not mouse DNA. All polymorphic markers were subsequently used to determine the affected haplotypes for each of the SCA5 families.

*Construction of BAC libraries from an affected SCA5 haploid cell line and shotgun DNA sequencing.* An incomplete *Hind III* digestion was performed on DNA from the haploid cell line containing the affected chromosome 11 and introduced into the pIndigoBAC-5 vector (Epicentre, Madison, WI), which was then used to prepare a BAC  
30

library of approximately 352,000 recombinant clones. The BAC libraries were screened by PCR using microsatellite markers and positive BAC clones were subsequently isolated by hybridization. Lark Technologies Inc. (Houston, TX) performed the shotgun sequencing and assembly. In brief, shotgun libraries were constructed for three BACs 5 (VI-C2, VI-C11, IV-H4), which spanned the region of haplotype conservation between the American and French SCA5 families, by subcloning the fragmented DNA into the pUC57 vector. Sequencing reactions of the three shotgun libraries were performed and subsequently analyzed on ABI3730xl DNA sequencers. The sequence data was assembled using the Phred-Phrap-Conse software (Gordon et al., Genome Res 8, 195- 10 202 (1998) and was subsequently BLASTED against specific genes using data available online through the UCSC Genome Bioinformatics and National Center for Biotechnology Information internet sites.

*SPTBN2 gene sequencing in SCA5 families and mutation screening in controls.*

Genomic DNA of affected French and German SCA5 patients was used to amplify 15 *SPTBN2* exons by PCR and the resulting products were sequenced. After the American and French mutations were identified, family members and 1,000 control chromosomes were screened for these deletion mutations by PCR. PCR was performed by labeling the 5' end of each forward primer with [ $\gamma$ -<sup>33</sup>P] ATP. The resulting products were separated on 4% denaturing polyacrylamide gels and visualized by autoradiography. Allele- 20 specific PCR analysis was used to screen for the German missense mutation. Two forward primers, one containing an altered nucleotide (C) at its 3'-end and the other containing a 19bp-tail at its 5'-end, were used in a single reaction to amplify both the mutant (shorter product) and normal (longer product) alleles, respectively. The resulting products were separated on 4% agarose gels and visualized by ethidium bromide. PCR 25 was subsequently performed on unrelated 1,000 control chromosomes to screen for the German mutation. The PCR primer sequences and conditions used for *SPTBN2* sequencing and mutation screening are shown in Table 1.

**Table 1. Primer sequences and PCR conditions for *SPTBN2* sequencing**

Exon(s)	Sequence (5'-3')	T <sub>a</sub>	M <sub>g</sub>	size (bp)
exon1-2	Forward: CTGCCTTCCGTGCTTCACTT	54	1	468
	Reverse: TCATGACGGAGCTGACAAAGC			
exon3	Forward: CCCTGCCAACACTGGGTITAG	54	1	282
	Reverse: GGTCCCCCTGGACACTTTC			
exon4	Forward: TGCCTGTCGTGTTCCCTGAG	54	1	395
	Reverse: TCCTCCACATCTTGTTGTTG			
exon5-6	Forward: ACACCAGGAGTTCCCTGTCCA	54	1	495
	Reverse: TGCTCCGAGTGGCTATTCCCT			
exon7	Forward: TTGGTGTGGGTTCTCTCTC	54	1	248
	Reverse: CACTGGTCCACCTCCCTGCT			
exon8-9	Forward: GAACTCTGGGAAGGCCCTGA	54	1	568
	Reverse: TCCCTGAAGGCTGTGCTAAT			
exon10	Forward: CCTCGTGGGGCTTTAATTCTG	54	1	228
	Reverse: ATGTGTGCAAGGCACTCTGG			
exon11	Forward: CCACCCCTGTCCCCTCCACTA	54	1	244
	Reverse: CCCAGTCTGACAGCCTAA			
exon12	Forward: AGAGGCCACTGTCCCCCTGGT	54	1	464
	Reverse: GCTGGTTCACACTCCACAGA			
exon13	Forward: GAAAAACGGCAGCCAGGTAG	54	1	279
	Reverse: GCTCTTGATGTGCTCCCTCC			
exon14	Forward: GGCTGGGTTAAGGCTCTGAC	57	1	990
	Reverse: AGGGGACTCACCCACAT			
exon15	Forward: GCTGCCCTCCCCAACATTAC	54	1	234
	Reverse: TCCCCATTGCTTCAATTTC			
exon16	Forward: GGAAAAGGCTTCAAACAGG	54	1	895
	Reverse: CCATCCGTCCCTCACATI			

		Forward: TGCTTGTGGTCACCTC Reverse: GGTTCCCTGTCACGTTA	54	1	395
exon18-19		Forward: GGTTAGCCAAGGGTCACAA Reverse: ACAAAAACCACGGTCCGTGAG	54	1	593
exon20		Forward: GGCTAATTGGGCACTTGA Reverse: CCCCTTCTCTGCTGTTCA	54	1	354
exon21		Forward: GCGGAAATGCCAGAGCTAACAA Reverse: GGAGATGGTCAATGCCAAAG	54	1	395
exon22		Forward: TGTCCCCACTCCCACTTAATC Reverse: AAAAACACGTCCAAGTCTGG	54	1	233
exon23-24		Forward: CTGACGGGTGTTACCATCG Reverse: AGCACITGAAGGCTCCACAIT	54	1	712
exon25		Forward: GAACAGACCGGAGGTCAAGAG Reverse: CTGTGGGTCCCTCCACTCTTC	61	2	328
exon26		Forward: TAACATCACGGCATGGTCTG Reverse: CCCTAGCTCCTGGGAACCT	54	1	498
exon27-28		Forward: CTGGAGTCCCCGGCT Reverse: AAGCAGAAAGCACCAGAA	54	1	599
exon29		Forward: TCACATCCTGGTGCTAACCTCA Reverse: CCTACTCTGGAACCCACAGG	61	3	201
exon30		Forward: CCACTCTGACCCACCATCTT Reverse: AAGCCAGCACGGTCAGG	54	1	300
exon31-33		Forward: CCCTCTTACACGCCAACCTTC Reverse: GACCCCTCGCCCTCACAGTA	54	1	541
exon34		Forward: GGTAGGGATCTCCCGTCTC Reverse: CCCTTGGCCAGAAGATGTA	54	1	374
exon35		Forward: AGATGGGAGCAGAACCTGAA Reverse: CTGGCCCTGGTTACTCCACCTC	54	1	392
exon36		Forward: TACGCTCTCACCCAGCAGCTA Reverse: CGCACACATCCAGTCITACC	59	2	243
exon37		Forward: CAGCTCACTTCTGCCTCTC Reverse: AGAGAGGGTGTGGTCAGGAA	57	1	998

Primer sequences and PCR conditions for SPTBN2 mutation screening

Mutation	Sequence (5'-3')	T <sub>a</sub>	Mg	size (bp)
For E532_M544del (American mutation)	Forward: AGCGCTTACCAAGACATCAAAG Reverse: CCCTTCGACTCTTGATCACCTCTT	54	1	222 (normal) 183 (mutant)
For L629_R634delinsW (French mutation)	Forward: GTGGCCAAAGCTAGAGCAGAG Reverse: CACCTCCCCAGAGGAAACCG	61	2	105 (normal) 90 (mutant)
For L253P (German mutation)	Forward: CACGACGTTGTAAACGACGAACCTGGGACTTACCAAAGCT (for normal) Forward: GAACTGGGACTTACCAAGCC (for mutant) Reverse: CCAAAGAAGCCCCCTGTATCA	55	1.5	177 (normal) 158 (mutant)

Primer sequences and PCR condition for RT-PCR analysis of the American SCA5 deletion

Purpose	Sequence (5'-3')	T <sub>a</sub>	Mg	size (bp)
For first-strand synthesis	CCTCAGGCTTCACCCACCTTC			
PCR-primer in exon 12	Forward: AGCGCTTACCAAGACATCAAAG	54	1	227 (normal)
PCR-primer in exon 13	Reverse: CAGGTCTCCACTCTGTCTTA			188 (mutant)

Primer sequences and PCR conditions to generate a β-III spectrin construct with the American SCA5 deletion

Purpose	Sequence (5'-3')	T <sub>a</sub>	Mg	size (bp)
To generate a PCR fragment including the American SCA5 deletion*	SPΔ39-1f: GTGTCCCCAGGACAACTTTGG SPΔ39-1r: ATCCAGTCCACCTGGACTGGATGGAAAGAGATG	54	1	260
SPΔ39-2f: CTCCCTCAAACCTGGACTGGATGGAAAGAGATG SPΔ39-2r: CTCCAGGGTGAGCTTCAGG		54	1.5	486
To introduce a myc-tag to the N-terminal coding region	myc-f1: CTCATCTCAGAAAGAGGATCTGAGCAGGCCGTCAACCC myc-f2: CGCGGGTACCAATGGAACAAAAACTCATCTCAGAAAGAGGATC myc-r: GAGGAGCCTCAGCAGGTTG	55	1	303 (f1-r) 328 (f2-r)

T<sub>a</sub>: annealing temperature (°C)Mg: MgCl<sub>2</sub> (mM)\* After PCR with SPΔ39 primer sets 1 and 2, a third PCR was performed at 54 °C with 1.5 mM MgCl<sub>2</sub> by SPΔ39-1f and SPΔ39-2r primers.

*RT-PCR analysis.* RNA was harvested from ~100mg of cerebellar autopsy tissue from an American SCA5 patient and a control individual using TRIzol (Invitrogen, Carlsbad, CA). First-strand synthesis was performed using the Invitrogen SuperScript™ 5 First-Strand Synthesis System for RT-PCR kit (Invitrogen, Carlsbad, CA) and a *SPTBN2* gene specific primer from exon 14. PCR primers flanking the American SCA5 deletion region were located in exons 12 and 13, respectively. The products were separated on a 2% agarose gel and visualized with ethidium bromide. The primers and conditions for RT-PCR analysis of the American SCA5 deletion are shown in Table1.

10            *Immunohistochemistry.* The autopsy tissue from an American SCA5 family member and a control individual, without neurological disease, and brains from control and SCA1 B05 transgenic mice were embedded in paraffin, and 5µm sections were prepared. These sections were incubated in 0.3% H<sub>2</sub>O<sub>2</sub> for 30 min to bleach endogenous peroxidase activity, then heated by a steamer in 10 mM citrate buffer at pH 6.0. Sections 15 were blocked in 5 % normal serum, derived from animals in which the secondary antibodies had been made. Slides were incubated at 4°C overnight with β-III spectrin or EAAT4 antibodies (Santa Cruz Biotechnology, Santa Cruz, CA) diluted at 1:500 or 1:100, respectively. Positive staining was visualized by the avidin-biotin-peroxidase complex method (Vector, Burlingame, CA) with diaminobenzidine as the chromogen and 20 counterstained with hemotoxylin.

20            *Immunoblot analysis.* Cerebellar tissue from an SCA5 American family member, human control, murine control and SCA1 B05 transgenic mice were used for Western analysis. Tissue was extracted with a Polytron homogenizer in RIPA lysis buffer (1X PBS, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS, 100µg/ml PMSF, 25 50KIU/ml aprotinin, 1mM sodium orthovanadate). To ensure the efficiency of protein extraction, the same cerebellar tissues were re-extracted in a stronger lysis buffer containing 8M urea, 4% SDS, 0.125M Tris-HCl (pH6.8), 12mM EDTA, 3% β-mercaptoethanol, and 1X protease inhibitors (Complete, Roche, Indianapolis, IN). To determine if EAAT4 was decreased in amount beyond that expected due to Purkinje cell 30 loss, the amount of protein loaded was normalized relative to the Purkinje cell specific protein calbindin. After solubilization, samples were separated by SDS-PAGE and transferred to a nitrocellulose membrane and incubated at 4°C overnight with EAAT4 or calbindin (Sigma-Aldrich, Saint Louis, MO) antibodies diluted at 1:200 or 1:6,000, respectively. The immunoblot was visualized with horseradish peroxidase-conjugated

secondary antibody and enhanced chemiluminescence (Amersham Biosciences, Uppsala, Sweden).

*Subcellular fractionation.* Subcellular fractionation analysis was performed as described elsewhere (Lee et al., Neuropharmacology 41, 680-692 (2001) with slight modifications. Briefly, cerebellar tissues (500mg each) from American SCA5 and control autopsy brains were resuspended by Polytron homogenization in 5ml of buffered sucrose (0.32M sucrose, 5mM Tris (pH7.5), 0.5mM CaCl<sub>2</sub>, 1mM MgCl<sub>2</sub>, and 1X protease inhibitors (Complete, Roche, Indianapolis, IN). Tissue was sheared by passage through an 18-gauge needle repeatedly, and the lysate was pelleted at 500 x g for 10 min (P1 fraction). The supernatant (S1) was separated into two 0.5-ml aliquots and all aliquots were centrifuged at 10,500 x g for 15 min. For one of the aliquots, the supernatant (S2) and pellet (P2) were isolated. For the other aliquot, the pellets from the 10,500 x g spin (P2) were resuspended and hypotonically lysed by the addition of 50µl of ice-cold H<sub>2</sub>O (with 1x protease inhibitors) and passage through an 18-gauge needle 10 times. This mixture was then centrifuged at 25,000 x g for 20 min, generating LS1 (supernatant) and LP1 (pellet) fractions. All pelletable fractions (P1, P2, and LP1) were resuspended in a lysis buffer containing 8M urea, 4% SDS, 0.125M Tris-HCl (pH6.8), 12mM EDTA, 3% β-mercaptoethanol, and 1X protease inhibitors. All resulting fractions were then analyzed by SDS-PAGE and Western blotting. Antibodies against proteins examined in subcellular fractionation analysis were used at the following dilutions: EAAT4 (1:200), GluRδ2 (1:1,000, BD Biosciences, San Jose, CA), and clathrin light chain (1:1,000, Synaptic Systems, Goettingen, Germany).

*Cloning of EAAT4 and β-III spectrin constructs, cell culture and transfection.*

Standard techniques were used in the construction of the β-III spectrin control and deletion constructs and the EAAT4-GFP construct. Briefly, a full-length *SPTBN2* pBluescript cDNA clone (KIAA0302, Kazusa DNA Research Institute) was re-cloned into the mammalian expression vector pcDNA3.1 (Invitrogen, Carlsbad, CA) and modified by PCR using overlapping primer sets (set1: SPΔ39-1f and SPΔ39-1r, and set 2: SPΔ39-2f and SPΔ39-2r). The American family deletion was created by generating separate PCR products (SPΔ39 primer sets 1 and 2) followed by a third PCR reaction (primers SPΔ39-1f and SPΔ39-2r) to generate the 39bp deletion mutation (SP-Δ39) found in the American kindred. These PCR products were then subcloned using BsmB I and Age I digestion. Subsequently, a myc-tag was introduced into both the wildtype (SP-WT) and mutant constructs immediately downstream of the ATG start codon by PCR (myc-f1

and myc-r, followed by myc-f2 and myc-r primers) and then subcloned using Kpn I and Pml I digestion. Sequencing was performed to verify the integrity of the tag and the entire cDNA and coding errors were fixed using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). The primer sequences and PCR conditions to generate the  $\beta$ -III spectrin constructs are shown in Table 1.

- The EAAT4-GFP construct was generated using primers containing the appropriate restriction enzyme recognition sites and an overlap extension PCR-based strategy. Resulting EAAT4 PCR products were cloned into the Eukaryotic expression vector pEGFP-C2 (Clontech), and coding regions were confirmed by sequencing.
- 10 HEK293 cells were transfected (0.5 $\mu$ g/dish) using FuGene 6 (Roche, Indianapolis, IN) following standard protocols. Cells were plated directly on glass bottom culture dishes (MatTek, Ashland, MA) and imaged 24 hours after transfection.

*TIRF microscopy and analysis.* Light from an ion laser was introduced into an inverted epifluorescence microscope (IX81, Olympus) and the light was focused at the 15 back focal plane of a TIRFM objective lens (PlanApo 60 $\times$ /1.45NA, Olympus). The transfected cells on the glass coverslip were maintained at 37°C using a temperature controller (Harvard apparatus) and pH 7.4 by 10 mM Hepes. Images were collected by an EM-charge-coupled-device camera (Olympus) operated with Metamorph 6.3 (Universal Imaging). Time laps images were acquired every 450 msec. Analysis, including tracking 20 (the single projection of different images) and area calculations, were performed using Metamorph. Each diffraction spot was filtered twice (High pass filter >3 pixel) and (Low pass filter <30 pixel). EAAT4 lateral movement images where superimposed to a single image to measure the total area of the transporter movement, while total trafficking distance of diffraction spots was calculated using the Metamorph tracking module.

25

## Results

The American family has two major branches that descend from the paternal grandparents of President Abraham Lincoln (Fig. 1). SCA5, referred to as "Lincoln Disease" by family members, is found among the descendants of President Lincoln's 30 paternal uncle Josiah and aunt Mary, indicating that one of President Lincoln's paternal grandparents carried the SCA5 mutation. These two branches of the family are shown in Figure 1. Clinical evaluations and DNA collection were performed on 299 family members, including 90 affecteds (onset 4-68yrs). Because the disease in some individuals is relatively mild and the clinical status of the President, his father Thomas,

and Thomas's descendants (all deceased since 1960) are unknown, the prior probability that the President inherited the SCA5 mutation is 25%. Recombinations were used to refine the critical region to 2.99 megabases containing ~100 genes (Fig. 2a). Haplotype comparisons between families identified a 255kb region of possible conservation between 5 the American and French families. Although this haplotype was also found in 3/84 (3.5%) control chromosomes, this region was prioritized because of the possibility that this conservation resulted from a common ancestral mutation. DNA from an affected chromosome-separated cell line known to contain the American SCA5 mutation, was used to construct a BAC library and clone contig of the region and shotgun sequencing of 10 patient-derived BAC clones (VI-C2, VI-C11, and IV-H4) spanning the area of haplotype conservation (Fig. 2b) was performed.

A 39-base pair deletion was found in exon 12 of the β-III spectrin gene (*SPTBN2*) which causes an in-frame 13 amino-acid deletion (p.E532\_M544del) within the third of 17 spectrin repeats (Fig. 2c,3a, Table 2). The mutation, which is detectable by PCR (Fig. 15 3a), was found in all 90 affected individuals (age of exam 7-80 yrs, mean 45 yrs) and 35 presymptomatic carriers (age of exam 13-67 yrs, mean 34 yrs).

**Table 2. Summary of DNA sequence variations of exons found in 3 BAC regions.**

Genes	Status	No of exons	No of seq. variations	Exon#	NCBI SNP ID
<i>MRPL11</i>	Reviewed	5	0	-	-
<i>PELI3</i>	Provisional	8	2	exon 6 exon 8	rs2277302 rs3179961
<i>DPP3</i>	Reviewed	18	3	exon 4 exon 17 exon 17	rs11550299 rs1671063 rs2305535
<i>BBS1</i>	Reviewed	17	4	exon 4 exon 14 exon 17 exon 17	rs2298806 rs3816492 rs8432 rs3741360
<i>AK126268</i>	Predicted	1	2	exon 1 exon 1	rs7116921 rs7116940
<i>ZDHC24</i>	Provisional	3	1	exon1	rs2305534
<i>ACTN3</i>	Reviewed	21	7	exon14 exon15 exon15 exon16 exon16 exon18 exon19	rs1671064 rs2305537 rs1815739 rs618838 rs7924602 unregistered <sup>a</sup> rs540874
<i>CTSF</i>	Reviewed	13	4	exon2 exon6 exon13 exon13	rs2075791 rs545009 rs572846 rs4576
<i>FLJ10786</i>	Predicted	1	0	-	-
<i>CCS</i>	Reviewed	8	1	exon8	rs1127145
<i>RBM14<sup>b</sup></i>	Validated	3	0	-	-
<i>MGC15912<sup>b</sup></i>	Predicted	1	0	-	-
<i>LOC440048<sup>b</sup></i>	Model	3	1	exon3	rs670900
<i>RBM4<sup>b</sup></i>	Provisional	5	0	-	-
<i>RBM30</i>	Predicted	4	0	-	-
<i>SPTBN2</i>	Provisional	37	1 <sup>c</sup>	exon14	rs4930388
<i>FLJ22531</i>	Predicted	6 <sup>d</sup>	0	-	-

a: synonymous SNP (AGG→AGA; Arg774)

b: found sequence gap between contigs

c: except for pathogenic SCA5 mutations

d: exons7-17 not included in IV-H4 BAC

Although the American and French families share a common haplotype, the 39-bp American deletion was not found in the French family. Similar to the American family, the French family has a short in-frame deletion in the same spectrin repeat consisting of a 5 15-base pair deletion in exon 14 (c.1886\_1900del; p.L629\_R634delinsW) (Fig. 2c, 3b). With the exception of the insertion of a tryptophan, this deletion does not disrupt the remainder of the open-reading frame (Fig. 3b). The French mutation was found in all six available affected individuals and one apparently presymptomatic carrier (age 24).

In the German family a T to C transition mutation (c.758T>C) in exon 7 that 10 causes a leucine to proline change (p.L253P) (Fig. 2c, 3c) was found in the calponin-homology domain containing the actin/ARP1 binding site. This region is highly conserved with the leucine 253 residue found in all five human  $\beta$ -spectrin proteins as well as chimp, mouse, rat, dog and fly (Fig. 3c). The German mutation co-segregated with the disease in 12 available affected individuals. None of the three SCA5 mutations were 15 found on 1,000 control chromosomes.

$\beta$ -III spectrin, a 2,390 amino-acid protein highly expressed in Purkinje cells (Ohara et al., Brain Res Mol Brain Res 57, 181-192 (1998)), (Stankewich et al., Proc. Natl. Acad. Sci. USA 95, 14158-14163 (1998)), was originally described as a protein associated with Golgi and vesicle membranes (Stankewich et al., Proc. Natl. Acad. Sci. 20 USA 95, 14158-14163 (1998)) and has been reported to bind to the dynactin subunit ARP1, suggesting a possible role in transport (Holleran et al., J. Biol Chem 276, 36598-36605 (2001)). Another function of  $\beta$ -spectrin is the stabilization of membrane proteins (Parkinson et al., Nat Genet 29, 61-65 (2001)); notably  $\beta$ -III spectrin stabilizes the Purkinje cell specific glutamate transporter EAAT4 (Jackson et al., Nature 410, 89-93 25 (2001)). RT-PCR analysis shows both normal and deleted  $\beta$ -III spectrin transcripts are expressed in affected cerebellar autopsy tissue (Fig. 3d) with immunohistochemistry showing staining of Purkinje cell bodies, dendrites and axons in both SCA5 and control cerebella, with marked Purkinje cell loss in SCA5 (Fig. 3e).

Western analysis was performed on cerebellar autopsy tissue to investigate 30 whether the 39-bp spectrin deletion mutation affects EAAT4. Protein levels of EAAT4 in SCA5 cerebellum extracted by Radio-Immunoprecipitation Assay (RIPA) buffer were dramatically reduced relative to calbindin, a Purkinje cell specific control (Fig. 4a). Surprisingly, when using a harsher extraction buffer (8M urea and 4% SDS),

approximately equal ratios of EAAT4/calbindin were seen in SCA5 and control (Fig. 4b) suggesting EAAT4 solubility or distribution is affected by mutant  $\beta$ -III spectrin.

Decreased EAAT4 transcript levels have been previously reported in SCA1 transgenic mice prior to Purkinje cell loss (Lin et al., Nat Neurosci 3, 157-163 (2000)), suggesting that loss or dysfunction of EAAT4 may be a common downstream molecular change. To determine if the extractability differences of EAAT4 in SCA5 is a non-specific change caused by Purkinje cell degeneration, EAAT4 extractability was examined in SCA1 transgenic mice with significant Purkinje cell loss (Fig. 4c,4d).

Consistent with previous reports reduced levels were found of EAAT4 by Western and in contrast to SCA5, EAAT4 levels were similarly reduced in RIPA and Urea extracts.

EAAT4 immunostaining of remaining Purkinje cells in SCA5 showed a consistent thinning of the dendritic arbor and darker staining of the cell body (Fig. 4e), while SCA1 transgenic animals showed uniform but lighter staining (Fig. 4f). These results indicate that the redistribution of EAAT4 in SCA5 is not caused by Purkinje cell degeneration and that EAAT4 is likely altered by different mechanisms in SCA1 and SCA5.

To further examine EAAT4 and to determine if mutant spectrin also causes changes in other membrane bound Purkinje cell proteins, subcellular fractionations of cerebellar tissue and subsequent Western analyses were performed (Fig. 5). Total protein loaded in the P1 and S1 fractions was determined by BCA protein assays, with the following amounts of protein in the respective lanes: P1 control (40.5  $\mu$ g), S1 control (5.5  $\mu$ g), P1 SCA5 (71.4  $\mu$ g), S1 SCA5 (3.9  $\mu$ g). Protein loading was also estimated by normalization of the Western blot membranes to clathrin light chain, a broadly expressed control protein known to cycle on and off plasma and vesicle membranes and to be abundant in membrane rich pelletable fractions. As expected, considerable enrichment of clathrin was observed in the predicted nuclear (P1), crude synaptosomal (P2), and enriched synaptosomal (LP1) fractions. More protein was loaded in SCA5 vs. control in the P1 (71.4 vs 40.5  $\mu$ g), P2 and LP1 fractions (see clathrin loading control) with slightly less protein in the SCA5 S1 (3.9 vs 5.5  $\mu$ g) fraction compared to control. Subcellular fractionations of EAAT4 and GluR $\delta$ 2 from SCA5 cerebellar extracts differ from control cerebellum. For example, because more protein was loaded in SCA5 P2 and LP1 fractions vs. control P2 and LP1 fractions (determined by clathrin), if EAAT4 in the SCA5 and control homogenates were fractionating in the same way, more EAAT4 would be expected in the overloaded SCA5 P2 and LP1 fractions. However, dramatically less EAAT4 is found in these SCA5 synaptosomal rich fractions (P2, LP1). Similar

redistribution of the GluRδ2 are found with markedly less than predicted amounts of GluRδ2 in the SCA5 P2 and LP1 fractions compared to control P2 and LP1. In contrast to control, the synaptic membrane proteins EAAT4 and GluRδ2 were not enriched in the synaptosomal fractions in SCA5 tissue, suggesting that mutant β-III spectrin affects the

5 cellular localization of these proteins.

To further characterize the physiological effects of mutant β-III spectrin on EAAT4 a series of controlled cell culture experiments were performed. HEK293 cells were transfected with eGFP-EAAT4 and total internal reflection fluorescence (TIRF) microscopy was used to follow the lateral movement of the glutamate transporter on the

10 cell's membrane. The glutamate transporters normally alternated within seconds between two main states: periods of rapid movement on the cell's membrane and restricted motion within a sub-micrometer area (Fig. 4g-i). When EAAT4 was expressed along with an empty control vector, almost 40% of the EAAT4 diffraction spots were actively moving at or near the plasma membrane (~4 microns), while the slow moving diffraction spots

15 were typically restricted to movements in a fixed small area (less than 1 micron) (Fig. 4g, Table 3). To further investigate the physiological relevance of the interaction between EAAT4 and wildtype β-III spectrin, EAAT4 was co-transfected with β-III spectrin and followed the trafficking of EAAT4. Consistent with previous biochemical studies (Jackson et al., Nature 410, 89-93 (2001)), co-expression of wildtype β-III spectrin

20 stabilized EAAT4 with only 5% of diffraction spots moving at or near the membrane, and none showing large lateral movements (>4 microns) (Fig. 4h, Table 1). However, in the presence of mutant β-III spectrin with the 39 bp deletion, the stabilization of EAAT4 was lost, and the transporter was highly motile with many lateral movements over 4 microns observed (Fig. 4i, Table 3). To confirm the specific interaction between EAAT4 and β-III

25 spectrin, β-III spectrin was co-transfected with EAAT3, another glutamate transporter also expressed in Purkinje cells. Neither wildtype (Table 3) nor mutant β-III spectrin had any substantial effect on EAAT3 stability. The lack of an effect on EAAT3 does not exclude the possibility that mutant β-III spectrin affects other membrane proteins. These studies however, provide evidence that mutant β-III spectrin can disrupt the stability of

30 EAAT4 and because altered expression of EAAT4 on the membrane is known to increase Purkinje cells to injury/degeneration it therefore may contribute to Purkinje cell degeneration in SCA5 (Welsh et al., Adv Neurol 89, 331-359 (2002)).

Table 3. Mutant β-III spectrin alters lateral trafficking of glutamate transporters.  
TIRF microscopy of HEK293 cells was performed and digital movies of the imaged cells were evaluated using Metamorph. Each diffraction spot was analyzed separately. For each condition 3-6 different experiments were recorded from different dishes and  
5 different days. The results are mean ± SD.

construct	Total diffraction	% of diffraction
	spots analyzed	spots not moving
eGFP-EAAT4 + empty vector	685	62.0±8.7
eGFP-EAAT4 + wildtype β-III spectrin	122	94.2±9.7
eGFP-EAAT4 + mutant β-III spectrin	375	67.5±4.4
eGFP-EAAT3 + empty vector	547	61.0±11.2
eGFP-EAAT3 + wildtype β-III spectrin	337	58.7±5.4

We report a novel mutational mechanism for spinocerebellar ataxia with the identification of three separate mutations in the β-III spectrin gene (*SPTBN2*) responsible  
10 for SCA5. The American and French families have similar but separate in-frame deletions within the third spectrin repeat, and are likely to disrupt the highly ordered triple-alpha-helical structure of the repeat changing the overall shape of the tetrameric alpha-beta-spectrin complex. Although it is possible that some feature of the shared haplotype between the American and French families led to similar microdeletions, it  
15 appears more likely that the shared haplotypes are coincidence as this haplotype is found on 3.5% of control chromosomes. The German family has a missense mutation in the calponin-homology domain, which may disrupt the ability of spectrin to bind to the actin cytoskeleton and similarly affect the stabilization of membrane proteins or cause alterations in transport by disrupting binding to ARP1 and the dynein motor complex  
20 (Holleran et al., J. Biol. Chem 276, 36598-36605)).

The cell fractionation studies suggest that mutant β-III spectrin (39bp deletion) affects localization of the synaptosomal proteins EAAT4 and GluRδ2. Interestingly, EAAT4 is also affected in SCA1 transgenic mice with the downregulation of transcript levels (Lin et al., Nat Neurosci 3, 157-163 (2000)) and (Serra et al., Hum Mol Genet 13,  
25 2535-2543 (2004)). Further evidence for the possible role of EAAT4 in ataxia comes

from intracisternal antisense knockdown experiments in rats which resulted in progressive ataxia (Raiteri et al., Prog Neurobiol 68, 287-309 (2002)). In addition, mutations in GluRδ2 cause ataxia in both *lurcher* and *hotfoot* mice (Lalouette et al., Genomics 50, 9-13 (1998)) and (Zuo et al., Nature 388, 769-773 (1997)). Loss of EAAT4 and GluRδ2 at 5 the plasma membrane in SCA5 could lead to glutamate signaling abnormalities, which over time could cause Purkinje cell death in SCA5.

The reported interaction of spectrin with the dynactin-dynein motor complex suggests that SCA5 mutations could affect protein trafficking as in other neurodegenerative diseases. These disorders include a dominantly inherited motor neuron disease caused by mutations in p150<sup>Glued</sup>, a subunit of dynactin (*DCTN1*) (Puls et 10 al., Nat Genet 33, 455-456 (2003)) and a motor neuronopathy caused by missense mutations in the mouse dynein heavy chain gene (*Dnchc1*) (Hafezparast et al., Science 300, 808-812 (2003)). In Huntington disease, alterations of the huntingtin/HAP1/p150<sup>Glued</sup> complex induce transport deficits and loss of neurotrophic 15 support contributing to neuronal toxicity (Gauthier et al., Cell 118, 127-138 (2004)), and axonal transport defects are found in Alzheimer's patients and murine models (Stokin et al., Science 307, 1282-1288)).

