(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 226 169 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
14.03.2007 Bulletin 2007/11

(21) Application number: **00972884.1**

(22) Date of filing: **31.10.2000**

(51) Int Cl.:
C07K 14/47 *(2006.01)* ***C12N 15/12*** *(2006.01)*
C12Q 1/68 *(2006.01)*

(86) International application number:
PCT/EP2000/010745

(87) International publication number:
WO 2001/032702 (10.05.2001 Gazette 2001/19)

(54) MOLECULAR STRUCTURE OF RHD NEGATIVE LOCUS

MOLEKULARSTRUKTUR DES RHD-NEGATIV GENORTES

STRUCTURE MOLECULAIRE DE i RDH /i NEGATIF

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **02.11.1999 EP 99121686**
31.05.2000 EP 00111696

(43) Date of publication of application:
31.07.2002 Bulletin 2002/31

(60) Divisional application:
06026286.2

(73) Proprietor: **DRK Blutspendedienst Baden-Württemberg GGmbH**
89081 Ulm (DE)

(72) Inventors:
- **FLEGEL, Willy, A.**
64807 Dieburg (DE)
- **WAGNER, Franz, F.**
89075 Ulm (DE)

(74) Representative: **Vossius & Partner**
Siebertstrasse 4
81675 München (DE)

(56) References cited:
WO-A-98/45712
WO-A-99/37763
WO-A-99/47555
US-A- 5 972 602

- **CARRITT, B. ET AL.: "Evolution of the Rh (rhesus) blood group genes: a 50 year old predication (partially) fulfilled" HUMAN MOLECULAR GENETICS., vol. 6, no. 6, 1997, pages 843-850, XP000998694 OXFORD UNIVERSITY PRESS, SURREY., GB ISSN: 0964-6906 cited in the application**
- **CARRITT, B. ET AL.: "Rh null phenotypes are not due to a gross deletion and can occur on different Rh genetic backgrounds" ANNALS OF HUMAN GENETICS, vol. 57, no. 4, 1993, pages 273-279, XP000926510**
- **OKUDA H ET AL: "SEQUENCE ANALYSIS OF THE SPACER REGION BETWEEN THE RHD AND RHCE GENES" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS., vol. 263, 24 September 1999 (1999-09-24), pages 378-383, XP000998148 ACADEMIC PRESS INC. ORLANDO, FL., US ISSN: 0006-291X cited in the application**
- **HU, G.: "Human small membrane protein 1 (SMP1) mRNA" EMBL SEQUENCE DATABASE, XP002170737 HEIDELBERG DE cited in the application**
- **LE VAN KIM CAROLINE ET AL: "Molecular cloning and primary structure of the human blood group RhD polypeptide." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 89, no. 22, 1992, pages 10925-10929, XP002134526 1992 ISSN: 0027-8424**
- **CHERIF-ZAHAR B ET AL: "Organization of the gene (RHCE) encoding the human blood group RhCcEe antigens and characterization of the promoter region." GENOMICS, vol. 19, no. 1, 1994, pages 68-74, XP000926511 ISSN: 0888-7543**

EP 1 226 169 B1

(Cont. next page)

- WAGNER, F.F. ET AL.: "RHD gene deletion occured in the Rhesus box" BLOOD, vol. 29, no. 12, 15 June 2000 (2000-06-15), pages 3662-3668, XP000926502 WASHINGTON US
- SINGLETON BELINDA K ET AL: "The presence of an RHD pseudogene containing a 37 base pair duplication and a nonsense mutation in Africans with the Rh D-negative blood group phenotype." BLOOD, vol. 95, no. 1, 1 January 2000 (2000-01-01), pages 12-18, XP001011309 ISSN: 0006-4971
- BLUNT, T. ET AL.: "Serotype switching in a partially deleted RHD gene" VOX SANGUINIS, vol. 67, no. 4, 1994, pages 397-401, XP000926512
- AVENT ND ET AL: "Evidence of genetic diversity underlying Rh D-, weak D (Du), and partial D phenotypes as determined by multiplex polymerase chain reaction analysis of the RHD gene" BLOOD, W.B. SAUNDERS, PHILADELPHIA, VA, US, vol. 89, no. 7, 1 April 1997 (1997-04-01), pages 2568-2577, XP002108680 ISSN: 0006-4971
- DATABASE EM_HUM [Online] EMBL; 13 August 1998 (1998-08-13) COBLEY. V. ET AL.: "Human DNA sequence clone RP3-469D22 on chromosome 1p35.1-36.13, contains the 5'part of the gene for RHCE." retrieved from EBI, accession no. HS469D22 Database accession no. AL031284 XP002178290 cited in the application

Description

[0001] The present invention relates to a nucleic acid molecular structure representing either the hybrid *Rhesus box,* or the upstream *Rhesus box* or the downstream *Rhesus box.* The invention further relates to the detection of nucleic acid molecules carrying a deletion of the RhD gene. The invention also relates to oligonucleotides, as identified in the claims. Additionally, the invention rotates to kits comprising or employing the above recited compounds of the invention.

[0002] The Rhesus D antigen (ISBT 004.001; RH1) is the most important blood group antigen determined by a protein. Anti-D remains the leading cause of hemolytic disease of the newborn (Filbey, Acta Obstet Gynecol Scand, 74:687,1995; Bowman, J, Semin Perinatol 21:39, 1997). Depending on the population, 3% to 25% of whites lack the antigen D (Mourant, The distribution of the human blood groups and other polymorphisms, London, Oxford University Press, 1976). Anti-D immunizations can occur readily in D-negative recipients (Urbaniak, Transfusion 21:64, 1981).

[0003] The antigens of the RH blood group are carried by proteins coded by two genes, *RHD* and *RHCE,* that are located at chromosomal position 1 p34.1 - 1 p36 (Cherif-Zahar, Hum. Genet. 86: 398, 1991; MacGeoch, Cytogenet. Cell Genet. 59:261, 1992) probably within less than a 450,000 base pair (bp) distance (Carritt, Hum. Mol. Genet. 6:843, 1997). Both genes encompass ten exons and their structures are highly homologous. The relative orientation of the genes, their distance, and the possibility of interspersed other genes were unknown (Flegel, Transfus. Med. 8:281, 1998). Very recently, Okuda et al. (Okuda, Biochem. Biophys. Res. Commun. 263:378, 1999) reported a sequence of about 11,000 bp, which was thought to represent the DNA segment between *RHD* and *RHCE.*

[0004] In whites, the vast majority of D-negative haplotypes is due to a deletion of the *RHD* gene: This deletion spans the whole *RHD* gene, because *RHD*-specific sequences ranging from exon 1 to the 3' untranslated region are absent (Gassner, Transfusion 37:1020, 1997). The exact extent of the deletion was uncertain, leaving open the possibility that neighboring genes were also affected.

[0005] The identification of the *RHD* gene as the molecular basis of the D antigen allowed RhD phenotype prediction by DNA typing (Flegel, Transfus. Med. 8:281, 1998; Lo, Lancet 341:1147, 1993). However, since the structure of the prevalent D-negative haplotype is unknown, a specific detection of the *RHD* deletion remained impossible and the discrimination of RHD^+/RHD^+ homozygous from RHD^+/RHD^- heterozygous individuals relied on indirect methods. This discrimination is of clinical interest in particular, because in D-negative mothers with an anti-D, the risk of an affected child is 100% with a RHD^+/RHD^+ father, but only 50% with a RHD^+/RHD^- father.

[0006] Several indirect approaches have been applied to determine the zygosity: (i) a simple guess based on the phenotype is correct in about 95% of cases, (ii) determination of the D antigen density which can be confounded by factors such as the presence of the C antigen, and (iii) several methods involving the parallel quantitative amplification of *RHD*- and *RHCE*-specific sequences (Cossu, Electrophoresis 17:1911, 1996; Döscher, Infusionsther. Transfusionsmed. 26(suppl 1):31, 1999 (abstr.)). These elaborate techniques may not be practical in routine laboratories. In addition, several investigators identified polymorphisms in the *RHCE* gene or neighboring sequences genetically linked to the lack of the *RHD* gene (Carritt, Hum. Mol. Genet. 6:843, 1997; Huang, Am. J. Hum. genet. 58: 133, 1996; Fujiwara, Hum: genet. 104:301, 1999; Onda, Gene 159:225,1995). This indirect approach relied on the linkage disequilibrium associating the *RHD* deletion with a polymorphism.

[0007] Furthermore, the utility of the *RHD* PCR is limited by the incomplete knowledge of presumably rare *RHD* positive alleles in RhD-negative. *RHD* positive alleles in RhD negative are caused by *RHD-CE-D* hybrid genes (Huang, Blood 88:2326-33, 1996; Faas, Transfusion 37:38-44, 1997, Faas, Transfusion 36:506-11, 1996), nonsense-mutations (Avent, Blood 89:2568-77, 1997), frameshifts (Andrews, Blood 92:1839-40, 1998; Cherif-Zahar Br. J. Haematol. 102:1263-70, 1998), or pseudogenes (Singleton, Blood 95:12-8, 2000). Such alleles are frequent in Africans (Faas, Transfusion 37: 38-44, 1997, Singleton, Blood 95:12-18, 2000) and Asians (Okuda, J. Clin. Invest. 100:373-9,1997) but rare in whites. Nevertheless, recent analyses (Avent, Blood 89:2568-77, 1997; Flegel, Transfus. Med. 8:281-302, 1998) suggested that even for whites these alleles are likely the leading cause of incorrect Rh phenotype prediction. Several observations in whites (Avent, Blood 89:2568-77, 1997; Hyland, Blood 84:321-4, 1994) indicated that these alleles clustered in the Cde and cdE haplotypes. Portions of the present invention have been published in F.F. Wagner et al. BLOOD 95 (2000), 3662-3668.

[0008] The most direct approach for analyzing the *RHD* locus on the molecular level would be PCR amplification spanning the *RHD* deletion site. Such an assay has, so far, not been available because the structure of the *RHD* locus in RhD positives and RhD negatives was incompletely understood.

[0009] Accordingly, the technical problem underlying the present invention was to provide means and methods for a reliable, nucleic acid based analysis of the Rhesus D locus. These means and methods should be, inter alia, suitable for the detection and/or discrimination of RHD^+/RHD^+ and RHD^+/RHD^- individuals.

[0010] The solution to said technical problem is achieved by providing the embodiments characterized in the claims.

[0011] Thus, the invention relates to a nucleic acid molecular structure representing either the hybrid *Rhesus box,* or the upstream *Rhesus box* or the downstream *Rhesus box,* the sequences of which are shown in Figures 7 to 9.

[0012] In the context of the present invention, the term "nucleic acid molecular structure" is defined as a linear DNA-

segment that comprises, in its broadest meaning, the combination of the above mentioned *Rhesus boxes* that co-determine said *Rhesus* gene locus. DNA sequences that give rise to the molecular structure of the invention include the following: The nucleotide sequence structure consists of the hybrid *Rhesus box* or either of the two *Rhesus boxes*.

[0013] The following sequences represent embodiments contained in the nucleic acid molecular structure of the invention.

[0014] The hybrid *Rhesus box* is represented in GenBank accession number AL252313 bases 33 to 9,180.

[0015] The two *Rhesus boxes* with intervening *RHD* gene consists of the upstream *Rhesus box,* represented in GenBank accession number AL252311 bases 34 to 9,175, the *RHD* gene and the downstream *Rhesus box* represented in GenBank accession number AL252312 bases 23 to 9,177 (see Fig. 7 to 9).

[0016] Whereas the upstream *Rhesus box* is located 5' of the *RHD* gene, the downstream *Rhesus box* is located between the *RHD* and *SMP1* genes in this structure of the present invention. The term "nucleic acid molecular structure" relates to DNA segments solely comprising the referenced *Rhesus boxes*.

[0017] For a better understanding of the claimed subject-matter, it is referred to figures 1 and 6, infra.

[0018] In accordance with the present invention, the term "nucleic acid molecular structure" comprises also any feasible derivative of the above referenced nucleic acid structure to which a nucleic acid probe may hybridize. In other words, the structure of the invention may be prepared by synthetic or semisynthetic means and thus consist of or comprise peptide nucleic acid. Said term also bears the meaning of a nucleic acid molecule.

[0019] In accordance with the present invention, the term *"Rhesus box"* describes upstream and downstream DNA segments that flank the *RHD* gene on the 5' and 3' end. The three *Rhesus boxes* are defined by their nucleotide sequences. The hybrid *Rhesus box* is represented in one embodiment in GenBank accession number AL252313 bases 33 to 9,180. The two *Rhesus boxes* with intervening *RHD* gene consists of the upstream *Rhesus box,* represented in one embodiment in GenBank accession number AL252311 1 bases 34 to 9,175 and the downstream *Rhesus box* represented in one embodiment in GenBank accession number AL252312 bases 23 to 9,177. As exemplified in the appended examples the *Rhesus boxes* are preferably approximately 9000bp long, having 98,6% identity and identical orientation. According to the present invention the upstream and downstream *Rhesus boxes* are at least 95% homologous. The length of these *Rhesus boxes* may vary. It is expected that the length of these *Rhesus boxes* may vary, because, among other structural features, multiple repetitive elements, some of them are organized in tandem arrays, are known to be prone to (array) elongation and deletion events. If such events occur the length of the *Rhesus boxes* may shrink to less than 1,000 nucleotides length or extend to more than 20,000 nucleotides length.

[0020] In accordance with the present invention the term "identity" refers to the determination of sequence identity using suitable alignment programs, such as BLAST.

[0021] As has been pointed out above, the diagnostic analysis of *RHD* negatives on the molecular level has so far been hampered by the fact, that the overall structure of the *RHD*/*RHCE* loci was unknown. It has now been surprisingly found, that the two genes, *RHD* and *RHCE,* have opposite orientation and face each other with their 3' ends. In accordance with the present invention it has further been found that the *RHD* gene is surrounded by two highly homologous *Rhesus boxes.* The physical distance between *RHD* and *RHCE* is about 30,000 bp and is filled with a *Rhesus box* and the *SMP1* gene. The breakpoints of the *RHD* deletion in the prevalent *RHD* negative haplotypes are located in the 1,463 bp identity region of the *Rhesus boxes.* Similar *RHD* deletion events may involve any other region within the highly homologous *Rhesus boxes.* Hence, a region of a breakpoint comprising an *RHD* deletion other than the common *RHD* deletion may be anticipated to occur anywhere within the *Rhesus boxes* as defined above.

[0022] The opposite orientation of the two *RH* genes explains the different character of hybrid genes in the MNS and RH blood group: The glycophorin genes encoding the MNS antigens occur in the same orientation (Onda, Gene 159: 225, 1995), and many recombinations may be explained as unequal crossing over resulting in single hybrid genes (Blumenfeld, Hum. Mutat 6:1999, 1995). Based on the surprising findings referred to above, the events on the molecular level that lead to *RHD* negatives can now be more fully understood. In the *RH* locus, the inversely oriented sequences are unlikely to trigger unequal crossing over, and if this event occurred, no functional hybrid gene would result. The conclusion that unequal crossing over at the *RH* gene locus is unlikely may explain that most *RH* hybrid genes are of *RHD-CE-D* or *RHCE-D-CE* type and involve stretches of homologous DNA positioned *in cis* as noted previously (Wagner, Blood 91:2157, 1998). Currently, the *RH* gene system is the only well investigated gene locus where the two genes have opposite orientation, rendering it a model system for the evolution of neighboring, oppositely oriented genes that are frequent throughout genomes.

[0023] Based on the structure of the *RH* gene locus (Fig. 1), a parsimonious model for the *RHD* gene deletion event is proposed (Fig. 6). Although the applicant does not wish to be bound to theory, the following is believed with regard to the generation of RhD negative. The *RHD* deletion may be explained by unequal crossing over triggered by the highly homologous *Rhesus boxes* embracing the *RHD* gene. The hybrid-type *Rhesus box* of RHD-negatives arises, when a crossover leading to a deletion event involving a breakpoint region within the identity region of the upstream and downstream *Rhesus boxes* takes place. Thus, the hybrid *RHD* box is characterized by a 5' portion derived from the upstream *RHD* box fused to a 3' portion from the downstream *RHD* box. In one preferred embodiment the breakpoint region is

903bp long. The sequence of this preferred hybrid *Rhesus box* is depicted in figure 4. In the specific embodiments described in the examples, said 903 bp breakpoint region in the *Rhesus boxes* is located in a 1,463 bp stretch of 99.9% homology resembling a THE-1B human transposable element and a L2 repetitive DNA element (Fig. 3). Interestingly, the >60,000 bp DNA segment that is deleted in the *RHD* negative haplotype consisted only of and contained all sequences that are duplicated in the *RHD* positive haplotype.

[0024] The findings of the present invention referred to herein above allow for the establishment of a number of easy to do or refined methods for the analysis of the genotype of an individual with regard to the *RH* gene locus. Examples of such methods are provided herein below.

[0025] While the molecular mechanism resulting in the prevalent *RHD* negative haplotype is now apparent, it is less clear how the much older duplication event gave rise to the structure of the *RH* genes in *RHD* positives. The duplication of the *Rhesus box* and the *RH* genes probably occurred as a single event, because the overall homology of the two *Rhesus boxes* is very similar to that of the *RH* genes. Without being bound by theory, it is tempting to speculate that the *RHD* duplication originate in causal connection with the insertion of the near full-length THE-1B transposon-like human element in duplicate. However, the open reading frame of the THE-1 B element probably was non-functional at the time of the duplication.

[0026] In a preferred embodiment of the present invention, said nucleic acid molecular structure is representative of the common *RHD* negative haplotypes as defined in the appended claims.

[0027] According to the present invention, the term "is representative of" relates to a nucleic acid molecular structure comprising all sequential and structural features to relate said structure to a group of molecular structures sharing said features. In the above preferred embodiment, said features give rise to the common *RHD* negative haplotype. In the present context this means preferably the deletion of the *RHD* gene encompassing the whole *RHD* gene and its 5' region, which are located between the upstream *Rhesus box* and the downstream *Rhesus box.*

[0028] In the present context this could also mean, for example, that all structures sharing a nonsense mutation, missense mutation, splice site mutation, partial deletion, partial insertion, partial inversion or a combination thereof within the *RHD* gene, which terminates or obliterates the expression of a protein product of the *RHD* gene, are representative of the *RHD* negative haplotype.

[0029] The term "haplotype" relates to a series of linked alleles within a defined region on a single maternal or paternal chromosome.

[0030] The term "common *RHD* negative haplotype" refers to any RhD antigen negative haplotype that comprises a hybrid *Rhesus box.* Preferably the DNA segment encompassing the whole *RHD* gene and its 5' region, which are located between the upstream *Rhesus box* and the downstream *Rhesus box,* is deleted.

[0031] The invention relates to a nucleic acid molecular structure, dubbed *Rhesus box,* which is flanking the breakpoint region of the *RHD* deletion in the common *RHD* negative haplotypes.

[0032] In accordance with the present invention the term "breakpoint region of the *RHD* deletion" describes a distinct DNA segment that is involved in an *RHD* deletion. As has been pointed out above, said deletion may be the result of an unequal crossing over event involving both the upstream and downstream *Rhesus boxes,* deleting interspersed sequences and finally giving rise to a nucleic acid molecular structure (the referenced *Rhesus boxes* for a better delimitation from the upstream and downstream *Rhesus box* also referred to as hybrid *Rhesus box*) wherein the 5' portion of the upstream *RHD* box is in close spatial proximity to the 3' portion of the downstream *RHD* box. As mentioned above and depicted in figures 6 and 3 this region can preferably be 903bp-long and be located in a 1,463bp stretch within the *Rhesus boxes*, having 99,9 % homology in this segment. In another preferred alternative said region is located downstream from said 903 bp fragment but is still contained within the 1463 bp stretch. Preferably, said fragment is 556 to 560bp long. The actual breakpoint may vary such that the contribution of the upstream *Rhesus box* and the downstream *Rhesus box* are different in different individuals. However, in accordance with the present invention, the breakpoint in any case occurs in the upstream and downstream *Rhesus boxes.*

[0033] The hybrid *Rhesus box* is particularly useful for the analysis of *RHD*-negative haplotypes. For example, oligonucleotides may be employed that hybridize to nucleic acid sequences comprising the breakpoint which arose as a result of the *RHD* deletion. It is to be understood that such oligonucleotides need to hybridize to a significant portion preferably encompassing 20 nucleotides, that is located 5' and 3' of the region of the actual breakpoint in order to be indicative of a deletion event. For example, when such an oligonucleotide is hybridized under stringent conditions such as 0.2 $\times$ SSC, 0.1 SDS at 65°C and the probe would be 943 nucleotides long, then the hybridizing region should include portions that hybridize 3' as well as 5' of the breakpoint.

[0034] For example, a *Rhesus box* or a part thereof encompassing the region of the breakpoint is amplified. Thereafter the amplification product is assayed in a sequence specific way by hybridization to an oligonucleotide of about six or more nucleotides length.

[0035] Preferably, a stretch of DNA representative of a *Rhesus box* or part thereof encompassing the region of the breakpoint is amplified using two primers. One primer may be located in the *Rhesus box* 5' of the identity region and is specific for both the upstream *Rhesus box* and the hybrid *Rhesus box.* The other primer may be located in the *Rhesus*

box 3' of the identity region and is specific for both the downstream *Rhesus box* and the hybrid *Rhesus box.* in this application, the presence of an amplification product of the expected size is indicative of the presence of a hybrid *Rhesus box* and hence, of the *RHD* deletion.

[0036] Another possible combination of primers is the following: One primer may be located in the *Rhesus box* 5' of the identity region and is specific for both the upstream *Rhesus box* and the hybrid *Rhesus box.* The other primer may be located in the *Rhesus box* 3'of the identity region. In this application, the presence of a hybrid *Rhesus box* is determined by examining the specificity of the parts of the amplification product pertaining to a DNA stretch of the *Rhesus box* 3'of the identity region. This may for example be effected by hybridization with an oligonucleotide that hybridizes to the hybrid *Rhesus box* and to the downstream *Rhesus box* but not to the upstream *Rhesus box,* or by digestion with a restriction enzyme that cuts the hybrid *Rhesus box* and the downstream *Rhesus box* but does not cut the upstream *Rhesus box,* or by digestion with a restriction enzyme that does not cut the hybrid *Rhesus box* and the downstream *Rhesus box* but cuts the upstream *Rhesus box,* or by nucleotide sequencing.

[0037] Another possible combination of primers is the following: One primer may be located in the *Rhesus box* 5' of the identity region. The other primer may be located in the *Rhesus box* 3' of the identity region and is specific for both the downstream *Rhesus box* and the hybrid *Rhesus box.* In this application, the presence of a hybrid *Rhesus box* is determined by examining the specificity of the parts of the amplification product pertaining to a DNA stretch of the *Rhesus box* 5' of the identity region. This may for example be effected by hybridization with a nucleotide that hybridizes to the hybrid *Rhesus box* and to the upstream *Rhesus box* but not to the downstream *Rhesus box,* or by digestion with a restriction enzyme that cuts the hybrid *Rhesus box* and the upstream *Rhesus box* but does not cut the downstream *Rhesus box,* or by digestion with a restriction enzyme that does not cut the hybrid *Rhesus box* and the upstream *Rhesus box* but cuts the downstream *Rhesus box,* or by nucleotide sequencing.

[0038] The hybrid *Rhesus box* may also serve as a diagnostic tool for the presence of the *RHD* deletion when analyzed by an anti-DNA antibody specific for one or more embodiments of the hybrid box, a fragment or derivative thereof such as an scFvFab or $F(ab')_2$ fragment or an aptamer etc. Thus, antibodies, fragments or derivatives thereof or such aptamers can be generated by the person skilled in the art according to conventional technology (see, for example, Harlow and Lane, "Antibodies, A Laboratory Manual", CSH Press 1988, Cold Spring Harbor).

[0039] In a preferred embodiment, the invention relates to a nucleic acid molecular structure representative of an *RHD* negative haplotype comprising an *RHD* gene deletion involving the upstream *Rhesus box,* the downstream *Rhesus box,* or both.

[0040] A nucleic acid molecular structure that is flanking the *Rhesus box* in the common *RHD* negative haplotypes can be used to derive primers for amplification reactions such as long range PCR for the molecular analysis of the *RHD* locus.

[0041] For example, a stretch of DNA representative of a *Rhesus box* and parts of their flanking regions or parts thereof encompassing the region of the breakpoint is amplified using two primers. One primer may be located in the 5' flanking region of the *Rhesus box*. Alternatively, this primer may be located in the *Rhesus box* 5' of the identity region and is specific for both the upstream *Rhesus box* and the hybrid *Rhesus box.* The other primer may be located in the *Rhesus box* 3'flanking region. Alternatively, this primer may be located in the *Rhesus box* 3' of the identity region and is specific for both the downstream *Rhesus box* and the hybrid *Rhesus box*. In this application the presence of an amplification product of the expected size is indicative of the presence of a hybrid *Rhesus box* and hence, of the *RHD* deletion.

[0042] Another possible combination of primers is the following: One primer may be located in the 5'flanking region of the *Rhesus box*. The other primer may be located in the *Rhesus box* 3' of the identity region. In this application, the presence of a hybrid *Rhesus box* is determined by examining the specificity of the parts of the amplification product pertaining to a DNA stretch of the *Rhesus box* 3' of the identity region. This may for example be effected by hybridization with an oligonucleotide that hybridizes to the hybrid *Rhesus box* and to the downstream *Rhesus box* but not to the upstream *Rhesus box,* or by digestion with an restriction enzyme that cuts the hybrid *Rhesus box* and the downstream *Rhesus box* but does not cut the upstream *Rhesus box,* or by digestion with a restriction enzyme that does not cut the hybrid *Rhesus box* and the downstream *Rhesus box* but cuts the upstream *Rhesus box,* or by nucleotide sequencing.

[0043] Another possible combination of primers is the following: One primer may be located in the *Rhesus box* 5' of the identity region. The other primer may be located in the 3' flanking region of the *Rhesus box.* In this application, the presence of a hybrid *Rhesus box* is determined by examining the specificity of the parts of the amplification product pertaining to a DNA stretch of the *Rhesus box* 5' of the identity region. This may for example be effected by hybridization with an oligonucleotide that hybridizes to the hybrid *Rhesus box* and to the upstream *Rhesus box* but not to the downstream *Rhesus box,* or by digestion with an restriction enzyme that cuts the hybrid *Rhesus box* and the upstream *Rhesus box* but does not cut the downstream *Rhesus box,* or by digestion with a restriction enzyme that does not cut the hybrid *Rhesus box* and the upstream *Rhesus box* but cuts the downstream *Rhesus box,* or by nucleotide sequencing.

[0044] In a preferred embodiment the nucleic acid molecular structure is representative of *RHD* positive haplotypes as defined in the appended claims.

[0045] The term "*RHD* positive haplotype" refers to any haplotype that comprises DNA sequences specific for the *RHD* gene.

[0046] In a preferred embodiment the invention relates to a nucleic acid molecular structure respresentative of the common *RHD* positive haplotype, as defined in the appended claims.

[0047] In another preferred embodiment the nucleic acid molecular structure is derived from a sample comprising an *RHD* positive haplotype that is serologically classified RhD negative, as defined in the appended claims.

[0048] In the context of the invention, the term "serologically classified RhD negative" describes a sample that has been tested for the presence of RhD antigen using, e.g., routine serological assays wherein the result of such assays was negative.

[0049] In a particularly preferred embodiment the sample that is classified RhD negative is obtained from a Caucasian population.

[0050] Several additional *RHD* positive alleles occurring in RhD negative individuals have previously been partly or fully characterized (Table 9). Three of these ten published *RHD* alleles represented *RHD-CE-D* hybrid alleles in which the *RHCE* specific stretch encompassed at least exons 4 to 7. For each of these three hybrid *RHD* alleles, alleles were found whose patterns would be compatible (Table 9). Out of the seven RhD negative patterns observed in the present study, six were compatible with such type of hybrid *RHD* allele. Seven out of ten published *RHD* alleles represented deletions, nonsense mutations or a pseudogene. None of these alleles occurred in this study, which may indicate that they are rare in whites.

[0051] In a further more preferred embodiment the nucleic acid molecular structure of the invention or a nucleic acid molecule being derived from the *RHD* gene correlates with a RhD-negative phenotype.

[0052] The invention also relates to a process to specifically detect a *RHD* negative haplotype in a sample by utilizing any structural feature or nucleotide sequence or both of the above-described nucleic acid molecular structure or combinations thereof with techniques known in the art, preferably amplification reactions, such as polymerase chain reaction (PCR), more preferably by PCR-RFLP, PCR-SSP or long-range PCR.

[0053] The described PCR-RFLP and long-range PCR methods utilize either *Rhesus box* sequences or *Rhesus box* flanking sequences. By utilizing the same DNA stretches or combinations thereof, other methods, like PCR-SSO or biochips, can be developed or applied.

[0054] In one embodiment of the present invention a process is provided to specifically detect a common *RHD* negative haplotype comprises the following steps:

(a) isolating the DNA from a blood sample of obtained from a blood donor;
(b) hybridizing at least two oppositely oriented primers under stringent conditions to the DNA so as to carry out a PCR;
(c) amplifying the target sequence;
(d) separating the amplification products on a gel; and
(e) analyzing the amplicons.

[0055] Said sample may or may be derived from blood, serum, sputum, feces or other body fluid. The sample to be analyzed may be treated as to extract, inter alia, nucleic acids. The isolation of DNA from preferably EDTA- or citrate agglutinated blood samples can be carryied out by a modified salting out methods, following the standard techniques as described in Gassner, Transfusion 37: 1020, 1997. The primers are preferably oligonucleotides that either occur naturally or in a purified restriction digest or are produced synthetically. The primers are preferably single stranded for a maximum of efficiency in the method of the present invention, and are preferably oligodeoxyribonucleotides. Purification of said primers is generally envisaged, prior to their use in the methods of the present invention, said purification comprising High Performance Liquid Chromatography (HPLC) or Polyacrylamide gel electrophoresis (PAGE), all technologies that are well known to the skilled artisan. Amplification methods such as PCR or LCR are well known in the art and described, for example in Flegel, Transfusion Medicine 8 (1998), 281-302; Maaskant, Transfusion 38 (1998), 1015-1021 and Legler, Transfusion (1996), 426-31.

[0056] According to the present invention a preferred method to detect the *RHD* deletion is performing PCR-RFLP using the expand high fidelity PCR-system and non-specific primers binding 5' of the end of the *Rhesus box* identity region as well as primers specific for the downstream *Rhesus box* and binding 3' of the end of the *Rhesus box* identity region. The PCR conditions involve preferably annealing at 65°C, extension for 10 min at 68°C. Thereafter, PCR amplicons are digested with PstI for 3h at 37°C and fragments resolved using 1% agarose gel. Additional preferred methods are further described in examples 8 and 9.

[0057] Another embodiment of the invention relates to a process to specifically detect a common *RHD* negative haplotype comprising the detection of the hybrid *Rhesus box.*

[0058] The detection of the hybrid *Rhesus box* provides the practitioner with an unambiguous result as regards the nature of the corresponding *RHD* allele. If the hybrid *Rhesus box* is detected, then the *RHD* gene is deleted. Detection of the hybrid *Rhesus box* is preferentially effected by using an oligonucleotide that specifically hybridizes to a region

comprising the breakpoint. The oligonucleotide used for hybridization must directly hybridize to that breakpoint and, in addition, hybridize to at least 943 nucleotides 5' and 3' of the breakpoint. Hybridization occurs preferably under stringent conditions such as 0.2 X SSC, 0.1% SDS at 65°C. The actual breakpoint within the hybrid *Rhesus box* may vary due to the exact nature of the putative crossover event. Accordingly, the hybrid *Rhesus box* may also be detected using a number of overlapping or non-overlapping oligonucleotides used for hybridization. The hybrid *Rhesus box* may also be detected using other protocols such as restriction analysis (preferably in combination with Southern blot analysis), or PCR technology, as described herein above.

[0059] Furthermore, another embodiment of the invention relates to a process to specifically detect a common *RHD* negative haplotype comprising assessing the nucleic acid molecular structure comprising the hybrid *Rhesus box* and the flanking regions thereof.

[0060] In accordance with the present invention, assessment of the molecular nucleic acid structure comprises analysis steps such as gel electrophoresis using either agarose gels or polyacrylamide gels, treatment with restriction-enzymes, blotting techniques, such as Southern or Northern blotting or related techniques, such as fluorescence-guided detection of hybridization and other techniques known in the art.

[0061] The present invention also relates to a process to specifically detect a *RHD* negative haplotype in a sample comprising the step of detecting any of the breakpoint regions mentioned in the present invention.

[0062] In a preferred embodiment the invention relates to the above-mentioned process wherein said detection or assessment comprises the determination of the length of a nucleic acid molecule comprising the hybrid *Rhesus box* or parts thereof.

[0063] Again, this preferred embodiment of the method of the invention utilizes standard separation techniques, such as gel electrophoresis or chromatography or standard techniques of nucleotide sequencing as known to a skilled artisan. Preferably the present invention utilizes a commercially available sequencing kit and an automatic sequencing machine from Applied Biosystems (ABI 373A or ABI 377), as further described in Example 5, for this purpose.

[0064] Another preferred embodiment of the invention relates to the above-mentioned process wherein said detection or assessment is effected by using PCR-RFLP, PCR-SSP or long-range PCR or a probe specifically hybridizing to the hybrid *Rhesus box,* preferably to the breakpoint or breakpoint region depicted in Figure 3 or 4, or hybridizing to the upstream or downstream *Rhesus box,* preferably by Southern blot analysis, gel electrophoresis, biochip-analysis, molecular weight determination or fluorescence.

[0065] According to the present invention the term "hybridizing to" relates to stringent or non-stringent conditions. The setting of conditions is well within the skill of the artisan and to be determined according to protocols described, for example, in Sambrook, loc. cit. or Hames and Higgins, "Nucleic acid hybridization, a practical approach", IRC Press, Oxford (1985). The detection of specifically hybridizing sequences will usually require stringent hybridizing and washing conditions such as 0.2 x SSC, 0.1% SDS at 65°C. Non-stringent hybridization conditions for the detection of homologous and not exactly complementary sequences may be set at 6x SSC, 1 % SDS at 50°C or 65°C. As is well known, the length of the probe and the composition of the nucleic acid to be determined constitute further parameters of the hybridization conditions.

[0066] Furthermore, the invention relates to a vector comprising the nucleic acid molecular structure of the invention.

[0067] The vector may be used for propagation and/or expression or may be designed for gene transfer or targeting purposes. Methods of producing such vectors are well known in the art. The same holds true for cloning the nucleic acids of the mutation into said vectors, as well as the propagation of vectors in suitable hosts, etc.

[0068] The vector may particularly be a plasmid, a cosmid, a virus or a bacteriophage used conventionally in genetic engineering that comprise the nucleic acid molecule of the invention. Expression vectors derived from viruses such as retroviruses, vaccinia virus, adeno-associated virus, herpes viruses, or bovine papilloma virus, may be used for delivery of the nucleic acid molecules or vector of the invention into targeted cell populations. Methods which are well known to those skilled in the art can be used to construct recombinant viral vectors; see, for example, the techniques described in Sambrook, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory (1989) N.Y. and Ausubel, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. (1989). Alternatively, the polynucleotides and vectors of the invention can be reconstituted into liposomes for delivery to target cells. The vectors containing the nucleic acid molecules of the invention can be transferred into the host cell by well-known methods, which vary depending on the type of cellular host. For example, calcium chloride transfection is commonly utilized for prokaryotic cells, whereas calcium phosphate treatment or electroporation may be used for other cellular hosts; see Sambrook, supra.

