FORM P8
(To be lodged in duplicate)

REPUBLIC OF SOUTH AFRICA
PATENTS ACT, 1978

PUBLICATION PARTICULARS AND ABSTRACT

[Section 32 (3) (a) - Regulations 22 (1) (g) and 31]

OFFICIAL APPLICATION NO	LODGING DATE	ACCEPTANCE DATE
21 2005/06283	22 5 Aug 2005	43 15 03 06

INTERNATIONAL CLASSIFICATION
51 C12Q G01N A61K

Not for publication
Classified by:

FULL NAME(S) OF APPLICANT(S)

71 **LANG, Florian**

FULL NAME(S) OF INVENTOR(S)

72 **LANG, Florian; BUSJAHN, Andreas**

TITLE OF INVENTION

54 **USE OF A NOVEL POLYMORPHISM IN THE HSGK1 GENE IN THE DIAGNOSIS OF HYPERTONIA AND USE OF THE SGK GENE FAMILY IN THE DIAGNOSIS AND THERAPY OF THE LONG QT SYNDROME**

EARLIEST PRIORITY CLAIMED	COUNTRY	NUMBER	DATE
NB - Use International abbreviation for country (See Schedule 4)	33 DE	31 103 05 213.5	32 7 Feb 2003

57	Abstract (not more than 150 words) and figure of the drawings to which the abstract refers, are attached.

Number of sheets	62

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
19. August 2004 (19.08.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/070057 A3

(51) Internationale Patentklassifikation[7]: C12Q 1/68, G01N 33/53, A61K 31/56, 38/18

(21) Internationales Aktenzeichen: PCT/EP2004/001051

(22) Internationales Anmeldedatum: 5. Februar 2004 (05.02.2004)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
103 05 213.5 7. Februar 2003 (07.02.2003) DE

(71) Anmelder und
(72) Erfinder *(für alle Bestimmungsstaaten mit Ausnahme von US)*: **LANG, Florian** [DE/DE]; Im Rotbad 52, 72076 Tübingen (DE).

(72) Erfinder; und
(75) Erfinder/Anmelder *(nur für US)*: **BUSJAHN, Andreas** [DE/DE]; Heine-Str. 18, 16341 Panketal (DE).

(74) Anwalt: **ISENBRUCK, Günter**; Isenbruck Bösl Hörschler Wichmann Huhn, Theodor-Heuss-Anlage 12 68165 Mannheim (DE).

(81) Bestimmungsstaaten *(soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart)*: 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, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

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

Veröffentlicht:
— *mit internationalem Recherchenbericht*
— *vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen*

[Fortsetzung auf der nächsten Seite]

(54) **Title:** USE OF A NOVEL POLYMORPHISM IN THE HSGK1 GENE IN THE DIAGNOSIS OF HYPERTONIA AND USE OF THE SGK GENE FAMILY IN THE DIAGNOSIS AND THERAPY OF THE LONG QT SYNDROME

(54) **Bezeichnung:** VERWENDUNG EINES NEUEN POLYMORPHISMUS IM HSGK1-GEN ZUR DIAGNOSE DER HYPERTONIE UND VERWENDUNG DER SGK-GENFAMILIE ZUR DIAGNOSE UND THERAPIE DES LONG-Q/T-SYNDROMS

(57) **Abstract:** The invention relates to the use of single- or double-stranded nucleic acids that contain a fragment of the hsgk in the diagnosis of hypertonia. The said fragment has a minimum length of 10 nucleotides/base pairs and the said fragment further comprises a polymorphism which is the result of the presence or absence of an insert of the nucleotide G in position 732/733 in intron 2 of the hsgk1 gene. The invention also relates to the use of the direct correlation between overexpression or the functional molecular modification of human homologues of the sgk family and the length of the Q/T time in the diagnosis of the Long QT syndrome, and to the use of the nucleic acid of a human homologue of the sgk gene family or of one of its fragments in the diagnosis of the Long QT syndrome. Polymorphisms of single nucleotides (single nucleotide polymorphisms = SNP) in the human homologues of the sgk gene family are especially useful in the diagnosis of a congenital predisposition for the Long QT syndrome. In another aspect, the invention relates to the use of a functional activator or a transcriptional factor which boosts expression of the genes of the sgk family for producing a drug for use in the therapy and/or the prophylaxis of the Long QT syndrome.

(57) **Zusammenfassung:** Die Erfindung betrifft die Verwendung einer einzel- oder doppelsträngigen Nukleinsäure enthaltend ein Fragment der hsgk zur Diagnose von Hypertonie, wobei das besagte Fragment mindestens 10 Nukleotide/Basenpaare lang ist und wobei das besagte Fragment weiterhin einen Polymorphismus umfasst, welcher sich aus der Anwesenheit bzw. Abwesenheit einer Insertion des Nukleotids G an Position 732/733 in Intron 2 des hsgk 1 Gens ergibt. Die Erfindung betrifft weiterhin die Verwendung der direkten Korrelation zwischen der Überexpression oder der funktionalen molekularen Modifikation von humanen Homologen der sgk-Familie und der Länge der Q/T-Zeit zur Diagnose des Long-Q/T-Syndroms, sowie die Verwendung der Nukleinsäure eines humanen Homologen der sgk-Genfamilie oder eines ihrer Fragmente zur Diagnose des Long-Q/T-Syndroms. Insbesondere lassen sich auch hier Polymorphismen einzelner Nukleotide (single nucleotide polymorphisms = SNP) in den humanen Homologen der sgk-Genfamilie zur Diagnose einer genetisch bedingten Prädisposition zum Long-Q/T-Syndrom einsetzen. In einem weiteren Aspekt betrifft die Erfindung die Verwendung eines funktionalen Aktivators oder eines Transkriptionsfaktors, der die Expression der Gene der sgk-Familie steigert, zur Herstellung eines Arzneimittels zur Therapie und/oder zur Prophylaxe des Long-Q/T-Syndroms.

WO 2004/070057 A3

Use of a novel polymorphism in the hsgk1 gene for the diagnosis of hypertension and use of the sgk gene family for the diagnosis and therapy of the long Q/T syndrome

The present invention relates to the use of a single-stranded or double-stranded nucleic acid containing an hsgk fragment for diagnosing hypertension, with said fragment being at least 10 nucleotides/base pairs in length and with said fragment furthermore comprising a polymorphism which ensues from the presence or absence of an insertion of the nucleotide G at position 732/733 in intron 2 of the hsgk1 gene.

The present invention furthermore relates to the use of the direct correlation between the overexpression or the functional molecular modification of human homologues of the sgk family and the length of the Q/T interval for diagnosing the long Q/T syndrome and also to the use of the nucleic acid of a human homologue of the sgk gene family or of one of its fragments for diagnosing the long Q/T syndrome. In particular, polymorphisms of individual nucleotides (single nucleotide polymorphisms = SNPs) in the human homologues of the sgk gene family can also, in the present case, be used for diagnosing a genetically determined predisposition for the long Q/T syndrome.

In a further aspect, the invention relates to the use of a functional activator or transcription factor which increases the expression of the genes of the sgk family for producing a pharmaceutical for the therapy and/or prophylaxis of the long Q/T syndrome.

Numerous extracellular signals lead to intracellular phosphorylation/dephosphorylation cascades for the purpose of ensuring rapid transfer of these signals from the plasma membrane and its receptors into the cytoplasm and the cell nucleus. The specificity of these reversible signal transfection cascades is made possible by a large number of individual proteins, in particular kinases, which transfer a phosphate group onto individual substrates.

The serum- and glucocorticoid-dependent kinase (sgk), which is a serine/threonine kinase whose expression is increased by serum and glucocorticoids, was initially cloned from rat mammary carcinoma cells (Webster et al., 1993). The human

version of sgk, i.e. hsgk1, was cloned from liver cells (Waldegger et al., 1997). It was found that the expression of hsgk1 is influenced by regulating the cell volume. It has as yet not been possible to demonstrate such a dependence on the cell volume as far as expression of the rat sgk is concerned. It has furthermore been found that the rat kinase stimulates the epithelial Na^+ channel (ENaC) (Chen et al., 1999; Naray-Pejes-Toth et al., 1999). The ENaC in turn plays a crucial role in renal Na^+ excretion. An increase in the activity of the ENaC leads to an increase in the renal retention of sodium ions and, in this way, to the development of hypertension, as WO02/074987 A2 demonstrates.

Finally, two further members of the human sgk human family, i.e. hsgk2 and hsgk3, have been cloned (Kobayashi et al., 1999), both of which genes are, like hsgk1 as well, activated by insulin and IGF1 by way of the PI3 kinase route. Electrophysiological experiments have shown that coexpression of hsgk2 and hsgk3 likewise results in a significant increase in the activity of the ENaC.

It is evident from DE 197 08 173 A1 that hsgk1 possesses substantial diagnostic potential in connection with many diseases in which changes in cell volume play a crucial pathophysiological role, such as hypernatremia, hyponatremia, diabetes mellitus, renal insufficiency, hypercatabolism, hepatic encephalopathy and microbial or viral infections.

WO 00/62781 reported that hsgk1 activates the endothelial Na^+ channel, thereby increasing renal Na^+ resorption. Since this increased renal Na^+ resorption is accompanied by hypertension, it was presumed, in this case, that an increase in the expression of hsgk1 would lead to hypertension while a reduction in the expression of hsgk1 would ultimately lead to hypotension.