Identifying additional mutations in *SPTBN2* that cause ataxia in families with unknown mutations will provide further insight into the functions of β-III spectrin and the 20 molecular mechanisms of neurodegenerative diseases. Specifically, it will be of interest to determine if mutations in *SPTBN2* also cause SCA20, a clinically distinct form of ataxia whose critical region includes *SPTBN2* (Knight et al., Brain 127, 1172-1181 (2004)). It will also be important to determine if mutations in *SPTBN5* or *SPTBN1*, 25 which map to the SCA11 and SCA25 critical regions respectively, also cause ataxia (Worth et al., Am J Hum Genet 65, 420-426 (1999)) and Stevanin et al., Ann Neurol 55, 97-104 (2004)). Consistent with the possibility that the β-spectrins may play additional 30 roles in disease, dominantly inherited mutations in a beta spectrin homologue cause an uncoordinated phenotype (*unc-70*) in *C. elegans* (Park et al., Genetics 113, 821-852 (1986)) and recessive mutations in the mouse spectrin beta 4 gene (*Spnb4*), an orthologue of human beta-IV spectrin (*SPTBN4*), cause a progressive ataxia with hind limb paralysis, deafness and tremor in *quivering* mice (*qv*) (Parkinson et al., Nat Genet 29, 61-65 (2001)).

The current estimate of 28 dominant ataxia loci provides an opportunity to use human genetics to define the fundamental causes and common molecular pathways

underlying this group of neurodegenerative diseases (Schols et al., Lancet Neurol 3, 291-304 (2004)). Interestingly, down-regulation of both  $\beta$ -III spectrin and EAAT4 transcripts found by micro-array analysis in two murine ataxia models, SCA1 transgenic and *staggerer* mice (Gold et al., Neuron 40, 1119-1131 (2003)) suggests the convergence  
5 pathogenic mechanisms triggered by distinct mutations. The identification of SCA5 mutations in a gene encoding a well known cytoskeletal protein will allow testing of specific hypotheses of disease pathogenesis involving destabilization of membrane proteins, glutamate dysregulation and vesicle trafficking deficits which will provide insight into the downstream molecular mechanisms common to SCA5 and other  
10 neurodegenerative diseases.

The history of ataxia in the Lincoln family raises the question of whether President Abraham Lincoln carried the SCA5 mutation. Historical descriptions suggest that the President had an uneven gait – an early sign of ataxia. On March 27, 1861, William Russell a reporter for the London Times wrote of Lincoln “Soon afterwards there entered, with a shambling, loose, irregular, almost unsteady gait, a tall, lank, lean man...”  
15 The identification of the SCA5 mutation makes it possible to unequivocally determine if President Lincoln carried the mutation using preserved artifacts containing his DNA. In 1991, the identification of a Marfan’s gene sparked debate concerning the testing of President Lincoln’s DNA to determine whether his tall stature could have resulted from  
20 that disease (McKusick., Nature 352, 279-281 (1991)). Unlike for Marfan’s syndrome, the Lincoln family history indicates President Lincoln was at risk of developing SCA5. Determining President Lincoln’s status relative to SCA5 would be of historical interest, and would increase public awareness of ataxia and neurodegenerative disease.

25 The complete disclosure of all patents, patent applications, and publications, and electronically available material (including, for instance, nucleotide sequence submissions in, e.g., GenBank and RefSeq, and amino acid sequence submissions in, e.g., SwissProt, PIR, PRF, PDB, and translations from annotated coding regions in GenBank and RefSeq) cited herein are incorporated by reference. The foregoing detailed description and  
30 examples have been given for clarity of understanding only. No unnecessary limitations are to be understood therefrom. The invention is not limited to the exact details shown and described, for variations obvious to one skilled in the art will be included within the invention defined by the claims.

Unless otherwise indicated, all numbers expressing quantities of components, molecular weights, and so forth used in the specification and claims are to be understood as being modified in all instances by the term "about." Accordingly, unless otherwise indicated to the contrary, the numerical parameters set forth in the specification and claims are approximations that may vary depending upon the desired properties sought to be obtained by the present invention. At the very least, and not as an attempt to limit the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits and by applying ordinary rounding techniques.

10 Notwithstanding that the numerical ranges and parameters setting forth the broad scope of the invention are approximations, the numerical values set forth in the specific examples are reported as precisely as possible. All numerical values, however, inherently contain a range necessarily resulting from the standard deviation found in their respective testing measurements.

15

All headings are for the convenience of the reader and should not be used to limit the meaning of the text that follows the heading, unless so specified.

What is claimed is:

1. A method comprising:
  - analyzing an *SCA5* polynucleotide; and
  - determining whether the *SCA5* polynucleotide comprises a mutation.
2. The method of claim 1 wherein a subject at risk of having *SCA5* has a mutation in an *SCA5* polynucleotide.
3. The method of claim 1 wherein a subject not at risk of having *SCA5* does not have a mutation in an *SCA5* polynucleotide.
4. The method of claim 1 wherein the *SCA5* polynucleotide is a genomic *SCA5* polynucleotide.
5. The method of claim 1 wherein the *SCA5* polynucleotide is a processed *SCA5* polynucleotide.
6. The method of claim 1 wherein the analyzing comprises amplification of the *SCA5* polynucleotide.
7. The method of claim 1 wherein the analyzing comprises hybridization of the *SCA5* polynucleotide to a second polynucleotide.
8. The method of claim 1 wherein the analyzing comprises sequencing a portion of the *SCA5* polynucleotide.
9. The method of claim 1 wherein the mutation is present in an exon.
10. The method of claim 9 wherein the *SCA5* polynucleotide encodes an *SCA5* polypeptide, and wherein the mutation in the *SCA5* polynucleotide results in an *SCA5* polypeptide having an amino acid sequence different than SEQ ID NO:2.

11. The method of claim 1 wherein the mutation is selected from a mutation in a nucleotide corresponding to exon 7 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 12 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 14 of the *SCA5* polynucleotide, or a combination thereof.

12. The method of claim 1 further comprising determining whether the subject displays at least one symptom of ataxia.

13. A method for identifying a subject not at risk for developing spinocerebellar ataxia type 5 comprising:

analyzing nucleotides of *SCA5* polynucleotide; and

determining if the polynucleotide comprises a mutation, wherein a subject not at risk of having *SCA5* does not have a mutation in an *SCA5* polynucleotide.

14. The method of claim 13 wherein the *SCA5* polynucleotide is a genomic *SCA5* polynucleotide.

15. The method of claim 13 wherein the *SCA5* polynucleotide is a processed *SCA5* polynucleotide.

16. The method of claim 13 wherein the analyzing comprises amplification of the *SCA5* polynucleotide.

17. The method of claim 13 wherein the analyzing comprises sequencing a portion of the *SCA5* polynucleotide.

18. The method of claim 13 wherein the *SCA5* polynucleotide encodes an *SCA5* polypeptide, and wherein the *SCA5* polypeptide has the amino acid sequence SEQ ID NO:2.

19. A method for identifying a subject at risk for developing spinocerebellar ataxia type 5 comprising:

analyzing nucleotides of *SCA5* polynucleotide; and  
determining if the polynucleotide comprises a mutation, wherein a subject  
at risk of having *SCA5* has a mutation in an *SCA5* polynucleotide.

20. The method of claim 19 wherein the *SCA5* polynucleotide is a genomic *SCA5* polynucleotide.

21. The method of claim 19 wherein the *SCA5* polynucleotide is a processed *SCA5* polynucleotide.

22. The method of claim 19 wherein the analyzing comprises amplification of the *SCA5* polynucleotide.

23. The method of claim 19 wherein the analyzing comprises sequencing a portion of the *SCA5* polynucleotide.

24. The method of claim 19 wherein the mutation is present in an exon.

25. The method of claim 24 wherein the *SCA5* polynucleotide encodes an *SCA5* polypeptide, and wherein the mutation in the *SCA5* polynucleotide results in an *SCA5* polypeptide having an amino acid sequence different than SEQ ID NO:2.

26. A method for determining whether a subject has spinocerebellar ataxia type 5 (*SCA5*), the method comprising analyzing an *SCA5* polynucleotide for a mutation, and determining whether the subject displays a symptom of *SCA5*, wherein having a mutation in an *SCA5* polynucleotide and having a symptom of *SCA5* indicates the subject has *SCA5*.

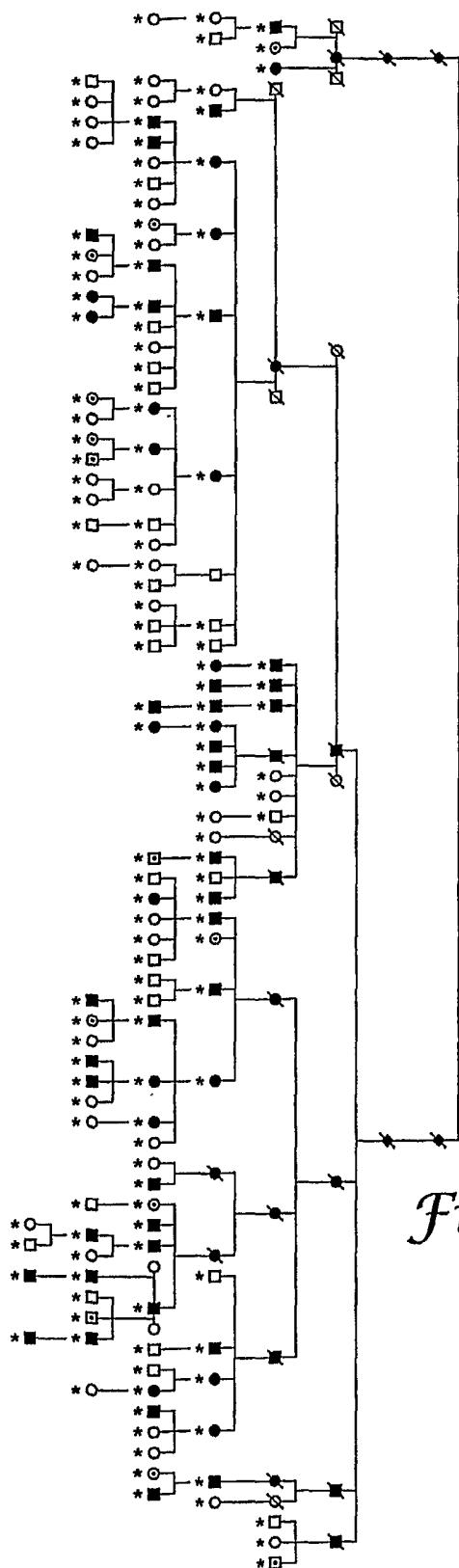
27. The method of claim 26 wherein the *SCA5* polynucleotide is a genomic *SCA5* polynucleotide.

28. The method of claim 26 wherein the *SCA5* polynucleotide is a processed *SCA5* polynucleotide.

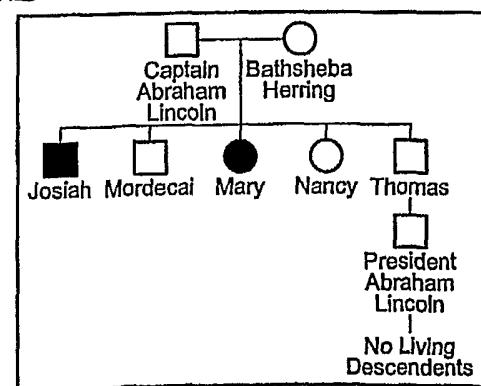
29. The method of claim 26 wherein the analyzing comprises amplification of the *SCA5* polynucleotide.
30. The method of claim 26 wherein the analyzing comprises sequencing a portion of the *SCA5* polynucleotide.
31. The method of claim 26 wherein the mutation is present in an exon.
32. The method of claim 31 wherein the *SCA5* polynucleotide encodes an *SCA5* polypeptide, and wherein the mutation in the *SCA5* polynucleotide results in an *SCA5* polypeptide having an amino acid sequence different than SEQ ID NO:2.
33. A kit for detecting an *SCA5* polynucleotide, comprising a primer pair that will amplify a portion of an *SCA5* polynucleotide.
34. An isolated polynucleotide comprising a mutant of SEQ ID NO:1 or a portion thereof.
35. The isolated polynucleotide of claim 34 wherein the mutation present in the polynucleotide is selected from a mutation in a nucleotide corresponding to exon 7 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 12 of the *SCA5* polynucleotide, a mutation in a nucleotide corresponding to exon 14 of the *SCA5* polynucleotide, or a combination thereof.
36. The isolated polynucleotide of claim 34 wherein the isolated polynucleotide is 15 to 500 nucleotides.
37. A vector comprising the isolated polynucleotide of claim 34.
38. A cell comprising the vector of claim 37.
39. A polynucleotide selected from SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID

NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, 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, SEQ ID NO:34, 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, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, or SEQ ID NO:77.

Fig. 1A

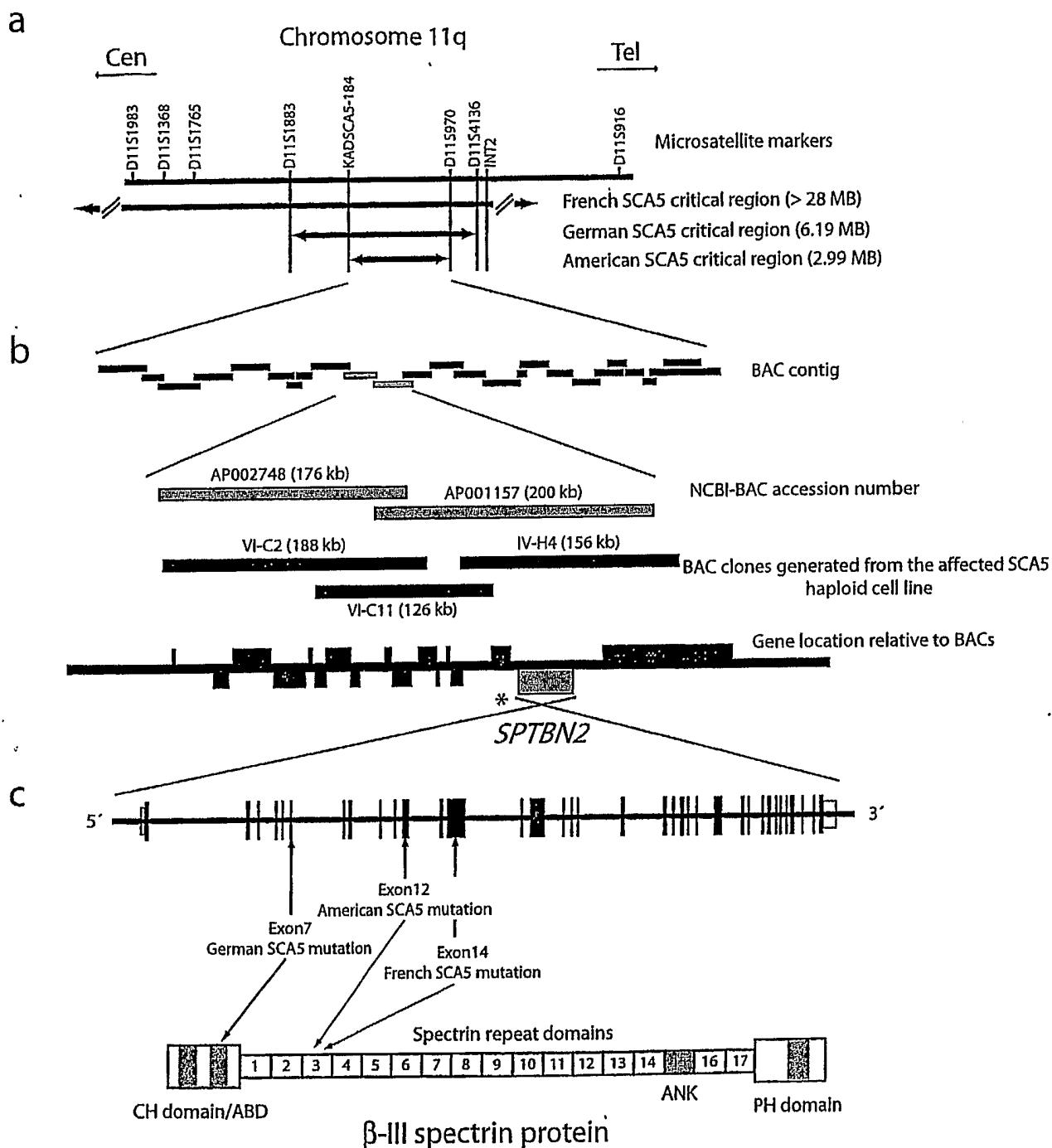


*Fig. 1B*

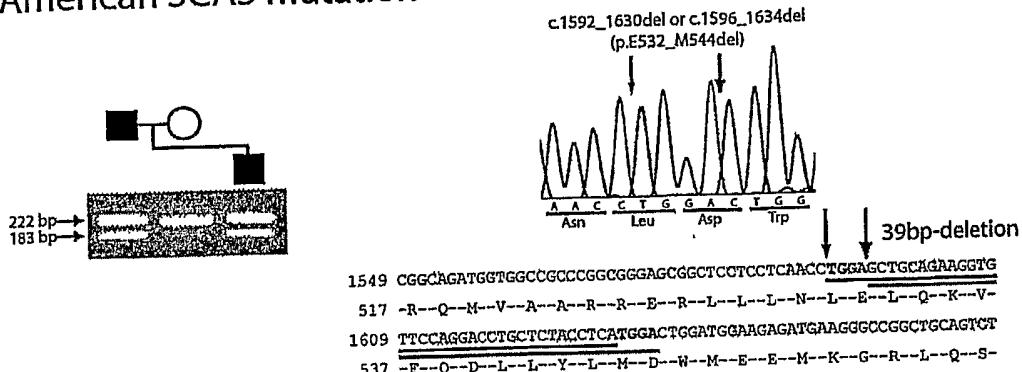


**SUBSTITUTE SHEET (RULE 26)**

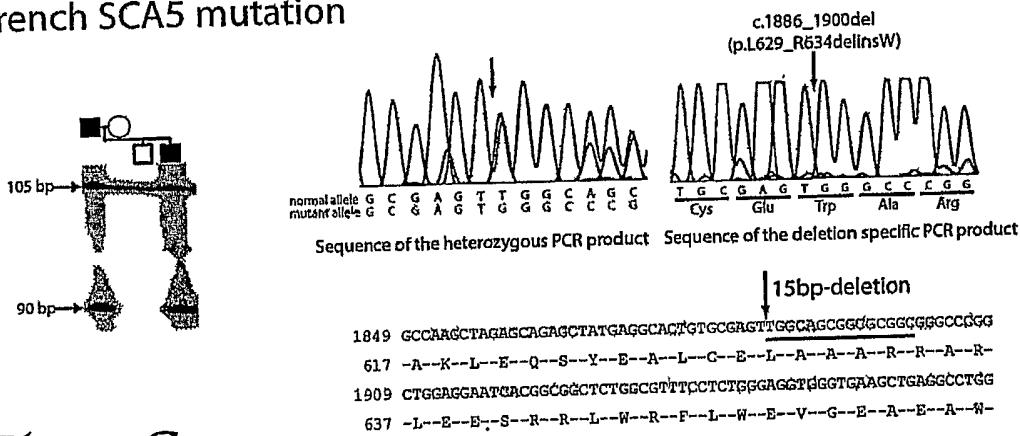
Fig. 2



**Fig. 3A**  
American SCA5 mutation



**Fig. 3B**  
French SCA5 mutation



**Fig. 3C**  
German SCA5 mutation

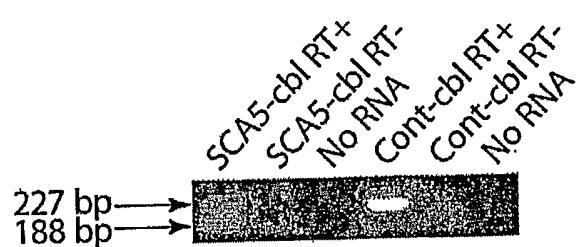
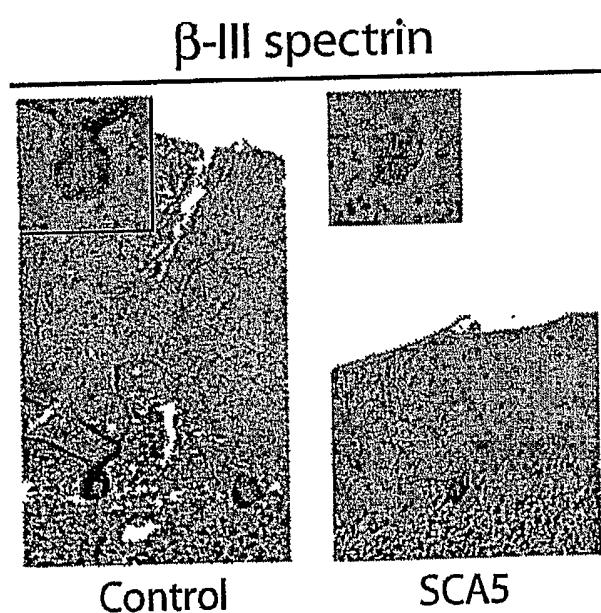


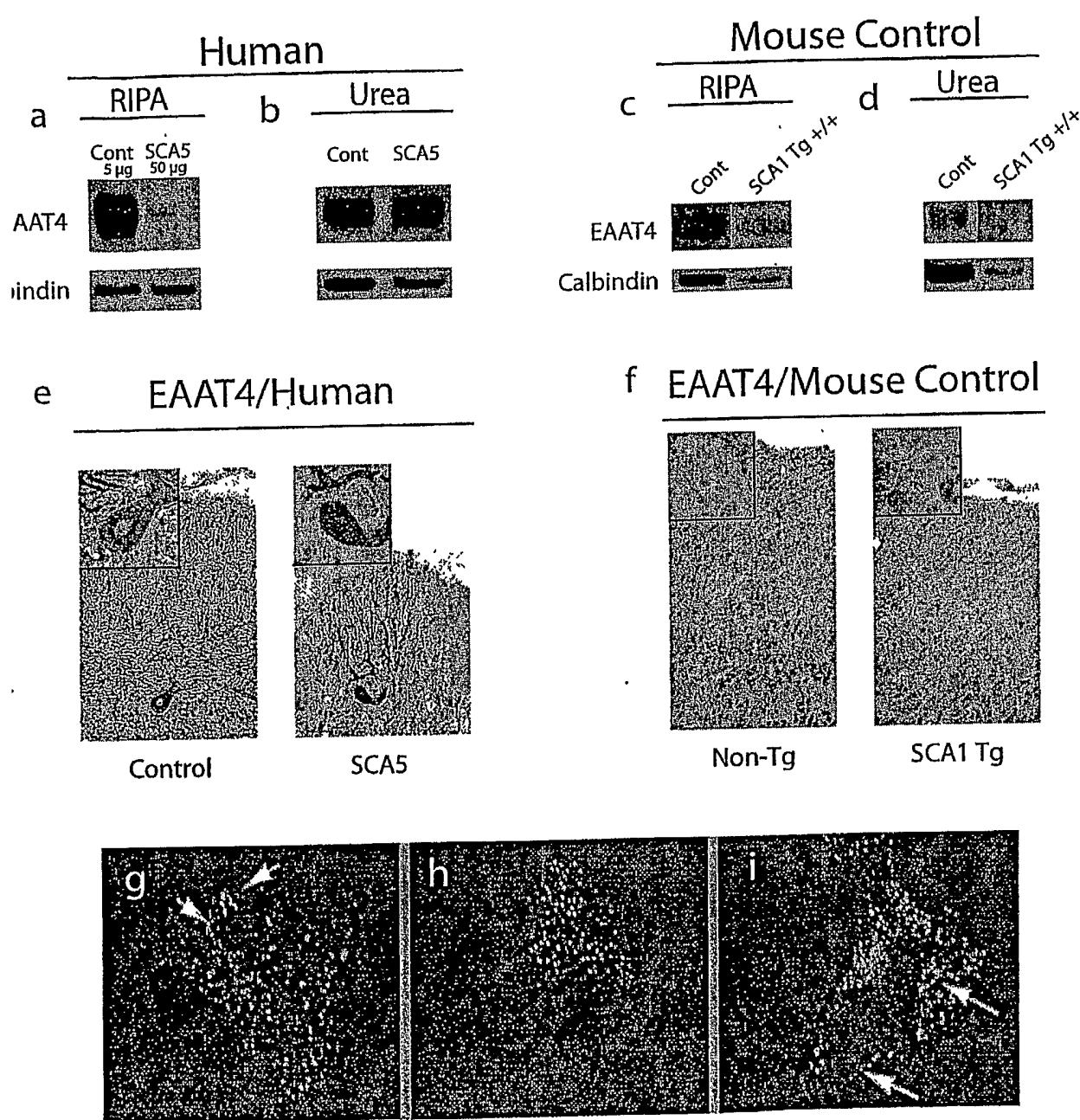
Conservation of calponin homology domain and leucine 253 among human beta spectrins

SPTBN1	SNAYNLQNQAFNLAEQHLGLTKI	LDPEDISVDH PDEKSIITYVVTYYHYFSKMKALAVEG
SPTBN2	CNAHYNLQNQAFNLAKEGLGLTKI	LDPEDISVDH PDEKSIITYVATYYHYFSKMKALAVEG
SEPTB	SNARKNLEHAFNVAEQLGLI	LDPEDISVDH PDEKSIITYVATYYHYFSKMKALAVEG
SPTBN4	SNANYNLQRARFTABQHLGLAR	LDPEDVNMEAPDEKSIITYVVSFHYFSKMKALAVEG
SPTBN5	DRPLHNLAFAFLVAEQELGIAC	LDPEDVAAAQPDERSIMTYVSLYHYCSRLHQGQTVO

Evolutionary conservation of calponin homology domain and leucine 253 among species

Mouse	Spnb3	LKKCNAHYNLQNQAFNLAKEGLGLTKI	LDPEDVNVDQPEDEKSIITYVATYYHYFSKMKALAVEG
Rat	Spnb3	LKKCNAHYNLQNQAFNLAKEGLGLTKI	LDPEDVNVDQPEDEKSIITYVATYYHYFSKMKALAVEG
Dog	SPTBN2	LKKCNAHYNLQNQAFNLAKEGLGLTKI	LDPEDVNVDQPEDEKSIITYVATYYHYFSKMKALAVEG
Human	SPTBN2	LKKCNAHYNLQNQAFNLAKEGLGLTKI	LDPEDVNVDQPEDEKSIITYVATYYHYFSKMKALAVEG
Chimp	SPTBN2	LKKCNAHYNLQNQAFNLAKEGLGLTKI	LDPEDVNVDQPEDEKSIITYVATYYHYFSKMKALAVEG
Worm	unc-70	LQKSNALYNLQSADFDTENQLGLAKELDAEDVNVDQPEDEKSIITYVVTYYHYFSKMKALAVEG	LQKSNALYNLQSADFDTENQLGLAKELDAEDVNVDQPEDEKSIITYVVTYYHYFSKMKALAVEG
Fly	beta-Spec	LSKTNIAHNLNNAFDVaedKLGLAKELDAEDVFVEHPDEKSIITYVVTYYHYFSKMKALAVEG	LSKTNIAHNLNNAFDVaedKLGLAKELDAEDVFVEHPDEKSIITYVVTYYHYFSKMKALAVEG

*Fig. 3D**Fig. 3E*

*Fig. 4*

*Fig. 5*

## SCA5

## Control

	P1	S1	P2	S2	LP1
EAAT4					
GluRδ2					
Clathrin, light					

*Fig. 6A*

CCACTGAGCAGCCAACCGCAGCCTCTGCCACAAGGAGAGCGGAGCACAGgtagggcaag  
 .....  
 aagacagaaggccccggtggtgagtggagggcttggccctggacacccctcactgg  
 ccacatttccttgcatGAGCAGGAAGCCGCCTACCACCATGAGCAGCACCGCTGTCACCC  
 .....-M--S---T--L--S--P-  
 ACAGACTTGACAGCTTGAAATCCAGGGCCAGTACAGTGACATCAACAACCGCTGGGAC  
 -T--D--F--D--S--L--E--I--Q--G--Q--Y--S--D--I--N--N--R--W--D--  
 CTTCCCTGACTCGGACTGGGACAATGACAGCAGCTCGGCCCGCCTCTTGAGAGGTCTCGC  
 -L--P--D--S--D--W--D--N--D--S--S--A--R--L--F--E--R--S--R--  
 ATTAAGGCTCTGGCAGGttaggtcagaggaggcgaggtggggatgtggaggaggc  
 -I--K--A--L--A--  
 tttgtcagctcgcatgaagagcccttaacttatggaaagactgactttcttctaa  
 gtggagacccagagcactctatcaattccctggtccccacatccagcctcaagaatagg  
 ctccaggccattcagaactttcccaagctttcccccaagcatagagacgcttactgg  
 cccctggactggctgagctgactaaaacagcttccctctggcctcagccccgggtca  
 tgccttcatccttcggactctcgctttgccactccttcggcctaccttc  
 tacctgtcatgccttggctgtcttgcattttgattccactgtcattttccacacctgttc  
 ccagtttattttccctctgtcccttcacactgtttccatgtccacccgttac  
 tcccattatctcatctgcctgttccctgaagcaggttccctccacccgtgtgggt  
 tctgcagccgcccatttccctgggtgtctctgcctgcctggcttccacttccct  
 gcagcccagggcctggctccgcctgtgagcggaaactcgtccaacggcctccgtct  
 tccggggaggagttgtcatggcaacaagagtgcggctgccacactgacaccaccaggc  
 tggcatggggctgactcacccacacaccccaaggccacacttccatggcttccct  
 ccatcccactttgtcattgcattgacactgtggctgcccagcccaaccccaacc  
 cagtcgtgggggtcccgactactctcaccgagagggtggtgctccgagtggtgggggt  
 tggccactagaagaagagagactccaaggggagggttccaaacctggccacatgttt  
 caggatcacccaggatgttttacaaatacaggatccaaagctcccaagtgttt  
 taagttagaatcttggtccagaactcaggatccaaagctcccaagtgttt  
 gagccagctcagcaccaggctgaggacccctgtggcagccatagctccaggaccc

*Fig. 6B*

ttctcatttctcctggcctctgtccttcgttgggtttcaggccggctgtgaaa  
 cactcctaccagtgtgtctatacatcagagaagacagaagggtgtgggttatggagaatgt  
 catgttgggaaaatgttccctggcactctgggtttggattcatgttagcctctagaactg  
 tcatgatgtctatgtgtacaacatgtgccacccaccatgttgacacatgacaaatg

rs11602953

ccatccaggactgtcaccatgcctgtgtacataggacacatgcctcccaggactctcacc  
 atgcctgtgtacataggacacatgccacccaggactgtcaccatgtctggctgcacataa  
 ctgggccacccageactgtcaccatgtctgtcacacatctgtgtgacctcgtctc  
 attgggtgttatctccaccactgaactggctctcctgaagaggcagctctgcatgggaa  
 tttgcctagcactgtgtctccagccagtaggtgtcaaataaacactgaatgaatgaagat  
 gaggaaggcaagtgtaggctgcacatgacaaacgctgaagcccagaggaagggcgctg  
 gtatatgtgcccagcagccccgcgtctggggcagggatggcctgggtggaggatggccct  
 ggtggccagagggcagggaaaggggctggcactggaggtggtgataagcaagtggaaagg  
 atggtaacagccctcgttcacagatgagggagtgaaggaaaatgcctcgcacagtgcct  
 ggctttcctctttgaggatttcctcgttccagattctttgtctgtttctgagctg  
 gggctcaaagcccacacatcctggagaaggacctcccagagtcacccagagaggggtcgc  
 cctcgctgttcctgcaaaccctctggaaacacaggagttgcattggggccctctcggtgg  
 gtctgtgcattgttagtaacttggggctgtggcgggtcttgaaggcgctggcagt  
 gcccctcccgccgccccgcagggcctggccctgccccccggccacactccgcctctc  
 tccctcccaacactgtccaacctgtttgtctggagccgcctgtctcctctcaccgcag  
 ccaggcgtgccccactctctgtcaggagttcgggacaccccgtaacttttgcca  
 ctccattgaaattcccaccattcttgaatcccctgacgctggactccagcctgaggttcca

rs551708

ctggggAACAGAAACTGGCCCTGCCAGCGGGGAGCATGGACTCTGGCCCTGGCTCT

rs7118311

gtAACTAGCAGCTGGCAGCATGAATGGAACGGAGTGAGGGGTGCACGGGAACGGCCT  
 gcatgaggccagcgagtttactatgaggctgtggaaaggcccacaacccggggccct  
 cctgcatttaccagctgtttcatcaaccctgctcagtatgccagcgtgctggaaaggacg  
 cttcaaaacagctgcaagggtgaggccccggggccctgtggggacgactgtgcctggcg  
 agagggacctgggtgtggggagagtggccggagctgaggacctggcacaggacacagtgt  
 gactccaggatgttgtggagacaacgtgatgcagcttagagggataattacaagag