[0069] Such vectors may comprise further genes such as marker genes which allow for the selection of said vector in a suitable host cell and under suitable conditions. Preferably, the nucleic acid molecule of the invention is operatively linked to expression control sequences allowing expression in prokaryotic or eukaryotic cells. Expression of said polynucleotide comprises transcription of the polynucleotide into a translatable mRNA. Regulatory elements ensuring expression in eukaryotic cells, preferably mammalian cells, are well known to those skilled in the art. They usually comprise regulatory sequences ensuring initiation of transcription and optionally poly-A signals ensuring termination of transcription and stabilization of the transcript. Additional regulatory elements may include transcriptional as well as translational

enhancers, and/or naturally-associated or heterologous promoter regions. Possible regulatory elements permitting expression in prokaryotic host cells comprise, e.g., the PL, lac, trp or tac promoter in E. coli, and examples for regulatory elements permitting expression in eukaryotic host cells are the AOX1 or GAL1 promoter in yeast or the CMV-, SV40- , RSV-promoter (Rous sarcoma virus), CMV-enhancer, SV40-enhancer or a globin intron in mammalian and other animal cells. Beside elements which are responsible for the initiation of transcription such regulatory elements may also comprise transcription termination signals, such as the SV40-poly-A site or the tk-poly-A site, downstream of the nucleic acid molecule. Furthermore, depending on the expression system used leader sequences capable of directing the polypeptide to a cellular compartment or secreting it into the medium may be added to the coding sequence of the polynucleotide of the invention and are well known in the art. The leader sequence(s) is (are) assembled in appropriate phase with translation, initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein, or a portion thereof, into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an C- or N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. In this context, suitable expression vectors are known in the art such as Okayama-Berg cDNA expression vector pcDV1 (Pharmacia), pCDM8, pRc/CMV, pcDNA1, pcDNA3 (In-vitrogene), or pSPORT1 (GIBCO BRL).

[0070] Preferably, the expression control sequences will be eukaryotic promoter systems in vectors capable of transforming or transfecting eukaryotic host cells, but control sequences for prokaryotic hosts may also be used.

[0071] As mentioned above, the vector of the present invention may also be a gene transfer or targeting vector. Gene therapy, which is based on introducing therapeutic genes into cells by ex-vivo or in-vivo techniques is one of the most important applications of gene transfer. Suitable vectors and methods for in-vitro or in-vivo gene therapy are described in the literature and are known to the person skilled in the art; see, e.g., Giordano, Nature Medicine 2 (1996), 534-539; Schaper, Circ. Res. 79 (1996), 911-919; Anderson, Science 256 (1992), 808-813; Isner, Lancet 348 (1996), 370-374; Muhlhauser, Circ. Res. 77 (1995), 1077-1086; Wang, Nature Medicine 2 (1996), 714-716; WO94/29469; WO 97/00957 or Schaper, Current Opinion in Biotechnology 7 (1996), 635-640, and references cited therein. The polynucleotides and vectors of the invention may be designed for direct introduction or for introduction via liposomes, or viral vectors (e.g. adenoviral, retroviral) into the cell.

[0072] Additionally, the invention relates to a non-human host transformed with the vector of the invention.

[0073] Appropriate hosts comprise transgenic animals, cells such as bacteria, yeast cells, animal, preferably mammalian cells, fungal cells or insect cells. Transformation protocols including transfection, microinjection, electroporation, etc., are also well known in the art.

[0074] Furthermore, the invention relates to an oligonucleotide hybridizing under stringent conditions to a portion of the nucleic acid molecular structure or the nucleic acid molecules of the invention, wherein said portion hybridizes to a breakpoint of the gene conversion as characterized in the appended claims.

[0075] In this embodiment of the invention, it is understood that the oligonucleotides hybridizes directly to the breakpoint. The setting of stringent hybridization conditions is well described, for example, in Sambrook et al, "Molecular Cloning, A Laboratory Handbook" CSH Press, Cold Spring Harbor 1989 or Hames and Higgins, "Nucleic acid hybridization, a practical approach", IRL Press, Oxford (1985). Thus, the detection of the specifically hybridizing sequences will usually require hybridization and washing conditions such as 0.2 x SSC, 0.1% SDS at 65°. As is well known, the length of the probe and the composition of the nucleic acid to be determined constitute further parameters of the stringent hybridization conditions. Preferably, the oligonucleotide is a deoxynucleotide. It is further preferred that the oligonucleotide comprises 12 to 50 nucleotides and more preferably 15 to 24 nucleotides. The hybridization to the breakpoint may be under stringent or non-stringent conditions. An example of non-stringent hybridization conditions is hybridization and washing at 50°C in 4 x SSC, 0,1% SDS.

[0076] Furthermore, the invention relates to a method for testing for the presence of a nucleic acid molecule encoding a mutant Rhesus D antigen or of a nucleic acid molecule carrying a deletion of the *RHD* gene as characterized by the nucleic acid molecular structure or the nucleic acid molecule of the invention in a sample comprising hybridizing the oligonucleotide of the invention under stringent conditions to nucleic acid molecules comprised in the sample obtained from a human and detecting said hybridization.

[0077] Preferably, the method of the invention further comprises digesting the product of said hybridization with a restriction endonuclease or subjecting the product of said hybridization to digestion with a restriction endonuclease and analyzing the product of said digestion.

[0078] This preferred embodiment of the invention allows by convenient means, the differentiation between an effective hybridization and a non-effective hybridization. For example, if the wild type *RHD* gene comprises an endonuclease restriction site, the hybridized product will be cleavable by an appropriate restriction enzyme whereas a mutated sequence will yield no double-stranded product or will not comprise the recognizable restriction site and, accordingly, will not be cleaved. The analysis of the digestion product can be effected by conventional means, such as by gel electrophoresis which may be optionally combined by the staining of the nucleic acid with, for example, ethidium bromide. Combinations with further techniques such as Southern blotting are also envisaged.

[0079] Detection of said hybridization may be effected, for example, by an anti-DNA double-strand antibody or by employing a labeled oligonucleotide. Conveniently, the method of the invention is employed together with blotting techniques such as Southern or Northern blotting and related techniques. Labeling may be effected, for example, by standard protocols and includes labeling with radioactive markers, fluorescent, phosphorescent, chemiluminescent, enzymatic labels, etc.

[0080] The invention additionally relates to a method for testing for the presence of a nucleic acid molecule encoding a mutant Rhesus D antigen or of a nucleic acid molecule carrying a deletion of the *RHD* gene as characterized by the nucleic acid molecular structure or the nucleic acid molecule of the invention in a sample comprising determining the nucleic acid sequence of at least a portion of the nucleic acid molecular structure of the invention, said portion encoding a breakpoint of said hybrid gene.

[0081] Preferably, the method of the invention further comprises, prior to determining said nucleic acid sequence, amplification of at least said portion of said nucleic acid molecular structure.

[0082] Furthermore, the invention relates to a method for testing for the presence of a nucleic acid molecule carrying a deletion of the *RHD* gene as characterized by the nucleic acid molecular structure of the invention in a sample comprising carrying out an amplification reaction using a set of primers that amplifies at least a portion of said sequence wherein at least one of the primers employed in said amplification reaction is the oligonucleotide as defined in the appended claims.

[0083] Preferably, amplification is effected by polymerase chain reaction (PCR). Other amplification methods such as ligase chain reaction may also be employed.

[0084] In a preferred embodiment of the method of the invention said PCR is PCR-RFLP, PCR-SSP or iong-range PCR.

[0085] Additionally, in another preferred embodiment of the invention the molecular weight of the amplification product is analyzed. Said analysis of the molecular weight utilizes standard techniques, such as agarose gel electrophoresis, SDS-PAGE, mass spectrometry such as MALDI-TOF for this purpose, which are well known to the person skilled in the art.

[0086] In one embodiment of the method of the invention, a method is provided that detects *RHD* positive alleles comprising the following steps:

(a) isolating DNA from a blood sample of a sample obtained from a blood donor;
(b) hybridizing at least two oppositely oriented primers under stringent conditions to the DNA so as to carry out a PCR;
(c) amplifying the target sequence;
(d) separating the amplification products on a gel; and
(e) analyzing the amplicons.

[0087] With regard to specific conditions to be applied in the various steps, it is referred to the corresponding description herein above.

[0088] In a preferred embodiment the *RHD* positive alleles are derived from a serologically RhD negative population. In another preferred embodiment the RhD-negative sample is selected from a Caucasian population.

[0089] Preferably, in the method of the invention said amplification or amplification reaction is or is effected by the polymerase chain reaction (PCR). Other amplification methods such as ligase chain reaction may also be employed.

[0090] Preferably, in the method of the invention said sample is blood, serum, plasma, fetal tissue, saliva, urine, mucosal tissue, mucus, vaginal tissue, fetal tissue obtained from the vagina, skin, hair, hair follicle or another human tissue.

[0091] Furthermore, the method of the invention preferably comprises the step of enrichment of fetal cells. This enrichment may be achieved by using appropriate antibodies, lectins or other reagents specifically binding fetal cells or by any technique attempting the differential separation of maternal and fetal cells, like by density gradients. Also preferably, in said method fetal DNA or mRNA from maternal tissue like peripheral blood, serum or plasma may be extracted, advantageously according to conventional procedures.

[0092] In an additional preferred embodiment of the method of the invention, said nucleic acid molecule or proteinaceous material from said sample is fixed to a solid support.

[0093] Preferably, said solid support is a chip.

[0094] The advantages of chips are well known in the art and need not be discussed herein in detail. These include the small size as well as an easy access of computer based analysis of analytes.

[0095] Furthermore, the present invention relates to the use of the nucleic acid molecular structure of the invention for the analysis of a negative or a positive Rhesus D phenotype.

[0096] The analysis can be effected, for example, on the basis of the methods described herein above.

[0097] The invention also relates to a method for determining whether a patient in need of a blood transfusion is to be transfused with RhD negative blood from a donor comprising the step of testing a sample from said patient for the presence of one or more nucleic acid molecular structures of the invention, wherein a positive testing for two different of said nucleic acid molecular structures or nucleic acid molecules is indicative of the need for a transfusion with Rh negative blood. Alternatively, a positive testing indicating the concomitant presence of two identical copies of one of said nucleic acid molecular structures or nucleic acid molecules is indicative of the need for a transfusion with Rh negative

blood.

[0098] Alternatively, a negative testing for the presence of the nucleic acid molecular structure representative of the common *RHD* negative haplotype with or without a negative testing for one or more nucleic acid molecular structures representative of the other *RHD* negative nucleic acid molecular structures of this invention permits the transfusion of blood that is typed as RhD positive. The invention has important implications for devising a transfusion therapy in humans. For example, it can now be conveniently tested whether the patient actually needs a transfusion with a RhD negative blood or whether such precautions need not be taken.

[0099] Furthermore, the invention relates to a method for determining whether blood of a donor may be used for transfusion to a patient in need thereof comprising the step of testing a sample from said donor for the presence of one or more of said nucleic acid molecular structures of the invention, wherein a negative testing for the nucleic acid molecular structures representative of the common *RHD* negative haplotype with or without a negative testing for one or more nucleic acid molecular structures or nucleic acid molecules representative of the other *RHD* negative haplotypes of this invention excludes the transfusion the donor's blood to a patient that is typed as RhD negative.

[0100] The samples referred to in the above recited methods may be samples that are referred to throughout the specification, such as blood, serum, etc.

[0101] As regards the guidelines for transfusing a patient on the basis of any of the above recited methods, the utmost care must be taken that suboptimal transfusion policy is avoided. The risk factor is always to be considered by the physician in charge. In all cases, the potential risk for the patient is to be minimized.

[0102] The invention also relates to a method for assessing of the risk of a RhD negative mother of conceiving or carrying an RhD positive fetus or of the risk of a mother having an anti-D titer of conceiving or carrying a fetus at risk to develop hemolytic disease of the newborn comprising assessing a sample obtained from the father of the fetus for the presence of one or more of said nucleic acid molecular structures or nucleic acid molecules of the invention, wherein a negative testing for nucleic acid molecular structures or nucleic acid molecules representative of the common *RHD* negative haplotype with or without a negative testing for one or more nucleic acid molecular structures or nucleic acid molecules representative of the other *RHD* negative haplotypes of this invention is indicative for a high risk of conceiving an RhD positive fetus.

[0103] In a preferred embodiment of the method of the present invention said nucleic acid molecular structure carries mutations or deletions.

[0104] Furthermore, the invention relates to a method for assessing the possibility or likelihood of a man being the father of a child by assaying a sample obtained from said man for the presence of one or more of said nucleic acid molecular structures of the invention, wherein the test results are used to determine the homozygosity for, the heterozygosity for or the absence of any nucleic acid molecular structures representative of the *RHD* negative haplotype of the present invention used to infer the possibility or likelihood of said man being the father of the child.

[0105] Hence and in summary, the present invention provides means and methods for the detection of *RHD* haplotypes, comprising common *RHD* negative haplotypes, as described, above, as well as presumably rare *RHD* positive allels in serologically RhD negative populations. Latter alleles, harbouring *RHD* sequences and therefore determined as *RHD*-positive, can comprise either *RHD*/*RHCE* hybrid genes, stop codons, splice site mutations or gene deletions, that terminate or reduce the RhD antigen expression. Carrying out the improved detection methods of the invention, it was surprisingly found, that several samples, determined as RhD negative in routine serology, could be identified having *RHD* positive alleles. Furthermore, some of those samples were even RhD antigen positive when performing a detection assay based on adsorption and elution, indicating that the molecular basis for the *RHD* positive alleles in RhD negatives is more heterogenous than anticipated. Advantageously, the disclosure content of the present invention now provides new and practicable nucleic acid amplification techniques to determine whether *RHD* specific sequences cause RhD positive or RhD negative phenotypes.

[0106] According to the present invention the term "polymorphism" relates to the existence in a population of more than one genetic structure or a gene of a haplotype or of a DNA segment. Nevertheless, sometimes such a genetic polymorphism does not always result in a differing phenotype, but may only be detected at the genetic level.

[0107] Furthermore, the invention relates to a kit comprising

(a) the oligonucleotide of the invention; and/or

(b) a pair of primers useful for carrying out the amplification reaction of the invention.

[0108] Parts of the kit can be packaged individually in vials or in combination in containers or multicontainer units. The kit of the present invention may be advantageously used for carrying out the method of the invention and could be, inter alia, employed in a variety of applications referred to above. The manufacture of the kits follows preferably standard procedures which are known to people skilled in the art.

[0109] Finally, the invention relates to the use of an oligonucleotide for hybridizing under stringent conditions to a

portion of the nucleic acid molecular structure of the invention wherein said portion hybridizes to a region involving the breakpoint of said hybrid gene or to the complementary portion thereof, wherein said region involving the breakpoint is characterized in that the 5' portion of the upstream Rhesus box is in close spatial proximity to the 3'portion of the downstream Rhesus box.

[0110] The figures show

Figure 1 Schematic structure of the *RH* gene locus. The positions and orientations of the genes and the *Rhesus boxes* are indicated by open arrows and triangles, respectively (Panel A). The exons are shown as vertical bars and their exon number is indicated. The two *RH* genes have opposite orientation, face each other with their 3' ends, and are separated by about 30,000 bp. A third gene, *SMP1,* has the same orientation as *RHD* and is positioned in between *RHD* and *RHCE.* The *RHD* gene is flanked on both sides by the two highly homologous *Rhesus boxes* (b). All exons are shorter than 200 bp with the exception of the *RHD* and *SMP1* 3' terminal exons. Data used to establish this structure (Panel B) include the extension of genomic sequences represented in the cDNAs (horizontal arrows), identities and homologies to genomic clones (bar a: identity with dJ465N24; b: homology of *RHD* to dJ469D22; c: homology of *RHD* 3' part to dJ465N24; d: identity with dJ469D22). The positions of three bridging PCR reactions are indicated. The correct position of a nucleotide stretch previously reported by Okuda et al. (Okuda, Biochem. Biophys. Res. Commun. 263:378, 1999) as "spacer" sequence between *RHD* and *RHCE* is indicated by the bar labeled s.

Figure 2 Chromosomal organization of the DNA regions located 5' to the *RHD* and *RHCE* genes. The proposed structure of the *RHCE* and *RHD* 5' flanking regions is depicted (Panel A). A total of 4,941 bp immediately 5' of the ATG start codons are homologous between the *RHCE* and *RHD* genes (vertically hatched bars). No homology is present further beyond this homology region (diagonally hatched bars). Two genomic clones, dJ469D22 and dJ465N24, were utilized for primer design. DJ469D22 comprises the full length of the depicted *RHCE* region, whereas dJ465N24 extends only 466 bp into the homology region. The positions of several PCR primers are indicated (a, rey14a; b, rend32; c, rey15a; d, re014; e, re011d). This proposed structure is supported by several PCR reactions (panel B). Forward priming was done with primer a (*RHCE* specific, lane 1 - 3), primer b (*RHD* specific, lane 4 - 6), and primer c (*RHCE* and *RHD* homology region, lane 7 - 9). Amplicons were lacking for primer a with *RHD* specific reverse primer e (lane 2) and for primer b with *RHD* negative DNA (lane 6). The other seven PCR reactions yielded amplicons of the predicted sizes in accordance with the genomic structure shown in panel A.

Figure 3 Chromosomal organization of the *Rhesus boxes.* The physical extension of the upstream *Rhesus box* (5' to *RHD)* is 9,145 bp (black bar). About 63% of the boxes' nucleotide sequence consists of repetitive DNA; the types of the repeat families are indicated. The overall homology between the upstream and downstream *Rhesus box* is 98.6%, but within an 1,463 bp identity region (horizontal arrows), there is only a single 4 bp insertion (double vertical line). A CpG-island (double-headed arrow) is located at the 3' end and is in the downstream *Rhesus box* (3' to *RHD*) adjacent to the *SMP1* promoter.

Figure 4 *RHD* gene deletion in the Rh negative haplotypes. Three 3,100 bp segments of the *Rhesus boxes* are shown. The upper line indicates the nucleotide sequence of the upstream *Rhesus box* in D-positives, the lower line the nucleotide sequence of the downstream *Rhesus box* in D-positives. The middle line gives the nucleotide sequence of the single *Rhesus box* carried by Rh negatives. Asterisks denote identical nucleotides. The *RHD* deletion occurred in a 903 bp segment of absolute identity that was part of a 1,463 bp identity region. The positions of primers rez7 and mb31 is shown (m indicates mismatch). *Pst*I restriction sites are indicated by carets (^). The three *Rhesus boxes* are deposited at EMBL under accession numbers AJ252311 (upstream *Rhesus box),* AJ252312 (downstream *Rhesus box*), and AJ252313 (hybrid *Rhesus box*).

Figure 5 Two technical procedures for specific detection of the *RHD* deletion in the common *RHD* negative haplotypes. A long-range PCR amplification with primers located in *non-Rhesus box* sequences (Panel A) and PCR-RFLP with primers located in the *Rhesus boxes* are shown (Panel B). The deduced genotypes are indicated. The primers of the long-range PCR were located 5' of the upstream *Rhesus box* (primer rez4) and in *SMP1* exon 1 (primer sr9). *RHD* negative haplotypes were detected specifically (Panel A, lane 1 - 6). DNA homozygous for the *RHD* gene was negative, because the PCR cannot amplify the 70,000 bp DNA stretch of the *RHD* gene. For the PCR-RFLP method, the PCR amplicons (primer rez7 and mb31) were digested with *Pst*I. In D-negatives, there are three *Pst*I sites in the amplicon (see Fig. 4) resulting in fragments of 1,888 bp, 564 bp, 397 bp, and 179 bp (lane 1 to 3). The downstream *Rhesus box* of D-positives lacks one *Pst*I-

site resulting in fragments of 1,888 bp, 744 bp, and 397 bp (lane 7 to 9). *RHD*$^{+}$/*RHD*$^{-}$ heterozygotes show both fragments of 744 and 564 bp (lane 4 to 6). The 564 bp fragment appears weaker because heterodimers are not cut by *Pst*I. Primer mb31 does not amplify the upstream *Rhesus box* of D-positives.

Figure 6 Model of the proposed mechanism causing the prevalent *RHD* negative haplotypes in whites. The physical structure of the *RHD* and *RHCE* gene locus is depicted (panel A). An unequal crossing over between the upstream and downstream *Rhesus boxes* can be triggered by their high homology (panel B). The breakpoint region in the *Rhesus boxes* was found to be of 100% homology for 903 bp (see Fig. 4). Resolving the crossed over chromosome yields the *RH* gene structure of the extant *RHD* negative haplotype (panel C).

Figure 7 DNA sequence of the hybrid *Rhesus box* of *RHD* negatives.

Figure 8 DNA sequence of the upstream *Rhesus box* of D-positives.

Figure 9 DNA sequence of the downstream *Rhesus box* of D-positives.

[0111] The examples illustrate the invention:

Example 1: Blood samples and DNA isolation

[0112] EDTA- or citrate-anticoagulated blood samples were collected from white blood donors and characterized as D negative in routine typing including an antiglobulin test with anti-D (Wissenschaftlicher Beirat der Bundesärztekammer; Paul-Ehrlich-Institut. Richtlinien zur Blutgruppenbestimmung und Bluttransfusion (Hämotherapie). Köln: Deutscher Ärzte-Verlag; 1996; Wagner, Infusionsther Transfusionsmed 22:285-90, 1995). If necessary, samples were collected at random for specific CcEe phenotypes. A total of 314 ccddee, 433 Ccddee, 271 ccddEe, 19 CcddEe, 24 CCddee, 1 CcddEE and 6 ccddEE samples were tested. DNA was isolated by a modified salting-out procedure as described in Gassner et al., Transfusion 37; 1020, 1997.

Example 2: Molecular work-up

[0113] All samples were tested by PCR-SSP for the presence of four different *RHD* specific polymorphisms located in the *RHD* promoter, intron 4, exon 7 and the 3' untranslated region of exon 10. 48 samples with at least one positive PCR reaction were detected (Table 5). Those samples were further investigated for the presence of *RHD* specific polymorphisms in exon 3, exon 4, exon 5, exon 6, exon 7, intron 7 and exon 9. Twenty-six samples showed one of eight distinct PCR patterns involving a mixture of positive and negative reactions (Table 6). Twenty-two samples were positive for all *RHD* specific polymorphisms investigated and were assigned to eight *RHD* alleles by *RHD* specific sequencing of the ten *RHD* exons from genomic DNA (Table 7). For each PCR pattern and each *RHD* allele, one sample was serologically investigated. The phenotypes were determined to represent weak D, partial D, and D_{el}, or confirmed as serologically D negative by adsorption/elution (Table 6 and 7).

Example 3: DNA database searches and analysis

[0114] The GenBank (http://www.ncbi.nlm.nih.gov/BLAST/) and the chromosome 1 database of the Sanger Center (http://www.sanger.ac.uk/cgi-bin/nph-Blast_Server.html) were searched with cDNA sequences representative of *RHD* (RhXIII, accession number X63097) and *RHCE* (RhVI, X63095) using the BLAST program. The 84,810 bp genomic clone dJ469D22 (GenBank accession number AL031284), the 129,747 bp genomic clone dJ465N24 (GenBank accession number AL031432) and the 2,234 bp *SMP1* cDNA (GenBank accession number AF081282) were identified. dJ469D22 represented a major fragment of the *RHCE* gene, starting 33,340 bp 5' of the *RHCE* start codon and ending 1,142 bp 3' of exon 9. In dJ465N24, an internal stretch of 1,418 bp located between position 120,158 and 121,568 was 96% homologous to the 3' end of the *RHD* cDNA. The 3' end of the *SMP1* cDNA was complementary to the 3' end of the *RHCE* cDNA with an overlap of 58 bp.

Example 4: PCR

[0115] If not mentioned otherwise, PCR reactions were done with 60°C annealing, 10 min extension at 68°C and denaturation at 92°C using the expand long template or the expand high fidelity PCR systems (Boehringer Mannheim, Mannheim, Germany) and the listed primers (Table 1). Three PCR reactions were used to bridge gaps in the 3' flanking regions of the *RH* genes. PCR 1 was done using primers rea7 and rend31 (PCR 2, rend32, sf1c; PCR 3 , rea7, sf3).

The structure of the 5' flanking regions was confirmed with PCR amplifications involving sense primers rend32, rey14a, rey15a and antisense primers re011d and re014. Intron 9 size was estimated to be about 9,000 bp based on PCR amplifications using rb10b and rr4 for *RHD* (re96 and rh7 for *RHCE*).

Example 5: Nucleotide sequencing

[0116] Nucleotide sequencing was performed with a DNA sequencing unit (Prism BigDye terminator cycle-sequencing ready reaction kit; ABI 373A, Applied Biosystems, Weiterstadt, Germany).

Example 6: Characterizing the *RH* gene locus

[0117] A physical structure of the *RH* genes' locus was derived (Fig. 1). This structure was deduced from the following considerations: **(i) 3' flanking regions.** The 3' flanking region of *RHD* was highly homologous to the 3' part of dJ465N24 (Fig. 1 B, region c). This homology continued beyond the end of the *RHD* cDNA and extended for at least 8,000 bp as proven by the fact that it was possible to obtain PCR amplicons (Fig. 1B, PCR 1). Sequences homologous to the 3' part of dJ465N24 were neighboring to the 5' region of the *SMP1* gene (Fig. 1B; PCR 2). The 3' end of the *SMP1* gene occurred immediately adjacent to the *RHCE* gene as indicated by the complementarity of the 3' ends of the respective cDNAs and confirmed by PCR (Fig. 1B, PCR 3). Further details of the *RHD* 3' flanking region (Rhesus *box*) and the *SMP1* gene are described below. **(ii) 5' flanking regions**. dJ469D22 comprised 33,340 bp 5' flanking region of *RHCE*. For *RHD*, a 466 bp homology between the 3' end of dJ465N24 and dJ469D22 indicated that dJ465N24 might represent the 5' flanking sequence of *RHD.* This assumption was proven by PCR (Fig. 2). **(iii) Analysis of YAC 38A-A10**. DNA from the YAC 38A-A10 (UK HGMP resource centre, Cambridge, UK) was isolated after a single growth phase by standard methods (http://hdkiab.wustl.edu/lab_manual/yeast). It was confirmed that this YAC contained *RH* DNA. Furthermore, shotgun cloning experiments indicated that some of its insert probably derived from the X chromosome (data not shown). This YAC had been known to contain *RHCE* exons 2 to 10 and *RHD* exons 1 to 10 (Carritt, Hum. Mol. Genet. 6:843, 1997) and was thus expected to contain the DNA segments interspersed between *RHD* and *RHCE.* The presence of DNA segments representative of different parts of the *RH* locus in this YAC was observed (Table 2). The results were concordant with the proposed structure of the *RH* locus shown in Fig. 1, Panel A.

Reference Example 1: Identification of *RHD* specific sequences in the *RHD* promoter

[0118] About 2,000 bp *RHD* promoter sequence was established by chromosomal walking (GenomeWalker kit, Clontech, Heidelberg, Germany). D-positive and D-negative samples were amplified using primers re04 and re11d (Table 1) and *RHD*- and *RHCE*-specific sequences established for 1,200 bp 5' of the start codon by sequencing with internal primers. A short deletion in the *RHD* gene was identified and used to develop the *RHD*-specific primer re011d. The 1,200 bp sequence including the *RHD* promoter has been deposited at EMBL under accession no. AJ252314.

Example 7: Characterization of *Rhesus boxes*

[0119] Two DNA segments of about 9,000 bp, located 5' and 3' of the *RHD* gene, were highly homologous, had identical orientation, and were designated "*Rhesus boxes*" (Fig. 4). The *Rhesus boxes* were amplified and sequenced using internal primers in two overlapping fragments using PCR primer pairs rez4/rend31 and rend32/re011d (upstream *Rhesus box*), rea7/rend31 and rend32/sr9 (downstream *Rhesus box*), and rez4/rend31 and rend32/sr9 (hybrid *Rhesus box* of RHD-negative). The upstream *Rhesus box* (5' of *RHD*) was about 9,142 bp long and ended about 4,900 bp 5' of the *RHD* start codon. The downstream *Rhesus box* (3' of *RHD*) was 9,145 bp long and started 104 bp after the *RHD* stop codon. The *Rhesus boxes* exactly embraced the part of *RHD* with homology to *RHCE.* The central portion of both *Rhesus boxes* contained an almost complete remnant of a transposon-like human element (THE-1B). The single open reading frame usually found in the THE-1B element was, however, abolished due to several nucleotide aberrations occurring in both *Rhesus boxes* in parallel, including a nonsense mutation in codon 4. While there was overall 98.6% homology between both *Rhesus boxes,* a 1,463 bp "identity region" located between positions 5,701 and 7,163 was completely identical with the single exception of a 4 bp T insertion in a poly T tract. downstream *Rhesus box*, in a distance between *RHD* and *RHCE* of about 30,000 bP (Fig. 1).

Example 8: Localization of the *RHD* gene deletion in the *RHD* negative haplotypes

[0120] It was reasoned that the homology of the two *Rhesus boxes* may have been instrumental for the mechanism of the *RHD* deletion in the common *RHD* negative haplotypes. The nucleotide sequence of the *Rhesus box* in *RHD* negative DNA was determined (Fig. 5). The single *Rhesus box* detected in *RHD* negatives had a hybrid structure. The

5' end of this *Rhesus box* represented a upstream *Rhesus box,* the 3' end a downstream *Rhesus box.* It was determined that the 903 bp breakpoint region of the *RHD* deletion was located in the identity region of the *Rhesus boxes* (Fig. 4, arrow pointing to left).

Example 9: Specific detection of the *RHD* deletion by PCR

[0121] Two PCR based methods were developed for specific detection of the *RHD* gene deletion occurring in the prevalent *RHD* negative haplotypes (Fig. 6). Long-range PCR-SSP was performed using the expand long template PCR system with buffer 3 and primers rez4 (5' of upstream *Rhesus box*) and sr9 (*SMP1* exon 1). Annealing was at 60°C and extension 20 min at 68°C. PCR amplicons were resolved using a 1% agarose gel. PCR-RFLP was performed using the expand high fidelity PCR system and primers rez7 (non-specific, 5' of *Rhesus box* identity region) and mb31 (specific for downstream *Rhesus box,* 3' of *Rhesus box* identity region). Annealing was at 65°C and extension 10 min at 68°C. PCR amplicons were digested with *Pst*I for 3 hrs at 37°C and fragments resolved using a 1 % agarose gel.
These techniques allowed the ready and direct detection of the common *RHD* negative haplotypes, even if they are *in trans* to *RHD* positive haplotypes. PCR-RFLP was further applied to a larger number of samples (Table 3). As expected, all 33 samples with known genotype were correctly typed. In 68 additional samples representative of the most common phenotypes, the results were consistent with the known haplotype frequencies in the population.

Example 10: *RHD* PCR-SSP

[0122] The PCR-SSP reactions (Table 4) were adapted and extended from a previously described *RHD* exon specific PCR-SSP method (Gassner, Transfusion 37:1020-6, 1997) and were triggered to work under identical thermocycling conditions. Concentrations of specific primers were 0.2 μM for all reactions with the exception of exon 6 (0.1 μM), intron 7 (0.4 μM) and exon 9 (0.4 μM). For most samples intron 4/exon 7 was tested as multiplex reaction containing 0.2 μM of exon 7 (primer set ga71/ga72) and 0.1 μM of intron 4 primers. Each reaction contained a set of HGH primers (Gassner, Transfusion 37:1020-6, 1997) as an internal control in concentrations of 0.05 μM for promoter, intron 4, and exon 7 with ga71/ga72; 0.075μM for exon 10; 0.1 μM for intron 7; 0.15 μM for exon 3, exon 4, exon 7 with rb26/re71, and exon 9; 0.2 μM for exon 5 and exon 6. Mg^{2+} concentration was 0.4 μM for intron 7 and for all other reactions 0.15 μM. For exon 6, 20 % solution Q (Qiagen, Hilden, Germany) was added.

Reference Example 2: Haplotype frequencies

[0123] For alleles observed more than once, their haplotype association was trivial. Alleles that were observed only once were assumed to be associated with the Cde or cdE haplotype rather than the cde haplotype, because no *RHD* positive allele was detected in any ccddee sample. An allele occurring in a single CcddEe sample was formally counted half for Cde and half for cdE. CCddee samples were assumed to harbour one aberrant and one normal Cde allele. The frequency of a given aberrant *RHD* allele in its haplotype was calculated as the number of observed samples divided by the number of the corresponding haplotypes under observation (500 Cde, 302 cdE). The population frequency of an *RHD* allele was calculated from the frequency of this allele in its haplotype and the known frequency of the haplotype in the local population (Wagner, Infusionsther. Transfusionsmed. 22:285-90, 1995). The haplotype frequencies were calculated for each PCR pattern and for each *RHD* allele (Table 8). In accordance with a previous study in England by Avent et al. (Avent, Blood 89:2568-77,1997), 4.9% of Cde haplotypes and 1.5% of cdE haplotypes were *RHD* positive in our population. As no *RHD* positive allele was detected among 314 ccddee samples, the frequency in the cde haplotype was less than 0.5 % (upper limit of one-sided 95% confidence interval, Poisson distribution). The three frequencies differed statistically significantly from each other ($p<0.05$; two sided Fisher's exact test for each pairwise comparison corrected according to Bonferoni-Holm). The population frequency of any D negative *RHD* positive haplotype was estimated to be 1:1,606. D_{el} alleles could only be observed in the presumed Cde haplotypes. About 3% of samples carrying antigen C that were typed D-negative in the blood bank routine represented D_{el}. The population frequency of D_{el} alleles was 1:3,030.