DE 100 421 37 also reported a similar connection between the overexpression or hyperactivity of the human homologues hsgk2 and hsgk3 and the hyperactivation of the ENaC, the increase in renal Na^+ resorption resulting therefrom and the hypertension which develops from this. Furthermore, this document already discussed the diagnostic potential of the kinases hsgk2 and hsgk3 with regard to essential hypertension.

WO02/074987 A2 discloses the connection between the occurrence of two different polymorphisms (single nucleotide polymorphism (SNP)) of individual nucleotides in the hsgk1 gene and a genetically determined predisposition for hypertension. In this case, the polymorphisms are a polymorphism in intron 6

(T→C) and a polymorphism in exon 8 (C→T) in the hsgk1 gene. In particular, it is evident from Table 5 in WO02/074987 A2 that there is a strong correlation disequilibrium between the two SNPs which had been analyzed: most CC carriers of the SNP in exon 8 are also intron 6 TT carriers (namely 64%) whereas only a few exon 8 TT carriers are also at the same time intron 6 CC carriers (only 2%). These correlations which had been observed between the occurrence of the polymorphisms in intron 6 and exon 8 substantiate the necessity of analyzing the genotype for the two polymorphisms (intron 6 and exon 8) in order to demonstrate a predisposition for hypertension, with this leading to a high degree of technical input and time consumption.

A need exists to provide a further polymorphism in the hsgk1 gene, the occurrence of which in one or the other version may correlate even better than the two known polymorphisms in exon 8 and intron 6 with the phenotypic occurrence of hypertension in the patient. In particular, the provision of a single SNP which correlates with the predisposition for hypertension and whose presence in one or the other version even has consequences for a functional molecular modification of the hsgk1 protein would be very advantageous.

This need was fulfilled by providing a novel polymorphism in the hsgk1 gene, which polymorphism comprises the insertion of the nucleotide G at position 732/733. It has been found that individuals which possess such an insertion of the nucleotide G at position 732/733 (InsG/InsG) occur more frequently and have a lower predisposition for developing hypertension. On the other hand, individuals which do not possess such an insertion at position 732/733 (WT/WT) occur more rarely and have a markedly higher predisposition for developing hypertension. According to the results obtained thus far, this correlation between the genotype, in regard to the polymorphism at position 732/733 in intron 2, and the predisposition to the development of hypertension appears to have a markedly higher significance than the corresponding correlations with regard to the polymorphisms in exon 8 and intron 6 (see Table 1).

Furthermore, it is to be assumed, on the basis of the prediction obtained using known prediction programs, that the expression of a specific splice variant of the hsgk1 gene depends on the presence or absence of the G insertion at position 732/733 in intron 2 of the hsgk1 gene. The expression of such a specific splice variant of the hsgk1 gene could result in a functional molecular modification of the hsgk1 protein, which leads to the hsgk1 activity being modified, in particular to the

hsgk1 activity being increased. The physiological consequences of this molecular modification of the hsgk1 protein, in particular an increase in the activity of the hsgk1, could then ultimately result in the development of the symptoms of hypertension.

In a first aspect, the invention relates to the use of an isolated single-stranded or double-stranded nucleic acid which comprises a fragment of the nucleic acid sequence as depicted in SEQ ID No. 1 or as depicted in SEQ ID No. 2 for diagnosing hypertension, with said fragment being at least 10 nucleotides/base pairs, preferably at least 15 nucleotides/base pairs, in particular at least 20 nucleotides/base pairs, in length and with said fragment comprising the polymorphism in intron 2 of the hsgk1 gene either with or without the insertion of the nucleotide G at position 732/733.

SEQ ID No. 1 describes the genomic DNA sequence of hsgk1 without the insertion of nucleotide G (or GTP) at position 732/733 in intron 2 of the hsgk1 gene, i.e. what is termed the "wild-type (WT)" sequence, and SEQ ID No. 2 describes the genomic DNA sequence of hsgk1 with the insertion of nucleotide G (or GTP) at position 732/733 in intron 2 of the hsgk1 gene, i.e. what is termed the "insertion G (InsG)" sequence.

In a second aspect, the present invention relates to a kit for diagnosing hypertension, which kit comprises at least one isolated single-stranded or double-stranded nucleic acid which comprises a fragment of the sequence as depicted in SEQ ID No. 1 or 2. In this connection, said fragment from SEQ ID No. 1 or 2 is at least 10 nucleotides/base pairs, preferably at least 15 nucleotides/base pairs, in particular at least 20 nucleotides/base pairs, in length. Furthermore, said fragment from SEQ ID No. 1 or 2 should comprise the polymorphism in intron 2 of the hsgk1 gene either with or without the insertion of the nucleotide G at position 732/733.

Alternatively, the kit for diagnosing hypertension can, in addition to, or instead of, the abovementioned single-stranded or double-stranded nucleic acid, also comprise at least one antibody which is directed against such a region of the hsgk protein whose presence in the hsgk1 protein depends on the presence of the insertion of the nucleotide G at position 732/733 in intron 2 of the corresponding encoding hsgk gene. If, for example, the presence of the G insertion at position 732/733 in the hsgk1 gene were to induce the splicing-out of an exon, an antibody which was directed against precisely this spliced-out protein region could be used for

detecting the polymorphism version of the individual. Such an antibody could be used, therefore, to diagnose a predisposition for developing hypertension.

In a third aspect, the invention relates to a method for diagnosing hypertension, which method comprises the following procedural steps:

a) withdrawing a body sample from an individual,

b) where appropriate, isolating and/or amplifying genomic DNA, cDNA or mRNA from the body sample according to a),

c) quantifying the alleles which possess an insertion of the nucleotide G at position 732/733 in intron 2 of the hsgk1 gene.

In step a), a body sample is withdrawn from a test individual which is preferably a mammal, in particular a human. In this diagnostic method according to the invention, the body samples from a patient which are preferably used are blood samples or saliva samples which comprise cellular material and which can be obtained from the patient with relatively little effort. However, other body samples which likewise comprise cells, such as tissue or cell samples or the like, can also be used.

In step b), standard methods (Sambrook J. and Russell D.W. (2001) Cold Spring Harbor, NY, CSHL Press) are used to prepare, where appropriate, and/or amplify, where appropriate, either genomic DNA or cDNA or else mRNA from the body sample from a). In this connection, it is possible to use any suitable methods which are familiar to the skilled person. It is also possible, where appropriate, to dispense with this DNA isolation step or DNA amplification step, in particular when use is made, in step c) of detection methods which themselves involve a PCR amplification step.

In step c), the number of alleles which possess an insertion of nucleotide G at position 732/733 in intron 2 of the hsgk1 gene is finally quantified. In this connection, those individuals which possess two WT alleles ought to have a predisposition for developing hypertension. The quantification/identification of the alleles with regard to the polymorphism at position 732/733 in intron 2 of the hsgk1 gene can be effected by using a variety of methods which are known to the skilled person. Some preferred methods are explained in more detail below. However, the quantification of the number of alleles which possess an insertion of nucleotide G at position 732/733 in intron 2 of the hsgk1 gene is not restricted to the following preferred methods which are described below.The genotype (or the number of alleles) can preferably be identified, with regard to the polymorphism at

position 732/733, by directly sequencing the DNA, preferably the genomic DNA, from the body sample at said position 732/733 in intron 2 of the hsgk1 gene. To do this, it is necessary to use known sequencing methods to make available, as sequencing primers, short oligonucleotides which possess sequences from the immediate vicinity of position 732/733 of the hsgk1 gene.

Any known methods which are based on hybridizing the genomic DNA from the body sample with specific hybridization probes constitute further methods, which are likewise preferred, for identifying the genotype (or for quantifying the number of alleles) with regard to the polymorphism at position 732/733.

Southern blotting is an example of such a hybridization method. If, for example, the presence of the G insertion at position 732/733 in intron 2 of the hsgk1 gene were to destroy or else form a cleavage site for a restriction endonuclease, it would be possible to use specific hybridization probes to detect nucleic acid fragments having lengths which differ from the corresponding fragment lengths in the WT allele. In this way, it would be possible to detect a genotype which was specific with regard to the polymorphism in question at position 732/733.

If, as a result of the presence or absence of the G insertion at position 732/733, an alternative splice variant which lacks a particular exon were to be expressed, it would also be possible to detect the genotype, with regard to the polymorphism at position 732/733 in question, using a specific hybridization probe from the exon which was missing in the splice variant.

Another example of a hybridization method is that of hybridizing the genomic DNA from the body sample with a labeled, single-stranded oligonucleotide which is preferably 15-25 nucleotides in length and which either does or does not possess a G insertion at position 732/733. Under very specific hybridization conditions, which can be tested experimentally for each individual oligonucleotide using known methods, it is possible to distinguish a completely hybridizing oligonucleotide from an oligonucleotide having one single base mismatch.

Other preferred methods for identifying the genotype (or for quantifying the number of alleles) with regard to the polymorphism at position 732/733 are, in particular, the PCR oligonucleotide elongation assay or the ligation assay.

