



(86) Date de dépôt PCT/PCT Filing Date: 1995/11/09
(87) Date publication PCT/PCT Publication Date: 1996/05/23
(45) Date de délivrance/Issue Date: 2011/07/05
(85) Entrée phase nationale/National Entry: 1997/05/06
(86) N° demande PCT/PCT Application No.: US 1995/014684
(87) N° publication PCT/PCT Publication No.: 1996/015268
(30) Priorités/Priorities: 1994/11/10 (US08/337,268);
1995/06/07 (US08/484,570)

(51) Cl.Int./Int.Cl. *C12N 15/12* (2006.01),
C07H 21/00 (2006.01), *C07K 14/705* (2006.01),
C07K 16/34 (2006.01), *C12Q 1/68* (2006.01),
G01N 33/543 (2006.01), *G01N 33/564* (2006.01)
(72) Inventeurs/Inventors:
LEE, SOOHEE, US;
REDMAN, COLVIN M., US
(73) Propriétaire/Owner:
NEW YORK BLOOD CENTER, INC., US
(74) Agent: PIASETZKI NENNIGER KVAS LLP

(54) Titre : METHODE ET KIT DE DIAGNOSTIC POUR DETERMINER LE GENOTYPE D'APRES LA METHODE DE
GROUPE SANGUIN DE KELL
(54) Title: DIAGNOSTIC METHOD AND KIT FOR DETERMINING KELL BLOOD GROUP GENOTYPE

(57) **Abrégé/Abstract:**

The invention provides a diagnostic method of determining Kell genotype by the identification of the molecular basis of a Kell polymorphism. Specifically, the invention provides a method for determining K1/K2 genotype with great accuracy, overcoming problems associated with traditional serological typing methods. The diagnostic method of the invention preferably employs amplification of K1/K2 nucleic acid sequences, and optionally employs differential cleavage of K1-K2-specific nucleic acid sequences by a restriction enzyme. Also provided are nucleic acid oligomers useful as probes or primers for the method of the invention. Furthermore, diagnostic kits for the determination of Kell genotype are provided.



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12Q 1/68, C12P 19/34, 21/06, C12N 1/20, 15/00, C07H 21/02, 21/04, C07K 5/00, 16/00	A1	(11) International Publication Number: WO 96/15268 (43) International Publication Date: 23 May 1996 (23.05.96)
(21) International Application Number: PCT/US95/14684 (22) International Filing Date: 9 November 1995 (09.11.95) (30) Priority Data: 08/337,268 10 November 1994 (10.11.94) US 08/484,570 7 June 1995 (07.06.95) US (71) Applicant: THE NEW YORK BLOOD CENTER, INC. [US/US]; 310 East 67th Street, New York, NY 10021 (US). (72) Inventors: LEE, Soohye; Unit #1604, 300 Winston Drive, Cliffside Park, NJ 07010 (US). REDMAN, Colvin, L.; 187 Renken Boulevard, Franklin Square, NY 11010 (US). (74) Agents: BARON, Ronald, J. et al.; Hoffmann & Baron, 350 Jericho Turnpike, Jericho, NY 11753 (US).		(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, LS, MW, SD, SZ, UG). Published <i>With international search report.</i>
(54) Title: DIAGNOSTIC METHOD AND KIT FOR DETERMINING KELL BLOOD GROUP GENOTYPE (57) Abstract The invention provides a diagnostic method of determining Kell genotype by the identification of the molecular basis of a Kell polymorphism. Specifically, the invention provides a method for determining K1/K2 genotype with great accuracy, overcoming problems associated with traditional serological typing methods. The diagnostic method of the invention preferably employs amplification of K1/K2 nucleic acid sequences, and optionally employs differential cleavage of K1-K2-specific nucleic acid sequences by a restriction enzyme. Also provided are nucleic acid oligomers useful as probes or primers for the method of the invention. Furthermore, diagnostic kits for the determination of Kell genotype are provided.		

DIAGNOSTIC METHOD AND KIT FOR
DETERMINING KELL BLOOD GROUP GENOTYPE

5 This invention was made with U.S. Government support under Grant HL35841 awarded by the NIH. The U.S. Government has certain rights in this invention.

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. application Serial Number 08/337,268 filed on November 10, 1994, now U.S. Patent No. 5,589,336.

10 BACKGROUND OF THE INVENTION

This invention relates to a method for determining Kell blood group genotype. More particularly, the invention relates to molecular genetic methods for the determination of *inter alia* K1/K2, K6/K7, K10/K(-10), and K3/K4/K21 genotypes.

15 The Kell blood group system is a well-known but complex group of blood antigens, comprising over 20 different related antigens. Of these, the antigen K1 (K, Kell) is known to be the strongest immunogen among the 23 known phenotypes. Serologically, the K1 sublocus has an allelic relationship
20 with a high frequency antigen K2 (k, Cellano). Approximately 9% of the population has the K1 red cell phenotype, and antibodies to K1 are developed in about 5% of persons receiving a single unit of incompatible blood (Ref. 1).

25 Hemolytic disease of the newborn (HDN) is usually associated with maternal alloimmunization to Rh(D), but K1 incompatibilities can also cause severe hemolytic disease in newborns(Refs. 2-7). K1 sensitization from a previous pregnancy can result in HDN and complications during subsequent pregnancies if the fetus is a K1 carrier. Similar

types of problems can arise in the more rare case of K2 sensitization. The identification of the fetal K1/K2 genotype would be of particular significance in situations in which the father is a K1/K2 heterozygote. Since the mother must be a homozygote in order to have been previously sensitized, there is a 50% chance that the fetus with a K:1,2 father is in danger of HDN. Because of this 50% risk, identification of the homozygosity or heterozygosity of the K1/K2 genotype of the fetus in these pregnancies becomes important in order to aid in their proper management. It is important to recognize, however, that the determination of parental genotype with respect to Kell antigens other than K1/K2 can also be of significance for managing pregnancies.

Kell inheritance is autosomal and codominant, and the gene for the Kell protein (KEL) has been mapped to chromosome 7q33 (Refs. 8-11). Kell antigens appear to be encoded in 5 sets of antithetical paired alleles expressing high and low prevalence antigens. Thus, K1 (K) and K2 (k) are products of alleles, as are K3 (Kp^a), K4 (Kp^b) and K21 (Kp^c); K6 (Js^a) and K7 (Js^b); K17 and K11; and K24 and K14. However, a number of high prevalence antigens such as K12, K13, K18 and K22 are independently expressed. Some Kell phenotypes are segregated by racial or ethnic groups. For example, K6, which was first reported in 1958, is more frequent in African-Americans, occurring in up to 19.5% in this group in a study performed in Seattle, Washington (Ref. 12). Also, K10 (U1^a) is more prevalent in Finns (Ref. 13). These various relationships, and their places in the Kell system, have been established through the years by serological analyses of informative families (Refs. 14-18).

A variety of studies, including a molecular cloning, established that Kell blood group antigens are carried on a 93 kDa type II glycoprotein (Refs. 19-25) found on the surface of red blood cells (Ref. 26). The Kell protein has a short, 46 amino acid, N-terminal domain in the cytoplasm, and a large C-

terminal portion, of 665 amino acids, on the external surface of the red cell. All of the carbohydrates are N-linked (Ref. 27), probably located in 5 sites, at asparagines 93, 115, 191, 345 and 627. Early biochemical studies suggested that Kell antigens reside on a protein whose conformation is largely dependent on disulfide bonds (Ref. 28). The Kell protein has 16 cysteine residues, one in the transmembrane region and 15 in the external portion (Ref. 25). Reduction of red cells by sulfhydryl reagents results in loss of Kell antigens and exposure of some neo-epitopes (Ref. 28).

The antigenic structure of the Kell protein has also begun to be mapped. A recently developed immunological test, MAIEA, which uses monoclonal antibodies to different Kell antigens, indicates that certain of the identified Kell antigens occur in spatially distinct regions of the glycoprotein (Ref. 26). For example, the K1/K2, and K6/K7 domains appear to be close together, while K3/K4 epitope is in a different location and K18 is in yet another protein domain (Ref. 29).

The determination of Kell genotype has heretofore been confined to inferences drawn from the detection and identification of Kell antigens. Such methods employ antibodies or other compounds which identify and interact with the Kell protein or portions thereof. For example, various antibodies specific for particular Kell antigens have been identified (Refs. 13, 22). Agglutination methods for detecting Kell protein and other blood group antigens are described in U.S. Patent Nos. 5,324,479, 5,302,512, 5,213,963, 4,560,647, 4,403,042, 4,358,436, and 4,148,607. These methods provide information about expressed protein profiles, not about protein molecular structure or the molecular genetic makeup of the individual. Zelinski et al. describe a method for indirectly inferring Kell genotype from linkage studies using marker genes (Ref. 30). Such methods, however, provide only limited and indirect information about Kell genotype.

Generally, these methods also require obtaining blood samples from subjects being examined. Determining fetal Kell phenotype also normally requires a fetal blood sample. This involves a potentially dangerous procedure in which the fetus is susceptible to hemorrhage and possibly death. None of the methods previously described discloses any method by which Kell genotype might be determined simply and directly on the basis of the *KEL* gene structure.

As a result, there exists a need for a method of safely and conveniently detecting Kell genotype. It would also be desirable to provide a method for determining Kell genotype in a fetus without the requirement for obtaining blood samples. A test based on DNA samples taken from amniotic cells would allow the clinician to avoid the risk of harm to the fetus and to more accurately predict the potential of anti-Kell-associated HDN.

SUMMARY OF THE INVENTION

The present invention provides a diagnostic method for the differential determination of Kell blood group genotype in a subject. The diagnostic method includes generating a characteristic nucleic acid product, from a source or sample of nucleic acid containing a Kell polymorphism locus, in an amount which is sufficient to enable the characterization of Kell genotype. The method may be employed to determine Kell genotype with respect to one or more Kell polymorphisms or Kell phenotypes. Preferably, the method is directed to detecting the Kell polymorphism loci which determine: K1/K2 genotype, K6/K7 genotype, K10/K(-10) genotype, and/or K3/K4/K21 genotype.

In a first preferred embodiment, the method includes: selectively cleaving a nucleic acid sample which includes a Kell polymorphism locus, preferably genomic DNA of the subject, to provide a nucleic acid product which includes one

or more nucleic acid fragments in an amount sufficient to characterize Kell genotype. Preferably, the method includes selectively cleaving nucleic acid which includes one or more of the K1/K2 polymorphism locus, the K6/K7 polymorphism locus, the K10/K(-10) polymorphism locus, and the K3/K4/K21 polymorphism locus.

The selective cleaving of DNA is accomplished by digesting the sample nucleic acid with a restriction enzyme which cleaves the DNA differentially based on a Kell polymorphism locus. Any suitable restriction enzyme may be employed as desired. In the case of K1/K2 polymorphism locus, a preferred restriction enzyme is *BsmI*. In the case of the K6/K7 polymorphism locus, preferred restriction enzymes include *MnII* and/or *DdeI*. In the case of K10 genotype,, a preferred restriction enzyme is *AccI*. In the case of the K3/K4/K21 polymorphism, preferred restriction enzymes include *NlaIII* and/or *PvuII*. Combinations of such enzymes may be employed when more than one Kell polymorphism is desired to be differentiated.

This diagnostic method preferably includes amplifying Kell DNA from a DNA sample obtained from a subject. Typically the amplification involves using a primer which amplifies only nucleic acid including the locus which determines a specific Kell polymorphism. Primers may be selected to selectively amplify DNA including particular Kell polymorphism loci, including the K1/K2 polymorphism locus, the K6/K7 polymorphism locus, the K10/K(-10) polymorphism locus, and/or the K3/K4/K21 polymorphism locus.

Preferably, the digesting of Kell nucleic acid by a restriction enzyme is performed following amplification of nucleic acid obtained from the subject, but the digesting may be performed preceding such amplification.

In a second preferred embodiment, the diagnostic method of the invention involves differentially amplifying a nucleic acid sample which includes a Kell polymorphism locus to provide a nucleic acid product which can be employed to characterize Kell genotype. Preferably, the sample nucleic acid operates as template permitting amplification by means of a primer which differentially amplifies nucleic acid encoding one specific allele of a set of alleles associated with a Kell polymorphism.

In the diagnostic methods of the invention which require a primer or primer set, the primer or primer set may include a single primer, or may include a plurality of primers, including, for example, a plurality of primer pairs. For example, in a preferred mode, the method of the invention may employ a primer which specifically amplifies K1 DNA (K1 allele-specific primer), a primer which specifically amplifies K2 DNA (K2 allele-specific primer), and two primers which each amplify both K1 and K2 DNA (allele-non-specific primers). In this mode, different sized PCR products are produced which correspond to a K1 DNA, K2 DNA and K1/K2 DNA. The pattern of PCR fragments serves to differentiate among K1/K1, K2/K2, and K1/K2 genotypes.

Primers are included which differentially amplify nucleic acid templates containing a K6 locus and/or a K7 locus. Primers are also included which differentially amplify nucleic acid templates containing a K10 locus and/or a K(-10) locus. Also, primers are included which differentially amplify nucleic acid templates associated with a K3 locus, a K4 locus, and/or a K21 locus.

An especially preferred oligonucleotide primer which is allele-specific for K1-containing nucleic acid comprises a nucleotide sequence selected from the group of nucleotide sequences including: ATA CTG ACT CAT CAG AAG TTT CAG CA (SEQ

IPEAUS 10 JUN 1996

ID NO:1), and ATA CTG ACT CAT CAG AAG TCT CAG CA (SEQ ID NO:2), and ATA CTG ACT CAT CAG AAG TGT CAG CA (SEQ ID NO:61).

5 An especially preferred oligonucleotide primer which is allele-specific for K2-containing nucleic acid has the nucleotide sequence: TGG ACT TCC TTA AAC TTT AAC TGA AC (SEQ ID NO:3).

10 Especially preferred oligonucleotide primers specific for K1- and K2-containing nucleic acids (but allele-non-specific) have nucleotide sequences selected from the group of nucleotide sequences including: TTT AGT CCT CAC TCC CAT GCT TCC (SEQ ID NO:4), and TAT CAC ACA GGT GTC CTC TCT TCC (SEQ ID NO:5), TTA GTC CTC ACT C-C CAT GCT TCC (SEQ ID NO:62), and TCA CAC AGG TGT CCT CTC TTC C (SEQ ID NO:63). The primers may be used in pairs for amplification.

15 An especially preferred oligonucleotide primer specific for K6- and K7-containing nucleic acids has nucleotide sequences selected from the group of nucleotide sequences including: CTC ACC TAG GCA GCA CCA ACC CTA (SEQ ID NO:64) and TTA CCT GGA GGG CAT GGT TGT CAC T (SEQ ID NO:65).

20 The diagnostic method of the invention preferably further includes the derivation of information from generated nucleic acid restriction fragments or amplification products. Preferably, the derivation of information includes separating the nucleic acid products to provide a pattern of
25 fragments which provides specific information characterizing Kell genotype. More preferably, the separation provides a pattern of nucleic acid products which provides specific information characterizing K1/K2 genotype, K6/K7 genotype, K10/K(-10) genotype, and/or K3/K4/K21 genotype. In addition,
30 it is preferred to detect some or all of the Kell nucleic acid products, such as by marking or staining one or more products. Products may be marked with a non-specific marker substance. Such substances may be used to stain some or all of the

products to permit their differentiation based on the pattern of their separation. Alternatively, the detecting step may include marking one or more of the Kell nucleic acid products with one or more hybridization probes in a selective or
5 specific fashion. Preferably, a hybridization probe(s) is employed which specifically marks one or more nucleic acid products which include a Kell polymorphism locus of interest, such as the K1 locus or the K2 locus.

10 The molecular genetic method of the invention generally involves obtaining nucleic acid, preferably genomic DNA, from a biological sample from a human subject or patient. A blood sample is typically preferred, but other types of tissue samples which contain erythroid tissue are useful. In a particularly preferred embodiment, the invention provides a
15 method for determining K1/K2 Kell blood group genotype in a fetus. In this embodiment, the preferred tissue sample includes a sample of amniotic fluid or chorionic villus.

The invention further provides Kell-based nucleic acid oligomers which include nucleotide sequences which are at
20 least substantially complementary to a polymorphic region of a Kell-based nucleic acid, such as Kell genomic DNA, mRNA or cDNA. In particular, the invention provides nucleic acid oligomers which include a nucleic acid sequence substantially complementary to K1 DNA, as well as oligomers which include a
25 nucleic acid sequence substantially complementary to K2 DNA. Preferably, the oligomers of the invention are exactly complementary to the region of interest.

Moreover, the invention provides nucleic acid oligomers which include a nucleic acid sequence substantially homologous
30 to a polymorphic region of a Kell-based nucleic acid, such as Kell genomic DNA, mRNA or cDNA. In particular, the invention provides nucleic acid oligomers which include a nucleic acid sequence substantially homologous to K1 DNA, as well as oligomers which include a nucleic acid sequence substantially

homologous to K2 DNA. These oligomers of the invention preferably include a nucleic acid sequence which is at least in part, exactly homologous to a target region of a Kell-based nucleic acid.

5 The nucleic acid oligomers of the invention may be used as hybridization probes to identify the presence of target nucleic acid sequences, through binding to or hybridizing with such target sequences. Accordingly, the nucleic acid oligomers may be detectably labeled by being linked to a
10 detectable marker moiety such as a fluorescent label, an electron dense substance, a reporter moiety, a specific or nonspecific binding moiety, or other detectable moiety such as is known in the art. Optionally, the oligomers of the invention may further include a reactive moiety permitting
15 cross-linking with a target nucleic acid sequence. Furthermore, the oligomers of the invention may be linked to a substrate material, for example, to a gel or resin to immobilize the oligomers.

20 The oligomers of the invention may also be employed as amplification primers which specifically bind to and cause elongation through a region of Kell nucleic acid comprising a Kell polymorphism locus. Preferred primers bind to or cause elongation through Kell nucleic acid which contains a K1/K2 locus, a K6/K7 locus, a K10/K(-10) locus, and/or a K3/K4/K21
25 locus.

30 Moreover, the identification of the locus characterizing the K1/K2 polymorphism now permits the preparation of Kell-based polypeptides which include amino acid sequences which are substantially homologous to the K1 domain or to the K2 domain of the Kell protein. The polypeptides may be derived from natural sources and substantially purified or may be synthesized in substantially pure form as desired. Such polypeptides may also be detectably labeled, may be attached to reactive moieties, and may be bound to a substrate in

accordance with methods known in the art. The polypeptides of the invention are useful as probes for, for example, detecting alloimmunization in a subject or patient. In such an assay, the polypeptide may comprise an amino acid sequence which is substantially homologous to K1 antigen and presents an immunologic profile which permits specific reaction with anti-K1 antibodies.

Thus in another embodiment, the invention provides a diagnostic method for detecting alloimmunization of a patient to a Kell antigen, preferably K1 antigen. In this embodiment, the method includes obtaining a blood sample from a patient or subject, and measuring a parameter of immune reactivity of the sample with a polypeptide probe. The polypeptide probe preferably includes an amino acid sequence which is substantially homologous to the K1/K2 domain of the Kell protein. Moreover, it is desirable that the polypeptide probe be specifically reactive with anti-K1 antibodies present in the sample. The invention also enables comparable methods for the detection of alloimmunization of a subject to other Kell antigens, including K6, K10, as well as K3 and K21.

In another embodiment, a diagnostic kit is provided for determining Kell blood group genotype in a sample of tissue from a patient. In this embodiment, the invention provides a diagnostic kit for determining Kell blood group genotype by detecting target nucleic acid sequences, such as sequences specific to K1 and K2. The kit includes amplification primers, i.e., oligonucleotides that bind to or cause elongation through sequences specific to K1 and K2. The in vitro kit further includes a container, such as a microtiter plate having a plurality of wells, having bound thereto oligonucleotide capture probes having nucleic acid sequences substantially complementary to the K1 and K2 target sequences.

Alternatively, the invention provides a diagnostic kit for determining Kell blood group genotype by detecting target nucleic acid sequences specific to particular Kell antigens, such as K1 and K2. In this embodiment, the kit includes:

- 5 (a) a primer set including first and second PCR primers wherein the first PCR primer is an oligonucleotide that will bind to or cause elongation through a sequence specific to K1 and the second PCR primer is an oligonucleotide that will bind to or cause elongation through a sequence specific to K2; and
- 10 (b) a container, such as a microtiter plate having a plurality of wells, having bound thereto oligonucleotide capture probes having a nucleic acid sequence substantially complementary to the target sequences.

15 The invention further provides expression recombinant vectors which carry Kell nucleic acid sequences. The invention provides expression vectors which include a nucleic acid sequence which encodes at least a part of the Kell protein including a part of the protein which encodes a site of Kell polymorphism. In particular, the invention provides

20 expression vectors which carry K1 cDNA permitting transformation of cells to produce transformed cells or transformants which express K1 protein on their cell surfaces. Vectors are also provided which carry K6 cDNA, as well as vectors which carry K3 or K21 cDNA, permitting transformation

25 of cells to express K6, as well as K3 and K21, respectively. The invention also provides a stable cell lines which have been modified (i.e., transformed) to express protein on its cell surface, as well as a method for transforming a cell line to express such protein. With respect to the K1 expression,

30 such protein preferably includes at least the K1 domain, and more preferably, the protein is K1 protein. Similarly expression of K6 as well as K3 and K21, the expressed protein preferably includes the K6, the K3 or the K21, domain

35 respectively; more preferably, the protein is K6, K3, or K21 protein, respectively. The invention further provides a method of making an antibody specific for a Kell antigen, by

inducing the generation of at least one antibody against a Kell antigen expressed by a transformed cell line produced according to the invention. Polyclonal antibodies are contemplated, but the antibody is preferably monospecific, such as a monoclonal antibody or an antigen-binding region thereof.

Accordingly, as a result of the invention, there is now provided a safe and convenient diagnostic method for differentially determining the Kell genotype of patients. In particular, there is now provided a method for determining Kell genotype in a fetus in utero without requiring the taking of a blood sample. A test based on determining Kell genotype, in particular K1/K2 genotype, from DNA obtained from amniotic fluid now allows the clinician to reduce the degree of risk to the fetus being tested. Moreover, the new diagnostic method permits the accurate prediction of the potential of anti-K-associated hemolytic disease of the newborn.

These and other advantages of the present invention will be appreciated from the detailed description and examples which are set forth herein. The detailed description and examples enhance the understanding of the invention, but are not intended to limit the scope of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred embodiments of the invention have been chosen for purposes of illustration and description, but are not intended in any way to restrict the scope of the present invention. The preferred embodiments of certain aspects of the invention are shown in the accompanying drawing, wherein:

Figure 1 shows a map of the Kell gene indicating the loci of the Kell exons, the relationships among the clones used to establish the exon locations, as well as restriction sites in the Kell gene.

Figures 2A-2S show the nucleotide sequences of the exons of the Kell gene.

Figure 3 shows the 5' flanking region and exon 1 of the Kell gene.

5 Figure 4 shows an autoradiogram illustrating the promoter activity of the 5' flanking region of the Kell gene.

Figure 5 shows a comparison of the 3' ends of the Kell gene obtained from various clones of the gene.

10 Figure 6 shows an ethidium bromide/agarose gel illustrating a differential separation of amplified DNA corresponding to specific fragments of the Kell gene.

Figure 7 shows a comparison of corresponding portions of K1 DNA and K2 DNA.

15 Figure 8 shows an ethidium bromide/agarose gel illustrating a differential separation of amplified DNA corresponding to K1 DNA and K2 DNA according to the method of the invention.

20 Figure 9 shows an ethidium bromide/agarose gel electrophoresis resolution of PCR products obtained according to the method of the invention.

Figure 10 shows a comparison of corresponding portions of K6 DNA and K7 DNA.

25 Figure 11 shows an ethidium bromide/agarose gel electrophoresis resolution of K6/K7 PCR products obtained according to the method of the invention.

Figure 12 shows an ethidium bromide/agarose gel electrophoresis resolution of K6/K7 PCR products obtained according to the method of the invention.

5 Figure 13 shows a comparison of corresponding portions of K10 DNA and K(-10) DNA.

Figure 14 shows a comparison of corresponding portions of K3 DNA, K4 DNA, and K21 DNA.

DETAILED DESCRIPTION OF THE INVENTION

10 The Kell blood group system is distinguished by its antigenic complexity (Ref. 18). Over 20 different antigens have been ascribed to the Kell system. A majority of the antigens are organized in 5 antithetical sets of high and low frequency antigens, with other Kell antigens being independently expressed. The molecular basis for this
15 antigenic diversity has not previously been understood. The present invention, which involves description of the organization of the *KEL* gene and the boundaries of its 19 exons, now enables the molecular characterization of the different Kell phenotypes.