Table 1. Primers

Primer	Nucleotide sequence	Localization	Position
rb10b	ggctaaatattttgatgaccaagtt	*RHD* cDNA	1,194 to 1,217
re011d	gcagccaacttcccctgtg	*RHD* promoter	-883 to -901
re014	gctctaccttggtcacctcc	dJ469D22	52,189 to 52,209
re04	aggtcacatccatttatcccactg	dJ469D22	53, 968 to 53,945
re11d	agaagatgggggaatctttttcct	dJ469D22	51,193 to 51,216

(continued)

Primer	Nucleotide sequence	Localization	Position
re96	ttgtgactgggctagaaagaaggtg	dJ469D22	242 to 216
rea7	tgttgcctgcatttgtacgtgag	*RHD* cDNA	1,311 to 1,333
rend31	ttctgtctgggttggggaggg	dJ465N24	128,649 to 128,629
rend32	ggaggggttaatatgggtggc	dJ465N24	127,355 to 127,375
rend8b1	tttgtcctggttgcctgtggtc	dJ465N24	69,296 to 69,274
rend8b2	caaatcctgttgactggtctcgg	dJ465N24	68,451 to 68,473
rend9a1	aacggctccatcacccctaaag	dJ465N24	50,008 to 49,987
rend9a2	cccactcctagataccaacccaag	dJ465N24	49,059 to 49,083
rey14a	ctttatgcactgcctcgttgaatc	dJ469D22	56,792 to 56,769
rey14b	ttgactggtgtggttgctgttg	dJ469D22	55,863 to 55,884
rey15a	gcagaaaggggagttgatgctg	dJ469D22	55,416 to 55,395
rey7	ctgacaaagttgagagcccactg	dJ469D22	62,324 to 62,346
rey8	ttaagcctacatccacatgctgag	dJ469D22	62,854 to 62,831
rez2	ccttggtctgccagaattttca	*RHD* cDNA	2738 to 2717
rez4	gtttggcatcataggagatttggc	dJ465N24	120,101 to 120,124
rez7	cctgtccccatgattcagttacc	dJ465N24	124,831 to 124,854
rh7	acgtacaaatgcaggcaac	*RHD* cDNA	1,330 to 1,312
mb31	ccttttttttgtttgtttttggcggtgc	downstream *Rhesus box*	6,710 to 6,684
rr4	agcttactggatgaccacca	*RHD* cDNA	1,541 to 1,522
sf1	gactgggggggaaaagcgcaatac	*SMP1* cDNA	142 to 164
sf1c	gtattgcgcttttccccccagtc	*SMP1* cDNA	164 to 142
sf3	tgacttgctctcatcccacatg	*SMP1* cDNA	1,696 to 1,717
sm19	gggcttgaagcaagtaaatggaag	*SMP1* intron 1	- 58 to -35
sr1	gctatcaatattttcttggttacagacac	*SMP1* cDNA	2,172 to 2,144
sr3	gttcactgccataagtcttcagtgc	*SMP1* cDNA	575 to 551
sr3kp	tggccgcactgaagacttatgg	*SMP1* cDNA	546 to 567
sr45	cagctgcatctatgataatccacc	*SMP1* cDNA	224 to 243
sr47	atggacaagtccgaggtgatag	*SMP1* cDNA	315 to 344
sr47c	atcacctcggacttgtccattc	*SMP1* cDNA	342 to 321
sr5	gcaatcagagatccaaaggccaac	*SMP1* cDNA	428 to 405
sr5c	gttggcctttggatctctgattgc	*SMP1* cDNA	405 to 428
sr55	gacatagtataccctggaattgctgt	*SMP1* cDNA	472 to 497
sr55c	acagcaattccagggtatactatgtc	*SMP1* cDNA	497 to 472
sr9	ctcccccgattttagccaagaa	*SMP1* cDNA	27 to 6

For the *RHD* promoter and the *RHD* cDNA, the positions refer to the distance from the A of the start codon. For introns, they refer to the distance from the intron/exon junction. For all other sequences including the *SMP1* cDNA, they refer to the distance from the start of the published sequences. The mismatches in primers rey14b, rnb31, and sf3 were inadvertently introduced. Primers re11d, re014 and re04 do not exactly match dJ469D22, because they were designed from our raw sequences covering the 5' flanking region of *RHD.*

Table 2. Presence of *RHD* flanking sequences in the YAC 38A-A 10

Primer				Amplicons obtained with		
				Genomic DNA		YAC
sense	antisense	Predicted position	Amplicon size	RHD^+	RHD^-	38A-A10
rend9a1	rend9a2 yes	*RHD* 5' flanking region	about 85,000 bp from ATG	948 bp	yes	yes

(continued)

Primer			Amplicon	Genomic DNA		YAC
sense	antisense	Predicted position	size	RHD^+	RHD^-	38A-A10
rend8b1	rend8b2 yes	*RHD* 5' flanking region	about 50,000 bp from ATG	845 bp	yes	yes
rea7.	rez2 yes	*RHD* 3' flanking region	about 1,500 bp from STOP	1,412 bp	yes	no
rend32	sr9 yes	*RHCE* 3' flanking region	about 20,000 bp from STOP	1,989 bp	yes	yes
sr1	sf3 yes	*RHCE* 3' flanking region	about 1,000 bp from STOP	477 bp	yes	yes
rey14b	rey14a no	*RHCE* 5' flanking region	about 5,300 bp from ATG	929 bp	yes	yes
rey7	rey8 no	*RHCE* 5' flanking region	about 10,000 bp from ATG	530 bp	yes	yes

Table 3. PCR-RFLP for the specific detection of the *RHD* deletion

			Number of samples with *RHD* genotype						
	Known genotype	Samples tested (n)	determined			expected*			
Phenotype			+/+	+/-	-/-	+/+	+/-	-/-	p¶
Known genotype									
ccddee	cde/cde	14	0	0	14	0	0	14	N.A.
CCddee	Cde/Cde†	5	0	0	5	0	0	5	N.A.
ccddEE	cdE/cdE†	1	0	0	1	0	0	1	N.A.
D variants	D/cde‡	9	0	9	0	0	9	0	N.A.
ccDEe	cDe/cDE§	4	4	0	0	4	0	0	N.A.
Common phenotypes									
CcDee		10	1	9	0	0.5	9.5	0	>0.4
ccDEe		10	0	10	0	0.3	9.7	0	>0.5
ccDee		10	1	9	0	0.5	9.5	0	>0.4
CCDee		10	9	1	0	9.5	0.5	0	>0.4
CcDEe		12	11	1	0	11	1	0	>0.5
ccDEE		10	10	0	0	9.2	0.8	0	>0.4
CCDEe		6	5	1	0	5.8	0.2	0	>0.1

*Expected number of RHD^+/RHD^+ and RHD^+/RHD^- samples based on known genotypes or the haplotype frequencies in the local population [41]

†*RHD*-negative in PCR.

‡ RHD^+/RHD^-, because a weak or partial D phenotype would be masked in a RHD^+/RHD^+ gentoype. These samples were weak D type 1 (n=2), type 2 (n=2), type 3 (n=2), type 4 (n=2) and D^{VII} (n=1).

§ Presence of two *RHD* genes differing in their polymorphic *Hae*III-site in intron 3 [42] demonstrated by PCR-RFLP.

N.A. - not applicable. Probabilities were calculated based on confidence limits of binomial distribution.

Table 4. *RHD* PCR-SSP

Region	Name* Reference	DNA sequence	Position		Polymorphisms detected	Amplicon size	
Promoter	re012	tccactttccacctccctgc	Promoter	-1,137 to -1,1197 bp deletion at - 1125		255	43
	re011d	gcagccaacttcccctgtg	Promoter	-883 to -901	4 bp deletion at -896		44
Exon 3	ga31 (D-3-383)	ttgtcggtgctgatctcagtgga	Exon 3	361 to 383	383 A	154	21
	rb21	aggtccctcctccagcac	Intron 3	28 to 11			42
Exon 4	ga41 (D-4-527)	acatgatgcacatctacgtgttcgc	Exon 4	503 to 527		123	21
	ga42 (D-4-602)	cagacaaactgggtatcgttgctg	Exon 4	625 to 602	602 C		21
Intron 4	re41	cgatacccagtttgtctgccatgc	Exon 4	608 to 631		226	this study
	rb12	tcctgaacctgctctgtgaagtgc	Intron 4	198 to 175	Intron 4 deletion in *RHD*		40
Exon 5	rb24	agacctttggagcaggagtg	Intron 4	-53 to -34		228	40
	ga51 (D-5-787)	ctgctcaccttgctgatcttccc	Intron 5/Exon 5		8 to 787	787 G	21
Exon 6	ga62 (D-6-826)	ttatgtgcacagtgcggtgttgg	Exon 6	804 to 826		133	21
	ga61 (D-6-916)	caggtacttggctcccccgac	Exon 6	936 to 916	916 G		21
Exon 7†	ga71 (D-7-967)	gttgtaaccgagtgctggggattc	Exon 7	944 to 967		123	21
	ga72 (D-7-1048)	tgccggctccgacggtatc	Exon 7	1,066 to 1,048	1,048 G		21
Exon 7†	rb26	aggggtgggtagggaatatg	Intron 6	-62 to -43		130	42
	re71	acccagcaagctgaagttgtagcc	Exon 7	1,008 to 985	985/986 GG		42
Intron 7	rb52	ccaggttgttaagcattgctgtacc	Intron 7	6,666 to 6,690	6,690 C		42
	rb51	gcatgacgtgttctgcctcttg	Intron 7	6,734 to 6,713	6,713 C	169	this study
							42
Exon 9	re83	gagattaaaaatcctgtgctcca	Intron 8	-56 to -34		119	42
	re94	cttggtcatcaaaatatttagcct	Exon 9	1,216 to 1,193	1,193 A		this study

(continued)

Region	Name* Reference	DNA sequence	Position		Polymorphisms detected	Amplicon size	
Exon 10	(3' UTR) rea7	tgttgcctgcatttgtacgtgag	3'UTR	1,310 to 1,333	*RHD*/*Rhesus box*	*23*	44
	rr4	agcttactggatgaccacca	3'UTR	1,541 to 1,522	junction		42

* Primer names in brackets are as described by Gassner et al. [21].
† Primer set ga71/ga72 was used for the screening, primer set rb26-re71 for *RHD* exon specific PCR-SSP.

Table 5. Population survey of known D negative blood donors screened by *RHD* PCR-SSP

Documented phenotype	Samples (n)	
	screened	PCR-SSP positive*
ccddee	314	0
Ccddee	433	34
ccddEe	271	5
CCddee	24	4
CcddEe	19	4
ccddEE	6	1
CcddEE	1	0
Total	1,068	48

Positive for at least one of four *RHD* specific polymorphisms tested (promoter, intron 4, exon 7 or 3' UTR).

Table 6. PCR patterns compatible with *RHD-RHCE-RHD* hybrid genes or partial *RHD* deletions in 25 D negative samples

PCR pattern	*RHD* specific PCR-SSP* P E3E4I4 E5E6E7I7 E9E10	Samples Possible cause†	Phenotype (n)	Documented	Confirmed	Haplotype	Reference‡
Pattern 1	+ - - - - - - - - +	*RHD-CE(3-9)-D*	11	Ccddee§	D negative	Cde	Whites [1,25], Africans [24]
Pattern 2	+ - - - - - - - + +	*RHD-CE(3-7)-D*	4	Ccddee	D negative	Cde	this study
Pattern 3	+ + - - - - - + +	*RHD-CE(4-7)-D*	3	ccddEe	D negative	cdE	Whites [16]
Pattern 4	+ + - - - - - + + +	*RHD-CE(4-7)-D*	1	CcddEe	D negative	n.k.¶	this study
Pattern 5	+ + - - - + + + + +	*RHD-CE(4-5)-D*	2	ccddEe$^{\alpha}$	partial D/$D_{el}$$^{\alpha}$	cDE	Whites [3,22,30,40]
Pattern 6	+ + + + + + + - - +	*RHD-CE(8-9)-D*	3	CCddee	D negative	Cde	Whites [21]
Pattern 7	- - - - - - - - - +	*RHCE(1-9)-D(10)*	1	ccddEe	D negative	cdE	this study
Pattern 8	- + - - - - - + + +	*RHD(1-3)-CE(4-7)-D*	1	CcddEe	D negative	Cde$^{\beta}$	Africans [5,15]

* P - Promoter, E3 - Exon 3, E4 - Exon 4, 14 Intron 4, E5 - Exon 5, E6 - Exon 6; E7 - Exon 7; 17 - Intron 7; E9 - Exon 9; E10 - Exon 10 (3' UTR)

† Assuming the presence of a single *RHD-CE-D* hybrid allele.

‡ Previously described alleles that fit PCR pattern and haplotype.

§ 11 samples: 9 Ccddee, 1 CCddee, 1 CcddEe

¶ n.k. - not known.

α 2 samples, 1 labeled CcddEe with D_{el} phenotype, 1 labeled ccddEe with partial D D^{VI} phenotype.

β Probably identical to Cde^{s} (see below).

Table 7. *RHD* alleles with single nucleotide substitutions in 22 D negative samples

Allele	Substitution Reference	Effect(s)	Samples (n)	Phenotype Documented	Phenotype Confirmed	Haplotype	
RHD(W16X) study	G->A at 48	Stop codon at codon 16	2	Ccddee	D negative	Cde	this
RHD(G486 (+1)A) study	g->a at 486+1	5' splice site intron 3 ACgt->ACat	3	Ccddee	D_{el}	CDe	this
RHD(G212V) study	G->T at 635	3' splice site intron 4 agGC->agTC Missense mutation G212V	1	Ccddee	D negative	Cde	this
RHD(C285Y) study[1]	G->A at 854	Missense mutation C285Y	1	ccddEe	partial D*	cDE	this
RHD(M295I)	G->T at 885	Missense mutation M2951	7	Ccddee	D_{el}	CDe†	42
RHD(Y330X) study	C->G at 985	Stop codon at codon 330	1	Ccddee	D negative	Cde	this
RHD(G1153 (+1)A) study	g->a at 1153+1	5' splice site intron 8 AGgt->AGat	1	Ccddee	D negative	Cde	this
RHD(G385A)	G->C at 1154	3' splice site intron 8 agGT->agCT Missense mutation G385A	1	CcddEe	weak D	cDE	42
RHD(K409K) study	G->A at 1227	5' splice site intron 9 AGgt->AAgt	5	Ccddee	D_{el}	CDe	this

* A detailed serologic analysis of this sample representing the partial D DIM has been published previously [43].
† The same allele occurring in a cDe haplotype has been described as weak D type 11.

Table 8. Estimated frequencies in population

PCR pattern/Allele	Frequency among Cde/ cdE	Frequency in population
Pattern 1	1:45	1:4,132
Pattern 2	1:125	1:11,364
Pattern 3	1:101	1:17,976
Pattern 4	1:500*	1:45,455*
Pattern 6	1:167	1:15,152
Pattern 7	1:302	1:53,929
Pattern 8	1:500	1:45,455
RHD(W16X)	1:250	1:22,727
RHD(G212V)	1:500	1:45,455
RHD(Y330X)	1:500	1:45,455

(continued)

PCR pattern/Allele	Frequency among Cde/ cdE	Frequency in population
RHD(G1153(+1)A)	1:500	1:45,455
Any D negative	1:20 / 1:67†	1:1,607
RHD(G486(+1)A)	1:167	1:15,152
RHD(M2951)	1:71	1:6,493
RHD(K409K)	1:100	1:9,091
Any D_{el}	1:33‡	1:3,030

* Assuming a Cde haplotype; a cdE haplotype would result in a frequency of 1:302 among cdE and 1:53,929 in the population. For statistics and sum frequencies, the haplotype was formally counted as 0.5 Cde and 0.5 cdE.
† 1:20 among Cde, 1:67 among cdE.
‡ 1:33 relative to the Cde haplotype.

Table 9. Previously described D negative, *RHD positive* alleles

Allele	Haplotype	Population	Possible match
RHD(Q41 X) [2]	Cde	Whites	not detected
RHD-CE(2-9)-D[1,24,25]	Cde	Whites [1,25], Blacks[24]	Pattern 1
RHD-CE(3:455-7)-D [5,15]	Cde[5]	Blacks	Pattern 8
RHD(488del4)[1]	Cde	Whites	not detected
RHD-CE(4-7)-D[16]	cdE	Whites	Pattern 3 or 4
RHDΨ[38]	cde	Blacks	not detected
RHD(600del) [10]	Cde	somatic mutation*	not detected
RHD (exon 5 variant) [8]	cde	not communicated	not detected
RHD(G314V) [34]	Cde	Japanese	not detected
RHD(exon 9 variant)	Cde	Whites	Pattern 6

* Allele acquired by somatic mutation in a woman with chronic myelogenic leucemia and restricted to the myeloid lineage

References

[0124]

1. Andrews KT, Wolter LC, Saul A, Hyland CA: The RhD- trait in a white patient with the RhCCee phenotype attributed to a four-nucleotide deletion in the *RHD* gene. Blood 92:1839, 1998

2. Avent ND, Martin PG, Armstrong-Fisher SS, Liu W, Finning KM, Maddocks D, Urbaniak SJ: Evidence of genetic diversity underlying Rh D negative, weak D (D^u) and partial D phenotypes as determined by multiplex PCR analysis of the RHD gene. Blood 89:2568, 1997

3. Avent ND, Liu W, Jones JW, Scott ML, Voak D, Pisacka M, Watt J, Fletcher A. Molecular analysis of Rh transcripts and polypeptides from individuals expressing the D^{VI} variant phenotype: an *RHD* gene deletion event does not generate all D^{VI}ccEe phenotypes. Blood 1997;89:1779-86.

4. Blumenfeld OO, Huang CH: Molecular genetics of the glycophorin gene family, the antigens for MNSs blood groups: multiple gene rearrangements and modulation of splice site usage result in extensive diversification. Hum Mutat 6:199, 1995

5. Blunt T, Daniels G, Carritt B: Serotype switching in a partially deleted *RHD* gene. Vox Sang 67:397, 1994

6. Bowman J: The management of hemolytic disease in the fetus and newborn. Semin Perinatol 21:39,1997

7. Carritt B, Kemp TJ, Poulter M: Evolution of the human RH (rhesus) blood group genes: a 50 year old prediction (partially) fulfilled. Hum Mol Genet 6:843, 1997

8. Carritt B, Steers FJ, Avent ND. Prenatal determination of fetal RhD type. Lancet 1994;344:205-6.

9. Cherif-Zahar B, Mattei MG, Le Van Kim C, Bailly P, Cartron JP, Colin Y: Localization of the human Rh blood group gene structure to chromosome region 1 p34.3-1 p36.1 by in situ hybridization. Hum Genet 86:398, 1991

10. Cherif-Zahar B, Bony V, Steffensen R, Gane P, Raynal V, Goosens D, Laursen JS, Varming K, Jersild C, Cartron JP. Shift from Rh-positive to Rh-negative phenotype caused by a somatic mutation within the RHD gene in a patient with chronic myelocytic leukaemia. Br J Haematol 1998;102:1263-70.

11. Colin Y, Cherif-Zahar B, Le Van Kim C, Raynal V, van Huffel V, Cartron J-P. Genetic basis of the RhD-positive and RhD-negative blood group polymorphism as determined by southern analysis. Blood 1991;78:2747-52.

12. Cossu G, Angius A, Gelfi C, Righetti PG: Rh D/d genotyping by quantitative polymerase chain reaction and capillary zone electrophoresis. Electrophoresis 17:1911; 1996

13. Daniels GL, Faas BH, Green CA, Smart E, Maaskant-van Wijk PA, Avent ND, Zondervan HA, von dem Borne AE, van der Schoot CE. The VS and V blood group
polymorphisms in Africans: a serologic and molecular analysis. Transfusion 1998;38:951-8.

14. Döscher A, Schunter F, Müller TH: Quantitativae PCR to assess *RHD* zygosity. Infusionsther Transfusionsmed 26(suppl 1):31, 1999 (abstr.)

15. Faas BHW, Becker EAM, Wildoer P, Ligthart PC, Overbeeke MAM, Zondervan HA, von dem Borne AEGK, van der Schoot CE: Molecular background of VS and weak C expression in blacks. Transfusion 37:38, 1997

16. Faas BHW, Beckers EAM, Simsek S, Overbeeke MAM, Pepper R, van Rhenen DJ, von dem Borne AEGK, van der Schoot CE: Involvement of Ser103 of the Rh polypeptides in G epitope formation. Transfusion 36:506, 1996

17. Filbey D, Hanson U, Wesstrom G: The prevalence of red cell antibodies in pregnancy correlated to the outcome of the newborn: a 12 year study in central Sweden. Acta Obstet Gynecol Scand 74:687, 1995

18. Flegel WA, Khull SR, Wagner FF. Primary anti-D immunization by weak D type 2 RBCs. Transfusion 2000; 40: 428 - 434

19. Flegel WA, Wagner FF, Müller TH, Gassner C: Rh phenotype prediction by DNA typing and its application to practice. Transfus Med 8:281, 1998

20. Fujiwara H, Okuda H, Omi T, Iwamoto S, Tanaka Y, Takahashi J, Tani Y, Minakami H, Araki S, Sato I, Kajii E: The STR polymorphisms in intron 8 may provide information about the molecular evolution of *RH* haplotypes. Hum Genet 104:301, 1999

21. Gassner C, Schmarda A, Kilga-Nogler S, Jenny-Feldkircher B, Rainer E, Müller TH, Wagner FF, Flegel WA, Schönitzer D: RhesusD/CE typing by polymerase chain reaction using sequence-specific primers. Transfusion 37: 1020, 1997

22. Huang CH. Human D^{VI} category erythrocytes: correlation of the phenotype with a novel hybrid RhD-CE-D gene but not an internally deleted RhD gene. Blood 1997;89:1834-9.

23. Huang CH, Reid ME, Chen Y, Coghlan G, Okubo Y: Molecular definition of red cell Rh haplotypes by tightly linked *Sph*I RFLPs. Am J Hum Genet 58:133, 1996

24. Huang CH: Alteration of RH gene structure and expression in human dCCee and DCW- red blood cells: phenotypic homozygosity versus genotypic heterozygosity. Blood 88:2326, 1996

25. Hyland CA, Wolter LC, Saul A. Three unrelated Rh D gene polymorphisms identified among blood donors with Rhesus CCee (r'r') phenotypes. Blood 1994; 84:321-4.

26. Kemp TJ, Poulter M, Carritt B: Microsatellite variation within the human *RHCE* gene. Vox Sang 77:159, 1999

27. Lo Y-MD, Bowell PJ, Selinger M, Mackenzie IZ, Chamberlain P, Gillmer MDG, Littlewood TJ, Fleming KA, Wainscoat JS: Prenatal determination of fetal RhD status by analysis of peripheral blood of rhesus negative mothers. Lancet 341:1147, 1993

28. Maaskant-van Wijk PA, Hemker MB, Douglas-Berger L, et al. Ethnic variability of the Rhesus system. Transfusion 1999;39:(10S)103S-4S (Abstract).

29. Maaskant-van Wijk PA, Faas BHW, de Ruijter JAM, Overbeeke MAM, von dem Borne AEGK, van Rhenen DJ, van der Schoot CE. Genotyping of RHD by multiplex polymerase chain reaction analysis of six *RHD*-specific exons. Transfusion 1998;38:1015-21.

30. Maaskant-van Wijk PA, Beckers EAM, van Rhenen DJ, Mouro I, Colin Y, Cartron J-P, Faas BHW, van der Schoot CE, Apoil PA, Blancher A, et al. Evidence that the RHD(VI) deletion genotype does not exist. Blood 1997; 90:1709-11.

31. MacGeoch C, Mitchell CJ, Carritt B, Avent ND, Ridgwell K, Tanner MJ, Spurr NK: Assignment of the chromosomal locus of the human 30-kDal Rh (rhesus) blood group-antigen-related protein (Rh30A) to chromosome region 1p36.13----p34. Cytogenet Cell Genet 59:261, 1992

32. Mourant AE, Kopec AC, Domaniewska-Sobczak K: The distribution of the human blood groups and other polymorphisms. London, Oxford University Press, 1976

33. Okuda H, Suganuma H, Tsudo N, Omi T, Iwamoto S, Kajii E: Sequence analysis of the spacer region between the *RHD* and *RHCE* genes. Biochem Biophys Res Commun 263:378, 1999

34. Okuda H, Kawano M, Iwamoto S, Tanaka M, Seno T, Okubo Y, Kajii E: The RHD gene is highly detectable in RhD-negative Japanese donors. J Clin Invest 100:373, 1997

35. Onda M, Fukuda M: Detailed physical mapping of the genes encoding glycophorins A, B and E, as revealed by P1 plasmids containing human genomic DNA. Gene 159:225, 1995

36. Reboul J, Gardiner K, Monneron D, Uze G, Lutfalla G: Comparative genomic analysis of the interferon/interleukin-10 receptor gene cluster. Genome Res 9:242, 1999

37. Salvignol I, Calvas P, Socha WW, Colin Y, Le Van Kim C, Bailly P, Ruffie J, Cartron JP, Blancher A: Structural analysis of the *RH*-like blood group gene products in nonhuman primates. Immunogenetics 41:271, 1995

38. Singleton BK, Green CA, Avent ND, Martin PG, Smart E, Daka A, Narter-Olaga EG, Hawthorne LM, Daniels G: The presence of an *RHD* pseudogene containing a 37 base pair duplication and a nonsense mutation in africans with the Rh D- negative blood group phenotype. Blood 95:12, 2000

39. Urbaniak SJ, Robertson AE: A successful program of immunizing Rh-negative male volunteers for anti-D production using frozen/thawed blood. Transfusion 21:64, 1981

40. Wagner FF, Gassner C, Müller TH, Schönitzer D, Schunter F, Flegel WA: Three molecular structures cause Rhesus D category VI phenotypes with distinct immunohematologic features. Blood 91:2157, 1998

41. Wagner FF, Kasulke D, Kerowgan M, Flegel WA: Frequencies of the blood groups ABO, Rhesus, D category VI, Kell, and of clinically relevant high-frequency antigens in South-Western Germany. Infusionsther Transfusionsmed 22:285, 1995

42. Wagner FF, Gassner C, Müller TH, Schönitzer D, Schunter F, Flegel WA: Molecular basis of weak D phenotypes. Blood 93:385, 1999

43. Wagner FF, Frohmajer A, Ladewig B, Eicher NI, Lonicer CB, Müller TH, Siegel MH, Flegel WA. Weak D alleles express distinct phenotypes. Blood 2000; 95: 2699 - 2708

44. Wagner FF, Flegel WA. *RHD* gene deletion occurred in the *Rhesus box*. Blood 2000;

45. Wissenschaftlicher Beirat der Bundesärztekammer; Paul-Ehrlich-Institut. Richtlinien zur Blutgruppenbestimmung und Bluttransfusion (Hämotherapie). Köln: Deutscher Ärzte-Verlag; 1996.

46. Legler TJ, Kohler M, Mayr WR, Panzer S, Ohto H, Fischer GF; Genotyping of the human platelet antigen systems 1 through 5 by multiplex polymerase chain reaction and ligation-based typing. Transfusion 1996 May; 36(5): 426-31

SEQUENCE LISTING

[0125]

<110> DRK Blutspendedienst Baden-Wuerttemberg gGmbH

<120> Molecular structure of RHD negative

<130> D 2600 PCT

<140> PCT/EP0O/1074
<141> 2000-10-31

<160> 66

<170> PatentIn Ver. 2.1

<210> 1
<211> 100
<212> DNA
<213> Homo sapiens

<400> 1

```
gcatgcgcga ctgagccggg tggatggtac tgctgcatcc gggtgtctgg aggctgtggc 60
cgttttgttt tcttggctaa aatcggggga gtgaggcggg                        100
```

<210> 2
<211> 26
<212> DNA
<213> Homo sapiens

<400> 2
actgcacgac ggggctggac tgacgt 26

<210> 3
<211> 45
<212> DNA
<213> Homo sapiens

<400> 3
agctgaaaaa aatgtctgga tttctagagg gcttgagatg ctcag 45

<210> 4
<211> 26
<212> DNA

<213> Homo sapiens

<400> 4
gcttccattg ctgctggtgt actagt 26

<210> 5
<211> 44
<212> DNA
<213> Homo sapiens

<400> 5
agtttttac aggctggtgg attatcatag atgcagctgt tatt 44

<210> 6
<211> 26
<212> DNA
<213> Homo sapiens

<400> 6
tatagcaacc atagccttcc taatgt 26

<210> 7
<211> 44
<212> DNA
<213> Homo sapiens

<400> 7
aggattaatg cagtatcgaa tggacaagtc cgaggtgata gtta 44

<210> 8
<211> 26
<212> DNA
<213> Homo sapiens

<400> 8
tgaaggttgt ctgggtcaaa acaggt 26

<210> 9
<211> 44
<212> DNA
<213> Homo sapiens

<400> 9
aggtgctcgc atttggcttt tcgttggttt catgttggcc tttg 44

<210> 10
<211> 26
<212> DNA
<213> Homo sapiens

<400> 10
tttttggagg ttatgttgct aaaggt 26

<210> 11
<211> 38
<212> DNA
<213> Homo sapiens

<400> 11
agaaaaagac atagtatacc ctggaattgc tgtatttt 38

<210> 12
<211> 26
<212> DNA
<213> Homo sapiens

<400> 12
tccagaatgc cttcatcttt tttggt 26

<210> 13
<211> 53
<212> DNA
<213> Homo sapiens

<400> 13
aggagggctg gtttttaagt ttggccgcac tgaagactta tggcagtgaa cac 53

<210> 14
<211> 249
<212> DNA
<213> Homo sapiens

<400> 14

agcatcatcc taatgaaact aaacatttat tttaaactta ttaaattgac tcttaaacta 60
agtttttagt ctttaatttt ttaatatcaa atctgtctct gaccttgttt cattatacat 120
aaggagcttt gctgtcatga gcgtttctca cgtacaaatg caggcaacag tgagaggaag 180
ttgtcttgtt tttgaacagg ccttgttttt cttggatgct tttgcttaaa atccaacagc 240
caaatgagg 249

<210> 15
<211> 3093
<212> DNA
<213> Homo sapiens

<400> 15

```
attcactatc acaagaacag cacgggtaag acctgtcccc atgattcagt tacctcccac 60
tgggtccctc ccacaacgca tgggaattca ggatgagatt tgggtgggga cacaaccaaa 120
ccctatcatt ccacccatgg cccctcccaa atttcatgtc ctcacatttc aaaaccaatc 180
acaccatccc aacagtccct caaagtctta aatgatttca gcattaactc aaaagtccac 240
agtctaatgt ctcatctgag acaaggcaag tcctttccat ttatgagcct ataaaatcca 300
aagcaagtta gttacttcct agatacaatg ggggtacagg cattgggtaa atacagccat 360
tccaaatggg ataaattggt caaaacaaag aggctacagg cccatgagag tccaaaatcc 420
agtggggcag tcaaatctta aagctccaaa atgatctcct ttgactccac atctcacatc 480
caggtcacgc agatggaagg ggtgggttcc catggtcttg ggcagctctg cccctgtacc 540
tttgcagggt acagcctccc tctcagctgc tttcatgggc tggcattgag tgtctgcagc 600
ttttccaggt acacggtgca agctgtcggt ggatctacca ttctggggtc tggaggacct 660
cttctcacag ctccactagg tggtgcccca gtagggactg tgtgtggggt ctctgacccc 720
acatttccct tctgcactgc cctggcagag gatctccatg agggccctgc tcctgcagca 780
aacttctgac tgggcatcca ggcatttccg cacatcctct ttaatctagg cgaaggtttc 840
caaaccccaa ttcttgactt ctgtgcactc gcagtctcaa caccacatgg aagctgtcaa 900
ggcttggggc ttgcactccc cgaagctaca gcccaagctc taccttgcct cccgtcagtc 960
atggttggga gtggctggga tgcagggcac caagtcccta ggctgcacac agcatgagga 1020
ccccgggcct ggccaacaaa accatttttt cctgatacct ctggacctgt gatgggaggg 1080
gttgccataa agacctctga catgccctgg agacattttc cccattgtct tgggaattag 1140
catttggctc ctgttactca tgcaaatttc tgcagccagc ttgaatttct cctcagaaaa 1200
tgggaatttt tcttttctat cacattgtca ggctgcaaat tttccgaact tttatgctct 1260
gcttccctta taaaactgaa tgtctttaac agcacccaag tcacctcttg aatgctttgc 1320
tgcttagaaa tttctcctgc cagatactct aaatcatctc tctgaagttc aaagttctac 1380
aaatatctcg tgcaggggca aaatgccgcc agtatctttg ctaaaacata acaagagtcc 1440
cctttgctcc agttcccaac aagttcctca tttccgtctg agaccacctc agcctatgga 1500
ctttattgtc cacagtgcta tcagcatttt gggcaaagcc attcaacaag tctctaggaa 1560
gttccaaact ttcccacatt tgcctgtctt cttctgagcc ctccaaactg ttccaaaccc 1620
tgcctgttac ccagttccaa agtcacatac ccatttttga gtatctacgg cagcacccca 1680
ctctactggt accaatttag ccactgaagt agttggagaa cagaagtaat agactctggt 1740
ttacattgta aaagcttctc tgtggctgct gtgtgaagaa aatatatgag aatgaagccc 1800
caagatgaag cagggacaca gttgcagtgg ttagagtaag aaatgctgct ggctggcact 1860
gaagtgatag cctggaggtt tgtgtgtgca catgcatgtg tatgtgtttt acgatagtag 1920
gcccaacaga tactgtaatc cacacttgtt ttttttttt gagacagagt ctcacctgtt 1980
gcctagacta gaatgcagtg gcacaatctt ggctcactac aacctccacc tcccaggttc 2040
aaacaatcct tgtgcttcag cctcccgagt agttgggatt acaggtgtgt gccaccgtgc 2100
ccagctatat tttttgtatt tttagcagag atgggatttt gccacattgg ccaggctggt 2160
cttgaactcc tggcctcaag caatcctccc accttagcct cccaaagtgc tgagccacca 2220

cacctggccg caactgattt ttaatcatga aatgacacat acatttaaaa aacccaatac 2280
ctataatatt cctggctagt actcttcaca tctatatcat caaaaacaaa gaaagtatgt 2340
gaaactgaca cagccaaggg gagactaagg agacataaca attaactgta atgtggtatt 2400
ctggagggga tcctggaaca gaaaaagaca ttaggcaaaa aactaaagaa atctgaataa 2460
aatgtggatg tcagttaata ataatgtatc atattagtcc agtaattgta acaaatatac 2520
cacaataatg aaagccatta attataggga aaatggaggg gttaatatgg gtggctggct 2580
tttgctattt ctagcagctc cattttatct gcaaaagaca aacattcatt aagtcccaaa 2640
aaggtaaaga atgacaaatt aagcatgtat cttattagta agagtaatat aaagatgctc 2700
actcctattt ataaatattt gacaatcatg ttaaggccac aaaagagaaa aaagggtagg 2760
ggcaaaaaac gcaaagagaa aggagttagt atcttttctc ccgcactcat tagctattaa 2820
aagaggatgt ttgtttaaag ctgctcagag ctggtaaact aatgttaagt cactaacggg 2880
aatttaaaag gtttcattaa gaactgcctg cactagattc ctccaccctg agacattaaa 2940
caatcacgat aaacctcctg agtggtaaga acttgtccat ttaaaaacag gctatagatt 3000
gtatcatgca gttttatcta ctaatcggct aatatcccgc caaaaacaaa aaaccccaaa 3060
gggatgaaag tttcatccat caaaggaaac aac                              3093
```