In the case of the PCR oligonucleotide elongation assay, it would be possible, for example, to provide an oligonucleotide which possesses the sequence of a

fragment from SEQ ID No. 2 and, at its 3' end, the G at the polymorphism position 732/733. When this oligonucleotide was hybridized with a sample fragment of the WT allele (without G insertion), it would not be possible to extend, and ultimately amplify, this fragment in a subsequent PCR reaction because of the mismatch at the 3' end. On the other hand, when this oligonucleotide was hybridized with an InsG allele, it would be possible, because of the perfect base pairing at the 3' end of the oligonucleotide, to achieve elongation and ultimately to obtain a PCR amplification product.

A ligation assay is ultimately based on the same principle as the PCR oligonucleotide elongation assay: only those double-stranded nucleic acid fragments which possess an exact base pairing at their end can be ligated to another double-stranded nucleic acid fragment. The appearance of a specific ligation product can therefore be made dependent on the presence or absence of the G insertion at position 732/733 in intron 2 of the hsgk1 gene.

In addition to the correlation of the polymorphism according to the invention with the predisposition for hypertension, a second correlation of the polymorphism according to the invention was surprisingly found with the length of what is termed the Q/T interval. Markedly shorter Q/T intervals are seen in individuals which possess a WT/WT genotype with regard to position 732/733 in intron 2 of the hsgk1 gene than in individuals which possess an InsG/InsG genotype. Heterozygous (WT/InsG) individuals possess intermediate Q/T intervals (see Table 3). A significantly extended Q/T interval leads to the development of what is termed the long Q/T syndrome, which can manifest itself in cardiac rhythm disturbances, by way of ventricular fibrillation through to sudden cardiac death. Individuals possessing the InsG/InsG genotype ought therefore to have a predisposition for developing the long Q/T syndrome.

Because of the direct correlation, which has been demonstrated, between the length of the Q/T interval and the genetic makeup of the hsgk1 gene, in particular between the length of the Q/T interval and the polymorphism at position 732/733 in intron 2 of the hsgk1 gene, it is to be assumed that nucleic acids of another human homologue of the sgk family are likewise suitable for diagnosing the long QT syndrome.

The Q, R and S waves which can be detected using an ECG measuring instrument constitute experimental values for assessing depolarization. The Q/T interval is defined as the time which is to be detected, using an ECG measuring instrument,

from the beginning of the propagation of the T wave (the appearance of the Q deflection) to the end of depolarization which is characterized by the end of the T wave. The Q/T interval therefore constitutes the time which elapses between the beginning of a new state of excitation of the heart and the return to the resting state. A markedly extended Q/T interval accordingly leads to cardiac rhythm disturbances and, ultimately, to the long Q/T syndrome which has already been mentioned.

The invention also relates, therefore, to the use of the direct correlation between the overexpression or functional molecular modification of human homologues of the sgk family, in particular of the hsgk1 gene, and the length of the Q/T interval for diagnosing the long QT syndrome.

A human homologue of the sgk family, which homologue encompasses, in the above sense, a functional molecular modification, is understood, in this connection, as being a homologue of the sgk family which is mutated in such a manner that the properties, in particular the catalytic properties or the substrate specificity, of the corresponding protein are altered.

The direct correlation, according to the invention, between the Q/T interval and the genetic makeup of the human homologues of the sgk family implies that it would be possible for individual mutations in the genes hsgk1, hsgk2 or hsgk3 to occur in individual patients, with these mutations modifying the level of expression or functional properties of the kinases hsgk1, hsgk2 or hsgk3 and, in this way, leading to a genetically occasioned prolongation of the Q/T interval and, ultimately, to a predisposition for the development of the long Q/T syndrome. Such mutations could, for example, occur in the regulatory gene regions or else in intron sequences of the sgk gene locus. On the other hand, the individual differences in the genetic makeup of the sgk locus could also affect the coding region of the gene. Mutations in the coding region could then, where appropriate, lead to a functional change in the corresponding kinase, as, for example, to the catalytic properties of the kinase being modified, with these modified properties also ultimately influencing the Q/T interval. Accordingly, both the above-described types of mutation could bring about a prolongation of the Q/T interval and thereby, ultimately, predisposition for development of the long Q/T syndrome.

The above-described mutations in the human homologues of the sgk family, which bring about the predisposition for development of the long Q/T syndrome in the patient, are as a rule what are termed single nucleotide polymorphisms (SNPs)

either in the exon region or in the intron region of these homologues. In their less frequently occurring version, termed the mutated version in that which follows, SNPs in the exon region of the hsgk genes can, where appropriate, give rise to amino acid substitutions in the corresponding hsgk protein and consequently lead to the kinase being functionally modified. In their mutated version, SNPs in the intron region or in regulatory sequences of the hsgk genes can, where appropriate, lead to a change in the level of expression of the corresponding kinase. However, SNPs in the intron region could also lead to a functional modification of the kinase if they affect the alternative splicing of the immature mRNA.

The invention also relates to the use of a single-stranded or double-stranded nucleic acid which comprises the sequence of a human homologue of the sgk family or one of its fragments, in particular the hsgk1 gene itself or one of its fragments, for diagnosing a predisposition for developing the long Q/T syndrome. In this connection, the single-stranded or double-stranded nucleic acid preferably has a length of at least 10 nucleotides/base pairs.

In addition to the abovementioned single-stranded or double-stranded nucleic acids, certain antibodies which are directed against substrates of the human homologues of the sgk family, in particular against substrates of hsgk1, are also suitable for diagnosing a predisposition for developing the long Q/T syndrome and hypertension. These diagnostic antibodies are preferably directed against an epitope of the human homologues of the sgk family, in particular of hsgk1, which contains the phosphorylation site of the substrate either in phosphorylated form or in unphosphorylated form.

In a preferred embodiment, the ubiquitin protein ligase Nedd4-2 (Acc No. BAA23711) is used as the substrate of the human homologue of the sgk family. This ubiquitin protein ligase is a protein which is specifically phosphorylated by the human homologues of the sgk family [Debonneville et al., Phosphorylation of Nedd4-2 by Sgk 1 regulates epithelial Na(+) channel cell surface expression. EMBO J., 2001; 20: 7052-7059; Snyder et al., Serum and glucocorticoid-regulated kinase modulates Nedd4-2-mediated inhibition of the epithelial Na(+) channel. J. Biol. Chem. 2002, 277: 5-8]. Phosphorylation sites for hsgk1 possess the consensus sequence (R X R X X S/T) where R is arginine, S is serine, T is threonine and X is any arbitrary amino acid. In Nedd4-2 (Acc No. BAA23711) there are two potential phosphorylation sites for hsgk1 which the abovementioned

consensus sequence fits: the serine at amino acid position 382 and the serine at amino acid position 468.

The abovementioned antibodies for diagnosing a predisposition for developing the long Q/T syndrome are therefore preferably directed against the substrate Nedd4-2 and, particularly preferably, against a region of the Nedd4-2 protein which possesses the sequence of the potential phosphorylation site for hsgk1, i.e. the consensus sequence (R X R X X S/T). In particular, these antibodies are directed against Nedd4-2 protein regions which encompass at least one of the two potential phosphorylation sites serine at amino acid position 382 and/or serine at amino acid position 468.

The invention furthermore relates to a kit for diagnosing the long QT syndrome or other diseases which manifest themselves in a prolongation of the Q/T interval. This kit for diagnosing the long QT syndrome preferably comprises antibodies which are directed against the human homologues of the sgk protein family or, in particular, nucleic acids which are able to hybridize, under stringent conditions, with the human homologues of the sgk gene family. The kit can also jointly comprise antibodies which are directed against the human homologues of the sgk protein family and nucleic acids which hybridize, under stringent conditions, with the human homologues of the sgk gene family. Particularly preferably, the kit according to the invention for diagnosing the long Q/T syndrome can also comprise antibodies which are directed against the hsgk1 protein or nucleic acids which are able to hybridize, under stringent conditions, with the hsgk1 gene.

In this connection, a hybridization under stringent conditions is understood as meaning a hybridization under those hybridization conditions, with regard to hybridization temperature and formamide content in the hybridization solution, which have been described in relevant specialist literature (Sambrook J. and Russell D.W. (2001) Cold Spring Harbor, NY, CSHL Press).

In particular, the diagnostic kit can comprise, as hybridization probes, single-stranded or double-stranded nucleic acids which possess a sequence as depicted in SEQ ID No. 1 or 2, which are at least 10 nucleotides/base pairs in length and which encompass the polymorphism at position 732/733 in intron 2 of the hsgk1 gene either with or without the insertion of the nucleotide G.

The diagnostic kit according to the invention provides, in particular, antibodies which are specifically directed against those regions of the hsgk1 protein whose

presence in the hsgk1 protein depends on the presence of the G insertion at position 732/733 in intron 2 of the hsgk1 gene. In particular, those regions which, due to the presence or absence of this G insertion in the immature mRNA, are spliced out alternatively, and are therefore not present in the mature mRNA and in the protein arising from it, are suitable for use as immunogenic epitopes against which diagnostic antibodies can be directed. Correspondingly, precisely those nucleic acid regions of the hsgk1 gene which are able to hybridize with such a gene region which is spliced out in dependence on the G insertion at position 732/733 are also suitable for use as diagnostic hybridization probes.

The kit for diagnosing the long Q/T syndrome can also preferably comprise, as specific hybridization probes, nucleic acid fragments which encompass the known SNPs in the hsgk1 gene, in particular the SNP in exon 8 (C2617T, D240D), the SNP in intron 6 (T2071C) and/or the SNP in intron 2 at position 732/733 (insertion of G).