20 The molecular basis of the different Kell phenotypes has not previously been determined. Having now studied the structure of the *KEL* gene, and identified for the first time the 19 exons which encode Kell protein, molecular bases of polymorphisms in the Kell genotype have now been determined.

25 In particular, the K1/K2 polymorphism and its locus has now been determined by sequencing the exons of K1/K1 DNA and comparing them to K2/K2 sequences. It was unexpectedly been found that base substitution in exon 6 of the K1 genotype predicts an amino acid change. This base substitution also

creates a restriction enzyme site which has been employed to test over 40 different samples to confirm that the base substitution identifies the K1 genotype.

5 As another of the Kell group of antigens, K6 (also known as Js^a) is a low frequency red cell antigen which has an allelic relationship with the high frequency K7 antigen. K6 appears in less than 1.0% of the total population, but up to 19.5% of African-Americans carry the K6 antigen. We have now determined the molecular basis of polymorphism in the K6/K7 polymorphism by sequencing the 19 exons of the Kell gene (*KEL*) of a K:6,-7 person. Comparisons of the sequence obtained from K:6,-7 DNA with that of K:6,7 DNA showed an unexpected base substitution in exon 17 that predicts an amino acid change from leucine to proline. This base change also causes a deletion of a restriction enzyme cleavage site. Another base change in K7 is silent in that it does not result in change in the expressed protein, yet it causes the deletion of another restriction enzyme cleavage site which can be exploited to differentiate K6/K7 genotype.

20 The K10 (U1^a) antigen is a low frequency Kell antigen for which there is no corresponding serologically detectable high frequency antigen. Thus, the wild-type phenotype (designated herein as K(-10) is characterized by a lack of the K10 antigen. It has now been determined that the molecular basis of the K10 polymorphism resides in exon 13. A point mutation at nt 1601 has been found to predict an amino acid change from glutamic acid to valine. This base change also creates a new restriction enzyme cleavage site which can be employed to determine K10 genotype according to the invention.

30 Another of the antithetical allelic sets in the Kell genotype is the K3/K4/K21 (Kp^a/Kp^b/Kp^c) set. Our study of the *KEL* gene has also unexpectedly revealed that the three alleles in this set may be differentiated on the basis of changes in two adjacent nucleotides in exon 8, each of which predicts a

discrete amino acid change at the same position in the amino acid sequence of the Kell protein. Each of the base changes also creates a new and distinct restriction enzyme cleavage site in the gene.

5 The most prevalent Kell phenotype is K:-1, 2, -3, 4, -6, 7, 9, 11, 12, 14, 18, 19, 22. We have now defined the 19 exons of the KEL gene of a person with common Kell phenotype. To determine the molecular basis of Kell polymorphisms we designed a series of primers which would amplify the 19 exons
10 of KELL. We then compared the DNA sequences of persons expressing various combinations of high and low frequency antigens to determine sequence variations potentially related to the observed variations in antigen expression.

 For example, we compared the DNA sequences of K1/K1 and
15 K2/K2 DNA. The only base change which would encode a different amino acid was found in exon 6, and changed a threonine to a methionine at a consensus N-glycosylation motif (Asn.X.Thr → Met). This change would prevent N-linked glycosylation at this site. Based on the amino acid sequence
20 of Kell protein of common phenotype, there are 6 possible N-glycosylation sites, i.e., asparagine residues 93, 115, 191, 345, 627 and 724. However, the asparagine at position 724 is probably not glycosylated because it is part of a sequence Asn.Pro.Ser and the presence of proline between asparagine and
25 serine/threonine inhibits N-glycosylation (Ref. 31). In any event, changing threonine to methionine at position 193 would prevent glycosylation at asparagine 191. Thus, K1 protein would be composed of at most 4 instead of 5 carbohydrate moieties. Lack of a carbohydrate side-chain may expose
30 different parts of the protein leading to immunogenicity. The loss of glycosylation in a red cell surface protein can lead to a change in blood group phenotype. The Webb glycophorin C variant also lacks an N-glycan (Refs. 32-33).

It has also now been observed that the point mutation from C to T in exon 6 also creates a new restriction enzyme site, 5'-GAATGCT-3', which can be cut by *BsmI*, a well known restriction endonuclease. The use of restriction enzyme digestion allows the differentiation of K1/K1 and K2/K2 homozygotes and of K1/K2 heterozygotes, permitting the development of a diagnostic genotype procedure.

While not wishing to be bound by theory, it is known that, in very rare cases, red blood cells will be observed which do not express any Kell antigens. In this phenotype, known as K_0 (null), the red cells appear not to have any Kell protein on the cell membrane. Yet, preliminary experiments in our laboratory indicate that two K_0 persons contain Kell mRNA in peripheral blood and that the sequence of the mRNA from the initiation ATG codon to the poly A tail is identical to mRNA obtained from persons with common Kell phenotype. This indicates that the base sequences in the 19 exons of K2 and of some K_0 persons are identical. Therefore, PCR amplification of exon 6 and genotyping by treatment with *BsmI* could indicate a K2 genotype in K_0 persons. Serological analysis easily detects the K_0 homozygote, but the K_0 heterozygotes would be serologically identified as K2 or K1, and *BsmI* analysis would provide no further discrimination. Presumably, the same would occur in the K_{mod} phenotypes in which expression of all Kell-related antigens are weakened. Nonetheless, for practical purposes, such ambiguous phenotypes are less clinically significant since K_0 and K_{mod} phenotypes are very rare. K_0 heterozygotes would be phenotypically, and therefore clinically, equivalent to K1 or K2 homozygotes. What is generally important in K1-sensitized pregnancies is to identify whether the fetus is a K1 carrier or not. The method of the invention now allows identification of Kell genotype, including identification of the K1 gene.

We have also now determined the molecular basis of polymorphism in the K6/K7 allelic set ("K6/K7 polymorphism") by sequencing the 19 exons of the Kell gene (*KEL*) of a K:6,-7 person. Comparisons of the sequence obtained from K:6,-7 DNA with that of K:6,7 DNA showed a T to C base substitution at nt 1910 in exon 17 that predicts an amino acid change from leucine in K7 to proline in K6. The T to C substitution at nt 1910 eliminated an *MnlI* restriction enzyme site found in K7. Analysis of a 111 bp PCR-amplified product of exon 17 spanning the T to C substitution was performed with *MnlI*. Eight unrelated persons with the K6 phenotype were tested and all contained the T to C substitution at nt 1910. Another base substitution from A to G at nt 2019 in exon 17 was also noted, but this point mutation did not encode a different amino acid. Even so, this substitution was found to create a new *DdeI* restriction enzyme site. Thus, these base substitutions permit diagnostic procedures for the differentiation of K6 from K7 according to the invention.

The molecular bases of the K3/K4/K21 polymorphism have also now been characterized. Two discrete changes in a single codon have been found to be associated with changes in protein expression from K4 to either of K3 or K21. The same changes also encode new and unique restriction cleavage sites which specifically characterize K3 and K21 nucleic acid as distinct from wild-type K4 nucleic acid. In the case of K3, a C 961 T substitution results in an Arg 281 Trp change. This base substitution also creates a new *NlaIII* restriction enzyme site. In the case of K21, a G 962 A substitution encodes Arg 281 Gln. This substitution also creates a new *PvuII* restriction enzyme site. Accordingly, the differentiation of K4 from K3 and K21 genotypes is possible by employing the diagnostic method of the invention.

For purposes of more clearly and accurately describing the invention herein, certain terminological conventions have been adopted in the following discussion. These conventions are intended to provide a practical means for enhancing
5 description of the invention, but are not intended to be limiting, and the skilled artisan will appreciate that other and additional interpretations may be implied.

As used herein, the term "restriction fragment length polymorphism" or "RFLP" refers to the differences in DNA
10 nucleotide sequences that are randomly distributed throughout the entire genome and that produce different restriction endonuclease patterns for different individuals upon digestion of genomic DNA. Likewise, "restriction site" refers to a site
15 in a nucleic acid at which a restriction enzyme cleaves the nucleic acid. Cleavage of a nucleic acid by a restriction enzyme generates pieces of nucleic acid termed "restriction fragments". When a restriction site is gained or lost as a
20 result of mutation or allelism, the restriction fragment pattern can vary, yielding information about the location of the site.

"Polymorphic" or "DNA polymorphism" refers to the condition in which two or more variations of a specific DNA sequence coexist in the same interbreeding population. Polymorphic differences are caused by differences in amino
25 acid sequence which may be due to point mutations, gene rearrangements or alternative splicing. "Alleles" are genes which are polymorphic variants of one another and occur alternatively at a given locus in the genome.

A gene encoding a Kell protein is designated a "Kell
30 gene". The region of a Kell gene which encodes a specific domain of a Kell protein is termed a "Kell locus". For example, it has been determined as a result of the present invention that the K1 antigen and the K2 antigen are encoded by alleles of the Kell gene. More specifically, the K1

antigen and the K2 antigen are alternate versions of the same domain of the Kell protein, in which a single amino acid substitution has been found to result in the difference between Kell protein including the K1 antigen and Kell protein including the K2 antigen. This single amino acid substitution has now been found to arise from a single nucleotide substitution in the Kell gene. Thus, the portion of the Kell gene which includes this site of nucleotide variation is designated the "K1/K2 locus". When the K1/K2 locus encodes K1 antigen this region of the Kell gene is designated the "K1 locus". When the K1/K2 locus encodes K2 antigen, this region of the Kell gene is designated the "K2 locus". The K1/K2 locus is also said to be the site or locus which determines or characterizes the "K1/K2 polymorphism", i.e., the observed phenotypic difference between the K1 and K2 antigens. Similarly, other Kell loci determine or characterize other Kell polymorphisms, such as the K6/K7, K10/K(-10), and K3/K4/K21 polymorphisms.

For purposes of this invention, Kell protein which includes a particular domain of interest may be designated according to that domain regardless of the other domains in the protein. Thus, if a Kell protein includes the domain responsible for K1 antigen, the protein may be designated "K1 protein", while Kell protein which includes the K2 antigen or K2 domain is designated "K2 protein". Using similar analysis, DNA encoding K1 protein may be termed "K1 DNA", while DNA encoding K2 protein may be termed "K2 DNA" regardless of whether or not the DNA further encodes other Kell domains. It is also to be understood that other Kell-related or Kell-based chromosomal DNA (including exons and introns and parts thereof), nucleic acid templates, nucleic acid transcripts, as well as cDNA, may be similarly designated as K1/K2, K1, or K2, depending on context or application. Similar considerations apply in the context of other antigens in the Kell system, such as the K6/K7, K10, and K3/K4/K21 antigens.

As a result of the present invention, it is now possible to design probes and/or primers comprising nucleic acid oligomers or oligonucleotides which can be employed to detect polymorphisms in the Kell gene. The probes or primers
5 suitable for this purpose preferentially hybridize with or are specific for a region of the Kell gene comprising a site or region in which a point change in nucleotide sequence causes a change in the Kell gene product characterizing a polymorphism. Accordingly, the probes or primers include those which
10 hybridize with higher frequency alleles as well as those which hybridize with lower frequency alleles. In certain applications, it is desirable to define the presence of heterozygosity. In such cases probe or primer combinations enabling differential detection of two or more alleles and/or
15 loci have been found to be useful.

For example, a probe of the invention may be said to bind to or hybridize with Kell DNA if it specifically recognizes a defined region of Kell DNA. Greater precision is intended when a probe is said to hybridize with a specific allele such
20 as a "K1 probe" which is understood to hybridize with K1 DNA.

A primer of the invention may be said to specifically amplify Kell DNA if it binds to or causes elongation through a defined region of Kell DNA. Thus, a primer specifically
25 amplifies K1/K2 DNA if it preferentially amplifies a region including the K1/K2 locus. Accordingly, a K1/K2 primer amplifies DNA including the K1/K2 locus but non-selectively amplifies both K1 DNA and K2 DNA. On the other hand, a primer is specific for a particular allele, such as K1 (K1 primer),
30 if it specifically amplifies only DNA associated with that allele, such as K1 DNA, while a primer specific for K2 (K2 primer) specifically amplifies only K2 DNA. Similar considerations apply to other polymorphisms in the Kell system.

It is particularly preferred that probes be capable of differentiating two alleles which differ by no more than a single nucleotide modification. It is known in the art that it is possible to control the specificity of hybridization by selectively constructing probes and primers and by adjusting hybridization or other experimental conditions. There may be situations, however, in which it would be desirable to employ probes which hybridize somewhat less selectively. Accordingly, it is within a particular context that a probe or primer according to the invention is said to be "substantially complementary" to a specific Kell sequence. That is, if the situation demands high precision, a probe or primer is substantially complementary to a target sequence if there exists only a very small probability that the oligomer will bind with a sequence other than the specific target sequence. In other situations, a probe or primer may be deemed to be substantially complementary to a target sequence if the probability of a unique hybridization is substantially less than 1.0. Thus, a probe or primer of the invention may be "substantially complementary" to a target region if it is either exactly complementary, or even only partially complementary, to the target region, depending on the required stringency of the method parameters.

Thus, the invention provides probes and primers which hybridize with parts or all of the Kell locus. Such probes are useful when the Kell gene or transcripts thereof are desired to be characterized. Moreover, the invention provides probes and primers which include part or all of one or more introns in the Kell locus in chromosomal DNA. Such probes are useful when the Kell gene itself is desired to be characterized and additional information is desired to be obtained about the structure of the gene in a particular individual.

The probes or primers of the invention may also include, as part of their nucleotide sequences, regions which are not

substantially complementary to any region adjacent to or near a target sequence. Thus, in any probe or primer of the invention, at least a part of the probe or primer is substantially complementary to a target segment.

5 "Amplification" refers to any molecular biology technique for detection of trace levels of a specific nucleic acid sequence by exponentially amplifying a template nucleic acid sequence. Techniques are known for the amplification of DNA as well as RNA, and for the generation of amplified RNA from
10 DNA and vice versa. In particular, amplification techniques include such techniques as polymerase chain reaction (PCR), and ligase chain reaction (LCR). The PCR is known to be a highly sensitive technique, and is in wide use. The LCR is known to be highly specific, and is capable of detecting point
15 mutations. In certain circumstances it is desirable to couple the two techniques to improve precision of detection. The PCR is a well-known technique and is described, for example, in Innis et al. (Ref. 34). The LCR is more recently developed and is described in Landegren et al. (Ref. 35) and Barany et
20 al. (Ref. 36). An LCR kit is available from Stratagene. Other amplification techniques may be expected to be employed according to the method of the invention.

 "Primer" refers to an oligonucleotide, whether natural, cloned, or synthetic, which is capable of initiating nucleic
25 acid synthesis for exponential amplification of a target or template nucleic acid sequence by an amplification technique. For PCR the primer acts as a point of initiation of DNA synthesis under conditions in which synthesis of a primer extension product complementary to a nucleic acid strand is
30 induced, i.e., in the presence of four different nucleoside triphosphates and an agent for polymerization (i.e., DNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature. For LCR the primer is capable of annealing to a target nucleic acid and of being ligated to
35 an adjacent primer to serve as a template for amplification.

Also for purposes of the LCR, the primer generally includes paired sets of adjacent, complementary oligonucleotides which can anneal to single stranded target molecules and ligate them together. For LCR amplification of DNA, the primers include
5 two sets of adjacent, complementary oligonucleotides.

A primer is preferably an oligodeoxyribonucleotide and is single stranded for maximum efficiency in amplification, but may also be double stranded. If double stranded, the primer is generally first treated to separate its strands before
10 being used to prepare extension products. Typically, LCR primers are double stranded. The exact length of a primer will depend on many factors, but typically ranges from 15 to 25 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid
15 complexes with the template. A primer need not reflect the exact sequence of the template but must be sufficiently complementary to hybridize with a template. An example of a non-complementary sequence which may be incorporated into the primer is a sequence which encodes a restriction enzyme
20 recognition site (see U.S. Pat. No. 4,800,159).

"Primer", as used herein, may refer to more than one primer, particularly in the case where there is some ambiguity in the information regarding one or both ends of the target region to be amplified. For instance, if a nucleic acid
25 sequence is inferred from a protein sequence, an operable "primer" may actually include a collection of primer oligonucleotides containing sequences representing some or all of the possible codon variations based on the degeneracy of the genetic code. One of the primer oligonucleotides in this
30 collection will be homologous with the end of the target sequence. Likewise, if a "conserved" region shows significant levels of polymorphism in a population, mixtures of primers can be prepared that will amplify adjacent sequences.

A primer or probe according to the invention can be labeled, if desired, by incorporating a label detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ³²P, fluorescent dyes, electron-dense reagents, reporter molecules such as enzymes (as commonly used in ELISAs), biotin, or haptens or proteins for which antisera or monoclonal antibodies are available. A label can also be used to "capture" the primer, so as to facilitate the immobilization of either the primer or amplified DNA on a solid support.

"Oligonucleotide" or "nucleic acid oligomer" refers to primers, probes, nucleic acid fragments to be detected, nucleic acid controls, and unlabeled blocking oligomers and is defined as a molecule comprised of two or more deoxyribonucleotides or ribonucleotides. The nucleic acid oligomers of the invention may be single-stranded oligomers of deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). The exact size of an oligonucleotide will depend upon many factors and the ultimate function or use of the oligonucleotide. The oligodeoxyribonucleotides and oligoribonucleotides may be obtained or derived by known methods from natural sources. Alternatively, the oligonucleotides may be produced synthetically according to methods known in the art. Such methods include, for example, cloning and restriction of appropriate sequences and direct chemical synthesis by a method such as the phosphotriester method of Narang et al. (Ref. 37); the phosphodiester method of Brown et al. (Ref. 38); the diethylphosphoramidite method of Beaucage et al. (Ref. 39); and the solid support method in U.S. Patent No. 4,458,066. It is preferred that the oligomers be at least substantially purified to avoid introduction of artifacts into the genotype determination method.

The invention also provides oligonucleotides which are structurally homologous to a part or all of genetic material related to or derived from the Kell gene. An oligomer is said

to be "substantially homologous" to another if, despite variations in the nucleotide sequence, the oligomer encodes an amino acid sequence which is phenotypically identical to the sequence to which it is being compared. Thus, sequences are not substantially homologous if, when expressed, the sequences result in shift in Kell phenotype. On the other hand, two sequences encoding the same locus or sublocus are substantially homologous if they are not identical but nonetheless encode sequences which correspond to parts or all of phenotypically non-differentiable variants of the Kell protein. Accordingly, sequences which encode K1 domain and sequences which encode the K2 domain are not substantially homologous, even though they differ, as has now been discovered, by a single nucleotide substitution. In contrast, variations in nucleotide sequence which do not result in an amino acid shift in the encoded gene product may be substantially homologous.

The method of the invention is a molecular genetic method which permits the determination of one or more aspects of the Kell genotype of a subject. The method generally involves obtaining DNA or other nucleic acid from a biological sample from a human subject or patient. Typically, the method requires a blood sample, but other types of tissue samples which contain erythroid tissue are useful. The method may be employed to test any individual for Kell genotype. In a particularly preferred embodiment, however, the invention provides a method for determining Kell blood group genotype in a fetus. In this embodiment, the preferred tissue sample includes a sample of amniotic fluid. It is understood, however, that such samples can be obtained from sample libraries or tissue archives such as blood banks, or from forensic evidence, etc. Accordingly, the method may be used on a unique or irreplaceable sample or may be used to screen large numbers of samples.

In general, the method of the invention includes generating at least one characteristic nucleic acid product from a nucleic acid sample. The sample nucleic acid is obtained or derived from a subject for whom Kell genotype is desired to be determined, and preferably includes a nucleotide sequence which contains a Kell polymorphism locus, although the exact sequence or structure need not be well defined beforehand. In preferred embodiments the nucleic acid sample includes genomic DNA of the subject.

The method of the invention generates nucleic acid product, characteristic of specific Kell genotype, in an amount sufficient to permit characterization of at least one aspect of Kell genotype. Preferably, the method of the invention generates characteristic nucleic acid product which permits the differentiation of alleles of a Kell polymorphism, such as K1 and K2; K6 and K7; and K3, K4, and K21; or which permits the differentiation of K10 from the wild-type K(-10).

In one preferred embodiment, the diagnostic method of the invention includes exposing Kell DNA to a restriction enzyme which differentially cleaves or digests nucleic acid containing a Kell polymorphism locus. Typically, the restriction enzyme generates a nucleic acid product which includes at least one restriction fragment which is distinctive in size and can be associated with a specific Kell allele on a consistent basis. For example, the restriction enzyme may digest one allele of a polymorphism set while failing to digest other alleles of the set. Alternatively, the restriction enzyme cleaves alleles in the set except for one which is left intact and is recognizable on that account. Numerous restriction enzymes are known in the art, and any restriction enzyme which produces at least one distinctive or characteristic fragment in this way may be employed according to this embodiment of the method of the invention.

For K1/K2 differentiation, a preferred restriction enzyme is *BsmI* which cuts at a restriction site which, it has now unexpectedly been found, occurs in K1 DNA but not in K2 DNA. Thus, *BsmI* cleaves K1 DNA while leaving K2 DNA intact. For the K6/K7 polymorphism, preferred restriction enzymes include *MnII* and *DdeI*, which unexpectedly cleave K7 DNA but not K6 DNA. For K10, a preferred restriction enzyme is *AccI*, which unexpectedly cleaves K10 DNA, but not wild-type K(-10) DNA at that locus. In the K3/K4/K21 group, preferred restriction enzymes include *NlaIII*, which unexpectedly cleaves only K3 (Kp^a) DNA, and/or *PvuII*, which unexpectedly cleaves only K21 (Kp^c) DNA; K4 (Kp^b) DNA remains uncleaved by either of these enzymes. Combinations of these or other enzymes may be employed as desired.

In another preferred embodiment, the diagnostic method of the invention involves differentially amplifying DNA including a Kell polymorphism locus. Preferably, the Kell DNA is amplified by means of a primer which differentially amplifies a specific allele in a set of alleles associated with a Kell polymorphism. For example, by carefully selecting an appropriate amplification primer, DNA containing a K1 locus may be amplified to produce a distinctive amplified product, while DNA containing a K2 locus is not amplified to any substantial degree. Alternatively, a set of primers may be employed which produce distinctive set of amplified products which may be analyzed to determine the presence or absence of alleles. For example, a K1 specific primer may produce a product characteristic of K1 genotype and a K2 primer may produce a product characteristic of the K2 genotype. Other alleles of other Kell polymorphisms may also be differentially amplified by selecting appropriate primers.

Any combination of primers may be employed according to this embodiment provided that the amplification product(s) permits characterization of a desired aspect of Kell genotype. Preferred primer sets include: primers which differentially

amplify K1 and K2 DNA; primers which differentially amplify K6 and K7 DNA; primers which differentially amplify K10 and K(-10) DNA; primers which differentially amplify K3, K4, and K21 DNA; or combinations thereof.

5 It is among the advantages of this approach that, by using appropriate primer combinations, specific fragments of DNA are amplified to provide products in an amount sufficient to characterize Kell genotype without any requirement for digesting the DNA with restriction enzymes. Nonetheless, it
10 is within the invention to employ primers which provide differentiating sets of DNA fragments, and to supplement the differentiation of products and consequent characterization of Kell genotype by means of selectively cleaving the fragments with restriction enzymes. This may be particularly helpful in
15 situations where some fragments are not completely separable or differentiable due to limitations of technology, while selective cleavage of the fragments resolves an ambiguity.

 The diagnostic methods of the invention typically yield a nucleic acid product, such as digested DNA or RNA fragments or
20 amplified DNA or RNA products, in amounts sufficient to permit the characterization of Kell genotype. The pieces of nucleic acid permit such characterization since they carry information in their sequences and/or sizes which permits at least some of the fragments to be differentiated from the others. The
25 differentiating characteristics of the fragments may be exploited by any methods known in the art.

 For example, fragments may be separated from each other on the basis of physical characteristics. A highly preferred differentiating characteristic is molecular weight. Specific
30 and characteristic molecular weights of the fragments are directly determined by site-specific cleavage by restriction enzymes. Particular primers may generate amplification products having specific molecular weights. The fragments may be separated from one another by methods such as

polyacrylamide gel electrophoresis which relies on molecular weight as well as net molecular charge for separation. Such techniques produce informative patterns of fragments permitting determination of Kell genotype.

5 Kell nucleic acid products may also be differentiated from one another on the basis of their sequence identity. Specific probes may be employed according to known methods to hybridize specifically or uniquely with specific fragments. By the differential labeling of several probes, the presence
10 or absence of particular Kell polymorphism-specific fragments may be established. Alternatively, following separation of fragments, their detection may be enabled by marking or staining some or all of the fragments specifically or non-specifically according to known methods. Various sequence
15 non-specific marker substances are known, such as ethidium bromide.