<210> 16

<211> 3099 hybrid Rhesus
<212> DNA
<213> Homo sapiens

<400> 16

attcactatc acaagaacag cacgggtaag acctgtcccc atgattcagt tacctcccac 60
tgggtccctc ccacaacgca tgggaattca ggatgagatt tgggtgggga cacaaccaaa 120
ccctatcatt ccacccatgg cccctcccaa atttcatgtc ctcacatttc aaaaccaatc 180
acaccatccc aacagtccct caaagtctta aatgatttca gcattaactc aaaagtccac 240
agtctaatgt ctcatctgag acaaggcaag tcctttccat ttatgagcct ataaaatcca 300
aagcaagtta gttacttcct agatacaatg ggggtacagg cattgggtaa atacagccat 360
tccaaatggg ataaattggt caaaacaaag aggctacagg cccatgagag tccaaaatcc 420
agtggggcag tcaaatctta aagctccaaa atgatctcct ttgactccac atctcacatc 480
caggtcacgc agatggaagg ggtgggttcc catggtcttg ggcagctctg cccctgtacc 540
tttgcagggt acagcctccc tctcagctgc tttcatgggc tggcattgag tgtctgcagc 600
ttttccaggt acacggtgca agctgtcggt ggatctacca ttctggggtc tggaggacct 660
cttctcacag ctccactagg tggtgcccca gtagggactg tgtgtggggt ctctgacccc 720
acatttccct tctgcactgc cctggcagag gatctccatg agggccctgc tcctgcagca 780
aacttctgac tgggcatcca ggcatttccg cacatcctct ttaatctagg cgaaggtttc 840
caaaccccaa ttcttgactt ctgtgcactc gcagtctcaa caccacatgg aagctgtcaa 900
ggcttggggc ttgcactccc cgaagctaca gcccaagctc taccttgcct cccgtcagtc 960
atggttggga gtggctggga tgcagggcac caagtcccta ggctgcacac agcatgagga 1020
ccccgggcct ggccaacaaa accatttttt cctgatacct ctggacctgt gatgggaggg 1080
gttgccataa agacctctga catgccctgg agacattttc cccattgtct tgggaattag 1140
catttggctc ctgttactca tgcaaatttc tgcagccagc ttgaatttct cctcagaaaa 1200
tgggaatttt tcttttctat cacattgtca ggctgcaaat tttccgaact tttatgctct 1260
gcttccctta taaaactgaa tgtctttaac agcacccaag tcacctcttg aatgctttgc 1320
tgcttagaaa tttctcctgc cagatactct aaatcatctc tctgaagttc aaagttctac 1380
aaatatctcg tgcaggggca aaatgccgcc agtatctttg ctaaaacata acaagagtcc 1440
cctttgctcc agttcccaac aagttcctca tttccgtctg agaccacctc agcctatgga 1500
ctttattgtc cacagtgcta tcagcatttt gggcaaagcc attcaacaag tctctaggaa 1560
gttccaaact ttcccacatt tgcctgtctt cttctgagcc ctccaaactg ttccaaaccc 1620
tgcctgttac ccagttccaa agtcacatac ccatttttga gtatctacgg cagcacccca 1680
ctctactggt accaatttag ccactgaagt agttggagaa cagaagtaat agactctggt 1740
ttacattgta aaagcttctc tgtggctgct gtgtgaagaa aatatatgag aatgaagccc 1800
caagatgaag cagggacaca gttgcagtgg ttagagtaag aaatgctgct ggctggcact 1860
gaagtgatag cctggaggtt tgtgtgtgca catgcatgtg tatgtgtttt acgatagtag 1920
gcccaacaga tactgtaatc cacacttgtt ttttttttct ttttgagaca gagtctcacc 1980
tgttgcctag actagaatgc agtggcacaa tcttggctca ctacaacctc cacctcccag 2040
gttcaaacaa tccttgtgct tcagcctccc gagtagttgg gattacaggt gtgtgccacc 2100
gtgcccagct atatttttttg tatttttagc agagatggga ttttgccaca ttggccaggc 2160
tggtcttgaa ctcctggcct caagcaatcc tcccacctta gcctcccaaa gtgctgagcc 2220
accacacctg gccgcaactg atttttaatc atgaaatgac acatacattt aaaaaaccca 2280
atacctataa tattcctggc tagtactctt cacatctata tcatcaaaaa caaagaaagt 2340
atgtgaaact gacacagcca aggggagact aaggagacat aacaattaac tgtaatgtgg 2400

```
tattctggag gggatcctgg aacagaaaaa gacattaggc aaaaaactaa agaaatctga 2460
ataaaatgtg gatgtcagtt aataataatg tatcatatta gtccagtaat tgtaacaaat 2520
ataccacaat aatgaaagcc attaattata gggaaaatgg aggggttaat atgggtggct 2580
ggcttttgct atttctagca gctccatttt atctacaaaa gacaaacatt cattaagtcc 2640
caaaaaggta aagaatgaca aattaagcat gtatcttatt agtaagagta atataaagat 2700
gctcactcat atttataaat atttgacaat gatgttaagg ccagaaaaga gaaaaaaggg 2760
taggggcaaa aaacgcaaag agaaaggagt tagtatcttt tctcccgcac tcattagcta 2820
ttaaaagagg atgtttgttt aaagctgctc agagctggta aactaatgtt aagtcactaa 2880
cgggaattta aaaggtttca ttaagaactg cctgcactag attcctccac cctgagacat 2940
taaacaatca cgataaacct cctgagtggt aagaacgtgt ccatttaaaa acaggctata 3000
gattgtcatg cagttttatc tactaatcgg ctaatgcacc gccaaaaaca aacaaaaaaa 3060
cccaaaggga tgaaagtttc atccatcaaa ggaaacaac                        3099
```

<210> 17
<211> 3099
<212> DNA
<213> Homo sapiens

<400> 17

attcactatc acgagaacag cacgggtaag acctgtcccc atgattcagt tacctcccac 60
tgggtccctc ccacaacgca tgggaattca ggatgagatt tgggtgggga cacaaccaaa 120
ccctatcatt ccacccatgg cccctcccaa atttcatgtc ctcacatttc aaaaccaatc 180
acaccatccc aacagtccct caaagtctta aatgatttca gcattaactc aaaagtccac 240
agtctaatgt ctcatctgag acaaggcaag tcctttccat ctatgagcct ataaaatcca 300
aagcaagtta attacttcct agatacaatg ggggtacagg cattgggtaa atacagccat 360
tccaaatggg ataaattggt caaaacaaag aggctacagg cccatgagag tccaaaatcc 420
agtggggcag tcaaatctta aagctccaaa atgatctcct cttgactcca catctcacat 480
ccaggtcacg cagatggaag gggtgggttc ccatggtctt gggcagctct gcccctgtac 540
ctttgcaggg tacagcctcc ctctcagctg ctttcatggg ctggcattga gtgtctgcaa 600
cttttccagg tacacggtgc aagctgtcgg tggatctacc attctggggt ctggaggacc 660
tcttctcaca gctccactag gtggtgcccc agtagggact gtgtgtgggg tctctgaccc 720
cacatttccc ttctgcactg ccctggcaga ggatctccat gagggccctg cccctgcagc 780
aaacttctgc ctgggcatcc aggcatttcc gcacatcctc tttaatctag gcgaaggttt 840
ccaaacccca gttcttgact tctgtgcact cgcagtctca acaccacatg gaagctgtca 900
aggcttgggg cttgcactcc ccgaagctac agcccaagct ctaccttgcc tcctgtcagt 960
catggttggg agtggctggg atgcagggca ccaagtccct aggctgcaca cagcatgagg 1020
accccgggcc tggccaacaa aaccattttt tcctgatatc tctggacctg tgatgggagg 1080
ggttgccata aagacctctg acatgccctg gagacatttt ccccattgtc ttgggaatta 1140
gcatttggct cctgttactc atgcaaattt ctgcagccag cttgaatttc tcctcagaaa 1200
atgggaattt ttcttttcta tcacattgtc aggctgcaaa ttttccgaac ttttatgctc 1260
tgcttccctt ataaaactga atgtctttaa cagcacccaa gtcacctctt gaatgctttg 1320
ctgcttagaa atttctcctg ccagatactc taaatcatct ctctgaagtt caaagttcta 1380
caaatatctc gtgcaggggc aaaatgccgc cagtatcttt gctaaaacat aacaagagtc 1440
ccctttgctc cagttcccaa caagttcctc atttccgtct gagaccacct cagcctatgg 1500
actttattgt ccacagtgct atcagcattt tgggcaaagc cattcaacaa gtctctagga 1560
agttccaaac tttcccacat ttgcctgtct tcttctgagc cctccaaact gttccaaacc 1620
ctgcctgtta cccagttcca aagtcacata cccatttttg agtatctacg gcagcacccc 1680
actctactgg taccaattta gccactgaag tagttggaga acagaagtaa tagactctgg 1740
tttacattgt aaaagcttct ctgtggctgc tgtgtgaaga aaatatatga gaatgaagcc 1800
ccaagatgaa gcagggacac agttgcagtg gttagagtaa gaaatgctgc tggctggcac 1860
tgaagtgata gcctggaggt ttgtgtgtgc acatgcatgt gtatgtgttt tacgatagta 1920
ggcccaacag atactgtaat ccacacttgt ttttttttt ttttgagac agagtctcac 1980
ctgttgccta gactagaatg cagtggcaca atcttggctc actacaacct ccacctccca 2040
ggttcaaaca atccttgtgc ttcagcctcc cgagtagttg ggattacagg tgtgtgccac 2100
cgtgcccagc tatatttttt gtatttttag cagagatggg attttgccac attggccagg 2160
ctggtcttga actcctggcc tcaagcaatc ctcccacctt agcctcccaa agtgctgagc 2220
caccacacct ggccgcaact gatttttaat catgaaatga cacatacatt taaaaaaccc 2280
aatacctata atattcctgg ctagtactct tcacatctat atcatcaaaa acaaagaaag 2340
tatgtgaaac tgacacagcc aaggggagac taaggagaca taacaattaa ctgtaatgtg 2400
gtattctgga ggggatcctg gaacagaaaa agacattagg caaaaaacta aagaaatctg 2460
aataaaatgt ggatgtcagt taataataat gtatcatatt agtccagtaa ttgtaacaaa 2520
tatacccaat aatgaaagcc attaattata gggaaaatgg aggggttaat atgggtggct 2580

ggcttttgct atttctagca gctccatttt atctacaaaa gacaaacatt cattaagtcc 2640
caaaaaggta aagaatgaca aattaagcat gtatcttatt agtaagagta atataaagat 2700
gctcactcat atttataaat atttgacaat gatgttaagg ccagaaaaga gaaaaaaggg 2760
taggggcaaa aaacgcaaag agaaaggagt tagtatcttt tctcccgcac tcattagcta 2820
ttaaaagagg atgtttgttt aaagctgctc agagctggta aactaatgtt aagtcactaa 2880
cgggaattta aaaggtttca ttaagaactg cctgcactag attcctccac cctgagacat 2940
taaacaatca cgataaacct cctgagtggt aagaacgtgt ccatttaaaa acaggctata 3000
gattgtcatg cagttttatc tactaatcgg ctaatgcacc gccaaaaaca aacaaaaaaa 3060
cccaaaggga tgaaagtttc atccatcaaa ggaaacaac 3099

<210> 18
<211> 9241
<212> DNA

<213> Homo sapiens

<400> 18

ctagaaaaca ctttgtcatt ttagaggtgt tatccaatgt tcgcgcaggc actggagtca 60
gagaaaatgg agttgaatcc tttctctgcc actctttgag gagaatctca ccatttatta 120
tgcactgtag aatacaacaa taaaatacag ccatgtacca cataacaaca tcttggtaaa 180
caacagactg catatatgat ggtggtcatc cagtaagcta aggttaattt attattattc 240
ccttttttttt ttctttttttt tgagatgtag tcttactctg tcacccaggc tagagtgcaa 300
tggcaccatc ttggctcact gcaacctctg cctcctgggt tcaagcgaat ctcctgcctc 360
agcctccgaa gtagctggga attacaggca cccaccacat ctggctaatt ttttgtattt 420
ttagtaaaga tggggtttca ccatgttggc caggctgatc tcaaactcct gacctcaagt 480
gatctgcctg cctcggcctc ccaaagtgct gggaccatag gcctgagcca ctgtgcccgg 540
ccttgtttgc ttttttaaca gttaacagtg tgctcataga aactgctttg acatgactgc 600
aatcatgtgc ttcatagaaa cttaattaga ttataccact agagtcttca gatttttata 660
ctttttttttt tgaaacggag tctcactctg tcaccaggct ggagtgcagt gccgcaatct 720
cagctcgccg caacctccgc ctcccaggtt caagtgattc tcctgcctca gcctcccgag 780
tagctgggat tacaagtgca cactaccacg cccagctaat ttttgcattt ttactagaca 840
gggtttcacc atgttggcta ggatagtttc accaggatct cttggcctca tgatcagcct 900
gcctcggcct cccaaagtgc tgggattaca ggtgtgagcc accgtgccca gcctatactt 960
ccctttttga ataccatttg gcgttttgaa gaattaacag ctttgtgaac gtggcagtgc 1020
ttgtgattca ggcttccact gagaccaagg ggagaacctg gttgcaggac aaacagacgg 1080
acagcgtgtg gcagtgttta aatgctcttc tgaaggctga tacgacagct ctctgtgcac 1140
tgattgtata cgcatcccaa gattatatta ttgttttcta ttgctatgtg tcacactttg 1200
ccaaacagga tgtggaaaat gaataagcgg ttttcttagg cacttcttaa cagacaattg 1260
gtcaaaatga actccattgc ttaagaaaca cataaacacc atttagtcac tgaatatagc 1320
tatatgtatg gttgctacta tggggaatct tgttttgcca attttctttg aaaattctgg 1380
cagaccaagg ttctttttgt ttacacaata cttgaaaaat aaaaatgaac aagccaacaa 1440
actaccaagt tttcacttac ataaatgtag ttacatacag aaaatgtgac tgtgaatttt 1500
ttctaggact tttaaactat aagcactatt tgcacgaaag agaaccaatc tatcaattac 1560
aaactcacat aattttacag attttttttt ccctacacag cacataaaac agaaggaatt 1620
tgaagccacc ctccaaacac aggggaagga ggctgtgtgt atatcctcat tgtctttcac 1680
attctaaggt ggttccactc agtgactgaa atccttaagt gttgtattag tcggcttggg 1740
ctaccataac agcagcttaa actgttgtta agccactcag acttaaacaa cagaaattta 1800
tttccttata gttctggagg ctggaagttc aaggtgccgg caaggctggt ttctggtgag 1860
acctctctcc ctgtcttgca gatggctgcc tcctccctgt gtcctcatag agcctgtcct 1920
ctgcttttac acttctggtg tcatcttcct ttttttttttt tttttgagac agagtctcgc 1980
tctatcgccc aggctggagt gcagtggccc gatcgatctc ggctcactgc aacctctgcc 2040
tcccaggttc aagcaattct cctgcctcag cctcccgagt agctgggact acaggtgccc 2100
gccatcatgt ctggctaatt tttgtatttt tagtagagac agggtttcac catattggcc 2160
aggctggtct ccaactcctg accttgtcat ctgcctgcct cggcctccca aagtgctagg 2220
attacaggcg tgagccaccg cacccggcct ctttctcttc ttataaggac accagtccta 2280
ttagattagg gctccaccct catgacctca tttgacctta actattattt ctttaaagca 2340
cctatttcca aatatagtca ctttaggggt tagggcttca aaatatgaat ctgagggaga 2400
tcaattcagt aaatagcagt agtcattaac ggacaatata tacaaagata atttcgtgat 2460
tactgtcctt atgcataaat gtcctcagtg ttccactgcc tttatccaga tttactatca 2520
caaagacttt gctctgagaa aaatgtgatt tctttctttt ttttttttttt ttgagacaga 2580
gtctcactct gtcacccagg ctggagtgca gtggtgcaat ctcggctcac tgcaatctcc 2640
gcctcccagg ttcacgccat tctcttgcct cagtctcccg agtagctggg cctacaggcg 2700
cccgccaccc tgcccagcta atttttttgta tttttagtag agacggggtt tcaccatgtt 2760

agccaggatg gtctcaatct cctgacctcg tgatccacct gcctcagcct cccaaagtgc 2820
tgggattaca ggcatgagcc accgcgccca gcagattttt ttttttttt ttttttttg 2880
agatggagtc ttgctgtgtt gcccagcctg gagtgcagtg ttatgatttt ggctcactgc 2940
aacctctgtc taccatgttc aagcgattct cccacctctg cctcccgtgt agctgggatc 3000
acaggcacac gccaccacac ctagctactt tttgtatttt tagtagaaat ggggtttcac 3060
catgttggcc aggatggtcc cgaactcctg acctcaagtg atcctcctgc ctcggcctcc 3120
caaagtgctg ggattacagg tgtgagccac tgtgcctggc caaaaatgtg atttcttatt 3180
tcccacattg ccaattccat ttcaattaac tataatagct atgtctattg agcactcaag 3240
tgtattctag aaactgttcc tgattctggg gatatatcca tgaatcaact atagtccctg 3300
ttattaagta atctgtagtc tgactaaacc attagaaatt taaaaaatgg ctactttcaa 3360
agacatcttg gagttcagga gtcccacact gcgaaccata ttacctaata atccaacctg 3420
cttgtaattc acttatttaa ccaatattta ttgagtgcca actttgagcc taagatacag 3480
cagtaaacaa atggataaag tccctgtcct catgaaactt gtattctaat ggaagaaaca 3540
gaaaacaaac agatatagga tgtaatatca ggtagggata aatactttga attcaaacaa 3600
aagtatacgt agtcagggtt cgccaaagag acacagccaa tcggatacat agatatataa 3660
aagagggttt atgagttaga aagggctcac atgattacag aggctgagaa gtcccacagc 3720
agattgtctg caagctggag acccagggat actggtagca tggctcagtc caagtcccaa 3780
agcctcagaa tcaggaaagc tgatgatata attcttagcc caaaggcctt agaaccccag 3840
cggtgacgga aaggctgatg taggtcctgg agtcctgaga cccaacagcc tgggatcctg 3900
aaatccaagg gcaggaatgg aagcgtgtat tccagctcca agagagtaag accaatttgc 3960
ctttcttccg tttttgtttc aagccacctg cacattgagg gcggatggtt ccctcttagt 4020
ccattcagtc atatatcaat ctcttctgga aataccctca cagacacact aacaaataat 4080
gcctttccag ttctctaggt attctttaat ccagtcaagc tgacacctaa aattaaccat 4140
cacaaaagtt aaggagaaag aagacaactt gtaggggagg ctgctatgca agacagtgtg 4200
tgaaggaagg gctctctgaa gaggttaata tctgagcaga gacttgaatg aagtgaagaa 4260
gtgagccatg tgggtatggg gaatacaact tccaggtaga gaagacaagt gtggtgtgta 4320
tcagggtcag caaagaagcc atgtgacaga gaagggtggg ccagggagag acggataagt 4380
gatctaactc ctgaggaggt ggcctggcca ggagctagag catgaagatc tcgtaggact 4440
ttattctgca aggtgaaaag ccattgtatt agtctgttca caaacccgag actaggcaat 4500
ttacaaaaga aagagaggtt taatggactt acagttccac atggctgggg aggcctcaca 4560
atcatggcga aaggcaatga ggagcaagtc acgtcttacg tggatggcag gcaaagacaa 4620
agacagcttg tgcagagaaa ctccccctta tagagccatc agatcctgtt agacttattc 4680
actatcacaa gaacagcacg ggtaagacct gtcccatga ttcagttacc tcccactggg 4740
tccctcccac aacgcatggg aattcaggat gagatttggg tggggacaca accaaaccct 4800
atcattccac ccatggcccc tcccaaattt catgtcctca catttcaaaa ccaatcacac 4860
catcccaaca gtccctcaaa gtcttaaatg atttcagcat taactcaaaa gtccacagtc 4920
taatgtctca tctgagacaa ggcaagtcct ttccatttat gagcctataa aatccaaagc 4980
aagttagtta cttcctagat acaatggggg tacaggcatt gggtaaatac agccattcca 5040
aatgggataa attggtcaaa acaaagaggc tacaggccca tgagagtcca aaatccagtg 5100
gggcagtcaa atcttaaagc tccaaaatga tctcctttga ctccacatct cacatccagg 5160
tcacgcagat ggaaggggtg ggttcccatg gtcttgggca gctctgcccc tgtacctttg 5220
cagggtacag cctccctctc agctgctttc atgggctggc attgagtgtc tgcagctttt 5280
ccaggtacac ggtgcaagct gtcggtggat ctaccattct ggggtctgga ggacctcttc 5340
tcacagctcc actaggtggt gccccagtag ggactgtgtg tggggtctct gaccccacat 5400
ttcccttctg cactgccctg gcagaggatc tccatgaggg ccctgctcct gcagcaaact 5460
tctgactggg catccaggca tttccgcaca tcctctttaa tctaggcgaa ggtttccaaa 5520
ccccaattct tgacttctgt gcactcgcag tctcaacacc acatggaagc tgtcaaggct 5580
tggggcttgc actccccgaa gctacagccc aagctctacc ttgcctcccg tcagtcatgg 5640
ttgggagtgg ctgggatgca gggcaccaag tccctaggct gcacacagca tgaggacccc 5700
gggcctggcc aacaaaacca ttttttcctg atacctctgg acctgtgatg ggaggggttg 5760
ccataaagac ctctgacatg ccctggagac attttcccca ttgtcttggg aattagcatt 5820
tggctcctgt tactcatgca aatttctgca gccagcttga atttctcctc agaaaatggg 5880
aatttttctt ttctatcaca ttgtcaggct gcaaattttc cgaactttta tgctctgctt 5940
cccttataaa actgaatgtc tttaacagca cccaagtcac ctcttgaatg ctttgctgct 6000
tagaaatttc tcctgccaga tactctaaat catctctctg aagttcaaag ttctacaaat 6060
atctcgtgca ggggcaaaat gccgccagta tctttgctaa aacataacaa gagtcccctt 6120
tgctccagtt cccaacaagt tcctcatttc cgtctgagac cacctcagcc tatggacttt 6180
attgtccaca gtgctatcag cattttgggc aaagccattc aacaagtctc taggaagttc 6240
caaactttcc cacatttgcc tgtcttcttc tgagccctcc aaactgttcc aaaccctgcc 6300
tgttacccag ttccaaagtc acatacccat ttttgagtat ctacggcagc accccactct 6360
actggtacca atttagccac tgaagtagtt ggagaacaga agtaatagac tctggtttac 6420
attgtaaaag cttctctgtg gctgctgtgt gaagaaaata tatgagaatg aagccccaag 6480
atgaagcagg gacacagttg cagtggttag agtaagaaat gctgctggct ggcactgaag 6540

tgatagcctg gaggtttgtg tgtgcacatg catgtgtatg tgttttacga tagtaggccc 6600
aacagatact gtaatccaca cttgtttttt ttttttttt gagacagagt ctcacctgtt 6660
gcctagacta gaatgcagtg gcacaatctt ggctcactac aacctccacc tcccaggttc 6720
aaacaatcct tgtgcttcag cctcccgagt agttgggatt acaggtgtgt gccaccgtgc 6780
ccagctatat tttttgtatt tttagcagag atgggatttt gccacattgg ccaggctggt 6840
cttgaactcc tggcctcaag caatcctccc accttagcct cccaaagtgc tgagccacca 6900
cacctggccg caactgattt ttaatcatga aatgacacat acatttaaaa aacccaatac 6960
ctataatatt cctggctagt actcttcaca tctatatcat caaaaacaaa gaaagtatgt 7020
gaaactgaca cagccaaggg gagactaagg agacataaca attaactgta atgtggtatt 7080
ctggagggga tcctggaaca gaaaaagaca ttaggcaaaa aactaaagaa atctgaataa 7140
aatgtggatg tcagttaata ataatgtatc atattagtcc agtaattgta acaaatatac 7200
ccaataatga aagccattaa ttatagggaa aatggagggg ttaatatggg tggctggctt 7260
ttgctatttc tagcagctcc attttatcta caaaagacaa acattcatta agtcccaaaa 7320
aggtaaagaa tgacaaatta agcatgtatc ttattagtaa gagtaatata aagatgctca 7380
ctcatattta taaatatttg acaatgatgt taaggccaga aaagagaaaa aagggtaggg 7440
gcaaaaaacg caaagagaaa ggagttagta tcttttctcc cgcactcatt agctattaaa 7500
agaggatgtt tgtttaaagc tgctcagagc tggtaaacta atgttaagtc actaacggga 7560
atttaaaagg tttcattaag aactgcctgc actagattcc tccaccctga gacattaaac 7620
aatcacgata aacctcctga gtggtaagaa cgtgtccatt taaaaacagg ctatagattg 7680
tcatgcagtt ttatctacta atcggctaat gcaccgccaa aaacaaacaa aaaaacccaa 7740
agggatgaaa gtttcatcca tcaaaggaaa caacagtcac cttggttccc atcccactca 7800
tatactgccg ccgtacatgt caatcagatg aacctgtgcg tatctcttaa cgacaattga 7860
cccacctttt taactgaagt gaaggggggt tctgctccgc gaccacttcc tggatctctc 7920
ccttcaccct ctgtgttctt tcgggtgcac catcgggtca aagccgcagc aacgccgtct 7980
ctgtgtgatc gcatgtgccc ttctgcacac gaccttcccc cgagagtgac cagctaccgg 8040
acaggcacca aggagggcta ccgagcacct cccggaccgg cggctgcagg atcgcgagcg 8100
cctccgctag ggagaccgca cgttgcgcct gtgcttcctg cggtggcgcc ttctgcaagg 8160
agacctcgac cctgctccct ctccggggct ggatctgact ccttgacggt gattccagac 8220
gcgagaccca aactgacggc ttctagaaga ggggcgagcc cggccgcaag tctttcacgt 8280
agctaagtca tcgttgcttc cggcttctta ccgttctccc ctttgtaaac ggttacctcc 8340
cgaaaaccca ggctctcctc caacagtggt tctcaagcga ggcgatcttc cccgggaggg 8400
gatatttggc aaagtctggg ggcatttttg gttcactggg gctgctactt gcatccactg 8460
ggtagaggcg ggggatgcag ctacacaacc tgcgaagcac gggacagcac cctccccaac 8520
ccagacagaa ttagccggcc caaaacctca gtagtgccca ggctgagaaa ccctgcctta 8580
aacaaacaac aaagaaaggc caagtcccat aagtgggtca ccgcgccgag actggggtcc 8640
acgggacacc ccagccacgc caagccggga agtccccgcc tcctggagct gaacccgccc 8700
ctctcccaga ggtggagctg cgggggggcgg gaacaggcac ggagaaaata aacaagacta 8760
aaaagtcctg agtagcgctg tgtggccgca aacctgaacc caccttttgc accacgcggg 8820
acccggcact cttcctgcca cccacccctg agagggctgc gcggccgacc ccagtactag 8880
aaaacactcg tcacctcact caagacgggt acgaaggcca acggacgcct tcctttagaa 8940
cgctcagcac acagagcaac ttctcacgcc tactctcaaa tggcgtactc caaactagca 9000
ctcccgacgt ccagctgtga acccagagcg gcggaaagcc cctgaaccca gcgcccgggc 9060
atgcgcagac gcgttgttgt ggtgggcgtg gctccctccg gacccggcgc cccgccctcc 9120
gccccgtgtc cgcatgcgcg actgagccgg gtggatggta ctgctgcatc cgggtgtctg 9180
gaggctgtgg ccgttttgtt ttcttggcta aaatcggggg agtgaggcgg gccggcgcgg 9240
c 9241

<210> 19
<211> 9236
<212> DNA
<213> Homo sapiens

<400> 19

```
ctagaaaaca ctttgtcatt ttagaggtgt tatccaatgt tcgcgcaggc actggagtca 60
gagaaaatgg agttgaatcc tttctctgcc actctttgag gagaatctca ccatttatta 120
tgcactgtag aatacaacaa taaaatacag ccatgtacca cataacaaca tcttggtaaa 180
caacagactg catatatgat ggtggtcatc cagtaagcta aggttaattt attattattc 240
ccttttttct ttctttttct tgagatgtag tcttactctg tcacccaggc tagagtgcaa 300
tggcaccatc ttggctcact gcaacctctg cctcctgggt tcaagcgaat ctcctgcctc 360
agcctccgaa gtagctggga attacaggca cccaccacat ctggctaatt ttttgtattt 420
ttagtaaaga tggggtttca ccatgttggc caggctgatc tcaaactcct gacctcaagt 480
gatctgcctg cctcggcctc ccaaagtgct gggaccatag gcctgagcca ctgtgcccgg 540
```

ccttgtttgc ttttttaaca gttaacagtg tgctcataga aactgctttg acatgactgc 600
aatcatgtgc ttcatagaaa cttaattaga ttataccact agagtcttca gatttttata 660
ctttttttt tgaaacggag tctcactctg tcaccaggct ggagtgcagt gccgcaatct 720
cagctcgccg caacctccgc ctcccaggtt caagtgattc tcctgcctca gcctcccgag 780
tagctgggat tacaagtgca cactaccacg cccagctaat ttttgcattt ttactagaca 840
gggtttcacc atgttggcta ggatagtttc accaggatct cttggcctca tgatcagcct 900
gcctcggcct cccaaagtgc tgggattaca ggtgtgagcc accgtgccca gcctatactt 960
ccctttttga ataccatttg gcgttttgaa gaattaacag ctttgtgaac gtggcagtgc 1020
ttgtgattca ggcttccact gagaccaagg ggagaacctg gttgcaggac aaacagacgg 1080
acagcgtgtg gcagtgttta aatgctcttc tgaaggctga tacgacagct ctctgtgcac 1140
tgattgcata cgcatcccaa gattatatta ttgttttcta ttgctatgtg tcacactttg 1200
ccaaacagga tgtggaaaat gaataagcgg ttttcttagg cacttcttaa cagacaattg 1260
gtcaaaatga actccattgc ttaagaaaca cataaacacc atttagtcac tgaatatagc 1320
tatatgtatg gttgctacta tggggaatct tgttttgcca attttctttg aaaattctgg 1380
cagaccaagg ttctttttgt ttacacaata cttgaaaaat aaaaatgaac aagccaacaa 1440
actaccaagt tttcacttac ataaatgtag ttacatacag aaaatgtgac tgtgaatttt 1500
ttctaggact tttaaactat aagcactatt tgcacgaaag agaaccaatc tatcaattac 1560
aaactcacat aattttacag attttttttt ccctacacag cacataaaac agaaggaatt 1620
tgaagccacc ctccaaacac aggggaagga ggctgtgtgt atatcctcat tgtctttcac 1680
attctaaggt ggttccactc agtgactgaa atccttaagt gttgtattag tcggcttggg 1740
ctaccataac agcagcttaa actgttgtta agccactcag acttaaacaa cagaaattta 1800
tttccttata gttctggagg ctggaagttc aaggtgccgg caaggctggt ttctggtgag 1860
acctctctcc ctgtcttgca gatggctgcc tcctccctgt gtcctcatag agcctgtcct 1920
ctgcttttac acttctggtg tcatcttcct ttttttttct tttttgagac agagtctcgc 1980
tctatcgccc aggctggagt gcagtggccc gatcgatctc ggctcactgc aacctctgcc 2040
tcccaggttc aagcaattct cctgcctcag cctcccgagt agctgggact acaggtgccc 2100
gccatcatgt ctggctaatt tttgtatttt tagtagagac agggtttcac catattggcc 2160
aggctggtct ccaactcctg accttgtcat ctgcctgcct cggcctccca aagtgctagg 2220
attacaggcg tgagccaccg cacccggcct ctttctcttc ttataaggac accagtccta 2280
ttagattagg gctccaccct catgacctca tttgacctta actattattt ctttaaagca 2340
cctatttcca aatatagtca ctttaggggt tagggcttca aaatatgaat ctgagggaga 2400
tcaattcagt aaatagcagt agtcattaac ggacaatata tacaaagata atttcgtgat 2460
tactgtcctt atgcataaat gtcctcagtg ttccactgcc tttatccaga tttactatca 2520
caaagacttt gctctgagaa aaatgtgatt tctttctttt ttttttttt ttgagacaga 2580
gtctcactct gtcacccagg ctggagtgca gtggtgcaat ctcggctcac tgcaatctcc 2640
gcctcccagg ttcacgccat tctcttgcct cagtctcccg agtagctggg cctacaggcg 2700
cccgccaccc tgcccagcta attttttgta tttttagtag agacggggtt tcaccatgtt 2760
agccaggatg gtctcaatct cctgacctcg tgatccacct gcctcagcct cccaaagtgc 2820
tgggattaca ggcatgagcc accgcgccca gcagattttt ttttttttt ttttttttg 2880
agatggagtc ttgctgtgtt gcccagcctg gagtgcagtg ttatgatttt ggctcactgc 2940
aacctctgtc taccatgttc aagcgattct cccacctctg cctcccgtgt agctgggatc 3000
acaggcacac gccaccacac ctagctactt tttgtatttt tagtagaaat ggggtttcac 3060
catgttggcc aggatggtcc cgaactcctg acctcaagtg atcctcctgc ctcggcctcc 3120
caaagtgctg ggattacagg tgtgagccac tgtgcctggc caaaaatgtg atttcttatt 3180
tcccacattg ccaattccat ttcaattaac tataatagct atgtctattg agcactcaag 3240
tgtattctag aaactgttcc tgattctggg gatatatcca tgaatcaact atagtccctg 3300
ttattaagta atctgtagtc tgactaaacc attagaaatt taaaaaatgg ctactttcaa 3360
agacatcttg gagttcagga gtcccacact gcgaaccata ttacctaata atccaacctg 3420
cttgtaattc acttatttaa ccaatattta ttgagtgcca actttgagcc taagatacag 3480
cagtaaacaa atggataaag tccctgtcct catgaaactt gtattctaat ggaagaaaca 3540
gaaaacaaac agatatagga tgtaatatca ggtagggata aatactttga attcaaacaa 3600
aagtatacgt agtcagggtt cgccaaagag acacagccaa tcggatacat agatatataa 3660
aagagggttt atgagttaga aagggctcac atgattacag aggctgagaa gtcccacagc 3720
agattgtctg caagctggag acccagggat actggtagca tggctcagtc caagtcccaa 3780
agcctcagaa tcaggaaagc tgatgatata attcttagcc caaaggcctt agaaccccag 3840
cggtgacgga aaggctgatg taggtcctgg agtcctgaga cccaacagcc tgggatcctg 3900
aaatccaagg gcaggaatgg aagcgtgtat tccagctcca agagagtaag accaatttgc 3960
ctttcttccg tttttgtttc aagccacctg cacattgagg gcggatggtt ccctcttagt 4020
ccattcagtc atatatcaat ctcttctgga aataccctca cagacacact aacaaataat 4080
gcctttccag ttctctaggt attctttaat ccagtcaagc tgacacctaa aattaaccat 4140
cacaaaagtt aaggagaaag aagacaactt gtaggggagg ctgctatgca agacagtgtg 4200
tgaaggaagg gctctctgaa gaggttaata tctgagcaga gacttgaatg aagtgaagaa 4260
gtgagccatg tgggtatggg gaatacaact tccaggtaga gaagacaagt gtggtgtgta 4320

