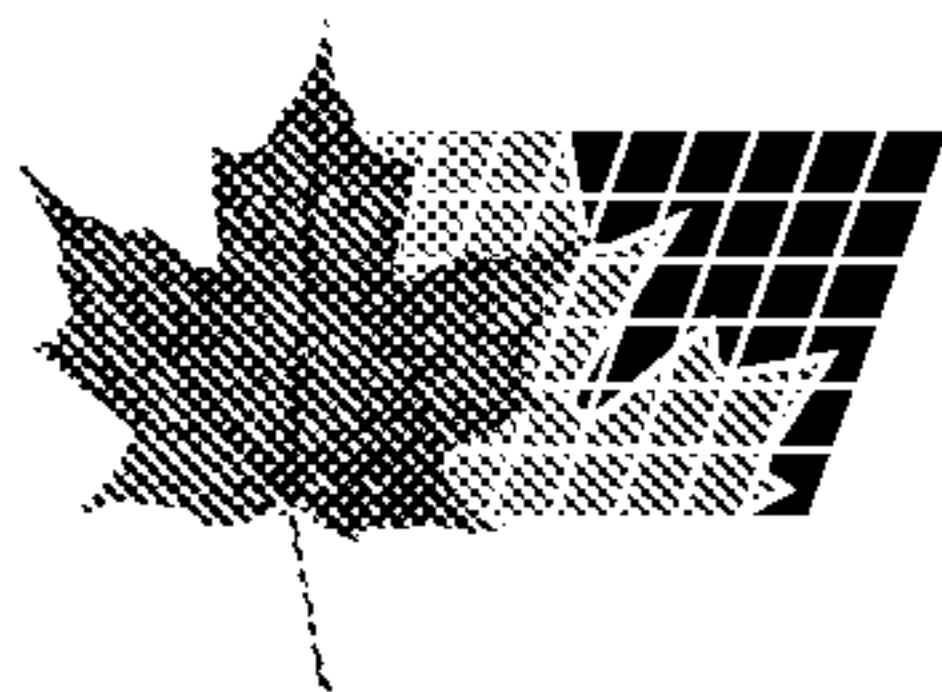




(72) ADAMS, Paul, CA

(72) CHAKRABARTI, Subrata, CA

(72) JEFFREY, Gary, AU

(71) LONDON HEALTH SCIENCES CENTRE, CA

(51) Int.Cl.<sup>6</sup> C07H 21/04, C12Q 1/68

(54) **METHODE DE DIAGNOSTIC DE L'HEMOCHROMATOSE  
HEREDITAIRE**

(54) **METHOD FOR DIAGNOSIS OF HEREDITARY  
HEMOCHROMATOSIS**

(57) A new polymorphism of the HFE gene is described, the presence of which in a subject can lead to misdiagnosis of the subject as a homozygote for the mutation causing hereditary hemochromatosis. A new more discriminating diagnostic method is provided.

**ABSTRACT**

A new polymorphism of the HFE gene is described, the presence of which in a subject can lead to misdiagnosis of the subject as a homozygote  
5 for the mutation causing hereditary hemochromatosis. A new more discriminating diagnostic method is provided.

## **METHOD FOR DIAGNOSIS OF HEREDITARY HEMOCHROMATOSIS**

### **Field of the Invention**

The invention relates to methods for detecting genetic mutations  
5 associated with hereditary hemochromatosis. More particularly, it relates to improved methods for detection of the mutation most commonly associated with the disease.

### **Background of the Invention**

10 Hereditary hemochromatosis (HH) is an inherited disorder of iron metabolism in which excess iron is accumulated in various organs of the body, gradually leading to destruction of these organs and resulting in liver cirrhosis, diabetes and other serious and debilitating diseases. Diagnosis of HH on the basis of symptoms is inadequate because by the time symptoms  
15 appear, organ damage has already occurred. It is therefore desirable to be able to detect subjects at risk for developing HH before the disease develops, so that treatment can be instituted.

The hereditary hemochromatosis gene, or HFE gene, was identified in 1996 as the gene which is the site of mutation in hereditary hemochromatosis  
20 (HH). 83% to 100% of the typical HH subjects have been found to be homozygous for a mutation of the HFE gene at nucleotide 5474, the mutation being 5474 G>A, resulting in the substitution of tyrosine for cysteine, C282Y, in the encoded protein (Feder et al., (1996), Nature Genetics, v. 13, pp. 399-408).

25 A less frequent mutation, 3514 C>G or H63D, has been found in 4% of HH subjects.

The earlier identification of a number of HH – associated genetic markers, followed by the identification of the 5474 G>A mutation, led to the development of diagnostic tests based on detection of these markers or of  
30 that mutation in the DNA of test subjects. Such tests are described, for example, in U.S.P. 5,705,343 and U.S.P. 5,712,098. These tests can be carried out on blood samples or on easily obtainable cells, such as those

from a buccal smear, thus providing non-invasive testing for symptomatic subjects and also opening up the possibility of screening asymptomatic subjects to determine whether or not they carry a disease-related mutation and, if so, whether they are homozygous or heterozygous for the mutation.