Hybridization probes may be employed to detect the presence or absence of specific Kell nucleic acids in a testable system. Diagnostic tests based on hybridization
20 techniques can be both extremely specific and highly sensitive, providing information characterizing Kell genotype. Such techniques are described, for example, in Glick BR and Pasternak JJ, Molecular Biotechnology: Principles and Applications of Recombinant DNA, ASM Press, Washington, D.C.,
25 pp. 192-201 (1994) (Ref. 40).

In a preferred embodiment, the diagnostic method includes:

- (a) obtaining DNA from a tissue sample from a patient;
- (b) amplifying the DNA by polymerase chain reaction
30 using a primer or primer set which specifically amplifies DNA characterizing or determining a specific Kell polymorphism;
- (c) selectively cleaving the amplified DNA by a restriction enzyme which differentially cleaves Kell DNA to produce a pattern of DNA fragments; and

(d) separating the DNA fragments according to molecular weight to form a pattern of DNA fragments,

wherein the pattern of DNA fragments provides specific information characterizing the Kell genotype of the patient with regard to the specific Kell polymorphism.

In another preferred embodiment, the diagnostic method includes:

(a) obtaining DNA from a tissue sample from a patient;

(b) amplifying the DNA by polymerase chain reaction using a primer or primer set comprising at least one primer which amplifies only DNA of one allele in a Kell allelic set; and

(c) separating the amplified DNA according to molecular weight to form a pattern of DNA fragments,

wherein the pattern of DNA fragments provides specific information characterizing the genotype of the patient with respect to the specific Kell polymorphism.

In still another embodiment, the diagnostic method includes:

(a) obtaining DNA from a tissue sample from a patient;

(b) amplifying the DNA by ligase chain reaction using a primer set which amplifies only DNA of a Kell allelic set to produce a pattern of DNA fragments; and

(c) separating the DNA fragments according to molecular weight to form a pattern of DNA fragments,

wherein the pattern of DNA fragments provides specific information characterizing the genotype of the patient with respect to the Kell polymorphism.

Moreover, in another embodiment, the diagnostic method includes:

(a) obtaining DNA from a tissue sample from a patient;

(b) selectively cleaving the DNA by a restriction enzyme which differentially cleaves DNA of at least one allele in a Kell allelic set to produce a pattern of DNA fragments;

(c) amplifying the pattern of DNA fragments by polymerase chain reaction using a primer or primer set which amplifies only DNA of the allelic set; and

(d) separating the amplified DNA according to molecular weight to form a pattern of DNA fragments,

wherein the pattern of DNA fragments provides specific information characterizing the genotype of the patient with respect to the Kell polymorphism.

In yet another embodiment, the invention provides a diagnostic method which includes:

(a) obtaining DNA from a tissue sample from a patient;

(b) exposing the DNA to a restriction enzyme which differentially cleaves K1 DNA and K2 DNA to produce a pattern of DNA fragments;

(c) separating the DNA fragments according to molecular weight to form a pattern of DNA fragments; and

(d) determining the presence of K1 DNA and/or K2 DNA using detectably labeled Kell cDNA probes;

wherein the pattern of DNA fragments provides specific information characterizing the K1/K2 genotype of the patient. Such methods include, for example, Southern blot methods such as are known in the art.

In an alternative embodiment, the invention provides a diagnostic method for detecting a target nucleic acid which is specific to K1 or K2. In this embodiment, the method includes:

(a) obtaining a nucleic acid fraction from a tissue sample from a patient;

(b) ascertaining the presence in the nucleic acid fraction of a target nucleic acid encoding at least a part of the Kell protein and including the site characterizing the K1/K2 polymorphism. In this method, the presence of the target nucleic acid is ascertained by means of a probe nucleic acid which includes a nucleic acid sequence which is known to specifically bind to or hybridize with K1 DNA or K2 DNA or

transcripts thereof. The probe nucleic acids useful in this method may be any of the Kell-based oligomers described herein, including those which are detectably labeled or attached to a substrate. This method may be used to detect
5 target nucleic acids which are chromosomal or genomic DNA, mRNA, and cDNA, as well as other related forms of nucleic acid encoding the K1/K2 domain of the Kell protein. Such methods include, for example, dot blot methods such as are known in the art.

10 In still another embodiment, the invention provides a method for the determination of Kell blood group genotype. In this embodiment, the method includes the steps of:

(a) selecting a probe nucleic acid sequence
substantially corresponding to at least part of a known Kell
15 exon or a transcript thereof, wherein the known Kell exon comprises a Kell polymorphism locus and codes for a specific Kell allele;

(b) contacting a sample nucleic acid sequence obtained or derived from a subject, the subject having an unknown
20 phenotype with respect to the Kell polymorphism locus, with the probe nucleic acid sequence under conditions which permit hybridization when the nucleic acid sequences are significantly complementary; and

(c) measuring an amount of hybridization between the
25 probe nucleic acid sequence and the sample nucleic acid sequence.

In this embodiment, the detection of an amount of hybridization corresponding to an amount resulting from significantly complementary sequences indicates that the
30 subject possesses the specific Kell allele under investigation. On the other hand, detection of an abnormally low amount of hybridization indicates that the subject lacks the specific Kell allele being investigated.

The invention further provides expression recombinant vectors which carry Kell nucleic acid sequences. Such vectors permit transformation of cells, particularly eukaryotic cells such as yeast and human cells, to cause such cells to express a heterologous protein product. The invention provides expression vectors which include a nucleic acid sequence which encodes at least a part of the Kell protein including a part of the protein which encodes a site of Kell polymorphism. In particular, the invention provides expression vectors which carry K1 cDNA permitting transformation of cells to produce transformed cells or transformants which express K1 protein on their cell surfaces. Vectors are also provided which carry K6 cDNA, as well as vectors which carry K10 cDNA, and vectors which carry K3 or K21 cDNA, permitting transformation of cells to express K6, as well as K10, and K3 and K21, respectively. The invention also provides stable cell lines which have been modified (i.e., transformed) to express protein on its cell surface, as well as a method for transforming a cell line to express such protein. With respect to the K1 expression, such protein preferably includes at least the K1 domain, and more preferably, the protein is K1 protein. Similarly, with respect to expression of K6, K10, as well as K3 and K21, the expressed protein preferably includes the K6, the K10, the K3 or the K21, domain, respectively; more preferably, the protein is K6, K10, K3, or K21 protein, respectively. Methods for preparing expression recombinant vectors, and for transforming cell lines using such vectors are known in the art, and are described generally in Glick BR, and Pasternak JJ, Molecular Biotechnology: Principles and Applications of Recombinant DNA ASM Press, Washington, D.C. (1994) Chapter 5 (Ref. 40).

The Kell genotype determination method of the invention is of particular utility in the determination of the fetal Kell genotype. The molecular genetic method described herein can easily be applied to DNA samples obtained from amniotic fetal cells and is useful in determining the Kell genotype of

the fetus. The methods of the invention are also useful in a variety of other situations in which molecular genetic information about a person is desired. For example, the methods of the invention are useful in situations in which it is desired to obtain information concerning the identity of an individual from forensic samples. Alternatively, the methods of the invention are useful for obtaining genetic information enabling the determination of paternity in those situations in which paternity is in doubt or dispute. In addition, the methods of the invention are useful for the determining the Kell genotype of a recipient of a blood transfusion, as well as for the screening of stored blood for Kell genotype, to avoid transfusion incompatibility. The peptide based methods of the invention are particularly useful in determining alloimmunization of a person to a Kell antigen, such as K1. Information about such alloimmunization would be expected to be useful in advising women about risk entailed in future pregnancies. Other applications of the methods of the invention will suggest themselves to the skilled artisan.

In addition, the genotype characterization method of the invention may be performed selectively so as to permit description of some or all of the Kell family of alleles. The method may also be performed in conjunction with the determination of phenotype of some or all of the Kell family of antigens. Thus, it is possible to confirm Kell genotype by employing serological testing in conjunction with genomic analysis using a restriction enzyme. Such serological testing steps may be performed either prior to, in parallel with, or after the prescribed genomic testing. Moreover, the planning of testing procedures may be based on preliminary partial results which rule out or imply the possibility of, for example, certain rare genotypes. It may also be desirable to perform genotype testing for other markers or blood groups together with the Rh genotyping method of the invention. For example, the genotyping method described in U.S. Patent No. 5,723,293

relating to the determination of Rh genotype, may be performed in conjunction with the method of the invention to provide counsel as to a broader spectrum of possible difficulties in pregnancy. In general, the diagnostic method, compositions and kits of the invention are expected to be employed in conjunction with a wide variety of known medical tests and assays to enhance and complement the information available to the clinician concerning a patient. See e.g., Walker RH et al., eds., Technical Manual, 10th ed., American Association of Blood Banks, Arlington, Virginia (1990) (Ref. 41).

The following examples are intended to assist in a further understanding of the invention. The particular materials and conditions employed are intended to be further illustrative of the invention and are not limiting upon the reasonable scope thereof.

For each of the examples described herein, the molecular biology techniques employed were performed generally in accordance with methods accepted in the art. See, for example, Sambrook et al. (Ref. 42), and Innis et al. (Ref. 34).

Example 1

The organization of the KEL gene was determined using the following procedure.

25 Screening of a Genomic Library

A human placental genomic DNA library constructed in EMBL3 Sp6/T7 was obtained from Clontech, Inc. (Palo Alto, CA). The genomic DNA library was constructed by partially digesting human placenta genomic DNA with *Sau3A*I and ligating the fragments into the *Bam*HI site of the vector. For screening, the λ phage library was grown in either NM528 or LE392 strain of *E. coli* and plated. The DNA was lifted on HybondTM-N+

membranes (Amersham Co., Arlington Heights, IL) and hybridized with Kell cDNA probes according to standard procedures (Ref. 43). The cDNA probes were labeled with ^{32}P by random primer extension using a commercial kit (Boehringer Mannheim Inc., Indianapolis, IN). The specific activity of the probes was approximately 1×10^8 cpm/ μg . Fourteen positive clones were isolated.

Characterization of Exons and Introns

Of the 14 positive clones identified, five ($\lambda 2$, $\lambda 8$, λB , λF , and λE) were selected for further characterization. Two of the clones ($\lambda 2$ and λE) span most of the Kell gene and cover about 21.5 Kb (see Figure 1). Clone $\lambda 8$ was studied because it extended more of the 5' flanking region and also served to reconfirm the sequences determined from clone $\lambda 2$. Clones λB and λF were used because they overlapped with λE , and in addition a small segment of λF overlapped with the 3' end of $\lambda 2$ (see Figure 1).

These five clones were initially mapped by digestion with *XhoI* plus *PstI* and with *XhoI* plus *EcoRI*, followed by Southern blot hybridization using different oligonucleotides specific for various locations in the Kell cDNA sequence. The genomic clones were then digested with the following enzymes: *PstI*, *EcoRI*, *BamHI*, *BglII* and *XhoI*, either individually or in combination. The individual gene fragments were subcloned into PUC18 (Gibco BRL, Gaithersburg, MD) or pGEM-3Zf(+) (Promega Co.) and sequenced on an automated 373A DNA sequencer (Applied Biosystems, Inc., Foster City, CA). The exons, determined in relation to the Kell cDNA sequence, and their flanking regions, were fully sequenced, as were some short introns. With one exception, the long introns were sized by PCR using primers from the flanking regions of the introns. The longest intron, between exons 10 and 11, was partially sized using primers from known intron sequences obtained from $\lambda 2$ and λE , using a YAC clone as template. (Seven yeast artificial chromosome (YAC) clones were separately isolated by

PCR specific for KEL exon 6 from a YAC resource enriched for chromosome 7 DNA.) This area of the intron was confirmed by PCR using human genomic DNA as template.

Restriction Enzyme Mapping and Southern Blot Analysis of the Genomic Clones

Subclones were mapped with *Bam*HI, *Eco*RI and *Pst*I. The different sizes of digested DNA were resolved by agarose gel electrophoresis. The DNA fragments were in some cases further analyzed by Southern blots using ³²P-labeled oligonucleotides derived from different locations of Kell cDNA and also identified with available sequences of the subcloned gene fragments. The combination of these procedures allowed a restriction map of *Bam*HI, *Eco*RI and *Pst*I to be constructed. Sequence analysis of the subcloned Kell gene fragments from these 5 genomic clones revealed that the human KEL gene contains 19 exons ranging in size from 63 bp (exon 7) to 288 bp (exon 19) which includes a 3' untranslated region.

Figure 1 shows the newly determined structure of the human Kell gene (*KEL*) with restriction enzyme sites. The position of exons are shown as dark boxes and the introns are shown as lines connecting the exons. The vertical lines of different heights mark the restriction enzyme sites in the Kell gene. The λ phase, genomic clones are shown at the bottom of the diagram (λ 2, λ 8, λ E, λ B and λ F). Two of the genomic clones from the commercial genomic library, clones λ B and λ F, gave inconsistent sequences when compared to other clones (λ 2 and λ 8). The 5' segments of clones λ B and λ F should be identical to the 3' segments of λ 2 and λ 8 but sequence analysis showed that they were different. The reason for this is not known, but could be due to ligation of DNA material during the preparation of the genomic library. These areas of difference are depicted in Figure 1 by broken lines. The restriction enzyme digested products which were subcloned and analyzed are shown above the genomic clones. Each sub-clone shows the restriction enzymes used and the genomic

clones from which it was derived. Four PCR-derived sub-clones (PCR25, PCR6, PCR1 and PCR20) are also shown. In parenthesis each PCR product shows the genomic clone from which it was amplified. A small segment of the large intron not covered by clones $\lambda 2$, $\lambda 8$ and λE was determined by sequencing PCR products (PCR6) from a YAC clone (yWSS679) and from the genomic DNA. This strategy was employed even though the λF clone covers the 3' end of the $\lambda 2$ clone. This allowed us to cover the entire Kell gene and to size the intron between exons 10 and 11 without ambiguity.

All exon/intron splice junctions were found to contain the 5' donor -gt and the 3' acceptor -ag sequences. The introns ranged in size from 93 bp to approximately 6 kb. there were 6 introns which were longer than 1 kb (Figs. 1 and 2). The long introns were not fully sequenced, but were sized by PCR. There was ambiguity in analyzing the intron between exons 10 and 11 because the 5' region of clones λB and λF , which were expected to overlap with the 3' region of clones $\lambda 2$ and $\lambda 8$, did not. These ambiguous areas of clones λB and λF are shown as dotted lines in Figure 1. Because of this uncertainty, the small gap of the gene not covered by clones $\lambda 2$ and λE was bridged by PCR amplification of YAC clone yWSS679 containing the Kell gene (Figure 1). The size of this PCR amplified region was further confirmed by PCR of genomic DNA obtained from a person of common Kell phenotype using the same primers used for the YAC clone.

All of the exons were sequenced and compared to that of a full length Kell cDNA isolated from a human bone marrow library obtained from Clontech Laboratories, Inc. The Kell phenotype of this library was unknown. Differences were noted in 4 bases in exon 3 as compared with the published sequence for Kell cDNA (Ref. 14). These differences were due to sequencing errors in the original study. The corrected sequences were submitted to EMBL/GenBank® Updates (National Center for Biotechnology Information, Bethesda, MD) under

accession No. M64934. One notable difference is a base substitution in the membrane-spanning region, which encodes a leucine instead of proline. The corrected base sequences are shown in bold type in Figure 2.

5 Figures 2A-2D illustrate the sequence of the individual exons encoding human Kell protein with the immediate intron flanking splice junction sequences. The base sequences of the individual exons are shown in capital letters and the flanking intron sequences in small letters. The amino acid sequences
10 encoded by the exons are shown above the base sequences. The numbers on the left-side, below the indicated exons, refer to the base numbers from the cDNA. The other numbers on the left and right sides indicate the amino acid residues and the bases, as previously described for the cDNA (Ref. 25). Intron
15 sizes are specified at the end of the exons. 5' and 3' splice sites are also shown.

 Since the Kell phenotype of the person from whom the genomic DNA library was constructed is unknown, we isolated RNA from peripheral blood of a person of known common
20 phenotype (K:-1,2,-3,4,-6,7). cDNA was prepared by RNA-PCR and sequenced according to methods known in the art. The deduced amino acid sequence was identical to that shown in Figure 2 as obtained from the Clontech genomic library. In the person of known phenotype, C to T base differences
25 occurred in two locations (nt 1656 and 1664), but these substitutions did not change the predicted amino acids.

 Of interest is that exon 1 includes the 5' untranslated region and codes for only initiation methionine. The single membrane spanning region is located in exon 3 and the
30 pentameric sequence (HELLH) which conforms with a consensus sequence (HEXXH) found in the active sites of zinc neutral endopeptidases (Ref. 43) is in exon 16 (Figure 2).

The coding region of the native protein is present in 18 of the 19 exons; the first exon contains the 5' untranslated region and the initiation methionine. Other examples having only the initiation codon in the first exon are known (Ref. 44). The single transmembrane region is encoded in exon 3 with exons 4 to 19 encoding most of the extracellular portion. The HEXXH motif, which is unique to zinc metallopeptidases (Ref. 29) is encoded in a 68 bp exon (exon 16). In addition to the consensus pentameric sequence, Kell has sequence and structural homology with a family of membrane-associated zinc neutral endopeptidases (EC24.11) (Ref. 25) of which the enkephalinases and CALLA are examples (Refs. 45-46). The CALLA gene is larger than the Kell gene, about 80kb, and is composed of 24 exons (Ref. 47). In CALLA the putative enzyme active site is encoded in exon 19 and a comparison of base sequences of exons 18 and 19 of CALLA with exons 15 and 16 of Kell shows 54.5% base identity.

Example 2

Transcription Initiation Site

The transcription initiation site of the *KEL* gene was determined by the rapid amplification of cDNA ends (5' RACE) employing a human fetal library (5' RACE Ready™ cDNA library, Clontech laboratories, Palo Alto, CA). The source of poly (A)⁺ RNA was normal liver pooled from 2 female Caucasian fetuses, 22 and 26 weeks of gestation. The single stranded cDNA library was made by reverse transcription, using random hexamers as primers. An oligonucleotide anchor 3'-NH₃-GGA GAC TTC CAA GGT CTT AGC TAT CAC TTA AGC AC-p-5' (SEQ ID NO:6) was ligated to the cDNA. For the 5' RACE, 2 sets of nested PCR reactions were carried out, using primers from 2 different locations on Kell cDNA (nt 132 and nt 178). In the first PCR, 5' RACE Ready™ cDNA was used as template. The primers used for one set (PCR1) included the anchor primer provided by Clontech (5'-CTG GTT CGG CCC ACC TCT GAA GGT TCC AGA ATC GAT AG-3') (SEQ ID NO:7) and Kell anti-sense primer (5'-CTC GGC TCT TCC TCA CTT TGG TCC-3', nt 132) (SEQ ID NO:8). For the

other set (PCR2), the primers included the Clontech anchor primer (SEQ ID NO:7) and Kell anti-sense primer (5'-CTC TTG GCT CCA GAG AGT TCC CAT-3', nt 178) (SEQ ID NO:9). For the second nested PCR the products of the first PCR reactions were used as template and the Clontech anchor primer (SEQ ID NO:7) was again used as sense primer for both parallel reactions. In PCR1 a Kell anti-sense primer (5'-CCC ACC TTC CAT CTG TCT ATC TTC-3', nt 109) (SEQ ID NO:10) was used. For PCR2, the Kell anti-sense, nt 132, (SEQ ID NO:8) was used.

PCR was performed in an automated thermocycler (Minicycler[®], MJ Research Inc. Watertown, MA) using an initial cycle of 94°C for 3 minutes, 62°C for 1 minute, and 72°C for 30 seconds. In cycles 2 to 30 the conditions were 94°C for 30 seconds, 62°C for 30 seconds and 72°C for 30 seconds. In the last cycle the polymerization step at 72°C was extended to 10 minutes. The final concentrations of reagents were 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 3 mM MgCl₂, 400 nM of each primer, 200 μM of each dNTP, 0.1% Triton[®]-X100, and 2.5 units of *tag* polymerase, in a final volume of 50 μL. The "hot start" method using Ampliwax[™] from Perkin Elmer (Branchburg, NJ) was employed. The PCR products were separated by electrophoresis in 0.8% low melting agarose gels, eluted and directly ligated to pT7-Blue(R) plasmid vector from Novagen (Madison, WI) and transformed in DH5αF' strain of *E. coli*.

Each primer yielded 3 different size products. These products were subcloned and sequenced. The largest product from each PCR reaction had a 5' end at 120 bp upstream from the initiation codon. The 2 shorter PCR products ended at 81 and 30 bp upstream from the initiation codon.

Figure 3 illustrates the nucleotide sequence of the KEL 5' flanking region showing exon 1 and possible cis regulatory elements. A 185 bp region upstream from the probable cap site is shown. The three possible transcription initiation sites are marked by ∇. Consensus sequences for GATA-1, Sp1 and a

CACCC region are boxed. The region -176 to -1, which was copied by PCR and placed in a CAT-expression vector is shown. The initiation methionine is in bold letters.

Three possible transcription initiation sites were found using the 5' RACE procedure and poly (A)⁺ RNA from fetal liver. The first of these cap sites, located 120 bp upstream from the initiation ATG, is also the 5' end of a cDNA cloned from a human bone cDNA library (Ref. 25) the other 2 sites are 81 and 30 bp upstream from the initiation codon. All 3 sites were obtained using 2 different Kell cDNA anti-sense primers. Although all 3 locations are purine bases and could act as transcription initiation sites. It is also possible that those at 81 and 30 bp upstream from the ATG are artifactual due to incomplete reverse transcription. This is unlikely given that random hexamers were used to prepare the cDNA library. However, secondary RNA structures can also cause premature termination of reverse transcription.

Example 3

Analysis of 5' Flanking Region

The constitution of the 5' flanking region was obtained by DNA sequencing of sub-clone λ 8, following digestion with *Pst*I (see Figure 1). The 5' flanking region was determined to span nucleotides -176 to -1.

Figure 5 shows a 185 bp sequence upstream from the first possible initiation transcription site. Analysis of this region, and of exon 1, indicates that there are no TATA sequences but several other possible transcription factor binding sites were noted. At least 2 GATA-1 sites are present close to a CACCC box. Sp1 and GATA-1 sequences are present in exon 1. The 5' flanking region contains purine-rich regions. These areas of interest are shown in Figure 3.

The transcriptional activity of the 5' flanking region from -176 to -1 (Figure 3) was determined by CAT assay in transfected K562 cells as compared to the promoter-less pCAT vector. In this procedure, a PCR product spanning nucleotides -176 to -1 relative to the first cap site was subcloned into pCAT basic vector obtained from Promega Co. Cells of the erythroleukemic cell line K562 were then transfected using the lipofectin® method. Construction of the CAT-vector, transfection and chloramphenicol acetyltransferase (CAT) activity were assayed following the protocol provided by Promega Co. CAT enzymatic activity was measured using [¹⁴C]-chloramphenicol (50-60 mCi/mmol, Amersham Co., Arlington Heights, IL). The reaction products were measured by liquid scintillation spectrometry and analyzing the reaction products by thin layer chromatography. As a negative control the pCAT basic vector without promoter was used. All cell extracts were normalized by protein analysis in order to compare values.

Introduced in front of a CAT reporter gene, the 5' flanking region exhibited promoter activity in the erythroleukemic cell line K562. pCAT vector with the 5' flanking region was found to express approximately 0.8 units of CAT activity per milligram of cell extract protein. A unit of CAT enzyme activity is defined as the amount of enzyme required to transfer 1 nmol of acetate of chloramphenicol in 1 minute at 37°C. Figure 4 shows CAT activity of the 5' flanking region. An autoradiogram, exposed for 48 hours, is shown. Lane 1 has pCAT vector with the 5' flanking region and lane 2 has pCAT basic vector without promoter. The radioactive butyrylated chloramphenicol is indicated as "bCm", and the unreacted chloramphenicol is indicated as "Cm".

The putative promoter region does not contain the typical TATA box usually located -25 to 30 nt relative to the cap site (Ref. 48). However, as in several erythroid specific genes, consensus sequences for GATA-1 factor were found in the

promoter region. In addition, possible Sp1 and GATA-1 binding sites were noted in exon 1. It is not known whether GATA-1 and Sp1 regulate KEL gene expression but GATA-1 is common in erythroid genes (Refs. 49-55) and is known to be expressed at low levels in hematopoietic progenitor cells and up-regulated during erythroid maturation (Refs. 56-58). If these transcription factors define erythroid tissue specificity it will be in agreement with our Northern blot studies which detected Kell transcripts in bone marrow and fetal liver but not in several non-erythroid tissues (Ref. 8).

Example 4

Analysis of the 3' End

Exon 19 is the largest Kell exon, encoding the C-terminal 53 amino acids and containing the 3' untranslated region with a polyadenylation signal 100 bp downstream of the termination codon. Previous Northern blot analysis showed that the major Kell transcript in bone marrow and fetal liver is 2.5 kb although smaller amounts of larger mRNAs, notably 6.6 kb, also were observed (Ref. 8). In originally cloning the Kell cDNA from a human bone marrow library, we isolated a cDNA with a large (3 kb) 3' untranslated region (Ref. 25).