tcagggtcag caaagaagcc atgtgacaga gaagggtggg ccagggagag acggataagt 4380
gatctaactc ctgaggaggt ggcctggcca ggagctagag catgaagatc tcgtaggact 4440
ttattctgca aggtgaaaag ccattgtatt agtctgttca caaacccgag actaggcaat 4500
ttacaaaaga aagagaggtt taatggactt acagttccac atggctgggg aggcctcaca 4560
atcatggcga aaggcaatga ggagcaagtc acgtcttacg tggatggcag gcaaagacaa 4620
agacagcttg tgcagagaaa ctccccctta tagagccatc agatcctgtt agacttattc 4680
actatcacaa gaacagcacg ggtaagacct gtccccatga ttcagttacc tcccactggg 4740
tccctcccac aacgcatggg aattcaggat gagatttggg tggggacaca accaaaccct 4800
atcattccac ccatggcccc tcccaaattt catgtcctca catttcaaaa ccaatcacac 4860
catcccaaca gtccctcaaa gtcttaaatg atttcagcat taactcaaaa gtccacagtc 4920
taatgtctca tctgagacaa ggcaagtcct ttccatttat gagcctataa aatccaaagc 4980
aagttagtta cttcctagat acaatggggg tacaggcatt gggtaaatac agccattcca 5040
aatgggataa attggtcaaa acaaagaggc tacaggccca tgagagtcca aaatccagtg 5100
gggcagtcaa atcttaaagc tccaaaatga tctcctttga ctccacatct cacatccagg 5160
tcacgcagat ggaaggggtg ggttcccatg gtcttgggca gctctgcccc tgtacctttg 5220
cagggtacag cctccctctc agctgctttc atgggctggc attgagtgtc tgcagctttt 5280
ccaggtacac ggtgcaagct gtcggtggat ctaccattct ggggtctgga ggacctcttc 5340
tcacagctcc actaggtggt gccccagtag ggactgtgtg tggggtctct gaccccacat 5400
ttcccttctg cactgccctg gcagaggatc tccatgaggg ccctgctcct gcagcaaact 5460
tctgactggg catccaggca tttccgcaca tcctctttaa tctaggcgaa ggtttccaaa 5520
ccccaattct tgacttctgt gcactcgcag tctcaacacc acatggaagc tgtcaaggct 5580
tggggcttgc actccccgaa gctacagccc aagctctacc ttgcctcccg tcagtcatgg 5640
ttgggagtgg ctgggatgca gggcaccaag tccctaggct gcacacagca tgaggacccc 5700
gggcctggcc aacaaaacca ttttttcctg atacctctgg acctgtgatg ggaggggttg 5760
ccataaagac ctctgacatg ccctggagac attttcccca ttgtcttggg aattagcatt 5820
tggctcctgt tactcatgca aatttctgca gccagcttga atttctcctc agaaaatggg 5880
aatttttctt ttctatcaca ttgtcaggct gcaaattttc cgaactttta tgctctgctt 5940
cccttataaa actgaatgtc tttaacagca cccaagtcac ctcttgaatg ctttgctgct 6000
tagaaatttc tcctgccaga tactctaaat catctctctg aagttcaaag ttctacaaat 6060
atctcgtgca ggggcaaaat gccgccagta tctttgctaa aacataacaa gagtcccctt 6120
tgctccagtt cccaacaagt tcctcatttc cgtctgagac cacctcagcc tatggacttt 6180
attgtccaca gtgctatcag cattttgggc aaagccattc aacaagtctc taggaagttc 6240
caaactttcc cacatttgcc tgtcttcttc tgagccctcc aaactgttcc aaaccctgcc 6300
tgttacccag ttccaaagtc acatacccat ttttgagtat ctacggcagc accccactct 6360
actggtacca atttagccac tgaagtagtt ggagaacaga agtaatagac tctggtttac 6420
attgtaaaag cttctctgtg gctgctgtgt gaagaaaata tatgagaatg aagccccaag 6480
atgaagcagg gacacagttg cagtggttag agtaagaaat gctgctggct ggcactgaag 6540
tgatagcctg gaggtttgtg tgtgcacatg catgtgtatg tgttttacga tagtaggccc 6600
aacagatact gtaatccaca cttgtttttt ttttttgaga cagagtctca cctgttgcct 6660
agactagaat gcagtggcac aatcttggct cactacaacc tccacctccc aggttcaaac 6720
aatccttgtg cttcagcctc ccgagtagtt gggattacag gtgtgtgcca ccgtgcccag 6780
ctatattttt tgtattttta gcagagatgg gattttgcca cattggccag gctggtcttg 6840
aactcctggc ctcaagcaat cctcccacct tagcctccca aagtgctgag ccaccacacc 6900
tggccgcaac tgatttttaa tcatgaaatg acacatacat ttaaaaaacc caatacctat 6960
aatattcctg gctagtactc ttcacatcta tatcatcaaa aacaaagaaa gtatgtgaaa 7020
ctgacacagc caaggggaga ctaaggagac ataacaatta actgtaatgt ggtattctgg 7080
aggggatcct ggaacagaaa aagacattag gcaaaaaact aaagaaatct gaataaaatg 7140
tggatgtcag ttaataataa tgtatcatat tagtccagta attgtaacaa atataccaca 7200
ataatgaaag ccattaatta tagggaaaat ggaggggtta atatgggtgg ctggcttttg 7260
ctatttctag cagctccatt ttatctgcaa aagacaaaca ttcattaagt cccaaaaagg 7320
taaagaatga caaattaagc atgtatctta ttagtaagag taatataaag atgctcactc 7380
ctatttataa atatttgaca atcatgttaa ggccacaaaa gagaaaaaag ggtaggggca 7440
aaaaacgcaa agagaaagga gttagtatct tttctcccgc actcattagc tattaaaaga 7500
ggatgtttgt ttaaagctgc tcagagctgg taaactaatg ttaagtcact aacgggaatt 7560
taaaaggttt cattaagaac tgcctgcact agattcctcc accctgagac attaaacaat 7620
cacgataaac ctcctgagtg gtaagaactt gtccatttaa aaacaggcta tagattgtat 7680
catgcagttt tatctactaa tcggctaata tcccgccaaa aacaaaaaac cccaaaggga 7740
tgaaagtttc atccatcaaa ggaaacaaca gtcaccttgg ttcccatctc actcatatac 7800
tgccgccgta catgtcaatc agatgaacct gtgcgtatct cttaatgaca attgacccac 7860
atttttgact gaagtgaaag ggggttctgc tccgcgacca cttcctggat ctccccctcc 7920
accctctgtg ttctttcggg tgcaccatcg ggtcaaagcc gcagcaacgc cgtctctgtg 7980
tgatcgcatg tgcccttctg cacacgacct tcccccgaga gtgaccagct accggacagg 8040
caccaaggag ggctaccgag cacctcccgg accggcggct gcaggatcgc gagcgcctcc 8100

```
gctagggaga ctgcacgttg cgcctgtgct tcctgcggtg gcgccttctg caaggagacc 8160
tcgaccctgc tccctctccg gggctggatc tgactccttg acggtgattc cagacgcgag 8220
acccaaactg acggcttcta gaagaggggc gagcccggcc gcaagtcttt cacgtagcta 8280
agtcatcgtt gcttccggct tcttaccgtt ctccccttg taaacggtta cctcccgaaa 8340
acccaggctc tcctccaaca gtggttctca agcgaggcga tcttccccgg gaggggatat 8400
ttggcaaagt ctgggggcat ttttggttca ctggggctgc tacttgcatc cactgggtag 8460
aggcggggga tgcagctaca caacctgcga agcacgggac agcaccctcc ccaacccaga 8520
cagaattagc cggcccaaaa cctcagtagt gcccaggctg agaaaccctg ccttaaacaa 8580
acaacaaaga aaagccaagt cccataagtg ggtcaccgcg ccgagactgg ggtccacggg 8640
acaccccagc cacgccaagc cgggaagtcc ccgcctcctg gagctgaacc cgcccctctc 8700
ccagaggtgg agctgcgggg ggcgggaaca ggcacggaga aaataaacaa gactaaaaag 8760
tcctgagtag cgctgtgtgg ccgcaaacct gaacccacct tttgcaccac gcgggacccg 8820
gcacgcttcc tgccacccac ccctgagagg gctgcgcggc cgaccccagt actagaaaac 8880
actcgtcacc tcaatcaaga cgggtacgaa ggccaacgga cgccttcctt tagaacgctc 8940
agcacacaga gcaacttctc acgcctactc tcaaatggcg tactccaaac tagcactccc 9000
gacgtccagc tgtgaaccca gagcggcgga aagcccctga acccagcgcc cgggcatgcg 9060
cagacgcgtt gttgtggtgg gcgtggctcc ctccggaccc ggcgccccgc cctccgcccc 9120
gtgtccgcat gcgcgactga gccgcggggg tggtactgct gcatccgggt gtctgaagat 9180
ccgatgaaat aacatatgca aaatgattgg gtccgtgatt ggcattccag aaatgg     9236
```

<210> 20
<211> 9238
<212> DNA
<213> Homo sapiens

<400> 20

aaacgctcat gacagcaaag tctccaatgt tcgcgcaggc actggagtca gagaaaatgg 60
agttgaatcc tttctctgcc actctttgag gagaatctca ccatttatta tgcactgtag 120
aatacaacaa taaaatacag ccatgtacca cataacaaca tcttggtaaa caacagactg 180
catatatgat ggtggtcatc cagtaagcta aggttaattt attattattc cttgtttttt 240
ttttttttt ttttttttg agatgtagtc ttactctgtc acccaggcta gagtgcaatg 300
gcaccatctt ggctcactgc aacctctacc tcctgggttc aagcaaatct cctgcctcag 360
cctccaaagt agctgggatt acaggcaccc accacatctg gctaatttt tgtattttta 420
gtaaagatgg ggtttcacca tgttggccag gctgatctca aactcctgac ctcaagtgat 480
ctgcccgcct cggcctccca aagtgctgga accacaggcc tgagccactg tgcccagcct 540
tggttgcttt tttaacagat aacagtgtgc tcatagaaac tgctttgaca tgactgcaat 600
catgtgcttc atagaaactt aattagatta taccactaga gtcttcagat ttttatactt 660
ttttttttg aaacggagtc tcactctgtc accaggctgg agtgcagtgc cgcaatctcg 720
gctcactgca acctccgcct cccaggttca agcaattctc ctgcctcagc ctcccgagta 780
gctgggatta caagtgcgca ctaccacacc cagctaattt ttgcattttt acttgacagg 840
gtttcaccat gttggctagg atagtttcac caggatctct tggcctcatg atcagcctgc 900
ctcggcctcc caaagtgctg ggattacagg tgtgagccac cgtgcccagc ctatacttcc 960
ctttttgaat accatttggc gttttgaaga attaacagct ttgtgaacgt ggcagtgctt 1020
gtgattcagg cttccattga gaccaagggg agaacctggt tgcaggacaa acagacggac 1080
agcgtgtggc agtgtttaaa tgctcttctg aaggctgata cgacagctct ctgtgcactg 1140
attgcatacg catcccaaga ttatattatt gttttctact gctatgtgtc acactttgcc 1200
aaacaggatg tggaaaatga ataagcggtt ttcttaggca cttcttaaca gacaattggt 1260
caaaatgaac tccattgctt aagaaacaca taaacaccat ttagtcactg aacatagcta 1320
tatgtatggt tgttactatg ggaaatcttg ttttgccaat tttctttgaa aattctggca 1380
gaccaaggtt ctttttgttt acataatact tgaaaaataa aaatgaacaa gctaacaaac 1440
taccaagttt tcacttacat aaatgtagtt gcatacagaa aatgtgactg tgaattaatt 1500
tttctaggac ttttaaacta taagcactat ttgcacaaaa gagaaccaat ctatcaatta 1560
caaactcaca taattttaca gatttttttt ttcctacaca gcacataaaa cagaaggaat 1620
ttgaagccac cctccaaaca caggggaagg aggctgtgtg tatatcctca ttgtctttca 1680
cattctaagg tggttccact cagtgactga aatccttaag cgttgtatta gtctgcttgg 1740
gctaccataa cagcagctta aactgttgtt tagccactca gacttaaaca acagaaattt 1800
atttccttat agttctggag gctggaagtt caaggtgccg gcaaggttgg tttctggtga 1860
gacctctctc cctgtcttgc agatggctgc ctcctccctg tgtcctcata gagcctgtct 1920
tctgctttta cacttctggt gtcatcttcc ttttttttt ttttttttt ttttgagaca 1980
gagtctcgct ctatcgccca ggctggagtg cagtggcccg atcgatctcg gctcactgca 2040
acctctgcct cccaggttca agcaattctc ctgcctcagc ctcccaagta gctgggacta 2100
caggtgcccg ccatcatgcc tggctaattt ttgtattttt agtagagaca gggtttcacc 2160

atattggcca ggctggtctc gaactcctga ccttgtcatc tgcctgcctc ggcctcccaa 2220
agtgctagga ttacaggcgt gagccaccgc acccggcctc ttcctcttct tataaggaca 2280
ccagtcctat tagattaggg ctccaccctc ataacctcat ttgaccttaa ctattatttc 2340
tttaaagcac ctatttccaa atatagtcac tttaggggtt agggcttcaa aagatgaatc 2400
tgagggagct caattcagta aatagcagta gtcattaatg gacaatgtat acaaagataa 2460
tttcgtgatt actgtcctta tgcataaacg tcctcagtgt tccactgcgt ttatccagat 2520
ttagtatcac aaagactttg ctctgagaaa aatgtgattt ttttttttt ttttttttg 2580
agacggagtc tcgctctgtt acccaggctg gagtgcagtg gcgcgatctc ggctcactgc 2640
aagctccgcc tcccgggttc acgccattct cctgcctcag cctccggagt agctgggact 2700
acaggcgccc gccactacgc ccggctaact tttttgtatt tttagtagag acggggtttc 2760
accatgttag ccaggatggt ctcaatctcc tgacctcgtg atccacctgc ctcagcctcc 2820
caaagtgctg ggattacagg catgagccac cgcgcccagc agatttttt ttttttttt 2880
gagatggagt cttgctctgt tgcccaacct ggagtgcagt gttatgattt tggctcactg 2940
caacctctac catgttcaag cgattctccc acctctgcct cccgtgtagc tgggatcaca 3000
ggcacacgcc accacaccta gctacttttt gtatttttag tagaaatggg gtttcaccat 3060
gttggccagg atggtcccga actcctgacc tcaagtgatc ctcctgcctc ggcctccaaa 3120
gtgctgggat tacaggtgtg agccactgtg cctggccaaa aatgtgattt cttatttccc 3180
acattgccaa ttccatttca attaactata atagctatgt ctattgagca ctcaagcgta 3240
ttctagaaac tgttcctgat tctggggata tatccatgaa tcaactatag tccctgttat 3300
taagtaatcc gtagtctgac taaaccatta gaaattaaaa aaaaatggct actttcaaag 3360
acatcttgga gttcaggagt cccacactgc gaaccatatt acctaataat ccaacctgct 3420
tgtaattcac ttatttaacc aatatttatt gagtgccaac tttgagccta agatacagca 3480
gtaaacaaat gcataaagtc cctgtcccca tgaaacttgt attctaatgg aaaaaacaga 3540
aaacaaacag atataggatg taatatcagg tagggataaa tactttgaat tcaaacaaaa 3600
gtatacgtag tcagggttcc ccagagagac acagccaatc gatacataga tatataagag 3660
agggtttatg agttagaaag ggctcacatg attacagagg ctgagaagtc ccacaacaga 3720
ttgtctgcaa gctggagacc cagggatact ggtagcatgg ctcagtccaa gtcccaaagt 3780
ctcagaatca ggaaagctga tgatataatt cttagcccaa aggccttaga accccagcgg 3840
tgacggaaag gctgatgtag gtcctggcgt cctgagaccc aacagcctgg gatcctgaaa 3900
tccaagggca ggaatggaag cgtgtattcc agctccaaga gagtaagacc aatttgcctt 3960
tcttccgttt ttgttccaag ccaactgcac gttgagggcg gatggttccc tcttagccca 4020
ttcagtcata tatcaatctc ttctggaaat accctcacag acacactaac aaataatgcc 4080
tttccagttc tctaggtatt ctttaatcca gtcaagctga cacctaaaat taaccatcac 4140
aaaagttaag gagaaagaag acaacttgta ggagaggctg ctatgcaaga cagtgtgtga 4200
aggaagggct ctctgaagag gttaatatct gagcagagac ttgaatgaag tgaagaagtg 4260
agccatgtgg gtatggggaa tacaacttcc aggtagagaa gacaagtgtg gtgtgtatca 4320
cggtcagcaa agaagccatg tgacagagaa gggtgggcca gggagagacg gataagtgat 4380
ctaactcctg aggaggtggc ctggccagga gctagagcat gaggatctcg taggatttta 4440
ttctgcaagg tgaaaagcca ttgtattagt ctgttcacaa accccagact aggcaattta 4500
caaaagaaag agaggtttaa tggacttaca gttccacatg gctggggagg cctcacaatc 4560
atggcgaaag gcaatgagga gcaagtcacg tcttacgtgg atggcaggca aagacaaaga 4620
cagcttgtgc agagaaactc ccccttatag agccatcaga tcctgttaga cttatcacta 4680
tcacgagaac agcacgggta agacctgtcc ccatgattca gttacctccc actgggtccc 4740
tcccacaacg catgggaatt caggatgaga tttgggtggg gacacaacca aaccctatca 4800
ttccacccat ggcccctccc aaatttcatg tcctcacatt tcaaaaccaa tcacaccatc 4860
ccaacagtcc ctcaaagtct taaatgattt cagcattaac tcaaaagtcc acagtctaat 4920
gtctcatctg agacaaggca agtcctttcc gtctatgagc ctataaaatc caaagcaagt 4980
taattacttc ctagatacaa tgggggtaca ggcattgggt aaatacagcc attccaaatg 5040
ggataaattg gtcaaaacaa agaggctaca ggcccatgag agtccaaaat ccagtggggc 5100
agtcaaatct taaagctcca aaatgatctc ctcttgactc cacatctcac atccaggtca 5160
tgcagatgga aggggtgggt tcccatggtc ttgggcagct ctgcccctgt acctttgcag 5220
ggtacagcct ccctctcagc tgctttcatg ggctggcatt gagtgtctgc aacttttcca 5280
ggtacacggt gcaagctgtc ggtggatcta ccattctggg gtctggagga cctcttctca 5340
cagctccact aggtggtgcc ccagtaggga ctgtgtgtgg ggtctctgac cccacatttc 5400
ccttctgcac tgccctggca gaggatctcc atgagggccc tgcccctgca gcaaacttct 5460
gcctgggcat ccaggcattt ccgcacatcc tctttaatct aggcgaaggt ttccaaaccc 5520
cagttcttga cttctgtgca ctcgcagtct caacaccaca tggaagctgt caaggcttgg 5580
ggcttgcact ccccgaagct acagcccaag ctctaccttg cctcctgtca gtcatggttg 5640
ggagtggctg ggatgcaggg caccaagtcc ctaggctgca cacagcatga ggaccccggg 5700
cctggccaac aaaaccattt tttcctgata tctctggacc tgtgatggga ggggttgcca 5760
taaagacctc tgacatgccc tggagacatt ttccccattg tcttgggaat tagcatttgg 5820
ctcctgttac tcatgcaaat ttctgcagcc agcttgaatt tctcctcaga aaatgggaat 5880
ttttcttttc tatcacattg tcaggctgca aattttccga acttttatgc tctgcttccc 5940

ttataaaact gaatgtcttt aacagcaccc aagtcacctc ttgaatgctt tgctgcttag 6000
aaatttctcc tgccagatac tctaaatcat ctctctgaag ttcaaagttc tacaaatatc 6060
tcgtgcaggg gcaaaatgcc gccagtatct ttgctaaaac ataacaagag tccccctttgc 6120
tccagttccc aacaagttcc tcatttccgt ctgagaccac ctcagcctat ggactttatt 6180
gtccacagtg ctatcagcat tttgggcaaa gccattcaac aagtctctag gaagttccaa 6240
actttcccac atttgcctgt cttcttctga gccctccaaa ctgttccaaa ccctgcctgt 6300
tacccagttc caaagtcaca tacccatttt tgagtatcta cggcagcacc ccactctact 6360
ggtaccaatt tagccactga agtagttgga gaacagaagt aatagactct ggtttacatt 6420
gtaaaagctt ctctgtggct gctgtgtgaa gaaaatatat gagaatgaag ccccaagatg 6480
aagcagggac acagttgcag tggttagagt aagaaatgct gctggctggc actgaagtga 6540
tagcctggag gtttgtgtgt gcacatgcat gtgtatgtgt tttacgatag taggcccaac 6600
agatactgta atccacactt gttttttttt ttttttgag acagagtctc acctgttgcc 6660
tagactagaa tgcagtggca caatcttggc tcactacaac ctccacctcc caggttcaaa 6720
caatccttgt gcttcagcct cccgagtagt tgggattaca ggtgtgtgcc accgtgccca 6780
gctatatttt ttgtattttt agcagagatg ggattttgcc acattggcca ggctggtctt 6840
gaactcctgg cctcaagcaa tcctcccacc ttagcctccc aaagtgctga gccaccacac 6900
ctggccgcaa ctgattttta atcatgaaat gacacataca tttaaaaaac ccaataccta 6960
taatattcct ggctagtact cttcacatct atatcatcaa aaacaaagaa agtatgtgaa 7020
actgacacag ccaaggggag actaaggaga cataacaatt aactgtaatg tggtattctg 7080
gaggggatcc tggaacagaa aaagacatta ggcaaaaaac taaagaaatc tgaataaaat 7140
gtggatgtca gttaataata atgtatcata ttagtccagt aattgtaaca aatataccca 7200
ataatgaaag ccattaatta tagggaaaat ggaggggtta atatgggtgg ctggcttttg 7260
ctatttctag cagctccatt ttatctacaa aagacaaaca ttcattaagt cccaaaaagg 7320
taaagaatga caaattaagc atgtatctta ttagtaagag taatataaag atgctcactc 7380
atatttataa atatttgaca atgatgttaa ggccagaaaa gagaaaaaag ggtaggggca 7440
aaaaacgcaa agagaaagga gttagtatct tttctcccgc actcattagc tattaaaaga 7500
ggatgtttgt ttaaagctgc tcagagctgg taaactaatg ttaagtcact aacgggaatt 7560
taaaaggttt cattaagaac tgcctgcact agattcctcc accctgagac attaaacaat 7620
cacgataaac ctcctgagtg gtaagaacgt gtccatttaa aaacaggcta tagattgtca 7680
tgcagtttta tctactaatc ggctaatgca ccgccaaaaa caaacaaaaa aacccaaagg 7740
gatgaaagtt tcatccatca aaggaaacaa cagtcacctt ggttcccatc ccactcatat 7800
actgccgccg tacatgtcaa tcagatgaac ctgtgcgtat ctcttaatga caattgaccc 7860
acctttttaa ctgaagtgaa gggggggttct gctccgcgac cacttcctgg atctctccct 7920
tcaccctctg tgttctttcg ggtgcaccat cgggtcaaag ccgcagcaac gccgtctctg 7980
tgtgatcgca tgtgcccttc tgcacacgac cttcccccga gagtgaccag ctaccggaca 8040
ggcaccaagg agggctaccg agcacctccc ggaccggcgg ctgcaggatc gcgagcgcct 8100
ccgctaggga gaccgcacgt tgcgcctgtg cttcctgcgg tggcgccttc tgcaaggaga 8160
cctcgaccct gctccctctc cggggctgga tctgactcct tgacggtgat tccagacgcg 8220
agacccaaac tgacggcttc tagaagaggg gcgagcccgg ccgcaagtct ttcacgtagc 8280
taagtcatcg ttgcttccgg cttcttaccg ttctcccctt tgtaaacggt tacctcccga 8340
aaacccaggc tctcctccaa cagtggttct caagcgaggc gatcttcccc gggaggggat 8400
atttggcaaa gtctgggggc atttttggtt cactggggct gctacttgca tccactgggt 8460
agaggcgggg gatgcagcta cacaacctgc gaagcacggg acagcaccct ccccaaccca 8520
gacagaatta gccggcccaa aacctcagta gtgcccaggc tgagaaaccc tgccttaaac 8580
aaacaacaaa gaaaggccaa gtcccataag tgggtcaccg cgccgagact ggggtccacg 8640
ggacacccca gccacgccaa gccgggaagt ccccgcctcc tggagctgaa cccgcccctc 8700
tcccagaggt ggagctgcgg ggggcgggaa caggcacgga gaaaataaac aagactaaaa 8760
agtcctgagt agcgctgtgt ggccgcaaac ctgaacccac cttttgcacc acgcgggacc 8820
cggcactctt cctgccaccc acccctgaga gggctgcgcg gccgacccca gtactagaaa 8880
acactcgtca cctcactcaa gacgggtacg aaggccaacg gacgccttcc tttagaacgc 8940
tcagcacaca gagcaacttc tcacgcctac tctcaaatgg cgtactccaa actagcactc 9000
ccgacgtcca gctgtgaacc cagagcggcg gaaagcccct gaacccagcg cccgggcatg 9060
cgcagacgcg ttgttgtggt gggcgtggct ccctccggac ccggcgcccc gccctccgcc 9120
ccgtgtccgc atgcgcgact gagccgggtg gatggtactg ctgcatccgg gtgtctggag 9180
gctgtggccg ttttgttttc ttggctaaaa tcgggggagt gaggcgggcc ggcgcggc 9238

<210> 21
<211> 1249
<212> DNA

<213> Homo sapiens

<400> 21

```
ctgatctaca taggaattgt tttcaagaca tttctgcatt cctctagtga cagggtgctc 60
actacctcat gagtatttca gtggacaact gtaatggtca ataaagtatc cactttccac 120
ctccctgcag ctcctggccc tggctttatt ctctggggct ccacacattc agtttacact 180
cagtggccag tggctgggac cattgtagaa aataaggaaa ctccaattcc ttccttcttt 240
tcttcctctt tcatctcttc ctccctctct acatccctct ctctcttcct tccttcctcg 300
acacttacca tgtaccagac cttctgccag gcacatggat gggagcacag gggaagttgg 360
ctgcagggtt agaactaagt cccaagcccc ctaaagctca tgccagggga ctggactgtc 420
cagtactgag ggatggggat gctgaggctg gtggccttcc tcaaatgcac tgtagtgccc 480
caggcagagt cctgggctgc cctgtgagga ggtgaccaga ggtagagcaa cttcacccta 540
aggctggatc aggatcccct ccaggttttt actagagcca aacccacatc tcctttctct 600
tctgccaccc ccccttaaaa tgcttagaaa cacatagatt taaatacaaa ttcaaatgta 660
agtaatttca actgtgtaac tatgaggagt cagttctacg tgggtcctat ctgtatcctc 720
cccagggctc agctccattc tttgctttca ttcattctca ttcaatacat tgttgttaag 780
agctcactgg gtgccctctc tgtcatgtag taaggtttta aaaagaaagc ctcttctgag 840
cttcagtttc cttattcata aaataggagt attgatccgt tccttgcttt tcttacaagg 900
atatgctgaa gatgactgaa gtacagagta aagaaggatt atgtttgggt gtcaaaggaa 960
tagaatgccc tctttcaaac tgagcacagc aggaacctgt aacaggaaca cagcaacttg 1020
ttgaatgaat gacaatattg gaaaacatac atttcctccc ctccccatca tagtccctct 1080
gcttccgtgt taactccata gagaggccag cacaaccagc cttgcagcct gagataaggc 1140
ctttggcggg tgtctccccc atcgctccct caagccctca agtaggtgtt ggagagaggg 1200
gtgatgcctg gtgctggtgg aacccctgca cagagacgga cacaggatg              1249
```

<210> 22
<211> 699
<212> DNA
<213> Homo sapiens

<400> 22

```
ctagagaggg aagtttttga aaattaaaca ctgtctaatt ttctgcaaag tttttattca 60
tgaattaaga gtatttccct tagtccatta ttcccaaggc aaatatggaa gtttgatcat 120
atgctaatca tactaaagct ggattctctt taagagattg agaaattaaa aggcaaaagc 180
tgatatatca tgtttagtta tactgtgagt cttataagaa gctgggaggc aaccccatta 240
actcaccaga atacagaact cagtctcaca acttaaatat aattcctctc aaacctttc 300
ctcaaagtta aattctgaaa ataatcttgt gattaagaga agaaggctgt ccaccaatgg 360
acttatctgt tatttcttcc ttattgtgag cttaatggca tgacaaagca gaggcaaaga 420
ggcatacatc aattcttcaa agtaggaagt caaaaaggtc agagcttcca cagcatggca 480
acagctttgc agatgcccac atcgtgatag ttgaaatagc aaagcccagc aaaggttaaa 540
gctgaaaatg ccaaaagccc tgccttggca gctttctgcg aggcatcccc atgaacatag 600
tcagtaacaa cttgtccaag gccccagtga ccatgaagag tgagggctgc agccagggaa 660
tagtccgtcg cagagcaagg attcaaataa gcagccgga                          699
```

<210> 23
<211> 699
<212> DNA
<213> Homo sapiens

<400> 23

```
ctagagaggg aagtttttga aaattaaaca ctgtctaatt ttctgcaaag tttttattca 60
tgaattaaga gtatttccct ttgtccatta ttcccaaggc aaatatggaa atttgatcat 120
gtactaatca tactaaagct ggattctctt taagagattg agaaattaaa aggcaaaagc 180
tgatatgtca tgtttagtta tattgtgagt cttataagaa gctgggaggc aaccccatta 240
actcaccaga atacagaact cagtctcaca acttaaatat aattcctctc aaacctttc 300
ctcaaagtta aattctgaaa ataatcttgt gattaagaga agaaggctgt ccaccaatgg 360
acttatctgt tatttcttcc ttattgtgag cttaatggca tgacaaagca gaggcaaaga 420
ggcatacatc aattcttcaa agtaggaagt caaaaaggtc agagcttcca cagcatggca 480
acagctttgc agatgcccac atcgtgatag ttgaaatagc aaagcccagc aaaggttaaa 540
gctgaaaatg ccaaaagccc tgccttggca gctttctgcg aggcatcccc atgaacatag 600
tcagtaacaa cttgtccaag gccccagtga ccatgaagag tgagggctgc agccagggaa 660
tagtccgtcg cagagcaagg attcaaataa gcagccgga                        699
```

<210> 24
<211> 699
<212> DNA
<213> Homo sapiens

<400> 24

```
ctagagaggg aagtttttga aaattaaaca ctgtctaatt ttctgcaaag tttttattca 60
tgaattaaga gtatttccct ttgtccatta ttcccaaggc aaatatggaa atttgatcat 120
gtactaatca taataaagct ggattctctt taagagattg agaaattaaa aggcaaaagc 180
tgatatatca tgtttagtta tattgtgagt cttataagaa gctgggaggc aaccccatta 240
actcaccaga atacagaact cagtctcaca acttagatat aattcctctc aaacctttc 300
ctcaaagatt aaattctgaa aataatcttg tgattaagag aagaaggctg tccaccaatg 360
ggcttatctg ttatttcttc cttattgtga gcttaatggc atgacaaagc agaggcaaag 420
aggcatacat caattcttca aagtaggaag tcaaaaaggt cagagcttcc acagcatggc 480
aacagctttg cagatgccca catcgtgata gttgaaatag caaagcccag caaaggttaa 540
agctgaaaat gccaaaagcc ctgccttggc agctttctgc gaggcatccc catgaacata 600
atcagtaaca acttgttcaa ggccccagtg accatgaaga gtgagggctg cagccgggaa 660
tagtccgtcg cagagcaagg attcaaataa gcagccgga                        699
```

<210> 25
<211> 994
<212> DNA
<213> Homo sapiens

<400> 25

```
ttgtatctct ttttacagct acctcccatt tcccttctat ttcaagctag taacacagtt 60
ttcttttaaa ttcatttatt taaatgtaaa aataagtcta tttggagaaa aaaaattttt 120
aatagcatct ctggaatgcc agtatggcta aattcatgaa tgttgtcctc aaatgctgaa 180
atctgggaag catctggcca agctttgtgg acaggcctgc ctagtttgaa tcccaagagc 240
cactcattcc gagccacaaa acattggaat tcttggttca cttccctaac ctgaacttgt 300
cctctgtgaa atagggacat taatagctca ctcacagggc tgctgtgagg acatgtgttg 360
agctgagggt ctcgccaggg gagaccctgt gcagggagac tgttatcatg gtgatggatt 420
tctgcttcat tcatttcttt ttccagacag catcatatag aatgagttgt ggggtggcag 480
tcagcaggtt tgggtttatc ctctattctg ccacttatta cttaaaaaaa aaaacccaac 540
ttatatagta taagctatat ccagaaaagt gcaaatatca tacaagtacc atttgatgaa 600
tcttctgata tccccacata accaacaccc agaacctctt cttgtctcat tccaggataa 660
ccactaacct gacttctaac agcatcagtc agttttgtct gtttttgtac attatatatg 720
tgatggtttg aatgtgtccc ccaaatttca tgtgctagaa acttaatcct tcaattcata 780
tgttgatgct atttggagga agggcctttg ggaagtaatt aggattagat aaggtcatgg 840
ggtgaggtat gatggcactg gtgacttata agaagagaaa gagaaatctg agctggcatg 900
ctcttgccct ctcaccgtgt gatgacttct ccatgtcatg atgcagcaag aaggccctca 960
ccagatggtg gcaccatgct tttggacttc ccag                              994
```

<210> 26
<211> 1000
<212> DNA
<213> Homo sapiens

<400> 26

```
ttgtatctct ttttacagct acctcccatt tcccttctat ttcaagctag taacacagtt 60
ttcttttaaa ttcatttatt taaatgtaaa aataagtcta tttggagaaa aaaaattttt 120
aatagcatct ctggaaggcc agtatggcta aattcatgaa tgttgtcctc aaatgctgaa 180
atctgggaag catctggcca agctttgtgg acaggcctgc ctagtttgaa tcccaagagc 240
cactcattcc gagccacaaa acattggaat tcttggttca cttccctaac ctgaacttgc 300
cctctgtgaa atagggacat taatagctca ctcacagggc tgctgtgagg acatgtgttg 360
agctgagggt ctggtcaggg gagaccctgt gcagggagac tgttatcatg gtgatggatt 420
tctgcttcat tcatttcttt ttccagacag catcatatag aatgagttgt ggggtggcag 480
tcagcaggtt tgggtttatc ctctattctg ccacttatta cttaaaaaaa ccccaaaaaa 540
cccaacttat atagtataag ctatatccag aaaagtgcaa atatcataca agtaccattt 600
gatgaatctt ctgatatccc cacataacca acacccagaa cctcttcttg tctcattcca 660
ggataaccac taacctgact tctaacagca tcagtcagtt ttgtctgttt ttgtacatta 720
tatatgtgat ggtttgaatg tgtcccccaa atttcatgtg ctggaaactt aatccttcaa 780
ttcatatgtt gatgctattt ggaggaaggg cctttgggaa gtaattagga ttagataagg 840
tcatggggtg aggtatgatg gcactggtga cttataagaa gagaaagaga aatctgagct 900
ggcatgctct tgccctctca ctgtgtgatg acttctccat gtcatgatgc agcaagaagg 960
ccctcaccag atggtggcac catgcttttg gacttcccag                      1000
```

<210> 27
<211> 999
<212> DNA
<213> Homo sapiens

<400> 27

ttgtatctct ttttacagct acctcccatt tcccttctat ttcaagctag taactcagtt 60
ttcttttaaa ttcaattatt taaatgtaaa aataagtcta tttggagaaa aaaaatttta 120
atagcatctc tggaatgcca gtatggctaa attcatgaat gttgtcctca aatgctgaaa 180
tctgggaagc atctggccaa gctttgtgga caggcctgcc tagtttgaat cccaagagcc 240
acccagtcca agccacaaaa cattggaatt cttggttcac ttccctaacc tgaacttgcc 300
ctctgtgaaa tagggacact aatagctcac tcacagggct gctgtgagga catgtgttga 360
gctgagggtc tcgccagggg agaccctgtg cagggagact gttatcatgg tgatggattt 420
ctgcttcatt catttctttt tccagacagc atcatataga atgagttgtg gggtggcagt 480
cagcaggttt gggtttatcc tctattctgc cacttattac ttaaaaaaac cccaaaaaac 540
ccaacttata tagtataagc tatatccaga aaagtgcaaa tatcatacaa gtaccatttg 600
atgaatcttc tgatatcccc acataaccaa cacccagaac ctcttcttgt ctcattccag 660
gataaccact aacctgactt ctaacagcat cagtcagttt tgtctgtttt tgtacattat 720
atatgtgatg gtttgaatgt gtcccccaaa tttcatgtgc tggaaactta atccttcaat 780
tcatatgttg atggtttttg gaggaagggc ctttgggaag taattaggat tagataaggt 840
catggggtga ggtatgatgg cactggtgac ttataagaag agaaagagaa atctgagctg 900
gcatgctctt gccctctcac tgtgtgatga cttctccatg tcatgatgca gcaagaaggc 960
cctcaccaga tggtggcacc atgcttttgg acttcccag 999