**SUBSTITUTE SHEET (RULE 26)**

9/33

*Fig. 6C*

rs11820790  
 gccggaaaatacacagatgtcaacggaatggggaaaggcttgacctgctggcagggacc  
 gtgtatgatggcatgggtgggtgtactgccggcctgggtgctgcatggagcaaagg  
 caactagccatggaggcgtgggggtgccagcacctgggaggcgccaggcagtgccctctg  
 aggtggatccggatctggcggtggagcgctggccaggacagttaggggaggcgggagggggt  
 gagccaggcccagggacctggagattgtgtttccaggccttgagcgtgtgtctgaatg  
 tgtacacggggcagtgtgtaggaagcaaggatgtctcgaaggaggatctgtgc  
 aaacaggcccaggttagatgtctgcgcaggtgttcactctggcacagtgcttaggt  
 gtgttttggaatactgttgtgtgtgtgttaaggatctaggatgtgtggcgt  
 cttgtgacgtgacacttggttctgggcttggcttgcccattttcccttagttgcca  
 ctaattctgtgtttctgactgtgtgccaaatggcctcacccagggtgcccagcgatgg  
 gggcagtgcccctggatccagcaatggcctgtgacccacgtccctggctcagatgg  
 gccctgtagacctgccggaggccccaaggcacttggttcatgggtctggggagcatg  
 gggccgtcccatgttattaggaccccaggtttgtgccagtatagtggctggagggtgaaa  
 gacacggagcgacgtgggtgtgtttgtcttatgggtgaggtgcttaggaatgtgag

rs4930389  
 attaaggatgagtcttcccggttctgagatgtctggtggagacggaagcaaagttg  
 ctcttcagaccccagccacttgggtggtccttggcacaggctctgtggaagtgg  
 cagaaatcctcatcaccagagcagttctgatgtcttcactttgtcacttagaacgttct  
 taattttatttcctacggaacctagtatttacttcccttcttagctccctccc  
 ctgtggctgggcagcgaggcattgccttaacttttttttagttgagattggtactctgc  
 ctggctttttccccatgttccctggctcaactgttttcccagtcccaactctcaac  
 atcttgttcttcttaagcaatgggtgttatgtcctggaactggatcccttgcagccag  
 ctgaatgtactgtccccacccccagaaggcagagggccctgcctttttttaagtttatcc  
 aaggacagtaagctgactcaggaggctgcagccctctataaccatcaagtcttaatgaccg  
 ctttggaagagccgctgcataatttagcccacacctttcctttgtctgcaacccaaag  
 agaacagccatacctcccacccggctgggcagcttgcctctcctttcgggccagatcaca  
 ccactttataacagcagctacccaagggccatttgtgccaggccaccaactgttttt

rs5792396  
 tataacttatctcattaatcaacacaatgacctgtaagtagaagtgtatgttccc  
 attttacagatgacacatttgggtgtatgtatgtaagtgttgaccacttccactg

**SUBSTITUTE SHEET (RULE 26)**

*Fig. 6D*

cttccacqactgccccatcccagcccctccctgcattgtcccatggttctgaggatgaaa

rs658538

gagacggagtcttgctctgtcgcccaggctggagtgcagtggcacaatctcggtccctg  
caagctccccctccaggttacgtcagtctgcctcagcctcccgagtagctggactac

rs657696

aggcgcggccactatgccggctaatttttgtaattttagtagagacgggtttcacc

rs10702473

qtqtttagccaggatggtctcgatctcctgaccttgtatccgcc\_cgccttgcttccaa

agtgctgggattacaggcatgagccaccacccaaaccttcgcctatttctgtctttca

rs9666734

tctttcttcttacaccttgcttatgagctcattgcaaggtaggctttcccccgtcat

cttccttcccttcctccttagaggccccgcttccatctctcaggcttccatcc

ccccccctqccaactggtgtttagaaggcaggcatactgaactcgagggacagaataaa

ccccatgttgccaccacagATGAACGAGAAGCTGTGCAGAAGAAAACCTTCACCAAGTGG

D--E--R--E--A--V--Q--K--K--T--F--T--K--W-

GTAAACTCGCACCTGGCCCCGGTCACGTGCCGGTGGGGGACCTGTACAGCAGCTCCGG  
-V--N--S--H--L--A--R--V--T--C--R--V--G--D--L--Y--S--D--L--R-

GACGGACGCAACCTGCTGAGGCTCCTCGAGGTGCTCTCGGGAGAGATACTGgtgagctgt  
-D--G--R--N--L--L--R--L--L--E--V--L--S--G--E--I--L-

rs12805133

ggtcatgaagggaagtgggggggccttagaaaaagtgtccaaaggggaccagctggaaac

rs11828633

agccctcagagagagtgacggcacagggcgaaagcagtggctcccttgctcagaggat  
gggttcatttcaagtgccttgccacttagattacataggagcattgcttcatggagaca

rs5792395

ggaacagtggccccattgaaaacattttttccacctactctgggtgcaattaatcc

cttggaaatcttaatctgcccaagcctcatctgttaagctcagctttggctccagccctgg

ggctgtgcctgtctgtttccctgaggttaggtcatggggaaagggtgtgcaduaggctggggcc

cacag CAAAGCGTACAAAGGGCCGATGC GGATCCACTGCCTGGAGAACGTGGACAAG  
-P--K--P--T--K--G--R--M--R--I--H--C--I--E--N--V--D--K-

11/33

*Fig. 6E*

-A--L--Q--F--L--K--E--Q--K--V--H--L--E--N--M--G--S--H--D--I--

GTGGACGGAAACCACCGACTGACCCTGGGCTGGTCTGGACCATCATCCTTCGATTCCAG  
 -V--D--G--N--H--R--L--T--L--G--L--V--W--T--I--I--L--R--F--Q--  
 gtaccccagcacactgtcacacagggtgtggttctgcacctggctctgcaccgcaagccg  
 ttcctgctggcccatgacacaacaaacacaaagatggaggatctagaaccttatgttagca  
 gctgccatggcggcgctatccaagcaccagatcagtggactccccctttcagcaggta  
 gaaaccagcttgggtgaatgtcacacccctgtggcaaattccctgtgttctggctgtgt  
 gttggcgtgtgcaatgggaccaggaccacagggtggccacaggagactggaagttta  
 ccaactggtccttaccacaggattggaaacttctagagggtctggtcactggcacca  
 attccctgccataaactctgcctgcctccccatggaacagtaccctcctcctca

rs529203

tgccccagtgtcctcatgcccatcacagcgttctgatcccccttcctccgggtct  
 gtcggcttacccctggagagccagcttttagagacccctggagagccagctctcagaaaat  
 gttcactttctgcctgcagttccccacctggaggaccaggaccctccactcagacaca  
 cccattttcagcatccatccctaattctgcctcaaattccctgcacccagtgttttgc  
 tggccctggcttaggttgaaggacccctggatgcctgtccacccattaccacagt  
 ctaaagagcagttccctgccatttatctgagcatctgtataacaccaggagttcgtc  
 cagggactgagggccctgttgccttccttagATCCAAGACATCAGTGAGACA  
 -I--Q--D--I--S--V--E--T-

GAAGACAACAAGGAGAAGAAGTCAGCCAAGGATGCCCTGCTCTGTGGTGCCAGATGAAG  
 -E--D--N--K--E--K--S--A--K--D--A--L--L--L--W--C--Q--M--K--

ACTGCAGGgtgaggacaccctggcgtgtggactggagggtcagtgaccccccaggctgt  
 -T--A--G

gctggctccaggaaccatgagctggtgactgcctggagtcatagaacagagctcagcc  
 ctactcaccatttctgaatctctgtttatagTTATCCCAACGTCAATGTACACAACCTTC  
 --Y--P--N--V--N--V--H--N--F--

ACCACCAAGCTGGAGAGATGGACTAGCTTCAACGCCATCGTGCATAAACACCGgtgagaa  
 -T--T--S--W--R--D--G--L--A--F--N--A--I--V--H--K--H--R

gatggggtagcaaacgcagcagactgtggctggctaatgacccctgggtgcacatgt  
 ggagacagagatgagatgacgcacagcagtcagacggagaaatagcactcggagcag  
 aaccagaacatggggtagggtagggtagggtagggtagggtagggtagggtagggtagggtag

SUBSTITUTE SHEET (RULE 26)

12/33

Fig. 6F

gtgtggggaggggctgtgagcagcagtgtgggttggatgcctacttctggacaaaatt

rs615536

gggtcggtgacttagggataaagattggtgtgggcccccccttcctcgtggatctcac  
atgtttctgatcccttctgtctgcctccccaccagGCCAGACCTGCTGGATTTGAG  
--P--D--L--L--D--F--E-

TCTCTGAAGAAGTGTAAATGCACACTATAATCTGCAGAACATTCAATCTGGCTGAAAAG  
-S--I--K--K--C--N--A--H--Y--N--L--Q--N--A--F--N--L--A--E--K-

GAACTGGGACTTACCAAGCTGCTGGATCCCGAAGgtggggccagagctatgtaaaaaga  
-E--I--G--I--T--K--I--L--D--P--E--

ggtaggttagagacaggagggtggaccagtggccaggggctgcggaggggaa  
catggagtaaccaccaactgatacagggttttggtgaaaaagagaatgttctgga  
attagaaaagtggtgatggttacacaacccctgttaataactgaaaaacacgtggttgctc  
acacccagcacttgggaggccaaggcaggaggatcacttgagcctaggagtttgtacc  
agcctgggcaacatagtgagaccccatgactaccctccccccaaaaaatatata  
tttacactgaaaaccactgaattatacattgaaagggtgaatttatggtgtgacaat

rs577595

catttctcaggtttttttttttaacaaaagactgttaggttaggaagaaggcagccca  
atgacagaaaqaqccctqqactgtgggttgtgtcccaggcagagcacagaggagctctgt

rs580024

gacatacagccagtggctagcctctgtggctagttcctcatgtctaaaagagaga  
atggctctctaaggcttctgaaattcttaggattctaaattaagctacatgttgctgg  
ccacggtgttacctgcctatagtccatgtctgaggctgaggcgagaggatcactt  
gagcccaggagttctaggctgcagtgggttatgctgattggatatctacactaaattcag  
cactggtgacctcctggagcaggggaccagcaggttacctaaggaggactgaaccagtc  
caggtcagaatggagcaggtaaaaactcccacactgatcagtagtggatcacaccgtc  
aatagccacagtactccagctggcaacagcgagacccatcttatTTTTaaaaca  
atTTTTaaaaaaaaaaattaaagctgcattacccactctcacccagaactggcctgt  
gtcaqctqaccttqctgattggttcctgtggttct`gtcataatgccatgagatcgatg

rs554305

153-161  
agatcagtaacagccactggcttttttttttttttggagacggagtctcactgttg  
cccagactggagtgcaatggcgcgatctcggctactgcaacccccgcctctggttca

**SUBSTITUTE SHEET (RULE 26)**

13/33

*Fig. 6G*

cggtctaattttgtattttagtagagacagggttccatgtggcaggctggtctc  
 taactcctgaccctcggtggcccccacctcgccctccaaatgcgcgtgtgataacagggtgt  
 gagcccccgccctgactgcgattaggctttatgtataacctggttctgcacttgatgt  
 gtttaatcttacaacaccccttggaaatcagtactgttgttatcctcagttatggaa  
 gagaagatggggcctataatggtaagcaacttgcccaggctccggcagatggagct  
 gagattggcggtcaggcaggctggctccagagcgtttccctccgctgcctctggaaag  
 tggagaccaatcttggtgcaaccagaacaggaggagagaaaggctgcctgtggaggt  
 gtccttaagttcactgggtggatgttactggcttggccattttccctgtttcca  
 ttctacctgttagccctatcacctggccctttactgcagaatgacaagtgtcagcgc  
 tagcaacagacataggcttcatgcagtgccttggagagcctctggtagcatggctgt  
 aattcagaacctggaaataaagctaagcattctgtacttccgagatgacttcagtg  
 tcctgcctgcccattgttccagaaataggcaagaggctctgggttcatagttacc  
 tttaggataaaatctggctattaaaagaaccacagattcaagttatggtcagcctc  
 agttgtgcgttggaaagccaattcggttggaaagacactccctccttggccacagctg  
 agctttaggatttaggttagaaacggcgtgatcacacctggagatgggtctgagagctgg  
 cctctgtccccactgagctgcaggatcacaggccatcgatctgagcatccttaggttc  
 aggaggcttagttcataccctccgtcccactgcctcaatccattgacttcttaacat  
 tattctagcctaaaagtccatgatccatccccacatagcaatgtgtttctttcttgct  
 tttctcccaaacaactgatttggctccacggggactctggtaaaaagcttccttg  
 tgcttctgggtttgtgctctgttggagggcccttgggtttccaggagaaggggaaaa  
 aaagatgagcttttggtgccctttcattgaatggtaagttgtcatttgctctaaa  
 aggagatggagagatgtcagccttgaggactggcaggcccagggttgggttaagtga  
 tgcaagttatggatgaactctggaggcctgacccaaatggcctcttgcaagACGTG  
 D--V-

AATGTGGACCAGCCAGATGAGAAGTCAATCATTACCTATGTGGCTACTTACTACCATTAC  
 -N--V--D--Q--P--D--E--K--S--I--I--T--Y--V--A--T--Y--H--Y--  
 Y-

TTCTCCAAGATGAAGGCCCTGGCGTGGAAAGGCAAGAGAATTGGCAAAGgtactgtccatg  
 -F--S--K--M--K--A--L--A--V--E--G--K--R--I--G--K-

ggcagttaggcataaaggccagaggaggcccggctgaggggtttactgccttagtgcaag  
 ggcagggtggagctgcaggactggccaggccctgtggactgtcacgtccctg

**SUBSTITUTE SHEET (RULE 26)**

14/33

*Fig. 6H*

tcttctgcctccagGTGCTGGACCATGCCATGGAGGCAGAGCGCCTGGTGGAGAAATAC  
 -V--L--D--H--A--M--E--A--E--R--L--V--E--K--Y--  
 GAGTC CCTGGCCTCGGAGCTGCTGCAGTGGATCGAGCAAACGATCGTGACCCCTCAATGAC  
 -E--S--L--A--S--E--L--L--Q--W--I--E--Q--T--I--V--T--L--N--D--  
 CGGCAGTTGGCCA ACTCCCTTAGCGGGGTCCAGAACCGAGCTGCAGTCCTCAACTCCTAC  
 -R--Q--L--A--N--S--L--S--G--V--Q--N--Q--L--Q--S--F--N--S--Y--  
 CGCACCGTGGAGAAGCCGCCAAgtaggtgtccctggggccccaccctccctgagctgt  
 -R--T--V--E--K--P--P--K  
 gctcccacgagaggaagcctaattagcacagccttcaggagggaaattggcagtaca  
 taaatgcagtggagatttccacaccagaaggcatccaaacaacatagttgataaaaaata  
 aaatttttaatgatttagtcgttttaaaaatcatgctgtggccggcatggctca  
 cgcctgtaatcccagcacttgggaggccaaggcggcgcatcaccgagaccagtctggc  
 caacatggtaaaacccatctctactgaggtcggagttttagaccagcctggccaacat  
 ggtgaaacccgtctccactaaaattacaaaaattagctggcatggtcacacgcc  
 tgtcatcccagctactcgggaggctgaggcaggagaaccacctgaacccggagacagag  
 gttgcagtgagccaagatcacgccactgcactccaacctgggtgacagagcaagactccg  
 tctcaaaaaacaaacaaacatgctgtggaaatgattgctatgatgtgttgagtca  
 atccatcagataagattactgatcaggtgtcatggcaacccaatcctagcataggataga  
 ggggactggacctgacaggcaggtcacaaccggcaacatgggtgtggctggtaacgtgag  
 aattgcagaggttcttatcatttcagtattttagcaattaactgttccaagtatgtgatt  
 gctgttaggcaattctgttcaagtacactgccaacacgtctgtgttatgctgtcatggagt  
 tctttggagttatatatcagctctgaaactaatattagtcagaaaacaattgttt  
 ctatacattccaggaaatgatttaaaggatgttacccttaaccttaggaagttgtcttctgga  
 cattaaacaaagatgttagctcaaggatgtttatcaacacacaaccactagaaaagaaaaat  
 aagtcatggcagccactgaaaataatattaaagaaatgcatttattgacatggaaaggt  
 gttcatgaaaaataagtggggaaataggttataaaagaattatttggctggcaccgt  
 ggctcatgcctgtaatcccagcacttgggagactgaggcgggtggatcacttgaggta  
 ggagtttagaccagcctggccagcatggtaaaacccatctctactaaaaatacataaa  
 tttagccaggtgtgggtgcattgcctgtaatcccagctactcgggaggctgaggcaagag

rs645307

attggcttgaacccgggaggcagagggttgcagcgaacccagatcgccaccactgcactcca  
 gcctgggtgacaaagcaagactctgttcaaaaataaaaatgatgttctttggta

**SUBSTITUTE SHEET (RULE 26)**

15/33  
*Fig. 6I*

gcccataaaaaatgtttctggcattccacataaaaaagtttatattgctttgttaagaagaaaagcaa  
 cacaggacaccatggaaagagtaaaagagtcottgtgaccctcggtggcttaattctg  
 accccaacccggcctcctgcctcagGTTTACCGAGAAAGGGAACTTGGAAAGTGCTGCTC  
 --F--T--E--K--G--N--L--E--V--L--L--  
 TTCACCATCCAGAGCAAGCTTCGGGCCAACAACAGAAGGTCTACACGCCCGCGAGGGC  
 --F--T--I--Q--S--K--L--R--A--N--N--Q--K--V--Y--T--P--R--E--G--  
 CGGCTCATCTGGACATCAACAAGgtccgtggctgccacaggccacccaccctcaggc  
 -R--L--I--S--D--I--N--K--  
 aggccctggcccagatgccttgcacacatccccaaaccggggccatgtcgaccccttacc  
 aagttctactatctgctgccccaaacttgaaactcgagcactctgcccagctgcccacact  
 gtgccagatgtgattctccatcctctcaggcacacggctccctgtcccttgcact  
 ctccatacgaggcacataaggaaattatgccccaaagtccctcagttgttaaacctgtc  
 cccaaactcaactatcccttcttttatttttccacaatattgttaagattactaaaagtaaa  
 aataagggttcgcacacagtcctgccaacacaaaggaaatcagatggttccactgttgc  
 ctccctccagtccttcagaagcttctacatggcttccacagactccagcgtgcct  
 gtttccatggacttaccatgtgcctccagccactgagccctccccaccctgtccctt  
 ccactaaccctgtccccaccccatagGCTTGGAGCGGCTGGAGAAGGCAGCACGAG  
 -A--W--E--R--L--E--K--A--E--H--E--  
 CGTGAGCTGGCCCTGCGCACCGAGCTCATCCGCCAGGAGAAGCTGGAGCAGCTGGCCGCC  
 -R--E--L--A--L--R--T--E--L--I--R--Q--E--K--L--E--Q--L--A--A--  
 CGCTTCGACCGCAAGGCTGCCATGCCGGAGACCTGGCTCAGCGAGAACAGCGCCTCGTG  
 -R--F--D--R--K--A--A--M--R--E--T--W--L--S--E--N--Q--R--L--V--  
 TCCCAGgttaggactttaggctctaggatgttaggctgtcagaactggagagagaca  
 -S--Q--  
 gggtagataagaagcccccgccccgggtggagagacaatgaaacacaaaatccgtcacctgg  
 taaaaaggcctagaggtccggagccaggggccagagggtggacaggagagagggtgg  
 tgagacaggatgggggtggatagagagggaaacttagagccaccgcattggagctggatc  
 ttgcaggccaaagtgcctgcaggacagaggcaggcactgtcccttggcccc  
 acacccctctccctgccccgacacagGACAACCTTGGCTGGAGCTGGCAGCTGTC  
 -D--N--F--G--L--E--L--A--A--V--  
 GAGGCAGCAGTACGGAAGCACGAAGCCATTGAGAGGGACATCGTGGCCTACAGCGGCCGG  
 -E--A--A--V--R--K--H--E--A--I--E--T--D--I--V--A--Y--S--G--R--  
 GTGCAGGCAGTGGACGCCGTGGCTGCAGAGCTGGCCGCCAGCGCTACCACGACATCAAG

16/33

Fig. 6J

-V--Q--A--V--D--A--V--A--A--E--L--A--A--E--R--Y--H--D--I--K  
CGCATGCCGCTGGCAGCACACGTGGCACGGCTCTGGGACTTCTGCGGCAGATGGTG  
-R--I--A--A--R--Q--H--N--V--A--R--L--W--D--F--L--R--Q--M--V  
GCCGCCGGCGGGAGCGGCTCCTCCTCAACCTGGAGCTGCAGAAGGTGTTCCAGGACCTG  
-A--A--R--R--E--R--L--L--N--L--E--L--Q--K--V--F--Q--D--L  
CTCTACCTCATGGACTGGATGGAAGAGATGAAGGtaccagtgaggcgtgctgggtggggt  
-L--Y--L--M--D--W--M--E--E--M--K--  
aagagtgtatcaagagtcgaggggccccacagtgggtgcgtccgcccgtctcgcccg  
atctctgtggagtgtgaaccagcacagggccctgtcccccaggtaaaaaactggcc  
cagggccctgtccacatggggttcgtgtccacatggaggaggctgtatgaaaaaccac  
gaccttccgtaaacaaaagagcaggggccataaaaccactgtatgaaaactggaacagg  
ggtgtgagggctggtcagacaaggcctctgagaaaataccatttaagccgagaccagaa  
tagtgcaggcagccagcacagagagctctcaggagagaatgttccggcagaaggaaac  
ccggcacaggccctcgccctggatcagcatggtggttcaggatcagaaggaaggccgg  
gggctgtgagcatgaggagtggggccatggactgagatttggatattcaaagtaacatggatc  
actaggccctacaggcctggtcagggatggatattcaaagtaacatggatc  
ttaaagtgatttgaaggggccaggcgcagtggtcaggatggatggatattcaaagtaacatggatc  
aggccaaggcaggcagatcacttgaggctcaggatggatggatggatattcaaagtaacatggatc  
gaaaccctgtctataaaaatacagaaattaaccagggtgcggtggttagcgcattgcctgt  
agtcccagctactcaggaggctgagacagcagaatcgcttgaacccaggaggcggagg  
gcagtgagccaagattgcaccactgcactccagcctaggatagagcaagactcagtca  
  
rs11227576  
aaaaaaaaaaaaaaagcagctaggcaaggtggctacgcctataatctaagactttgg  
gaggccgaggcggcggatcacctgaggtcaggagtccaaagaccaggcctggcaacatgg  
tgaaacccatctgtctataaaaatacaaaaattagctgagtggtggcacacgcctgt  
atcccagctactcaggaggctgaggaaggagaattgcttggacctggagggtggagg  
cagtgaactgagattgtgactgcactccagcctgcactacaggagcggactccatct  
caaaataaataaataaaataaaatgtattttagaagggtttgagagggggagagattt  
accttatgtttgtaaaaaaagtcaactttggctactctgaagaaggcctgtggcaggcag  
atgacagcggagaccgttagaggttgccttccaggcggcagcggctaggagg  
ggattggggagagaatggatggatggatggatggatggatggatggatggatggatgg  
SHEET 26

# **SUBSTITUTE SHEET (RULE 26)**

17/33

*Fig. 6K*

gaggggtgtcgatgattctgaggctgagaagattagtcacagcaaagcttccaagagcagg  
 agtttagtaaacattgttccacagcagggagttcaggataaggaggtcggtgc  
 ggcagaactgggtatctcaagtatgttgagcaagcaggttcacctgggtcaaca  
 cctggggccgtggagaagttgaagaaaaacgcagccaggtagcgcgtctgtgc  
 caggcggcatgtgcacttcatgtgtgcctggcag  
 GGCGGGCTGCAGTCTCAGGACCTGGCAGGCACCTAGCAGGAGTGGAGGACCTGCTGCAG  
 -G--R--L--Q--S--Q--D--L--G--R--H--L--A--G--V--E--D--L--L--Q--  
 CTGCACGAGCTGGTGGAGGCAGACATGCCGTGCAGGCCGAGAGGGTGC  
 GGGCCGTCAAGC  
 -L--H--E--L--V--E--A--D--I--A--V--Q--A--E--R--V--R--A--V--S--  
 GCCTCTGCCCTGCGCTTCTGCAACCCAGGGAAAGgtgagaagtcagcgaaggcactggag  
 -A--S--A--L--R--F--C--N--P--G--K--  
 agggagggctggaaaggagcacatcaagagccgaggtgaaagggttggaaacctgggg  
 acggaaaagatggtgccaaacgatggtgactgagctaaagccagggaggaaagtccaag  
 agagtgtggcaggagcagccgaggctgggtaaggctctgaccctctcctgtga  
 ctttctcagAGTATAGACCTTGCACCCGCAGCTGGTGTGGAGCGGGTGGCCAAGCTA  
 E--Y--R--P--C--D--P--Q--L--V--S--E--R--V--A--K--L--  
 GAGCAGAGCTATGAGGCACTGTGCGAGTTGGCAGCGGCCGGCGGGCCGGCTGGAGGAA  
 -E--Q--S--Y--E--A--L--C--E--L--A--A--A--R--R--A--R--L--E--E--  
 TCACGGCGGCTCTGGCTTCCCTGGAGGTGGTGAAGCTGAGGCCTGGTGC  
 GGGAG  
 -S--R--R--L--W--R--F--L--W--E--V--G--E--A--E--A--W--V--R--E--  
 CAGCAGCACCTCCTGGCTCAGCCGACACGGGCCGAGACCTGACCGGTGCCCTCCGCCTG  
 -Q--Q--H--L--A--S--A--D--T--G--R--D--L--T--G--A--L--R--L--  
 CTCAACAAAGCACACAGCCCTGCAGGGCGAGATGAGCGGCCGGCTGGGGCCCTGAAGCTC  
 -L--N--K--H--T--A--L--R--G--E--M--S--G--R--L--G--P--L--K--L--  
 ACCCTGGAGCAGGGCCAGCAGTTGGTGGCCAGGGTCACCCCTGGGCAAGCCAGGCCTCT  
 -T--L--E--Q--Q--L--V--A--E--G--H--P--G--A--S--Q--A--S--  
 GCCCGTGCAGCTGAACCTCAAGCCCAGTGGAGCGGCTAGAGGCCCTGGCCGAGGAGCGT  
 -A--R--A--A--E--L--Q--A--Q--W--E--R--L--E--A--L--A--E--E--R--  
 GCCCAGCGGCTGGCCAAGCCGCCAGCCTCTACCAGTTCCAGGCCGATGCAAACGACATG  
 -A--Q--R--L--A--Q--A--A--S--L--Y--Q--F--Q--A--D--A--N--D--M--  
 GAGGCCTGGTTGGTTGACGCAGTGCCTGGTGTCCAGCCCCGAGCTGGGCACGACGAG  
 -E--A--W--L--V--D--A--L--R--L--V--S--S--P--E--L--G--H--D--E--  
 TTCTCCACGCAGGCTTGCGCTTACGCTGCGCTGGTGTCCAGCCCCGAGCTGGGCACGACGAG

18/33

Fig. 6L

-F--S--T--O--A--L--A--R--O--H--R--A--L--E--E--E--I--R--S--H-

CGGCCAACCTGGACGCCTGAGGAACAGGCAGCAGCCCTGCCCCCCACACTGAGCCGC  
-R--P--T--L--D--A--L--R--E--Q--A--A--A--L--P--P--T--L--S--R-  
rs4930388

R

ACGCCCCGAGGTGCAG~~G~~GCCGGGTGCCACCCCTGGAGCGGCAC~~T~~ACGAGGGAGCTGCAGGCC  
-T--P--E--V--Q--S--R--V--P--T--L--E--R--H--Y--E--E--L--Q--A-

CGGGCAGCGAGCGAGCGCGGGCCTTGGAGGCAGCCCTGGCGCTACACCATGCTCAGC  
-R--A--G--E--R--A--R--A--L--E--A--A--L--A--L--Y--T--M--L--S--

GAGGGCCGGGGCTGTGGA  
-E--A--G--A--C--G--L--W--V--E--E--K--E--Q--W--L--N--G--L--A-

CTGCCCTGAACGCCCTGGAGGACCTGGAGGTCTGCGAGCAGAGgtaggccccctcaggctct  
-L--P--E--R--L--E--D--L--E--V--V--Q--Q--R

rs3741359

```
agtgggaccagcctgggaggtgggggtggggggccaggatgtggtggtgagtcctc  
cataaacttcctgcctcacccctttagtcttaatggtgtccatttctagtttaacaa  
aaaatgttaaccatactcacaagtagagacattcaccacaaacctctacaccgtca  
cccagattcagtcattgccaacatcttgcctattggttctgcctttaaaaagca  
aatccccaaagattagcacatccctcctacccctttagggtgtgcccttaagaaataca  
aggaagcggccgggcacggtggtcacacctgtaatcccagcactttgggaggccgaggc  
aggcggatcacgaggtcaggaggtcgagaccatccctggctaacatggtaaacgcgtct  
caactaaaaataaaaaattagccgggcacagtgggggcgcctgttagtcccagctact  
cgggaggctgaggcaggagactggcatgaacctgggaggcggagcttgcagtgagccgag
```

rs11286358

rs12807346 rs12807677



20/33

*Fig. 6N*

gcatgacagttattctgtttggggagttgtgggttccttgcatacgacaga  
 tactatgaacctcataggctgagaggcagagactcctgcagataagcctcgtgaa  
 ggagacaccctcctgggtgttaagagaggtggacttagtgactggcacatcgctgc  
 ctcccacaattcacgcctggcttcacccttcagGTTCGAGACCCTGGAGCCTGAA  
 --F--E--T--L--E--P--E-

ATGAACACCCTGCAGCACAAATCACCGCGGTGAATGACATTGCCGAGCAGTTACTGAAG  
 -M--N--T--L--A--A--Q--I--T--A--V--N--D--I--A--E--Q--L--L--K-

GCCAACCCCCCAGGCAAAGACCGCATGTCAACACCCAGGAGCAGCTAACCAACAGgtgg  
 -A--N--P--P--G--K--D--R--I--V--N--T--Q--E--Q--L--N--H--R

rs2276140      rs532439  
gtttgggagggcaggaccaggaaactgacagaaaaatgaagcaatggggatggcagttag  
 aggcaagggttttagggcttggaaagggtggctacaaaggaggaagcaaaccagtctggaa

rs2276139  
tatgttgggaaggaaaaggatgaaagagatgagagagaggggtcaaggtggctgaagc

rs11227575  
ggctgtggccagaaagggaggagtgaagggagctaccaagagagagaagcaggaa  
 gaagcttccaaacaggcctggccagggcaggaagctgaaccccccgtctcagGTGG  
 --W-

CAGCAGTTCGGCGTCTGGCAGACGGCAAGAAGGCAGCTCTCACCTCAGCCCTGAGCATH  
 -Q--Q--F--R--R--L--A--D--G--K--K--A--A--L--T--S--A--L--S--I-

CAGAACTACCACTTAGAGTGCACGGAGACCCAGGCCTGGATGAGAGAGAAGACCAAAGTC  
 -Q--N--Y--H--L--E--C--T--E--T--Q--A--W--M--R--E--K--T--K--V--