The correlation, which has been demonstrated within the context of the invention, between the genetic makeup of the genes of the hsgk1 gene family and the length of the Q/T interval also makes it possible to use functional activators, or positive transcription regulators, of the sgk family therapeutically for treating the long Q/T syndrome and similar diseases which are likewise accompanied by a prolonged Q/T interval. In this connection, a "functional activator" is understood as being a substance which activates the physiological function of the corresponding kinase of the sgk family. A "positive transcription regulator" is understood as being a substance which activates the expression of the corresponding kinase of the sgk family.

The invention consequently also relates to the use of a functional activator, or a positive transcription regulator, of a human homologue of the sgk family, in particular of hsgk1, for lowering the Q/T interval and, in particular, for therapy and/or prophylaxis of the long QT syndrome. Known functional activators and/or positive transcription regulators of the human homologues of the sgk family, in particular of hsgk1, are glucocorticoids, mineralocorticoids, aldosterone, gonadotropins and a number of cytokines, in particular TGF-β.

The invention therefore furthermore relates to the use of substrates selected from the group of substances comprising glucocorticoids, mineralocorticoids, aldosterone, gonadotropins and cytokines, in particular TGF-β, for producing a pharmaceutical for the therapy and/or prophylaxis of the long QT syndrome. The

invention also relates to a pharmaceutical which comprises a substance selected from the abovementioned group of substances for the therapy and/or prophylaxis of the long Q/T syndrome.

The present invention is explained in detail by means of the following examples.

Example 1

A correlation study, in which the genotype of the hsgk1 gene in different patients (twins) was compared with the systolic and diastolic blood pressure values which were measured in these patients, and then statically evaluated, was carried out within the context of the present invention.

75 dizygotic pairs of twins were used for the correlation analysis (Busjahn et al., J. Hypertens 1996, 14: 1195-1199; Busjahn et al., Hypertension 1997, 29: 165-170). The experimental subjects all belonged to the German-Caucasian race and originated from different parts of Germany. Blood was taken from the pairs of twins, and from their parents, for the purpose of verifying the dizygotism and for further molecular genetic analyses. Each of the experimental subjects underwent a prior medical examination. None of the experimental subjects was known to be suffering from any chronic medically recognized disease. After 5 min, the blood pressure of the test subject, whose was in the sitting position, was measured by a trained physician using a standardized mercury sphygmomanometer (2 measurements at a time interval of 1 min). The mean of the two measurements was used as the blood pressure value.

The advantage of using dizygotic twins for correlation studies is that they are of the same age and that the external influences on their phenotypes are to be judged as being minimal (Martin et al., Nat Genet 1997, 17: 387-392).
The importance of studies on twins for the elucidation of complex genetic diseases was recently described by Martin et al., 1997.

The dizygotism of the pairs of twins was confirmed by using the polymerase chain reaction (PCR) to amplify five microsatellite markers. In this analysis of microsatellite markers, deoxyribonucleic acid (DNA) fragments are amplified by PCR using specific oligonucleotides which contain regions which are highly variable in different human individuals. The high degree of variability in these regions of the genome can be detected by means of slight differences in sizes of the amplified fragments, resulting, when there is diversity at the corresponding gene locus, in double bands, i.e. what are termed microsatellite bands, being formed after the PCR products have been subjected to gel-electrophoretic fractionation (Becker et al., J. Reproductive Med 1997, 42: 260-266).

For the purpose of carrying out a molecular genetic analysis of the target gene, in the present case the hsgk1 gene, three further microsatellite marker regions

(d6s472, d6s1038, d6s270) in the immediate vicinity of the hsgk1 locus were amplified by PCR and then compared with the corresponding samples from the other twin and the parents. In this way, it was possible to decide whether the twins had inherited alleles, from their parents, which were identical or different with regard to the allele under investigation. The correlation analysis was carried out using the structural equation modelling (SEM) model (Eaves et al., Behav Genet 1996, 26: 519-525; Neale, 1997: Mx: Statistical modeling, Box 126 MCV, Richmond, VA 23298: Department of Psychiatry. 4th edition). This model is based on variance-covariances matrices of the test pairs which are characterized by the probability that they possess either no, one or two identical alleles. The variance with regard to the phenotype was divided into a variance which is based on the genetic background of all the genes (A), a variance which is based on the genetic background of the target gene (Q), in this case the hsgk1 gene, and the variance due to external influences (E).

$$VAR = A^2+Q^2+E^2$$

The covariance of a test pair was defined for the three possible allele combinations IBD_0, IBD_1 and IBD_2 (IBD = identical by descent; 0, 1 or 2 identical alleles) as follows:

$$COV(IBD_0)=0.5\ A^2 \qquad COV(IBD_1)=0.5\ A^2+0.5\ Q^2 \qquad COV(IBD_2)=0.5\ A^2+Q^2$$

In order to evaluate the correlation between the genetic makeup of the hsgk1 locus and the blood pressure of the test subject, the differences between models which do and, respectively, do not take into account the genetic variance with regard to the target gene hsgk1 were calculated as a χ^2 statistic. For each pair and each gene locus, the allele ratios were calculated by means of the so-called multipoint model (MAPMAKER/SIBS; Kruglyak et al., Am J Hum Genet 1995, 57: 439-454) based on the parental genotypes.

The greater informative value of the analytical method which is based on a variance-covariance evaluation, as compared with the above-described χ^2 statistic (S.A.G.E. Statistical Analysis for Genetic Epidemiology, Release 2.2. Computer program package, Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA, 1996) was recently confirmed in a simulation study (Fulker et al., Behav Gen 1996, 26: 527-532). An error probability of $p < 0.01$ was accepted in order to ensure a significant correlation

with regard to the criteria of Lander and Kruglyak (Lander et al., Nat Genet 1995, 11: 241-246).

Table 1 shows the results of this correlation study,

Table 1:

Phenotype	max χ^2	p
Systolic blood pressure value (lying)	4.44	0.04
Diastolic blood pressure value (lying)	14.36	0.0002
Systolic blood pressure value (sitting)	5.55	0.019
Diastolic blood pressure value (sitting)	4.92	0.027
Systolic blood pressure value (standing)	1.91	0.17
Diastolic blood pressure value (standing)	4.83	0.028

As can be seen from Table 1, the low values for the ascertained error probabilities p, which do not exceed, or only slightly exceed, the accepted error probability of $p < 0.01$, prove that there is a direct correlation between the genetic variance with regard to the hsgk1 gene locus and the phenotypically ascertained variance of the measured blood pressure.

Example 2

The genomic organization of the hsgk1 gene has already been described (Waldegger et al., Genomics, 51, 299 [1998]), http://www.ensembl.org/Homo_sapiens/geneview?gene= ENSG00000118515).

In order to identify SNPs whose occurrences are relevant for a predisposition for developing hypertension, the SNPs in the hsgk1 gene which were published in databases were first of all investigated in order to determine whether they are genuine SNPs, and not simple sequencing errors, and whether the SNPs are sufficiently polymorphic in order to form the basis for a diagnostic detection of a predisposition for hypertension. The SNP rs 1057293 in exon 8, which concerns the replacement of a C with a T (http://www.ensembl.org/Homo_sapiens/snpview?snp=1057293; http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?type=rs&rs=1057293) and a second SNP, which is located in the hsgk1 gene, at a distance of precisely 551 bp from the first

SNP, in the donor splice site of intron 6 to exon 7 and concerns the replacement of a T with a C, had already been located in this way.

Example 3

Blood samples were taken from a random sample of the 75 pairs of twins. After the genomic DNA of the hsgk1 gene had been amplified from the blood samples by means of PCR, the exons and introns (but not the promoter region) of the hsgk1 gene were sequenced directly and completely using suitable sequencing primers. When the sequences of the hsgk1 genes which originated from different test subjects were compared, a further polymorphism in intron 2, consisting of the insertion of an additional nucleotide G in position 732/733, was noted. Furthermore, the presence or absence of this G insertion at position 732/733 in the hsgk1 genes of the individual test subjects exhibited a significant correlation with the blood pressure which was measured in the individual test subjects: on average, InsG/InsG genotypes exhibited significantly lower systolic and diastolic blood pressure values than did the less frequent WT/WT genotypes as well as the heterozygous WT/InsG genotypes (see Table 3). By contrast, other polymorphisms in the hsgk1 gene exhibited a correlation with the measured blood pressure which was less significant (e.g. intron 6 (C2071T) and exon 8 (T2617C, D240D)) or else no correlation with the measured blood pressure (e.g. intron 3 position Ins 13 + xT, T1300-1312 and intron 4 (C1451T) and intron 7 position 2544delA), as Table 2 shows.

The ECG values, which were likewise measured on the test subjects, also showed that there was a marked correlation of the Q/T intervals, which were determined for the individual test subjects, with the genotype of the test subjects with regard to the polymorphism in intron 2 at position 732/733 of the hsgk1 gene: in this connection, test subjects possessing the less frequent WT/WT genotype exhibited markedly shorter Q/T intervals than heterozygous WT/InsG test subjects, while these latter in turn exhibited significantly shorter Q/T intervals than did test subjects possessing the more frequent InsG/InsG genotype (see Table 3). Longer Q/T intervals increase the danger of contracting cardiac rhythm disturbances, such as, in particular, the long Q/T syndrome. Consequently, inverse correlations are found between the genotype of the polymorphism in intron 2 at position 732/733 of the hsgk1 gene and a predisposition for the long Q/T syndrome, on the one hand, and a predisposition for hypertension, on the other hand. These correlations can in each case be used for the diagnosis, therapy and prophylaxes of hypertension and the long Q/T syndrome.