<210> 28
<211> 25
<212> DNA
<213> Artificial Sequence

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

<400> 28
ggctaaatat tttgatgacc aagtt 25

<210> 29
<211> 19
<212> DNA
<213> Artificial Sequence

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

<400> 29
gcagccaact tcccctgtg 19

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

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

<400> 30
gctctacctt ggtcacctcc 20

<210> 31
<211> 24
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial primer

<400> 31
aggtcacatc catttatccc actg 24

<210> 32
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial sequence: artificial primer

<400> 32
agaagatggg ggaatctttt tcct 24

<210> 33
<211> 25
<212> DNA
<213> Artificial Sequence

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

<400> 33
ttgtgactgg gctagaaaga aggtg 25

<210> 34
<211> 23
<212> DNA
<213> Artificial Sequence

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

<400> 34
tgttgcctgc atttgtacgt gag 23

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

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

<400> 35
ttctgtctgg gttggggagg g 21

<210> 36
<211> 21
<212> DNA
<213> Artificial Sequence

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

<400> 36

ggaggggtta atatgggtgg c 21

<210> 37
<211> 22
<212> DNA
<213> Artificial Sequence

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

<400> 37
tttgtcctgg ttgcctgtgg tc 22

<210> 38
<211> 23
<212> DNA
<213> Artificial Sequence

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

<400> 38
caaatcctgt tgactggtct cgg 23

<210> 39
<211> 22
<212> DNA
<213> Artificial Sequence

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

<400> 39
aacggctcca tcacccctaa ag 22

<210> 40
<211> 24
<212> DNA
<213> Artificial Sequence

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

<400> 40
cccactccta gataccaacc caag 24

<210> 41
<211> 24
<212> DNA
<213> Artificial Sequence

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

<400> 41
ctttatgcac tgcctcgttg aatc 24

<210> 42

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

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

<400> 42
ttgactggtg tggttgctgt tg 22

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

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

<400> 43
gcagaaaggg gagttgatgc tg 22

<210> 44
<211> 23
<212> DNA
<213> Artificial Sequence

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

<400> 44
ctgacaaagt tgagagccca ctg 23

<210> 45
<211> 24
<212> DNA
<213> Artificial Sequence

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

<400> 45
ttaagcctac atccacatgc tgag 24

<210> 46
<211> 22
<212> DNA <213> Artificial Sequence

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

<400> 46
ccttggtctg ccagaatttt ca 22

<210> 47
<211> 24
<212> DNA
<213> Artificial Sequence

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

<400> 47
gtttggcatc ataggagatt tggc 24

<210> 48
<211> 23
<212> DNA
<213> Artificial Sequence

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

<400> 48
cctgtcccca tgattcagtt acc 23

<210> 49
<211> 19
<212> DNA
<213> Artificial Sequence

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

<400> 49
acgtacaaat gcaggcaac 19

<210> 50
<211> 27
<212> DNA
<213> Artificial Sequence

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

<400> 50
cctttttttg tttgtttttg gcggtgc 27

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

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

<400> 51
agcttactgg atgaccacca 20

<210> 52
<211> 23
<212> DNA
<213> Artificial Sequence

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

<400> 52
gactgggggg aaaagcgcaa tac 23

<210> 53
<211> 23
<212> DNA
<213> Artificial Sequence

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

<400> 53
gtattgcgct tttcccccca gtc 23

<210> 54
<211> 22
<212> DNA
<213> Artificial Sequence

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

<400> 54
tgacttgctc tcatcccaca tg 22

<210> 55
<211> 24
<212> DNA
<213> Artificial Sequence

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

<400> 55
gggcttgaag caagtaaatg gaag 24

<210> 56
<211> 29
<212> DNA
<213> Artificial Sequence

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

<400> 56
gctatcaata ttttcttggt tacagacac 29

<210> 57
<211> 25
<212> DNA
<213> Artificial Sequence

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

<400> 57
gttcactgcc ataagtcttc agtgc 25

<210> 58
<211> 22
<212> DNA
<213> Artificial Sequence

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

<400> 58
tggccgcact gaagacttat gg 22

<210> 59
<211> 24
<212> DNA
<213> Artificial Sequence

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

<400> 59
cagctgcatc tatgataatc cacc 24

<210> 60
<211> 22
<212> DNA
<213> Artificial Sequence

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

<400> 60
atggacaagt ccgaggtgat ag 22

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

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

<400> 61
atcacctcgg acttgtccat tc 22

<210> 62
<211> 24
<212> DNA
<213> Artificial Sequence

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

<400> 62
gcaatcagag atccaaaggc caac 24

<210> 63
<211> 24
<212> DNA

<213> Artificial Sequence

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

<400> 63
gttggccttt ggatctctga ttgc 24

<210> 64
<211> 26
<212> DNA
<213> Artificial Sequence

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

<400> 64
gacatagtat accctggaat tgctgt 26

<210> 65
<211> 26
<212> DNA
<213> Artificial Sequence

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

<400> 65
acagcaattc cagggtatac tatgtc 26

<210> 66
<211> 22
<212> DNA
<213> Artificial Sequence

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

<400> 66
ctcccccgat tttagccaag aa 22

1

17

Claims

1. A nucleic acid molecular structure representing either the *Rhesus hybrid box* or the upstream *Rhesus box* or the downstream *Rhesus box,* the sequence of which is shown respectively in SEQ ID NOs 18 to 20.

2. The nucleic acid molecular structure of claim 1 representative of the common *RHD* negative haplotypes, wherein said *RHD* negative haplotype refers to any RhD antigen negative haplotype that comprises a *hybrid Rhesus box.*

3. The nucleic acid molecular structure of claim 1 representative of an *RHD* negative haplotype comprising an *RHD* gene deletion involving the upstream *Rhesus box,* the downstream *Rhesus box* or both.

4. The nucleic acid molecular structure of claim 1 representative of the common *RHD* positive haplotypes, wherein

said *RHD* positive haplotype refers to any haplotype that comprises DNA sequences specific for the *RHD* gene.

5. The nucleic acid molecular structure of claim 1 derived from a sample comprising an *RHD* positive haplotype that is serologically classified RhD negative, wherein said *RHD* positive haplotype that is serologically classified RhD negative describes a sample that has been tested for the presence of RhD antigen by using for example routine serological assays wherein the result of such assays was negative.

6. The nucleic acid molecular structure of claim 5, wherein said sample is selected from a Caucasian population.

7. A process to specifically detect a *RHD* negative haplotype caused by a nonsense mutation, missense mutation, splice site mutation, partial deletion, partial insertion, partial inversion or a combination thereof within the *RHD* gene, which terminates or obliterates the expression of a protein product of the *RHD* gene in a sample by utilizing either the *hybrid Rhesus box* or the upstream *Rhesus box* or the downstream *Rhesus box,* the sequence of which is shown respectively in SEQ ID NOs 18 to 20 or a structural feature of any one of claims 2 to 6 with techniques known in the art, preferably by PCR-RFLP, PCR-SSP or long-range PCR.

8. A vector comprising the nucleic acid molecular structure of any one of claims 1 to 6.

9. A non-human host transformed with the vector of claim 8.

10. An oligonucleotide hybridizing under stringent conditions to a portion of the nucleic acid molecular structure of any one of claims 1 to 6, wherein said portion hybridizes to a region involving the breakpoint of said hybrid gene or to the complementary portion thereof, wherein said region involving the breakpoint is **characterized in that** the 5' portion of the upstream *RHD* box is in close spatial proximity to the 3' portion of the downstream *RHD* box and wherein said oligonucleotide encompasses 20 nucleotides and is located 5' and 3' of the actual breakpoint.

11. A method for testing for the presence of a nucleic acid molecule carrying a deletion of the *RHD* gene as **characterized by** the nucleic acid molecular structure of any one of claims 1 to 6 in a sample comprising hybridizing the oligonucleotide of claim 10 under stringent conditions to nucleic acid molecules comprised in the sample obtained from a human and detecting said hybridization.

12. The method of claim 11 further comprising digesting the product of said hybridization with a restriction endonuclease and analyzing the product of said digestion.

13. A method for testing for the presence of a nucleic acid molecule carrying a deletion of the *RHD* gene as **characterized by** the nucleic acid molecular structure of any one of claims 1 to 6 in a sample comprising determining the nucleic acid sequence of at least a portion of the nucleic acid molecular structure of any one of claims 1 to 6, said portion encoding a breakpoint of said hybrid gene.

14. The method of claim 13 further comprising, prior to determining said nucleic acid sequence, amplification of at least said portion of said nucleic acid structure.

15. A method for testing for the presence of a nucleic acid molecule carrying a deletion of the *RHD* gene as **characterized by** the nucleic acid molecular structure of any one of claims 1 to 6 in a sample comprising carrying out an amplification reaction using a set of primers that amplifies at least a portion of said sequence, wherein at least one of the primers employed in said amplification reaction is an oligonucleotide hybridizing under stringent conditions to a portion of the nucleic acid molecular structure of any one of claims 1 to 6, wherein said portion hybridizes to a region involving the breakpoint of said hybrid gene or to the complementary portion thereof, wherein said region involving the breakpoint is **characterized in that** the 5' portion of the upstream *RHD* box is in close spatial proximity to the 3' portion of the downstream *RHD* box.

16. The method of claims 14 or 15, wherein said amplification is effected by or said amplification reaction is the polymerase chain reaction (PCR).

17. The method of claim 16, wherein said PCR is PCR-RFLP, PCR-SSP or long-range PCR.

18. The method of any one of claims 14 to 17, wherein the molecular weight of the amplification product is analyzed.

19. A method for testing for the presence of the nucleic acid molecular structure of claims 5 or 6 encoding *RHD* positive alleles comprising the following steps:

(a) isolating the DNA from a blood sample or a sample obtained from a blood donor;
(b) hybridizing at least two oppositely oriented primers under stringent conditions to the DNA so as to carry out a PCR;
(c) amplifying the target sequence;
(d) separating the amplification products on a gel; and
(e) analyzing the amplicons.

20. The method of claim 19, wherein said *RHD* positive alleles are derived from a serologically RhD negative sample.

21. The method of claim 19 or 20, wherein said sample is selected from a Caucasian population.

22. The method of any one of claims 7 and 11 to 21, wherein said sample is blood, serum, plasma, fetal tissue, saliva, urine, mucosal tissue, mucus, vaginal tissue, fetal tissue obtained from the vagina, skin, hair, hair follicle or another human tissue.

23. The method of claim 22 comprising enrichment of fetal cells or extraction of fetal DNA or mRNA from maternal tissue, like peripheral blood, serum or plasma.

24. The method of any one of claims 7 and 11 to 23 wherein said nucleic acid molecule or proteinaceous material from said sample is fixed to a solid support.

25. The method of claim 24, wherein said solid support is a chip.

26. Use of the nucleic acid molecular structure of any one of claims 1 to 6 for the analysis of a negative or a positive Rhesus D phenotype.

27. A method for determining whether a patient in need of a blood transfusion is to be transfused with RhD negative blood from a donor comprising the step of testing a sample from said patient for the presence of one or more nucleic acid molecular structures of any one of claim 2 and claims 5 and 6, wherein a positive testing for two different of said nucleic acid molecular structures is indicative of the need for a transfusion with Rh negative blood.

28. A method for determining whether blood of a donor may be used for transfusion to a patient in need thereof comprising the step of testing a sample from said donor for the presence of one or more nucleic acid molecular structures of any one of claims 1 to 6, wherein a negative testing for the nucleic acid molecular structure of claim 2 with or without a negative testing for one or more nucleic acid molecular structures of claim 4 to 6 excludes the transfusion of the donor's blood to a patient that is typed as RhD negative.

29. A method of assessing of the risk of a RhD negative mother of conceiving or carrying an RhD positive fetus or of the risk of a mother having an anti-D titer of conceiving or carrying a fetus at risk to develop hemolytic disease of the newborn comprising assessing a sample obtained from the father of the fetus for the presence of one or more nucleic acid molecular structures as defined in any one of claims 1 to 6.

30. The method of claim 29, wherein said nucleic acid molecular structure carries mutations or deletions.

31. A method for assessing the possibility or likelihood of a man being the father of a child by assaying a sample obtained from said man for the presence of one or more nucleic acid molecular structures of any one of claims 1 to 6, wherein the test results are used to determine the homozygosity for, the heterozygosity for or the absence of any nucleic acid molecular structure or a nucleic molecule of claims 2 and 4 to 6 used to infer the possibility or likelihood of said man being the father of the child.

32. Kit comprising

(a) the oligonucleotide of claim 10; and/or
(b) a pair of primers useful for carrying out the amplification reaction of any one of claims 15 to 17.

33. Use of an oligonucleotide for hybridizing under stringent conditions to a portion of the nucleic acid molecular structure of any one of claims 1 to 6, wherein said portion hybridizes to a region involving the breakpoint of said hybrid gene or to the complementary portion thereof, wherein said region involving the breakpoint is **characterized in that** the 5' portion of the upstream *Rhesus* box is in close spatial proximity to the 3' portion of the downstream *Rhesus* box.

34. A method for testing for the presence of the nucleic acid molecular structure of claim 3 representative of a common *RHD* negative haplotype comprising the following steps:

(a) isolating the DNA from a blood sample or a sample obtained from a blood donor;
(b) hybridizing at least two oppositely oriented primers under stringent conditions to the DNA so as to carry out a PCR;
(c) amplifying the target sequence;
(d) separating the amplification products on a gel; and
(f) analyzing the amplicons.

Patentansprüche

1. Nucleinsäuremolekülstruktur, welche entweder die *Rhesus Hybrid Box,* die stromaufwärts gelegene *Rhesus-Box* oder die stromabwärts gelegene *Rhesus-Box* ist, und deren Sequenz in den SEQ ID NOs: 18 bis 20 gezeigt ist.

2. Nucleinsäuremolekülstruktur nach Anspruch 1, welche die gemeinsamen RHD-negativen Haplotypen repräsentiert, wobei sich der RHD-negative Haplotyp auf jeden RhD Antigen-negativen Haplotyp bezieht, der eine *Hybrid-Rhesus-Box* beinhaltet.

3. Nucleinsäuremolekülstruktur nach Anspruch 1, welche einen *RHD*-negativen Haplotypen repräsentiert, umfassend eine *RHD*-Gen-Deletion in der stromaufwärts gelegenen *Rhesus-Box,* der stromabwärts gelegenen *Rhesus-Box* oder in beiden.

4. Nucleinsäuremolekülstruktur nach Anspruch 1, welche die gemeinsamen *RHD*-positiven Haplotypen repräsentiert, wobei sich der *RHD*-positive Haplotyp auf jeden Haplotyp bezieht, der DNA-Sequenzen enthält, die spezifisch für das *RHD*-Gen sind.

5. Nucleinsäuremolekülstruktur nach Anspruch 1, abgeleitet aus einer Probe, welche einen *RHD*-positiven Haplotypen umfasst, welcher serologisch als RhD-negativ eingestuft wird, wobei sich der *RHD*-positive Haplotyp, welcher serologisch negativ eingestuft wird, auf eine Probe bezieht, welche auf das Vorhandensein des RhD-Antigens mittels routinemäßigen serologischen Tests getestet wurde, und für den die Ergebnisse dieser Tests negativ waren.

6. Nucleinsäuremolekülstruktur nach Anspruch 5, wobei die Probe aus einer kaukasischen Population stammt.

7. Verfahren zum spezifischen Nachweis eines *RHD*-negativen Haplotypen verursacht durch eine Nonsense-Mutation, Missense-Mutation, Spleißstellen-Mutation, teilweise Deletion, teilweise Insertion, teilweise Inversion oder eine Kombination dieser im *RHD*-Gen, wodurch die Expression eines Proteinproduktes des *RHD*-Gens in einer Probe beendet oder gestört wird, unter Verwendung der *Hybrid-Rhesus-Box* oder der stromaufwärts gelegenen *Rhesus-Box* oder der stromabwärts gelegenen *Rhesus-Box,* deren Sequenzen jeweils in SEQ ID NOs: 18 bis 20 dargestellt sind, oder eines strukturellen Merkmales nach einem der Ansprüche 2 bis 6 mit Hilfe von Techniken, welche im Stand der Technik bekannt sind, vorzugsweise mittels PCR-RFLP, PCR-SSP oder Long-Range-PCR.

8. Vektor, enthaltend eine Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6.

9. Nicht-menschlicher Wirt, welcher mit dem Vektor nach Anspruch 8 transformiert ist.

10. Oligonucleotid, welches unter stringenten Bedingungen mit einem Teil der Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6 hybridisiert, wobei dieser Teil mit einer Region, welche den Schnittpunkt des Hybridgens einbezieht oder deren komplementären Strang hybridisiert; wobei die Region, welche den Schnittpunkt einbezieht, **dadurch gekennzeichnet ist, dass** der 5'-Teil der stromaufwärts gelegenen *RHD*-Box in großer räumlicher Nähe zum 3'-Teil der stromabwärts gelegenen *RHD*-Box liegt, und worin das Oligonucleotid 20 Nucleotide umfasst und sich 5' und 3' vom tatsächlichen Schnittpunkt befindet.

11. Verfahren zum Testen des Vorhandenseins eines Nucleinsäuremoleküls, welches eine Deletion des *RHD*-Gens wie **gekennzeichnet durch** die Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6 trägt, in einer Probe, umfassend das Hybridisieren des Oligonucleotides nach Anspruch 10 unter stringenten Bedingungen an Nucleinsäuremoleküle, welche in einer Probe enthalten sind, die einem Menschen entnommen wurde, und Nachweis der Hybridisierung.

12. Verfahren nach Anspruch 11, des weiteren umfassend Verdau des Hybridisierungsproduktes mit einer Restriktionsendonuclease und Analyse des Produktes dieses Verdaus.

13. Verfahren zum Testen des Vorhandenseins eines Nucleinsäuremoleküls, welches eine Deletion des *RHD*-Gens wie **gekennzeichnet durch** die Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6 trägt, in einer Probe, umfassend das Bestimmen der Nucleinsäuresequenz von wenigstens einem Teil der Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6, wobei dieser Teil den Schnittpunkt des Hybridgenes codiert.

14. Verfahren nach Anspruch 13, des weiteren umfassend Amplifikation wenigstens eines Teiles der Nucleinsäuremolekülstruktur vor dem Bestimmen der Nucleinsäuresequenz.

15. Verfahren zum Testen des Vorhandenseins eines Nucleinsäuremoleküls, welches eine Deletion des *RHD*-Gens wie **gekennzeichnet durch** die Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6 trägt, in einer Probe, umfassend das Ausführen einer Amplifikationsreaktion unter Verwendung eines Primersatzes, welcher wenigstens einen Teil der Sequenz amplifiziert, wobei wenigstens einer der in der Amplifikationsreaktion verwendeten Primer ein Oligonucleotid ist, welches unter stringenten Bedingungen mit einem Teil der Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6 hybridisiert, wobei dieser Teil mit einer Region hybridisiert, welche den Schnittpunkt des Hybridgens beinhaltet, oder mit dem komplementären Teil davon, wobei die Region, welche den Schnittpunkt, **dadurch gekennzeichnet ist**, dass der 5'-Teil der stromaufwärts gelegenen *RHD*-Box in großer räumlicher Nähe zum 3'-Teil der stromabwärts gelegenen *RHD*-Box liegt.

16. Verfahren nach Anspruch 14 oder 15, wobei die Amplifikation mittels einer Polymerasekettenreaktion (PCR) erfolgt oder eine PCR ist.

17. Verfahren nach Anspruch 16, wobei die PCR eine PCR-RFLP, PCR-SSP oder eine Long-Range-PCR ist.

18. Verfahren nach einem der Ansprüche 14 bis 17, wobei das Molekulargewicht des Amplifikationsproduktes analysiert wird.

19. Verfahren zum Testen des Vorhandenseins einer Nucleinsäuremolekülstruktur nach Anspruch 5 oder 6, welche *RHD*-positive Allele codiert, umfassend die folgenden Schritte:

(a) Isolieren der DNA aus einer Blutprobe oder aus einer Probe eines Blutspenders;
(b) Hybridisieren von wenigstens zwei entgegengesetzt ausgerichteten Primern unter stringenten Bedingungen an die DNA zur Ausführung einer PCR;
(c) Amplifizieren der Zielsequenz;
(d) Auftrennen der Amplifikationsprodukte auf einem Gel; und
(e) Analysieren der Amplikons.

20. Verfahren nach Anspruch 19, wobei die *RHD*-positiven Allele aus einer serologisch *RhD*-negativen Probe stammen.

21. Verfahren nach den Ansprüchen 19 oder 20, wobei die Probe aus einer kaukasischen Population ausgewählt wurde.

22. Verfahren nach einem der Ansprüche 7 und 11 bis 21, wobei die Probe Blut, Serum, Plasma, fötales Gewebe, Speichel, Urin, Schleimhautgewebe, Schleim, Scheidengewebe, fötales Gewebe aus der Scheide, Haut, Haar, Haarfollikel oder anderes menschliches Gewebe ist.

23. Verfahren nach Anspruch 22 umfassend eine Anreicherung von fötalen Zellen oder die Extraktion von fötaler DNA oder mRNA von mütterlichem Gewebe, wie beispielsweise peripherem Blut, Serum oder Plasma.

24. Verfahren nach einem der Ansprüche 7 und 11 bis 23, wobei das Nucleinsäuremolekül oder proteinhaltige Material der Probe auf einem festen Träger fixiert wird.

25. Verfahren nach Anspruch 24, wobei der feste Träger ein Chip ist.

26. Verwendung der Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6 für die Analyse eines negativen oder positiven Rhesus D-Phenotypen.

27. Verfahren zur Ermittlung, ob einem Patienten, welcher eine Bluttransfusion benötigt, RhD-negatives Blut von einem Spender übertragen werden kann, umfassend den Schritt des Testens einer Probe des Patienten auf die Anwesenheit einer oder mehrerer Nucleinsäuremolekülstrukturen nach einem der Ansprüche 2 und 5 und 6, wobei ein positiver Test für zwei verschiedene dieser Nucleinsäuremolekülstrukturen anzeigt, dass eine Transfusion mit Rhnegativem Blut notwendig ist.

28. Verfahren zur Ermittlung, ob Blut eines Spenders für eine Bluttransfusion auf einen Patienten geeignet ist, der dessen bedarf, umfassend den Schritt des Testens der Probe des Spenders auf die Anwesenheit einer oder mehrerer Nucleinsäuremolekülstrukturen nach einem der Ansprüche 1 bis 6, wobei ein negativer Test für die Nucleinsäurestruktur nach Anspruch 2 mit oder ohne einem negativen Test für eine oder mehrere Nucleinsäuremolekülstrukturen nach Anspruch 4 bis 6 die Verwendung des Blutes des Spenders für eine Transfusion auf einen Patienten, welcher RhD-negativ ist, ausschließt.

29. Verfahren zur Einschätzung des Risikos, dass eine RhD-negative Mutter einen RhD-positiven Fötus empfängt oder trägt, oder des Risikos, dass eine Mutter, welche einen Anti-D Titer aufweist, einen Fötus empfängt oder trägt, welcher einem Risiko der Entwicklung einer hämolytischen Erkrankung von Neugeborenen ausgesetzt ist, umfassend das Testen einer Probe vom Vater des Fötus auf die Anwesenheit einer oder mehrerer Nucleinsäuremolekülstrukturen nach einem der Ansprüche 1 bis 6.

30. Verfahren nach Anspruch 29, wobei die Nucleinsäuremolekülstruktur Mutationen oder Deletionen trägt.

31. Verfahren zur Einschätzung der Möglichkeit oder Wahrscheinlichkeit, dass ein Mann Vater eines Kindes ist, durch Testen einer Probe des Mannes auf die Anwesenheit einer oder mehrerer Nucleinsäuremolekülstrukturen nach einem der Ansprüche 1 bis 6, wobei die Testergebnisse genutzt werden um Homozygotie, Heterozygotie oder die Abwesenheit einer beliebigen Nucleinsäuremolekülstruktur oder eines Nucleinsäuremoleküles nach Anspruch 2 und 4 bis 6 zu bestimmen, um die Möglichkeit oder Wahrscheinlichkeit, dass der Mann der Vater des Kindes ist, abzuleiten.

32. Kit beinhaltend

(a) das Oligonucleotid nach Anspruch 10; und/oder
(b) ein Primerpaar für die Ausführung einer Amplifikationsreaktion nach einem der Ansprüche 15 bis 17.

33. Verwendung eines Oligonucleotids zur Hybridisierung unter stringenten Bedingungen an einen Teil der Nucleinsäuremolekülstruktur nach einem der Ansprüche 1 bis 6, wobei dieser Teil mit einer Region hybridisiert, welche den Schnittpunkt des Hybridgenes umfasst, oder an eine komplementäre Region davon, wobei diese Region, welche den Schnittpunkt beinhaltet,
dadurch gekennzeichnet ist, dass der 5'-Teil der stromaufwärts gelegenen *Rhesus*-Box sich in großer räumlicher Nähe zum 3'-Teil der stromabwärts gelegenen *Rhesus*-Box befindet.

34. Verfahren zum Testen des Vorhandenseins einer Nucleinsäuremolekülstruktur nach Anspruch 3, welche den gemeinsamen *RHD*-negativen Haplotyp repräsentiert, umfassend die folgenden Schritte:

(a) Isolieren der DNA aus einer Blutprobe oder aus einer Probe eines Blutspenders;
(b) Hybridisieren von wenigstens zwei entgegengesetzt ausgerichteten Primern unter stringenten Bedingungen an die DNA zur Ausführung einer PCR;
(c) Amplifizieren der Zielsequenz;
(d) Auftrennen des Amplifikationsproduktes auf einem Gel; und
(e) Analysieren der Amplikons.

Revendications

1. Structure moléculaire d'acide nucléique représentant soit la boîte hybride Rhésus, soit la boîte Rhésus en amont, soit la boîte Rhésus en aval, dont la séquence est représentée respectivement dans SEQ ID NOs : 18 à 20.

2. Structure moléculaire d'acide nucléique selon la revendication 1, représentative des haplotypes *RHD* négatifs communs, un dit haplotype *RHD* négatif se référant à tout haplotype négatif pour l'antigène RhD qui comprend une boîte Rhésus hybride.

3. Structure moléculaire d'acide nucléique selon la revendication 1, représentative d'un haplotype *RHD* négatif comprenant une délétion du gène *RHD* impliquant la boîte Rhésus en amont, la boîte Rhésus en aval ou les deux.

4. Structure moléculaire d'acide nucléique selon la revendication 1, représentative des haplotypes *RHD* positifs communs, un dit haplotype *RHD* positif se référant à tout haplotype qui comprend des séquences d'ADN spécifiques du gène *RHD.*

5. Structure moléculaire d'acide nucléique selon la revendication 1 dérivée d'un échantillon comprenant un haplotype *RHD* positif qui est classé RhD négatif sur le plan sérologique, ledit haplotype *RHD* positif qui est classé RhD négatif sur le plan sérologique décrivant un échantillon qui a été testé vis-à-vis de la présence de l'antigène RhD en utilisant, par exemple, des dosages sérologiques de routine, le résultat de tels dosages ayant été négatif.

6. Structure moléculaire d'acide nucléique selon la revendication 5, dans laquelle ledit échantillon est choisi à partir d'une population Caucasienne.

7. Procédé pour détecter spécifiquement un haplotype *RHD* négatif provoqué par une mutation non sens, mutation faux sens, mutation de site d'épissage, délétion partielle, insertion partielle, inversion partielle ou une combinaison de celles-ci dans le gène *RHD*, qui interrompt ou supprime l'expression d'un produit protéique du gène *RHD* dans un échantillon en utilisant soit la boîte Rhésus hybride, soit la boîte Rhésus en amont, soit la boîte Rhésus en aval, dont la séquence est représentée respectivement dans SEQ ID NOs : 18 à 20, soit une caractéristique structurale selon l'une quelconque des revendications 2 à 6, avec des techniques connues dans le domaine, de préférence par PCR-RFLP, PCR-SSP ou PCR longue distance.

8. Vecteur comprenant la structure moléculaire d'acide nucléique selon l'une quelconque des revendications 1 à 6.

9. Hôte non humain transformé avec le vecteur selon la revendication 8.

10. Oligonucléotide s'hybridant dans des conditions stringentes à une partie de la structure moléculaire d'acide nucléique selon l'une quelconque des revendications 1 à 6, ladite partie s'hybridant à une région comportant le point de cassure dudit gène hybride ou à la partie complémentaire de celle-ci, ladite région comportant le point de cassure étant **caractérisée en ce que** la partie 5' de la boîte *RHD* en amont est à proximité spatiale étroite de la partie 3' de la boîte *RHD* en aval et ledit oligonucléotide couvrant 20 nucléotides et étant localisé en 5' et 3' du point de cassure réel.

11. Méthode pour tester la présence d'une molécule d'acide nucléique portant une délétion du gène *RHD* telle que **caractérisée par** la structure moléculaire d'acide nucléique selon l'une quelconque des revendications 1 à 6 dans un échantillon, comprenant l'hybridation de l'oligonucléotide selon la revendication 10 dans des conditions stringentes à des molécules d'acide nucléique comprises dans l'échantillon obtenu à partir d'un humain et la détection de ladite hybridation.

12. Méthode selon la revendication 11, comprenant en outre la digestion du produit de ladite hybridation avec une endonucléase de restriction et l'analyse du produit de ladite digestion.

13. Méthode pour tester la présence d'une molécule d'acide nucléique portant une délétion du gène *RHD* telle que **caractérisée par** la structure moléculaire d'acide nucléique selon l'une quelconque des revendications 1 à 6 dans un échantillon, comprenant la détermination de la séquence d'acide nucléique d'au moins une partie de la structure d'acide nucléique selon l'une quelconque des revendications 1 à 6, ladite partie codant un point de cassure dudit gène hybride.

14. Méthode selon la revendication 13, comprenant en outre, avant la détermination de ladite séquence d'acide nucléi-

que, l'amplification d'au moins ladite partie de ladite structure d'acide nucléique.

15. Méthode pour tester la présence d'une molécule d'acide nucléique portant une délétion du gène *RHD* telle que **caractérisée par** la structure moléculaire d'acide nucléique selon l'une quelconque des revendications 1 à 6 dans un échantillon, comprenant la réalisation d'une réaction d'amplification en utilisant un ensemble d'amorces qui amplifie au moins une partie de ladite séquence, dans laquelle au moins l'une des amorces utilisées dans ladite réaction d'amplification est un oligonucléotide s'hybridant dans des conditions stringentes à une partie de la structure moléculaire d'acide nucléique selon l'une quelconque des revendications 1 à 6, ladite partie s'hybridant à une région comportant le point de cassure dudit gène hybride ou à la partie complémentaire de celle-ci, ladite région comportant le point de cassure étant **caractérisée en ce que** la partie 5' de la boîte *RHD* en amont est à proximité spatiale étroite de la partie 3' de la boîte *RHD* en aval.

16. Méthode selon la revendication 14 ou 15, dans laquelle ladite amplification est réalisée par, ou ladite réaction d'amplification est, l'amplification en chaîne par polymérase (PCR).

17. Méthode selon la revendication 16, dans laquelle ladite PCR est une PCR-RFLP, PCR-SSP ou PCR longue distance.

18. Méthode selon l'une quelconque des revendications 14 à 17, dans laquelle le poids moléculaire du produit d'amplification est analysé.

19. Méthode pour tester la présence d'une structure moléculaire d'acide nucléique selon la revendication 5 ou 6 codant des allèles positifs vis-à-vis de *RHD*, comprenant les étapes suivantes :

(a) isolement de l'ADN provenant d'un échantillon de sang ou d'un échantillon obtenu à partir d'un donneur de sang ;
(b) hybridation d'au moins deux amorces orientées de manière opposée, dans des conditions stringentes, à l'ADN de manière à réaliser une PCR ;
(c) amplification de la séquence cible ;
(d) séparation des produits d'amplification sur un gel ; et
(e) analyse des amplicons.

20. Méthode selon la revendication 19, dans laquelle lesdits allèles positifs vis-à-vis de *RHD* sont dérivés à partir d'un échantillon RhD négatif sur le plan sérologique.

21. Méthode selon la revendication 19 ou 20, dans laquelle ledit échantillon est choisi à partir d'une population Caucasienne.

22. Méthode selon l'une quelconque des revendications 7 et 11 à 21, dans laquelle ledit échantillon est du sang, sérum, plasma, tissu foetal, salive, urine, tissu muqueux, mucus, tissu vaginal, tissu foetal obtenu à partir du vagin, peau, cheveu, follicule capillaire ou autre tissu humain.

23. Méthode selon la revendication 22, comprenant l'enrichissement en cellules foetales ou l'extraction de l'ADN ou l'ARNm foetal, par rapport à du tissu maternel, comme le sang périphérique, le sérum ou le plasma.

24. Méthode selon l'une quelconque des revendications 7 et 11 à 23, dans laquelle ladite molécule d'acide nucléique ou ledit matériel protéique provenant dudit échantillon est fixé à un support solide.

25. Méthode selon la revendication 24, dans laquelle ledit support solide est une puce.

26. Utilisation de la structure moléculaire d'acide nucléique selon l'une quelconque des revendications 1 à 6, pour l'analyse d'un phénotype Rhésus D négatif ou positif.

27. Méthode pour déterminer si un patient nécessitant une transfusion sanguine doit être transfusé avec du sang négatif vis-à-vis de RhD provenant d'un donneur, comprenant l'étape de test d'un échantillon provenant dudit patient vis-à-vis de la présence d'une ou plusieurs structures moléculaires d'acide nucléique selon l'une quelconque de la revendication 2 et des revendications 5 et 6, dans laquelle un test positif pour deux desdites structures moléculaires d'acide nucléique différentes indique le besoin d'une transfusion avec du sang négatif vis-à-vis de Rh.

28. Méthode pour déterminer si le sang d'un donneur peut être utilisé pour la transfusion d'un patient le nécessitant, comprenant l'étape de test d'un échantillon provenant dudit donneur vis-à-vis de la présence d'une ou plusieurs structures moléculaires d'acide nucléique selon l'une quelconque des revendications 1 à 6, dans laquelle un test négatif vis-à-vis de la structure moléculaire d'acide nucléique selon la revendication 2, avec ou sans un test négatif vis-à-vis d'une ou plusieurs structures moléculaires d'acide nucléique des revendications 4 à 6, exclut la transfusion du sang du donneur à un patient qui est typé RhD négatif.

29. Méthode pour évaluer le risque pour une mère RhD négative de concevoir ou porter un foetus RhD positif, ou du risque pour une mère ayant un titre anti-D de concevoir ou porter un foetus risquant de développer une maladie hémolytique du nouveau-né, comprenant l'évaluation d'un échantillon obtenu à partir du père du foetus vis-à-vis de la présence d'une ou plusieurs structures moléculaires d'acide nucléique telles que définies dans l'une quelconque des revendications 1 à 6.

30. Méthode selon la revendication 29, dans laquelle ladite structure moléculaire d'acide nucléique porte des mutations ou des délétions.