5       The present inventors have found that one of the previously described and commonly used tests for the 5474 G>A mutation leads to a high percentage of false positives and they have devised a new and more accurate test for detection of the 5474 G>A mutation.

## 10    Summary of the Invention

      The genomic DNA sequence of the HFE gene is available in GenBank, Accession No. Z 92910. All nucleotide positions of the HFE gene referred to herein are based on numbering from the first nucleotide of the methionine start codon (underlined in Figure 1) as 1.

15       Most studies have determined the prevalence of hemochromatosis through the use of restriction enzyme digestion assays or by oligonucleotide ligation assay (OLA) of amplified HFE genomic DNA obtained by polymerase chain reaction (PCR) using as sense primer,  
5'TGGCAAGGGTAAACAGATCC and as antisense primer, 5'-  
20   CTCAGGCACTCCTCTCAACC (Feder et al., supra). Other genotype assays have been developed using the same PCR conditions and primers in combination with restriction enzyme digestion, since the 5474 G>A mutation produces a Sna B1 site and a Rsa 1 site (Jawinska et al., (1996), Nature Genetics, v. 14, pp. 249-251; Jouanelle et al., (1966), Nature Genetics, v. 14,  
25   pp. 251-252).

      The inventors conducted a population screening of 5211 voluntary blood donors, using the Rsa 1 restriction enzyme digestion assay, and identified thirty-one apparent 5474A homozygotes. Further assessment of these apparent homozygotes showed that many had no phenotypic evidence  
30   of HH; for example, they had normal transferrin saturation and normal serum ferritin. These thirty-one subjects were further examined by genomic DNA sequence analysis and only sixteen were confirmed to be 5474A

homozygotes. Fifteen were heterozygous for the 5474 G>A mutation. As heterozygotes are not at risk for HH, it is important that such subjects are not misidentified and treated as if they were at risk for HH.

The inventors have found, unexpectedly, that each of the fifteen  
5 misdiagnosed 5474A heterozygotes was also heterozygous for a previously unrecognized single base-pair polymorphism at nucleotide position 5569, the polymorphism being 5569 G>A. Nucleotide position 5569 is located within the binding region of the antisense primer used in the previously described diagnostic assays, leading to failure of the primer to bind and a false positive  
10 result.

The inventors have devised a new diagnostic assay for determining the presence or absence of the 5474 G>A mutation of the HFE gene, using an antisense primer, 5'-TACCTCCTCAGGCACTCCTC, which flanks the 5474 nucleotide position but does not span the 5569 nucleotide position. The  
15 fifteen subjects who appeared to be homozygous by the Rsa 1 restriction enzyme assay but heterozygous by direct DNA sequencing were confirmed to be 5474A heterozygotes by the new diagnostic assay. The newly described base-pair polymorphism, 5569 G>A, is not associated with HH and was found in 21% of normal subjects (n = 113).

20 In accordance with one embodiment, the invention provides a method to determine whether a subject is homozygous for the hereditary hemochromatosis (HFE) gene mutation 5474 G>A, the method comprising:

- (a) obtaining a genomic DNA sample from the subject;
- (b) amplifying the DNA sample by polymerase chain reaction  
25 (PCR), using suitable oligonucleotide primers to amplify a segment of the gene including nucleotide position 5474 to give an amplified product, wherein the primers do not span nucleotide position 5569 of the HFE gene; and
- (c) subjecting the amplified product from step (b) to a suitable  
30 assay to detect a component thereof indicative of mutant allele amplification and a component thereof indicative of normal allele amplification,

wherein the detection of a component indicative of mutant allele amplification but no detection of a component indicative of normal allele amplification indicates that the subject is homozygous for the mutation 5474 G>A.

Step (c) may comprise subjecting the amplified product from step (b) to  
5 digestion with a restriction enzyme which recognizes the mutant 5474A allele sequence and analysing the digestion product for at least one nucleotide fragment indicative of mutant allele amplification and at least one fragment indicative of normal allele amplification, wherein the detection of the at least one nucleotide fragment indicative of mutant allele amplification but no  
10 detection of the at least one fragment indicative of normal allele amplification indicates that the subject is homozygous for the mutation 5474 G>A.

If the restriction enzyme employed is Rsa 1, and the primers described in the examples herein are used, the digestion product may be analysed for nucleotide fragments of 146 bp and 117 bp. The presence of a single  
15 nucleotide fragment of 117 bp indicates that the subject is homozygous for the mutation 5474 G>A.

In accordance with a further embodiment, there is provided a method for diagnosing hereditary hemochromatosis (HH) in a subject, the method comprising determining whether the genome of the subject has 0, 1 or 2  
20 alleles of a mutant form of the HFE gene.

Using the primers of the invention and Rsa 1, and looking for the presence of a 146 bp band and a 117 bp band, the presence of only a 117 band is indicative of two alleles having a 5474 G>A mutation, the presence of a 146 bp band and a 117 bp band is indicative of one allele having the  
25 mutation and the presence of only a 146 bp band is indicative of neither allele having the mutation.