To determine the 3' end structure of the Kell gene, RNA was isolated from peripheral blood as described by Goosens et al. (Ref. 59) and cDNA was prepared by reverse transcription and PCR amplification using a Perkin Elmer RNA PCR kit (Roche Molecular Systems, Inc., Branchburg, NJ). First strand synthesis was initiated using an anchored oligo d(T)₁₆ primer. PCR amplification of the 3' end of Kell cDNA was performed using the anchor primer and an oligonucleotide primer from the coding sequence of Kell cDNA. The anchor antisense primer used was oligo 5'-GACTCGAGTCGACAACGTT(T)₁₆-3' (SEQ ID NO:11) and the sense primer from the 3' end coding sequence of Kell cDNA was 5'-ATGGGGAGACTGTCCTG-3' (SEQ ID NO:12).

A PCR product of about 400 bp was obtained and was sub-cloned and sequenced. The 3' end sequences of different sub-clones are shown in Figure 5. The base sequences, prior to the poly A region, of cDNA clones from a human bone marrow library (top) and from peripheral blood of a person with common Kell phenotype (bottom) are shown. (A) indicates the poly A region.

The 3' end sequences of four different subclones, obtained from peripheral blood are shown in the bottom portion of Figure 5. All four of the subclones had identical sequences from the termination codon to the polyadenylation signal (AATAAA). At the 3' end, the base sequences differ slightly in length before the start of the poly A sequence. The distance between the polyadenylation signal and the cleavage site is known to be variable (Ref. 60). Similar variations in 3' end sequences were observed in 3 different Kell cDNAs obtained from a human bone marrow cDNA library (Clones 185, 182 and 190).

Only one of the cDNA clones, the original full-length cDNA obtained from the bone marrow library (Clone 191), contained a large 3' untranslated region. The 3 kb untranslated fragment from this cDNA clone does not hybridize with human genomic DNA indicating that it is foreign DNA. This foreign fragment is preceded by an *EcoRI* site and a possible linker and may have been artificially inserted during the preparation of the library.

Sequences of the 3' segments of transcripts in peripheral blood by RNA PCR only detected short 3' untranslated regions, which varied slightly in length before the poly A tail. Similar sequences were obtained when other cDNAs from the human bone marrow library were analyzed. The larger Kell transcripts could not be amplified using the RNA PCR method because of their length, but Northern analysis indicates their presence. The occurrence of multiple transcripts with larger

3' untranslated region is not uncommon and although their function is not well understood, the 3' untranslated regions are thought to play roles in regulation of expression (Refs. 61-62).

5

Examples 5-7

DNA Preparation

In Examples 5-7, DNA was prepared from peripheral blood obtained from persons whose Kell phenotypes have been determined serologically. DNA was prepared either from 1 to 5
10 ml of whole blood collected with anticoagulants or, when red cells were removed by centrifugation at $1000 \times g$ for 10 mins, from the resulting buffy coat. In both cases, the procedure described by John et al. (Ref. 63) to prepare DNA was used. Some of the DNA samples were obtained from Canadian Hutterites
15 and serological studies of these families indicated that they were homozygous for either K1 or K2. In these samples DNA was isolated as described by Zelinski et al. (Ref. 30).

Polymerase Chain Reaction

For Examples 5-7, the polymerase chain reaction was
20 performed as follows: Denaturation, annealing and polymerization steps were performed in an automated thermocycler (Minicycler[®], MJ Research Inc., Watertown, MA). An initial cycle of 94°C for 3 min., 62°C for 1 min., and 72°C for 30 seconds. In cycles 2 through 30 the conditions were
25 94°C for 30 seconds, 62°C for 30 seconds, and 72°C for 30 seconds. In the last cycle the polymerization step at 72°C was extended to 10 minutes. The final concentrations of reagents were 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 3 mM MgCl₂, 350 nM of each primer, 200 μ M of each dNTP, 0.1% Triton[®]-X100,
30 100-200 ng genomic DNA and 2.5 units of *tag* polymerase in a final volume of 100 μ L. The "hot start" method using Ampliwax[™] from Perkin Elmer (Branchburg, NJ) was employed. The amplified PCR products were separated by electrophoresis on 0.8% agarose gels and detected by ethidium bromide
35 staining.

Restriction Enzyme Digestion

BsmI was added directly to the final PCR reaction mixture. The PCR mixture (10 μ L) was optimized by *BsmI* digestion by adding 2 μ L of 35 mM $MgCl_2$, 1 μ L of 10 mM mercaptoethanol or 10 mM DTT, 1 μ L of 10X *BsmI* buffer, 4 μ L of water, and 2 μ L of 5 units/mL *BsmI*. The mixture were incubated for 90 min. at 65°C. The DNA in the reaction mixture was analyzed by electrophoresis in 0.8% agarose.

Other Reagents

Taq DNA polymerase and dNTPs were purchased from Promega Co. (Madison, WI). X-Gal was purchased from Appligene Inc. (Pleasanton, CA). The DNA 1 kb ladder standards, DH5 α F' strain *E. coli* competent cells, and the low melting agarose were purchased from Bethesda Research Laboratories (Gaithersburg, MD). T4 DNA ligase and *BsmI* were purchased from New England Biolabs (Beverly, MA). pT7 blue (R) plasmid vector was purchased from Novagen (Madison, WI). Quick Spin[™] Column (G-50 Sephadex[®]) for DNA purification was purchased from Boehringer Mannheim, Inc. (Indianapolis, IN).

Example 5Comparison of K1 and K2 DNA Sequences

Nine pairs of forward and reverse primers were used to amplify the 19 KEL exons. These primer pairs were selected to cover the open reading frames of the exons. The sequences of the primers and the target exons (spanning regions) are shown in Table 1. Genomic DNA from a homozygous K1 person was used as template DNA. The PCR products were separated by electrophoresis on 0.8% agarose gels and stained with ethidium bromide. In all cases single products ranging in size from 0.48 to 1.5 kb were obtained. Figure 6 illustrates the PCR amplification of the KEL exons. Lanes 1 and 11 in Figure 6 are 1 kb DNA ladder standards. Lanes 2 to 10 contain PCR products of the primer pairs PCR 1, 2, 3, 4, 5, 6, 7, 8 and 9, in that order.

Table 1. Primers Used in PCR of KEL Exons

PCR	Primers	SEQ ID NOS:	Target Exons	Product (kb)
1	5'-CAG TCC TCC GAA TCA GCT CCT AGA-3'	13	1*	0.48
	5'-CTC TTG GCT CCA GAG AGT TCC CAT-3'	14		
2	5'-GAA GGT GGG GAC CAA AGT GAG GAA-3'	15	2*, 3, 4	1.0
	5'-ACA GGG TTT GGA GCA GTC ATG GTC-3'	16		
3	5'-TTT AGT CCT CAC TCC CAT GCT TCC-3'	4	5, 6	0.74
	5'-TAT CAC ACA GGT GTC CTC TCT TCC-3'	5		
4	5'-ATA TTC CCC ACC TCC CCA CAC CTG-3'	17	7, 8, 9	0.8
	5'-ATC TAC GGT GCT CAG GCT CTC CTC-3'	18		
5	5'-GGA AGC ATG GGA GTG AGG ACT AAA-3'	19	10	1.2
	5'-TGG CAT CCA TGG TAC CTC ATG GAA-3'	20		
6	5'-GAG GCT TTT GAA ACC CCA GGA TGA-3'	21	11	1.3
	5'-TTC CCC AGC CAC CTG CCA TCT CAT-3'	22		
7	5'-CCC TTT TCC AAG GGT CAG AAG CTG-3'	23	12, 13, 14, 15	1.4
	5'-GGG CTT ATT TGA CCC CCA GAA TCT-3'	24		
8	5'-CCT AAT CCC TGG ATG CCT GCC TGT-3'	25	14*, 15, 16, 17, 18	1.5
	5'-CAG TGA GGA CAT CTG CAG AAG AGG-3'	26		
9	5'-TCC TGT GGA CCC TCC CCC TTC AAT-3'	27	19*	0.33
	5'-GGG CGG AAG CCA AGT GCC AGC TTT T-3'	28		

The first primer is the forward sequence and the second is the antisense primer. The 2G bases at the 5' end of antisense primer of PCR are not in the gene and were added to create an SmaI site for another purpose. Some PCR products did not cover the entire exons and those exons are marked with asterisks. The PCR products of exons 1 (PCR1) and 19 (PCR 9) contain the entire translated regions. Exon 2 was spanned by overlapping PCR products (PCR 1 & 2) and exon 14 was covered by PCR 7.

Example 6

The PCR products obtained in Example 1 were sequenced using the following protocol: The PCR products, separated by electrophoresis on low 2.8% melting agarose, were eluted, ligated to pT7 Blue(R) plasmid vector and transformed in DH5 α F' strain of *E. coli*. Plasmid DNA was prepared on a small scale by the alkali lysis method and purified with a Quick Spin™ (Sephadex G50) column. Standard molecular biology procedures were employed, such as are described in detail in Sambrook et al. (Ref. 42). DNA sequencing was performed by an automated system (Applied Biosystems, Model 373A, Version 1.2.0).

The sequenced products were compared to the previously described sequence of K2 cDNA (Ref. 25). It was unexpectedly found that the sequence of K1 DNA encodes an amino acid sequence identical to that of K2 DNA, except for a single base change. This difference between K1 and K2 occurs as a shift from C to T at nucleotide 701 in exon 6.

The PCR product PCR3 (Table 1), which spans exons 5 and 6, had a single base difference when compared to K2 DNA. The base sequences of a portion of exon 6, and the amino acids which they encode, are shown in Figure 7. The sequence of K2 DNA is on the top and K1 DNA on the bottom. In the PCR3 product, which was 740 bp in length, there was a single cytosine (C) to thymine (T) substitution, corresponding to position 701 of the Kell cDNA. The C to T substitution is marked in bold letters and is highlighted with an arrow. This change predicts a threonine to methionine change at a consensus N-glycosylation site (Asn.X.Thr). An N-linked glycosylation motif in K2, and the disrupted motif in K1, are underlined. Since this single base change was the only difference between K1 and K2 encoding a change in an amino acid, this observation suggested that the threonine to

methionine change (Thr → Met) at position 193 would prevent N-glycosylation at asparagine 191 in proteins expressing K1, thus identifying the K1/K2 polymorphism.

Example 7

5 Confirmation of K1/K2 Polymorphism by BsmI Analysis

10 The C to T substitution at nucleotide position 701 (nt 701) in exon 6 creates a new sequence including the sequence 5'-GAATGCT-3', which is known to define a restriction enzyme site specific for BsmI. As a result, treatment of the 740 bp PCR product which spans exons 5 and 6 (PCR3) with BsmI was hypothesized to provide a means by which to differentiate K1/K2 genotypes. In a PCR method employing digestion with BsmI, the K1/K1 genotype should yield 2 fragments, of 540 and 200 bp; the K2/K2 genotype should yield the uncut 740 bp PCR product and K1/K2 heterozygotes should yield 3 fragments of 15 740, 540 and 200 bp. To confirm that this base change does in fact identify the K1 genotype, this region was analyzed in 42 different persons of known K1/K2 phenotype.

20 The PCR3 primer pair (Table 1) was used to generate PCR products (740 bp) by PCR amplification of DNA obtained from various Kell phenotypes. After amplification, the samples were treated with BsmI and separated by electrophoresis on 0.8% agarose gels.

Figure 8 illustrates the results obtained in the above-described experiment. The gel lanes are defined as follows:

<u>Lane</u>	<u>Phenotype</u>	<u>Lane</u>	<u>Phenotype</u>
1	DNA Standard	11	K:-1,2,3,-4
5 2	Untreated PCR3	12	K:-1,2,6,-7
3	K:1,-2	13	McLeod
4	K:1,-2	14	K:-1,2,10
5	K:-1,2	15	K:-1,2,14,24
6	K:1,2	16	K:1,2
10 7	K:1,2	17	K:1,-2
8	K:-1,2	18	K:-1,2
9	K ₀	19	K:-1,2
10	K:1,2	20	DNA Standard

Lanes 1 and 20 contain 1 kb ladder DNA standards. Lane 2 is untreated PCR3 from K1 homozygotes. Lanes 3 to 19 are treated with *BsmI*. All K2 homozygote (K:-1,2) samples gave uncut 740 bp products. K1 homozygotes (K:1,-2) yielded only 540 and 200 bp fragments. K1/K2 heterozygotes (K:1,2) had 3 bands, the uncut 740 and the smaller 540 and 200 bp products.

Of the 42 DNA samples tested, 12 were either K1 or K2 homozygotes, 6 were K1/K2 heterozygotes and 24 were K:-1,2 phenotypes, but contained low prevalence or rare Kell phenotypes. These included K3, K6, K10, K₀, K14/K24 heterozygote, and a McLeod phenotype. In 40 of the 42 cases the *BsmI* genotyping agreed with the Kell phenotypes which were determined by serological analysis of red cells. In two cases, genotyping identified one of the samples as K:-1,2 homozygote and the other as a K:1,2 heterozygote. These two samples were, however, serologically identified as having "weak" K1 phenotypes.

None of the other low prevalence or rare Kell phenotypes listed above had the C to T base substitution in exon 6, indicating that this change is specific for the K1/K2 polymorphism.

Example 8

In a method of differentially determining K1/K2 genotype, the polymerase chain reaction was employed to test samples of genomic DNA using a unique primer mixture. The primer mixture included the following primers:

MK1R	ATA CTG ACT CAT CAG AAG TTT CAG CA	(SEQ ID NO:1)
MK2F	TGG ACT TCC TTA AAC TTT AAC TGA AC	(SEQ ID NO:3)
EI5F	TTT AGT CCT CAC TCC CAT GCT TCC	(SEQ ID NO:4)
EI6R	TAT CAC ACA GGT GTC CTC TCT TCC	(SEQ ID NO:5)

The MK1R primer is specific for K1 DNA, producing a PCR product 540 bp in length. The MK2F primer is specific for K2 DNA, and produces a PCR product 240 bp in length. The other primers EI5F and EI6R are K1/K2 specific, producing PCR products 740 bp in length, and are used as an internal control.

The primer mix included the primers in the following concentrations: 20 ng/ μ L MK1R; 20 ng/ μ L MK2F; 30 ng/ μ L EI5F; and 30 ng/ μ L EI6R. EI5F and EI6R primers are used in the PCR for two reactions, while the MK1R and MK2F primers are each used for only one reaction. The concentrations of the primers were adjusted to ensure sufficient quantities of the EI5F and EI6R primers for the duration of the reaction.

Blood samples were obtained from human volunteers. DNA was isolated from leukocytes present in the buffy coat of blood samples using the method described elsewhere herein.

Reaction tubes were set up including one Ampliwax™ and 25 μ L of a first reagent cocktail including: 2.5 μ L 10X buffer (Promega Co., Madison, WI); 4 μ L dNTP; 3 μ L 25 mM MgCl₂; 7 μ L primer mix; and 8.5 μ L H₂O, to give a preliminary reaction mixture volume of 25 μ L. These preliminary mixtures were incubated at 80°C for 5 min. and cooled to room temperature for 5 min. before proceeding. To each of the tubes were added 23

μ L of a second reagent cocktail including: 2.5 μ L 10X buffer; 3 μ L 25 mM $MgCl_2$; 17 μ L H_2O ; and 0.5 μ L *taq* DNA polymerase (Promega Co., 5 units/ μ L) and 100 ng of isolated genomic DNA in 2 μ L H_2O .

5 The PCR reaction was performed using the following cycling program. The initial cycle comprised 94°C for 3 minutes, 62°C for 1 minute, and 72°C for 30 seconds. Cycles 2-30 each included 94°C for 30 seconds, 62°C for 30 seconds, and 72°C for 30 seconds. In the last cycle, the
10 polymerization step was 72°C was extended to 10 minutes.

 Following completion of the PCR, the amplified products were resolved using ethidium bromide/agarose gel electrophoresis as described in the art.

 The results are of this procedure are shown in Figure 9.
15 Lanes 1 and 11 are identical control samples including mixed DNA fragments of known size. Lanes 2-4 represent internal control samples from subjects in whom the K1/K2 genotype was already known. Control Lane 2 shows the presence of 540 bp and 740 bp fragments, indicating that the subject is
20 homozygous K1 (K:1,-2). Control Lane 3 shows the presence of 200 bp and 740 bp fragments, indicating that the subject is homozygous K2 (K:-1,2). Control Lane 4 shows the presence of each of the K1/K2 fragments, i.e., 200 bp, 540 bp, and 740 bp, indicating that the subject is heterozygous (K:1,2).

25 Lanes 5-7 in Figure 9 represent samples from a family of test subjects: Lanes 5 and 6 represent parents and Lane 7 represents the fetus. A comparison of the test bands with the control bands reveals that the parent represented in Lane 5 is heterozygous (K:1,2) and that the parent represented in Lane 6
30 is homozygous K2 (K:-1,2). The fetus is clearly homozygous K2 (K:-1,2).

Lanes 8-10 in Figure 9 represent samples from another family of test subjects. The first parent, represented in Lane 8, is heterozygous (K:1,2), while the second parent, represented in Lane 9, is homozygous K2 (K:-1,2). The fetus, represented in Lane 10, is identifiable as being homozygous K2 (K:-1,2).

The experimental data provided in Examples 7 and 8 clearly demonstrate the diagnostic efficacy of the method of the invention. The examples show that individuals of each of the three K1/K2 genotypes can be distinguished positively by virtue of the differences between K1 DNA and K2 DNA. More particularly, the newly identified locus of K1/K2 polymorphism can be exploited by means of amplification of Kell DNA using K1- and K2-specific nucleic acid probes, as well as by differential digestion of K1/K2 DNA with a restriction enzyme.

Example 9

The molecular basis for the K6/K7 polymorphism has now unexpectedly been found, showing that it is due to a point mutation in exon 17 leading to proline (K6) being encoded instead of leucine (K7) at amino acid position 597 of the Kell protein. This amino acid substitution is in close proximity to the zinc-binding putative enzyme active site of Kell protein.

DNA Preparation.

Leukocyte DNA was prepared from peripheral blood of eight unrelated individuals with K6 phenotype, as described in Examples 5-7, based on a procedure described by John et al. (Ref. 63).

Polymerase Chain Reaction (PCR).

The forward and reverse primers used to copy and amplify the 19 exons of the *KEL* gene were, for the most part, the same as those previously described elsewhere herein. In a few instances, other primers were designed to produce smaller PCR

products to facilitate sequencing. The forward and reverse primers used to amplify an 807 bp segment from exons 17 and 18 are 5'-CTCACCTAGGCAGCACCAACCCTA-3' (SEQ ID NO:64) and 5'-CAGTGAGGACATCTGCAGAAGAGG-3' (SEQ ID NO:26), respectively.

5 The PCR conditions used were the same as previously described in Examples 5-7 above, except that the number of cycles was increased to 35.

DNA Sequencing.

10 The PCR products were separated by gel electrophoresis on 0.8% low melting agarose gels. After elution the DNA was either directly sequenced or subcloned into pT7Blue(R)T plasmid vector (Novagen, Madison, Wisconsin), and transformed in DH5 α F' strain of *E. coli* (Gibco BRL Life Technologies, Gaithersburg, Maryland). If subcloned, the plasmid DNA was 15 prepared in a small scale by the alkali lysis method and purified with a Quick Spin[™] (Sephadex[™] G50) column. Standard molecular biology procedures were employed. DNA sequencing was performed by an automated system (Model 373A, Version 1-2.0, Applied Biosystems, Foster City, CA).

20 Restriction Enzyme Digestion of PCR Product.

A short 111 bp sequence of exon 17, predicted to contain the point mutation for the K6 genotype, was copied and amplified by PCR. Only a short segment of exon 17 was copied because exon 17 contains other neighboring *MnII* cleavage 25 sites. The following oligonucleotide primers were used for the PCR reaction, 5'-CTCACCTAGGCAGCACCAACCCTA-3' (SEQ ID NO:64) (forward primer, same as described above), and 5'-TTACCTGGAGGGCATGGTTGTCCT-3' (SEQ ID NO:65) (reverse primer). The PCR product (10 μ L of final reaction mixture) 30 was treated with 10 Units of *MnII* (New England Biolabs, Beverly, Massachusetts) and incubated at 37°C for 90 minutes. The final reaction volumes was 20 μ L. The DNA was then separated by electrophoresis on 0.7% agarose gels, supplemented with 1.65% GelTwin[™] brand surfactant (J.T. 35 Baker, Philipsburg, New Jersey).

Comparison of K6 and K7 DNA sequences.

The base sequences of the PCR products corresponding to the 19 exons of K:6,-7 DNA were compared to a "wild-type" DNA (K:-6,7), which expresses the high incidence Kell phenotypes including K7 (Ref. 43). Two discrete base substitutions were noted, both occurring in exon 17. The changes are shown graphically in Figure 10. At nt 1910, a T to C change was observed which results in a proline being encoded instead of leucine at amino acid position 597. Another base substitution, A to G, was observed at nt 2019, but this change does not cause the encoding of a different amino acid.

Confirmation of K6/K7 Polymorphism by *MnII* Digestion.

The T to C substitution at nt 1910 eliminates a recognition site for the restriction endonuclease *MnII*. Thus, wild-type K7 DNA is cleaved at this point when exposed to *MnII*, while the low frequency K6 DNA is not affected. This distinct feature may be used to analyze for this point mutation and to determine whether this substitution identifies the K6 genotype. A 111 bp sequence, which spans this region of exon 17, was obtained by PCR of genomic DNA from eight unrelated K:6,-7 and K:-6,7 persons, and treated with *MnII*. The K7 (wild-type) DNA contains the *MnII* site and upon treatment with *MnII* yields a 91 bp and 20 bp products. On the other hand, the K6 PCR product is not cleaved by *MnII*, and the 111 bp product remains uncut. Thus, a heterozygote yields both the uncut PCR product and the two smaller bands.

The results are shown in Figure 11. PCR products from seven unrelated K6 phenotype persons all showed 111 bp bands. Four examples are shown (Lanes 3, 6, 7, 9). PCR products from all K7 phenotype persons showed a smaller 91 bp product (Lanes 4 and 8), but the 20 bp product, because of its small size, was difficult to detect by ethidium bromide staining. The K:6,7 heterozygotes (Lanes 5 and 10) showed both the uncut 111 bp band and the smaller 91 bp band.

Chimpanzee red cells carry the K6 (Js^a) antigen (Ref. 27). A single DNA sample from a chimpanzee was analyzed and it also yielded the 111 bp uncut PCR product, a pattern which is consistent with a K:6,-7 genotype. This result is shown as Lane 11 in the gel shown in Figure 11.

Example 10

The K6/K7 polymorphism was explored further and a second restriction enzyme cleavage site was identified in association with the point mutation at nt 2019. While this is a silent mutation, it may be exploited according to the invention since it is consistently associated with the expressed mutation at nt 1910.

DNA was prepared from peripheral blood of the seven individuals identified in Example 9, as described therein. The PCR and DNA sequencing were performed as described in Example 9. The 807 bp PCR product was digested with *DdeI*. Ten microliters (10 μ L) of the PCR product was incubated with 10-20 units of *DdeI* (New England Biolabs, Beverly, Massachusetts) at 37°C for 90 minutes, and the enzyme-digested mixture was separated on a 0.7% agarose gels supplemented with 1.65% GelTwin™, as described in Example 9.

The PCR product was found to have several *DdeI* sites, and multiple bands were obtained upon digestion. DNA from four K6 phenotype individuals yielded three prominent bands of 454, 183, and 170 bp (Figure 12, Lanes 3, 6, 7, and 9). By contrast, the K& phenotypes yielded four bands (266, 188, 183, and 170 bp), because the 454 bp band was further digested to 266 and 188 bp products (Lanes 4 and 8). the 188 and 183 bp products are not well separated, and sometimes migrate together. Two heterozygotes (Lanes 5 and 10) could be distinguished from the K6 and K7 homozygotes, because they yielded five bands (454, 266, 188, 183, and 170 bp).

One sample of chimpanzee DNA was also digested using *DdeI*. As noted above in Example 9, chimpanzee red cells carry type K:6,-7. As shown in Figure 12, the 454 bp product was uncut by *DdeI* (Lane 11), again consistent with a K:6,-7 genotype.

A comparison of the base sequence of the coding region of the *KEL* gene of K6 and K7 DNA showed only 2 base differences, both occurring in exon 17. See Figure 10. Only one of these base substitutions encoded a different amino acid. A T to C change at nt 1910 changed a leucine to a proline at amino acid residue 597. The other base substitution at nt 2019 is a silent mutation which retains leucine at that position. Both of the base substitutions at nt 1910 and nt 2019 are common to K:6,-7 genotypes, since analysis of the seven unrelated individuals described in Examples 9 and 10 showed these base changes. As expected, K:6,7 heterozygotes possessed both the wild-type and the K:6,-7 sequences.