Table 2:

SNP/DNA No.	intron 2 position insG 732^733	intron 3 position ins13+xT T1300^1312	intron 4 C1451T	intron 6 C2071T	intron 7 position delA 2544delA	exon 8 T2617C, D240D
1899	wt/wt	ins13+xT	C/C	C/T	wt/wt	T/C
2022	wt/wt	ins13+xT	C/C	C/C	wt/wt	C/C
2094	insG/wt	ins13+xT	C/C	C/C	wt/wt	T/T
1902	insG/wt	ins13+xT	C/C	T/T	wt/wt	C/C
2041	wt/wt	ins13+xT	C/C	C/C	wt/wt	C/C
2108	insG/wt	ins13+xT	C/C	C/T	wt/wt	T/C
1921	insG/wt	ins13+xT	C/C	C/T	delA/wt	C/C
2048	insG/wt	ins13+xT	C/C	T/T	wt/wt	C/C
2115	wt/wt	ins13+xT	C/C	C/T	wt/wt	T/C
1934	insG/wt	ins13+xT	C/C	C/T	wt/wt	T/C
2049	insG/wt	ins13+xT	C/C	C/T	wt/wt	C/C
2133	insG/insG	ins13+xT	C/C	T/T	wt/wt	C/C
1944	wt/wt	ins13+xT	C/C	C/T	wt/wt	C/C
2072	insG/insG	ins13+xT	C/C	T/T	wt/wt	C/C
2159	insG/wt	ins13+xT	C/C	T/T	wt/wt	C/C
1983	wt/wt	ins13+xT	C/C	C/C	wt/wt	T/C
2076	insG/wt	ins13+xT	C/C	C/T	wt/wt	T/C
2166	wt/wt	ins13+xT	C/C	C/C	wt/wt	T/C
2011	wt/wt	ins13+xT	C/C	C/C	wt/wt	T/C
2084	insG/wt	ins13+xT	C/C	C/T	wt/wt	C/C
2278	wt/wt	ins13+xT	C/C	C/C	wt/wt	T/T
2020	insG/insG	ins13+xT	C/C	T/T	wt/wt	C/C
2085	wt/wt	ins13+xT	C/C	C/T	wt/wt	T/C
2338	insG/insG	ins13+xT	C/T	T/T	wt/wt	C/C

Table 3:

Measured quantity/genotype	wt/wt	wt/ins	ins/ins	Signifi-cance
(Mean ± standard deviation)	n=7	n=14	n=7	

Systolic blood pressure	123 ± 17	116 ± 10	117 ± 15	< 0.05
Diastolic blood pressure	73 ± 14	70 ± 9	72 ± 9	n.s.
Q/T interval	403 ± 13	411 ± 17	428 ± 10	< 0.05

The sequence listings attached hereto are to be considered as an integral part of the present specification.

"Comprises/comprising" when used in this specification is taken to specify the presence of stated features, integers, steps or components but does not preclude the presence or addition of one or more other features, integers, steps or components or groups thereof.

As enclosed to IPER

Amended Patent Claims

1. The use of an isolated single-stranded or double-stranded nucleic acid comprising a fragment of the nucleic acid sequence as depicted in SEQ ID No. 1 or as depicted in SEQ ID No. 2 for diagnosing hypertension in vitro, wherein said fragment is at least 10 nucleotides/base pairs in length and said fragment comprises the polymorphism in intron 2 of the hsgk1 gene either with or without the insertion of the nucleotide G at position 732/733.

2. A kit for quantitatively diagnosing hypertension, comprising at least one isolated single-stranded or double-stranded nucleic acid as defined in claim 1.

3. A kit for quantitatively diagnosing hypertension, comprising at least one antibody directed against a region of the hsgk protein, wherein the presence of said region in the hsgk1 protein depends on the presence of an insertion of the nucleotide G at position 732/733 in intron 2 of the encoding hsgk gene.

4. A method for diagnosing hypertension in vitro, comprising the following procedural steps:
   a) withdrawing a body sample,
   b) where appropriate, isolating and/or amplifying genomic DNA, cDNA or mRNA from the body sample according to a),
   c) quantifying the alleles which possess an insertion of the nucleotide G at position 732/733 in intron 2 of the hsgk1 gene.

5. The method as claimed in claim 4, wherein the body sample from step a) is selected from the group consisting of blood, saliva, tissue and cells.

6. The method as claimed in claim 4 or 5, wherein the alleles are quantified according to step c) by directly sequencing the genomic DNA or cDNA which has been isolated from the body sample.

7. The method as claimed in any one of claims 4 to 6, wherein the alleles are quantified according to step c) by specifically hybridizing the genomic DNA or cDNA which has been isolated from the body sample.

8. The method as claimed in any one of claims 4 to 7, wherein the alleles are quantified according to step c) by means of a PCR oligo elongation assay or a ligation assay.

9. The use of the direct correlation between the overexpression or functional molecular modification of human homologues of the sgk family and the length of the Q/T interval for diagnosing the long QT syndrome in vitro.

10. The use of the single-stranded or double-stranded nucleic acid comprising the sequence of a human homologue of the sgk family or one of its fragments having a length of at least 10 nucleotides/base pairs for diagnosing the long QT syndrome in vitro.

11. The use as claimed in claim 9 or 10, wherein the human homologue of the sgk family is the hsgk1 gene.

12. The use as claimed in claim 11, wherein the nucleic acid the hsgk1 gene or of one of its fragments possesses a length of at least 10 nucleotides/base pairs and said nucleic acid comprises the polymorphism at position 732/733 in intron 2 of the hsgk1 gene either with or without the insertion of the nucleotide G.

13. The use of an antibody directed against Nedd 4-2 having the Acc. No. BAA23711 for diagnosing in vitro a predisposition for developing the long Q/T syndrome, with the antibody being directed against an epitope of the human homologue which contains the phosphorylation site either in phosphorylated form or in unphosphorylated form.

14. The kit for diagnosing the long QT syndrome, comprising antibodies which are directed against the human homologues of the sgk protein family, or single-stranded or double-stranded nucleic acid fragments which are at least 10 nucleotides/base pairs in length and which are able to hybridize, under stringent conditions, with the human homologues of the sgk gene family, or comprising these antibodies and nucleic acids jointly.

15. The kit as claimed in claim 14, wherein the human homologue of the sgk family is the hsgk1 gene.

16. The kit as claimed in claim 15, which comprises nucleic acid fragments, as specific hybridization probes, which comprise at least one of the SNPs in the hsgk1 gene in exon 8 (C2617T, D240D), in intron 6 (T2071C) or that in intron 2 at position 732/733 (6 insertion).

17. The use of a functional activator, or of a positive transcription regulator, of a human homologue of the sgk family, for lowering the Q/T interval.

18. The use as claimed in claim 17, wherein the sgk family is hsgk1.

19. The use as claimed in claim 17 or 18, wherein the functional activator or positive transcription regulator is selected from the group consisting of glucocorticoids, mineralocorticoids, aldosterone, gonadotropins and cytokines, in particular TGF-β.

20. The use of substances selected from the group consisting of glucocorticoids, mineralocorticoids, aldosterone, gonadotropins and cytokines, for producing a pharmaceutical for the therapy and/or prophylaxis of the long QT syndrome.

21. The use as claimed in claim 20, wherein the cytokines are TGF-β.

22. A pharmaceutical comprising at least one substance from the group of substances consisting of mineralocorticoids, aldosterone, gonadotropins and cytokines, for the therapy and/or prophylaxis of the long QT syndrome.

23. The pharmaceutical as claimed in claim 22, wherein the cytokines are TGF-β.

24. The use of an isolated single-stranded or double stranded nucleic acid as claimed in any one of claims 1 and 10 to 12, substantially as hereinbefore described or exemplified.

25. The use of an isolated single-stranded or double stranded nucleic acid including any new and inventive integer or combination of integers, substantially as herein described.

26. A kit as claimed in any one of claims 2, 3, and 14 to 16, substantially as hereinbefore described or exemplified.

27. A kit according to the invention including any new and inventive integer or combination of integers, substantially as herein described.

28. The method according to the invention for diagnosing hypertension in vitro, substantially as hereinbefore described or exemplified.

29. The method for diagnosing hypertension in vitro including any new and inventive integer or combination of integers, substantially as herein described.

30. The use of an antibody as claimed in claim 13, substantially as hereinbefore described or exemplified.

31. The use of an antibody according to the invention including any new and inventive integer or combination of integers, substantially as herein described.

32. The use of a functional activator as claimed in any one of claims 17 to 19, substantially as hereinbefore described or exemplified.

33. The use of a functional activator according to the invention including any new and inventive integer or combination of integers, substantially as herein described.

34. The use of substances as claimed in claim 20 or 21, substantially as hereinbefore described or exemplified.

35. The use of substances according to the invention including any new and inventive integer or combination of integers, substantially as herein described.