31. Méthode pour évaluer la possibilité ou la probabilité pour un homme d'être le père d'un enfant en dosant un échantillon obtenu à partir dudit homme vis-à-vis de la présence d'une ou plusieurs structures moléculaires d'acide nucléique selon l'une quelconque des revendications 1 à 6, dans laquelle les résultats de test sont utilisés pour déterminer l'homozygotie pour, l'hétérozygotie pour, ou l'absence d'une quelconque structure moléculaire d'acide nucléique ou d'une molécule d'acide nucléique selon les revendications 2 et 4 à 6 utilisées pour conclure à la possibilité ou à la probabilité pour ledit homme d'être le père de l'enfant.

32. Kit comprenant

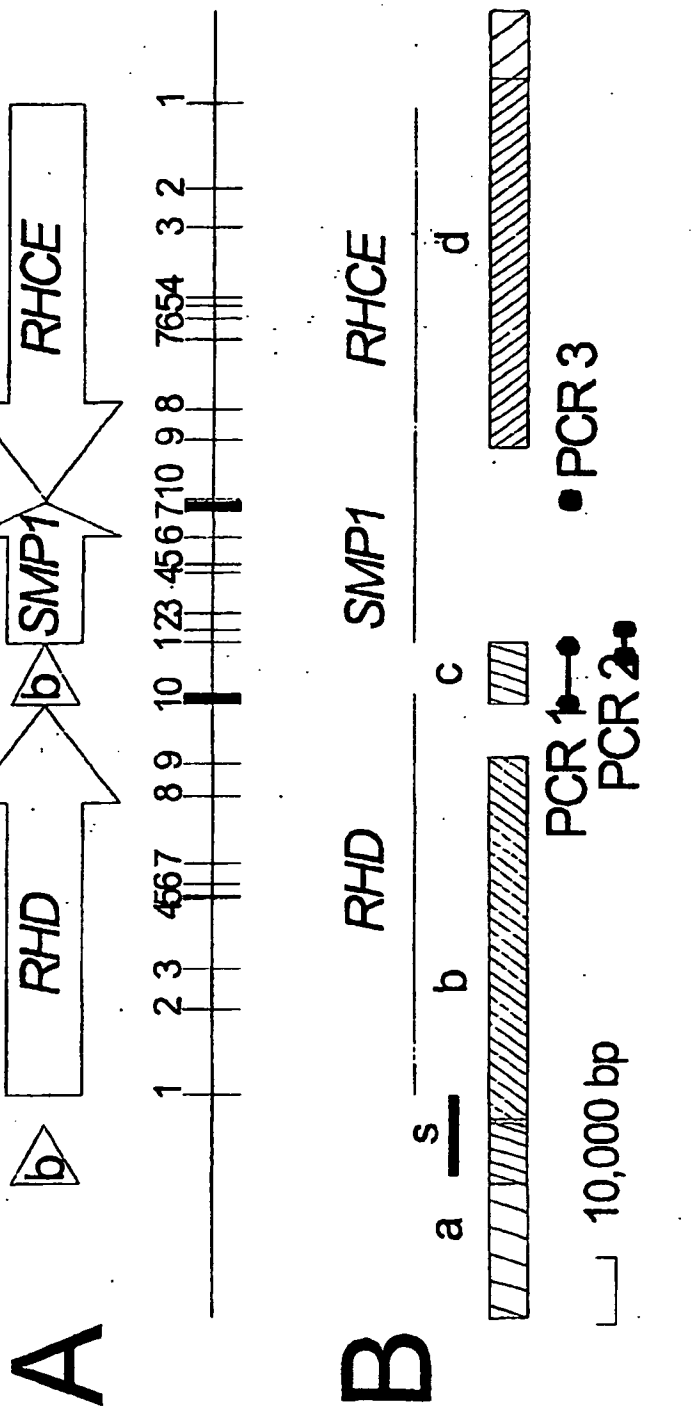
(a) l'oligonucléotide selon la revendication 10 ; et/ou
(b) une paire d'amorces utile pour réaliser la réaction d'amplification selon l'une quelconque des revendication 15 à 17.

33. Utilisation d'un oligonucléotide pour l'hybridation, dans des conditions stringentes, à une partie d'une structure moléculaire d'acide nucléique selon l'une quelconque des revendications 1 à 6, dans laquelle ladite partie s'hybride à une région comprenant le point de cassure dudit gène hybride ou à la partie complémentaire de celle-ci, ladite région comprenant le point de cassure étant **caractérisée en ce que** la partie 5' de la boîte Rhésus en amont est à proximité spatiale étroite de la partie 3' de la boîte Rhésus en aval.

34. Méthode pour tester la présence de la structure moléculaire d'acide nucléique selon la revendication 3, représentative d'un haplotype *RHD* négatif commun, comprenant les étapes suivantes :

(a) isolement de l'ADN provenant d'un échantillon de sang ou d'un échantillon obtenu à partir d'un donneur de sang ;
(b) hybridation d'au moins deux amorces orientées de manière opposée dans des conditions stringentes à l'ADN de manière à réaliser une PCR ;
(c) amplification de la séquence cible ;
(d) séparation des produits d'amplification sur un gel ; et
(f) analyse des amplicons.

Fig. 1


A
RHD
SMP1
RHCE
b
1 2 3 4 5 6 7 8 9 10 1 2 3 4 5 6 7 8 9 10 1
B
RHD
SMP1
RHCE
a
s
b
c
d
PCR 1
PCR 2
PCR 3
10,000 bp

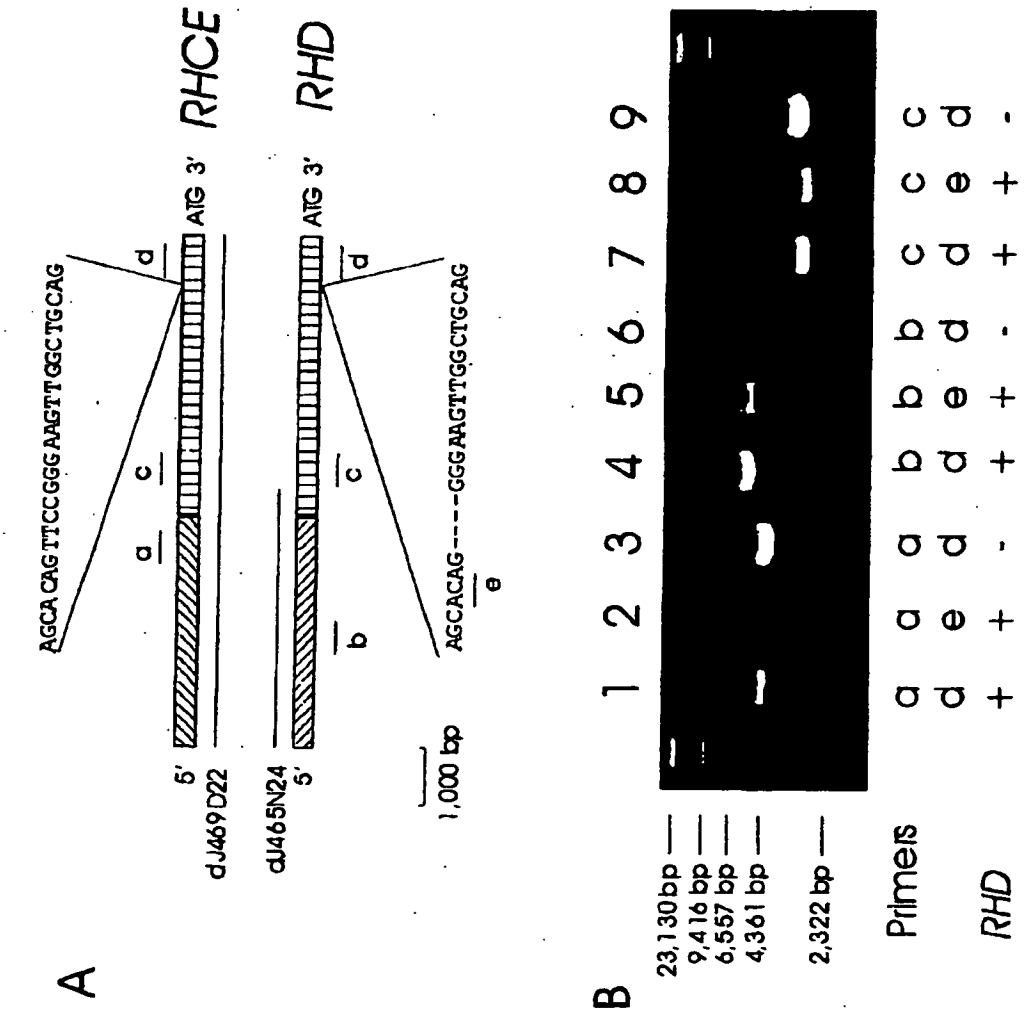

A
AGCA CAG TTCCGGG AAGTTGGCTGCAG
RHCE
RHD
ATG 3'
5'
dJ469D22
dJ465N24
AGCACAG----GGGAAGTTGGCTGCAG
1,000 bp
B
1 2 3 4 5 6 7 8 9
23,130 bp
9,416 bp
6,557 bp
4,361 bp
2,322 bp
Primers
RHD

Fig. 2


cpg
Identity
MER20
L1MC3
1,000 bp
Other
Alu
MLT
L2
THE1B
Tigger3

Fig.3

Fig.4

```
attcactatcacaagaacagcacgggtaa
**************************
**-**********g*************

  -------rez7----------->
gacctgtcccatgattcagttacct
*************************
*************************

cccactgggtcccctcccacaacgcatgggaattcaggatgagattt
***********************************************
***********************************************

gggtggggacacaaccaaaccctatcattccacccatggcccctcccaaa
***************************************************
***************************************************

ttcatgtcctcacatttcaaaaccaatcacaccatcccaacacagtccctc
***************************************************
***************************************************

aaagtcttaaatgatttcagcattaactcaaaagtccacagtctaatgtc
***************************************************
***************************************************
```

Fig.4 **cont.**

```
tcatctgagacaaggcaagtcctttccatttatgagcctataaaatccaa
**********************************************
**********************************c***********
agcaagttagttacttcctagatacaatgggggtacaggcattgggtaaa
**********************************************
**********a***********************************
tacagccattccaaatgggataaattggtcaaaacaaagaggctacaggc
**********************************************
**********************************************
ccatgagagtccaaaatccagtgggcagtcaaatcttaaagctccaaaat
**********************************************
**********************************************
gatctcct-ttgactccacatctcacatccaggtcacgcagatggaagg
*********-**********************************
**********c*********************************
ggtgggttcccatggtcttgggcagctctgccccctgtaccttttgcagggt
**********************************************
**********************************************
```

Fig.4 **cont.**

```
acagcctccctctcagctgctttcatgggctggcattgagtgtctgcagc
**************************************************
**************************************************a*
                                                  ^^^^^^

ttttccaggtacacggtgcaagctgtcggtggatctaccattctggggtc
**************************************************
**************************************************

tggaggacctcttctcacagctccactaggtggtgccccagtagggactg
**************************************************
**************************************************

tgtgtggggtctctgaccccacatttcccttctgcactgccctggcagag
**************************************************
**************************************************

gatctccatgagggccctgctcctgcagcaaacttctgactgggcatcca
**************************************************
**********************c**********************c**********
                      ^^^^^^

ggcatttccgcacatcctctttaatctaggcgaaggtttccaaaccccaa
**************************************************
**************************************************g
```

Fig. 4 cont.

```
ttcttgacttctgtgcactcgcagtctcaacaccacatggaagctgtcaa
**************************************************
**************************************************

ggcttggggcttgcactccccgaagctacagcccaagctctaccttgcct
**************************************************
**************************************************

cccgtcagtcatggttgggagtggctgggatgcagggcaccaagtcccta
**************************************************
**t***********************************************

ggctgcacacagcatgaggaccccgggcctggccaacaaaaccatttttt
**************************************************
**************************************************

        +-- breakpoint region ->
        +-- identity region ->
cctgatacctctggacctgtgatgggaggggttgccataaagacctctga
**************************************************
*******t******************************************

catgccctggagacattttccccattgtcttgggaattagcatttggctc
**************************************************
**************************************************
```

Fig.4 **cont.**

```
ctgttactcatgcaaatttctgcagccagcttgaatttctcctcagaaaa
**************************************************
**************************************************
                      ^^^^^^

tgggaatttttctttctatcacattgtcaggctgcaaattttccgaactt
**************************************************
**************************************************

ttatgctctgcttcccttataaaactgaatgtctttaacagcacccaag
**************************************************
**************************************************

tcacctcttgaatgctttgctgcttagaaatttctcctgccagatactct
**************************************************
**************************************************

aaatcatctctctgaagttcaaagttctacaaatatctcgtgcaggggca
**************************************************
**************************************************

aaatgccgccagtatctttgctaaaacataacaagagtccccttgctcc
**************************************************
**************************************************
```

Fig.4 **cont.**

agttcccaacaagttcctcatttccgtctgagaccacctcagcctatgga
****
****

ctttattgtccacagtgctatcagcattttgggcaaagccattcaacaag
****
****
tctctaggaagttccaaactttcccacatttgcctgtcttcttctgagcc
****
****

ctccaaactgttccaaaccctgcctgttacccagttccaaagtcacatac
****
****

ccatttttgagtatctacggcagcaccccactctactggtaccaatttag
****
****

ccactgaagtagttggagaacagaagtaatagactctggtttacattgta
****
****

aaagcttctctgtggctgctgtgtgaagaaaatatatgagaatgaagccc
****
****

Fig.4 **cont.**

```
caagatgaagcagggacacagttgcagtggttagagtaagaaatgctgct
**************************************************
**************************************************

ggctggcactgaagtgatagcctggaggtttgtgtgtgcacatgcatgtg
**************************************************
**************************************************

tatgtgttttacgatagtaggcccaacagat
*******************************
*******************************

  <-- breakpoint region ----+
actgtaatccacacttgtttttttttttt----ga
*******************************tttt**
*******************************tttt**

gacagagtctcacctgttgcctagactagaatgc
**********************************
**********************************

agtggcacaatcttggctcactacaacctccacctcccaggttcaaacaatc
**************************************************
**************************************************
```

Fig.4 **cont.**

```
cttgtgcttcagcctcccgagtagttgggattacaggtgtgtgccacc
**************************************************
**************************************************

gtgcccagctatatttttgtatttttagcagagatgggattttgccacat
**************************************************
**************************************************

tggccaggctggtcttgaactcctggcctcaagcaatcctcccacctta
**************************************************
**************************************************

gcctcccaaagtgctgagccaccacacctggccgcaactgattttaatc
**************************************************
**************************************************

atgaaatgacacatacatttaaaaaacccaatacctataatattcctggc
**************************************************
**************************************************

tagtactcttcacatctatatcatcaaaaacaaagaaagtatgtgaaact
**************************************************
**************************************************
```

Fig.4 **cont.**

```
gacacagccaaggggagactaaggagacataacaattaactgtaatgtgg
**************************************************
**************************************************

tattctggaggggatcctggaacagaaaaagacattaggcaaaaaactaa
**************************************************
**************************************************

agaaatctgaataaaatgtggatgtcagttaataataatgtatcatatta
**************************************************
**************************************************

    <-- identity region-+
gtccagtaattgtaacaaatataccacaataatgaaagccattaattata
**************************-************************
**************************-************************

gggaaaatggaggggttaatatgggtggctggcttttgctatttctagcag
**************************************************
**************************************************

ctccattttatctgcaaaagacaaacattcattaagtcccaaaaaggtaaa
************a*************************************
************a*************************************
```

Fig.4 **cont.**

```
gaatgacaaattaagcatgtatcttattagtaagagtaatataaagatg
**********************************************
**********************************************

ctcactcctatttataaatatttgacaatcatgttaaggccacaaaagag
*******a**********************g**********g*******
*******a**********************g**********g*******

aaaaaagggtagggggcaaaaaacgcaaagagaaaggagttagtatctttt
**********************************************
**********************************************

ctcccgcactcattagctattaaaagaggatgtttgtttaaagctgctca
**********************************************
**********************************************

gagctggtaaactaatgttaagtcactaacgggaatttaaaaggtttcat
**********************************************
**********************************************

taagaactgcctgcactagattcctccaccctgagacattaaacaatcac
**********************************************
**********************************************
```

Fig.4 **cont.**

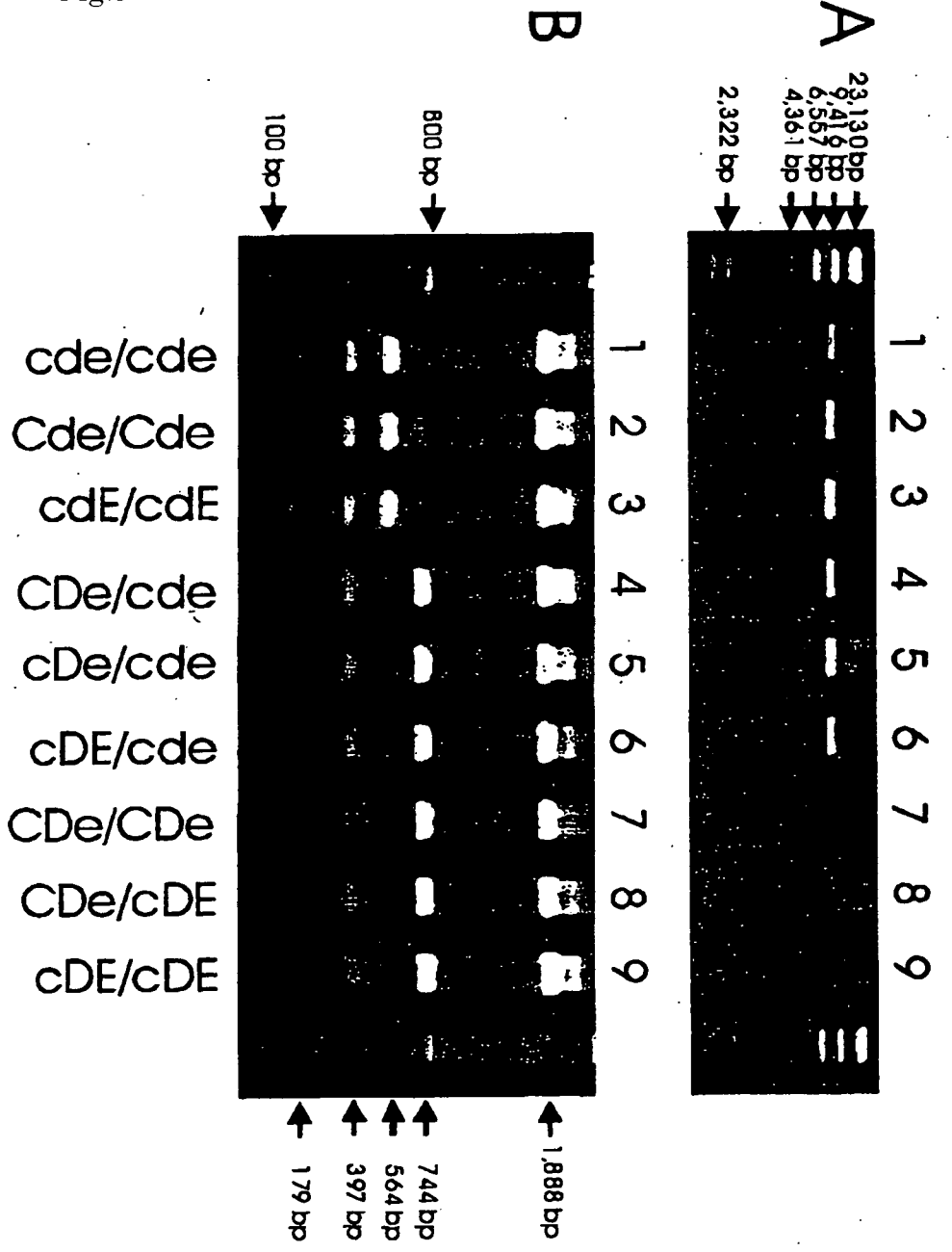
```
gataaacctcctgagtggtaagaacttgtccatttaaaaacaggctatag
**************************g************************
**************************g************************

attgtatcatgcagttttatctactaatcggctaa
******--******************************
******--******************************

tatcccgccaaaaac----aaaaaaccccaaa
*gca**********aaac******a******
*gca**********aaac******a******
 <--------rnb31-----------mm

gggatgaaagtttcatccatcaaaggaaacaac
*********************************
*********************************
```

Fig.5


A
23,130 bp
9,416 bp
6,557 bp
4,361 bp
2,322 bp
1 2 3 4 5 6 7 8 9
B
100 bp
800 bp
cde/cde
Cde/Cde
cdE/cdE
CDe/cde
cDe/cde
cDE/cde
CDe/CDe
CDe/cDE
cDE/cDE
1 2 3 4 5 6 7 8 9
179 bp
397 bp
564 bp
744 bp
1,888 bp

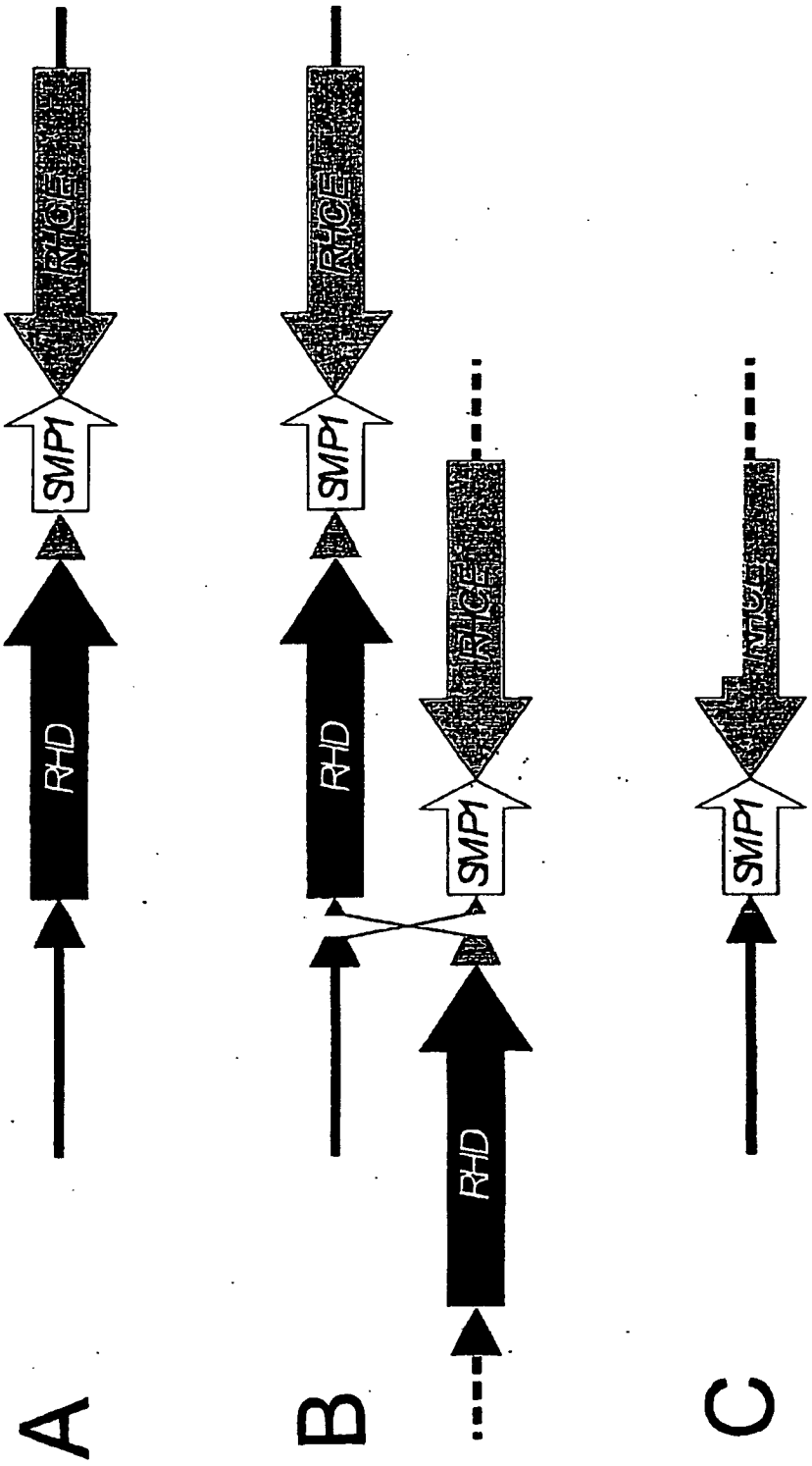


Fig.6

Fig.7

Hybrid Rhesus box of RHD negatives

5' ctagaaaacactttgtcattttagaggtgtta
(start of Rhesus box)
tccaatgttcgcgcaggcactggagtcagagaaaatggagttgaatcctttctctgccactc
tttgaggagaatctcaccatttattatgcactgtagaatacaacaataaaatacagccatgt
accacataacaacatcttggtaaacaacagactgcatatatgatggtggtcatccagtaagc
taaggttaatttattattattccctttttttttttctttttttttgagatgtagtcttactctg
tcacccaggctagagtgcaatggcaccatcttggctcactgcaacctctgcctcctgggttc
aagcgaatctcctgcctcagcctccgaagtagctgggaattacaggcacccaccacatctgg
ctaatttttttgtatttttagtaaagatggggtttcaccatgttggccaggctgatctcaaac
tcctgacctcaagtgatctgcctgcctcggcctcccaaagtgctgggaccataggcctgagc
cactgtgcccggccttgtttgctttttttaacagttaacagtgtgctcatagaaactgctttg
acatgactgcaatcatgtgcttcatagaaacttaattagattataccactagagtcttcaga
ttttttatactttttttttttgaaacggagtctcactctgtcaccaggctggagtgcagtgccg
caatctcagctcgccgcaacctccgcctcccaggttcaagtgattctcctgcctcagcctcc
cgagtagctgggattacaagtgcacactaccacgcccagctaatttttgcatttttactaga
cagggtttcaccatgttggctaggatagtttcaccaggatctcttggcctcatgatcagcct
gcctcggcctcccaaagtgctgggattacaggtgtgagccaccgtgcccagcctatacttcc
ctttttgaataccatttggcgttttgaagaattaacagctttgtgaacgtggcagtgcttgt
gattcaggcttccactgagaccaaggggagaacctggttgcaggacaaacagacggacagcg
tgtggcagtgtttaaatgctcttctgaaggctgatacgacagctctctgtgcactgattgta
tacgcatcccaagattatattattgttttctattgctatgtgtcacactttgccaaacagga
tgtggaaaatgaataagcggttttcttaggcacttcttaacagacaattggtcaaaatgaac
tccattgcttaagaaacacataaacaccatttagtcactgaatatagctatatgtatggttg
ctactatggggaatcttgttttgccaatttttctttgaaaattctggcagaccaaggttcttt
ttgtttacacaatacttgaaaaataaaaatgaacaagccaacaaactaccaagttttcactt
acataaatgtagttacatacagaaaatgtgactgtgaatttttttctaggacttttaaactat
aagcactatttgcacgaaagagaaccaatctatcaattacaaactcacataatttttacagat
ttttttttccctacacagcacataaaacagaaggaatttgaagccaccctccaaacacaggg
gaaggaggctgtgtgtatatcctcattgtctttcacattctaaggtggttccactcagtgac
tgaaatccttaagtgttgtattagtcggcttgggctaccataacagcagcttaaactgttgt
taagccactcagacttaaacaacagaaatttatttccttatagttctggaggctggaagttc
aaggtgccggcaaggctggtttctggtgagacctctctccctgtcttgcagatggctgcctc
ctccctgtgtcctcatagagcctgtcctctgcttttacacttctggtgtcatcttccttttt
ttttttttttgagacagagtctcgctctatcgcccaggctggagtgcagtggcccgatcga
tctcggctcactgcaacctctgcctcccaggttcaagcaattctcctgcctcagcctcccga
gtagctgggactacaggtgcccgccatcatgtctggctaatttttgtatttttagtagagac
agggtttcaccatattggccaggctggtctccaactcctgaccttgtcatctgcctgcctcg
gcctcccaaagtgctaggattacaggcgtgagccaccgcacccggcctctttctcttcttat
aaggacaccagtcctattagattagggctccaccctcatgacctcatttgaccttaactatt
atttctttaaagcacctatttccaaatatagtcactttaggggttagggcttcaaaatatga
atctgagggagatcaattcagtaaatagcagtagtcattaacggacaatatatacaaagata
atttcgtgattactgtccttatgcataaatgtcctcagtgttccactgcctttatccagatt
tactatcacaaagactttgctctgagaaaaatgtgatttctttctttttttttttttttga
gacagagtctcactctgtcacccaggctggagtgcagtggtgcaatctcggctcactgcaat
ctccgcctcccaggttcacgccattctcttgcctcagtctcccgagtagctgggcctacagg

Fig.7 **cont.**

Hybrid Rhesus box of RHD negatives

cgcccgccaccctgcccagctaatttttgtatttttagtagagacggggtttcaccatgtt
agccaggatggtctcaatctcctgacctcgtgatccacctgcctcagcctcccaaagtgctg
ggattacaggcatgagccaccgcgcccagcagatttttttttttttttttttttgagat
ggagtcttgctgtgttgcccagcctggagtgcagtgttatgattttggctcactgcaacctc
tgtctaccatgttcaagcgattctcccacctctgcctcccgtgtagctgggatcacaggcac
acgccaccacacctagctacttttgtatttttagtagaaatggggtttcaccatgttggcc
aggatggtcccgaactcctgacctcaagtgatcctcctgcctcggcctcccaaagtgctggg
attacaggtgtgagccactgtgcctggccaaaaatgtgatttcttatttcccacattgccaa
ttccatttcaattaactataatagctatgtctattgagcactcaagtgtattctagaaactg
ttcctgattctggggatatatccatgaatcaactatagtccctgttattaagtaatctgtag
tctgactaaaccattagaaatttaaaaaatggctactttcaaagacatcttggagttcagga
gtcccacactgcgaaccatattacctaataatccaacctgcttgtaattcacttatttaacc
aatatttattgagtgccaactttgagcctaagatacagcagtaaacaaatggataaagtccc
tgtcctcatgaaacttgtattctaatggaagaaacagaaaacaaacagatataggatgtaat
atcaggtagggataaatactttgaattcaaacaaaagtatacgtagtcagggttcgccaaag
agacacagccaatcggatacatagatatataaaagagggtttatgagttagaaagggctcac
atgattacagaggctgagaagtcccacagcagattgtctgcaagctggagacccagggatac
tggtagcatggctcagtccaagtcccaaagcctcagaatcaggaaagctgatgatataattc
ttagcccaaaggccttagaaccccagcggtgacggaaaggctgatgtaggtcctggagtcct
gagacccaacagcctgggatcctgaaatccaagggcaggaatggaagcgtgtattccagctc
caagagagtaagaccaatttgcctttcttccgtttttgtttcaagccacctgcacattgagg
gcggatggttccctcttagtccattcagtcatatatcaatctcttctggaaataccctcaca
gacacactaacaaataatgcctttccagttctctaggtattctttaatccagtcaagctgac
acctaaaattaaccatcacaaaagttaaggagaaagaagacaacttgtaggggaggctgcta
tgcaagacagtgtgtgaaggaagggctctctgaagaggttaatatctgagcagagacttgaa
tgaagtgaagaagtgagccatgtgggtatggggaatacaacttccaggtagagaagacaagt
gtggtgtgtatcagggtcagcaaagaagccatgtgacagagaagggtgggccagggagagac
ggataagtgatctaactcctgaggaggtggcctggccaggagctagagcatgaagatctcgt
aggactttattctgcaaggtgaaaagccattgtattagtctgttcacaaacccgagactagg
caatttacaaaagaaagagaggtttaatggacttacagttccacatggctggggaggcctca
caatcatggcgaaaggcaatgaggagcaagtcacgtcttacgtggatggcaggcaaagacaa
agacagcttgtgcagagaaactccccttatagagccatcagatcctgttagacttattcac
tatcacaagaacagcacgggtaagacctgtccccatgattcagttacctcccactgggtccc
tcccacaacgcatgggaattcaggatgagatttgggtggggacacaaccaaaccctatcatt
ccacccatggcccctcccaaatttcatgtcctcacatttcaaaaccaatcacaccatcccaa
cagtccctcaaagtcttaaatgatttcagcattaactcaaaagtccacagtctaatgtctca
tctgagacaaggcaagtcctttccatttatgagcctataaaatccaaagcaagttagttact
tcctagatacaatgggggtacaggcattgggtaaatacagccattccaaatgggataaattg
gtcaaaacaaagaggctacaggcccatgagagtccaaaatccagtggggcagtcaaatctta
aagctccaaaatgatctcctttgactccacatctcacatccaggtcacgcagatggaagggg
tgggttcccatggtcttgggcagctctgcccctgtacctttgcagggtacagcctccctctc
agctgctttcatgggctggcattgagtgtctgcagcttttccaggtacacggtgcaagctgt
cggtggatctaccattctggggtctggaggacctcttctcacagctccactaggtggtgccc
cagtagggactgtgtgtggggtctctgaccccacatttcccttctgcactgccctggcagag
gatctccatgagggccctgctcctgcagcaaacttctgactgggcatccaggcatttccgca