The pentameric consensus sequence, HEXXH, which Kell protein shares with a family of zinc endopeptidases occurs at amino acid residues 581 to 585. Thus, this putative active site is in close proximity to residue 597 which has the leucine (K7) to proline (K6) change. We cannot predict whether this amino acid change influences the putative enzymatic activity of Kell, but its proximity cannot be ignored. It is also of interest that the leucine to proline change occurs at a residue which is located between two cysteine residues (see Figure 10) yielding a Cys.Pro.Ala.Cys sequence in K6, and a Cys.Leu.Ala.Cys sequence in K7. Because of their distinctive molecular structure, proline residues can induce local structural changes in proteins, and can change the geometry of helices. In this case, because it is situated between two cysteine residues, proline could cause rearrangements of disulfide bonds, thus markedly changing the secondary and tertiary structure of the protein and exposing new epitopes. A computer assisted Chou-Fasman analysis of the

secondary structure of Kell protein indicates that substituting proline for leucine at residue 597 would, in fact, alter the turns predicted for that region.

5 It should be noted that Kell has 16 cysteine residues, with 15 of these residues occurring in the extracellular domain, and one occurring in the membrane-spanning region. This indicates extensive folding of the protein, which is in agreement with early biochemical studies showing inactivation of Kell antigen by reducing reagents. Of all the Kell
10 antigens, K6 and K7 are the ones which are inactivated with the lowest concentrations (1-2 mM) of dithiothreitol (DTT). Other Kell antigens such as K1, K2, K3 and K21 require much higher concentrations (100-200 mM) of DTT for inactivation.

15 An immunological procedure, termed MAIEA, has been used to map the spatial relationships between Kell epitopes. This method has predicted that K6/K7 is close to the K1/K2 epitope in the native Kell protein. The point mutation which distinguishes the K1/K2 polymorphism occurs at amino acid residue 193, while that for K6/K7 is linearly distant,
20 occurring at amino acid 597. This apparent discrepancy again points to the folded nature of Kell protein and that the epitopes can be far removed from the points of mutation.

The experimental data provided in Examples 9 and 10 again show the diagnostic efficacy of the method of the invention.
25 Specifically, individuals of having various K6/K7 genotypes can be distinguished positively on the basis of differences between K6 DNA and K7 DNA. More particularly, the newly identified locus of K6/K7 polymorphism can be exploited by means of amplification of Kell DNA using K6- and K7-specific
30 nucleic acid probes, as well as by differential digestion of K6/K7 DNA with a restriction enzyme.

Example 11

Further studies revealed that a single nucleotide substitution at nt 1601 in exon 13 of the *KEL* gene is associated with K10 genotype. The molecular basis of K10 is shown in Figure 13, which provides the nucleotide and amino acid sequences of wild-type Kell DNA in the K10 polymorphism locus, as well as the point mutation (A 1601 T) which causes a change in amino acid sequence (Glu 494 Val). The point mutation at nt 1601 results in the K10 phenotype.

RFLP analysis was performed in general accordance with the methods described elsewhere herein. It was found that the nucleotide substitution creates a new cleavage site for the restriction enzyme *AccI*, permitting differential detection of amplification products from K10 and wild-type DNA and consequent differential determination of K10 genotype.

Example 12

Analysis of the K3/K4/K21 polymorphism locus was performed to establish whether RFLP methods could be used to determine K3/K4/K21 genotype. Studies revealed unexpectedly that two discrete nucleotide substitutions in a single codon are associated with the polymorphism of the K3/K4/K21 locus. The molecular basis of K4 is shown in Figure 14, which provides the nucleotide and amino acid sequences of K4 DNA, as well as the point mutations which causes changes in amino acid sequence. The point mutation at nt 961 (C → T) cause a change in amino acid sequence (Arg 281 Trp), resulting in the K3 phenotype. The point mutation at nt 962 (G → A) cause a change in amino acid sequence (Arg 281 Gln), resulting in the K21 phenotype.

RFLP analysis was performed in general accordance with the methods described elsewhere herein. It was found that the C 961 T nucleotide substitution creates a new cleavage site for the restriction enzyme *NlaIII*, permitting differential detection of amplification products from K3 and wild-type K4

DNA and consequent differential determination of K3/K4 genotype. It was also found that the G 962 A nucleotide substitution creates a new cleavage site for the restriction enzyme *PvuII*, permitting differential detection of amplification products from K21 and wild-type K4 DNA and consequent differential determination of K21/K4 genotype. Together, the discovery of these restriction sites now enables a method for directly determining K3/K4/K21 genotype.

Thus, while there have been described what are presently believed to be the preferred embodiments of the present invention, those skilled in the art will realize that other and further embodiments can be made without departing from the spirit of the invention, and it is intended to include all such further modifications and changes as come within the true scope of the claims set forth herein.

REFERENCES

1. Giblett ER, "A critique of the theoretical hazard of inter- vs. intra-racial transfusion" Transfusion 1:233 (1961).
2. Mayne KM, Bowell PJ, and Pratt GA, "The significance of anti-Kell sensitization in pregnancy" Clin Lab Haematol 12:379-385 (1990).
3. Duguid JKM, and Bromilow IM, "Haemolytic disease of the newborn due to anti-K" Vox Sang 58:69 (1990).
4. Moncharmont P, Juron-Dupraz F, Doillon M, Vignal M, and Debeaux V, "A case of hemolytic disease of the newborn infant due to anti-K (Cellano)" Acta Haematol 85:45 (1991).
5. Constantine G, Fitzgibbon N, and Weave JB, "Anti-Kell in pregnancy" Brit J Obs Gyn 98:943 (1991).

6. Leggat HM, Gibson JM, Barron SL, and Reid MM, "Anti-Kell in pregnancy" Brit J Obs Gyn 98:162 (1991).
7. Bowman JM, Pollock JM, Manning FA, Harman CK, and Menticoglou S, "Maternal Kell blood group alloimmunization" Obstetrics and Gynecology 79:239 (1992).
8. Lee S, Zambas E, Marsh WL, and Redman CM, "The human Kell blood group gene maps to chromosome 7q33 and its expression is restricted to erythroid cells" Blood 81:2804 (1993).
9. Zelinski T, Coghlan G, Myal Y, Shiu RPC, Phillips S, White L, and Lewis M, "Genetic linkage between the Kell blood group system and prolactin-inducible protein loci: Provisional assignment of KEL to chromosome 7" Ann Hum Genet 55:137 (1991).
10. Purohit KR, Weber JL, Ward LJ, and Keats BJB, "The Kell blood group locus is closed to the cystic fibrosis locus on chromosome 7" Hum Genet 89:457 (1992).
11. Murphy MT, Morrison N, Miles JS, Fraser RH, Spurr NK, and Boyd E, "Regional chromosomal assignment of the Kell blood group locus (KEL) to chromosome 7q35-q35 by fluorescence in situ hybridization: Evidence for the polypeptide nature of antigenic variations" Hum Genet 91:585 (1993).
12. Giblett ER, and Chase J, "Js^a a 'new' red cell antigen found in negroes: evidence for an eleventh blood group system" Brit J Haematol 5:319-26 (1959).
13. Furuhielm U, Nevanlinna HR, Nurkka R, Gavin J, Tippet P, Gooch A, and Sanger R, "The blood group antigen U1^a (Karhula)" Vox Sang 15:118-24 (1968).

14. Marsh WL, "Blood groups of human red cells in clinical practice of blood transfusion" (Petz, LD and Swisher SN, eds) pp. 79-130, Churchill-Livingstone, New York (1981).
- 5 15. Marsh WL and Redman CM, "Recent developments in the Kell blood group system" Trans Med Rev 1:4 (1987).
16. Marsh WL and Redman CM, "The Kell blood group system: a review" Transfusion 30:151 (1990).
- 10 17. Redman CM and Marsh WL, "The Kell antigens and McLeod red cells" in Protein Blood Group Antigens of the Human Red Cell:structure, Function and Clinical Significance. (Agre PC and Cartron JP, eds.) pp 53-69, Johns Hopkins University Press, Baltimore, MD (1992).
- 15 18. Redman CM and Marsh WL, "The Kell blood group system and the McLeod phenotypes" Seminars in Hematology 40:309 (1993).
19. Redman CM, Marsh WL, Mueller KA, Avellino GP, and Johnson CL, "Isolation of Kell-active protein from the red cell membrane" Transfusion 24:176 (1984).
- 20 20. Wallas C, Simon R, Sharpe MA, and Byler C, "Isolation of Kell-reactive protein from red cell membranes" Transfusion 26:173 (1986).
- 25 21. Redman CM, Avellino G, Pfeffer SR, Mukherjee TK, Nichols M, Rubinstein P, and Marsh WL, "Kell blood group antigens are part of a 93,000 Dalton red cell membrane protein" J Biol Chem 261:9521 (1986).
22. Jaber A, Blanchard D, Goossens D, Bloy C, Lambin P, Rouger P, Salmon C, and Cartron JP, "Characterization of blood group Kell (K1) antigen with a human monoclonal antibody" Blood 73:1597 (1989).

23. Jaber A, Loirot MJ, Willem C, Bloy C, Cartron JP, and Blanchard D, "Characterization of murine monoclonal antibodies directed against the Kell blood group glycoprotein" Brit J Haematol 79:311 (1991).
- 5 24. Parson SF, Gardner B, and Anstee DJ, "Monoclonal antibodies against Kell glycoprotein: Serology, immunochemistry and quantification of antigen sites" Transfusion Medicine 3:137 (1993).
- 10 25. Lee S, Zambas E, Marsh WL, and Redman CM, "Molecular cloning and primary structure of Kell blood group protein" Proc Natl Acad Sci (USA) 88:6353 (1991).
26. Petty AC, Daniels GL, and Tippet P, "Application of the MAIEA assay to the Kell blood group system" Vox Sang 66:216 (1994).
- 15 27. Redman CM, Lee S, Ten Bokkel Huinink D, Rabin, BI, Johnson CL, Oyen R, and Marsh WL, "Comparison of human and chimpanzee Kell blood group system" Transfusion 239:486-490 (1989).
- 20 28. Branch D, Muensch H, Sy Siok Hian A, and Petz D, "Disulfide bonds are a requirement of Kell and Cartwright (Yt^a) blood group antigen integrity" Br J Haematol 54:573-578 (1993).
- 25 29. Jongeneel CX, Bouvier J, and Bairoch A, "A unique signature identifies a family of zinc-dependent metallopeptidases" FEBS Letters 242:211 (1989).
30. Zelinski T, "The use of DNA restriction fragment length polymorphisms in conjunction with blood group serology" Transfusion 31:762-770 (1991).

31. Bause E, "Structural requirements of N-glycosylation proteins" Biochem J 209:331-336 (1983).
32. Telen MJ, Le Van Kim C, Guizzo ML, Cartron JP, and Colin Y, "Erythrocyte Webb-type glycophorin C variant lacks N-glycosylation due to an asparagine to serine substitution" Am J Hematol 37:51-52 (1991).
33. Chang S, Reid M, Conboy J, Kan Y, and Mohandas N, "Molecular characterization of erythrocyte glycophorin C variants" Blood 77:644-648 (1991).
34. Innis MA, Gelfand DH, Sninsky JJ, and White TJ, PCR Protocols: A Guide to Methods and Applications, Academic Press, Inc., San Diego (1990).
35. Landegren U, et al., Science 241:1077 (1988).
36. Barany F, PCR Methods and Applications 1:5 (1991).
37. Narang et al., Meth Enzymol 68:90 (1979).
38. Brown et al., Meth Enzymol 68:109 (1979).
39. Beaucage et al., Tetrahedron Lett 22:1859 (1981).
40. Glick BR and Pasternak JJ, Molecular Biotechnology: Principles and Applications of Recombinant DNA, American Society for Microbiology Press, Washington, D.C. (1994).
41. Walker RH et al., eds., Technical Manual, 10th ed., American Association of Blood Banks, Arlington, Virginia (1990).
42. Sambrook J, Fritsch EF, and Maniatis T, Molecular Cloning: A Laboratory Manual, 2d. ed., Cold Spring Harbor Laboratory, NY (1989).

43. Lee S, Zambas E, Green ED, and Redman C, "Organization of the gene encoding the human Kell blood group protein" Blood 85:912-16 (1995).
- 5 44. Gerard NP, Bao L, Yiao-Ping H, Eddy RL, Shows TB, and Gerard C, "Characterization of the human C5a receptor gene" Biochemistry 32:1243 (1993).
- 10 45. Letarte M, Vera R, Tran JBL, Addis JBL, Ouizuka RJ, Quackenbush EJ, Jongeneel CY, and McInnes RR, "Common acute lymphocytic leukemia antigen is identical to neutral endopeptidase" J Exp Med 168:1247 (1988).
- 15 46. Shipp MA, Vijayaraghavan J, Schmitt EV, Masteller EL, D'Adamio L, Heish LB, and Reinherz EL, "Common acute lymphoblastic antigen (CALLA) is active neutral endopeptidase 24.11 (enkephalinase): Direct evidence by cDNA transfection analysis" Proc Natl Acad Sci USA 86:297 (1989).
- 20 47. D'Adamio L, Shipp MA, Masteller EL, and Reinherz EL, "Organization of the gene encoding common acute lymphoblastic leukemia antigen (neutral endopeptidases 24.11): Multiple mini exons and separate 5' untranslated region" Proc Natl Acad Sci USA 86:7103 (1989).
48. Bucher P, and Trifonov EN, "Compilation and analysis of eukaryotic POL II promoter sequences" Nucleic Acids Res. 14:10009 (1986).
- 25 49. Evans T, Reitman M, and Felsenfeld G, "An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes" Proc Natl Acad Sci, USA 85:5976 (1988).

50. Plum M, Frampton J, Wainwright H, Walker M, Macleod K, Goodwin G, and Harrisson P, "GATAAG: A cis-control region binding an erythroid specific nuclear factor with a role in globin and non-globin gene expression" Nucleic Acids Res. 17:73 (1989).
51. Philipsen S, Talbot D, Fraser P, and Grosvelt F, "The β -globin dominant control region: Hypersensitive site 2" EMBO J. 9:2159 (1990).
52. Beaupain D, Elouet JF, and Romeo PH, "Initiation of transcription of erythroid promoter of the porphobilinogen deaminase gene is regulated by a cis-acting sequence around the cap site" Nucleic Acids Res. 18:6509 (1990).
53. Maouche L, Tournamille C, Hattab C, Boffa G, Cartron JP, and Chretien S, "Cloning of the gene encoding the human erythropoietin receptor" Blood 78:2557 (1991).
54. Rahuel C, Vinit M-H, Lemarchandel V, Cartron JP, and Romeo P-H, "Erythroid specific activity of the glycophorin B promoter requires GATA-1 mediated displacement of a repressor" EMBO J. 11:4095 (1992).
55. Cherif-Zahar B, LeVan Kim C, Rouillac C, Raynal V, Cartron JP, and Colin YP, "Organization of the gene (RHCE) encoding the human blood group RhCcEe antigens and characterization of the promoter region" Genomics 19:68 (1994).
56. Crotta S, Nicolis S, Ronchi A, Ottolenghy S, Ruzzi L, Shimada Y, Migliaccio AR, and Migliaccio G, "Progressive inactivation of the expression of the erythroid transcription factor in GM - and G-CSF-dependent cell lines" Nucleic Acid Res. 18:6864 (1990).

57. Orkin SH, Tsai SF, Zan LI, Martin D, and Whitelaw E, "The erythroid-specific transcription factor GATA-1: Structure and expression" in Regulation of Hemoglobin Switching, (Stamatoyannopoulos G, Nienhuis AW, eds) Alan R Liss, New York. pp 310 (1992).
58. Crossley M, and Orkin SH, "Regulation of the β globin locus" Current Opinion in Genetics and Development 3:232 (1993).
59. Goosens M and Kan YW, "DNA analysis in the diagnosis of hemoglobin disorders" Methods Enzymol 76:805 (1981).
60. Wahle E and Keller W, "The biochemistry of 3' end cleavage and polyadenylation of messenger RNA precursors" Ann Rev Biochem 61:419 (1992).
61. Rastinejad F and Blau HM, "Genetic complementation reveals a novel regulatory role for 3' untranslated regions in growth and differentiation" Cell 72:903 (1993).
62. Jackson RJ, "Cytoplasmic regulation of mRNA function: The importance of the 3' untranslated region" Cell 74:9 (1993).
63. John, SWM, Weitzner G, Rozen R, and Sriver CR, "A rapid procedure for extracting genomic DNA from leukocytes" Nucleic Acids Res 19:408 (1991).

Sequence Listing

(1) GENERAL INFORMATION:

- (i) APPLICANT: Lee, Soohie
Redman, Colvin L.
- (ii) TITLE OF INVENTION: Diagnostic Method and Kit for
Determining Kell Blood Group
Genotype
- (iii) NUMBER OF SEQUENCES: 65
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Ronald J. Baron, Esq.
Hoffmann & Baron
 - (B) STREET: 350 Jericho Turnpike
 - (C) CITY: Jericho
 - (D) STATE: New York
 - (E) COUNTRY: USA
 - (F) ZIP: 11753
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 - (B) COMPUTER: IBM PC Compatible
 - (C) OPERATING SYSTEM: MS-DOS
 - (D) SOFTWARE: WordPerfect 6.1
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/484,570
 - (B) FILING DATE: 07-JUN-1995
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/337,268
 - (B) FILING DATE: 11-OCT-1994
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Baron, Ronald J.
 - (B) REGISTRATION NUMBER: 29,281
 - (C) REFERENCE/DOCKET NUMBER: 454-3 PCT
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (516) 822-3550
 - (B) TELEFAX: (516) 822-3582

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

WO 96/15268

PCT/US95/14684

- (A) LENGTH: 26 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATACTGACTC ATCAGAAGTT TCAGCA

26

(3) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 26 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATACTGACTC ATCAGAAGTC TCAGCA

26

(4) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 26 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGGACTTCCT TAAACTTTAA CTGAAC

26

(5) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTTAGTCCTC ACTCCCATGC TTCC

24

(6) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TATCACACAG GTGTCCTCTC TTCC

24

(7) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGAGACTTCC AAGGTCTTAG CTATCACTTA AGCAC

35

(8) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTGGTTCGGC CCACCTCTGA AGGTTCCAGA ATCGATAG

38

(9) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTCGGCTCTT CCTCACTTTG GTCC

24

(10) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTCTTGGCTC CAGAGAGTTC CCAT

24

(11) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCCACCTTCC ATCTGTCTAT CTTC

24

(12) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GACTCGAGTC GACAACGTTT TTTTTTTTTT TTTT

35

(13) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ATGGGGAGAC TGTCTG

17

(14) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CAGTCCTCCG AATCAGCTCC TAGA

24

(15) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTCTTGGCTC CAGAGAGTTC CCAT

24

(16) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAAGGTGGGG ACCAAAGTGA GGAA

24

(17) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACAGGGTTTG GAGCAGTCAT GGTC

24

(18) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATATTCCCCA CCTCCCCACA CCTG

24

(19) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATCTACGGTG CTCAGGCTCT CCTC

24

(20) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GGAAGCATGG GAGTGAGGAC TAAA

24

(21) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TGGCATCCAT GGTACCTCAT GGAA

24

(22) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GAGGCTTTTG AAACCCAGG ATGA

24

(23) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTCCCCAGCC ACCTGCCATC TCAT

24

(24) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCCTTTTCCA AGGGTCAGAA GCTG

24

(25) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGGCTTATTT GACCCCCAGA ATCT

24

(26) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CCTAATCCCT GGATGCCTGC CTGT

24

(27) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CAGTGAGGAC ATCTGCAGAA GAGG

24

(28) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCCTGTGGAC CCTCCCCCTT CAAT

24

(29) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GGGCGGAAGC CAAGTGCCAG CTTTT

25

(30) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AATAAAAGCT GGCACCTGGC TTCCGCCGGA ATTC

34

(31) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AATAAAAGCT GGCACCTGGC TTCCG

25

(32) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AATAAAAGCT GGCACCTGGC TTCC

24

(33) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AATAAAAGCT GGCACCTGGC TTC

23

(34) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AATAAAAGCT GGCACCTGG

19

(35) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

AATAAAAGCT GGCACCTTGGC TTCC

24

(36) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AATAAAAGCT GGCACCTTGGC TTCCGCTTGT CTCT

34

(37) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

AATAAAAGCT GGCACCTTGGC TTCCGCTTGT CTCTT

35

(38) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:

WO 96/15268

PCT/US95/14684

- (A) LENGTH: 148 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GAAGTGCCCC TTCTCCAGGA TCAAGGAACT GGGGCGGGGG GTGTTTCCTG	50
GACCCCAAGTC CTCCGAATCA GCTCCTAGAG TGGAACCAGG AAGGATTCTG	100
GAGCCACAGA AGATAGACAG ATG GTAAGTCCCC TTTTGGAGTC AGAGG	148
Met	

(39) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 128 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CTCCTTCTCC CTCCACTCAC TTCAG GAA GGT GGG GAC CAA AGT GAG GAA	49
Glu Gly Gly Asp Gln Ser Glu Glu	
1 5	
GAG CCG AGG GAA CGC AGC CAG GCA GGT GGA ATG GGA ACT CTC TGG	94
Glu Pro Arg Glu Arg Ser Gln Ala Gly Gly Met Gly Thr Leu Trp	
10 15 20	
AGC CAA GAG GTAAGTGGCC TCCTCTCCTG GGTCT	128
Ser Gln Glu	
25	

(40) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 nucleotides
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TTTCACCTCT	TGGTTCCTCC	CACAG	AGC	ACT	CCA	GAA	GAG	AGG	CTG	CCC	49				
			Ser	Thr	Pro	Glu	Glu	Arg	Leu	Pro					
			1				5								
GTG	GAA	GGG	AGC	AGG	CCA	TGG	GCA	GTG	GCC	AGG	CGG	GTG	CTG	ACA	94
Val	Glu	Gly	Ser	Arg	Pro	Trp	Ala	Val	Ala	Arg	Arg	Val	Leu	Thr	
	10					15					20				
GCT	ATC	CTG	ATT	TTG	GGC	CTG	CTC	CTT	TGT	TTT	TCT	CTC	CTT	TGT	139
Ala	Ile	Leu	Ile	Leu	Gly	Leu	Leu	Leu	Cys	Phe	Ser	Leu	Leu	Cys	
	25					30					35				
TTT	TCT	GTG	CTT	TTG	TTC	TAC	AAC	TTC	CAG	AAC	TGT	GGC	CCT	C	182
Phe	Ser	Val	Leu	Leu	Phe	Tyr	Asn	Phe	Gln	Asn	Cys	Gly	Pro		
	40					45					50				
GTAAGCAAGA	TCCCAGACCC	CCCAA													207

(41) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 227 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CCCAGCTCTG	AGCTTTTCCC	CACAG	GC	CCC	TGT	GAG	ACA	TCT	GTG	TGT	48				
			Arg	Pro	Cys	Glu	Thr	Ser	Val	Cys					
			1				5								
TTG	GAT	CTC	CGG	GAT	CAT	TAC	CTG	GCC	TCT	GGG	AAC	ACA	AGT	GTG	93
Leu	Asp	Leu	Arg	Asp	His	Tyr	Leu	Ala	Ser	Gly	Asn	Thr	Ser	Val	
	10					15					20				
GCC	CCC	TGC	ACC	GAC	TTC	TTC	AGC	TTT	GCC	TGT	GGA	AGG	GCC	AAA	138

WO 96/15268

PCT/US95/14684

Ala	Pro	Cys	Thr	Asp	Phe	Phe	Ser	Phe	Ala	Cys	Gly	Arg	Ala	Lys	
25						30					35				
GAG	ACC	AAT	AAT	TCT	TTT	CAG	GAG	CTT	GCC	ACA	AAG	AAC	AAA	AAC	183
Glu	Thr	Asn	Asn	Ser	Phe	Gln	Glu	Leu	Ala	Thr	Lys	Asn	Lys	Asn	
40						45					50				
CGA	CTT	CGG	AGA	ATA	CTG	G	GTGAGGAAAG	CAGGGTGGAA	GATGC						227
Arg	Leu	Arg	Arg	Ile	Leu										
55															