ATCGAGTCCACCCAGGGCTAGGCAACGATCTGGCTGGGTGCTGGCCCTGCAGCGCAAG  
 -I--E--S--T--Q--G--L--G--N--D--L--A--G--V--L--A--L--Q--R--K--

CTGGCCGGCACGGAGCGGGACCTGGAGGCCATCGCCGCCGGGTGGCGAAGTGACTCGA  
 -L--A--G--T--E--R--D--L--E--A--I--A--A--R--V--G--E--L--T--R--  
 rs506028

Y  
 GAGGCCAAATGCCCTGGCTGCCGCCATCCGCTCAGGCAGGGCCATCAACGCCGGCTG  
 -E--A--N--A--L--A--A--G--H--P--A--Q--A--V--A--I--N--A--R--L--

AGAGAGGTGCAGACCGGCTGGAGGACCTCAGGGCCACCATGCGCGTCGAGAAGAGTCG  
 -R--E--V--Q--T--G--W--E--D--L--R--A--T--M--R--R--E--S--

CTGGGGAGGCAGGGCTGCAGGACTTCTTGCAGCTGGATGACTTCCAGGCCTGG  
 -L--G--E--A--R--R--L--Q--D--F--L--R--S--L--D--F--Q--A--W--

*Fig. 6O*

CTAGGCCGCACTCAGACTGCTGTGGCCTCTGAAGAAGGGCCGGCCACCCTGCCTGAGGCA  
-L--G--R--T--Q--T--A--V--A--S--E--E--G--P--A--T--L--P--E--A-

GAGGCCCTCCTGGCCAACATGCAGCCCTGCAGGGAGAGGTGGAGCGGGCCCAGAGCGAG  
-E--A--L--L--A--Q--H--A--A--L--R--G--E--V--E--R--A--Q--S--E-

TATAGCCGGCTGCGAGCCCTGGCGAGGAGGTGACCCGGACCAGGCTGACCCCCAGTGC  
-Y--S--R--L--R--A--L--G--E--E--V--T--R--D--Q--A--D--P--Q--C-

CTCTTCCTACGACAGCGACTGGAGGCCCTGGGAACGGCTGGGAGGAGCTGGCCGAATG  
-L--F--L--R--Q--R--L--E--A--L--G--T--G--W--E--E--L--G--R--M-

TGGGAGAGCCGGCAAGGTGCGCTGGCCAGGCCACGGCTTCCAGGGATTCTGCAGGAT  
-W--E--S--R--Q--G--R--L--A--Q--A--H--G--F--Q--G--F--L--R--D-

GCTCGTCAGGCTGAGGGCGTGCTCAGCAGCCAGtgcggccaaagtcccaagc  
-A--R--Q--A--E--G--V--L--S--S--Q-

aggaggaagagcaaagttagggacccgggaaatgtgaaggagcaggatggcaggaagga

rs11227574

catgctagcaaaatggggcagcgccagtggtcacacctgaaattccagcactttggaggg  
ccaaagttaggaggatcacttgaggctgagaatatccagaccacactggcaacatggcaa  
gaccttgtctcacaaaaaaaatttttaagaaaatagaagaattttaaaaagaaaaaa  
tgggagccagacaggatggctcacacttgtactcccagtatttggaggccgaggcagg  
agaataacttgagctcaggagttgagaccagccctggcaacatagtgagacccccatct  
ctataaaaaaaaattactggacatgggtgtcatgcctgtagctccagctactgg  
gaggctgaggctggatcaactggagccaggagttgaggctgcagtgagctatgatat  
gccactgcactccaacctggccacagaatgaaaccctctcaaaaaaaaaaaaaagaa  
aaaagaagaaaaatgagaatgaaaaagacgtgaatataattactaaaactgactttag  
aagaaatagaaagcccaggttgtcatgtactgttaatggaatcaggagctagaagtga  
gatagaacaggatttggctgggaatggaaggccctccaccagctccctgtgact  
ttctgaggctccatgtggctggcagccctgtctcagactctctggctaccct  
ccctgtctgttggccctaccctctcagatttgcctgggtgggtccctcttaggg  
gggtgaattgtctggggaaaatgagctgaatgtcatccctcccacacagGAATATGTT  
-E--Y--V-

22/33

*Fig. 6P*

-L--S--H--T--E--M--P--G--T--L--Q--A--A--D--A--A--I--K--K--L-

GAGGACTTCATGAGCACCATGGACGCCAATGGGAAACGGATCCACGGGCTCCTGGAGGCT  
-E--D--F--M--S--T--M--D--A--N--G--E--R--I--H--G--L--L--E--A-

GGCCGCCAGCTGGTATCTGAAGGCAACATCCACGCCGACAAGATTGGGGAAAGGCAGAC  
-G--R--O--L--V--S--E--G--N--I--H--A--D--K--I--R--E--K--A--D-

TCCATTGAGAGGGAGgtctgatgaggacagtccatgaatttagggttcccagggggaaatcg  
-S--I--E--R--R

gagaaacagggtgacctcaaagataaacgtggcacaggaaacccacagatgggcaggag  
ctgacagagaagttagagggaaagaactaagtggtagggctggagattccacccc  
caaccagggtctaaaaggaagtcaggattcctggtagcctcatgtgctcccgaaaggcg  
ttattcccctagaagaaatggaggcccctaggttagccaaagggtcacaatccttcaca  
gcaaattccagagttcacaagagggtgtcggtccagGCACAAGAAGAATCAAGACGCA  
--H---K---K---N---Q---D---A-

GCGCAGCAATTCTGGGCCGCTTCGGGACAACCGGGAGCAGCAGCATTTCTGCAAGAT  
-A--Q--Q--F--L--G--R--L--R--D--N--R--E--Q--Q--H--F--L--Q--D-

TGTCACGAGgtgaggctccctggggccccggatattccctagccatcccttctcacct  
-C---H---E-

tgaggcctagaataagtccagcacaaggtaaccggagactgtgagccccttcatggcttctt  
ccccaggcgcccacttctcctggctcactgtggccctgctttatgccccctctccct  
cccttgaaaaatgtccctgtttatgttttgtccagCTGAAGCTCTGGATCGACGAGAAG  
-L--K--L--W--I--D--E--K-

ATGCTGACAGCCCAGGACGTGTCCTATGACGAGGCCGCACCTGCATACTAAGTGGCAG  
-M--L--T--A--Q--D--V--S--Y--D--E--A--R--N--L--H--T--K--W--Q-

AAGCACCCAGGCATTCTGGCCGAGCTGGCTGCCAACAAAGACTGGCTGGACAAGGTGGAC  
-K--H--O--A--F--M--A--E--L--A--A--N--K--D--W--L--D--K--V--D-

AAGgtgagcagtgtggggctgccttggcagagtccccatggtacggggagg  
-K-

23/33

*Fig. 6Q*

rs11823888  
aaaccttgtctggaaagggctggatggtcaatatctgaggcttgagtc  
cagttctactacaactactcatctctgtcgtagtgc  
rs11227573  
agtgaatgaatgaacatggctttccaaataaaactttatgtaaaaacaggc  
gccagattaccctgtggctgtatgttagtctgccaatc  
cctggttttgtctccacccagctgtcccaagagtcaagg  
atctgctcagcaagtatccctacagaacatggta  
cagggctacgaaaagacactttccattc  
accagtgtgcttcaagtgttccattc  
ccaaaggcattgtttaggc  
rs12419099  
ctgaaggaacaactgttagccaggatatgaggcatatgc  
cagacagaatctgatgagtgctgtgagagtggcccc  
agaggagatatggaaattctagaagg  
gaatttggacaagtggagatatggagg  
aaaaccagtggtaggc  
acagaatgtgc  
accaccacatgc  
cagtgcttagc  
aggcagtgt  
gtgtccccgg  
rs11825713  
gttaccccatgggcatcagggccaacccattctgc  
gaacacagaacacagcacttc  
gctccctgtctgt  
aatttgc  
cctcaag  
gtgaag  
tgtgg  
tcg  
SUBSTITUTE SHEET (RULE 26)

24/33

*Fig. 6R*

gaacagcagttagaaggatttaaacccgtgaagagtgtcgataagaacacaca  
 aggacttggattattgccacacttcaggtctctactgcccgttaaggcccctggctgc  
 tctcagcctctaaagttcaggctaatttggcacttgacatttctgtcctgccttcca

gGAAGGGCGAGAGCTCACCTTGAGAAGGCCAGAGCTGAAAGCCCTGGTGTGGAGAAGCTG  
 -E--G--R--E--L--T--L--E--K--P--E--L--K--A--L--V--S--E--K--L-

AGAGACCTGCACAGGCGCTGGGACGAGCTGGAGACCACCCAAGGCCAAGGCCCGCAGC  
 -R--D--L--H--R--W--D--E--L--E--T--T--Q--A--K--A--R--S-

CTCTTGATGCCAACCGAGCTGAGCTGTTGCCAGAGCTGCTGTGCCCTGGAGAGCTGG  
 -L--F--D--A--N--R--A--E--L--F--A--Q--S--C--C--A--L--E--S--W-

CTGGAGAGCCTGCAGGCCAGCTGCACACTGGATGACTACGGCAAGGACCTCACCAGCGTC  
 -L--E--S--L--Q--A--Q--L--H--S--D--D--Y--G--K--D--L--T--S--V-

AACATCCTGCTCAAGAACGCAGCAGgtgtgctgtggccctttagatggggatggtaacagc  
 -N--I--L--L--K--K--Q--Q-

agaagaaaagggtctgcagcttcaagatttggaggccagctgaggcctggcagataaca  
 ctttcactagcattccagagtcatttttggcagccagtatcagaatctctaggaa  
 tatgtcgcagtcagggttagagtttaggagacagacccctaataatttgcgtatgagaac  
 ttcaagtaaaaagaaatattaatttaggccaggcacagtggctcatgcctgtaatctcagca  
 ctggaggccaaggcaggcaggtcgctttaggtcaggatcaagaccgcctggcca  
 acatggtaaaacccgtctctactaaaaataaaaaattagccaggcgtggcgcacg  
 cctgttatcccagctacttggaggctgagggtgaggaggatagcttgcataacccag  
 gagacagaggttcagtgagccgtatgcacatccactccgcctggcaagccct  
 gtctcaaaaaaaaaacagaagtataaatttagtaaaagggtgtaaagagaactctatagc  
 aggggtcagaaaacttttcaagcaccagatagtaaattttggggcttcatggg  
 tcgtaaggctatgtcacagccctcaactctgcccttgcataaggaaaagcagctgcggat  
 tatatgtaaaataagggtatgccagggtgtggctcacacatgtaatccaaacactt  
 tggaggccagggtggccggatcaattgagtcaggatggacttagcctggcaac  
 atggcaaaacccatctctacaaaaactacaaaaattagctggcgtgggtgggtgt  
 gcctgttagtcccagctactcaggaggctgagggtgagaagattttgagcctggagagtc  
 gaggctgcagtggctgtgatcgtgcactgcactgcagcctgggtggcagagcagact

**SUBSTITUTE SHEET (RULE 26)**

*Fig. 6S*

tgtctaaaaatttttaaaaagggggtgcagctgtggtccaataaaaactttataaggc  
taggcagtggctggacttgctggccctgctctgaagaatgcagaatggaaagtgttagg  
gagcagctgacacacctgggttggaggatccccaggaaggagcaagatggctgagtct  
cagaccccttggagagggcgaggcccactgaatggcctggaaatttgcgtgggtatca  
caggccagagcatgtctgaagccagcaagccaggctgccagcaggaagtggccactag  
agtgcagcaaacttgcagagagtaagagccaccggatcttgcgcactgtggacagt  
tttgcgttaggaggaagaaaagcaccttggaaaccgtggaaatgccccacatctgctaggca  
tttgtgtgtccctccagccctctgctgacaaggcttagattgtgcctgtggcaaagg  
aggaatgttccagggccagttccagtgtctcaaagcagggccacgatggattgggat  
ggagagacaatgaattgataatggcacagtttatcacactaccctggatttttaatttc  
aatcttttagattcgccgggtgtgggtcaggttttataagttatattgtgtatgctg  
aagttgggtgttaattgaacccatcacccatgttagtgagcacagtaaccaataggatt  
tttcaacccctgcccattggattctgtcatataaagtgcagatccactgtccagcc  
catgaggggatgtctccctgggtgagctgcagtgtgagcagccacagtggctggata  
gagactgggtgtctgtggcatgaaggcgaaatgcagagctaacatgtccattc  
ctcctggcggttgcagATGCTGGAATGGGAGATGGCTGTGAGAGAGAAGGAGGTGGAG  
-M--L--E--W--E--M--A--V--R--E--K--E--V--E--  
rs471334

Y

GCAATCCAGGCCAGG[CAAAGCACTGGCCCAGGAGGACCAGGGTGCAGGGGAGGTGGAG  
-A--I--Q--A--Q--A--K--A--L--A--Q--E--D--Q--G--A--G--E--V--E-

AGAACCTCGAGGGCCGTGGAGGAGAAGTTCAGGGCCTGTGCCAGCCATGCAGGGAACGC  
-R--T--S--R--A--V--E--K--F--R--A--L--C--Q--P--M--R--E--R-

TGCCGGCGCCTGCAGGCTTCTCGCGAGCAGCACCAAGTTCCACCGCGATGTGAAAGATGAG  
-C--R--R--L--Q--A--S--R--E--Q--H--Q--F--H--R--D--V--E--D--E-

ATTgtgagtcaactggggccaaggacggcaagctgccccagccatgtggttctccagcct  
-I-

ccctcctggatgccaggagatgccagcaggcttattccctttctttggcattga  
ccatctccctataggagacttggagatgcctcccagaaccagagatgactgtcccca  
cacacaggcggtagccccaggtgtccccactccactaatcagtcctgtgtccct

tgccctctggcctccactgaccccttccatTTGTGGGTGACAGAGCGGCTG  
**SUBSTITUTE SHEET (RULE 26)**

*Fig. 6T*

CCCATGGCCAGCTCCATGGAGCATGGCAAGGACCTGCCAGCGTCCAGCTTCTCATGAAG  
 -P--M--A--S--S--M--E--H--G--K--D--L--P--S--V--Q--L--L--M--K--  
 AAAAACAGgtgaggcagaggctgaaggcaaaagagaagtcccaggagcctgccagac  
 -K--N--Q-  
 ttggacgtgttttcttaagaccaggccccctcgatgctgagtgtaaccctggacttc  
 agtggctgtcttcctcaccttggaggggtgggtcctctcagtaggagatagtggggct  
 gggccggctgacgggtgttaccatcgccccaggctaggagagggagcagaggacccag  
 gaaggagagaggcacaggggtgaagggtggtctccgggagcacgtgggctggggcagga  
 ctttcagcatttctttctgtggccatggcag  
 ACCCTGCAGAAAGAGATTCAAGGCCATGAGCCCCGGATCGCGGACCTGAGGGAGCGGCAG  
 -T--L--Q--K--E--I--Q--G--H--E--P--R--I--A--D--L--R--E--R--Q--  
 CGTGCTCTAGGTGCAGCAGCAGCAGGTCCAGAGCTGGCTGAGCTGCAGGAAATGTGGAAA  
 -R--A--L--G--A--A--A--G--P--E--L--A--E--L--Q--E--M--W--K--  
 CGCCTGGGCCACGAGCTGAACTTCGAGGGAAAGCGACTGGAGGATGCCCTGCGAGCCCAG  
 -R--L--G--H--E--L--E--R--G--K--R--L--E--D--A--L--R--A--Q--  
 CAGTTCTACCGCGATGCCCGAGGCCGAGGCCCTGGATGGCGAGCAGGAATTACACATG  
 -Q--F--Y--R--D--A--A--E--A--E--A--W--M--G--E--Q--E--L--H--M--  
 ATGGGCCAGGAGAAGGCCAAGGtgcaggccaggacagagccactgtatgtgaccagttc  
 -M--G--Q--E--K--A--K--  
 tgccctccctgacctgatgtggatgccactgtccctcccccagGATGAGCTGAGTGCC  
 -D--E--L--S--A--  
 rs639938  
 R  
 CAGGCAGAGGTGAAGAAGCACCAGGTGCTGAGCAAGCCCTGGCGACTACGCGCAGACC  
 -Q--A--E--V--K--K--H--Q--V--L--E--Q--A--L--A--D--Y--A--Q--T--  
 ATCCACCAGCTGGCGGCCAGCAGCCAGGACATGATTGACCACGAGCACCCAGAGAGGtgg  
 -I--H--Q--L--A--A--S--S--Q--D--M--I--D--H--E--H--P--E--S  
 rs11227572  
 gtgcagcggcagccccggccagccctgggggtggagccggctgcaggaacaggaagggtca  
 gggaaatgtggagcccttcagtgtgtgcacggagcccttctagaaagctggAACACAGGG  
 tggcgagctgttggagactcagagggacagggtccacagaacagaccggagggtcaga  
 gctacacccctaagtcccacagtgcctccctcactttcttcagCACTGGATA~~TCCATC~~  
 --T--R--I--S--I-

*Fig. 6U*

CGCCAAGCCCAGGTGGACAAGCTGTATGCCGGCCTGAAGGGAGCTGGCTGGAGAGCGGCCGG  
-R--Q--A--Q--V--D--K--L--Y--A--G--L--K--E--L--A--G--E--R--R-

GAGCGCCTGCAGGAGCACCTCCGGCTGTGCCAGCTCCGCCGCGACTGGATGACCTGGAA  
-E--R--L--Q--E--H--L--R--C--Q--L--R--R--E--L--D--D--L--E-

CAGTGGATCCAGGAGCGCGAGGTGGTGGCGGCCTCCCACGAGCTGGGCCAGGACTACGAG  
-Q--W--I--Q--E--R--E--V--V--A--A--S--H--E--L--G--Q--D--Y--E-

CATGTGACTgtgagtgttagggagggcacccagctcagatcaaccgtggaaagagtggagg  
-H--V--T-

acccacagggagactaggacctagtcccaggcagagcactgagggctaagggcaagac  
caggctgagcaggcactgtcctcctgggtttgaggtatatgattgaagaggccggcat  
aggctcacacctgtaatcctagcaccttggaggctgagatggaggattgcttgagtcc  
aggagttcaagaccgcctggcaacatagtgagacccccatctctacaaaaaaaaattt  
tttttattagccaggcacagccgtgcatgcctgttaggccaactacttaggaggctgag  
gtgggaggatctttagcctggaggtcgacactgcagtgattgccttagtgcactcca  
gagcaagaccatatcatcttaatacatacacacacacaatttgaagaattgtacag  
aaaggggacatgaagctgagctgagacaagggcaaacagggaaattatcactcacacttt  
caacagaattcaacaaaacggagtatgatccacccccccgacagttttggggcttggacc  
ctcagtcgtacagaaatatggaccatagtttaggtatccttgtgaagcctaagggtcc  
caagcaccttggaaacgccttgcgtgggggtggaggagtgctgtcaaatacagggtcagtg  
ggaagcagctgcacctactcctaaccacagggaaagaaacctgtcccgccccatgcag  
cggggagaggtgggttctgagttggagtgaggactcgctccatcaggcaggaccaca  
tccccctaacatcacggcatggctgtccatctgcttcttaccagATGCTCCGAGACAAA  
-M--L--R--D--K-

TTCCGAGAGTTCTCCGGACACAAGCACCATCGGTCAAGGAGCGCGTAGATAGCGCCAAT  
-F--R--E--F--S--R--D--T--S--T--I--G--Q--E--R--V--D--S--A--N-

GCGCTGGCCAATGGCTCATGGCTGGGGCATGCTGCACGGCCACCGTGGCGAGTGG  
-A--L--A--N--G--L--I--A--G--G--H--A--A--R--A--T--V--A--E--W--  
rs623022  
Y

AAGGACAGTCTCAAGAGGCCTGGCTGACCTGCTTGAGCTGCTGGACACACGGGGTCAG  
-K--D--S--L--N--E--A--W--A--D--L--E--L--D--T--R--G--Q-

28/33

*Fig. 6V*

-V--L--A--A--A--Y--E--L--Q--R--F--L--H--G--A--R--Q--A--L--A-

CGGGTGCAGCACAAGCAGCAGCAGCTCCGGACGGACTGGCCGACCTAACGCTGCC  
-R--V--Q--H--K--Q--Q--L--P--D--G--T--G--R--D--L--N--A--A-

GAGGCCCTGCAGCGCCGACACTGTGCCTACGAGCATGACATTCAAGGCCCTCAGCCCCCAG  
-E--A--L--Q--R--R--H--C--A--Y--E--H--D--I--Q--A--L--S--P--Q-

rs12804382

gtctgacccaagcatgaggaggtggggaggcttagaggaggccatggggaaatgccgg  
ggagagttcccaggagctaggtagagctcagaaatgctggaggacggcactttcga  
ggccatttcagatgagaaattgctccaagaaaacatctctgttggccaggcgccgt  
ggctctcgctgtaatcccagcactttggaggccgaggaggcggatcatctgagattg  
ggagttcgagaccgcctgaccaacatggagaaaccttctactaaaaataaaaaat  
tagccaggcatggtgccatgcctgtaatccttagctactcaggaggctgaggcaggaga  
atcacttgaacccaggaggcggagggttgcagttagtgcagatccctccactgcactccag  
cctggcaacaagagcggaaactccgtctaaaaagaaaacatctgtggcaggggac  
atctagccactgcattcacactttgtatgacaagccccctttagagagccatgtcttc  
agggttcccagtggccaggagactttcatgctgcgttaatcagcctgttagtgccttc  
ccaggctttagcttaggctcacactcagtaagcatggatccttccttcaaccactagaa  
ccatcagccctgcctgaatccctttcaggctaagttagcacgttcttagcaatctcca  
aggcttatcacatcctagacaggcctctgaaagttcaactggcaatggctttccag  
atcctccccagttaaagcccagggtgttagatgagacccaacaggcagagtgtggccag  
ccctgccttcatcttcccagcacccactccctggctctccctgtttttgg  
gtgagacaggcaggaggtgagagtggcagccaggactgaggcaggctgcggctcttgg  
gtcccccgctgcctgactctgaccggggctccgcagGTCCAGCAGGTGCAGGACGAC  
-V--Q--V--Q--D--D-

GGCCACCGGCTCCAGAAGGCCTACGCTGGAGACAAGGCTGAGGAGATGGCCGCCACATG  
-G--H--R--L--Q--K--A--Y--A--G--D--K--A--E--I--G--R--H--M-

CAGGCCGTGGCCGAGGCCTGGGCCAGCTTCAGGAAAGCTCTGCCGCCGCCAGCTG  
-Q--A--V--A--E--A--W--A--Q--L--Q--G--S--S--A--A--R--R--Q--L--

CTGCTGGACACCACAGACAAGTTCGGCTTCTCAAGGCTGTCGGAAACTGATGCTCTGG  
-L--L--D--T--D--K--F--R--F--K--A--V--R--E--L--M--L--W--

29/33

*Fig. 6W*

ATGGATGAGGTCAACCTGCAGATGGATGCCAGGAGCGTCCCCGgtgagaatcccagggc  
 -M--D--E--V--N--L--Q--M--D--A--Q--E--R--P--R

tcagggcctcacgggttgtgtggctggctgtgatgggctgtggcagcaggtgatgc  
 tgtttctgcctccccccagGGATGTGTCTCCGCAGATCTAGTCATCAAGAACAGCAA  
 --D--V--S--S--A--D--L--V--I--K--N--Q--Q-

GGCATCAAGGCAGAGATAAGGGCCCGGGCAGACCGCTTCCTCCTGCATCGACATGGGG  
 -G--I--K--A--E--I--E--A--R--A--D--R--F--S--S--C--I--D--M--G-

AAGGAGCTGGCCAGGAGCCACTATGCCGGAGGAGgtgggtgaggcctgggtggcc  
 -K--E--L--L--A--R--S--H--Y--A--A--E--E-

ggccattctcaactgtcgagtgcatacgctgacttctggtggtggcttcctc  
 attctcttttcaagaaaccccttcatctccctcagtgccattcacgggttctcg  
 cattctggaaatctctcccaagaaccgtctgccttgcacatcagaaatcactggtctg  
 ggccaggcatggtggtcactgcctgtaatcccacactttgggaggcccaggcggcaga  
 tcacctgaggtcacgagttcgagaccagcctggcctccatgggtgtggctggctgt

rs508996  
 gatgaaaacatggtaaaccccggttccactaaaaacacaaaaattagccggcgtgg  
 cgggcacctgtaatcccagctactcggaggctgaggcaggagaattgcttgaacctgg  
 ggtggaggttgcagttagctgagattgcgccactgcattccagccttggtgataagagcg  
 aaactccgtctaaaaaaaaaaaaaaaagaaagaaagaaatcacctgtccccctggatgactcc  
 ccagggggcgtggaaagatctcacatcctggtgtaactcatatctctgccccccggcccc  
 cagATCTCAGAGAAGCTGTCTCAGCTGCAGGCACGGCGCCAGGAGACAGCTGAGAAAGTGG  
 -I--S--E--K--L--S--Q--L--Q--A--R--R--Q--E--T--A--E--K--W-

CAGGAGAAGATGGACTGGCTTCAGCTGGgtgagctgccaaggggcccccaggccctgtgg  
 -Q--E--K--M--D--W--L--Q--L--

ggagtggggggcatcctgcaccctgtgggttccagagttagtgagacttaggaaccctgg  
 tgtgaaactcacatgcccatttcgtgtgcctggcacagctctggggcagtggctt  
rs2276138  
 ctctgtgtccctgttcttgagggttgcacatggccactgtgcctgcactggccactctga  
rs2276137  
 cccaccatcttcctgcaaccccccgtccatgtggcagTTTGGAGGTGCTTGTGTTGGAAGA  
 V--L--E--V--L--V--F--G--R-

GATGCAGGGATGGCAGAGGCCTGGCTCTGCAGCCAGGAGCCACTGGTGCGCAGCCGTGAG  
 -D--A--G--M--A--E--A--W--L--C--S--Q--E--P--L--V--R--S--A--E-

SUBSTITUTE SHEET (RULE 26)

*Fig. 6X*

CTGGGTTGCACGGTCGACGAAGTTGAGAGCCTCATCAAGCGGCACGAGGCCTTCCAGAAG  
 -L--G--C--T--V--D--E--V--E--S--L--I--K--R--H--E--A--F--Q--K-

TCAGCAGTGGCCTGGAGGGAGCGATTCTGTGCGCTGGAGAAGCTTACTGCGgtgagggac  
 -S--A--V--A--W--E--E--R--F--C--A--L--E--K--L--T--A-

acaggaccccgatgccactccaacctgctccctgacctgtgctggcttcgttttgars2276136  
 gaagacatgtcttctcttcccactgcaccagtggccccagatgtgagctgagagtgc  
 cattgtcaccacaccccttaggagggtcagttccctgggtatgccgagacaccccttc  
 ccgcctttgttctgtcccccaccatcttcaaattctgttctccatattcatattca  
 tcaaatttagagttgagacattcagctccctgtttccctggccctttacacgcaacc  
 ttctacacgctgaatgcagcataaaaacagCTAGAGGAGCGGGAGAAGGAGCGAAAG  
 -L--E--E--R--E--K--E--R--K-

AGAAAGAGGGAGGAGGAGGAGCAGCGGGAAACAGCCGCCTGCTCCGAACCCACAGCCAGT  
 -R--K--R--E--E--E--R--R--K--Q--P--P--A--P--E--P--T--A--S-

GTCCTCCAGGGGACCTGGTGGCGGCCAGACAGCTTCTGACACCACCTGGGACGGgtga  
 -V--P--P--G--D--L--V--G--Q--T--A--S--D--T--W--D--G  
 gagccaggatgcctggtaggaggaggcggtgagcccaggccacccaggactaatgat  
 tctgtgttgccttggcatccagAACCCAGCCACGCCACCATCCACACAAGCA  
 --T--Q--P--R--P--P--P--S--T--Q--A-

CCCAGTGTAAATGGAGTCTGCACAGATGGAGAGCCCTCACAGgtgacccactgtccctc  
 -P--S--V--N--G--V--C--T--D--G--E--P--S--Q--  
 tgtccccatcgagtcgtggccatccccgcacatcctttacagattttgtc  
 ctgcagCCCCCTGCTGGGACAACAGAGACTTGAGCACAGCAGCTCCCCGAAGGGCCGgt  
 -P--L--L--G--Q--R--L--E--H--S--S--F--P--E--G--P--  
 gagttccctgcaagtgtgggtataactgtgaggcgaagggtccagaggggtggtg  
 agtgcgggtggggagttactgggatgggatggagaggcagaggctcacaggcagctt  
 gggggcaggaagaccaactcctggacacggagcttccctggcacccaggtagggatctcc  
 cgtctcaaccccttgcactgacactgattccccccagGGACCTGGCTCAGGGGACGAA  
 -G--P--G--S--G--D--E-

GCCAATGGGCCCGGGGAGAGAGGGCAGACCGGACTGGGGCCGGCCCATCTGCAATG  
 -A--N--G--P--R--G--E--Q--T--R--G--P--A--P--S--A--M-

*Fig. 6Y*

CCCCAGAGCAGGTCTACCGAGTCAGCCATGCTGCCACCCTGCCGCCTCGAGGCCAGAG  
-P--Q--S--R--S--T--E--S--A--H--A--A--T--L--P--P--R--G--P--E-

CCATCTGCCAGGAGCAGATGGAGGGATGCTGTGCCGCAAGCAGGAGATGGAGGCCTTC  
-P--S--A--Q--E--Q--M--E--G--M--L--C--R--K--Q--E--M--E--A--F-

GGGAAGAAGGCTGCCAACAGgtacagcctctctggagcctgctctcagagggcacttccc  
-G--K--K--A--A--N--R

cagagcctctgccagatagagggaggatgcccttagagtacatcttctggcaaagg  
gttaggacttgggaccagagcggggcctcaggggaggaccagagggtgtgaagacccgtggc  
ctaaagatgggagcagaactggaagtccctaggacacccaagagggctccaggttgcgggc  
gccactgaggccggccagtcagcaccccgctccctcgcaGTCCTGGCAGAACGTGTACTGT  
--S--W--Q--N--V--Y--C-

GTCCTGCCGTGGAGCCTCGGCTTTACAAGGATGCCAAGGCAGGCCAGCGCGGGAGTG  
-V--L--R--R--G--S--L--G--F--Y--K--D--A--K--A--A--S--A--G--V-

CCATACCACGGAGAAGTGCCTGTCAGCCTGGCCAGGGCCCAGGGCAGCGTCGCCTTGAT  
-P--Y--H--G--E--V--P--V--S--L--A--R--A--Q--G--S--V--A--F--D-

TACCGAAAGCGCAAACATGCTTCAAGCTGGGtaggaacagggAACAGTgctctcgga  
-Y--R--K--R--K--H--V--F--K--L--G  
tgggaggagagtggagtgacacaggtgagccatgagtcaggtccagaggaggaggag  
tcctgtaaggagcctgagtggagtaaccaggccagccacttgggatagtgttagatgagg  
gaggccggagattctgggtctccacaccaaggggagcaggagaaccaagccaggcct  
gacggctgccaatgtcaaggtaaaaattaccagggtggaaatccaaaggttagggatc  
tgggaagacctccaggggctgtccctgcccgcagcaagcagccggagcaggaggcccg  
ggctgggtggaagtgagctccccctgcctctggggccagtcagaaaggatgtgcttc  
cacagtctgggaagctgaagaatgtgcagagcgtgtctgggtccggctctgagggtgca  
gccagatgtcccagcctgggtttgggtcatgattacgcttcaccagcagctacctggca  
gatccggattctcagctctgcctgtggcctctctcccaacagCTTACAGGATGGAAA  
--L--Q--D--G--K-

GAATATTATTCCAGGCCAAGGATGAGgtgagotgtccttcgtgttccctctgtccgt  
-E--Y--L--F--Q--A--K--D--E-

ccattccagaagcttccagctgcagactcccctttccctgtcctcccttttct

ggtgtttgtccttggtaagactggatgtgtgcgcacggccgcaccaggccgactccct

**SUBSTITUTE SHEET (RULE 26)**

Fig. 6Z

gtcttaaggccggccacattctcctaatacgatgaaacagtcagtcacttctgcctc  
ctcttacacttccctgtgtccactgcggccaactcagcacagtgtccctgaagctgatt  
gagggtcttccatccacagGCAGAGATGAGCTCGTGGCTACGGGTGGTGAATGCAGCC  
-A--E--M--S--S--W--L--R--V--V--N--A--A-

ATTGCCACAGCGTCTGCCTCTGGAGAGCCTGAAGAGCCGGTGCTGCCAGCACCA  
-I--A--T--A--S--S--A--S--G--E--P--E--E--P--V--V--P--S--T--T-

CGGGGCATGACCCGGGCCATGACCATGCCCGGAGTGTCAACCCTGGGGCTGAGGGGCCT  
-R--G--M--T--R--A--M--T--M--P--P--V--S--P--V--G--A--E--G--P-

GTTGTGCTCCGCAGCAAAGACGGCAGAGAACGAGAGCGAGAAAAACGCTTCAGCTCTTT  
-V--V--L--R--S--K--D--G--R--E--R--E--R--E--K--R--F--S--F--F-

AAGAAGAACAAAGTAGTTGGGGCAAGGTCCCAGGCCACTCCCTCCCTCGGTCAGGAAA  
-K--K--N--K--\*-.....