Fig.7 **cont.**

Hybrid Rhesus box of RHD negatives

catcctctttaatctaggcgaaggtttccaaaccccaattcttgacttctgtgcactcgcag
tctcaacaccacatggaagctgtcaaggcttggggcttgcactccccgaagctacagcccaa
gctctaccttgcctcccgtcagtcatggttgggagtggctgggatgcagggcaccaagtccc
taggctgcacacagcatgaggaccccgggcctggccaacaaaaccatttttcctgatacct
ctggacctgtgatgggaggggttgccataaagacctctgacatgccctggagacatttccc
cattgtcttgggaattagcatttggctcctgttactcatgcaaatttctgcagccagcttga
atttctcctcagaaaatgggaatttttctttctatcacattgtcaggctgcaaattttccg
aacttttatgctctgcttcccttataaaactgaatgtctttaacagcacccaagtcacctct
tgaatgctttgctgcttagaaatttctcctgccagatactctaaatcatctctctgaagttc
aaagttctacaaatatctcgtgcaggggcaaaatgccgccagtatctttgctaaaacataac
aagagtccccttgctccagttcccaacaagttcctcatttccgtctgagaccacctcagcc
tatggactttattgtccacagtgctatcagcattttgggcaaagccattcaacaagtctcta
ggaagttccaaactttcccacatttgcctgtcttcttctgagccctccaaactgttccaaac
cctgcctgttacccagttccaaagtcacatacccatttttgagtatctacggcagcacccca
ctctactggtaccaatttagccactgaagtagttggagaacagaagtaatagactctggttt
acattgtaaaagcttctctgtggctgctgtgtgaagaaaatatatgagaatgaagccccaag
atgaagcagggacacagttgcagtggttagagtaagaaatgctgctggctggcactgaagtg
atagcctggaggtttgtgtgtgcacatgcatgtgtatgtgttttacgatagtaggcccaaca
gatactgtaatccacacttgttttttttttttttgagacagagtctcacctgttgcctag
actagaatgcagtggcacaatcttggctcactacaacctccacctcccaggttcaaacaatc
cttgtgcttcagcctcccgagtagttgggattacaggtgtgtgccaccgtgcccagctatat
tttttgtatttttagcagagatgggattttgccacattggccaggctggtcttgaactcctg
gcctcaagcaatcctcccaccttagcctcccaaagtgctgagccaccacacctggccgcaac
tgatttttaatcatgaaatgacacatacatttaaaaaacccaatacctataatattcctggc
tagtactcttcacatctatatcatcaaaaacaaagaaagtatgtgaaactgacacagccaag
gggagactaaggagacataacaattaactgtaatgtggtattctggaggggatcctggaaca
gaaaaagacattaggcaaaaaactaaagaaatctgaataaaatgtggatgtcagttaataat
aatgtatcatattagtccagtaattgtaacaaatatacccaataatgaaagccattaattat
agggaaaatggaggggttaatatgggtggctggcttttgctatttctagcagctccatttta
tctacaaaagacaaacattcattaagtcccaaaaaggtaaagaatgacaaattaagcatgta
tcttattagtaagagtaatataaagatgctcactcatatttataaatatttgacaatgatgt
taaggccagaaaagagaaaaaagggtaggggcaaaaaacgcaaagagaaaggagttagtatc
ttttctcccgcactcattagctattaaaagaggatgtttgtttaaagctgctcagagctggt
aaactaatgttaagtcactaacgggaatttaaaaggtttcattaagaactgcctgcactaga
ttcctccaccctgagacattaaacaatcacgataaacctcctgagtggtaagaacgtgtcca
tttaaaaacaggctatagattgtcatgcagttttatctactaatcggctaatgcaccgccaa
aaacaaacaaaaaaacccaaagggatgaaagtttcatccatcaaaggaaacaacagtcacct
tggttcccatcccactcatatactgccgccgtacatgtcaatcagatgaacctgtgcgtatc
tcttaacgacaattgacccacctttttaactgaagtgaaggggggttctgctccgcgaccac
ttcctggatctctcccttcaccctctgtgttctttcgggtgcaccatcgggtcaaagccgca
gcaacgccgtctctgtgtgatcgcatgtgcccttctgcacacgaccttcccccgagagtgac
cagctaccggacaggcaccaaggagggctaccgagcacctcccggaccggcggctgcaggat
cgcgagcgcctccgctagggagaccgcacgttgcgcctgtgcttcctgcggtggcgccttct
gcaaggagacctcgaccctgctccctctccggggctggatctgactccttgacggtgattcc
agacgcgagacccaaactgacggcttctagaagaggggcgagcccggccgcaagtctttcac

Fig.7 **cont.**

Hybrid Rhesus box of RHD negatives

gtagctaagtcatcgttgcttccggcttcttaccgttctccccctttgtaaacggttacctcc
cgaaaacccaggctctcctccaacagtggttctcaagcgaggcgatcttccccgggagggga
tatttggcaaagtctgggggcatttttggttcactggggctgctacttgcatccactgggta
gaggcgggggatgcagctacacaacctgcgaagcacgggacagcaccctccccaacccagac
agaattagccggcccaaaacctcagtagtgcccaggctgagaaaccctgccttaaacaaaca
acaaagaaaggccaagtcccataagtgggtcaccgcgccgagactggggtccacgggacacc
ccagccacgccaagccgggaagtccccgcctcctggagctgaacccgcccctctcccagagg
tggagctgcgggggggcgggaacaggcacggagaaaataaacaagactaaaaagtcctgagta
gcgctgtgtggccgcaaacctgaacccacctttgcaccacgcgggacccggcactcttcct
gccacccacccctgagagggctgcgcggccgaccccagtactagaaaacactcgtcacctca
ctcaagacgggtacgaaggccaacggacgccttcctttagaacgctcagcacacagagcaac
ttctcacgcctactctcaaatggcgtactccaaactagcactcccgacgtccagctgtgaac
ccagagcggcggaaagcccctgaacccagcgcccgggcatgcgcagacgcgttgttgtggtg
ggcgtggctccctccggacccggcgccccgccctccgccccgtgtccgcatgcgcgactgag
ccgggtggatggtactgctgcatccgggtgtctg
(end of Rhesus box)
gaggctgtggccgttttgttttcttggctaaaatcgggggagtgaggcgggccggcgcggc
3'

Fig.8

Upstream Rhesus box of D-positives

5' ctagaaaacactttgtcattttagaggtgtta
(start of Rhesus box)
tccaatgttcgcgcaggcactggagtcagagaaaatggagttgaatcctttctctgccactc
tttgaggagaatctcaccatttattatgcactgtagaatacaacaataaaatacagccatgt
accacataacaacatcttggtaaacaacagactgcatatatgatggtggtcatccagtaagc
taaggttaatttattattattccctttttttttttctttttttttgagatgtagtcttactctg
tcacccaggctagagtgcaatggcaccatcttggctcactgcaacctctgcctcctgggttc
aagcgaatctcctgcctcagcctccgaagtagctgggaattacaggcacccaccacatctgg
ctaatttttttgtatttttagtaaagatggggtttcaccatgttggccaggctgatctcaaac
tcctgacctcaagtgatctgcctgcctcggcctcccaaagtgctgggaccataggcctgagc
cactgtgcccggccttgtttgctttttttaacagttaacagtgtgctcatagaaactgctttg
acatgactgcaatcatgtgcttcatagaaacttaattagattataccactagagtcttcaga
ttttttatactttttttttttgaaacggagtctcactctgtcaccaggctggagtgcagtgccg
caatctcagctcgccgcaacctccgcctcccaggttcaagtgattctcctgcctcagcctcc
cgagtagctgggattacaagtgcacactaccacgcccagctaatttttgcatttttactaga
cagggtttcaccatgttggctaggatagtttcaccaggatctcttggcctcatgatcagcct
gcctcggcctcccaaagtgctgggattacaggtgtgagccaccgtgcccagcctatacttcc
ctttttgaataccatttggcgttttgaagaattaacagctttgtgaacgtggcagtgcttgt
gattcaggcttccactgagaccaaggggagaacctggttgcaggacaaacagacggacagcg
tgtggcagtgtttaaatgctcttctgaaggctgatacgacagctctctgtgcactgattgca
tacgcatcccaagattatattattgttttctattgctatgtgtcacactttgccaaacagga
tgtggaaaatgaataagcggttttcttaggcacttcttaacagacaattggtcaaaatgaac
tccattgcttaagaaacacataaacaccatttagtcactgaatatagctatatgtatggttg
ctactatggggaatcttgttttgccaatttttctttgaaaattctggcagaccaaggttcttt
ttgtttacacaatacttgaaaaataaaaatgaacaagccaacaaactaccaagttttcactt
acataaatgtagttacatacagaaaatgtgactgtgaatttttctaggacttttaaactat
aagcactatttgcacgaaagagaaccaatctatcaattacaaactcacataattttacagat
ttttttttccctacacagcacataaaacagaaggaatttgaagccaccctccaaacacaggg
gaaggaggctgtgtgtatatcctcattgtctttcacattctaaggtggttccactcagtgac
tgaaatccttaagtgttgtattagtcggcttgggctaccataacagcagcttaaactgttgt
taagccactcagacttaaacaacagaaatttatttccttatagttctggaggctggaagttc
aaggtgccggcaaggctggtttctggtgagacctctctccctgtcttgcagatggctgcctc
ctccctgtgtcctcatagagcctgtcctctgcttttacacttctggtgtcatcttcctttt
ttttttttttgagacagagtctcgctctatcgcccaggctggagtgcagtggcccgatcga
tctcggctcactgcaacctctgcctcccaggttcaagcaattctcctgcctcagcctcccga
gtagctgggactacaggtgcccgccatcatgtctggctaatttttgtatttttagtagagac
agggtttcaccatattggccaggctggtctccaactcctgaccttgtcatctgcctgcctcg
gcctcccaaagtgctaggattacaggcgtgagccaccgcacccggcctctttctcttcttat
aaggacaccagtcctattagattagggctccaccctcatgacctcatttgaccttaactatt
atttctttaaagcacctatttccaaatatagtcactttaggggttagggcttcaaaatatga
atctgagggagatcaattcagtaaatagcagtagtcattaacggacaatatatacaaagata
atttcgtgattactgtccttatgcataaatgtcctcagtgttccactgcctttatccagatt
tactatcacaaagactttgctctgagaaaaatgtgatttctttctttttttttttttttga
gacagagtctcactctgtcacccaggctggagtgcagtggtgcaatctcggctcactgcaat
ctccgcctcccaggttcacgccattctcttgcctcagtctcccgagtagctgggcctacagg

Fig.8 **cont.**

Upstream Rhesus box of D-positives

cgcccgccaccctgcccagctaatttttgtatttttagtagagacggggtttcaccatgtt
agccaggatggtctcaatctcctgacctcgtgatccacctgcctcagcctcccaaagtgctg
ggattacaggcatgagccaccgcgcccagcagatttttttttttttttttttttttgagat
ggagtcttgctgtgttgcccagcctggagtgcagtgttatgattttggctcactgcaacctc
tgtctaccatgttcaagcgattctcccacctctgcctcccgtgtagctgggatcacaggcac
acgccaccacacctagctacttttttgtatttttagtagaaatgggtttcaccatgttggcc
aggatggtcccgaactcctgacctcaagtgatcctcctgcctcggcctcccaaagtgctggg
attacaggtgtgagccactgtgcctggccaaaaatgtgatttcttatttcccacattgccaa
ttccatttcaattaactataatagctatgtctattgagcactcaagtgtattctagaaactg
ttcctgattctggggatatatccatgaatcaactatagtccctgttattaagtaatctgtag
tctgactaaaccattagaaatttaaaaaatggctactttcaaagacatcttggagttcagga
gtcccacactgcgaaccatattacctaataatccaacctgcttgtaattcacttatttaacc
aatatttattgagtgccaactttgagcctaagatacagcagtaaacaaatggataaagtccc
tgtcctcatgaaacttgtattctaatggaagaaacagaaaacaaacagatataggatgtaat
atcaggtagggataaatactttgaattcaaacaaaagtatacgtagtcagggttcgccaaag
agacacagccaatcggatacatagatatataaaagagggtttatgagttagaaagggctcac
atgattacagaggctgagaagtcccacagcagattgtctgcaagctggagacccagggatac
tggtagcatggctcagtccaagtcccaaagcctcagaatcaggaaagctgatgatataattc
ttagcccaaaggccttagaaccccagcggtgacggaaaggctgatgtaggtcctggagtcct
gagacccaacagcctgggatcctgaaatccaagggcaggaatggaagcgtgtattccagctc
caagagagtaagaccaatttgcctttcttccgttttgtttcaagccacctgcacattgagg
gcggatggttccctcttagtccattcagtcatatatcaatctcttctggaaatacctcaca
gacacactaacaaataatgcctttccagttctctaggtattctttaatccagtcaagctgac
acctaaaattaaccatcacaaaagttaaggagaaagaagacaacttgtaggggaggctgcta
tgcaagacagtgtgtgaaggaagggctctctgaagaggttaatatctgagcagagacttgaa
tgaagtgaagaagtgagccatgtgggtatggggaatacaacttccaggtagagaagacaagt
gtggtgtgtatcagggtcagcaaagaagccatgtgacagagaagggtgggccagggagagac
ggataagtgatctaactcctgaggaggtggcctggccaggagctagagcatgaagatctcgt
aggactttattctgcaaggtgaaaagccattgtattagtctgttcacaaacccgagactagg
caatttacaaaagaaagagaggtttaatggacttacagttccacatggctggggaggcctca
caatcatggcgaaaggcaatgaggagcaagtcacgtcttacgtggatggcaggcaaagacaa
agacagcttgtgcagagaaactcccccttatagagccatcagatcctgttagacttattcac
tatcacaagaacagcacgggtaagacctgtccccatgattcagttacctcccactgggtccc
tcccacaacgcatgggaattcaggatgagatttgggtggggacacaaccaaaccctatcatt
ccacccatggcccctcccaaatttcatgtcctcacatttcaaaaccaatcacaccatcccaa
cagtccctcaaagtcttaaatgatttcagcattaactcaaaagtccacagtctaatgtctca
tctgagacaaggcaagtcctttccatttatgagcctataaaatccaaagcaagttagttact
tcctagatacaatgggggtacaggcattgggtaaatacagccattccaaatgggataaattg
gtcaaaacaaagaggctacaggcccatgagagtccaaaatccagtggggcagtcaaatctta
aagctccaaaatgatctcctttgactcaacatctcacatccaggtcacgcagatggaagggg
tgggttcccatggtcttgggcagctctgcccctgtacctttgcagggtacagcctcctctc
agctgctttcatgggctggcattgagtgtctgcagcttttccaggtacacggtgcaagctgt
cggtggatctaccattctggggtctggaggacctcttctcacagctccactaggtggtgccc
cagtagggactgtgtgtggggtctctgaccccacatttcccttctgcactgccctggcagag
gatctccatgagggccctgctcctgcagcaaacttctgactgggcatccaggcatttccgca

Fig.8 **cont.**

Upstream Rhesus box of D-positives

catcctctttaatctaggcgaaggtttccaaaccccaattcttgacttctgtgcactcgcag
tctcaacaccacatggaagctgtcaaggcttggggcttgcactccccgaagctacagcccaa
gctctaccttgcctcccgtcagtcatggttgggagtggctgggatgcagggcaccaagtccc
taggctgcacacagcatgaggaccccgggcctggccaacaaaaccatttttcctgatacct
ctggacctgtgatgggaggggttgccataaagacctctgacatgccctggagacattttccc
cattgtcttgggaattagcatttggctcctgttactcatgcaaatttctgcagccagcttga
atttctcctcagaaaatgggaatttttctttctatcacattgtcaggctgcaaattttccg
aacttttatgctctgcttcccttataaaactgaatgtctttaacagcacccaagtcacctct
tgaatgctttgctgcttagaaatttctcctgccagatactctaaatcatctctctgaagttc
aaagttctacaaatatctcgtgcaggggcaaaatgccgccagtatctttgctaaaacataac
aagagtcccctttgctccagttcccaacaagttcctcatttccgtctgagaccacctcagcc
tatggactttattgtccacagtgctatcagcattttgggcaaagccattcaacaagtctcta
ggaagttccaaactttcccacatttgcctgtcttcttctgagccctccaaactgttccaaac
cctgcctgttacccagttccaaagtcacatacccatttttgagtatctacggcagcacccca
ctctactggtaccaatttagccactgaagtagttggagaacagaagtaatagactctggttt
acattgtaaaagcttctctgtggctgctgtgtgaagaaaatatatgagaatgaagccccaag
atgaagcagggacacagttgcagtggttagagtaagaaatgctgctggctggcactgaagtg
atagcctggaggtttgtgtgtgcacatgcatgtgtatgtgttttacgatagtaggcccaaca
gatactgtaatccacacttgtttttttttttgagacagagtctcacctgttgcctagacta
gaatgcagtggcacaatcttggctcactacaacctccacctcccaggttcaaacaatccttg
tgcttcagcctcccgagtagttgggattacaggtgtgtgccaccgtgcccagctatattttt
tgtatttttagcagagatgggattttgccacattggccaggctggtcttgaactcctggcct
caagcaatcctcccaccttagcctcccaaagtgctgagccaccacacctggccgcaactgat
ttttaatcatgaaatgacacatacatttaaaaaacccaatacctataatattcctggctagt
actcttcacatctatatcatcaaaaacaaagaaagtatgtgaaactgacacagccaagggga
gactaaggagacataacaattaactgtaatgtggtattctggaggggatcctggaacagaaa
aagacattaggcaaaaaactaaagaaatctgaataaaatgtggatgtcagttaataataatg
tatcatattagtccagtaattgtaacaaatataccacaataatgaaagccattaattatagg
gaaaatggaggggttaatatgggtggctggcttttgctatttctagcagctccatttttatct
gcaaaagacaaacattcattaagtcccaaaaaggtaaagaatgacaaattaagcatgtatct
tattagtaagagtaatataaagatgctcactcctatttataaatatttgacaatcatgttaa
ggccacaaaagagaaaaaagggtaggggcaaaaaacgcaaagagaaaggagttagtatcttt
tctcccgcactcattagctattaaaagaggatgtttgtttaaagctgctcagagctggtaaa
ctaatgttaagtcactaacgggaatttaaaaggtttcattaagaactgcctgcactagattc
ctccaccctgagacattaaacaatcacgataaacctcctgagtggtaagaacttgtccattt
aaaaacaggctatagattgtatcatgcagttttatctactaatcggctaatatcccgccaaa
aacaaaaaaccccaaagggatgaaagtttcatccatcaaaggaaacaacagtcaccttggtt
cccatctcactcatatactgccgccgtacatgtcaatcagatgaacctgtgcgtatctctta
atgacaattgacccacatttttgactgaagtgaaaggggttctgctccgcgaccacttcct
ggatctccccctccaccctctgtgttcttcgggtgcaccatcgggtcaaagccgcagcaac
gccgtctctgtgtgatcgcatgtgcccttctgcacacgaccttcccccgagagtgaccagct
accggacaggcaccaaggagggctaccgagcacctcccggaccggcggctgcaggatcgcga
gcgcctccgctagggagactgcacgttgcgcctgtgcttcctgcggtggcgccttctgcaag
gagacctcgaccctgctccctctccggggctggatctgactccttgacggtgattccagacg
cgagacccaaactgacggcttctagaagaggggcgagcccggccgcaagtctttcacgtagc

Fig.8 **cont.**

Upstream Rhesus box of D-positives

taagtcatcgttgcttccggcttcttaccgttctccccttttgtaaacggttacctcccgaaa
acccaggctctcctccaacagtggttctcaagcgaggcgatcttccccgggaggggatattt
ggcaaagtctgggggcatttttggttcactggggctgctacttgcatccactgggtagaggc
gggggatgcagctacacaacctgcgaagcacgggacagcaccctccccaacccagacagaat
tagccggcccaaaacctcagtagtgcccaggctgagaaaccctgccttaaacaaacaacaaa
gaaaagccaagtcccataagtgggtcaccgcgccgagactggggtccacgggacaccccagc
cacgccaagccgggaagtccccgcctcctggagctgaacccgcccctctcccagaggtggag
ctgcgggggggcgggaacaggcacggagaaaataaacaagactaaaaagtcctgagtagcgct
gtgtggccgcaaacctgaacccacctttgcaccacgcgggacccggcacgcttcctgccac
ccacccctgagagggctgcgcggccgaccccagtactagaaaacactcgtcacctcaatcaa
gacgggtacgaaggccaacggacgccttcctttagaacgctcagcacacagagcaacttctc
acgcctactctcaaatggcgtactccaaactagcactcccgacgtccagctgtgaacccaga
gcggcggaaagcccctgaacccagcgcccgggcatgcgcagacgcgttgttgtggtgggcgt
ggctccctccggacccggcgccccgccctccgccccgtgtccgcatgcgcgactgagccgcg
ggggtggtactgctgcatccgggtgtctg
(end of Rhesus box)
aagatccgatgaaataacatatgcaaaatgattgggtccgtgattggcattccagaaatgg
3'

Fig.9

Upstream Rhesus box of D-positives

5' ctagaaaacactttgtcattttagaggtgtta
(start of Rhesus box)
tccaatgttcgcgcaggcactggagtcagagaaaatggagttgaatcctttctctgccactc
tttgaggagaatctcaccatttattatgcactgtagaatacaacaataaaatacagccatgt
accacataacaacatcttggtaaacaacagactgcatatatgatggtggtcatccagtaagc
taaggttaatttattattattccctttttttttttctttttttgagatgtagtcttactctg
tcacccaggctagagtgcaatggcaccatcttggctcactgcaacctctgcctcctgggttc
aagcgaatctcctgcctcagcctccgaagtagctgggaattacaggcacccaccacatctgg
ctaatttttgtatttttagtaaagatggggtttcaccatgttggccaggctgatctcaaac
tcctgacctcaagtgatctgcctgcctcggcctcccaaagtgctgggaccataggcctgagc
cactgtgcccggccttgtttgctttttaacagttaacagtgtgctcatagaaactgctttg
acatgactgcaatcatgtgcttcatagaaacttaattagattataccactagagtcttcaga
ttttatactttttttttgaaacggagtctcactctgtcaccaggctggagtgcagtgccg
caatctcagctcgccgcaacctccgcctcccaggttcaagtgattctcctgcctcagcctcc
cgagtagctgggattacaagtgcacactaccacgcccagctaatttttgcatttttactaga
cagggtttcaccatgttggctaggatagtttcaccaggatctcttggcctcatgatcagcct
gcctcggcctcccaaagtgctgggattacaggtgtgagccaccgtgcccagcctatacttcc
ctttttgaataccatttggcgttttgaagaattaacagctttgtgaacgtggcagtgcttgt
gattcaggcttccactgagaccaaggggagaacctggttgcaggacaaacagacggacagcg
tgtggcagtgtttaaatgctcttctgaaggctgatacgacagctctctgtgcactgattgca
tacgcatcccaagattatattattgttttctattgctatgtgtcacactttgccaaacagga
tgtggaaaatgaataagcggttttcttaggcacttcttaacagacaattggtcaaaatgaac
tccattgcttaagaaacacataaacaccatttagtcactgaatatagctatatgtatggttg
ctactatggggaatcttgttttgccaattttctttgaaaattctggcagaccaaggttcttt
ttgtttacacaatacttgaaaaataaaaatgaacaagccaacaaactaccaagttttcactt
acataaatgtagttacatacagaaaatgtgactgtgaatttttctaggacttttaaactat
aagcactatttgcacgaaagagaaccaatctatcaattacaaactcacataatttacagat
tttttttccctacacagcacataaaacagaaggaatttgaagccaccctccaaacacaggg
gaaggaggctgtgtgtatatcctcattgtctttcacattctaaggtggttccactcagtgac
tgaaatccttaagtgttgtattagtcggcttgggctaccataacagcagcttaaactgttgt
taagccactcagacttaaacaacagaaatttatttccttatagttctggaggctggaagttc
aaggtgccggcaaggctggtttctggtgagacctctctccctgtcttgcagatggctgcctc
ctccctgtgtcctcatagagcctgtcctctgcttttacacttctggtgtcatcttccttttt
ttttttttttgagacagagtctcgctctatcgcccaggctggagtgcagtggcccgatcga
tctcggctcactgcaacctctgcctcccaggttcaagcaattctcctgcctcagcctcccga
gtagctgggactacaggtgcccgccatcatgtctggctaatttttgtatttttagtagagac
agggtttcaccatattggccaggctggtctccaactcctgaccttgtcatctgcctgcctcg
gcctcccaaagtgctaggattacaggcgtgagccaccgcacccggcctctttctcttcttat
aaggacaccagtcctattagattagggctccaccctcatgacctcatttgaccttaactatt
atttctttaaagcacctatttccaaatatagtcactttaggggttagggcttcaaaatatga
atctgagggagatcaattcagtaaatagcagtagtcattaacggacaatatatacaaagata
atttcgtgattactgtccttatgcataaatgtcctcagtgttccactgcctttatccagatt
tactatcacaaagactttgctctgagaaaaatgtgatttctttctttttttttttttttga
gacagagtctcactctgtcacccaggctggagtgcagtggtgcaatctcggctcactgcaat
ctccgcctcccaggttcacgccattctcttgcctcagtctcccgagtagctgggcctacagg

Fig.9 **cont.**

Upstream Rhesus box of D-positives

cgcccgccaccctgcccagctaatttttttgtatttttagtagagacggggtttcaccatgtt
agccaggatggtctcaatctcctgacctcgtgatccacctgcctcagcctcccaaagtgctg
ggattacaggcatgagccaccgcgcccagcagatttttttttttttttttttttttgagat
ggagtcttgctgtgttgcccagcctggagtgcagtgttatgattttggctcactgcaacctc
tgtctaccatgttcaagcgattctcccacctctgcctcccgtgtagctgggatcacaggcac
acgccaccacacctagctacttttttgtatttttagtagaaatggggtttcaccatgttggcc
aggatggtcccgaactcctgacctcaagtgatcctcctgcctcggcctcccaaagtgctggg
attacaggtgtgagccactgtgcctggccaaaaatgtgatttcttatttcccacattgccaa
ttccatttcaattaactataatagctatgtctattgagcactcaagtgtattctagaaactg
ttcctgattctggggatatatccatgaatcaactatagtccctgttattaagtaatctgtag
tctgactaaaccattagaaatttaaaaaatggctactttcaaagacatcttggagttcagga
gtcccacactgcgaaccatattacctaataatccaacctgcttgtaattcacttatttaacc
aatatttattgagtgccaactttgagcctaagatacagcagtaaacaaatggataaagtccc
tgtcctcatgaaacttgtattctaatggaagaaacagaaaacaaacagatataggatgtaat
atcaggtagggataaatactttgaattcaaacaaaagtatacgtagtcagggttcgccaaag
agacacagccaatcggatacatagatatataaaagagggtttatgagttagaaagggctcac
atgattacagaggctgagaagtcccacagcagattgtctgcaagctggagacccagggatac
tggtagcatggctcagtccaagtcccaaagcctcagaatcaggaaagctgatgatataattc
ttagcccaaaggccttagaaccccagcggtgacggaaaggctgatgtaggtcctggagtcct
gagacccaacagcctgggatcctgaaatccaagggcaggaatggaagcgtgtattccagctc
caagagagtaagaccaatttgcctttcttccgtttttgtttcaagccacctgcacattgagg
gcggatggttccctcttagtccattcagtcatatatcaatctcttctggaaataccctcaca
gacacactaacaaataatgcctttccagttctctaggtattctttaatccagtcaagctgac
acctaaaattaaccatcacaaaagttaaggagaaagaagacaacttgtaggggaggctgcta
tgcaagacagtgtgtgaaggaagggctctctgaagaggttaatatctgagcagagacttgaa
tgaagtgaagaagtgagccatgtgggtatggggaatacaacttccaggtagagaagacaagt
gtggtgtgtatcagggtcagcaaagaagccatgtgacagagaagggtgggccagggagagac
ggataagtgatctaactcctgaggaggtggcctggccaggagctagagcatgaagatctcgt
aggactttattctgcaaggtgaaaagccattgtattagtctgttcacaaacccgagactagg
caatttacaaaagaaagagaggtttaatggacttacagttccacatggctggggaggcctca
caatcatggcgaaaggcaatgaggagcaagtcacgtcttacgtggatggcaggcaaagacaa
agacagcttgtgcagagaaactccccttatagagccatcagatcctgttagacttattcac
tatcacaagaacagcacgggtaagacctgtccccatgattcagttacctcccactgggtccc
tcccacaacgcatgggaattcaggatgagatttgggtggggacacaaccaaaccctatcatt
ccacccatggcccctcccaaatttcatgtcctcacatttcaaaaccaatcacaccatcccaa
cagtccctcaaagtcttaaatgatttcagcattaactcaaaagtccacagtctaatgtctca
tctgagacaaggcaagtcctttccatttatgagcctataaaatccaaagcaagttagttact
tcctagatacaatgggggtacaggcattgggtaaatacagccattccaaatgggataaattg
gtcaaaacaaagaggctacaggcccatgagagtccaaaatccagtggggcagtcaaatctta
aagctccaaaatgatctcctttgactccacatctcacatccaggtcacgcagatggaagggg
tgggttcccatggtcttgggcagctctgcccctgtacctttgcagggtacagcctcctctc
agctgctttcatgggctggcattgagtgtctgcagcttttccaggtacacggtgcaagctgt
cggtggatctaccattctggggtctggaggacctcttctcacagctccactaggtggtgccc
cagtagggactgtgtgtggggtctctgaccccacatttcccttctgcactgccctggcagag

Fig.9 **cont.**

Upstream Rhesus box of D-positives

gatctccatgagggccctgctcctgcagcaaacttctgactgggcatccaggcatttccgca
catcctctttaatctaggcgaaggtttccaaaccccaattcttgacttctgtgcactcgcag
tctcaacaccacatggaagctgtcaaggcttggggcttgcactccccgaagctacagcccaa
gctctaccttgcctcccgtcagtcatggttgggagtggctgggatgcagggcaccaagtccc
taggctgcacacagcatgaggaccccgggcctggccaacaaaaccatttttcctgatacct
ctggacctgtgatgggaggggttgccataaagacctctgacatgccctggagacattttccc
cattgtcttgggaattagcatttggctcctgttactcatgcaaatttctgcagccagcttga
atttctcctcagaaaatgggaatttttctttttctatcacattgtcaggctgcaaattttccg
aacttttatgctctgcttcccttataaaactgaatgtctttaacagcacccaagtcacctct
tgaatgctttgctgcttagaaatttctcctgccagatactctaaatcatctctctgaagttc
aaagttctacaaatatctcgtgcaggggcaaaatgccgccagtatctttgctaaaacataac
aagagtccccttgctccagttcccaacaagttcctcatttccgtctgagaccacctcagcc
tatggactttattgtccacagtgctatcagcattttgggcaaagccattcaacaagtctcta
ggaagttccaaactttcccacatttgcctgtcttcttctgagccctccaaactgttccaaac
cctgcctgttacccagttccaaagtcacatacccatttttgagtatctacggcagcacccca
ctctactggtaccaatttagccactgaagtagttggagaacagaagtaatagactctggttt
acattgtaaaagcttctctgtggctgctgtgtgaagaaaatatatgagaatgaagccccaag
atgaagcagggacacagttgcagtggttagagtaagaaatgctgctggctggcactgaagtg
atagcctggaggtttgtgtgtgcacatgcatgtgtatgtgttttacgatagtaggcccaaca
gatactgtaatccacacttgtttttttttttgagacagagtctcacctgttgcctagacta
gaatgcagtggcacaatcttggctcactacaacctccacctcccaggttcaaacaatccttg
tgcttcagcctcccgagtagttgggattacaggtgtgtgccaccgtgcccagctatattttt
tgtatttttagcagagatgggattttgccacattggccaggctggtcttgaactcctggcct
caagcaatcctcccaccttagcctcccaaagtgctgagccaccacacctggccgcaactgat
ttttaatcatgaaatgacacatacatttaaaaaacccaatacctataatattcctggctagt
actcttcacatctatatcatcaaaaacaaagaaagtatgtgaaactgacacagccaagggga
gactaaggagacataacaattaactgtaatgtggtattctggaggggatcctggaacagaaa
aagacattaggcaaaaaactaaagaaatctgaataaaatgtggatgtcagttaataataatg
tatcatattagtccagtaattgtaacaaatataccacaataatgaaagccattaattatagg
gaaaatggaggggttaatatgggtggctggcttttgctatttctagcagctccattttatct
gcaaaagacaaacattcattaagtcccaaaaaggtaaagaatgacaaattaagcatgtatct
tattagtaagagtaatataaagatgctcactcctatttataaatatttgacaatcatgttaa
ggccacaaaagagaaaaaagggtaggggcaaaaaacgcaaagagaaaggagttagtatcttt
tctcccgcactcattagctattaaaagaggatgtttgtttaaagctgctcagagctggtaaa
ctaatgttaagtcactaacgggaatttaaaaggtttcattaagaactgcctgcactagattc
ctccaccctgagacattaaacaatcacgataaacctcctgagtggtaagaacttgtccattt
aaaaacaggctatagattgtatcatgcagttttatctactaatcggctaatatcccgccaaa
aacaaaaaaccccaaagggatgaaagtttcatccatcaaaggaaacaacagtcaccttggtt
cccatctcactcatatactgccgccgtacatgtcaatcagatgaacctgtgcgtatctctta
atgacaattgacccacatttttgactgaagtgaaaggggttctgctccgcgaccacttcct
ggatctccccctccaccctctgtgttctttcgggtgcaccatcgggtcaaagccgcagcaac
gccgtctctgtgtgatcgcatgtgcccttctgcacacgaccttcccccgagagtgaccagct
accggacaggcaccaaggagggctaccgagcacctcccggaccggcggctgcaggatcgcga
gcgcctccgctagggagactgcacgttgcgcctgtgcttcctgcggtggcgccttctgcaag

Fig.9 **cont.**

Upstream Rhesus box of D-positives

gagacctcgaccctgctccctctccggggctggatctgactccttgacggtgattccagacg
cgagacccaaactgacggcttctagaagaggggcgagcccggccgcaagtctttcacgtagc
taagtcatcgttgcttccggcttcttaccgttctccccttttgtaaacggttacctcccgaaa
acccaggctctcctccaacagtggttctcaagcgaggcgatcttccccgggaggggatattt
ggcaaagtctgggggcatttttggttcactggggctgctacttgcatccactgggtagaggc
gggggatgcagctacacaacctgcgaagcacgggacagcaccctccccaacccagacagaat
tagccggcccaaaacctcagtagtgcccaggctgagaaaccctgccttaaacaaacaacaaa
gaaaagccaagtcccataagtgggtcaccgcgccgagactggggtccacgggacaccccagc
cacgccaagccgggaagtccccgcctcctggagctgaacccgcccctctcccagaggtggag
ctgcgggggcgggaacaggcacggagaaaataaacaagactaaaaagtcctgagtagcgct
gtgtggccgcaaacctgaacccacctttgcaccacgcgggacccggcacgcttcctgccac
ccacccctgagagggctgcgcggccgaccccagtactagaaaacactcgtcacctcaatcaa
gacgggtacgaaggccaacggacgccttcctttagaacgctcagcacacagagcaacttctc
acgcctactctcaaatggcgtactccaaactagcactcccgacgtccagctgtgaacccaga
gcggcggaaagcccctgaacccagcgcccgggcatgcgcagacgcgttgttgtggtgggcgt
ggctccctccggacccggcgccccgccctccgccccgtgtccgcatgcgcgactgagccgcg
ggggtggtactgctgcatccgggtgtctg
(end of Rhesus box)
aagatccgatgaaataacatatgcaaaatgattgggtccgtgattggcattccagaaatgg
3'

专利名称(译)	Rhd阴性基因座的分子结构		
公开(公告)号	EP1226169B1	公开(公告)日	2007-03-14
申请号	EP2000972884	申请日	2000-10-31
[标]申请(专利权)人(译)	DRK BLUTSPENDEDIENST巴符州公益有限责任公司		
申请(专利权)人(译)	DRC血液服务巴符州公益有限责任公司		
当前申请(专利权)人(译)	DRC血液服务巴符州公益有限责任公司		
[标]发明人	FLEGEL WILLY A WAGNER FRANZ F		
发明人	FLEGEL, WILLY, A. WAGNER, FRANZ, F.		
IPC分类号	C07K14/47 C12N15/12 C12Q1/68 G01N33/53 A61K39/395 A61P15/00 C07K14/705 C07K16/18 C12N5/10 C12N7/00 C12N15/09 C12P21/02 C12P21/08 C12Q1/02 G01N33/566		
CPC分类号	A61P15/00 C07K14/705 C12Q1/6883 C12Q2600/156 C12Q2600/16 C12Q2600/172 C12N15/52		
代理机构(译)	法思博事务所		
优先权	1999121686 1999-11-02 EP 2000111696 2000-05-31 EP		
其他公开文献	EP1226169A2		
外部链接	Espacenet		

摘要(译)

本发明涉及代表Rhesus基因座的核酸分子结构，其包含RHD，SMP1和RHCE基因和/或恒河猴盒，优选混合恒河猴盒，上游恒河猴盒和/或下游恒河猴盒。 。此外，本发明涉及特异性检测常见RHD阴性单倍型的方法。本发明还涉及D阴性个体中RHD阳性降低型的检测。已经鉴定了RHD基因中的各种突变，其允许开发诊断工具。本发明还涉及寡核苷酸，其特异性杂交杂交盒，优选断裂点或断裂点区域或上游和下游恒河猴盒。另外，本发明涉及包含或使用上述本发明化合物的试剂盒。