(42) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 181 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TTTAGTCCTC	ACTCCCATGC	TTCCTTTCTA	G	AG	GTC	CAG	AAT	TCC	TGG						
				Glu	Val	Gln	Asn	Ser	Trp	48					
				1				5							
CAC	CCA	GGC	TCT	GGG	GAG	GAG	AAA	GCC	TTC	CAG	TTC	TAC	AAC	TCC	93
His	Pro	Gly	Ser	Gly	Glu	Glu	Lys	Ala	Phe	Gln	Phe	Tyr	Asn	Ser	
			10					15					20		
TGC	ATG	GAT	ACA	CTT	GCC	ATT	GAA	GCT	GCA	GGG	ACT	GGT	CCC	CTC	138
Cys	Met	Asp	Thr	Leu	Ala	Ile	Glu	Ala	Ala	Gly	Thr	Gly	Pro	Leu	
			25					30					35		
AGA	CAA	GTT	ATT	GAG	GAG	GTGAGAAAAG	TTGGGATATT	AACTT							181
Arg	Gln	Val	Ile	Glu	Glu										
			40												

(43) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 197 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TCAGCCCCCT	CTCTCTCCTT	TAAAG	CTT	GGA	GGC	TGG	CGC	ATC	TCT	GGT	49				
			Leu	Gly	Gly	Trp	Arg	Ile	Ser	Gly					
			1				5								
AAA	TGG	ACT	TCC	TTA	AAC	TTT	AAC	CGA	ACG	CTG	AGA	CTT	CTG	ATG	94
Lys	Trp	Thr	Ser	Leu	Asn	Phe	Asn	Arg	Thr	Leu	Arg	Leu	Leu	Met	
	10					15					20				
AGT	CAG	TAT	GGC	CAT	TTC	CCT	TTC	TTC	AGA	GCC	TAC	CTA	GGA	CCT	139
Ser	Gln	Tyr	Gly	His	Phe	Pro	Phe	Phe	Arg	Ala	Tyr	Leu	Gly	Pro	
	25					30					35				
CAT	CCT	GCC	TCT	CCA	CAC	ACA	CCA	GTC	ATC	CAG	GTGAGGGATG				182
His	Pro	Ala	Ser	Pro	His	Thr	Pro	Val	Ile	Gln					
	40					45									
CACTGGCGAA	GACAC														197

(44) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 113 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TCTCTCCAGT	CTCTCTTGTG	CCCAG	ATA	GAC	CAG	CCA	GAG	TTT	GAT	GTT	49			
			Ile	Asp	Gln	Pro	Glu	Phe	Asp	Val				
			1				5							
CCC	CTC	AAG	CAA	GAT	CAA	GAA	CAG	AAG	ATC	TAT	GCC	CAG	GTAAG	93
Pro	Leu	Lys	Gln	Asp	Gln	Glu	Gln	Lys	Ile	Tyr	Ala	Gln		
	10					15					20			
ATGGCACATG	GACAAAGGCC													113

(45) INFORMATION FOR SEQ ID NO:44:

						1					5					
TTG	TCC	TGC	TTG	CAA	GCG	ACA	TTC	ACA	CCG	ATG	TCC	CTG	AGC	CCT		94
Leu	Ser	Cys	Leu	Gln	Ala	Thr	Phe	Thr	Pro	Met	Ser	Leu	Ser	Pro		
	10					15					20					
TCT	CAG	TCC	CTC	GTG	GTC	CAT	GAC	GTG	GAA	TAT	TTG	AAA	AAC	ATG		139
Ser	Gln	Ser	Leu	Val	Val	His	Asp	Val	Glu	Tyr	Leu	Lys	Asn	Met		
	25					30					35					
TCA	CAA	CTG	GTG	GAG	GAG	ATG	CTG	CTA	AAG	CAG	AG	GTTCGCCGCA				184
Ser	Gln	Leu	Val	Glu	Glu	Met	Leu	Leu	Lys	Gln	Arg					
	40					45					50					
GGTGGGATTG	GGGAG															199

(47) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 180 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GTGTGGGTCT	CTTTTGTCTC	CATAG	G	GAC	TTT	CTG	CAG	AGC	CAC	ATG		47				
				Asp	Phe	Leu	Gln	Ser	His	Met						
			1				5									
ATC	TTA	GGG	CTG	GTG	GTG	ACC	CTT	TCT	CCA	GCC	CTG	GAC	AGT	CAA		92
Ile	Leu	Gly	Leu	Val	Val	Thr	Leu	Ser	Pro	Ala	Leu	Asp	Ser	Gln		
	10					15					20					
TTC	CAG	GAG	GCA	CGC	AGA	AAG	CTC	AGC	CAG	AAA	CTG	CGG	GAA	CTG		137
Phe	Gln	Glu	Ala	Arg	Arg	Lys	Leu	Ser	Gln	Lys	Leu	Arg	Glu	Leu		
	25					30					35					
ACA	GAG	CAA	CCA	CCC	ATG	GTGAGGAGAG	GAGCGGGTGT	ATTTG								180
Thr	Glu	Gln	Pro	Pro	Met											
	40															

(48) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:

GCGTCCTGGC TGGTG

149

(50) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 128 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TAGCCTCTTC	TGTGTCTCTC	TCCAG	GTT	GCT	CAA	CTG	CAG	GTG	GAG	ATG	49				
			Val	Ala	Gln	Leu	Gln	Val	Glu	Met					
			1				5								
GGG	GCT	TCA	GAA	TGG	GCC	CTG	AAG	CCA	GAG	CTG	GCC	CGA	CAA	GAA	94
Gly	Ala	Ser	Glu	Trp	Ala	Leu	Lys	Pro	Glu	Leu	Ala	Arg	Gln	Glu	
	10					15					20				
TAC	AAC	GAT	GTGGGTCCCT	GTGTTTCCA	GCTCC										128
Tyr	Asn	Asp													
	25														

(51) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 151 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AAGTCACCTC	CTGCCTCTTC	CCCAG	ATA	CAG	CTT	GGA	TCG	AGC	TTC	CTG	49				
			Ile	Gln	Leu	Gly	Ser	Ser	Phe	Leu					
			1				5								
CAG	TCT	GTC	CTG	AGC	TGT	GTC	CGG	TCC	CTC	CGA	GCT	AGA	ATT	GTC	94
Gln	Ser	Val	Leu	Ser	Cys	Val	Arg	Ser	Leu	Arg	Ala	Arg	Ile	Val	
	10					15					25				

WO 96/15268

PCT/US95/14684

CAG AGC TTC TTG CAG CCT CAC CCC CAA CAC AG GTATGACAGC AGGGG 141
 Gln Ser Phe Leu Gln Pro His Pro Gln His Arg
 30 35

AGACACAGGC 151

(52) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 161 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GAGTTCACAT GTCCTCTTCC CACAG G TGG AAG GTG TCC CCT TGG GAC 47
 Trp Lys Val Ser Pro Trp Asp
 1 5

GTC AAT GCT TAC TAT TCG GTA TCT GAC CAT GTG GTA GTC TTT CCA 92
 Val Asn Ala Tyr Tyr Ser Val Ser Asp His Val Val Val Phe Pro
 10 15 20

GCT GGA CTC CTC CAA CCC CCA TTC TTC CAC CCT GGC TAT CCC AG 136
 Ala Gly Leu Leu Gln Pro Pro Phe Phe His Pro Gly Tyr Pro Arg
 25 30 35

GTATGGGTCA CTCTGTAAGG GTAGG 161

(53) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 118 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GTCAAATAAG CCCTTGTCTC CCTAG A GCC GTG AAC TTT GGC GCT GCT 47

Ala Val Asn Phe Gly Ala Ala
1 5

GGC	AGC	ATC	ATG	GCC	CAC	GAG	CTG	TTG	CAC	ATC	TTC	TAC	CAG	CTC	92
Gly	Ser	Ile	Met	Ala	His	Glu	Leu	Leu	His	Ile	Phe	Tyr	Gln	Leu	
		10					15					20			

T GTGGGTAACA GGGGCCACTG GGAGG 118

(54) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 220 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TGTTCTCTTG	TCCCATTTTC	AACAG	TA	CTG	CCT	GGG	GGC	TGC	CTC	GCC	48
				Leu	Leu	Pro	Gly	Gly	Cys	Leu	Ala
				1			5				

TGT	GAC	AAC	CAT	GCC	CTC	CAG	GAA	GCT	CAC	CTG	TGC	CTG	AAG	CGC	93
Lys	Asp	Asn	His	Ala	Leu	Gln	Glu	Ala	His	Leu	Cys	Leu	Lys	Arg	
	10					15					20				

CAT	TAT	GCT	GCC	TTT	CCA	TTA	CCT	AGC	AGA	ACC	TCC	TTC	AAT	GAC	138
His	Tyr	Ala	Ala	Phe	Pro	Leu	Pro	Ser	Arg	Thr	Ser	Phe	Asn	Asp	
	25					30					35				

TCC	CTC	ACA	TTC	TTA	GAG	AAT	GCT	GCA	GAC	GTT	GGG	GGG	CTA	GCC	183
Ser	Leu	Thr	Phe	Leu	Glu	Asn	Ala	Ala	Asp	Val	Gly	Gly	Leu	Ala	
	40					45					50				

ATC	GCG	CTG	CAG	GTATGCAAGT	GTCAAGGGCC	ACAGT	220
Ile	Ala	Leu	Gln				
	55						

(55) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 146 nucleotides
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

CCCTTCTCTA	CCCACCCCTA	CCCAG	GCA	TAC	AGC	AAG	AGG	CTG	TTA	CGG	49				
			Ala	Tyr	Ser	Lys	Arg	Leu	Leu	Arg					
			1				5								
CAC	CAT	GGG	GAG	ACT	GTC	CTG	CCC	AGC	CTG	GAC	CTC	AGC	CCC	CAG	94
His	His	Gly	Glu	Thr	Val	Leu	Pro	Ser	Leu	Asp	Leu	Ser	Pro	Gln	
10						15					20				
CAG	ATC	TTC	TTT	CGA	AGC	TAT	GCC	CAG	GTAGGCAGCG	GCCACCTCCC	141				
Gln	Ile	Phe	Phe	Arg	Ser	Tyr	Ala	Gln							
25						30									
GCCAC											146				

(56) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 313 nucleotides

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

TTCAATAACC	TCTCTTCCTG	CTCAG	GTG	ATG	TGT	AGG	AAG	CCC	AGC	CCC	49				
			Val	Met	Cys	Arg	Lys	Pro	Ser	Pro					
			1				5								
CAG	GAC	TCT	CAC	GAC	ACT	CAC	AGC	CCT	CCA	CAC	CTC	CGA	GTC	CAC	94
Gln	Asp	Ser	His	Asp	Thr	His	Ser	Pro	Pro	His	Leu	Arg	Val	His	
10						15					20				
GGG	CCC	CTC	AGC	AGC	ACC	CCA	GCC	TTT	GCC	AGG	TAT	TTC	CGC	TGT	139
Gly	Pro	Leu	Ser	Ser	Thr	Pro	Ala	Phe	Ala	Arg	Tyr	Phe	Arg	Cys	
25						30					35				
GCA	CGT	GGT	GCT	CTC	TTG	AAC	CCC	TCC	AGC	CGC	TGC	CAG	CTC	TGG	184
Ala	Arg	Gly	Ala	Leu	Leu	Asn	Pro	Ser	Ser	Arg	Cys	Gln	Leu	Trp	

40	45	50	
TAACTTGGTT	ACCAAAGATG	CCACAGCACA	GAAATATCGA CCAACACCTC 234
CCTGGTCACA	TCCATGGAAT	CAGAGCAAGA	TTTCCTTTCT GCTTCTGTTC 284
CAAAAATAAA	AGCTGGCACT	TGGCTTCCG	313

(57) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 327 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

GTCACAGTGC	AAGACAAAAG	GAGCAGACCA	AGGGCAAGAT	TGCTTGGGGA	50
GTGAAGACTC	CCTCCCTCTT	CTCCCCTGAG	AAGCTGAGAT	AAAGGGGGAG	100
GAGAAGCCTG	GGTGCCCCCC	ACTGATAAGC	AGGCTCCACC	CAGAGGCCAG	150
TCCTGTGTGT	CTGGGGACAA	GGCGAAAGAG	CAGCAGAAGT	GCCCCTTCTC	200
CAGGATCAAG	GAAGTGGGGC	GGGGGGTGTT	TCCTGGACCC	CAGTCCTCCG	250
AATCAGCTCC	TAGAGTGGAA	CCAGGAAGGA	TTCTGGAGCC	ACAGAAGATA	300
GACAGAGTGT	AAGTCCCCTT	TTGGAGT			327

(58) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

TGG	ACT	TCC	TTA	AAC	TTT	AAC	CGA	ACG	CTG	AGA	CTT	CTG	ATG	AGT	45
Trp	Thr	Ser	Leu	Asn	Phe	Asn	Arg	Thr	Leu	Arg	Leu	Leu	Met	Ser	
1				5					10					15	

(59) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TGG	ACT	TCC	TTA	AAC	TTT	AAC	CGA	ATG	CTG	AGA	CTT	CTG	ATG	AGT	45
Trp	Thr	Ser	Leu	Asn	Phe	Asn	Arg	Met	Leu	Arg	Leu	Leu	Met	Ser	
1				5					10					15	

(60) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

His	Glu	Leu	Leu	His
1				5

(61) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal fragment

(ix) FEATURE:

- (A) NAME/KEY: variable residues
- (B) LOCATION: 3-4
- (C) OTHER INFORMATION: /note= "consensus sequence found in active sites of zinc neutral peptidases"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

His Glu Xaa Xaa His
1 5

(62) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

ATACTGACTC ATCAGAAGTG TCAGCA .

26

(63) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

TTAGTCCTCA CTCNCCATGC TTCC

24

(64) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TCACACAGGT GTCCTCTCTT CC

22

(65) INFORMATION FOR SEQ ID NO:64:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

CTCACCTAGG CAGCACCAAC CCTA

24

(66) INFORMATION FOR SEQ ID NO:65:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

TTACCTGGAG GGCATGGTTG TCACT

25

**THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE
PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:**

1. A diagnostic method for determining K1/K2 genotype of a subject,
5 comprising:
generating a nucleic acid product from a nucleic acid sample comprising a
K1/K2 polymorphism locus, wherein the locus is in exon 6 of the gene which
encodes the Kell protein,
wherein K1 genotype of the Kell protein is encoded by exons 1 through
10 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that
nucleotide 78 in SEQ. ID. NO. 42 is thymine,
wherein K2 genotype of the Kell protein is encoded by exons 1 through
19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively,
wherein said generating step comprises digesting said nucleic acid sample
15 with a restriction enzyme which differentially cleaves K1 nucleic acid,
wherein said restriction enzyme comprises *BsmI*,
wherein the product is generated in an amount sufficient to characterize
K1/K2 genotype.
- 20 2. The diagnostic method of Claim 1, wherein said generating step further
comprises amplifying nucleic acid.
3. The diagnostic method of Claim 2, wherein said amplifying step comprises
amplifying said nucleic acid sample prior to said selectively cleaving step.
25
4. The diagnostic method of Claim 2, wherein said amplifying step comprises
amplifying said nucleic acid product following said selectively cleaving step.
5. The diagnostic method of Claim 2, wherein said amplifying step comprises
30 amplifying nucleic acid by polymerase chain reaction, ligase chain reaction,
or a combination thereof.
6. A diagnostic method for determining K1/K2 genotype of a subject,
comprising:

generating a nucleic acid product which is characteristic of K1/K2 genotype from a nucleic acid sample comprising a K1/K2 polymorphism locus, wherein the locus is in exon 6 of the gene which encodes the Kell protein,

5 wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine,

 wherein K2 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively,

10 wherein said generating step comprises differentially amplifying said nucleic acid sample using an allele-specific primer which differentially amplifies alleles of a K1/K2 polymorphism, wherein said allele-specific primer includes the nucleotide sequence selected from the group consisting of: ATA CTG ACT CAT CAG AAG TTT CAG CA (SEQ ID NO:1); ATA
15 CTG ACT CAT CAG AAG TCT CAG CA (SEQ ID NO:2); ATA CTG ACT CAT CAG AAG TGT CAG CA (SEQ ID NO:61); TGG ACT TCC TTA AAC TTT AAC TGA AC (SEQ ID NO:3); TTT AGT CCT CAC TCC CAT GCT TCC (SEQ ID NO:4); TAT CAC ACA GGT GTC CTC TCT TCC (SEQ ID NO:5); TTA GTC CTC ACT CNC CAT GCT TCC (SEQ ID
20 NO:62); and TCA CAC AGG TGT CCT CTC TTC C (SEQ ID NO:63) to provide said characteristic nucleic acid product,

 wherein the product is generated in an amount sufficient to characterize K1/K2 genotype.

25 7. The diagnostic method of Claim 6, wherein said differentially amplifying step comprises differentially amplifying said nucleic acid sample by polymerase chain reaction, ligase chain reaction, or a combination thereof.

30 8. The diagnostic method of Claim 1 or Claim 6, further comprising deriving information characterizing K1/K2 genotype from said nucleic acid product, wherein said nucleic acid product comprises one or more nucleic acid fragments, and wherein said information deriving step comprises separating said one or more nucleic acid fragments to provide a fragment pattern,

wherein said fragment pattern provides specific information characterizing K1/K2 genotype of the subject.

- 5 9. The diagnostic method of Claim 8, wherein said information deriving step further comprises detecting one or more of the nucleic acid fragments in the fragment pattern.
- 10 10. The diagnostic method of Claim 9, wherein said detecting step includes marking said one or more nucleic acid fragments with a non-specific marker substance.
- 15 11. The diagnostic method of Claim 9, wherein said detecting step includes specifically marking said one or more nucleic acid fragments with one or more hybridization probes, and wherein each of said hybridization probes has specificity for a single nucleic acid fragment.
- 20 12. The diagnostic method of Claim 11, wherein each of said hybridization probes is detectably labeled.
- 25 13. The diagnostic method of Claim 1 or 6, further comprising determining a non-Kell genotype of said subject.
14. The diagnostic method of Claim 1 or 6, further comprising determining a Kell phenotype of said subject.
- 30 15. The diagnostic method of Claim 14, wherein said determining step comprises determining a Kell phenotype of said subject by serological testing.
16. The diagnostic method of Claim 1, comprising using a DNA sample derived from a subject.
17. The diagnostic method of Claim 16, wherein said DNA sample comprises genomic DNA.

18. The diagnostic method of Claim 16, wherein said DNA sample is derived from a biological sample containing erythroid tissue of the subject.
19. The diagnostic method of Claim 18, wherein said biological sample
5 comprises amniotic fluid or chorionic villus and wherein said subject is a fetus *in utero*.
20. The diagnostic method of Claim 18, wherein said biological sample comprises a blood sample.
21. A nucleic acid oligomer, comprising a nucleic acid sequence which specifically binds to K1 polymorphism locus,
10 wherein the locus is in exon 6 of the gene which encodes the Kell protein, and
15 wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine.
22. The nucleic acid oligomer of Claim 21, wherein said oligomer is detectably
20 labeled.
23. The nucleic acid oligomer of Claim 21, wherein said oligomer is attached to a substrate.
24. The nucleic acid oligomer of Claim 21, wherein said nucleic acid sequence is
25 exactly complementary to said K1 polymorphism locus nucleic acid.
25. A nucleic acid oligomer comprising a nucleic acid sequence which encodes K1 polymorphism locus, wherein the amino acid sequence of the
30 polymorphic locus is shown in SEQ. ID. NO. 58.
26. The nucleic acid oligomer of Claim 25, wherein said oligomer is detectably labeled.

27. The nucleic acid oligomer of Claim 25, wherein said oligomer is attached to a substrate.
28. A Kell-based polypeptide comprising an amino acid sequence which
5 comprises the K1 domain shown in SEQ. ID. NO. 58.
29. The Kell-based polypeptide of Claim 28, wherein said polypeptide is detectably labeled.
- 10 30. The Kell-based polypeptide of Claim 28, wherein said polypeptide is attached to a substrate.
31. An expression recombinant vector which transforms a cell to cause
expression of K1 protein, comprising an expression vector which includes
15 SEQ. ID. NO. 37 to SEQ. ID. NO. 55, wherein nucleotide 78 in SEQ. ID. NO. 42 is thymine.
32. A method of producing a cell line transformed to express K1 protein,
comprising transforming a cell by means of the expression recombinant
20 vector of Claim 31 to produce a transformed cell, and establishing a stable transformed cell line derived from said transformed cell.
33. A transformed cell line produced according to the method of Claim 32.
- 25 34. A method of making an antibody specific for K1 protein, comprising:
inducing the generation of at least one antibody against K1 antigen
expressed by a transformed cell line produced according to Claim 32, wherein
the K1 antigen comprises the amino acid sequence of at least SEQ. ID. NO.
58.
- 30 35. A diagnostic method for the differential determination of K1 and/or K2
genotype in a patient, comprising:
amplifying a DNA sample obtained from a patient using a primer which
amplifies only K1 DNA and/or K2 DNA, including the locus which

determines K1/K2 polymorphism, to identify the presence of amplified DNA corresponding to K1 DNA and/or K2 DNA,

digesting said DNA using a restriction enzyme which differentially cleaves K1 DNA and K2 DNA, and

5 separating said amplified DNA according to molecular weight to form a pattern of fragments, wherein said pattern of fragments provides specific information diagnostic of the K1/K2 genotype of the patient,

wherein the locus is in exon 6 of the gene which encodes the Kell protein,

10 wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine,

wherein K2 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively.

15

36. The diagnostic method of claim 35, wherein said primer amplifies K1 DNA and K2 DNA.

20 37. The diagnostic method of claim 35, wherein said at least one primer comprises at least one primer which amplifies only K1 DNA.

38. The diagnostic method of claim 35, wherein said at least one primer comprises at least one primer which amplifies only K2 DNA.

25 39. The diagnostic method of claim 35, wherein said at least one primer comprises at least one primer which amplifies only K1 DNA and at least one primer which amplifies only K2 DNA.

30 40. The diagnostic method of claim 35, wherein said at least one primer includes a primer specific for K1 comprising a nucleotide sequence selected from the group consisting of: ATA CTG ACT CAT CAG AAG TTT CAG CA (SEQ ID NO:1), and ATA CTG ACT CAT CAG AAG TCT CAG CA (SEQ ID NO:2).

41. The diagnostic method of claim 35, wherein said at least one primer includes a primer specific for K2 comprising the nucleotide sequence: TGG ACT TCC TTA AAC TTT AAC TGA AC (SEQ ID NO:3).
- 5 42. The diagnostic method of claim 35, wherein said at least one primer includes a primer specific for K1 DNA and/or K2 DNA comprising a nucleotide sequence selected from the group consisting of: TTT AGT CCT CAC TCC CAT GCT TCC, (SEQ ID NO:4), and TAT CAC ACA GGT GTC CTC TCT TCC, (SEQ ID NO:5).
- 10 43. The diagnostic method of claim 35, wherein said amplifying includes amplifying by polymerase chain reaction, ligase chain reaction or a combination thereof.
- 15 44. The diagnostic method of claim 35, wherein said digesting follows said amplifying.
45. The diagnostic method of claim 35, wherein said digesting precedes said amplifying.
- 20 46. The diagnostic method of claim 35, wherein said restriction enzyme is *BsmI*.
47. The diagnostic method of claim 35, wherein said DNA sample is obtained from a biological sample containing erythroid tissue of the patient.
- 25 48. The diagnostic method of claim 47, wherein said biological sample comprises amniotic fluid and wherein said patient is a fetus in utero.
49. The diagnostic method of claim 47, wherein said biological sample comprises a blood sample.
- 30 50. The diagnostic method of claim 35, further comprising serologically typing a blood sample from said patient to determine Kell phenotype.

51. A nucleic acid oligomer comprising a nucleic acid sequence which specifically binds to a region of Kell DNA diagnostic of K1 DNA, wherein the region is in exon 6 of the gene which encodes the Kell protein,
 wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine.
52. The nucleic acid oligomer of claim 51, wherein said region of Kell DNA is specific to K1 DNA.
53. The nucleic acid oligomer of claim 51, wherein said nucleic acid sequence is detectably labeled.
54. The nucleic acid oligomer of claim 51, wherein said nucleic acid sequence is attached to a substrate.
55. The nucleic acid oligomer of claim 51, wherein said nucleic acid sequence is exactly complementary to said region of Kell DNA.
56. A nucleic acid oligomer having a nucleic acid sequence which encodes a region of K1 protein, which region includes the K1 domain shown in SEQ. ID. NO. 58.
57. The nucleic acid oligomer of claim 56, wherein said nucleic acid sequence is specific to K1 DNA.
58. The nucleic acid oligomer of claim 56, wherein said nucleic acid sequence is detectably labeled.
59. The nucleic acid oligomer of claim 56, wherein said nucleic acid sequence is attached to a substrate.