36 The pharmaceutical as claimed in claim 22 or 23, substantially as hereinbefore described or exemplified.

37. The pharmaceutical according to the invention including any new and inventive integer or combination of integers, substantially as herein described.

SEQUENCE LISTING

<110> Lang, Florian

<120> Verwendung eines neuen Polymorphismus im hsgk1-Gen zur Diagnose der Hypertonie und Verwendung der sgk-Genfamilie zur Diagnose und Therapie des Long-Q/T-Syndroms

<130> L62136

<160> 2

<170> PatentIn version 3.1

<210> 1
<211> 5704
<212> DNA
<213> Homo sapiens

<220>
<221> exon
<222> (36)..(155)
<223>

<220>
<221> exon
<222> (303)..(378)
<223>

<220>
<221> exon
<222> (806)..(881)
<223>

<220>
<221> exon
<222> (1317)..(1421)
<223>

<220>
<221> exon
<222> (1526)..(1609)
<223>

<220>
<221> exon
<222> (1725)..(1856)
<223>

<220>
<221> exon
<222> (2106)..(2218)
<223>

<220>
<221> exon
<222> (2560)..(2683)
<223>

<220>
<221> exon
<222> (3141)..(3236)
<223>

<220>
<221> exon
<222> (3652)..(3807)
<223>

<220>
<221> exon
<222> (3915)..(4004)
<223>

<220>
<221> exon
<222> (4349)..(5526)
<223>

<220>
<221> Intron
<222> (156)..(302)
<223>

<220>
<221> Intron
<222> (379)..(805)
<223>

<220>
<221> Intron
<222> (882)..(1316)
<223>

<220>
<221> Intron
<222> (1422)..(1525)
<223>

<220>
<221> Intron
<222> (1610)..(1724)
<223>

<220>
<221> Intron
<222> (1857)..(2105)
<223>

<220>
<221> Intron
<222> (2219)..(2559)
<223>

<220>
<221> Intron
<222> (2684)..(3140)
<223>

<220>
<221> Intron
<222> (3237)..(3651)
<223>

<220>
<221> Intron
<222> (3808)..(3914)
<223>

<220>
<221> Intron
<222> (4005)..(4348)
<223>

<220>
<221> mutation
<222> (732)..(733)
<223> Insertion of G at the position 732/733 in certain genotypes

<220>
<221> mutation
<222> (1299)..(1300)
<223> Homopolymorph insertion of 13 x T in certain genotypes

<220>
<221> mutation
<222> (1451)..(1451)
<223> C/T-exchange at the position 1451 in certain genotypes

<220>
<221> mutation
<222> (2071)..(2071)
<223> C/T-exchange at the position 2071 in Intron 6 in certain genotype
s

<220>
<221> mutation
<222> (2543)..(2544)
<223> Insertion of A at the position 2543/2544 in Intron 7 in certain genotypes

<220>
<221> mutation
<222> (2617)..(2617)
<223> T/C-exchange at the position 2617 in Exon 8 in certain genotypes (resulting in no amino acid exchange, D240D)