CTGCCAGGGACAGTCGACAGGGACCGCCCTTTGTCAAGGACAATGCCCTGCTGCTAGGGT

CTGGTGGCCAAGGTCAACCCATCACCGAGGAAGTGTCACTGGGGACGAGTCCATGTTCCCAA

rs693564

TCTGCCACCCCCCCCAGCTTCTCCCTTCCCCAGCCTCGTGCCTCTGTCCCTCACCAACG

GTGTGGACAGTGCCGCACCCTAACATAGGCCATGTGGGGAGTGGCTGCCCTGCCTCAG

.....

rs11828658

ACCAACCCACCAACCTCTCACACCTGAAACCAACCCCCCTGGGGGGCAGAAGTGCCAGATGACAG

TCAAGAGGCGCAGGAGCCCTCCCCACCCCCACCCCTGTAACCTCCAGCTGCCACTCCAT

.....

rs1064352                          rs12575562

Y Y  
CTCCAGTGCTCTCAATGGCTCCAGGTGTGTTGTCGGGGACAGCCACCGCCTTGAGTC  
SUBSTITUTE SHEET (RULE 26) . . . . .

*Fig. 6AA*

TGGCCAAGGAGGTGATTAAACAGCTCAGCTTCTC

.....

110.02600201.txt  
SEQUENCE LISTING

<110> REGENTS OF THE UNIVERSITY OF MINNESOTA  
<120> IDENTIFICATION OF A GENE ASSOCIATED WITH  
SPINOCEREBELLAR ATAXIA TYPE 5 AND METHODS OF USE  
<130> 110.02600201  
<140>  
<141>  
<150> 60/655,172  
<151> 2005-02-22  
<160> 77  
<170> PatentIn version 3.3  
<210> 1  
<211> 36147  
<212> DNA  
<213> Homo sapiens  
<220>  
<221> CDS  
<222> (160)..(316)  
  
<220>  
<221> CDS  
<222> (5419)..(5570)  
  
<220>  
<221> CDS  
<222> (6005)..(6178)  
  
<220>  
<221> CDS  
<222> (6993)..(7084)  
  
<220>  
<221> CDS  
<222> (7229)..(7309)  
  
<220>  
<221> CDS  
<222> (7654)..(7769)  
  
<220>  
<221> CDS  
<222> (10371)..(10483)  
  
<220>  
<221> CDS  
<222> (10631)..(10818)  
  
<220>  
<221> CDS  
<222> (12381)..(12498)  
  
<220>  
<221> CDS  
<222> (13101)..(13259)  
  
<220>  
<221> CDS  
<222> (13582)..(13884)  
  
<220>  
<221> CDS  
<222> (15563)..(15716)

110.02600201.txt

<220>  
<221> CDS  
<222> (15932)..(16802)

<220>  
<221> CDS  
<222> (19679)..(19816)

<220>  
<221> CDS  
<222> (20118)..(20874)

<220>  
<221> CDS  
<222> (21792)..(21994)

<220>  
<221> CDS  
<222> (22318)..(22408)

<220>  
<221> CDS  
<222> (22615)..(22761)

<220>  
<221> CDS  
<222> (24860)..(25123)

<220>  
<221> CDS  
<222> (27037)..(27261)

<220>  
<221> CDS  
<222> (27539)..(27628)

<220>  
<221> CDS  
<222> (27954)..(28214)

<220>  
<221> CDS  
<222> (28300)..(28430)

<220>  
<221> CDS  
<222> (28660)..(28864)

<220>  
<221> CDS  
<222> (29742)..(30116)

<220>  
<221> CDS  
<222> (31116)..(31360)

<220>  
<221> CDS  
<222> (31456)..(31594)

<220>  
<221> CDS  
<222> (32219)..(32303)

<220>  
<221> CDS  
<222> (32550)..(32746)

<220>  
<221> CDS

110.02600201.txt

&lt;222&gt; (33088)..(33230)

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (33320)..(33395)

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (33481)..(33531)

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (33752)..(33972)

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (34232)..(34405)

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (34959)..(35001)

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (35295)..(35525)

&lt;400&gt; 1

ccactgagca gccaaccgca gcctctggcc acaaggagag cggagcacag gtagggcaag 60

aagacagaag gccccgggtgg tgagtggagg gcttgggttgt ccctggacac ccctcactgg 120

ccacattctc cttgcagggag caggaagccg cctaccacc atg agc agc acg ctg 174  
Met Ser Ser Thr Leu  
1 5tca ccc aca gac ttt gac agc ttg gaa atc cag ggc cag tac agt gac 222  
Ser Pro Thr Asp Phe Asp Ser Leu Glu Ile Gln Gly Gln Tyr Ser Asp  
10 15 20atc aac aac cgc tgg gac ctt cct gac tcg gac tgg gac aat gac agc 270  
Ile Asn Asn Arg Trp Asp Leu Pro Asp Ser Asp Trp Asp Asn Asp Ser  
25 30 35agc tcg gcc cgc ctc ttt gag agg tct cgc att aag gct ctg gca g 316  
Ser Ser Ala Arg Leu Phe Glu Arg Ser Arg Ile Lys Ala Leu Ala  
40 45 50

gtgaggtcag aggagggtcg aggtgggggg atgctggagg aggcttgtc agctcgcat 376

gaagagccct ttaactttat ggaaagactg acttttctt ctaagtggga gacccagagc 436

actctatcaa ttcccggtc cccacatcca gcctcaagaa taggctccag gccattcaga 496

actttccccca gctttccccca ccaagcatag agacgcttca ctggccctg gactggctga 556

gctgactaaa acagttccc tctctggcct cagccccgg gtcatgtccc ttcatccttc 616

ggcactctcg ctcttgccac tccttctcc gcctacctt cccttacctg tcatgccctt 676

gttgtgtctt ctttgatttc cactgtcatt tcccacacct gttccagtc tattattccc 736

ctctgctccc tctccacatcg ttcccttccac ctgcccacatg ttactccat tatctcatct 796

gcctgttccc tgaaggcagg tccctccacc cgctgctgct ggtttctgca gccgcccatt 856

ctccctgggt tgctctctgc ctgccttggc tcctactttt ccctgcagcc cagggcctgg 916

gctctccgct ctgagcggaa actcgccaa cggcctccct gtctccggg ggaggagttg 976

tcatggcaac aagagtgcgg ctgcccacatcg acaccacccag tggctggcat ggggctgact 1036

110.02600201.txt  
 cacccacaca ccccaaggcc acctaccccg tctggcttcc tcctccatcc cactcttgc 1096  
 attgccattg acctgctggc tgcccagccc agcccaaccc aaccaggatct gaggggtccc 1156  
 agctactctc accgagaggt ggtgctccga gtgggggtgg gggttggcca ctagaagaag 1216  
 agagactcca agggggaggt tcccaaacct ggccacatgt tttacaggat cacccaggg 1276  
 tgtttttac aaatacaggt ttccaagccc ctccctagac acaataagtt agaatcttg 1336  
 tccagaactc aggattttc aaaagctccc caagtgattc tttttagcca gctcagcacc 1396  
 agtctgagga ccccttgtg gcagccatag cttccagga cccctctca ttttctcctg 1456  
 ggcctctcg tccttcttgt tgggtcttca ggccggctgt gaaacactcc taccagtgt 1516  
 tctatacatc agagaagaca gaagggtgtgg gttatggaga atgtcatgtt tggggaaaat 1576  
 gttccctggca ctctgggtt ggattcatgt agcctctaga actgtcatga tgtctatgt 1636  
 tgtacaacat gtgccaccca ccatgtttgc acacatgaca aatgccatcc aggactgtca 1696  
 ccaygcctgt gtacatagga cacatgcctc ccaggactct caccatgcct gtgtacatag 1756  
 gacacatgcc acccaggact gtcaccatgt ctggctgcac ataacttggg ccacccagca 1816  
 ctgtcaccat gtctgtgcac acatctgtgt gacctcgtgc tctcattggg tgtttatctc 1876  
 caccactgaa ctgggctctc ctgaagaggc agctctgcat gggatttgc tagcactgt 1936  
 tctccagcca gtaggtgtca aataaacact gaatgaatga agatgaggaa gggcaagtgt 1996  
 aggctgcaca tgacaaacgc tgaagcccag aggaaggcg cctggatat gtcggccagca 2056  
 gcggggcgtc tgggggcagg gatggcctgg tggaggatgg ccctgggtgg ccagaggc 2116  
 gggaaaggggc tggcactgga ggtggtgata agcaagtgga aaggatggta acagccctcg 2176  
 tttcacagat gagggagtga aggaaaatgc ctcgcacagt gcctggcttt tcctcttga 2236  
 ggatttcctc gtttccagat ttctcttgc ctgtttctga gctggggctc aaagcccaca 2296  
 catcctggag aaggacctcc cagagtcacc cagaggggg tcgcctcg 2356  
 aacctctggg aaacaggagt ttgcattggg ggccctctc gtgggtcctg tgcattgtt 2416  
 ggtaacttgg ggcctgatgg gcgggtctt aaggcgctgg cagtgccctc ctccgccc 2476  
 ccgcagggcc tggccctgcc cccgcccccc acactccg 2536  
 gtccaaacctg tttgtctgga gccgcctgtc tcctctcacc gcagccaggc gtgc 2596  
 tctctgtctc aggagttcg 2656  
 ccattcttga atccccctgac gctggactcc agcctgaggt tccacttggg aacagaaact 2716  
 ggccccctgc crgcggggag catggactct gggccctggg ctctgttaact agcagctggc 2776  
 agcatgaatg ggaacggagt gagsaggg 2836  
 ttttactatg aggctgtgga aggggcccac aaccccg 2896  
 gctttcatca accctgctca gtatgccagc gtgctggaaag gacgcttcaa acagctgcaa 2956  
 ggtgaggccc gggggcccccc tgggggacg actgtgcctg ggcgagagg 3016  
 ggggagagtg gccggagctg aggacctggc acaggacaca gtgtgactcc agtgtttgt 3076  
 ttgggagaca acgtgatgca gcttagaggg gataattaca agaggccg 3136  
 atgtcaacgg aatggggaaa ggcttggacc tgctggcagg gaccgtgtat gatggcatgg 3196

## 110.02600201.txt

gtgggtgtga	ctgccgggcc	ttgggggtgc	gcatggagca	aaggcaacta	gccatggagg	3256
gcgtgggggt	gccagcacct	ggggaggcgc	cagcagtgcc	tctgaggtgg	atccggatct	3316
ggcggtggag	cgctggccag	gacagttagg	ggaggcggag	gggtgagcca	ggcccaggga	3376
cctggagatt	gtgtttcca	ggcctctgag	cgtgtgtctg	aatgtgtaca	cggggcagt	3436
gtgtgtggaa	gcaaaggag	tgtctcgaag	gagaggatct	gtgcaaacag	gcccaggtag	3496
atgtctgcgc	aggtgtgtgt	tcactctggc	acaagtgcct	aggtgtgttt	ttgggaatca	3556
ctgttgtgtg	tgtgtgtgt	aggatctagg	gatgctgtgg	gcgtcttgc	acgtgacact	3616
tggtttctgg	tgcttggcct	tgcccatttt	cctcctagtt	gccactaatt	ctgtgtcttc	3676
tgactgtgtg	ccaaaatggc	ctcacccca	ggtgcccagc	agtagggcag	tgcccccgtga	3736
tccagcaatg	gcctgtgacc	cacgctcctc	ctggctcaga	tggggcctgt	agacctgccc	3796
ggaggccccca	aggcaccttgc	gtttcatggg	tgctggggag	catggggccg	tcccatgtta	3856
tttaggacccc	aggtttggtg	ccagtagatg	ggctggaggt	gaaagacacg	gagcgacgtg	3916
ggtgtgttgc	tgtgtcttat	gggtgaggtg	cttaggaatg	tgagattaag	ggatgagtc	3976
ttccggtttc	tgasatgctg	gtgctgagac	ggaagcaaaa	gtgactcttc	agacccca	4036
cacttgggtt	gggtccttgc	acaggctct	ctgtggaaa	gtggcagaaa	tcctcatcac	4096
cagagcagtt	tctgatgctt	tcactttgtc	acttagaacg	ttcttaattt	ttatttcct	4156
acggaaccta	gtatttactt	ccccccttca	gtagctcccc	tccctgtgg	gctggcagc	4216
agggcattgc	ctaacttttt	tttgagattt	gtacctaact	ctgcctggct	ttttttcccc	4276
catgtttcct	ggcctcactg	ttttcccagt	tcccaactct	caacatcttgc	ttcttctaag	4336
caatgggtg	tttatgtcct	ggaactggat	ccccttgca	ccagctgaat	gtactgtccc	4396
caaccccccag	aagcagagga	gcctgcctt	tcttaagtt	atccaaggac	agtaagctga	4456
ctcaggaggc	ctgcagcctc	tataccatca	agtccataatg	accgcttgg	agagccgctg	4516
caataatttt	agcccacacc	ttctcccttgc	tctgcaacca	aaagagaaca	gccataacctc	4576
ccaccggcct	ggcagctct	tgctctcctt	cgggcccagat	cacaccactt	tataacagca	4636
gctcaccaag	gagccatttt	gtgccaggca	ccaaactagtt	gttttatata	ccttatctca	4696
tttaatcaac	acaatgacct	gttgaagtag	aaagtgtatgt	tcccatttta	cagatgagac	4756
attggggctc	tgtatgaatga	agtcaatttgc	ccagcattcc	actgcttcca	cgactgcccc	4816
atcccagccc	ctccctgcat	gtcccatgg	ttctgaggat	gaaaggctt	catccttct	4876
tctcaaagcc	tatttcttt	tytttcttt	tctttttttt	tttgagacg	gagtcttgct	4936
ctgtcgccca	ggctggagtg	cagtggcaca	atctcggtc	cctgcaagct	ccccctcccc	4996
gttacacgtc	agtctgcctc	agcctccga	gtagctggga	ctacaggcgc	ccgcccactat	5056
gcccggctaa	ttttkgtaa	tttagtaga	gacggggttt	caccgtgtt	gccaggatgg	5116
tctcgatctc	ctgaccttgt	gatccggcc	ccttggcttc	ccaaagtgc	gggattacag	5176
gcatgagcca	ccacacccaa	cttcgccta	tttctgtct	ttcatcttcc	ttctttacct	5236
tgcttatgag	ctcattgcaa	gttacggc	ttcctccgg	tcatcttcc	tcccttcc	5296

110.02600201.txt

ctccttagag gaccccgctt cccatctctc aggcttcctcc atccccacc cctgccaact	5356
ggtgtttaga agcaggcata ctgaactcga gggacagaaaa taaacccatg ttgccaccac	5416
ag at gaa cga gaa gct gtg cag aag aaa acc ttc acc aag tgg gta Asp Glu Arg Glu Ala Val Gln Lys Lys Thr Phe Thr Lys Trp Val 55 60 65	5462
aac tcg cac ctg gcc cggtgc acg tgc cggtgg gac ctg tac agc Asn Ser His Leu Ala Arg Val Thr Cys Arg Val Gly Asp Leu Tyr Ser 70 75 80	5510
gac ctc cgg gac gga cgc aac ctg ctg agg ctc ctc gag gtg ctc tcg Asp Leu Arg Asp Gly Arg Asn Leu Leu Arg Leu Leu Glu Val Leu Ser 85 90 95	5558
gga gag ata ctg gtgagctgtg gtcatgaagg gaagtgggg ggcctwagaa Gly Glu Ile Leu 100	5610
aagtgtccaa ggggaccagc tggaagaaca gccctcagag agagtgcgg cacagggcgg gaagcagtgg ctccctctgc tcagagratg gggtcatttc aagtgccttt gccacttaga ttacatagga gcattgcttc atggagacag gaacagtggc cccattgaaa acatttcttt tcccacctac tctgggtgca attaattccc ttggaatct taatctgcc aagcctcatc tgtaagctca gcttggctc cagccctgg gctgtgcctg tctgtgttcc tgaggttagt catgggaaag gtgtgcaaag ctggccctg tgaaggctgg ggaagaaggt ggaagccac agtcccggtgc catttgccc ccgtccctcc acag cca aag cct aca aag ggc cgc Pro Lys Pro Thr Lys Gly Arg 105 110	5670
atg cgg atc cac tgc ctg gag aac gtg gac aag gca ctg cag ttc ctc Met Arg Ile His Cys Leu Glu Asn Val Asp Lys Ala Leu Gln Phe Leu 115 120 125	5730
aag gag cag aaa gtg cac ttg gaa aac atg ggc tcc cat gac att gtg Lys Glu Gln Lys Val His Leu Glu Asn Met Gly Ser His Asp Ile Val 130 135 140	5790
gac gga aac cac cga ctg acc ctt ggg ctg gtc tgg acc atc atc ctt Asp Gly Asn His Arg Leu Thr Leu Gly Leu Val Trp Thr Ile Ile Leu 145 150 155	5850
cga ttc cag gtacccagc acactgtcac acagggtgtg gttcctgccc Arg Phe Gln 160	5910
gggtctgtca cgcgaagccg ttccctgtgg cccatgacac aacaaacaca aagatggagg atctagaacc ttatgttagca gctgcccattgg cggcgctatc caagcaccag atcagtggac tcccctctt cagcaggta gaaaccagct tgggtgaatg tcacacccct gtggcaaatt ccctgtgttc tggctgtgt gttggcgtg tgcaatggga ccaggaccac aggttggcc acaggagact ggaagttta ccaactggtc cttcaccaca gggatttggg aacttctaga gggtctgtc actggcacca attccctgcc ataaactctg cctgcctccc ccaatggAAC agtcaccctc ctccctctca tgccccagtg tcctcatgcy atcacagcgt tcctgatcac ccccctcctt cccgggttct gtcggcattca cttggagag ccagttta gagaccttgg agagccagct ctcagaaaat gttcactctt ctgcctgcag ttccccacct ggagggacca gcaccctcca ctcagacaca cccattcttc agcatcctac cctaattctgc tgcctcaaatt	6025
	6073
	6121
	6169
	6218
	6278
	6338
	6398
	6458
	6518
	6578
	6638
	6698
	6758
	6818

## 110.02600201.txt

cccttgcacc cagtgtttg tggccctgg cctaggttga aaggaccta ggctatgcct	6878
gtccacccca ttaccacagt ctaaagagca gttcctgcc atttatctga gcacatctctga	6938
taacaccagg agttcctgtc cagggactga gggccctgct tgcttccttc ctag atc Ile	6995
caa gac atc agt gtg gag aca gaa gac aac aag gag aag aag tca gcc	7043
Gln Asp Ile Ser Val Glu Thr Glu Asp Asn Lys Glu Lys Lys Ser Ala	
165 170 175	
aag gat gcc ctg ctt ctg tgg tgc cag atg aag act gca gg	7084
Lys Asp Ala Leu Leu Leu Trp Cys Gln Met Lys Thr Ala Gly	
180 185 190	
gtgaggacac cctgggcgtg tggcactgga gggtcagtga ccccccaggct gtgctggc	7144
ccaggaacca tgagctggtg actgcccctgg agtcatagaa cagagctcag ccctactcac	7204
cattttctga atctctgttt atag t tat ccc aac gtc aat gta cac aac ttc Tyr Pro Asn Val Asn Val His Asn Phe	7256
195 200	
acc acc agc tgg aga gat gga cta gct ttc aac gcc atc gtg cat aaa	7304
Thr Thr Ser Trp Arg Asp Gly Leu Ala Phe Asn Ala Ile Val His Lys	
205 210 215	
cac cg gtgagaagat ggggttagca aacgcagcag agctgctgg tcggctaattg	7359
His Arg	
acctgggtg cacatgtgga gacagagatg agatgagctg cacagcagtc agacggaagg	7419
aatagcactc ggagcagaac cagaacatgg ggggtgcaggg ctgcgcagag cacttgtat	7479
gggaagcagg ggaggtggtg tggggagggg ctgtgagcag cagtgtgggt ttgggatgcc	7539
tacttctgga caaaaattggg tcggtgactt agggataaaag attgggtgtgg gtttcccttt	7599
cctcttcrag atctcacatg tttctgatcc cttctgtctt gcctccccac ccag g cca Pro 220	7657
gac ctg ctg gat ttt gag tct ctg aag aag tgt aat gca cac tat aat	7705
Asp Leu Leu Asp Phe Glu Ser Leu Lys Lys Cys Asn Ala His Tyr Asn	
225 230 235	
ctg cag aat gca ttc aat ctg gct gaa aag gaa ctg gga ctt acc aag	7753
Leu Gln Asn Ala Phe Asn Leu Ala Glu Lys Glu Leu Gly Leu Thr Lys	
240 245 250	
ctg ctg gat ccc gaa g gtggggccag agctatgtga aaaagaggtg gtagggtaga	7809
Leu Leu Asp Pro Glu	
255	
gacaggaggt ggaccagtgg ttgccagggg ctgcggaggg ggaagacatg gagtaaccac	7869
caactgatac aggggcttct ttgggtgggg aagagaatgt tctggaatta gaaagtggtg	7929
atggttacac aacttgtaa atatactgaa aaacacgtgg ttgctcacac ccagcacttt	7989
gggaggccaa ggcaggagga tcacttgagc ctaggagttt ggtaccagcc tgggcaacat	8049
agtgagaccc catgactacc cctccccccc aaaaaaaaaata tatacttta cactgaaaac	8109
cactgaatta tatacattga aaggtgaat tttatggtgt gacaatcatt tctcaggttt	8169
tttttttttt taacaaaaga ctgttaggtag gaaraaaggc agcccaagtga cagaaagagc	8229

110.02600201.txt	
cctggactgt	gggttggtc ccaggcagag cacagaggag ctctgtgaca tacagccagt
ggcttagcctc	tctgtggctc agttycctca tgtctaaaaa gagagaattt gctctctaagg
gctcttctga	aattcttagga ttctaaatta agctacatgt tgctggccac ggtggcacct
gcctatagtc	ctagctgcta gtgaggctga ggcgagagga tcacttgagc ccaggagttc
taggctgcag	tgggttatgc tgattggata tctacactaa attcagcact ggtgacctcc
tgggagcagg	ggaccagcag gttacctaag gaggactgaa ccagtccagg tcagaatgga
gcaggtcaaa	actcccacac tgatcagtag tggatcaca cctgtcaata gccacagtac
tccagcttgg	gcaacagcga gaccatctc tattttttta aaacaaattt ttaagaaaaaa
aaattaagct	gcatgacctt ccactctcac ccagaactgg cctgttagtca gctgaccctt
ctgattggtt	cctgtggttc tggtcataat gccgatgaga tcgatgagat cagtaacagc
cmctgggctc	ttttttttt ttttgagacg gagtctcact ctgttgcac gactggagtg
caatggcgcg	atctcggctc actgcaacct ccgcctcctg gttcaagca attctcctgc
ctcagcctct	ttagtagctg agattacagg cgcattccac cacgccccgc taattttgt
attttagta	gagacaggggt ttcaccatgt tggtcaggct ggtctcta ac tcctgacctc
gtggtccgccc	cacctcggcc tcccaatgcg ctgtgataac aggtgtgagc ccccgccct
gactgcgatt	aggctttat ttgataacctg gttctgcact ttagtgttt taatcttac
aacacccctt	tgaaatcagt actgttgta tcctcagttt atggaagaga agatgggggc
ctataatggg	taagcaactt gcccagggtcc tccggcagat ggagctgaga ttggcgttca
ggcagcctgg	ctccagagcc tggccctcc gctgcctctg ggaaagtgg accaatctt
ggtgcaacca	gaacagggag gagagaaagg ctcgcctgct ggaggtgctc cttaaagttca
ctgggtggat	gttcaactggt ctggccatt ttctccctgt tttccattct acctgttagcc
ctatcacctg	ggcctttta ctgcagaatg acaagtgtca gcgcgttagc aacagacata
ggcttcatgc	agtgccttggggcctc tgggttagcat ggctgtaatt cagaacctgg
aaataaagct	aagcattctg taactttccg agatgacttc agtggctcct gcctgcccatt
gtatccatgc	aaataggcaa gaggctctg gtctcatagt cttaccttta ggataaaatc
tggcctattt	aaaagaaccc acagattca agttatggtc agcctcagg ttgcgttggaa
agccaattcg	gcttggaaaga cactccctc ctggccca cagctgagct ttaggattga
ggttagaaac	ggcgtgatca cacctggaga tgggtgctgag agctggcctc tggccctact
gagctgcagg	gatcacaggg cccatcgatc tgagcatcct aggttcagga ggcttagttc
ataccttccg	tcccactgcc cctcaatcca ttgacttctt aaacattatt ctagcctaaa
agtccatgat	ccatccccac atagcaatgc tggccctttt cttgcctttt tccccaaac
aactgatttt	ggtctccacg gggaaactctg ggtaaaagct tccttgct tctggggttt
gtgctctgct	tggagggccc ttcttgttc caggagaagg ggaaaaaaag atgagctttt
tggtgcctt	ttcatttgaa tggtaagtt gtcatttgc tctaaaagga gatggagaga
tgtcagccctt	ggaggactgg gcaggcccag ggtttggggt aagtgtatgca agttatggat
gaacttctgg	gaggcctgac ccaaattggtc ctcttttgc g ac gtg aat gtg gac

110.02600201.txt

Asp	Val	Asn	Val	Asp
				260

cag cca gat gag aag tca atc att acc tat gtg gct act tac tac cat Gln Pro Asp Glu Lys Ser Ile Ile Thr Tyr Val Ala Thr Tyr Tyr His 265	270	275	10432
tac ttc tcc aag atg aag gcc ctg gcc gtg gaa ggc aag aga att ggc Tyr Phe Ser Lys Met Lys Ala Leu Ala Val Glu Gly Lys Arg Ile Gly 280	285	290	10480
aag gtactgtcca tggcagtag gcataaaggc cagaggaggc ccggctgagg Lys 295			10533
ggtcttactg ccctagtgca agggcagggt ggagctgcag gactgggcca gggaccctgt ggctggact gtcacgtccc tgtcttctgc ctcccaag gtg ctg gac cat gcc atg Val Leu Asp His Ala Met 300			10593
gag gca gag cgc ctg gtg gag aaa tac gag tcc ctg gcc tcg gag ctg Glu Ala Glu Arg Leu Val Glu Lys Tyr Glu Ser Leu Ala Ser Glu Leu 305	310	315	10648
ctg cag tgg atc gag caa acg atc gtg acc ctc aat gac cg <sup>g</sup> cag ttg Leu Gln Trp Ile Glu Gln Thr Ile Val Thr Leu Asn Asp Arg Gln Leu 320	325	330	10696
gcc aac tcc ctt agc ggg gtc cag aac cag ctg cag tcc ttc aac tcc Ala Asn Ser Leu Ser Gly Val Gln Asn Gln Leu Gln Ser Phe Asn Ser 335	340	345	10744
tac cgc acc gtg gag aag ccg ccc aa gtaggtgtcc ctggggcccc Tyr Arg Thr Val Glu Lys Pro Pro Lys 350	355		10792
accctccct gagctgtgct cccacgagag gaagcctaaa ttagcacagc cttcagggag ggaaatttgg cagtagataa atgcagtgga gatttccac accagaagca tccaaacaac atagtgata caaaataaaa ttttttaatg attagtcgtt tttaaaaatc atgctgtgg ccgggcatgg tggctcacgc ctgtaatccc agcactttgg gaggccaagg cgggcgcac accgagacca gtctggccaa catggtaaa cccatctct actgaggtcg ggagtttgag accagcctgg ccaacatggt gaaacccgt ctccactaaa attacaaaaa aattagctgg gcatggtggc acacgcctgt catcccagct actcgggagg ctgaggcagg agaaccacct gaacccggga gacagaggtt gcagtgagcc aagatcacgc cactgcactc caacctgggt gacagagcaa gactccgtct caaaaaacaa acaaacatgc tgtggaaatg attgctatga tgtgttgagt cagtgcgatc catcagataa gattactgt caggtgtcat ggcaacccaa tccttagcata ggatagaggg gactggacct gacaggcagg tcacaaccgg caacatgggt gtggctggta acgtgagaat tgcagaggtt cttatcattt cagtattta gcaattaact gttccaaagta tgtgattgct gttaggcaat tctgttcagt acctgccaaa acgtctgtgt ttatgctgtc atggagttct tttggagttt atatatatca gctctgaaac taatattagt ccagaaaaca attgtttcta tacattccag ggaatgattt aagggttacc ctttaaccta ggaagttgtc ttctggacat taaacaaaga tgtagttca aggtgttta tcaacacaac cactagaaaaa gaaaaataag tcatggcag ccactgaaaa taatattaaa gaaatgcatt		10838	
			10889
			10958
			11018
			11078
			11138
			11198
			11258
			11318
			11378
			11438
			11498
			11558
			11618
			11678
			11738
			11798
			11858

110.02600201.txt

tattgacatg gaaaggtgtt cataaaaat aagtggggga aataggttat aaaagaatta	11918
tttggctgg gcaccgtggc tcatgcctgt aatcccagca ctttggaga ctgaggcggg	11978
tggatcactt gaggtcagga gtttgagacc agcctggcca gcatggtaaa accccatctc	12038
tactaaaaat acataaaatta gccaggtgtg gtggtgcatg cctgtaatcc cagctactcg	12098
ggaggctgag gcaagagatt ggcttgaacc cgggaggcag agrttgcagc gaacccagat	12158
cgcaccactg cactccagcc tgggtgacaa agcaagactc tgtctcaaaa taaaataaaa	12218
tgatgttctt ttggtagcc catatttct ggcttccaca taaaaagttt atattgttt	12278
tgtaagaaga aaagcaacac aggacaccca ttttggaaagag taaagagtcc ttgtgaccct	12338
cgtggcttt aattctgacc ccaacccggc ctccgcctc ag g ttt acc gag aaa Phe Thr Glu Lys 360	12393
ggg aac ttg gaa gtg ctg ctc ttc acc atc cag agc aag ctt cg <sub>g</sub> gcc Gly Asn Leu Glu Val Leu Leu Phe Thr Ile Gln Ser Lys Leu Arg Ala 365 370 375	12441
aac aac cag aag gtc tac acg ccc cgc gag ggc cg <sub>g</sub> ctc atc tcg gac Asn Asn Gln Lys Val Tyr Thr Pro Arg Glu Gly Arg Leu Ile Ser Asp 380 385 390	12489
atc aac aag gtccgtggct gcccacagggc cacccacccct cagggcagggc Ile Asn Lys 395	12538
cctggcccag atgccttgca cacatcccc aacccggggc catgtcgacc ttcaccaagt	12598
tctactatct gctgccccaa cttgaaaactc gagcactctg cccagctgcc cacactgtgc	12658
cagatgtgat tctccatcc tctcaggcac acggctccct gtccctttgt ctgactctcc	12718
atacgaggtc acataaggaa attatgcccc aagttcctca gcttgtaaa cctgtcccc	12778
actcactatc cttttttta ttttttac aatattgtaa gattactaaa agtaaaaata	12838
aggttcgcac acagtcctgc caacacaaag gaatcagatg gttccactg ttctgtcc	12898
ctccagtcct ttccagaag cttctacatg gcttcacaga ctccagcgc tgccctgtt	12958
cctgtttcca cgcccaggcg tctcaccac tgcattccac ggatccacca gaccgagctt	13018
ccatggact taccatgtgc ctccagcccc actgagccct ccccacccctg tcccttccac	13078
taaccctgtc cccacccat ag gct tgg gag cg <sub>g</sub> ctg gag aag gcg gag cac Ala Trp Glu Arg Leu Glu Lys Ala Glu His 400 405	13130
gag cgt gag ctg gcc ctg cgc acc gag ctc atc cgc cag gag aag ctg Glu Arg Glu Leu Ala Leu Arg Thr Glu Leu Ile Arg Gln Glu Lys Leu 410 415 420	13178
gag cag ctg gcc gcc cgc ttc gac cgc aag gct gcc atg cg <sub>g</sub> gag acc Glu Gln Leu Ala Ala Arg Phe Asp Arg Lys Ala Ala Met Arg Glu Thr 425 430 435	13226
tgg ctc agc gag aac cag cgc ctc gtg tcc cag gtaggacttg aggctcctag Trp Leu Ser Glu Asn Gln Arg Leu Val Ser Gln 440 445 450	13279
gatgcttagg ctggtcagaa ctgggagaga gacagggtgg ataagaagcc ccgcgggtg	13339
gagagacaat gaaacacaaa atccgtcacc tggtaaaaaa ggcctagagg tccgggagcc	13399
agggccaga gggtggaca ggagagaggg ttggtgagac agatggggg tggatagag	13459