60. A nucleic acid primer comprising a primer which specifically binds to or causes elongation through K1 DNA, including the locus which determines K1 polymorphism,
 wherein the locus is in exon 6 of the gene which encodes the Kell protein,
 5 and
 wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine.
- 10 61. The nucleic acid primer of claim 60, wherein said primer specifically binds to or causes elongation through K1 DNA.
62. The nucleic acid primer of claim 60, wherein said primer comprises a K1 primer which specifically binds to or causes elongation through K1 DNA.
- 15 63. An isolated Kell-based polypeptide, which is specifically reactive with an antibody specific for a protein comprising the amino acid sequence of SEQ. ID. No. 58, said polypeptide comprising a K1/K2 polymorphism locus, wherein the locus is in exon 6 of the gene which encodes the Kell protein,
 20 wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine, or
 wherein K2 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively.
- 25 64. The isolated Kell-based polypeptide of claim 63, wherein said polypeptide is detectably labeled.
65. The isolated Kell-based polypeptide of claim 63, wherein said polypeptide is
 30 attached to a substrate.
66. The isolated Kell-based polypeptide of claim 63, wherein said polypeptide has an amino acid sequence which includes a methionine residue at a position corresponding to amino acid 192 of the Kell protein.

67. The isolated Kell-based polypeptide of claim 63, wherein said polypeptide has an amino acid sequence identical to at least a portion of the Kell protein including the K1 domain shown in SEQ. ID. NO. 58.
- 5 68. A diagnostic method for detecting alloimmunization of a patient to K1 antigen comprising the amino acid sequence of at least SEQ. ID. NO. 58, said method comprising:
- measuring a parameter of immune reactivity of a blood sample of a patient with a peptide probe,
- 10 wherein said peptide probe comprises an amino acid sequence which is specifically reactive with antibodies specific for a protein comprising the amino acid sequence of SEQ. ID. NO. 58 in the sample, wherein the amino acid sequence is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in
- 15 SEQ. ID. NO. 42 is thymine.
69. A diagnostic method for the detection of a target nucleic acid encoding at least a part of the K1 domain shown in SEQ. ID. NO. 58 or the K2 domain shown in SEQ. ID. NO. 57 of a Kell protein, comprising:
- 20 ascertaining the presence, in a nucleic acid fraction obtained from a patient, of a target nucleic acid encoding at least a part of the K1 or K2 domain of a Kell protein by means of a nucleic acid probe which is known to specifically bind to a nucleic acid sequence encoding the K1 or K2 domain of a Kell protein,
- 25 wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine, and
- wherein K2 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively.
- 30 70. The diagnostic method of claim 69, wherein said nucleic acid probe specifically binds to a nucleic acid sequence encoding the K1 domain of Kell protein.

71. The diagnostic method of claim 69, wherein said nucleic acid probe specifically binds to a nucleic acid sequence specifically encoding the K2 domain of Kell protein.
- 5 72. The diagnostic method of claim 69, wherein said nucleic acid probe is detectably labeled.
73. The diagnostic method of claim 69, wherein said nucleic acid probe is attached to a substrate.
- 10 74. The diagnostic method of claim 69, wherein said target nucleic acid is genomic DNA.
75. The diagnostic method of claim 69, wherein the target nucleic acid is mRNA.
- 15 76. The diagnostic method of claim 69, wherein the target nucleic acid is cDNA.
77. A diagnostic kit for the differential determination of K1/K2 genotype, comprising:
- 20 (a) a primer specific for a region of K1 DNA and/or K2 DNA including the locus which determines the K1/K2 polymorphism; and
- (b) a container,
- wherein the region is in exon 6 of the gene which encodes the Kell protein,
- 25 wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine, and
- wherein K2 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively.
- 30 78. The diagnostic kit of claim 77, wherein said container comprises a biological specimen container for housing said primer.

79. A recombinant expression vector which transforms a cell to cause expression of a polypeptide including the K1 domain, comprising an expression vector which includes a nucleic acid sequence encoding at least a part of K1 protein including at least the K1 domain shown in SEQ. ID. NO. 58.
- 5
80. A method of producing a cell line transformed to express a K1 polypeptide or protein, comprising transforming a cell with the recombinant expression vector of claim 79 to produce a transformed cell, and establishing a stable transformed cell line derived from said transformed cell.
- 10
81. A transformed cell line produced according to the method of claim 80.
82. A diagnostic method for determining Kell genotype of a subject, comprising:
- 15 (a) obtaining a nucleic acid sample from a subject;
- (b) generating a nucleic acid fragment from said sample, the detection of which permits the differentiation of alleles of a K1/K2 polymorphism;
- (c) detecting said nucleic acid fragment; and
- (d) determining the genotype of said subject with respect to said K1/K2 polymorphism based on said detecting,
- 20 wherein the K1/K2 polymorphism is in exon 6 of the gene which encodes the Kell protein,
- wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine,
- 25 wherein K2 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively.
83. The diagnostic method of claim 82, wherein said generating step comprises selectively cleaving said nucleic acid sample to provide at least one cleavage fragment whose detection permits the differentiation of alleles of the K1/K2 polymorphism; and wherein said detecting step comprises detecting said at least one cleavage fragment.
- 30

84. The diagnostic method of claim 83, wherein said selectively cleaving step comprises digesting said nucleic acid sample with at least one restriction enzyme which cleaves said nucleic acid sample differentially based on a nucleotide sequence difference at a K1/K2 polymorphism locus.
- 5
85. The diagnostic method of claim 84, wherein said at least one restriction enzyme is *BsmI*.
86. The diagnostic method of claim 83, wherein said generating step further
10 comprises amplifying a Kell nucleic acid using a primer to provide amplified K1/K2 nucleic acid.
87. The diagnostic method of claim 86, wherein said generating step comprises amplifying said Kell nucleic acid prior to said selectively cleaving step.
- 15
88. The diagnostic method of claim 86, wherein said generating step comprises amplifying said Kell nucleic acid following said selectively cleaving step.
89. The diagnostic method of claim 86, wherein said amplifying step comprises
20 amplifying nucleic acid by polymerase chain reaction, ligase chain reaction, or a combination thereof.
90. The diagnostic method of claim 82, wherein said generating step comprises:
25 differentially amplifying said nucleic acid sample using an allele-specific primer which differentially amplifies alleles of a K1/K2 polymorphism to provide said nucleic acid fragment.
91. The diagnostic method of claim 90, wherein said differentially amplifying step
30 comprises differentially amplifying said nucleic acid sample by polymerase chain reaction, ligase chain reaction, or a combination thereof.
92. The diagnostic method of claim 90, wherein said allele-specific primer includes a primer comprising a nucleotide sequence selected from the group consisting of:
ATA CTG ACT CAT CAG AAG TTT CAG CA (SEQ ID NO:1), ATA CTG

ACT CAT CAG AAG TCT CAG CA (SEQ ID NO:2), and ATA CTG ACT CAT CAG AAG TGT CAG CA (SEQ ID NO:61).

5 93. The diagnostic method of claim 90, wherein said allele-specific primer includes a primer comprising the nucleotide sequence: TGG ACT TCC TTA AAC TTT AAC TGA AC (SEQ ID NO:3).

10 94. The diagnostic method of claim 90, wherein said primer includes a primer comprising a nucleotide sequence selected from the group consisting of: TTT AGT CCT CAC TCC CAT GCT TCC (SEQ ID NO:4), TAT CAC ACA GGT GTC CTC TCT TCC (SEQ ID NO:5), TTA GTC CTC ACT CNC CAT GCT TCC (SEQ ID NO:62), and TCA CAC AGG TGT CCT CTC TTC C (SEQ ID NO:63).

15 95. The diagnostic method of claim 82, wherein said nucleic acid product comprises one or more nucleic acid fragments; and wherein said generating step comprises: separating said one or more nucleic acid fragments to provide a fragment pattern, wherein said separating step provides a fragment pattern which permits differentiation of K1/K2 genotype.

20

96. The diagnostic method of claim 95, wherein said detecting step comprises detecting one or more of the nucleic acid fragments in the fragment pattern.

25 97. The diagnostic method of claim 96, wherein said detecting step includes marking said one or more nucleic acid fragments with a sequence non-specific marker.

98. The diagnostic method of claim 96, wherein said detecting step includes specifically marking said one or more nucleic acid fragments with one or more K1/K2 allele-specific hybridization probes.

30

99. The diagnostic method of claim 98, wherein each of said hybridization probes is detectably labeled.

100. The diagnostic method of claim 82, further comprising determining a non-K1/K2 genotype of said subject.
- 5 101. The diagnostic method of claim 82, further comprising determining a K1/K2 phenotype of said subject.
- 10 102. The diagnostic method of claim 101, wherein said determining step comprises determining a K1/K2 phenotype of said subject by serological testing.
- 15 103. The diagnostic method of claim 82, further comprising obtaining a DNA sample from a subject.
104. The diagnostic method of claim 103, wherein said DNA sample comprises genomic DNA.
- 20 105. The diagnostic method of claim 103, wherein said obtaining step comprises obtaining a DNA sample from a biological sample containing erythroid tissue of the subject.
- 25 106. The diagnostic method of claim 105, wherein said biological sample comprises amniotic fluid or chorionic villus and wherein said subject is a fetus *in utero*.
107. The diagnostic method of claim 105, wherein said biological sample comprises a blood sample.
- 30 108. A nucleic acid primer set, comprising at least two nucleic acid oligomers that specifically binds to or causes elongation through a region of Kell DNA comprising K1 DNA,
wherein the primer set comprises a nucleic acid oligomer that specifically binds to or causes elongation through Kell nucleic acid comprising a K1 polymorphism locus,

wherein the locus is in exon 6 of the gene which encodes the Kell protein, and

wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine.

109. A Kell-based polypeptide having an amino acid sequence which comprises a region of a Kell protein comprising at least the K1 domain shown in SEQ. NO. 58.
110. The Kell-based polypeptide of claim 109, wherein said polypeptide is detectably labeled.
111. The Kell-based polypeptide of claim 109, wherein said polypeptide is attached to a substrate.
112. The Kell-based polypeptide of claim 109, wherein said amino acid sequence includes a methionine residue at amino acid 193 of the Kell protein.
113. The Kell-based polypeptide of claim 109, wherein said amino acid sequence includes a proline residue at amino acid 597 of the Kell protein.
114. The Kell-based polypeptide of claim 109, wherein said amino acid sequence includes a valine residue at amino acid 494 of the Kell protein.
115. The Kell-based polypeptide of claim 109, wherein said amino acid sequence includes a tryptophan residue at amino acid 281 of the Kell protein.
116. The Kell-based polypeptide of claim 109, wherein said amino acid sequence includes a glutamine residue at amino acid 281 of the Kell protein.
117. A diagnostic method for detecting alloimmunization of a patient to a K1/K2 antigen comprising the amino acid sequence of at least SEQ. ID. NO. 58, the method comprising:

measuring a parameter of immune reactivity of a blood sample of a patient with a peptide probe,

wherein said peptide probe comprises an amino acid sequence of at least a portion of a Kell protein, which portion comprises the K1 domain shown in SEQ. ID. NO. 58, and

wherein said peptide probe is specifically reactive with anti-K1/K2 antibodies in the sample.

118. A diagnostic kit for the differential determination of a Kell genotype, comprising:
- (a) a primer which specifically amplifies a region of Kell DNA including at least a K1/K2 polymorphism locus, wherein the amino acid sequence of the polymorphic locus is shown in SEQ. ID. NO. 58; and
 - (b) a container.

119. The diagnostic kit of claim 118, wherein said kit further comprises at least one restriction endonuclease which selectively cleaves Kell nucleic acid comprising at least one of said polymorphism loci to provide nucleic acid fragments whose detection permits the differential identification of a K1 and/or K2 genotype.

120. A diagnostic kit for determining a Kell genotype by detecting target nucleic acid sequences specific to Kell alleles, said kit comprising:
- (a) a primer set including at least two PCR primers wherein each of the PCR primers is an oligonucleotide that specifically binds to or causes elongation through a sequence specific to one of at least two Kell alleles, wherein the at least two Kell alleles differ at the K1/K2 polymorphism locus,
- wherein the K1/K2 polymorphism locus is in exon 6 of the gene which encodes the Kell protein,
- wherein K1 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively, except that nucleotide 78 in SEQ. ID. NO. 42 is thymine,

wherein K2 genotype of the Kell protein is encoded by exons 1 through 19 defined by SEQ. ID. NO. 37 to SEQ. ID. NO. 55, respectively; and

5 (b) a microtiter plate having a plurality of wells and having bound thereto oligonucleotide capture probes which specifically hybridize to said target sequences.

121. A recombinant expression vector which transforms a cell to cause expression of Kell protein, comprising a nucleic acid sequence encoding at least a part of
10 a Kell protein, said part comprising the K1 domain shown in SEQ. ID. NO. 58.

122. A method of producing a cell line transformed to express a Kell protein, comprising transforming a cell by means of the recombinant expression
15 vector of claim 121 to produce a transformed cell, and establishing a stable transformed cell line derived from said transformed cell.

123. A transformed cell line produced according to the method of claim 122.

20 124. A nucleic acid molecule, having a nucleic acid sequence, which encodes at least a fragment of a Kell protein, which fragment includes at least the K1 domain shown in SEQ. ID. NO. 58.

25 125. A nucleic acid molecule, having a nucleic acid sequence according to claim 124, comprising a Kell cDNA which encodes a Kell protein which includes at least the K1 domain shown in SEQ. ID. NO. 58.

30 126. The nucleic acid primer set of claim 111, wherein the primer set comprises:
a K1 primer comprising a nucleotide sequence selected from the group consisting of: ATA CTG ACT CAT CAG AAG TTT CAG CA (SEQ ID NO:1), and ATA CTG ACT CAT CAG AAG TCT CAG CA (SEQ ID NO:2), and
a K2 primer comprising the nucleotide sequence: TGG ACT TCC TTA AAC TTT AAC TGA AC (SEQ ID NO:3).

FIG-1

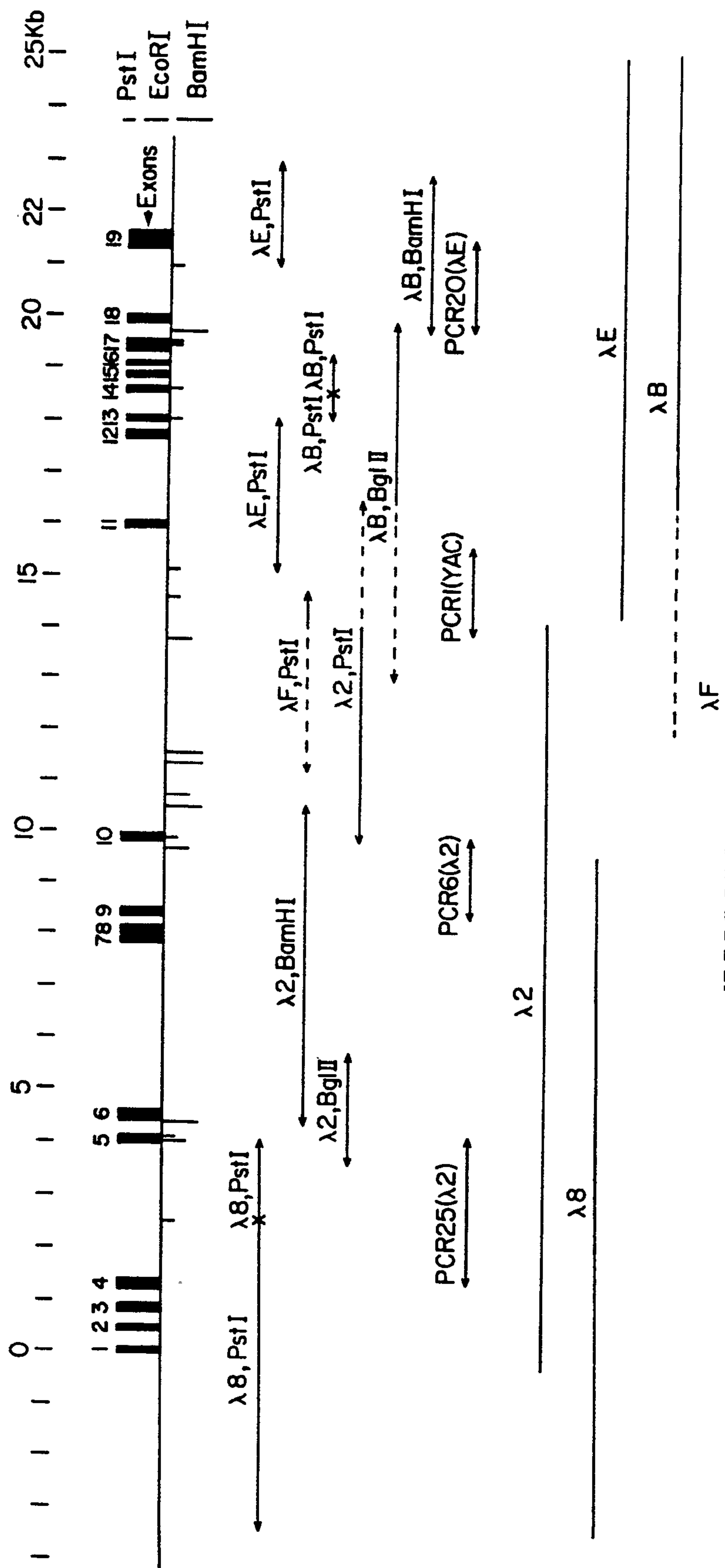


FIG-2A

WO 96/15268

CA 02204636 1997-05-06

PCT/US95/14684

EXON 1 (SEQ ID NO:37)
1-123

29	CTGGGGCGGGGGTGTTCCTGGACCCAGTCCTCCGAATCAGCTCCTAGAGTGAACC	28
88	AGGAAGGATTCTGGAGCCACAGAAGATAGACAG ATG gtaagtccttggagtcagagg.....0.34kb	87

Met

2/31

RECTIFIED SHEET (RULE 91)

FIG-2B

WO 96/15268

CA 02204636 1997-05-06

PCT/US95/14684

EXON 2 (SEQ ID NO:38)
124-201

3/31

5	133	Asp Gln Ser Glu Glu Glu Glu Glu	ctccttctccctccactcacttcag	Glu Gly Gly	4
		GAC CAA AGT GAG GAA GAG GAG GAG		GAA GGT GGG	132
20	178	Met Gly Thr Leu Trp Ser Gln Glu			
		ATG GGA ACT CTC TGG AGC CAA GAG	gtaagtggcctcctcctggtct.....		0.29kb

FIG-2C

EXON 3 (SEQ ID NO:39)
202-343

	tttcacctctcttggttcctccacacag	Ser Thr Pro Glu Glu Arg Leu	34
		AGC ACT CCA GAA GAG AGG CTG	222
35	Pro Val Glu Gly Ser Arg Pro Trp Ala Val Ala Arg Arg Val Leu		49
223	CCC GTG GAA GGG AGC AGG CCA TGG GCA GTG GCC AGG CGG GTG CTG		267
50	Thr Ala Ile Leu Ile Leu Gly Leu Leu Cys Phe Ser Val Leu		64
268	ACA GCT ATC CTG ATT TTG GGC CTG CTC CTT TGT TTT TCT GTG CTT		312
65	Leu Phe Tyr Asn Phe Gln Asn Cys Gly Pro A		
313	<u>TTG TTC TAC AAC TTC CAG AAC TGT GGC CCT C gtaagcaagatccccca</u>		

AMENDED SHEET

4/31

CA 02204636 1997-05-06

IP/EA/US 14684
10 JUN 1996

FIG-2D

WO 96/15268

CA 02204636 1997-05-06

PCT/US95/14684

EXON 4 (SEQ ID NO:40)
344-520

5/31

80	Ser Val Cys Leu Asp Leu Arg Asp His Tyr Leu Leu Ala Ser Glu Thr	79
358	TCT GTG TGT TTG GAT GAT CGG GAT CAT CAT TAC CTG GCC TCT GGT GAG ACA	357
95	Thr Ser Val Ala Pro Cys Thr Asp Phe Phe Ser Ser Phe Ala Cys Gly	94
403	ACA AGT GTG GCC ACC TGC ACC GAC TTC TTC TTC AGC TTT GCC TGT GGA	402
110	Arg Ala Lys Glu Thr Asn Asn Ser Phe Gln Glu Leu Ala Thr Lys	109
448	AGG GCC AAA GAG ACC AAT AAT TCT TTT CAG GAG CTT GCC ACA AAG	447
125	Asn Lys Asn Arg Leu Arg Arg Ile Leu G	124
493	AAC AAA AAC CGA CTT CGG AGA ATA CTG G	492

gtgaggaaagcaggggtggaagatgc..... ~2.6kb

RECTIFIED SHEET (RULE 91)

FIG-2E

EXON 5 (SEQ ID NO:41)
521-645

	tttagtcctcactcccatgcttccttcttag	lu	Val	Gln	Asn	Ser	Trp	139
			GTC	CAG	AAT	TCC	TGG	537
140	His Pro Gly Ser Gly Glu Glu Lys Ala Phe Gln Phe Tyr Asn Ser							154
538	CAC CCA GGC TCT GGG GAG GAG AAA GCC TTC CAG TTC TAC AAC TCC							582
155	Cys Met Asp Thr Leu Ala Ile Glu Ala Ala Gly Thr Gly Pro Leu							169
583	TGC ATG GAT ACA CTT GCC ATT GAA GCT GCA GGG ACT GGT CCC CTC							627
170	Arg Gln Val Ile Glu Glu							
628	AGA CAA GTT ATT GAG GAG	gtgagaaaagtgggatatattaactt.....0.33kb						

6/31

FIG-2F

WO 96/15268

EXON 6 (SEQ ID NO:42)
646-792

185	Trp	Thr	Ser	Leu	Asn	Phe	Leu	Gly	Gly	Gly	Trp	Arg	Ile	Ser	Gly	Lys	184
673	TGG	ACT	TCC	TTA	AAC	TTT	AAC	CGA	ACG	CTG	AGA	CTT	CTG	ATG	AGT	AAA	672
200	Gln	Tyr	Gly	His	Phe	Pro	Phe	Phe	Arg	Ala	Tyr	Leu	Gly	Pro	His		214
718	CAG	TAT	GGC	CAT	TTC	CCT	TTC	TTC	AGA	GCC	TAC	CTA	GGA	CCT	CAT		762
215	Pro	Ala	Ser	Pro	His	Thr	Pro	Val	Ile	Gln							
763	CCT	GCC	TCT	CCA	CAC	ACA	CCA	GTC	ATC	CAG	gtgagggatgcactggcgaagacac....~3.2kb						

7/31

FIG-2G

WO 96/15268

CA 02204636 1997-05-06

PCT/US95/14684

8/31

EXON 7 (SEQ ID NO:43)
793-855

230	Phe	Asp	Val	Pro	Leu	Lys	Gln	Asp	Gln	Glu	Gln	Pro	Glu	229
808	TTT	GAT	GTT	CCC	CTC	AAG	CAA	GAT	CAA	GAA	CAG	CCA	GAG	807
245	Gln													244
853	CAG	gtaagatggcacatggacaaaggcc.....0.093kb												852

tctctccagtcctctcttctgtgccag Ile Asp Gln Pro Glu
ATA GAC CAG CCA GAG

RECTIFIED SHEET (RULE 91)

[illegible]

EXON 9 (SEQ ID NO:45)
1045-1193

[illegible]

RECTIFIED SHEET (RULE 91)

FIG-2J

WO 96/15268

CA 02204636 1997-05-06

PCT/US95/14684

11/31

EXON 10 (SEQ ID NO:46)
1194-1323

365	Met	Ile	Leu	Gly	Leu	Val	Val	Thr	Leu	Ser	Leu	Gln	Ser	His	364
1213	ATG	ATC	TTA	GGG	CTG	GTG	GTG	ACC	CTT	TCT	CCA	GCC	CTG	CAC	1212
380	Gln	Phe	Gln	Glu	Ala	Arg	Arg	Lys	Leu	Ser	Gln	Lys	Leu	Arg	379
1258	CAA	TTC	CAG	GAG	GCA	CGC	AGA	AAG	CTC	AGC	CAG	AAA	CTG	Glu	1257
395	Leu	Thr	Glu	Gln	Pro	Pro	Met								394
1303	CTG	ACA	GAG	CAA	CCA	CCC	ATG	gtgaggagaggagcggtgtatttg.....~6kb							1302

FIG-2K

WO 96/15268

CA 02204636 1997-05-06

PCT/US95/14684

12/31

EXON 11 (SEQ ID NO:47)
1324-1434

	actcattccagctt	gtgtctccatag	Pro	Ala	Arg	Pro	Arg	Trp	Met	Lys	409	
			CCT	GCC	CGC	CCA	CGA	TGG	ATG	AAG	1347	
410	Cys	Val	Glu	Glu	Thr	Gly	Thr	Phe	Phe	Glu	424	
1348	TGC	GTG	GAG	GAG	ACA	GGC	ACG	TTC	TTC	GAG	1392	
									ACG	CTG		
425	Leu	Phe	Val	Arg	Glu	Ala	Phe	Gly	Pro	Ser		
1393	TTG	TTT	GTT	CGT	GAG	GCC	TTT	GGC	CCG	AGC		
									ACC	CGA		
									AGT	GCT		
										gta		
	tgtgagagctct	tcccagccca										