<400> 1
ggccgagcgc gcggcctggc gcacgatacg ccgag ccg gtc ttt gag cgc taa 53

cgt ctt tct gtc tcc ccg cgg tgg tga tga cgg tga aaa ctg agg ctg 101

cta agg gca ccc tca ctt act cca gga tga ggg gca tgg tgg caa ttc 149

tca tcg gtgagtgcag gaatcttgcg ggacttctgc tccaggagac gcaaagtgga 205

aatttttga aagtcccgga tcagattagt gtgtgtggcg ccgggacgtt atgaagccgt 265

ctaaacgttt ctttatttct cctccttcta tccacag ctt tca tga agc aga gga 320

gga tgg gtc tga acg act tta ttc aga aga ttg cca ata act cct atg 368

cat gca aac a gtaagttcag accggattga ggaaataact agtatagttt 418

gaatttgcca gcggtaaaca ttctcatcac ggcgtttatc gggaaggcga agacttcttc 478

tggggtgggg atctcatttc tccttaaatt ctaatatatt tgacacattt taaacattaa 538

agttaatttg ctgatttggc ttgaactgga gatgtaagat aaatggttcg tgttggccga 598

attcacgctt tctccatgag caacaatcct tatttctgta tttaatgggg tttattattt 658

tctttaactg actaatgtat tggggtattt tcagtttaaa cagtgaatta tcgggtagaa 718

gtcggtagag ccagaaactc acttttgatg ttggtgtgcc ccctagtggc gagctggatt 778

ctaaatcgtg ccctttattc cctgcag cc ctg aag ttc agt cca tct tga aga 831

tct ccc aac ctc agg agc ctg agc tta tga atg cca acc ctt ctc ctc 879

ca gtaagttttt gtatgtgccg tgcatctgtg gagaactgta agggagtcag 931

ttagtattcc tacattaatg gattaaaata gcatttctag aaattagtat caaggcagga 991

atgcttcatt atgcataaca gtgatataaa tatttaagta ttgagtcaga gtattatttt 1051

tatttttttc ctgggcatat tttacctcaa gtggttattt taaaaggcat atttcataaa 1111

aaggttttat ctgtctgaaa caacatgact gtgtgcagtt tccatactca tttgaaatgt 1171

gatgaaatgt agttttgaat gtttatagat gtatggtcat ttgcatcagt catttgtaga 1231

tgtaacattt tctacatcgt ttatgttata gatgtcttcc tttgaagcaa tggtattaaa 1291

agaaattcct agccaagtcc ttctc a gca aat caa cct tgg ccc gtc gtc caa 1344

tcc tca tgc taa acc atc tga ctt tca ctt ctt gaa agt gat cgg aaa 1392

ggg cag ttt tgg aaa ggt aat ttc aaa tc tgaagatctt ttggtacact 1441

tccttcatgt cctcttttat attctccctg gatgaggatc gaaaaatgat ttttttaaat 1501

tgaaatttca ggttcttcta gcaa g aca caa ggc aga aga agt gtt cta tgc 1553

agt caa agt ttt aca gaa gaa agc aat cct gaa aaa gaa aga ggt gag 1601

atg tgc tt gatggggctg gcattggcgg tagacactcc ttgaataatc 1649

ttgattctgg aatgttggtg ccagttgaac atgccactaa atctgaatcg tcattttcct 1709

aggagaagca tatta t gtc gga gcg gaa tgt tct gtt gaa gaa tgt gaa 1758

gca ccc ttt cct ggt ggg cct tca ctt ctc ttt cca gac tgc tga caa 1806

att gta ctt tgt cct aga cta cat taa tgg tgg aga ggt gag cag ggg 1854

gg atagaagtca actcttagtg tctctgcaca gcctgctttg ttttagtttg 1906

agaaaaaagt tttcaaagat ttttggtggg gagaatgtta ccagaattag catttccttc 1966

aacctgtcag gttatagtta atagattact tggggccact tcctgcagtt gttcttttgc 2026

tgtgtatgtc aaaactaatt aaattacatt gcgcaaccca gaatgacttt gttctgtctc 2086

ctgcagttgt tctaccatc t cca gag gga acg ctg ctt cct gga acc acg 2136

ggc tcg ttt cta tgc tgc tga aat agc cag tgc ctt ggg cta cct gca 2184

ttc act gaa cat cgt tta tag gta agc ctg aga g ctcttcaggc 2228

taccagtttt ggtataaagg agacgtagca ctggctgttt catagggcct taaaataatt 2288

tgtgtttatt tgcaacttgg ttcgctaaaa ccagatcccc tagcacgtga gctggcttga 2348

cttaagtgcc aaggggaac agccaagtag gattgtgcct aatccagaat agatgagcag 2408

aacaagggct ccttttttct tcactacaca actacagtga acctaaatgc ctctaatacc 2468

ttagcaatta tctttaagag gatatcttat gaagtgaaat taacttgtgc aactactttt 2528

ctttcacttt tttacagaga cttaaaacca g ag aat att ttg cta gat tca 2579

cag gga cac att gtc ctt act gat ttc gga ctc tgc aag gag aac att 2627

gaa cac aac agc aca aca tcc acc ttc tgt ggc acg ccg gag gta ggc 2675

gct gtc tt ggtttggtgc ctggtttacc cccgccttcc aagagagaga 2723

tgtacaatca tgcacttaac taccaaaaag agtaaactcc tctcagagac ttcttaatac 2783

agttcagtgc aaataaaata catttgctgt ttgatgtagc atgagaaatc ccaagtcctt 2843

ctgttccttt actgaaaagt agctgtttgt aagtaagatc tgcatcataa aaactttcta 2903

atcctaagta agagatatca agtgccagca gtttcctaaa tgtcagtaca cataggtagc 2963

cagtcaccct caaaaagtcc agcagtttta tcaggaagga atctaaagat atctatcttc 3023

caagctggct ctgggtctct cagctttttc aaactaaatg tgtggtcgtg ggattgcttg 3083

ctttcgcagg ttctaaacgc tgtttccctg gtctgttttt cagtatctcg cacctga g 3141

gtg ctt cat aag cag cct tat gac agg act gtg gac tgg tgg tgc ctg 3189

gga gct gtc ttg tat gag atg ctg tat ggc ctg gtg agt ggc aca tt 3236

gggaaccact ggaacactgc ctgctcccta caatattgcc ttcacacagc aaaagcagct 3296

aagaggcata ttggttattt tatagttcat aagaataatc acttacctgg ttcttttgtg 3356

catttcacat tttactagat aggaccacat tgaacctgtg tggtggtgaa aaactaccac 3416

ttattaacat ctacccccta ccctccacac acacacacac aaacacacac acgggttgca 3476

aagtagacac ttaaatagca agggaaaaga aagcattgag gtggggagag tttctcaaat 3536

cgagcctaat atttattgcc gtttatatct ttttctctac tggtaatgtg tgccatatga 3596

aacttccaat taagtctaaa gtaattttcc ccttctttca gccgcctttt tatag c 3652

cga aac aca gct gaa atg tac gac aac att ctg aac aag cct ctc cag 3700

ctg aaa cca aat att aca aat tcc gca aga cac ctc ctg gag ggc ctc 3748

ctg cag aag gac agg aca aag cgg ctc ggg gcc aag gat gac ttc gtg 3796

agt gat gtt tt cctgtcctcc tgggccggcc gggacgtgca ctagacctcc 3847

ctgcccttat tgaatgcacc tgtctaaatt aatcttgggt ttcttatcaa cagatggaga 3907

ttaagag t cat gtc ttc ttc tcc tta att aac tgg gat gat ctc att aat 3957

aag aag att act ccc cct ttt aac cca aat gtg gtg agt atc tgt ct 4004

ctcttctaag tatagagaag ccaagcgatt tattttaatt cagaattgtc tgggggaggg 4064

ttggaaggaa tacattggca gatgttttct ccataaacct gttattttac ctacatagac 4124

acatttatca attcgaagca ccaaaaggca acaagtgaac attattctta tgtttaactg 4184

tgtgtagcct tttgagattt tgtgcttgaa gtgggtgatt atggaagttg atataagact 4244

taaacttggt atttaaagcc tggtcaagat ttccctgtcc tgtgtctagt gtgagttctt 4304

gacaagagtg tttttccctt cccgtcacag agtgggccca acga g cta cgg cac 4358

ttt gac ccc gag ttt acc gaa gag cct gtc ccc aac tcc att ggc aag 4406

tcc cct gac agc gtc ctc gtc aca gcc agc gtc aag gaa gct gcc gag 4454

gct ttc cta ggc ttt tcc tat gcg cct ccc acg gac tct ttc ctc tga 4502

acc ctg tta ggg ctt ggt ttt aaa gga ttt tat gtg tgt ttc cga atg 4550

ttt tag tta gcc ttt tgg tgg agc cgc cag ctg aca gga cat ctt aca 4598

aga gaa ttt gca cat ctc tgg aag ctt agc aat ctt att gca cac tgt 4646

tcg ctg gaa ttt ttt gaa gag cac att ctc ctc agt gag ctc atg agg 4694

ttt tca ttt tta ttc ttc ctt cca acg tgg tgc tat ctc tga aac gag 4742

cgt tag agt gcc gcc tta gac gga ggc agg agt ttc gtt aga aag cgg 4790

acc tgt tct aaa aaa ggt ctc ctg cag atc tgt ctg ggc tgt gat gac 4838

gaa tat tat gaa atg tgc ctt ttc tga aga gat tgt gtt agc tcc aaa 4886

gct ttt cct atc gca gtg ttt cag ttc ttt att ttc cct tgt gga tat 4934

gct gtg tga acc gtc gtg tga gtg tgg tat gcc tga tca cag atg gat 4982

ttt gtt ata agc atc aat gtg aca ctt gca gga cac tac aac gtg gga 5030

cat tgt ttg ttt ctt cca tat ttg gaa gat aaa ttt atg tgt aga ctt 5078

ttt tgt aag ata cgg tta ata act aaa att tat tga aat ggt ctt gca 5126

atg act cgt att cag atg cct aaa gaa agc att gct gct aca aat att 5174

tct att ttt aga aag ggt ttt tat gga cca atg ccc cag ttg tca gtc 5222

aga gcc gtt ggt gtt ttt cat tgt tta aaa tgt cac ctg taa aat ggg 5270

cat tat tta tgt ttt ttt ttt tgc att cct gat aat tgt atg tat tgt 5318

ata aag aac gtc tgt aca ttg ggt tat aac act agt ata ttt aaa ctt 5366

aca ggc tta ttt gta atg taa acc acc att tta atg tac tgt aat taa 5414

cat ggt tat aat acg tac aat cct tcc ctc atc cca tca cac aac ttt 5462

ttt tgt gtg tga taa act gat ttt ggt ttg caa taa aac ctt gaa aaa 5510

tat tta cat ata ttg t gtcatgtgtt attttgtata ttttggttaa gggggtaatc 5566

atgggttagt ttaaaattga aaaccatgaa aatcctgctg taatttcctg cttagtggtt 5626

tgctccaaca gcagtggttt ctgactccag ggagtatagg atggcttaag ccaccacgtc 5686

caggccttta gcagcatt 5704

<210> 2
<211> 5701
<212> DNA
<213> Homo sapiens

<220>
<221> exon
<222> (36)..(155)
<223>

<220>
<221> exon
<222> (303)..(378)
<223>

<220>
<221> exon
<222> (807)..(882)
<223>

<220>
<221> exon
<222> (1318)..(1422)
<223>

<220>
<221> exon
<222> (1527)..(1610)
<223>

<220>
<221> exon
<222> (1726)..(1857)
<223>

<220>
<221> exon
<222> (2107)..(2219)
<223>

<220>
<221> exon
<222> (2561)..(2684)
<223>

<220>
<221> exon
<222> (3142)..(3237)
<223>

<220>
<221> exon
<222> (3653)..(3808)
<223>

<220>
<221> exon
<222> (3916)..(4005)
<223>

<220>
<221> exon
<222> (4350)..(5527)
<223>

<220>
<221> Intron
<222> (156)..(302)
<223>

<220>
<221> Intron
<222> (379)..(806)
<223>

<220>
<221> Intron
<222> (883)..(1317)
<223>

<220>
<221> Intron
<222> (1423)..(1526)
<223>

<220>
<221> Intron
<222> (1611)..(1725)
<223>

<220>
<221> Intron
<222> (1858)..(2106)
<223>

<220>
<221> Intron
<222> (2220)..(2560)
<223>

<220>
<221> Intron
<222> (2685)..(3141)
<223>

<220>
<221> Intron
<222> (3238)..(3652)
<223>

<220>
<221> Intron
<222> (3809)..(3915)
<223>

<220>
<221> Intron
<222> (4006)..(4349)
<223>

<220>
<221> mutation
<222> (733)..(733)
<223> Deletion of the nucleotide G at the position 733 in certain genot
ypes

<220>
<221> mutation
<222> (1300)..(1301)
<223> Homopolymorph insertion of 13 x T in certain genotypes

<220>
<221> mutation
<222> (1452)..(1452)
<223> C/T-exchange at the position 1452 in certain genotypes

<220>
<221> mutation
<222> (2072)..(2072)

<223> C/T-exchange at the position 2072 in Intron 6 in certain genotype
s

<220>
<221> mutation
<222> (2544)..(2545)
<223> Insertion of A at the position 2544/2545 in Intron 7 in certain g
enotypes

<220>
<221> mutation
<222> (2618)..(2618)
<223> T/C-exchange at the position 2618 in Exon 8 in certain genotypes
(resutling in no amino acid exchange, D240D)