## 110.02600201.txt

agggaactta gagccaccgc atggagctgg gatcttgcag gccaaagtgcc ctgcaggaca	13519
gaggagcagg cagaggcact gtcccttggt ccccacacccct cctctccctg ccccccacac	13579
ag gac aac ttt ggg ctg gag ctg gca gct gtc gag gca gca gta cg	13626
Asp Asn Phe Gly Leu Glu Leu Ala Ala Val Glu Ala Ala Val Arg	
455 460 465	
aag cac gaa gcc att gag acg gac atc gtg gcc tac agc ggc cg g	13674
Lys His Glu Ala Ile Glu Thr Asp Ile Val Ala Tyr Ser Gly Arg Val	
470 475 480	
cag gca gtg gac gcc gtg gct gca gag ctg gcc gca gag cgc tac cac	13722
Gln Ala Val Asp Ala Val Ala Ala Glu Leu Ala Ala Glu Arg Tyr His	
485 490 495	
gac atc aag cgc atc gcc gct cg g cag cac aac gtg gca cg ctc tgg	13770
Asp Ile Lys Arg Ile Ala Ala Arg Gln His Asn Val Ala Arg Leu Trp	
500 505 510	
gac ttc ttg cgg cag atg gtg gcc gcc cg g g g cgg ctc ctc ctc	13818
Asp Phe Leu Arg Gln Met Val Ala Ala Arg Arg Glu Arg Leu Leu Leu	
515 520 525	
aac ctg gag ctg cag aag gtg ttc cag gac ctg ctc tac ctc atg gac	13866
Asn Leu Glu Leu Gln Lys Val Phe Gln Asp Leu Leu Tyr Leu Met Asp	
530 535 540 545	
tgg atg gaa gag atg aag gtaccagtga ggctgtctgg gtggggtaag	13914
Trp Met Glu Glu Met Lys	
550	
agtgtatcaag agtcgagggg gccccacagt gggtgcgtcc gcccgtctgc tcggccgatc	13974
tctgtggagt gtgaaccagc acagggccct gtcccccagtt gcagggaaaca aagtaaacag	14034
ggccctgtcc acatggggtt cgtgtccaca tggaggaggc tgatgacaaa acccacggac	14094
ctttccgtaa acaaaaagagc aggggccata aaaccactga tgaaactgga acagggttgt	14154
gtgaggggct ggtcagacaa ggctctgag aaaataccat ttaagccgag accagaatag	14214
tgccaggagc cagcacagag agctctcagg agagaatgtt ccggcagaa ggaacagccg	14274
gcacaggccc tcggcctgga tcagcatggt gtgttcagga tcagaaggaa ggccgggtgg	14334
ctgtgagcat gaggagtggg ggccatggac tgagattgg gagggcacgg ccagccact	14394
agggcattac aggcctggc agggagttt gattttattt aaagtaacat ggatccattt	14454
aagtgattt aaggggccag gcgcagtggc tcatgcctgt aatcccagct ctttgggagg	14514
ccaaggcagg cagatcactt gaggtcagga gttcaagacc agcctggcca acatggtaaa	14574
accccggttc tataaaaata cagaaattaa ccaggtgcgg tggtagcgc tgcctgttagt	14634
cccagctact caggaggctg agacagcaga atcgcttcaa cccaggaggc ggaggttgca	14694
gtgagccaag attgcaccac tgcactccag cctaggtgat agagcaagac tcagtcaaaa	14754
aaaaaaaaaa aagcagtya ggcaagggtgg ctcacgccta taatcctaag actttgggag	14814
gccgaggcgg gcggatcacc tgaggtcagg agtccaagac cagcctggcc aacatggtaaa	14874
aaccccatct ctgctaaaaa tacaaaaatt agctgagtgt ggtggcacac gcctgttaatc	14934
ccagctagtc aggaggctga ggaaggagaa ttgcttggac ctgggaggtg gaggttgcag	14994
tgaactgaga ttgtgctact gcactccagc ctgcactaca ggagcgcagac tccatctcaa	15054

110.02600201.txt													
aataaataaa	taaataaaagt	gattagaagg	tttgagagg	gggagagatt	tgattgtacc							15114	
ttatgttgtt	aaaaaaagtca	cttggctac	tctgaagaag	gcctgtggc	aggcagcatg							15174	
acagcgagac	cgttagaggt	tgctgccatc	ttccaggcgg	gcagcggcta	ggagggtgtga							15234	
ttggggagag	atgtggaagg	aaggctgctg	ggtgctgggg	cggagcggtg	gcggcaggag							15294	
ggtgtcgatg	attctgaggc	tgagaagatt	agtcacagca	aagcttccaa	gagcaggagt							15354	
tagttagaac	attgttccac	agcaggggag	ttcaggataa	ggaggtcggg	gtgccagggc							15414	
agaactgggg	tatctcaagt	atgtgtgagc	aagcaggttt	ctcacctggg	tgcaacacct							15474	
ggggccgtgg	agaagttgga	agaaaaacgc	agccaggtta	gcgcgtctgt	gctcaggcgg							15534	
gcatgtgcac	ttcatgtgtg	cctggcag	ggc cgg	ctg cag	tct cag	gac ctg	Gly Arg	Leu Gln	Ser Gln	Asp Leu		15586	
							555						
ggc agg cac cta gca gga	gtg gag gac	ctg ctg cag	ctg cac gag	ctg Gly	Arg His	Leu Ala Gly	Val Glu Asp	Leu Leu Gln	Leu His Glu	Leu Leu		15634	
560	565	570	575										
gtg gag gca gac atc	gcc gtg cag gcc	gag agg gtg	cgg gcc gtc	agc Val	Glu Ala Asp	Ile Ala Val	Gln Ala Glu	Arg Val Arg Ala	Val Ser			15682	
	580	585	590										
gcc tct gcc ctg cgc	ttc tgc aac cca	ggg aaa g	gtgagaagtc									15726	
Ala Ser Ala Leu Arg	Phe Cys Asn Pro	Gly Lys											
595	600												
agcgaaggca	ctggagaggg	agggctggg	aaggagcaca	tcaagagccg	aggtggaagg							15786	
gttggaaac	ctggggacgg	gaaagatggt	ggccaaacga	tggtaactga	gctaaagcca							15846	
gggagggaaag	tccaagagag	tgtggtgca	ggagcagcc	ggagggctgg	gttaaggctc							15906	
tgaccctctc	ctgtgacttt	ctcag ag tat	aga cct tgc	gac ccg cag	ctg Glu	Tyr Arg Pro	Cys Asp Pro	Gln Leu				15957	
		605	610										
gtg tcg gag cgg	gtg gcc aag cta	gag cag	agc tat	gag gca	ctg ctg	Val Ser Glu Arg	Val Ala Lys	Glu Gln	Ser Tyr	Glu Ala	Leu Cys	16005	
	615	620	625										
gag ttg gca gcg	gcf cgf cgg	gcc cgf ctg	gag gaa	tca cgf	cgf ctc	Glu Leu Ala	Ala Ala Arg	Ala Arg	Glu Ser	Arg Arg	Leu	16053	
	630	635	640										
tgg cgt ttc ctc	tgg gag gtg	ggt gaa	gct gag	gcc tgg	gtg cgg	gag Trp Arg	Phe Leu Trp	Glu Val	Gly Ala	Glu Trp	Val Arg	Glut	16101
	645	650	655										
cag cag cac ctc	ctg gcc tca	gcc gac acg	ggc cga	gac ctg	acc ggt	Gln Gln His	Leu Leu Ala	Ser Ala	Asp Thr	Gly Arg	Asp Leu	Thr Gly	16149
	660	665	670										
gcc ctc cgc ctg	ctc aac aag	cac aca	gcc ctg	cgf ggc	gag atg	Ala Leu Arg	Leu Asn Lys	His Thr	Ala Leu	Arg Gly	Glu Met	Ser	16197
	680	685	690										
ggc cgg ctg	ggg ccc ctg	aag ctc acc	ctg gag	cag ggc	gag atg	Gly Arg	Leu Gly	Pro Leu	Lys Leu	Thr Leu	Glu Gln	Gln Leu	16245
	695	700	705										
gtg gcc gag	ggt cac cct	ggg gca	agc cag	gcc tct	gcc cgt	gca gct	Val Ala Glu	Gly His Pro	Gly Ala Ser	Gln Ala Ser	Ala Arg	Ala Ala	16293
	710	715	720										
gaa ctc caa qcc	caq taq qaq	caq cta	qaq qcc	cta acc	acc qaq	qaq cat						16341	

## 110.02600201.txt

Glu	Leu	Gln	Ala	Gln	Trp	Glu	Arg	Leu	Glu	Ala	Leu	Ala	Glu	Glu	Arg		
725						730										16389	
gcc	cag	cg	ctg	gcc	caa	gcc	gcc	agc	ctc	tac	cag	ttc	cag	gcc	gat		
Ala	Gln	Arg	Leu	Ala	Gln	Ala	Ala	Ser	Leu	Tyr	Gln	Phe	Gln	Ala	Asp		
740					745					750					755		
gca	aac	gac	atg	gag	gcc	tgg	ttg	gtt	gac	gca	ctg	cgc	ctg	gtg	tcc		
Ala	Asn	Asp	Met	Glu	Ala	Trp	Leu	Val	Asp	Ala	Leu	Arg	Leu	Val	Ser		
760					765					770					775		
agc	ccc	gag	ctg	ggg	cac	gac	gag	ttc	tcc	acg	cag	gct	cta	gcc	agg		
Ser	Pro	Glu	Leu	Gly	His	Asp	Glu	Phe	Ser	Thr	Gln	Ala	Leu	Ala	Arg		
775					780					785					790		
cag	cat	cg	gcc	ctg	gag	gag	gag	att	cga	agc	cac	cg	cca	acc	ctg		
Gln	His	Arg	Ala	Leu	Glu	Glu	Glu	Ile	Arg	Ser	His	Arg	Pro	Thr	Leu		
790					795					800					805		
gac	gcc	ttg	agg	gaa	cag	gca	gca	gcc	ctg	ccc	ccc	aca	ctg	agc	cgc		
Asp	Ala	Leu	Arg	Glu	Gln	Ala	Ala	Ala	Leu	Pro	Pro	Pro	Thr	Leu	Ser		
805					810					815					820		
acg	ccc	gag	gtg	cag	rgc	cg	gtg	ccc	acc	ctg	gag	cg	cac	tac	gag		
Thr	Pro	Glu	Val	Gln	Xaa	Arg	Val	Pro	Thr	Leu	Glu	Arg	His	Tyr	Glu		
820					825					830					835		
gag	ctg	cag	gcc	cg	gca	gg	gag	cga	g	cg	gg	cc	ttg	gag	gca	gcc	
Glu	Leu	Gln	Ala	Arg	Ala	Gly	Glu	Arg	Ala	Arg	Ala	Leu	Glu	Ala	Ala		
840					845					850					855		
ctg	g	ctc	ta	cc	atg	ctc	agc	gag	gcc	gg	gcc	tgt	gga	ctc	tgg		
Leu	Ala	Leu	Tyr	Thr	Met	Leu	Ser	Glu	Ala	Gly	Ala	Cys	Gly	Leu	Trp		
855					860					865					870		
gtg	gag	gag	aag	gag	cag	tgg	ctc	aac	ggg	ctg	gcc	ctg	cct	gaa	cgc		
Val	Glu	Glu	Lys	Glu	Gln	Trp	Leu	Asn	Gly	Leu	Ala	Leu	Pro	Glu	Arg		
870					875					880					885		
ctg	gag	gac	ctg	gag	gtc	gtg	cag	cag	ag	gtaggccc	c	tc	tc	tc	tc		
Leu	Glu	Asp	Leu	Glu	Val	Val	Gln	Gln	Arg								
885					890											16822	
gtgggaccag	ccttgggagg	tgggggtggk	ggggccagga	tgtgggttgt	gagtccctcc											16882	
ataaaacttcc	tgcctcaccc	ctttgagtct	taatggttgt	ccatttctag	ttttaacaaa											16942	
aaatgttaac	catactcaca	agtagagaca	tttcaccaca	aaccctctta	cacccgtcac											17002	
ccagattcag	tcattgccaa	catcttgctc	tatttggttc	tctgcccttt	taaaaagcaa											17062	
atccccaga	ttagcacatc	cctcctacct	tcttcagggt	gtgcccttta	agaaaatacaa											17122	
ggaagcggcc	gggcacggtg	gctcacacct	gtaatcccag	cactttggga	ggccgaggca											17182	
ggcggatcac	gaggtcagga	ggtcgagacc	atcctggcta	acatggtgaa	acgctgtctc											17242	
aactaaaaat	acaaaaatt	agccggcac	agtggggggc	gcctgttagtc	ccagctactc											17302	
gggaggctga	ggcaggagac	tggcatgaac	ctgggaggcg	gagcttgcag	tgagccgaga											17362	
tcgtgccact	gctctccagc	cttggcgaca	gagcgagact	gtctcaaaaa	aaaaaaaaaa											17422	
aaaaagaaaa	gaaatacag	gaagcattct	tacctacaat	atattat	tttcaattgtct											17482	
gtgtggactt	aaaatagctg	aggcttgaa	gttataaacc	tttttagacag	agacactaaa											17542	
atagtttat	ttacagaata	aatctcaa	tcatttattc	aataataat	cttaaaaatt											17602	
tttttataga	accacaatac	cattatcaca	tagaaca	tttagaattt	cttttttttt											17662	

110.02600201.txt

tgagatggag tctcaactcg tcacccaggc tggagcgcag tggycgatc tsggctcact	17722
gcatgctctg cctccgggt tccagcgatt ctccgcctc agcctccaa gtagctggga	17782
ctacaggcac ccaccacat gcccagctaa tttttgtat ttttttatta gtagagacag	17842
ggtttcacca tgtagcgag gatggtctcg atctcctgac ctcgtatcc acccggcttg	17902
gcctccaaa gtactggat tacaggcatg agccactgcg cygggtcagg aatttcttga	17962
tatgatctat gatcaatgtg taatcaaatt tcctggtgtt taaaaaaatr taattaggtc	18022
aggcctggtg gtcacacact gtaatcccag cactttggga ggccaaggca ggagtatcac	18082
ttgattccaa gagttcataa ccagcctggg caacatagtg agactccacc tctacaaaat	18142
tttttttta attagccagg catggtggca ggcgcctgta gtcccagcta ctggggaggc	18202
tgaggcagga ggattgcttg agccagagg tcgaggctgc agttagctga gatcgtgcc	18262
ctgcactcca acctgggtga cagagtaaga ccctgtctca aaaaaataaa taaaataaa	18322
aatgttaatta aagattttt ttagtttca aggtatccta atgttagaggt taacaaactg	18382
caacccttgg gccaaatcca gcctgctgcc tgtttctgta aataaagttt ttttggttt	18442
ttttttttt tttttttttt gagacggagt cttgctctgt cccccaggct ggagtgcagk	18502
ggcacaatct tggctcactg caagctccgc cgccgggtt cacgccattc tcctgcctca	18562
gcctcccaag tagctggac tacaggcacc cgccaccaca cctggctaat ttttgtatt	18622
tttagtagag acgaggtttc mcccgtttag ccaggatggt ctcgatctcc tgaccttgc	18682
atccacccat ctcagcctcc caaagtgctg ggattacagg cgtgagccac catgcccgg	18742
ctgtaaataa agtttactg gaacagagct acactcgccc tttgcgtat tgcccgtggc	18802
tgcttcagg ctacagcagc agggctgaat agttgccca gcagctgcac ggctcacgaa	18862
gcctataata ttaccatca agcccttcac agaaaacatt tgcagaccct gcccatttag	18922
aacatcaaata gtcggccagg cgaggtggct cacgcctgta atcccagcac tttggggaggc	18982
caaggcgggt ggagcatgag gtcaggagtt caagaccagc ctggccaaga tagtcaaaca	19042
cccgctcaca ccaaaaatac aaaaattagc taggcttggt ggcgggcacc tgtaatccca	19102
gctattcggg aagcggaggc agaaaattgc ttgaaccgg gaggcggagg ttgcattgag	19162
tcgagattgc accactacac tccagcctgg gtgacagagc gagactccat ctcaaaaaca	19222
aaacataaaa tgtccctgga agaagttatc cttcacattt tcatcaactca gacaaatatt	19282
tttggttatt acttcatttt ctaacacata ttagtggat gagaccaga taggcagcta	19342
aggggaagt cccaggaggc tgagcagctc tgataaaaggg ttggcccat gccagggtct	19402
agtttgactt gtgtgtctgt tcctgacctt taccttcac catctttga gcatttctag	19462
cgatgacagt tattctgctt gtttgggg agttgtgggt tccttccttg atgcacagat	19522
actatgaacc tcataggctg agaggcagag agactcctgc agataagcct cctggtaag	19582
gagacaccct cctgggtgtt taagagaggt ggacttaggt ggactggca catcgctgcc	19642
tcccacaatt cacgcctggc tcttccaccc tctcag g ttc gag acc ctg gag cct Phe Glu Thr Leu Glu Pro 895	19697
gaa atg aac acc ctt gca gca caa atc acc gcg gtg aat gac att gcc	19745

## 110.02600201.txt

Glu Met Asn Thr Leu Ala Ala Gln Ile Thr Ala Val Asn Asp Ile Ala  
900 905 910 915

gag cag tta ctg aag gcc aac ccc cca ggc aaa gac cgc att gtc aac 19793  
Glu Gln Leu Leu Lys Ala Asn Pro Pro Gly Lys Asp Arg Ile Val Asn  
920 925 930

acc cag gag cag ctc aac cac ag gtgggttgg raggcagga ccrggaaact 19846  
Thr Gln Glu Gln Leu Asn His Arg  
935

gacagaaaaa tgaagcaatg gggatggcag tgagaggcag gttttgttagg gcctggaagg 19906

gtggctacaa aggaggaagc aaaccagtct ggaatatgyt ggggaaggaa aaaggatgaa 19966

agagatgaga gagggggtca aggtggtctg aagcrgctgc tggccagaaa gggaggagtg 20026

aaggggagct accaagagag agagaagcag ggaagaagct tccaaacagg cctggccagg 20086

gcaggaagct gaaccttccc cctgctctca g g tgg cag cag ttt cgg cgt ctg 20139  
Trp Gln Gln Phe Arg Arg Leu  
940 945

gca gac ggc aag aag gca gct ctc acc tca gcc ctg agc atc cag aac 20187  
Ala Asp Gly Lys Lys Ala Ala Leu Thr Ser Ala Leu Ser Ile Gln Asn  
950 955 960

tac cac tta gag tgc acg gag acc cag gcc tgg atg aga gag aag acc 20235  
Tyr His Leu Glu Cys Thr Glu Thr Gln Ala Trp Met Arg Glu Lys Thr  
965 970 975

aaa gtc atc gag tcc acc cag ggc cta ggc aac gat ctg gct ggg gtg 20283  
Lys Val Ile Glu Ser Thr Gln Gly Leu Gly Asn Asp Leu Ala Gly Val  
980 985 990

ctg gcc ctg cag cgc aag ctg gcc ggc acg gag cgg gac ctg gag 20328  
Leu Ala Leu Gln Arg Lys Leu Ala Gly Thr Glu Arg Asp Leu Glu  
995 1000 1005

gcc atc gcc gcc cgg gtg ggc gaa ctg act cga gag gca aat gcc 20373  
Ala Ile Ala Ala Arg Val Gly Glu Leu Thr Arg Glu Ala Asn Ala  
1010 1015 1020

ctg gct gcc ggc cat ccc gct cag gca gyg gcc atc aac gcc cgg 20418  
Leu Ala Ala Gly His Pro Ala Gln Ala Xaa Ala Ile Asn Ala Arg  
1025 1030 1035

ctg aga gag gtg cag acc ggc tgg gag gac ctc agg gcc acc atg 20463  
Leu Arg Glu Val Gln Thr Gly Trp Glu Asp Leu Arg Ala Thr Met  
1040 1045 1050

cgg cgt cga gaa gag tcg ctg ggg gag gcg cgg cgg ctg cag gac 20508  
Arg Arg Arg Glu Glu Ser 1060 Leu Gly Glu Ala Arg Arg Leu Gln Asp  
1055 1065

ttc ttg cgc agc ttg gat gac ttc cag gcc tgg cta ggc cgc act 20553  
Phe Leu Arg Ser Leu Asp Asp Phe Gln Ala Trp Leu Gly Arg Thr  
1070 1075 1080

cag act gct gtg gcc tct gaa gaa ggg ccg gcc acc ctg cct gag 20598  
Gln Thr Ala Val Ala Ser Glu Glu Gly Pro Ala Thr Leu Pro Glu  
1085 1090 1095

gca gag gcc ctc ctg gcc caa cat gca gcc ctg cgg gga gag gtg 20643  
Ala Glu Ala Leu Leu Ala Gln His Ala Ala Leu Arg Gly Glu Val  
1100 1105 1110

gag cgg gcc cag agc gag tat agc cgg ctg cga gcc ctg ggc gag 20688  
Glu Arg Ala Gln Ser Glu Tyr Ser Arg Leu Arg Ala Leu Gly Glu  
1115 1120 1125

## 110.02600201.txt

gag	gtg	acc	cgg	gac	cag	gct	gac	ccc	cag	tgc	ctc	ttc	cta	cga	20733
Glu	Val	Thr	Arg	Asp	Gln	Ala	Asp	Pro	Gln	Cys	Leu	Phe	Leu	Arg	
1130					1135					1140					
cag	cga	ctg	gag	gcc	ctg	gga	act	ggc	tgg	gag	gag	ctg	ggc	cga	20778
Gln	Arg	Leu	Glu	Ala	Leu	Gly	Thr	Gly	Trp	Glu	Glu	Leu	Gly	Arg	
1145					1150				1155						
atg	tgg	gag	agc	cgg	caa	ggg	cgc	ctg	gcc	cag	gcc	cac	ggc	ttc	20823
Met	Trp	Glu	Ser	Arg	Gln	Gly	Arg	Leu	Ala	Gln	Ala	His	Gly	Phe	
1160					1165				1170						
cag	gga	tcc	ctg	cgg	gat	gct	cgt	cag	gct	gag	ggc	gtg	ctc	agc	20868
Gln	Gly	Phe	Leu	Arg	Asp	Ala	Arg	Gln	Ala	Glu	Gly	Val	Leu	Ser	
1175					1180				1185						
agc	cag	gtgaaagtcc	agggcaaagt	cccaaggcagg	aggaagagca	aagttagggac									20924
Ser	Gln														
1190															
ccggggaaat	gtgaaggagc	aggatggca	ggaaggacat	gctagcaaaa	tggggcagcr										20984
cagtggttca	cacctgaaat	tccagcactt	tgggaggcca	aagtaggagg	atcacttgag										21044
gctgagaata	tccagaccaa	cctggcaac	atggcaagac	cttgtctcta	caaaaaaatt										21104
tttttaagaa	aatagaagaa	ttttttaaaa	agaaaaatgg	gagccagaca	ggatggctca										21164
cacttgtact	cccagtattt	tgggaggccg	aggcaggaga	ataacttgag	ctcaggagtt										21224
tgagaccagc	ctgggcaaca	tagtgagacc	cccatctcta	tgaaaaaaaaa	aaattaactg										21284
gacatggtgg	tgcattgcctg	tagtccagc	tactggggag	gctgaggctg	gtggatcact										21344
ggagccagga	gtttgaggct	gcagtgagct	atgatatgcc	actgcactcc	aacctgggcc										21404
acagaatgaa	accctctctc	aaaaaaaaa	aaaagaaaaa	agaaggaaaa	atgagaatga										21464
aaaagacgtg	aatataattt	actaaaactg	actttagaag	aaatagaaag	cccagtttgt										21524
catgtaactg	ttaaatggaa	tcaggagcta	gaagtgagat	agaacaggat	ttgggctggg										21584
gaatggaagg	tccttcccac	ccagcttccc	tgtgactttc	tgaggctccc	atgctggctg										21644
gcagcctccc	tgtcttcaga	gctctctgg	ctaccctccc	tgtctgcttg	ttggtcccta										21704
cctctcagat	ttgccccctgg	gtgggtccct	cctagggggg	tgaattgtgc	tggggaaaat										21764
gagctgaatg	tcatccctcc	cacacag	gaa tat gtt	ctg tct cac acg gag											21815
			Glu Tyr Val	Leu Ser His Thr Glu											
1195															
atg	cca	ggg	aca	ctc	cag	gct	gct	gat	gct	gcc	att	aaa	aaa	ctg	21860
Met	Pro	Gly	Thr	Leu	Gln	Ala	Ala	Asp	Ala	Ala	Ile	Lys	Lys	Leu	
1200					1205					1210					
gag	gac	tcc	atg	agc	acc	atg	gac	gcc	aat	ggg	gaa	cg	atc	cac	21905
Glu	Asp	Phe	Met	Ser	Thr	Met	Asp	Ala	Asn	Gly	Glu	Arg	Ile	His	
1215					1220					1225					
ggg	ctc	ctg	gag	gct	ggc	cgc	cag	ctg	gta	tct	gaa	ggc	aac	atc	21950
Gly	Leu	Leu	Glu	Ala	Gly	Arg	Gln	Leu	Val	Ser	Glu	Gly	Asn	Ile	
1230					1235					1240					
cac	gcc	gac	aag	att	cgg	gaa	aag	gca	gac	tcc	att	gag	agg	ag	21994
His	Ala	Asp	Lys	Ile	Arg	Glu	Lys	Ala	Asp	Ser	Ile	Glu	Arg	Arg	
1245					1250					1255					
gtctgatgag	gacagtccat	gaatttagggt	tcccaggggg	gaatcgagaga	aacagggtga										22054
cctcaaagat	aaacgtggca	cagggaaaccc	acagatgggg	caggagctga	cagagaagta										22114

## 110.02600201.txt

gaggggaaga actaagtgg tggagaggc tggagattc caccacaac cagggctaaa 22174  
 aggaagtcag gattcctgg tagcctcatt gtgctcccg aaggcgttat tcccctagaa 22234  
 gaaatggagg cccctaggtt agccaaagg tcacaatcct ttcacagcaa attccagagt 22294  
 ttcacaagag ggtgtcggttc cag g cac aag aag aat caa gac gca gcg cag 22345  
                   His Lys Lys Asn Gln Asp Ala Ala Gln  
                   1260                   1265  
 caa ttt ctg ggc cgt ctt cg gac aac cg gag cag cag cat ttc 22390  
   Gln Phe Leu Gly Arg Leu Arg Asp Asn Arg Glu Gln Gln His Phe  
   1270               1275               1280  
 ctg caa gat tgt cac gag gtgaggctcc ctggggcccc gggatattcc 22438  
   Leu Gln Asp Cys His Glu  
   1285  
 ctagccatcc ctttctcacc ttgagcctag aataagtcca gcacaaggta ccggagactg 22498  
 tgagccctt catggcttct tcccaaggcg cccacttctc ctggctcaact gtggccctgc 22558  
 ttttatgccc ccctctcccc tcccttgaa atgtccctgt tttatgtttg gtccag 22614  
 ctg aag ctc tgg atc gac gag aag atg ctg aca gcc cag gac gtg 22659  
   Leu Lys Leu Trp Ile Asp Glu Lys Met Leu Thr Ala Gln Asp Val  
   1290               1295               1300  
 tcc tat gac gag gcc cgc aac ctg cat act aag tgg cag aag cac 22704  
   Ser Tyr Asp Glu Ala Arg Asn Leu His Thr Lys Trp Gln Lys His  
   1305               1310               1315  
 cag gca ttc atg gcc gag ctg gct gcc aac aaa gac tgg ctg gac 22749  
   Gln Ala Phe Met Ala Glu Leu Ala Ala Asn Lys Asp Trp Leu Asp  
   1320               1325               1330  
 aag gtg gac aag gtgagcagtg ctgtggggc tgcccttg cagagtcccc 22801  
   Lys Val Asp Lys  
   1335  
 catggtacgg gggagggcct ggctccagga cgtggtttt gtcatggta gagattgtgg 22861  
 ggttctggc ccacagctcc atggtgaaa gttatggcct ctggaaaca aacgtcttg 22921  
 ttggagacaa agagtggac agtgagtcca gaaccttgat aaagtacaaa ttagcaggt 22981  
 gcgaaactgc agcctggatt taacgccata actgttcctg agttcaaagc aggatggc 23041  
 tcctcagcag gctctgcaaa ctttgtctgg aaaggcctg ggytggatgg tcaatatctg 23101  
 aggcttgca gtccaaacag tttctactac aactactcat ctctgtcggtt gttagtgc 23161  
 agcagcctca gagaataart gaatgaatga acatggctgt ttcccaataa aactttat 23221  
 gtaaaaaacag gcagtggcc agattacccc gtgggctgta tgttagtctgc caatcctt 23281  
 tttcagtc tgccgtccct ggttcttggt ctcccaccca gctgttcctc aagagtcaag 23341  
 gacagataaa gctttatatc tgctcagcaa gtatccctac agaacatggt acgtgctca 23401  
 cagtctgcga ggcactgcag ggtctacgaa aagacacttt cctttctca agggcatgat 23461  
 ggttggaca aaaggagacc agtgtgctt caagtgttc cattcggtca gcaaataatt 23521  
 acttggtgct actatatcca aagcattgta tttaggcccgt cgtggggtcc gaaaggatc 23581  
 ccatggtagg ctccaccytc aaggaacaac actgttagcca ggatatgagg catatatgca 23641  
 ggttagctaa tcgaatccag agcagaatct gatgagtgtt gttggatgttccccc 23701