FIG-2L

WO 96/15268

EXON 12 (SEQ ID NO:48)
1435-1533

RECTIFIED SHEET (RULE 91)

440	Met	Lys	Leu	Phe	Thr	Ala	Ile	Arg	Asp	Ala	Leu	Ile	Thr	Arg	Leu	Ala	439
1438	ATG	AAA	TTA	TTC	ACT	GCG	ATC	CGG	GAT	GCC	CTC	ATC	ACT	CGC	CTC		1437
455	Arg	Asn	Leu	Pro	Trp	Met	Asn	Glu	Glu	Thr	Gln	Asn	Met	Ala	Gln		454
1483	AGA	AAC	CTT	CCC	TGG	ATG	AAT	GAG	GAG	ACC	CAG	AAC	ATG	GCC	CAG		1482
470	Asp	Lys															469
1528	GAC	AAG	gtc	agg	cc	agg	cg	tc	ct	gg	ct	gg	tg	...	0.24kb		1527

ctgtccctggacctcactcccacag GCC

13/31

PCT/US95/14684

CA 02204636 1997-05-06

FIG-2M

EXON 13 (SEQ ID NO:49)
1534-1611

483
1567

482
1566

497
1611

14/31

ctgtgtctctctctccag

Val Ala Gln Leu Gln Val Glu Met Gly Ala Ser

GTT GCT CAA CTG CAG GTG GAG ATG GGG GCT TCA

Glu Trp Ala Leu Lys Pro Glu Leu Ala Arg Gln Glu Tyr Asn Asp

GAA TGG GCC CTG AAG CCA GAG CTG GCC CGA CAA GAA TAC AAC GAT

gtgggtccctgtgtgtttccagctcc.....0.44kb

tagcctctt

FIG-2N

EXON 14 (SEQ ID NO:50)
1612-1712

500
1618

515
1663

530
1708

499
1617

514
1662

529
1707

Ile Gln
ATA CAG

aagtcacctcctgcctctcttccccag

Leu Gly Ser Ser Phe Leu Gln Ser Val Leu Ser Cys Val Arg Ser
CTT GGA TCG AGC TTC TCT GTC GTC CTG CTG AGC TGT GTC CGG TCC

Leu Arg Ala Arg Ile Val Gln Ser Phe Leu Gln Pro His Pro Gln
CTC CGA GCT AGA ATT GTC CAG AGC TTC TTG CAG CCT CAC CCC CAA

His Ar
CAC AG gtagacagcaggggagacacaggc....0.19kb

15/31

EXON 15 (SEQ ID NO:51)
1713-1823

		gagttcacatgtcctcttcc											
545 1753	g cacag G	Trp	Lys	Val	Ser	Pro	Trp	Asp	Val	Asn	Ala	Tyr	Tyr Ser
		TGG	AAG	GTG	TCC	CCT	TGG	GAC	GTC	AAT	GCT	TAC	TAT TCG
545 1753	Val Ser	Asp	His	Val	Val	Val	Phe	Pro	Ala	Gly	Leu	Leu	Gln Pro
		TCT	CAT	GTG	GTA	GTC	TTT	CCA	GCT	GGA	CTC	CTC	CAA CCC
560 1798	Pro Phe	Phe	His	Pro	Gly	Tyr	Pro	Ar					
		TTC	CAC	CCT	GGC	TAT	CCC	AG	gtaggggtcactctgtaaggtagg.....0.15kb				

EXON 16 (SEQ ID NO:52)
1824-1891

575	Ala	Gly	Ser	Ile	Met	Ala	His	Glu	Leu	Leu	His	Ile	Phe	Tyr	Gln	574
1843	GCT	GGC	AGC	ATC	ATG	GCC	CAC	GAG	CTG	TTG	CAC	ATC	TTC	TAC	CAG	1842
590	Leu	L														589
1888	CTC	T	gtgggtaacaggggccactgggagg													1887

gtcaataaagcccttggtctccctag A GCC GTG AAC TTT GGC GCT

g Ala Val Asn Phe Gly Ala

FIG-2Q

WO 96/15268

CA 02204636 1997-05-06

PCT/US95/14684

EXON 17 (SEQ ID NO:53)
1892-2061

18/31

	tggtctctgtgtcccatTTTca														
	eu	Leu	Pro	Gly	Gly	Cys	Leu	Ala	Cys	Asp	Asn	His	Ala	Leu	604
	acag	TA	CTG	CCT	GGG	GGC	TGC	CTC	TGT	GAC	AAC	CAT	GCC	CTC	1932
605	Gln	Glu	Ala	His	Leu	Cys	Leu	Lys	Arg	His	Tyr	Ala	Ala	Phe	Pro
1933	CAG	GAA	GCT	CAC	CTG	TGC	CTG	AAG	CGC	CAT	TAT	GCT	GCC	TTT	CCA
	Leu	Pro	Ser	Arg	Thr	Ser	Phe	Asn	Asp	Ser	Leu	Thr	Phe	Leu	Glu
620	TTA	CCT	AGC	AGA	ACC	TCC	TTC	AAT	GAC	TCC	CTC	ACA	TTC	TTA	GAG
1978															
	Asn	Ala	Ala	Asp	Val	Gly	Gly	Leu	Ala	Ile	Ala	Leu	Gln		
635	AAT	GCT	GCA	GAC	GTT	GGG	GGG	CTA	GCC	ATC	GCG	CTG	CAG	gtatgca	
2023															
	agtgtcaagggccacagt....0.35kb														

RECTIFIED SHEET (RULE 91)

FIG-2R

WO 96/15268

EXON 18 (SEQ ID NO:54)
2062-2157

650	Ser	Lys	Arg	Leu	Leu	Arg	His	His	Gly	Glu	Thr	Val	Leu	Pro	Ser	Ala	Tyr	649
2068	AGC	AAG	AGG	CTG	TTA	CGG	CAC	CAT	GGG	GAG	ACT	GTC	CTG	CCC	AGC	GCA	TAC	2067
665	Leu	Asp	Leu	Ser	Pro	Gln	Gln	Ile	Phe	Phe	Arg	Ser	Tyr	Ala	Gln			664
2113	CTG	GAC	CTC	AGC	CCC	CAG	CAG	ATC	TTC	TTT	CGA	AGC	TAT	GCC	CAG			2112
																		679
																		215

gtaggcagcgccacccctccggccac...~1.3kb

RECTIFIED SHEET (RULE 91)

FIG-2S

EXON 19 (SEQ ID NO:55)
2158-2445

ttcaataaacctctcttccctgctcag

2158	Val Met Cys Arg Lys Pro Ser Pro Gln Asp Ser His Asp Thr His	694
	GTG ATG TGT AGG AAG CCC AGC CCC CAG GAC TCT CAC GAC ACT CAC	2202
695	Ser Pro Pro His Leu Arg Val His Gly Pro Leu Ser Ser Thr Pro	709
2203	AGC CCT CCA CAC CTC CGA GTC CAC GGG CCC CTC AGC AGC ACC CCA	2247
710	Ala Phe Ala Arg Tyr Phe Arg Cys Ala Arg Gly Ala Leu Leu Asn	724
2248	GCC TTT GCC AGG TAT TTC CGC TGT GCA CGT GGT GCT CTC TTG AAC	2292
725	Pro Ser Ser Arg Cys Gln Leu Trp ***	
2293	CCC TCC AGC CGC TGC CAG CTC TGG TAACTTGGTTACCAAGATGCCACAGC	2343
2344	ACAGAAATATCGACCACACCTCCCTGGTCAATCCATGGAATCAGAGCAAGATTTCCT	2402
2403	TTCTGCTTCTGTCCAAAAATAAAAGCTGGCAGCTTGGCTTCCG	

AMENDED SHEET

20/31

CA 02204636 1997-05-06

PCT/US 95 / 14684
IPEA/US 10 JUN 1996

FIG-3

(SEQ ID NO:56)

-176

-185 gtcacagtgc_aagacaaaaggagcagacccaaggccaagattgcttggggagtgaagactc

-125 cctccctcttctccccctgagaagctga_gataaa_gggggaggagaagcctgggtgcccccc

-65 a_dtgataa_gcaggct_ccaccc_agaggccagtcctgtgtgtctgggacaaaggcgaaagag

-5 cagca_gaa_gtg_ccc_ctt_ctc_cag_gat_caa_gga_act_ggg_gcg_ggg_ggt

+43 gtt_tcc_tgg_acc_cca_gtc_ctc_cga_atc_agc_tcc_tag_agt_gga_acc_agg

+81 aag_gat_tct_gga_gcc_aca_gaa_gat_aga_cag_atg_gtaagtcccccttttgaggt

22/31

FIG-4

bCm 

bCm 

Cm  

1

2

FIG-5

<u>Bone marrow cDNA library</u>	<u>Clone No.</u>
AAATAAAGCTGGCACTTGGCTTCCG-CCGGAATTC-3kb ext	191
AAATAAAGCTGGCACTTGGCTTCCG (A)	185
AAATAAAGCTGGCACTTGGCTTCC (A)	182
AAATAAAGCTGGCACTTGGCTTC (A)	190
<u>Peripheral blood</u>	
AAATAAAGCTGGCACTTGG (A)	23
AAATAAAGCTGGCACTTGGCTTCC (A)	22
AAATAAAGCTGGCACTTGGCTTCCGCTTGCTCT (A)	19
AAATAAAGCTGGCACTTGGCTTCCGCTTGCTCTT (A)	21

24/31

FIG-6

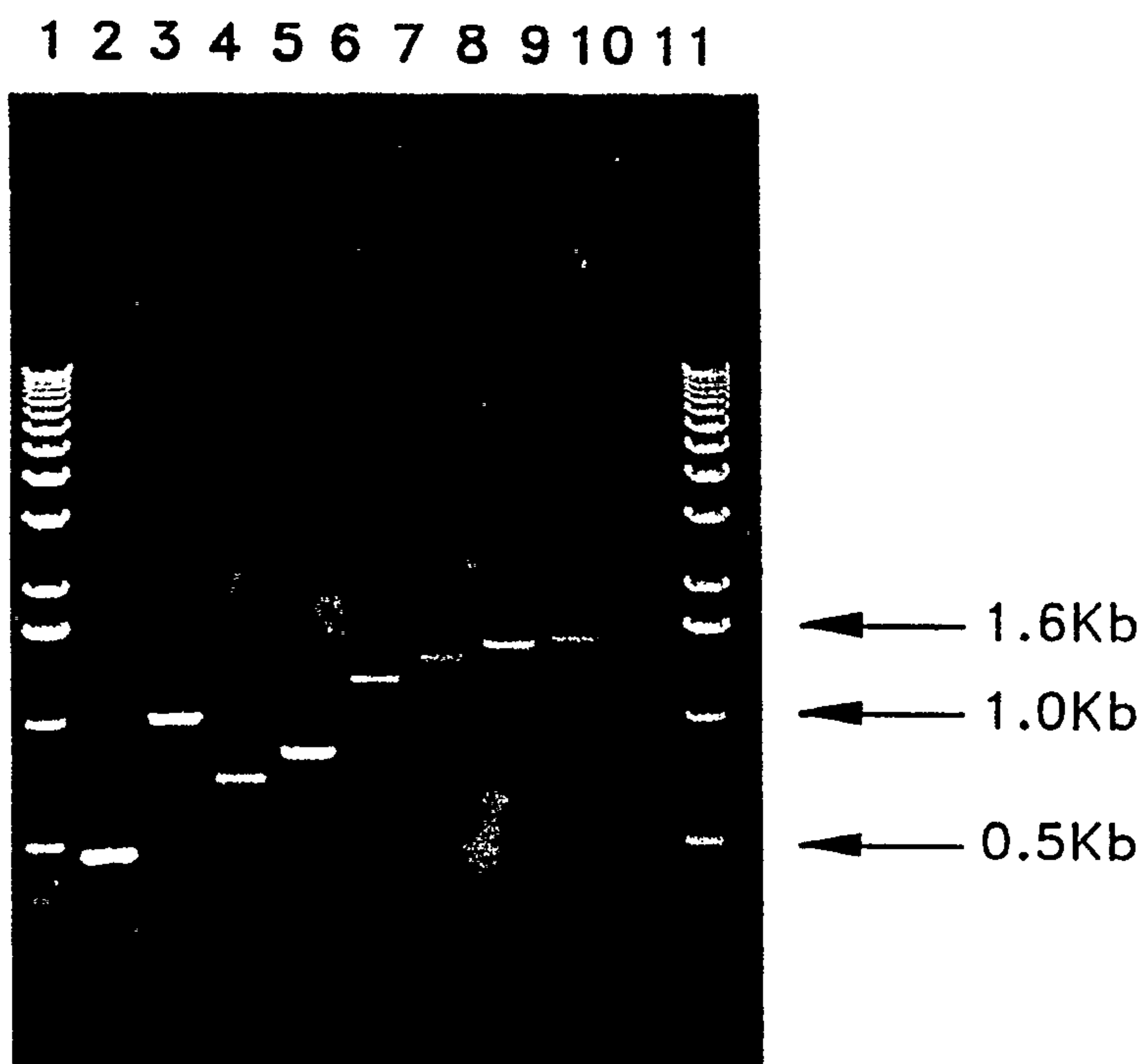


FIG-7

K2 (SEQ ID NO:57)

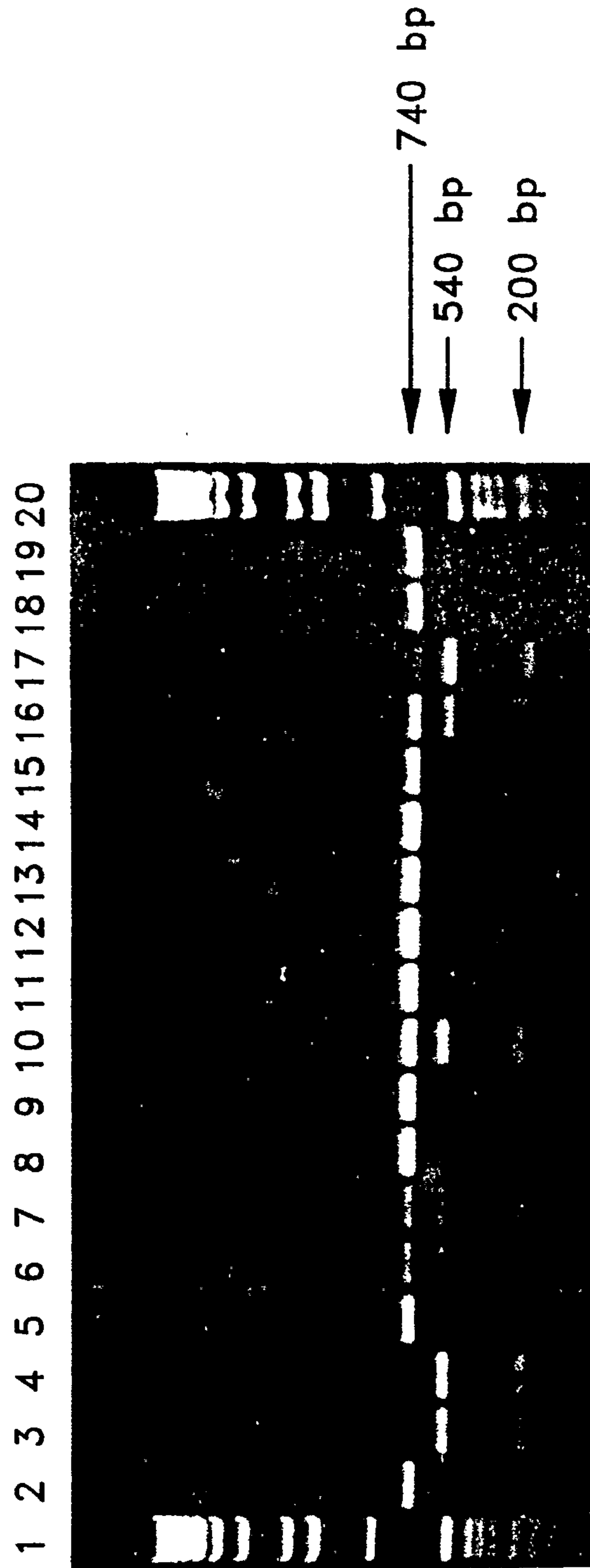
185	Trp	Thr	Ser	Leu	Asn	Phe	Asn	Arg	Thr	Leu	Arg	Leu	Leu	Met	Ser	199
	TGG	ACT	TCC	TTA	AAC	TTT	AAC	CGA	ACG	CTG	AGA	CTT	CTG	ATG	AGT	

K1 (SEQ ID NO:58)

TGG	ACT	TCC	TTA	AAC	TTT	AAC	CGA	ATG	CTG	CTG	AGA	CTT	CTG	ATG	AGT
Trp	Thr	Ser	Leu	Asn	Phe	Asn	Arg	Met	Leu	Arg	Leu	Leu	Met	Ser	

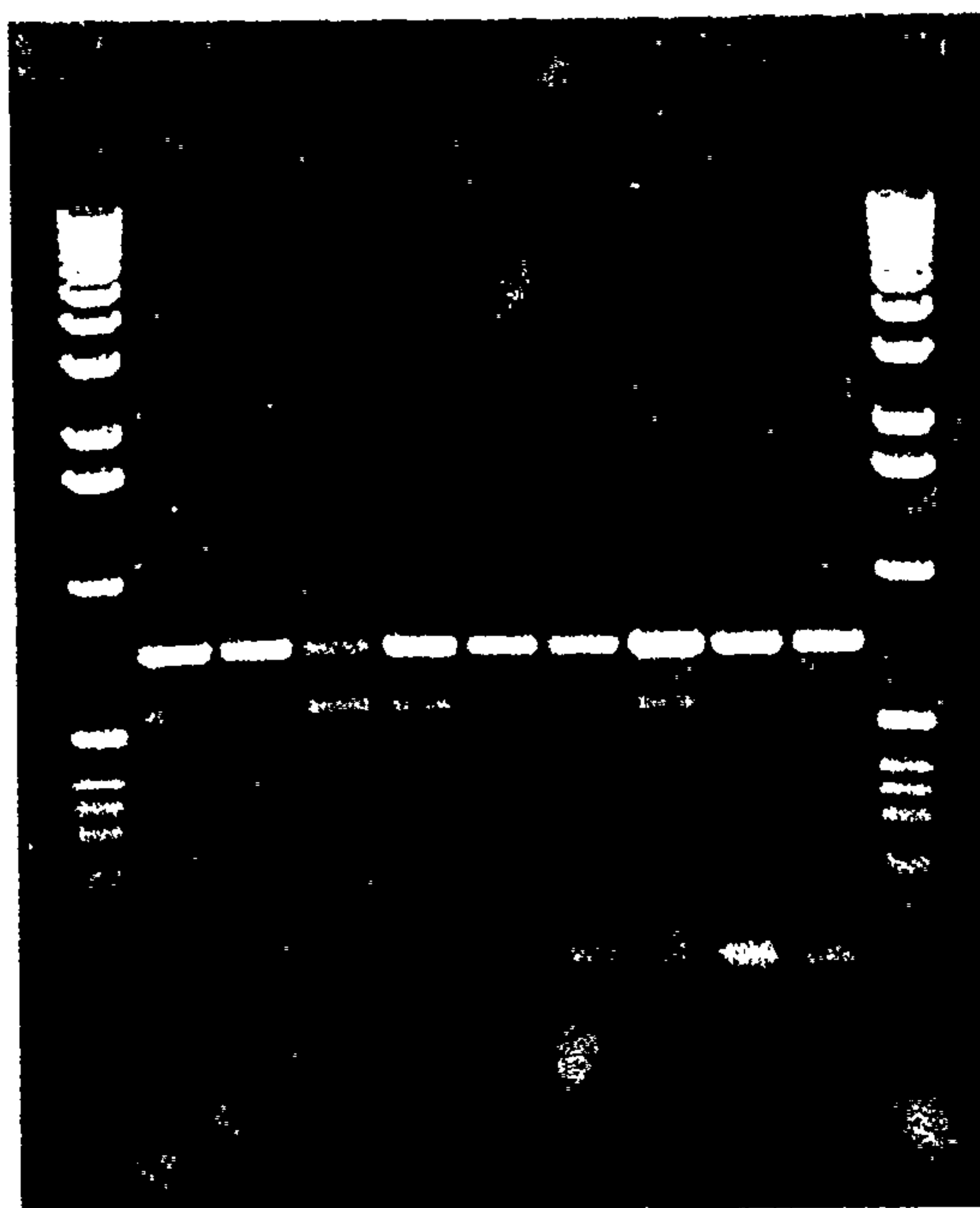
26/31

FIG-8



27/31

FIG-9



K 6.-7 genotype

.....A GCC GTG AAC TTT GGC GCT																	
575	Ala	Gly	Ser	Ile	Met	Ala	His	Glu	Leu	Leu	His	Ile	Phe	Tyr	Gln	589	
1843	GCT	GGC	AGC	ATC	ATG	GCC	CAC	GAG	CTG	TTG	CAC	ATC	TTC	TAC	CAG	1887	
Exon 16 Exon 17																	
590	Leu	Leu	Leu	Pro	Gly	Gly	Cvs	Leu	Ala	Cvs	Asp	Asn	His	Ala	Leu	604	
1888	CTC	TTA	CTG	CCT	GGG	GGC	TGC	CTC	GCC	TGT	GAC	AAC	CAT	GCC	CTC	1932	
								CCC		Pro							
605	Gln	Glu	Ala	His	Leu	Cvs	Leu	Lys	Arg	His	Tyr	Ala	Ala	Phe	Pro	619	
1933	CAG	GAA	GCT	CAC	CTG	TGC	CTG	AAG	CGC	CAT	TAT	GCT	GCC	TTT	CCA	1977	
620	Leu	Pro	Ser	Arg	Thr	Ser	Phe	Asn	Asp	Ser	Leu	Thr	Phe	Leu	Glu	634	
1978	TTA	CCT	AGC	AGA	ACC	TCC	TTC	AAT	GAC	TCC	CTC	ACA	TTC	TTA	GAG	2022	
														TTG			
635	Asn	Ala	Ala	Asp	Val	Gly	Gly	Leu	Ala	Ile	Ala	Leu	Gln				649
2023	AAT	GCT	GCA	GAC	GTT	GGG	GGG	CTA	GCC	ATC	GCG	CTG	CAG				

Application number / numéro de demande: 2204636

Figures: 11-12

Pages: 29/31

Unscannable items
received with this application
(Request original documents in File Prep. Section on the 10th floor)

Documents reçu avec cette demande ne pouvant être balayés
(Commander les documents originaux dans la section de préparation des dossiers au
10^{ème} étage)

FIG-13

Molecular basis of K10 (U^a)

EXON 13
1534-1611

AMENDED SHEET

	ctgtgtctctctcag	Val Ala Gln Leu Gln Val Glu Met Gly Ala Ser	484
		GTT GCT CAA CTG CAG GTG GAG ATG GGG GCT TCA	1566
485	Glu Trp Ala Leu Lys Pro Glu Leu Ala Arg Gln		
1567	GAA TGG GCC CTG AAG CCA GAG CTG GCC CAA	Glu GAA GTA Val K10(Ul ^a)	499 1611
	gtggggtccctgtgttttccagctcc.....0.44kb		

New restriction enzyme site created in K10 genotype: AccI

31/31

FIG-14

Molecular basis of Kp^a, Kp^b and Kp^c

EXON 8

856-1044 (Wild-Type Kp^b)

856	Ile Phe Arg Glu Tyr Leu Thr Tyr Leu Asn Gln Leu Gly Thr	259
	cag ATC TTT CGG GAA TAC CTG ACT TAC CTG AAT CAG CTG GGA ACC	897
260	Leu Leu Gly Gly Asp Pro Ser Lys Val Gln Gln His Ser Ser Leu	274
898	TTG CTG GGA GGA GAC CCA AGC AGC AAG GTG CAA GAA CAC TCT TCC TTG	942
275	Ser Ile Ser Ile Thr Ser	289
943	TCA ATC TCC ATC ACT TCA	987
	Arg Leu Phe Gln Phe Phe Leu Arg Pro Leu	
	CGG CTG CTG TTC CAG CAG TTT CTG AGG CCC CTG	
	TGG	
	Trp Kp ^a	
	CAG Kp ^c	
	Gln	
290	Glu Gln Arg Arg Ala Gln Gly Lys Leu Phe Gln Met Val Thr Ile	304
988	GAG CAG CGG CGG GCA CAG GGC GGC AAG CTC CTC TTC CAG ATG GTC ACT ATC	1032
305	Asp Gln Leu Lys	
1033	GAC CAG CTC AAG gtgcctggaactg	

Kp^a; New NlaIII restriction enzyme site created.Kp^c; New PvuII restriction enzyme site created.

AMENDED SHEET