<400> 2
ggccgagcgc gcggcctggc gcacgatacg ccgag ccg gtc ttt gag cgc taa 53

cgt ctt tct gtc tcc ccg cgg tgg tga tga cgg tga aaa ctg agg ctg 101

cta agg gca ccc tca ctt act cca gga tga ggg gca tgg tgg caa ttc 149

tca tcg gtgagtgcag gaatcttgcg ggacttctgc tccaggagac gcaaagtgga 205

aatttttga aagtcccgga tcagattagt gtgtgtggcg ccgggacgtt atgaagccgt 265

ctaaacgttt ctttatttct cctccttcta tccacag ctt tca tga agc aga gga 320

gga tgg gtc tga acg act tta ttc aga aga ttg cca ata act cct atg 368

cat gca aac a gtaagttcag accggattga ggaaataact agtatagttt 418

gaatttgcca gcggtaaaca ttctcatcac ggcgtttatc gggaaggcga agacttcttc 478

tggggtgggg atctcatttc tccttaaatt ctaatatatt tgacacattt taaacattaa 538

agttaatttg ctgatttggc ttgaactgga gatgtaagat aaatggttcg tgttggccga 598

attcacgctt tctccatgag caacaatcct tatttctgta tttaatgggg tttattattt 658

tctttaactg actaatgtat tggggtattt tcagtttaaa cagtgaatta tcgggtagaa 718

gtcggtagag ccaggaaact cacttttgat gttggtgtgc cccctagtgg cgagctggat 778

tctaaatcgt gccctttatt ccctgcag cc ctg aag ttc agt cca tct tga 829

aga tct ccc aac ctc agg agc ctg agc tta tga atg cca acc ctt ctc 877

ctc ca gtaagttttt gtatgtgccg tgcatctgtg gagaactgta agggagtcag 932

ttagtattcc tacattaatg gattaaaata gcatttctag aaattagtat caaggcagga 992

atgcttcatt atgcataaca gtgatataaa tatttaagta ttgagtcaga gtattatttt 1052

tatttttttc ctgggcatat tttacctcaa gtggttattt taaaaggcat atttcataaa 1112

aaggttttat ctgtctgaaa caacatgact gtgtgcagtt tccatactca tttgaaatgt 1172

gatgaaatgt agttttgaat gtttatagat gtatggtcat ttgcatcagt catttgtaga 1232

tgtaacattt tctacatcgt ttatgttata gatgtcttcc tttgaagcaa tggtattaaa 1292

agaaattcct agccaagtcc ttctc a gca aat caa cct tgg ccc gtc gtc caa 1345

tcc tca tgc taa acc atc tga ctt tca ctt ctt gaa agt gat cgg aaa 1393

ggg cag ttt tgg aaa ggt aat ttc aaa tc tgaagatctt ttggtacact 1442

tccttcatgt cctcttttat attctccctg gatgaggatc gaaaaatgat ttttttaaat 1502

tgaaatttca ggttcttcta gcaa g aca caa ggc aga aga agt gtt cta tgc 1554

agt caa agt ttt aca gaa gaa agc aat cct gaa aaa gaa aga ggt gag 1602

atg tgc tt gatggggctg gcattggcgg tagacactcc ttgaataatc 1650

ttgattctgg aatgttggtg ccagttgaac atgccactaa atctgaatcg tcattttcct 1710

aggagaagca tatta t gtc gga gcg gaa tgt tct gtt gaa gaa tgt gaa 1759

gca ccc ttt cct ggt ggg cct tca ctt ctc ttt cca gac tgc tga caa 1807

att gta ctt tgt cct aga cta cat taa tgg tgg aga ggt gag cag ggg 1855

gg atagaagtca actcttagtg tctctgcaca gcctgctttg ttttagtttg 1907

agaaaaaagt tttcaaagat ttttggtggg gagaatgtta ccagaattag catttccttc 1967

aacctgtcag gttatagtta atagattact tggggccact tcctgcagtt gttcttttgc 2027

tgtgtatgtc aaaactaatt aaattacatt gcgcaaccca gaatgacttt gttctgtctc 2087

ctgcagttgt tctaccatc t cca gag gga acg ctg ctt cct gga acc acg 2137

ggc tcg ttt cta tgc tgc tga aat agc cag tgc ctt ggg cta cct gca 2185

ttc act gaa cat cgt tta tag gta agc ctg aga g ctcttcaggc 2229

taccagtttt ggtataaagg agacgtagca ctggctgttt catagggcct taaaataatt 2289

tgtgtttatt tgcaacttgg ttcgctaaaa ccagatcccc tagcacgtga gctggcttga 2349

cttaagtgcc aaggggggaac agccaagtag gattgtgcct aatccagaat agatgagcag 2409

aacaagggct cctttttttct tcactacaca actacagtga acctaaatgc ctctaatacc 2469

ttagcaatta tctttaagag gatatcttat gaagtgaaat taacttgtgc aactactttt 2529

ctttcacttt tttacagaga cttaaaacca g ag aat att ttg cta gat tca 2580

cag gga cac att gtc ctt act gat ttc gga ctc tgc aag gag aac att 2628

gaa cac aac agc aca aca tcc acc ttc tgt ggc acg ccg gag gta ggc 2676

gct gtc tt ggtttggtgc ctggtttacc cccgccttcc aagagagaga 2724

tgtacaatca tgcacttaac taccaaaaag agtaaactcc tctcagagac ttcttaatac 2784

agttcagtgc aaataaaata catttgctgt ttgatgtagc atgagaaatc ccaagtcctt 2844

ctgttccttt actgaaaagt agctgtttgt aagtaagatc tgcatcataa aaactttcta 2904

atcctaagta agagatatca agtgccagca gtttcctaaa tgtcagtaca cataggtagc 2964

cagtcaccct caaaaagtcc agcagtttta tcaggaagga atctaaagat atctatcttc 3024

caagctggct ctgggtctct cagctttttc aaactaaatg tgtggtcgtg ggattgcttg 3084

ctttcgcagg ttctaaacgc tgtttccctg gtctgttttt cagtatctcg cacctga g 3142

gtg ctt cat aag cag cct tat gac agg act gtg gac tgg tgg tgc ctg 3190

gga gct gtc ttg tat gag atg ctg tat ggc ctg gtg agt ggc aca tt 3237

gggaaccact ggaacactgc ctgctcccta caatattgcc ttcacacagc aaaagcagct 3297

aagaggcata ttggttattt tatagttcat aagaataatc acttacctgg ttcttttgtg 3357

catttcacat tttactagat aggaccacat tgaacctgtg tggtggtgaa aaactaccac 3417

ttattaacat ctacccccta ccctccacac acacacacac aaacacacac acgggttgca 3477

aagtagacac ttaaatagca agggaaaaga aagcattgag gtggggagag tttctcaaat 3537

cgagcctaat atttattgcc gtttatatct ttttctctac tggtaatgtg tgccatatga 3597

aacttccaat taagtctaaa gtaattttcc ccttctttca gccgcctttt tatag c 3653

cga aac aca gct gaa atg tac gac aac att ctg aac aag cct ctc cag 3701

ctg aaa cca aat att aca aat tcc gca aga cac ctc ctg gag ggc ctc 3749

ctg cag aag gac agg aca aag cgg ctc ggg gcc aag gat gac ttc gtg 3797

agt gat gtt tt cctgtcctcc tgggccggcc gggacgtgca ctagacctcc 3848

ctgcccttat tgaatgcacc tgtctaaatt aatcttgggt ttcttatcaa cagatggaga 3908

ttaagag t cat gtc ttc ttc tcc tta att aac tgg gat gat ctc att aat 3958

aag aag att act ccc cct ttt aac cca aat gtg gtg agt atc tgt ct 4005

ctcttctaag tatagagaag ccaagcgatt tattttaatt cagaattgtc tgggggaggg 4065

ttggaaggaa tacattggca gatgttttct ccataaacct gttattttac ctacatagac 4125

acatttatca attcgaagca ccaaaaggca acaagtgaac attattctta tgtttaactg 4185

tgtgtagcct tttgagattt tgtgcttgaa gtgggtgatt atggaagttg atataagact 4245

taaacttggt atttaaagcc tggtcaagat ttccctgtcc tgtgtctagt gtgagttctt 4305

gacaagagtg tttttccctt cccgtcacag agtgggccca acga g cta cgg cac 4359

ttt gac ccc gag ttt acc gaa gag cct gtc ccc aac tcc att ggc aag 4407

tcc cct gac agc gtc ctc gtc aca gcc agc gtc aag gaa gct gcc gag 4455

gct ttc cta ggc ttt tcc tat gcg cct ccc acg gac tct ttc ctc tga 4503

acc ctg tta ggg ctt ggt ttt aaa gga ttt tat gtg tgt ttc cga atg 4551

ttt tag tta gcc ttt tgg tgg agc cgc cag ctg aca gga cat ctt aca 4599

aga gaa ttt gca cat ctc tgg aag ctt agc aat ctt att gca cac tgt 4647

tcg ctg gaa ttt ttt gaa gag cac att ctc ctc agt gag ctc atg agg 4695

ttt tca ttt tta ttc ttc ctt cca acg tgg tgc tat ctc tga aac gag 4743

cgt tag agt gcc gcc tta gac gga ggc agg agt ttc gtt aga aag cgg 4791

acc tgt tct aaa aaa ggt ctc ctg cag atc tgt ctg ggc tgt gat gac 4839

gaa tat tat gaa atg tgc ctt ttc tga aga gat tgt gtt agc tcc aaa 4887

gct ttt cct atc gca gtg ttt cag ttc ttt att ttc cct tgt gga tat 4935

gct gtg tga acc gtc gtg tga gtg tgg tat gcc tga tca cag atg gat 4983

ttt gtt ata agc atc aat gtg aca ctt gca gga cac tac aac gtg gga 5031

cat tgt ttg ttt ctt cca tat ttg gaa gat aaa ttt atg tgt aga ctt 5079

ttt tgt aag ata cgg tta ata act aaa att tat tga aat ggt ctt gca 5127

atg act cgt att cag atg cct aaa gaa agc att gct gct aca aat att 5175

tct att ttt aga aag ggt ttt tat gga cca atg ccc cag ttg tca gtc 5223

aga gcc gtt ggt gtt ttt cat tgt tta aaa tgt cac ctg taa aat ggg 5271

cat tat tta tgt ttt ttt ttt tgc att cct gat aat tgt atg tat tgt 5319

ata aag aac gtc tgt aca ttg ggt tat aac act agt ata ttt aaa ctt 5367

aca ggc tta ttt gta atg taa acc acc att tta atg tac tgt aat taa 5415

cat ggt tat aat acg tac aat cct tcc ctc atc cca tca cac aac ttt 5463

ttt tgt gtg tga taa act gat ttt ggt ttg caa taa aac ctt gaa aaa 5511

tat tta cat ata ttg t gtcatgtgtt attttgtata ttttggttaa gggggtaatc 5567

atgggttagt ttaaaattga aaaccatgaa aatcctgctg taatttcctg cttagtggtt 5627

tgctccaaca gcagtggttt ctgactccag ggagtatagg atggcttaag ccaccacgtc 5687

caggccttta gcag 5701