110.02600201.txt

gcctgggaa taaggagaga ggagatatgg aaattctaga agcagcaaca tccagggcag 23761  
 tccatggagg atgagtagaa tttggacaag tggagatatg gagaaaaaga tgtttccagg 23821  
 tagagaaaac caaacagaaa accagtggta ggcagtcct gtttggagc aagggggta 23881  
 gtgagtccag cagcagcaca gaatgtgtc agggccccat ggaagaagag gtggagaga 23941  
 agccccggga agtcctgacc accacatgcg tgtgcttcca agtgttgatt gtgctcagg 24001  
 ttgtgaagct acccaggcag tgctagcctt tcatgctctc catcctgaca gccacagcca 24061  
 cctgcaaaat gcctgaagg cagtgctgg gtttgttgc ccctagaagc agaccctgaa 24121  
 acagtattc tagggaggtg tccccggaa tgatcagtag ggggtggag ccatgagaag 24181  
 ggacagaagt catgccagg tccccatgg gcatcaggmg cccaaccctc atggggcatt 24241  
 ctgcgggggt gtagatggaa cacagaacac agcaactcta cagccacggc acaaggatgc 24301  
 tgtgtttta tccatcagct ccctgtctgt cactggttga gggctgcttc tgggcttgg 24361  
 aactctgtgg ctcttccaat ttgccttcca cacagttcaa gagagcacac ttccctcacc 24421  
 accccagaca cagaaatcct caagcagaaa agctgaaatgg gggaaacca ggtgggactc 24481  
 aacttgctat cactgaagt aagagtaatg agttgagaaa tacaaaatga tggcaaataa 24541  
 ctcaagtcaa tgtgcagtgt ggtgccagaa accacaggca ccaaggaacc caaactaaag 24601  
 agagtcaag ggcttcctcg ttgaaatgct gttttagctg ggtcttgaag gaggagtgg 24661  
 gttcatcagc agtgaggaa cagcagtttta gaaggattt taaacctccg tgaagagtgt 24721  
 gtcgataaga acacacaagg acttggattt attgcccaca cttcaggtct ctactgccc 24781  
 taaggccctt ggtctgctct cagcctctaa agttcaggct aatttggca ctttgacatt 24841  
 ttctgtcctg cttccag gaa ggg cga gag ctc acc ctt gag aag cca gag 24892  
 Glu Gly Arg Glu Leu Thr Leu Glu Lys Pro Glu  
 1340 1345

ctg aaa gcc ctg gtg tcg gag aag ctg aga gac ctg cac agg cgc 24937  
 Leu Lys Ala Leu Val Ser Glu Lys Leu Arg Asp Leu His Arg Arg  
 1350 1355 1360

tgg gac gag ctg gag acc acc acc caa gcc aag gcc cgc agc ctc 24982  
 Trp Asp Glu Leu Glu Thr Thr Thr Gln Ala Lys Ala Arg Ser Leu  
 1365 1370 1375

ttt gat gcc aac cga gct gag ctg ttt gcc cag agc tgc tgt gcc 25027  
 Phe Asp Ala Asn Arg Ala Glu Leu Phe Ala Gln Ser Cys Cys Ala  
 1380 1385 1390

ctg gag agc tgg ctg gag agc ctg cag gcc cag ctg cac tcg gat 25072  
 Leu Glu Ser Trp Leu Glu Ser Leu Gln Ala Gln Leu His Ser Asp  
 1395 1400 1405

gac tac ggc aag gac ctc acc agc gtc aac atc ctg ctc aag aag 25117  
 Asp Tyr Gly Lys Asp Leu Thr Ser Val Asn Ile Leu Leu Lys Lys  
 1410 1415 1420

cag cag gtgtgctgtg ggcctttgat gggatggtg aacagcagaa gaaaggggct 25173  
 Gln Gln  
 1425

gcagctttca agatttggga ggccagctga ggcctggcag ataacacctt cactagcatt 25233  
 tcccagagtc atttcttgg gcagccagta tcagaatctc tagggatatg tcgcagtcag 25293  
 gtttagagtt aggagacaga ccccttaata atttgacgat gagaacttca gtaaaaaagaa 25353

## 110.02600201.txt

atattaatta ggccaggcac agtggctcat gcctgtata tcagcacttt gggaggccaa 25413  
ggcaggcagg tcgcttgagg tcaggagttc aagaccagcc tggccaacat ggtgaaaccc 25473  
cgtctctact aaaaatacaa aaattagcca ggcgtggtgg cgcacgcctg ttatcccagc 25533  
tacttgggag gctgagggct gagggaggag gatagcttga acccaggaga cagaggttgc 25593  
agtgagccgt gatcatgcca tcgcactcca gcctggcaa gcctctgtct caaaaaaaaa 25653  
aacagaagta taaatttagta aaaggtggta aagagaactc tatagcaggg gtcagaaaaac 25713  
ttttttcagt aaagcaccag atagtaaatt tttggggctt catgggtcgt aaggctatg 25773  
tcacagccct tcaactctgc ccttgtaagg aaaagcagct gcggattata tgtaaaataa 25833  
ggggtatggc caggtgttgt ggctcacaca tgtaatccca acactttggg aggccgaggt 25893  
ggccggatca cttgagtcca ggagtttga gactgcctg ggcaacatgg caaaaacccca 25953  
tctctacaaa aactacaaaa attagctggg cgtggtgttg gtgtgtgcct gtagtcccag 26013  
ctactcagga ggctgaggtg agaagattgt ttgagcctgg agagtcgagg ctgcagtggg 26073  
ctgtgatcgt gccactgcac tgcagcctgg gtggcagagc gagacttgc tcaaaaaatt 26133  
ttttaaaaag gggtcagct gtggccaat aaaaactta taaggctagg cagtggctg 26193  
gacttgctgg tccctgctct gaagaatgca gaatggaaag tgttagggagc agctgacacc 26253  
tctgggttg aggagtcccc caggaaggag caagatggct gagtctcaga cttccttgg 26313  
gagggcggag gcccactgaa tggcctggaa atttgctggg gtatcacagg ccagagcatg 26373  
tctgaagcca gcaagccaag gctggccagc aggaagtggc cactagagtg cagcaaaact 26433  
tgccagagag taagagccac cggatctctt ggcactgtg gacagtttg ctgtaggagg 26493  
aagaaaagca ctttggaaacc gtggaatgcc ccacatctgc taggcattgt ggtgtccctc 26553  
cagccccctc tgctgacaag gcctagattt tgcctgctgg caaaggagga atgtttccag 26613  
ggcccgatcc cagtgctca aagcagggcc acgatggatt gggatggag agacaatgaa 26673  
ttgataatgg gcacagttt cacactaccc tggattttt aatttcaatc ttttagattc 26733  
gcgggtgtgg gtgcaggattt gttataagtt atattgtgt atgctgaagt ttggggtgta 26793  
attgaaccca tcacccatgt agtgagcaca gtacccaata gggattttt caacccctgc 26853  
ccttggattc ttgtgcatat aaagttcgag atccactgat ccagcccatg agggatgct 26913  
ctccctgggt gagctgcagt gatgagcagc cacagtggc tggatagaga ctggttggtg 26973  
tctctgctgg gcatgaaggc ggaaatgcag agctaacatg tccttcctcc tggcggttg 27033  
cag atg ctg gaa tgg gag atg gct gtg aga gag aag gag gtg gag 27078  
Met Leu Glu Trp Glu Met Ala Val Arg Glu Lys Glu Val Glu  
1430 1435 1440  
gca atc cag gcc cag gyc aaa gca ctg gcc cag gag gac cag ggt 27123  
Ala Ile Gin Ala Gin Xaa Lys Ala Leu Ala Gln Glu Asp Gln Gly  
1445 1450 1455  
gca ggg gag gtg gag aga acc tcg agg gcc gtg gag gag aag ttc 27168  
Ala Gly Glu Val Glu Arg Thr Ser Arg Ala Val Glu Glu Lys Phe  
1460 1465 1470  
agg gcc ttg tgc cag ccc atg cgg gaa cgc tgc cgg cgc ctg cag 27213  
Arg Ala Leu Cys Gin Pro Met Arg Glu Arg Cys Arg Arg Leu Gln

110.02600201.txt

1475	1480	1485	
gct tct cgc gag cag cac cag ttc cac cgc gat gtg gaa gat gag Ala Ser Arg Glu Gln His Gln Phe His Arg Asp Val Glu Asp Glu 1490	1495	1500	27258
att gtgagtcact gggccaagg acggcaagct gcccccagcc atgtggttct Ile			27311
ccagcctccc tcctggatgc cagggagatg ccagcagggc tctattccct cttctcttg gcattgacca tctccctat agggagactt ggagatgcct cccagaacca gagatgactg ttccccacac acagggcggt agcccaaggt gtcccaactc ccactaatca gtccctgctg cttgccttgc cctctggcc tccactgacc ccctttcct cttccag ttg tgg gtg Leu Trp Val			27371 27431 27491 27547
aca gag cgg ctg ccc atg gcc agc tcc atg gag cat ggc aag gac Thr Glu Arg Leu Pro Met Ala Ser Ser Met Glu His Gly Lys Asp 1505	1510	1515	27592
ctg ccc agc gtc cag ctt ctc atg aag aaa aac cag gtgaggcaga Leu Pro Ser Val Gln Leu Leu Met Lys Lys Asn Gln 1520	1525	1530	27638
ggctgaaggc aaaagagaag ttcccaggag cctgcccaga cttggacgtg tttttctta agaccaggcc cccctcgatg ctgagtgtaa ccctggactt cagtggtgtt ctttcctcac cttgggaggg tgggtcctct cagtaggaga tagtgggggc tggcggctg acgggtgtta ccatcgcacc cccaggctag gagagggagc agaggaccca ggaaggagag aggcacaggg gtgaagggtg gtctccggga gcacgtgggg ctggggcagg actttcagca ttttctttc tgtggccatg ggcag acc ctg cag aaa gag att cag ggc cat gag ccc Thr Leu Gln Lys Glu Ile Gln Gly His Glu Pro 1535	1540	27698 27758 27818 27878 27938 27986	
cgg atc gcg gac ctg agg gag cg <sup>g</sup> cag cgt gct cta ggt gca gca Arg Ile Ala Asp Leu Arg Glu Arg Gln Arg Ala Leu Glu Ala Ala 1545	1550	1555	28031
gca gca ggt cca gag ctg gct gag ctg cag gaa atg tgg aaa cgc Ala Ala Gly Pro Glu Leu Ala Glu Leu Gln Glu Met Trp Lys Arg 1560	1565	1570	28076
ctg ggc cac gag ctg gaa ctt cga ggg aag cga ctg gag gat gcc Leu Gly His Glu Leu Glu Leu Arg Gly Lys Arg Leu Glu Asp Ala 1575	1580	1585	28121
ctg cga gcc cag cag ttc tac cgc gat gcc gcc gag gcg gag gcc Leu Arg Ala Gln Gln Phe Tyr Arg Asp Ala Ala Glu Ala Glu Ala 1590	1595	1600	28166
tgg atg ggc gag cag gaa tta cac atg atg ggc cag gag aag gcc Trp Met Gly Glu Gln Glu Leu His Met Met Gly Gln Glu Lys Ala 1605	1610	1615	28211
aag gtgagggcca ggacagagcc cagtgtatgt gaccagttct gccctccct Lys			28264
gacctgatgc tggatgccac tgtcccttcc cccag gat gag ctg agt gcc cag Asp Glu Leu Ser Ala Gln	1620		28317
gca gag gtg aag aag cac cag gtg ctr gag caa gcc ctg gcc gac			28362

## 110.02600201.txt

Ala	Glu	Val	Lys	Lys	His	Gln	Val	Leu	Glu	Gln	Ala	Leu	Ala	Asp	
1625						1630				1635					
tac	gcg	cag	acc	atc	cac	cag	ctg	gcg	gcc	agc	agc	cag	gac	atg	28407
Tyr	Ala	Gln	Thr	Ile	His	Gln	Leu	Ala	Ala	Ser	Ser	Gln	Asp	Met	
1640						1645				1650					
att	gac	cac	gag	cac	cca	gag	ag	gtgggtgcag	yggcagcccg						28450
Ile	Asp	His	Glu	His	Pro	Glu	Ser								
1655						1660									
gcccagcctg	ggggtggagc	cggctgcagg	aacaggaagg	tgcagggaaat	gtggagcctt										28510
cagtgtgtg	tgcacggagc	cttctagaaa	gctggaacac	agggtggcg	agctgtttgg										28570
agactcagag	ggacagggct	ccacagaaca	gaccggaggt	cagagctaca	cccctaagtc										28630
ccacagtgcc	tccctcaactc	ttcttgcag	c act	cg	ata	tcc	atc	cg	caa						28681
			Thr	Arg	Ile	Ser	Ile	Arg	Gln						
						1665									
gcc	cag	gtg	gac	aag	ctg	tat	gcc	ggc	ctg	aag	gag	ctg	gct	gga	28726
Ala	Gln	Val	Asp	Lys	Leu	Tyr	Ala	Gly	Leu	Lys	Glu	Leu	Ala	Gly	
1670						1675				1680					
gag	cgg	cgg	gag	cgc	ctg	cag	gag	cac	ctc	cgg	ctg	tgc	cag	ctc	28771
Glu	Arg	Arg	Glu	Arg	Leu	Gln	Glu	His	Leu	Arg	Leu	Cys	Gln	Leu	
1685						1690				1695					
cgc	cgc	gag	ctg	gat	gac	ctg	gaa	cag	tgg	atc	cag	gag	cgc	gag	28816
Arg	Arg	Glu	Leu	Asp	Asp	Leu	Glu	Gln	Trp	Ile	Gln	Glu	Arg	Glu	
1700						1705				1710					
gtg	gtg	gcf	gcc	tcc	cac	gag	ctg	ggc	cag	gac	tac	gag	cat	gtg	28861
Val	Val	Ala	Ala	Ser	His	Glu	Leu	Gly	Gln	Asp	Tyr	Glu	His	Val	
1715						1720				1725					
act	gtgagtgtag	ggagggcacc	cagctcagat	caaccgtggg	aagagtggag										28914
Thr															
1730															
gaccacagg	gagacttagga	cctagtccca	ggcagagcac	tgaggggcta	agggcaaga										28974
ccaggctgag	caggcactgt	cctcctggtt	ttgaggtata	tgatttgaag	aggccggca										29034
taggctcaca	cctgtaatcc	tagcacctt	ggaggctgag	atgggaggat	tgcttgagtc										29094
caggagttca	agaccagcct	ggcaacata	gtgagacccc	ccatctctac	aaaaaaaaatt										29154
tttttttatt	agccaggcac	agccgtgcat	gcctgttaggc	ccaactactt	aggaggctga										29214
ggtgggagga	tctcttgagc	ctgggaggtc	gacactgcag	tgattgcgct	agtgcactcc										29274
agagcaagac	cctatcatct	ctaaatacat	acacacacac	acaatttgaa	gaatttgtaca										29334
gaaaggggac	atgaagctga	gctgagacaa	gggcaaacag	gaaatttatca	ctcacacttt										29394
tcaacagaat	tcaacaaaac	ggagttatgtat	ccaccccccc	gacagtttgg	gggcttggac										29454
cctcagtcgt	acagaaatat	ggaccatagt	tagggtatcc	ttgtctgaag	cctaagggtc										29514
ccaagcacct	tgaaaacgcct	ttgctgggg	tggagggagt	ctgtcaaata	cagggtcagt										29574
gggaagcagc	tgcacctact	ccttaaccac	aggaaagaaa	cctgtcccgg	cccccatgca										29634
gcggggagag	gtgggttctg	agtttggagt	ggaggactcg	ctccatcagg	cagggaccac										29694
atccccctaac	atcacggcat	ggtctgtcca	tctgcttcct	ctaccag	atg ctc cga										29750
					Met Leu Arg										

110.02600201.txt

gac	aaa	ttc	cga	gag	ttc	tcc	cg <sup>g</sup>	gac	aca	agc	acc	atc	ggt	cag	29795
Asp	Lys	Phe	Arg	Glu	Phe	Ser	Arg	Asp	Thr	Ser	Thr	Ile	Gly	Gln	
1735		1740					1745								
gag	cgc	gta	gat	agc	gcc	aat	g <sup>c</sup> g	ctg	gcc	aat	gg <sup>g</sup>	ctc	att	gct	29840
Glu	Arg	Val	Asp	Ser	Ala	Asn	Ala	Leu	Ala	Asn	Gly	Leu	Ile	Ala	
1750					1755						1760				
gg <sup>g</sup>	g <sup>g</sup> c	cat	gct	gca	cg <sup>g</sup>	g <sup>g</sup> c	acc	gt <sup>g</sup>	gcc	gag	tg <sup>g</sup>	aag	gac	agt	29885
Gly	Gly	His	Ala	Ala	Arg	Ala	Thr	Va <sup>l</sup>	Ala	Glu	Trp	Lys	Asp	Ser	
1765					1770						1775				
ctc	aay	gag	gcc	tgg	gct	gac	ct <sup>g</sup>	ctt	gag	ct <sup>g</sup>	ct <sup>g</sup>	gac	aca	cg <sup>g</sup>	29930
Leu	Asn	Glu	Ala	Trp	Ala	Asp	Leu	Leu	Glu	Leu	Leu	Asp	Thr	Arg	
1780					1785						1790				
gg <sup>g</sup>	cag	gt <sup>g</sup>	ct <sup>g</sup>	gcc	g <sup>g</sup> c	g <sup>g</sup> c	ta <sup>c</sup>	gag	ct <sup>g</sup>	cag	cg <sup>c</sup>	ttc	ct <sup>g</sup>	cac	29975
Gly	Gln	Val	Leu	Ala	Ala	Ala	Tyr	Glu	Leu	Gln	Arg	Phe	Leu	His	
1795					1800						1805				
gg <sup>g</sup>	gca	cgc	caa	gcc	ct <sup>g</sup>	g <sup>g</sup> c	cg <sup>g</sup>	gt <sup>g</sup>	cag	cac	aag	cag	cag	cag	30020
Gly	Ala	Arg	Gln	Ala	Leu	Ala	Arg	Val	Gln	His	Lys	Gln	Gln	Gln	
1810					1815						1820				
ctt	ccg	gac	gg <sup>g</sup>	act	g <sup>g</sup> c	cg <sup>c</sup>	gac	ct <sup>c</sup>	aa <sup>c</sup>	g <sup>c</sup> t	gg <sup>g</sup>	gag	gg <sup>g</sup> c	ct <sup>g</sup>	30065
Leu	Pro	Asp	Gly	Thr	Gly	Arg	Asp	Leu	Asn	Ala	Ala	Glu	Ala	Leu	
1825					1830						1835				
cag	cgc	cga	cac	tgt	g <sup>g</sup> c	ta <sup>c</sup>	gag	cat	gac	att	cag	gg <sup>g</sup> c	ct <sup>c</sup>	ag <sup>c</sup>	30110
Gln	Arg	Arg	His	Cys	Ala	Tyr	Glu	His	Asp	Ile	Gln	Ala	Leu	Ser	
1840					1845						1850				
ccc	cag	gt <sup>c</sup> tcgaccca	agcatgagga	gg <sup>g</sup> tggggag	gcttagagga	ggccatgggg									30166
Pro	Gln														
1855															
gaaatgccrg	ggagagttcc	caggagctag	ggtagagctc	agaaatgctg	ggaggacggc										30226
actctttcga	gggccatttc	gagatgagga	aattgctcca	agaaaacatc	tctgtttggc										30286
caggcgcggt	ggctctcgcc	tgtaatccca	gcactttggg	aggccgagga	gggcggatca										30346
tctgagattg	ggagttcgag	accagcctga	ccaacatgga	gaaacccctt	ctctactaaa										30406
aatacaaaat	tagccaggca	tggtggcgca	tgcctgtaat	cctagctact	caggaggctg										30466
aggcaggaga	atcacttgaa	cccaggaggc	ggaggttgca	gtgagctgag	atcccctccac										30526
tgcactccag	cctgggcaac	aagagcgaaa	ctccgtctca	aaaagaaaaga	aaacatctct										30586
ggcaggggac	atctagccac	tgcattcaca	cttttgcata	caagccccct	tgagagccca										30646
tgatgtcttc	aggggttccc	agtggccagg	agactcttca	tgctgcgcta	aatcagcctg										30706
tttagtgcttc	ccaggcttga	gctctaggct	cacactcagt	aagcatggat	ctttccttca										30766
accactagaa	cccatcagcc	cctgcctgaa	tccccttttc	aggctaagta	gcacgttcta										30826
gcaatctcca	aggcttatca	catcctagac	aggcctctga	aagttcaact	ggccaatggc										30886
tcttttccag	atccccc	agttaagccc	aggtgtctag	atgagaccca	acagggcaga										30946
gtgtggccag	cccccgcctt	catctttccc	agcaccccca	ctccctggct	cctccccctg										31006
cttcttttgt	gtgagacagg	caggaggtga	gagtggcagc	caggactgag	gcaggctgct										31066
gctccttgga	gtcccccgct	ctgcctgact	ctgacccggg	gctccgcag	gtc	cag	cag								31124
					Val	Gln	Gln								
gtg	cag	gac	gac	ggc	cac	cg <sup>g</sup>	ctc	cag	aag	g <sup>c</sup> c	ta <sup>c</sup>	g <sup>c</sup> t	g <sup>g</sup> a	gac	31169

110.02600201.txt

val	Gln	Asp	Asp	Gly	His	Arg	Leu	Gln	Lys	Ala	Tyr	Ala	Gly	Asp	
1860						1865						1870			
aag	gct	gag	gag	atc	ggc	cgc	cac	atg	cag	gcc	gtg	gcc	gag	gcc	31214
Lys	Ala	Glu	Glu	Ile	Gly	Arg	His	Met	Gln	Ala	Val	Ala	Glu	Ala	
1875						1880						1885			
tgg	gcc	cag	ctt	cag	gga	agc	tct	gcc	gcc	cgc	cgg	cag	ctg	ctg	31259
Trp	Ala	Gln	Leu	Gln	Gly	Ser	Ser	Ala	Ala	Arg	Arg	Gln	Leu	Leu	
1890						1895						1900			
ctg	gac	acc	aca	gac	aag	ttc	cgc	ttc	ttc	aag	gct	gtc	cgg	gaa	31304
Leu	Asp	Thr	Thr	Asp	Lys	Phe	Arg	Phe	Phe	Lys	Ala	Val	Arg	Glu	
1905						1910						1915			
ctg	atg	ctc	tgg	atg	gat	gag	gtc	aac	ctg	cag	atg	gat	gcc	cag	31349
Leu	Met	Leu	Trp	Met	Asp	Glu	Val	Asn	Leu	Gln	Met	Asp	Ala	Gln	
1920						1925						1930			
gag	cgt	ccc	cg	gtgagaatcc	cagggctcag	ggcctccacg	gttgtgtgg								31400
Glu	Arg	Pro	Arg												
1935															
ctgggctgtg	atggggctgt	gggcagcagg	tgatgctgtt	tctgcctccc	cccaag	g									31456
gat	gtg	tcc	tcc	gcg	gat	cta	gtc	atc	aag	aac	cag	caa	ggc	atc	31501
Asp	Val	Ser	Ser	Ala	Asp	Leu	Val	Ile	Lys	Asn	Gln	Gln	Gly	Ile	
1940						1945						1950			
aag	gca	gag	ata	gag	gcc	cgg	gca	gac	cgc	ttc	tcc	tcc	tgc	atc	31546
Lys	Ala	Glu	Ile	Glu	Ala	Arg	Ala	Asp	Arg	Phe	Ser	Ser	Cys	Ile	
1955						1960						1965			
gac	atg	ggg	aag	gag	ctg	ctg	gcc	agg	agc	cac	tat	gcg	gcc	gag	31591
Asp	Met	Gly	Lys	Glu	Leu	Leu	Ala	Arg	Ser	His	Tyr	Ala	Ala	Glu	
1970						1975						1980			
gag	gtgggtgagg	cctgggtggc	cgggccattc	tcactgtgca	gtgctgaacg										31644
Glu															
ctgacttctt	ggtggcttc	tgcttcctct	cattctctct	ttgcaagaaa	cctttcattc										31704
tctccctca	gtgccattca	cgggtttctc	gcattctgga	aatctctccc	cagaaccgtc										31764
tgccttctga	catcagaaat	cactggtcct	gggccaggca	tggtggtca	cgcctgtaat										31824
cccacgactt	tgggaggccc	aggcgggcag	atcacctgag	gtcacgagtt	cgagaccagc										31884
ctggcctcca	tgggttgtgt	ggctgggctg	tgatgaaaca	tggtgaaacc	ccrtttccac										31944
taaaaaacaca	aaaatttagcc	gggcgtggtg	gcgggcacct	gtaatccag	ctactcggag										32004
gctgaggcag	gagaattgct	tgaacctggg	aggtggaggt	tgcagtgagc	tgagattgcg										32064
ccactgcatt	ccagccttgg	tgataagagc	gaaactccgt	ctcaaaaaaa	agaaagaaag										32124
aaagaaatca	cctgtcccc	tggatgactc	cccagggggcc	gtggaaagat	ctcacatcct										32184
ggtgctaact	catactcttg	ccccccgccc	ccag	atc	tca	gag	aag	ctg	tct						32236
Ile	Ser	Glu	Lys	Leu	Ser										
1985															
cag	ctg	cag	gca	cgg	cgc	cag	gag	aca	gct	gag	aag	tgg	cag	gag	32281
Gln	Leu	Gln	Ala	Arg	Arg	Gln	Glu	Thr	Ala	Glu	Lys	Trp	Gln	Glu	
1990						1995						2000			
aag	atg	gac	tgg	ctt	cag	ctg	g	gtgagctgcc	aaggggggccc	caggccctgt					32333
Lys	Met	Asp	Trp	Leu	Gln	Leu									
2005						2010									

110.02600201.txt

ggggagtgcc gggcatcctg caccctgtgg gttccagagt aggtgagact aggaacctcg 32393  
ggtgtgaaac tcacgatgcc caatcttcgt gctgcctggc acagctctgg ggcagtggct . 32453  
ttctctgtgt ccctgttctt gaggttgccr tggccactgt gccctgcccag tggccactct 32513  
gaccaccat ctccctgcaa ccccygatcc tgccag tt ttg gag gtg ctt gtg 32566  
Val Leu Glu Val Leu Val  
2015

ttt gga aga gat gca ggg atg gca gag gcc tgg ctc tgc agc cag 32611  
 Phe Gly Arg Asp Ala Gly Met Ala Glu Ala Trp Leu Cys Ser Gln  
 2020 2025 2030 2030

gag cca ctg gtg cgc agc gct gag ctg ggt tgc acg gtc gac gaa 32656  
 Glu Pro Leu Val Arg Ser Ala Glu Leu Gly Cys Thr Val Asp Glu  
 2035 2040 2045

gtt gag agc ctc atc aag cg<sup>g</sup> cac gag g<sup>c</sup>c ttc c<sup>a</sup>g a<sup>a</sup>g t<sup>c</sup>a g<sup>c</sup>a 32701  
 Val Glu Ser Leu Ile Lys Arg His Glu Ala Phe Gln Lys Ser Ala  
 2050 2055 2060

gtg gcc tgg gag gag cga ttc tgt gcg ctg gag aag ctt act gcg 32746  
 Val Ala Trp Glu Glu Arg Phe Cys Ala Leu Glu Lys Leu Thr Ala  
 2065 2070 2075

gtgaggggaca caggaccccg gatgccact ccaacctqct cccctqacct qtqctqqctt 32806

ctgcttggcg aagacatgtc ttctctttcc ccactgcacc agtggcccccc agatgtgagc 32866

tgagagtgcc attgtcacca caccctttagg aggtgtcagt tccccctgggt gatgccgaga 32926

caccttctcc cgctttgttt cttgtccccca ccatcttctc aaattctgtt ctccctttat 32986

catatttcat caaatttagag ttgagacatt tcagctctcc ctgtttccct ggcccttta 33046

cacgcaacct tctacacgct gaatgcagca tttttgaca g cta gag gag cg 33099  
Leu Glu Glu Arg  
2080

gag aag gag cga aag aga aag agg gag gag gag gag cgg cgg aaa  
 Glu Lys Glu Arg Lys Arg Lys Arg Glu Glu Glu Arg Arg Arg Lys  
 2085 2090 2095 33144

cag ccg cct gct ccc gaa ccc aca gcc agt gtg cct cca ggg gac 33189  
 Gln Pro Pro Ala Pro Glu Pro Thr Ala Ser Val Pro Pro Gly Asp  
 2100 2105 2110

ctg gtg ggc ggc cag aca gct tct gac acc acc tgg gac gg 33230  
 Leu Val Gly Gly Gln Thr Ala Ser Asp Thr Thr Trp Asp Gly  
 2115 2120 2125

gtgagagcca ggatgcctgg gttaggaggag gcggctgagc ccaggccacc cagggactaa 33290

tgattctgtg ttgccttgg tcatccag a acc cag cca cgg cca cca cca 33341  
                  Thr Gln Pro Arg Pro Pro Pro Pro  
                  2130

tcc aca caa gca ccc agt gtt aat gga gtc tgc aca gat gga gag 33386  
 Ser Thr Gln Ala Pro Ser Val Asn Gly Val Cys Thr Asp Gly Glu  
 2135 2140 2145

ccc tca cag gtgaccccac tgtccctctg tgccccatc ggagtcgtag 33435  
Pro Ser Gln  
2150

ccccctccacc cccgcacatc cttttacaga ttcttgtcct tgcag ccc ctg ctg gga 33492  
Pro Leu Leu Gly

caa cag aga ctt gag cac agc aqc ttc ccc gaa ggg ccg gtqagttccc 33541

110.02600201.txt

Gln	Gln	Arg	Leu	Glu	His	Ser	Ser	Phe	Pro	Glu	Gly	Pro		
2155			2160						2165					
ctgcaagtgt ggtgttata actgtgaggc gaagggtcca gaggggtgg tgagtgcggg													33601	
tggggagta ctggggatgg gatgggagag ggcagaggct cacaggcagc ttggggcag													33661	
gaagaccaac tcctggacac ggagcttcct ggcacccagg ttagggatct cccgtctcaa													33721	
cctttacac tgacactgat tccccccag gga cct ggc tca ggg gac gaa													33772	
Gly Pro Gly Ser Gly Asp Glu														
2170														
gcc aat ggg ccc cgg gga gag agg cag acc cg													33817	
Ala	Asn	Gly	Pro	Arg	Gly	Glu	Arg	Gln	Thr	Arg	Thr	Arg	Gly	Pro
2175			2180						2185					
gcc cca tct gca atg ccc cag agc agg tct acc gag tca gcc cat													33862	
Ala	Pro	Ser	Ala	Met	Pro	Gln	Ser	Arg	Ser	Thr	Glu	Ser	Ala	His
2190			2195						2200					
gct gcc acc ctg ccg cct cga ggc cca gag cca tct gcc cag gag													33907	
Ala	Ala	Thr	Leu	Pro	Pro	Arg	Gly	Pro	Glu	Pro	Ser	Ala	Gln	Glu
2205			2210						2215					
cag atg gag ggg atg ctg tgc cgc aag cag gag atg gag gcc ttc													33952	
Gln	Met	Glu	Gly	Met	Leu	Cys	Arg	Lys	Gln	Glu	Met	Glu	Ala	Phe
2220			2225						2230					
ggg aag aag gct gcc aac ag gtacagcctc tctggagcct gctctcagag													34002	
Gly	Lys	Lys	Ala	Ala	Asn	Arg								
2235			2240											
ggcacttccc cagagcctct gcccagatag agggagggat gcccattaga gtacatctc													34062	
tgggcaaagg gtaggacttg ggaccagagc ggggcctcag gggaggacca gagggtgtga													34122	
agaccgtggc ctaaagatgg gagcagaact ggaagtccta ggacacccaa gagggctcca													34182	
ggttgcgggc gccactgagg ccggccagtc agcaccgcgt ccctcgca g tcc tgg													34238	
Ser Trp														
cag aac gtg tac tgt gtc ctg cgg cgt ggg agc ctc ggc ttt tac													34283	
Gln	Asn	Val	Tyr	Cys	Val	Leu	Arg	Arg	Gly	Ser	Leu	Gly	Phe	Tyr
2245				2250					2255					
aag gat gcc aag gca gcc agc gcg gga gtg cca tac cac gga gaa													34328	
Lys	Asp	Ala	Lys	Ala	Ala	Ser	Ala	Gly	Val	Pro	Tyr	His	Gly	Glu
2260			2265						2270					
gtg cct gtc agc ctg gcc agg gcc cag ggc agc gtc gcc ttt gat													34373	
Val	Pro	Val	Ser	Leu	Ala	Arg	Ala	Gln	Gly	Ser	Val	Ala	Phe	Asp
2275			2280						2285					
tac cga aag cgc aaa cat gtc ttc aag ctg gg gtaggaacag													34415	
Tyr	Arg	Lys	Arg	Lys	His	Val	Phe	Lys	Leu	Gly				
2290			2295											
ggaacagtgc tctcggatg ggaggaggt tggagtgac acaggtgagc catgagtcag													34475	
gtccagaggg agggagttc ctgttaaggag cctgagtggta gtaaccaggc cagccactt													34535	
gggatagtgt agatgaggaa ggcggagatt ctgggtgtct ccacaccaag gggagcagga													34595	
gaaccaagac ccaggcctga cggctgccaa tgtcaaggtg aaaaattacc cagggtgaaa													34655	
aatccaaagg tagggatctg ggaagacctc caggggcctg tccctgcctg ccagcaagca													34715	
gccggagcag gaggcccgaa ctgggtggaa agttagctcc ccctgcctct gggccagtc													34775	

110.02600201.txt

agaaaggatg tgcttctgca cagtctgggg aagctgaaga atgtgcagag cgtgtcttgt 34835  
 ccggctctga gggctgcagc cagatgtccc agcctggtgt tgggtcatga ttacgctctc 34895  
 accagcagct acctggcaga tccggattct cagctctgcc ctgtggcctc tctctcccaa 34955  
 cag c tta cag gat gga aaa gaa tat tta ttc cag gcc aag gat gag 35001  
 Leu Gln Asp Gly Lys Glu Tyr Leu Phe Gln Ala Lys Asp Glu  
 2300 2305 2310

gtgagctgtc cttcggttcc ctctctgtcc gtgccattcc agaagcttcc agctgcagac 35061  
 tccccttctt tccctgtcct ccctttttt cctggtcttg tcctttgtgg taagactgga 35121  
 tgtgtgcac ggccgcacca ggccgcactc cctgtctaag ccggccacat tctcctaata 35181  
 gcatgaaaca gtcagctcac tttctgcctc ctcctttac acttccctgc tgtccactgc 35241  
 ggccaactca gcacagtgtc cttgaagctg attgagggtc tttccatcca cag gca 35297  
 Ala

gag atg agc tcg tgg cta cgg gtg gtg aat gca gcc att gcc aca 35342  
 Glu Met Ser Ser Trp Leu Arg Val Val Asn Ala Ala Ile Ala Thr  
 2315 2320 2325

gcg tct tct gcc tct gga gag cct gaa gag ccg gtg gtg ccc agc 35387  
 Ala Ser Ser Ala Ser Gly Glu Pro Glu Glu Pro Val Val Pro Ser  
 2330 2335 2340

acc acc cg<sup>g</sup> ggc atg acc cgg gcc atg acc atg ccc cca gtg tca 35432  
 Thr Thr Arg Gly Met Thr Arg Ala Met Thr Met Pro Pro Val Ser  
 2345 2350 2355

ccc gtc ggg gct gag ggg cct gtt gtg ctc cgc agc aaa gac ggc 35477  
 Pro Val Gly Ala Glu Gly Pro Val Val Leu Arg Ser Lys Asp Gly  
 2360 2365 2370

aga gaa cga gag cga gaa aaa cgc ttc agc ttc ttt aag aag aac 35522  
 Arg Glu Arg Glu Arg Glu Lys Arg Phe Ser Phe Phe Lys Lys Asn  
 2375 2380 2385

aag tagttggggg caaggtccca ggccaactcc ctccctccgt tcagggaaact 35575  
 Lys  
 2390

gccagggaca gtcgacaggg accgcccctt tgcaggaca actgcctgct gctagggtct 35635  
 gttgccaagg tcaacccatc accaggaact gtcactgggg acgagtccat gttcccaagg 35695  
 rcagcccttc tcttctgctg tttaattcca gactggtgtt gggacccagg taaccccttc 35755  
 tcccccccccc gccgacttct cccctttccc cagcctcg<sup>t</sup> cctctgtccc tcaccacgg 35815  
 gtggacagtg ccgcaccctc aacataggcc atgtggggag tggctgcccc tgccctcagg 35875  
 tcattctcctt gccatgygag ggcactcgcc ttctgccttc tggttcctca cccctcagac 35935  
 cagccaggaa cctctcagag ctgaagcagg ccctggggc agaagtgc<sup>a</sup> gatgacagtc 35995  
 agaggcgcag gagccctccc tcccccccccc caccctgtaa ctccagctgc cactccatct 36055  
 ccagytgctc tcaatggctt ccaggtgtgt tgytcgggga cagccaccgc cttgagtctg 36115  
 gccaaggagg tgattaaaca gctcagttc tc 36147

<210> 2  
 <211> 2390  
 <212> PRT  
 <213> Homo sapiens

<220>  
<221> MOD\_RES  
<222> (825)  
<223> Gly or Ser

<220>  
<221> MOD\_RES  
<222> (1034)  
<223> Ala or Val

<220>  
<221> MOD\_RES  
<222> (1446)  
<223> Ala or Val

<400> 2  
Met Ser Ser Thr Leu Ser Pro Thr Asp Phe Asp Ser Leu Glu Ile Gln  
1 5 10 15  
Gly Gln Tyr Ser Asp Ile Asn Asn Arg Trp Asp Leu Pro Asp Ser Asp  
20 25 30  
Trp Asp Asn Asp Ser Ser Ala Arg Leu Phe Glu Arg Ser Arg Ile  
35 40 45  
Lys Ala Leu Ala Asp Glu Arg Glu Ala Val Gln Lys Lys Thr Phe Thr  
50 55 60  
Lys Trp Val Asn Ser His Leu Ala Arg Val Thr Cys Arg Val Gly Asp  
65 70 75 80  
Leu Tyr Ser Asp Leu Arg Asp Gly Arg Asn Leu Leu Arg Leu Leu Glu  
85 90 95  
Val Leu Ser Gly Glu Ile Leu Pro Lys Pro Thr Lys Gly Arg Met Arg  
100 105 110  
Ile His Cys Leu Glu Asn Val Asp Lys Ala Leu Gln Phe Leu Lys Glu  
115 120 125  
Gln Lys Val His Leu Glu Asn Met Gly Ser His Asp Ile Val Asp Gly  
130 135 140  
Asn His Arg Leu Thr Leu Gly Leu Val Trp Thr Ile Ile Leu Arg Phe  
145 150 155 160  
Gln Ile Gln Asp Ile Ser Val Glu Thr Glu Asp Asn Lys Glu Lys Lys  
165 170 175  
Ser Ala Lys Asp Ala Leu Leu Trp Cys Gln Met Lys Thr Ala Gly  
180 185 190  
Tyr Pro Asn Val Asn Val His Asn Phe Thr Thr Ser Trp Arg Asp Gly  
195 200 205  
Leu Ala Phe Asn Ala Ile Val His Lys His Arg Pro Asp Leu Leu Asp  
210 215 220  
Phe Glu Ser Leu Lys Lys Cys Asn Ala His Tyr Asn Leu Gln Asn Ala  
225 230 235 240  
Phe Asn Leu Ala Glu Lys Glu Leu Gly Leu Thr Lys Leu Leu Asp Pro  
245 250 255  
Glu Asp Val Asn Val Asp Gln Pro Asp Glu Lys Ser Ile Ile Thr Tyr  
260 265 270  
Val Ala Thr Tyr Tyr His Tyr Phe Ser Lys Met Lys Ala Leu Ala Val  
275 280 285

110.02600201.txt

Glu Gly Lys Arg Ile Gly Lys Val Leu Asp His Ala Met Glu Ala Glu  
 290 295 300

Arg Leu Val Glu Lys Tyr Glu Ser Leu Ala Ser Glu Leu Leu Gln Trp  
 305 310 315 320

Ile Glu Gln Thr Ile Val Thr Leu Asn Asp Arg Gln Leu Ala Asn Ser  
 325 330 335

Leu Ser Gly Val Gln Asn Gln Leu Gln Ser Phe Asn Ser Tyr Arg Thr  
 340 345 350

Val Glu Lys Pro Pro Lys Phe Thr Glu Lys Gly Asn Leu Glu Val Leu  
 355 360 365

Leu Phe Thr Ile Gln Ser Lys Leu Arg Ala Asn Asn Gln Lys Val Tyr  
 370 375 380

Thr Pro Arg Glu Gly Arg Leu Ile Ser Asp Ile Asn Lys Ala Trp Glu  
 385 390 395 400

Arg Leu Glu Lys Ala Glu His Glu Arg Glu Leu Ala Leu Arg Thr Glu  
 405 410 415

Leu Ile Arg Gln Glu Lys Leu Glu Gln Leu Ala Ala Arg Phe Asp Arg  
 420 425 430

Lys Ala Ala Met Arg Glu Thr Trp Leu Ser Glu Asn Gln Arg Leu Val  
 435 440 445

Ser Gln Asp Asn Phe Gly Leu Glu Leu Ala Ala Val Glu Ala Ala Val  
 450 455 460

Arg Lys His Glu Ala Ile Glu Thr Asp Ile Val Ala Tyr Ser Gly Arg  
 465 470 475 480

Val Gln Ala Val Asp Ala Val Ala Ala Glu Leu Ala Ala Glu Arg Tyr  
 485 490 495

His Asp Ile Lys Arg Ile Ala Ala Arg Gln His Asn Val Ala Arg Leu  
 500 505 510

Trp Asp Phe Leu Arg Gln Met Val Ala Ala Arg Arg Glu Arg Leu Leu  
 515 520 525

Leu Asn Leu Glu Leu Gln Lys Val Phe Gln Asp Leu Leu Tyr Leu Met  
 530 535 540

Asp Trp Met Glu Glu Met Lys Gly Arg Leu Gln Ser Gln Asp Leu Gly  
 545 550 555 560

Arg His Leu Ala Gly Val Glu Asp Leu Leu Gln Leu His Glu Leu Val  
 565 570 575

Glu Ala Asp Ile Ala Val Gln Ala Glu Arg Val Arg Ala Val Ser Ala  
 580 585 590

Ser Ala Leu Arg Phe Cys Asn Pro Gly Lys Glu Tyr Arg Pro Cys Asp  
 595 600 605

Pro Gln Leu Val Ser Glu Arg Val Ala Lys Leu Glu Gln Ser Tyr Glu  
 610 615 620

Ala Leu Cys Glu Leu Ala Ala Ala Arg Arg Ala Arg Leu Glu Glu Ser  
 625 630 635 640

Arg Arg Leu Trp Arg Phe Leu Trp Glu Val Gly Glu Ala Glu Ala Trp  
 645 650 655

Val Arg Glu Gln Gln His Leu Leu Ala Ser Ala Asp Thr Gly Arg Asp  
 660 665 670

## 110.02600201.txt

Leu Thr Gly Ala Leu Arg Leu Leu Asn Lys His Thr Ala Leu Arg Gly  
 675 680 685  
 Glu Met Ser Gly Arg Leu Gly Pro Leu Lys Leu Thr Leu Glu Gln Gly  
 690 695 700  
 Gln Gln Leu Val Ala Glu Gly His Pro Gly Ala Ser Gln Ala Ser Ala  
 705 710 715 720  
 Arg Ala Ala Glu Leu Gln Ala Gln Trp Glu Arg Leu Glu Ala Leu Ala  
 725 730 735  
 Glu Glu Arg Ala Gln Arg Leu Ala Gln Ala Ala Ser Leu Tyr Gln Phe  
 740 745 750  
 Gln Ala Asp Ala Asn Asp Met Glu Ala Trp Leu Val Asp Ala Leu Arg  
 755 760 765  
 Leu Val Ser Ser Pro Glu Leu Gly His Asp Glu Phe Ser Thr Gln Ala  
 770 775 780  
 Leu Ala Arg Gln His Arg Ala Leu Glu Glu Glu Ile Arg Ser His Arg  
 785 790 795 800  
 Pro Thr Leu Asp Ala Leu Arg Glu Gln Ala Ala Ala Leu Pro Pro Thr  
 805 810 815  
 Leu Ser Arg Thr Pro Glu Val Gln Xaa Arg Val Pro Thr Leu Glu Arg  
 820 825 830  
 His Tyr Glu Glu Leu Gln Ala Arg Ala Gly Glu Arg Ala Arg Ala Leu  
 835 840 845  
 Glu Ala Ala Leu Ala Leu Tyr Thr Met Leu Ser Glu Ala Gly Ala Cys  
 850 855 860  
 Gly Leu Trp Val Glu Glu Lys Glu Gln Trp Leu Asn Gly Leu Ala Leu  
 865 870 875 880  
 Pro Glu Arg Leu Glu Asp Leu Glu Val Val Gln Gln Arg Phe Glu Thr  
 885 890 895  
 Leu Glu Pro Glu Met Asn Thr Leu Ala Ala Gln Ile Thr Ala Val Asn  
 900 905 910  
 Asp Ile Ala Glu Gln Leu Leu Lys Ala Asn Pro Pro Gly Lys Asp Arg  
 915 920 925  
 Ile Val Asn Thr Gln Glu Gln Leu Asn His Arg Trp Gln Gln Phe Arg  
 930 935 940  
 Arg Leu Ala Asp Gly Lys Lys Ala Ala Leu Thr Ser Ala Leu Ser Ile  
 945 950 955 960  
 Gln Asn Tyr His Leu Glu Cys Thr Glu Thr Gln Ala Trp Met Arg Glu  
 965 970 975  
 Lys Thr Lys Val Ile Glu Ser Thr Gln Gly Leu Gly Asn Asp Leu Ala  
 980 985 990  
 Gly Val Leu Ala Leu Gln Arg Lys Leu Ala Gly Thr Glu Arg Asp Leu  
 995 1000 1005  
 Glu Ala Ile Ala Ala Arg Val Gly Glu Leu Thr Arg Glu Ala Asn Ala  
 1010 1015 1020  
 Leu Ala Ala Gly His Pro Ala Gln Ala Xaa Ala Ile Asn Ala Arg Leu  
 1025 1030 1035 1040  
 Arg Glu Val Gln Thr Gly Trp Glu Asp Leu Arg Ala Thr Met Arg Arg

110.02600201.txt  
 1045                    1050                    1055

Arg Glu Glu Ser Leu Gly Glu Ala Arg Arg Leu Gln Asp Phe Leu Arg  
 1060                    1065                    1070

Ser Leu Asp Asp Phe Gln Ala Trp Leu Gly Arg Thr Gln Thr Ala Val  
 1075                    1080                    1085

Ala Ser Glu Glu Gly Pro Ala Thr Leu Pro Glu Ala Glu Ala Leu Leu  
 1090                    1095                    1100

Ala Gln His Ala Ala Leu Arg Gly Glu Val Glu Arg Ala Gln Ser Glu  
 1105                    1110                    1115                    1120

Tyr Ser Arg Leu Arg Ala Leu Gly Glu Glu Val Thr Arg Asp Gln Ala  
 1125                    1130                    1135

Asp Pro Gln Cys Leu Phe Leu Arg Gln Arg Leu Glu Ala Leu Gly Thr  
 1140                    1145                    1150

Gly Trp Glu Glu Leu Gly Arg Met Trp Glu Ser Arg Gln Gly Arg Leu  
 1155                    1160                    1165

Ala Gln Ala His Gly Phe Gln Gly Phe Leu Arg Asp Ala Arg Gln Ala  
 1170                    1175                    1180

Glu Gly Val Leu Ser Ser Gln Glu Tyr Val Leu Ser His Thr Glu Met  
 1185                    1190                    1195                    1200

Pro Gly Thr Leu Gln Ala Ala Asp Ala Ala Ile Lys Lys Leu Glu Asp  
 1205                    1210                    1215

Phe Met Ser Thr Met Asp Ala Asn Gly Glu Arg Ile His Gly Leu Leu  
 1220                    1225                    1230

Glu Ala Gly Arg Gln Leu Val Ser Glu Gly Asn Ile His Ala Asp Lys  
 1235                    1240                    1245

Ile Arg Glu Lys Ala Asp Ser Ile Glu Arg Arg His Lys Lys Asn Gln  
 1250                    1255                    1260

Asp Ala Ala Gln Gln Phe Leu Gly Arg Leu Arg Asp Asn Arg Glu Gln  
 1265                    1270                    1275                    1280

Gln His Phe Leu Gln Asp Cys His Glu Leu Lys Leu Trp Ile Asp Glu  
 1285                    1290                    1295

Lys Met Leu Thr Ala Gln Asp Val Ser Tyr Asp Glu Ala Arg Asn Leu  
 1300                    1305                    1310

His Thr Lys Trp Gln Lys His Gln Ala Phe Met Ala Glu Leu Ala Ala  
 1315                    1320                    1325

Asn Lys Asp Trp Leu Asp Lys Val Asp Lys Glu Gly Arg Glu Leu Thr  
 1330                    1335                    1340

Leu Glu Lys Pro Glu Leu Lys Ala Leu Val Ser Glu Lys Leu Arg Asp  
 1345                    1350                    1355                    1360

Leu His Arg Arg Trp Asp Glu Leu Glu Thr Thr Thr Gln Ala Lys Ala  
 1365                    1370                    1375

Arg Ser Leu Phe Asp Ala Asn Arg Ala Glu Leu Phe Ala Gln Ser Cys  
 1380                    1385                    1390

Cys Ala Leu Glu Ser Trp Leu Glu Ser Leu Gln Ala Gln Leu His Ser  
 1395                    1400                    1405

Asp Asp Tyr Gly Lys Asp Leu Thr Ser Val Asn Ile Leu Leu Lys Lys  
 1410                    1415                    1420

## 110.02600201.txt

Gln Gln Met Leu Glu Trp Glu Met Ala Val Arg Glu Lys Glu Val Glu  
 1425 1430 1435 1440  
 Ala Ile Gln Ala Gln Xaa Lys Ala Leu Ala Gln Glu Asp Gln Gly Ala  
 1445 1450 1455  
 Gly Glu Val Glu Arg Thr Ser Arg Ala Val Glu Glu Lys Phe Arg Ala  
 1460 1465 1470  
 Leu Cys Gln Pro Met Arg Glu Arg Cys Arg Arg Leu Gln Ala Ser Arg  
 1475 1480 1485  
 Glu Gln His Gln Phe His Arg Asp Val Glu Asp Glu Ile Leu Trp Val  
 1490 1495 1500  
 Thr Glu Arg Leu Pro Met Ala Ser Ser Met Glu His Gly Lys Asp Leu  
 1505 1510 1515 1520  
 Pro Ser Val Gln Leu Leu Met Lys Lys Asn Gln Thr Leu Gln Lys Glu  
 1525 1530 1535  
 Ile Gln Gly His Glu Pro Arg Ile Ala Asp Leu Arg Glu Arg Gln Arg  
 1540 1545 1550  
 Ala Leu Gly Ala Ala Ala Ala Gly Pro Glu Leu Ala Glu Leu Gln Glu  
 1555 1560 1565  
 Met Trp Lys Arg Leu Gly His Glu Leu Glu Leu Arg Gly Lys Arg Leu  
 1570 1575 1580  
 Glu Asp Ala Leu Arg Ala Gln Gln Phe Tyr Arg Asp Ala Ala Glu Ala  
 1585 1590 1595 1600  
 Glu Ala Trp Met Gly Glu Gln Glu Leu His Met Met Gly Gln Glu Lys  
 1605 1610 1615  
 Ala Lys Asp Glu Leu Ser Ala Gln Ala Glu Val Lys Lys His Gln Val  
 1620 1625 1630  
 Leu Glu Gln Ala Leu Ala Asp Tyr Ala Gln Thr Ile His Gln Leu Ala  
 1635 1640 1645  
 Ala Ser Ser Gln Asp Met Ile Asp His Glu His Pro Glu Ser Thr Arg  
 1650 1655 1660  
 Ile Ser Ile Arg Gln Ala Gln Val Asp Lys Leu Tyr Ala Gly Leu Lys  
 1665 1670 1675 1680  
 Glu Leu Ala Gly Glu Arg Arg Glu Arg Leu Gln Glu His Leu Arg Leu  
 1685 1690 1695  
 Cys Gln Leu Arg Arg Glu Leu Asp Asp Leu Glu Gln Trp Ile Gln Glu  
 1700 1705 1710  
 Arg Glu Val Val Ala Ala Ser His Glu Leu Gly Gln Asp Tyr Glu His  
 1715 1720 1725  
 Val Thr Met Leu Arg Asp Lys Phe Arg Glu Phe Ser Arg Asp Thr Ser  
 1730 1735 1740 1745  
 Thr Ile Gly Gln Glu Arg Val Asp Ser Ala Asn Ala Leu Ala Asn Gly  
 1745 1750 1755 1760  
 Leu Ile Ala Gly Gly His Ala Ala Arg Ala Thr Val Ala Glu Trp Lys  
 1765 1770 1775  
 Asp Ser Leu Asn Glu Ala Trp Ala Asp Leu Leu Glu Leu Leu Asp Thr  
 1780 1785 1790  
 Arg Gly Gln Val Leu Ala Ala Ala Tyr Glu Leu Gln Arg Phe Leu His  
 1795 1800 1805

## 110.02600201.txt

Gly Ala Arg Gln Ala Leu Ala Arg Val Gln His Lys Gln Gln Gln Leu  
 1810 1815 1820  
 Pro Asp Gly Thr Gly Arg Asp Leu Asn Ala Ala Glu Ala Leu Gln Arg  
 1825 1830 1835 1840  
 Arg His Cys Ala Tyr Glu His Asp Ile Gln Ala Leu Ser Pro Gln Val  
 1845 1850 1855  
 Gln Gln Val Gln Asp Asp Gly His Arg Leu Gln Lys Ala Tyr Ala Gly  
 1860 1865 1870  
 Asp Lys Ala Glu Glu Ile Gly Arg His Met Gln Ala Val Ala Glu Ala  
 1875 1880 1885  
 Trp Ala Gln Leu Gln Gly Ser Ser Ala Ala Arg Arg Gln Leu Leu Leu  
 1890 1895 1900  
 Asp Thr Thr Asp Lys Phe Arg Phe Phe Lys Ala Val Arg Glu Leu Met  
 1905 1910 1915 1920  
 Leu Trp Met Asp Glu Val Asn Leu Gln Met Asp Ala Gln Glu Arg Pro  
 1925 1930 1935  
 Arg Asp Val Ser Ser Ala Asp Leu Val Ile Lys Asn Gln Gln Gly Ile  
 1940 1945 1950  
 Lys Ala Glu Ile Glu Ala Arg Ala Asp Arg Phe Ser Ser Cys Ile Asp  
 1955 1960 1965  
 Met Gly Lys Glu Leu Leu Ala Arg Ser His Tyr Ala Ala Glu Glu Ile  
 1970 1975 1980  
 Ser Glu Lys Leu Ser Gln Leu Gln Ala Arg Arg Gln Glu Thr Ala Glu  
 1985 1990 1995 2000  
 Lys Trp Gln Glu Lys Met Asp Trp Leu Gln Leu Val Leu Glu Val Leu  
 2005 2010 2015  
 Val Phe Gly Arg Asp Ala Gly Met Ala Glu Ala Trp Leu Cys Ser Gln  
 2020 2025 2030  
 Glu Pro Leu Val Arg Ser Ala Glu Leu Gly Cys Thr Val Asp Glu Val  
 2035 2040 2045  
 Glu Ser Leu Ile Lys Arg His Glu Ala Phe Gln Lys Ser Ala Val Ala  
 2050 2055 2060  
 Trp Glu Glu Arg Phe Cys Ala Leu Glu Lys Leu Thr Ala Leu Glu Glu  
 2065 2070 2075 2080  
 Arg Glu Lys Glu Arg Lys Arg Lys Arg Glu Glu Glu Arg Arg Lys  
 2085 2090 2095  
 Gln Pro Pro Ala Pro Glu Pro Thr Ala Ser Val Pro Pro Gly Asp Leu  
 2100 2105 2110  
 Val Gly Gln Thr Ala Ser Asp Thr Thr Trp Asp Gly Thr Gln Pro  
 2115 2120 2125  
 Arg Pro Pro Pro Ser Thr Gln Ala Pro Ser Val Asn Gly Val Cys Thr  
 2130 2135 2140  
 Asp Gly Glu Pro Ser Gln Pro Leu Leu Gly Gln Gln Arg Leu Glu His  
 2145 2150 2155 2160  
 Ser Ser Phe Pro Glu Gly Pro Gly Ser Gly Asp Glu Ala Asn  
 2165 2170 2175  
 Gly Pro Arg Gly Glu Arg Gln Thr Arg Thr Arg Gly Pro Ala Pro Ser

	<b>110.02600201.txt</b>		
2180	2185	2190	
Ala Met Pro Gln Ser Arg Ser Thr Glu Ser Ala His Ala Ala Thr Leu			
2195	2200	2205	
Pro Pro Arg Gly Pro Glu Pro Ser Ala Gln Glu Gln Met Glu Gly Met			
2210	2215	2220	
Leu Cys Arg Lys Gln Glu Met Glu Ala Phe Gly Lys Lys Ala Ala Asn			
2225	2230	2235	2240
Arg Ser Trp Gln Asn Val Tyr Cys Val Leu Arg Arg Gly Ser Leu Gly			
2245	2250	2255	
Phe Tyr Lys Asp Ala Lys Ala Ala Ser Ala Gly Val Pro Tyr His Gly			
2260	2265	2270	
Glu Val Pro Val Ser Leu Ala Arg Ala Gln Gly Ser Val Ala Phe Asp			
2275	2280	2285	
Tyr Arg Lys Arg Lys His Val Phe Lys Leu Gly Leu Gln Asp Gly Lys			
2290	2295	2300	
Glu Tyr Leu Phe Gln Ala Lys Asp Glu Ala Glu Met Ser Ser Trp Leu			
2305	2310	2315	2320
Arg Val Val Asn Ala Ala Ile Ala Thr Ala Ser Ser Ala Ser Gly Glu			
2325	2330	2335	
Pro Glu Glu Pro Val Val Pro Ser Thr Thr Arg Gly Met Thr Arg Ala			
2340	2345	2350	
Met Thr Met Pro Pro Val Ser Pro Val Gly Ala Glu Gly Pro Val Val			
2355	2360	2365	
Leu Arg Ser Lys Asp Gly Arg Glu Arg Glu Arg Glu Lys Arg Phe Ser			
2370	2375	2380	
Phe Phe Lys Lys Asn Lys			
2385	2390		

<210> 3  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 3  
 ctgccttcct gcttcacttt 20

<210> 4  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 4  
 tcatgacgag ctgacaaagc 20

<210> 5  
 <211> 20  
 <212> DNA

110.02600201.txt

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 5

ccctgccaaac tggtgtag

20

&lt;210&gt; 6

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 6

ggtcccccttg gacactttc

20

&lt;210&gt; 7

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 7

tgcctgtctg tgttcctgag

20

&lt;210&gt; 8

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 8

tcctccatct ttgtgttgt tg

22

&lt;210&gt; 9

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 9

acaccaggag ttcctgtcca

20

&lt;210&gt; 10

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 10

tgctccgagt gctattcctt

20

&lt;210&gt; 11

&lt;211&gt; 20

110.02600201.txt

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 11  
ttggtgtggg tttcctcttc

20

<210> 12  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 12  
cactggtcca ctcctgtct

20

<210> 13  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 13  
gaacctctgg gaggcctga

19

<210> 14  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 14  
tccctgaagg ctgtgctaatt

20

<210> 15  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 15  
cctcggtggc tttaattctg

20

<210> 16  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 16  
atgtgtgcaa ggcatactgg

19

<210> 17

110.02600201.txt

<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 17  
ccaccctgtc ctttccacta

20

<210> 18  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 18  
cccagttctg accagcctaa

20

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

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 19  
agaggcactg tcccttggt

19

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

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 20  
gctggttcac actccacaga

20

<210> 21  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 21  
aaaaaacgca gccaggttag

20

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

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 22  
gctcttgatg tgctccttcc

20

110.02600201.txt

<210> 23  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 23  
ggctgggtta aggctctgac 20

<210> 24  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 24  
agggactcac cacccacat 19

<210> 25  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 25  
gctgcctccc acaattcac 19

<210> 26  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 26  
tccccattgc ttcattttc 20

<210> 27  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 27  
ggaagaagct tccaaacagg 20

<210> 28  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 28  
ccatcctgct cttcacatt 20

110.02600201.txt

<210> 29  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 29  
tgcttggttcctaccc 20

<210> 30  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 30  
ggtttcctgt gccacgttta 20

<210> 31  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 31  
ggtagccaa agggtcacaa 20

<210> 32  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 32  
acaaaaacca cgtcctggag 20

<210> 33  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 33  
ggctaatttg ggcactttga 20

<210> 34  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 34  
cccctttctt ctgctgttca 20

<210> 35  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 35  
gcggaaatgc agagctaaca 20

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

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 36  
ggagatggtc aatgccaaag 20

<210> 37  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 37  
tgtccccact cccactaatc 20

<210> 38  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 38  
aaaaacacgt ccaagtctgg 20

<210> 39  
<211> 19  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 39  
ctgacgggtg ttaccatcg 19

<210> 40  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 40

110.02600201.txt

agcaactgaag gctccacatt

20

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

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 41  
gaacagaccg gaggtcagag

20

<210> 42  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 42  
ctgtgggtcc tccactcttc

20

<210> 43  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 43  
taacatcacg gcatggtctg

20

<210> 44  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 44  
ccctagctcc tggaaactct

20

<210> 45  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 45  
cttggagtcc cccgctct

18

<210> 46  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

110.02600201.txt

<400> 46  
aagcagaaag ccaccaagaa 20

<210> 47  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 47  
tcacatcctg gtgctaactc a 21

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

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 48  
cctactctgg aacccacagg 20

<210> 49  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 49  
ccactctgac ccaccatctt 20

<210> 50  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 50  
aagccagcac aggtcagg 18

<210> 51  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 51  
ccctcttaca cgcaaccttc 20

<210> 52  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 52  
gacccttcgc ctcacagttt 20

<210> 53  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 53  
ggtagggat ctcccggtc 20

<210> 54  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 54  
cccttgccc agaagatgtt 20

<210> 55  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 55  
agatgggagc agaactggaa 20

<210> 56  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 56  
ctggccttgt tactccactt 20

<210> 57  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 57  
tacgctctca ccagcagcta 20

<210> 58  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>

110.02600201.txt

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 58

cgcacacatc cagtcttacc

20

&lt;210&gt; 59

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 59

cagtcactt tctgcctcct

20

&lt;210&gt; 60

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 60

agagaggctg tggtcaggaa

20

&lt;210&gt; 61

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 61

agcgttacca cgacatcaa

20

&lt;210&gt; 62

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 62

ccctcgactc ttgatcactc tt

22

&lt;210&gt; 63

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 63

gtggccaagg tagagcagag

20

&lt;210&gt; 64

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

110.02600201.txt

<220>  
<223> Description of Artificial Sequence: Synthetic primer  
  
<400> 64  
cacctcccaag agggaaacg 18

<210> 65  
<211> 39  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer  
  
<400> 65  
cacgacgttg taaaacgacg aactgggact taccaagct 39

<210> 66  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer  
  
<400> 66  
gaactgggac ttaccaagcc 20

<210> 67  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer  
  
<400> 67  
ccaaagaagg ccctgtatca 20

<210> 68  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer  
  
<400> 68  
cctcagcttc acccacctc 19

<210> 69  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer  
  
<400> 69  
agcgctacca cgacatcaag 20

<210> 70  
<211> 20  
<212> DNA  
<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 70

caggccctcc actcctgcta

20

&lt;210&gt; 71

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 71

gtgtcccaagg acaactttgg

20

&lt;210&gt; 72

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 72

atccagtcca gtttgaggag gagccgctcc

30

&lt;210&gt; 73

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 73

ctccctcaacc tggactggat ggaagagatg

30

&lt;210&gt; 74

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 74

ctccagggtg agcttcagg

19

&lt;210&gt; 75

&lt;211&gt; 39

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 75

ctcatcttag aagaggatct gagcagcacg ctgtcaccc

39

&lt;210&gt; 76

&lt;211&gt; 44

&lt;212&gt; DNA

110.02600201.txt

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: synthetic primer

&lt;400&gt; 76

cgcgggtacc accatggaac aaaaactcat ctcagaagag gatc

44

&lt;210&gt; 77

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 77

gaggagcctc agcaggttg

19