The invention also provides an isolated polynucleotide comprising at least 12 to 12146 nucleotides of the human HFE gene (Sequence ID NO:1), wherein the polynucleotide includes nucleotide position 5569 and wherein  
30 nucleotide 5569 is A or a complement thereof.

The invention further includes a DNA primer pair that flanks nucleotide position 5569 of the HFE gene and a composition for use in diagnosing HH,

the composition comprising one or more primers adapted to amplify a segment of a HFE gene wherein at least one of the primers is complementary to a sequence within 100 bp downstream of nucleotide position 5569 of the HFE gene.

- 5           The invention further provides a kit for diagnosis of HH comprising nucleotide primers adapted to distinguish between 5474A homozygotes and 5474A/5569A heterozygotes.

### **Summary of the Drawings**

- 10           Certain embodiments of the invention are described, reference being made to the accompanying drawings, wherein:

Figure 1 shows the genomic DNA sequence of the HFE gene (GenBank Accession No. Z 92910: Sequence ID NO:1).

- 15           Figure 2A shows a portion of the genomic DNA sequence of the HFE gene flanking position 5569;

Figure 2B shows the newly described 5569 polymorphism.

- 20           Figure 3 shows a schematic representation of the HFE gene and the position of PCR primers designed for the amplification of genomic DNA containing the 5474A (C282Y) mutation in exon 4. The binding site and nucleotide sequence of the intron 3 sense primer is shown. The binding site and nucleotide sequence of the previously known antisense primer is also shown. The binding site of this antisense primer a includes the 5569 G>A polymorphism in intron 4. The binding site of the antisense primer b of the present invention excludes this base pair polymorphism.

- 25           Figure 4 shows a scan of a gel of a Rsa 1 digest of amplified DNA from four 5474A/5569A compound heterozygote individuals (1 to 4).

- (a)    PCR amplification using intron 3 sense primer and intron 4 antisense primer that spans the 5569 G>A polymorphism site;
- (b)    PCR amplification using intron 3 sense primer and intron 4 antisense primer that does not span the 5569 G>A polymorphism site;
- 30

Cys: 146 bp fragment; Tyr: 117 bp fragment

**Detailed Description of the Invention**

DNA samples were obtained from 5,211 voluntary blood donors after informed consent. Thirty-one apparent 5474A homozygotes were initially identified using Rsa 1 enzyme digestion assay and amplified DNA obtained using primers as described by Feder et al. (supra). Direct DNA sequencing was also performed, since many of the apparent 5474A homozygotes had a normal serum transferrin saturation. Only 16 were confirmed to be 5474A homozygous by sequence analysis and the remaining 15 were heterozygous for this mutation. All fifteen 5474A heterozygotes were homozygous for 3514C (wild type). It was found by the inventors that each of the fifteen 5474A heterozygotes were also heterozygous for a previously undescribed 5569 G>A single nucleotide polymorphism of the HFE gene. The relevant portion of the sequence is shown in Figures 2A and 2B. The 5569A polymorphism was not present in the sixteen 5474A homozygotes.

The newly discovered 5569A polymorphism is located five nucleotides within the 3' end of the binding site for the antisense primer designed for use in the previously described amplification of the genomic DNA surrounding the 5474A mutation. It is likely that when the 5569A polymorphism is present, it disrupts the binding of the 3' end of the antisense primer and prevents Taq polymerase synthesis of the antisense DNA strand. Analysis of a 5474A/5569A compound heterozygote therefore results in amplification of the 5474A-containing allele alone and gives a false positive homozygous result using any PCR based assay that uses a primer which overlaps the 5569 position. Using an antisense primer (5'TACCTCCTCAGGCACTCCTC) that did not bind the 5569 G>A site, and repeating the Rsa 1 genotype assay on the fifteen 5474A/5569A compound heterozygotes, the inventors clearly identified that all fifteen subjects were 5474A heterozygotes. All 5474A/5569A compound heterozygotes had a transferrin saturation in the normal range (mean 31%, range 20–40%).

Interestingly, DNA from two 5569A homozygotes which initially did not amplify with the previously described primers were only later identified using

the new assay.

Use of the above-mentioned antisense primer of the present invention results in a 393 bp product. Rsa 1 digestion of this product from either a normal allele or a 5569A mutant allele yields a 146 bp band. On the other hand, when a 5474A mutant allele product is reacted with Rsa 1 there is a 117 bp band rather than a 146 bp band. Thus, the method of the present invention allows one to determine whether an individual is homozygous or heterozygous for the 5474A mutation. The same principle will apply to digestion with any restriction enzyme which has a unique site generated as a result of a mutation within the amplified region. It is the identification of the polymorphism at the 5569<sup>th</sup> bp that allows for the development of the improved method of the present invention.

It is clearly apparent to one skilled in the art that PCR primers other than the specific primer outlined above could be used to distinguish between homozygotes and heterozygotes with the proviso that the primer does not span the 5569 position. For example, primers that bind upstream or downstream of the 5569 site could be used to amplify the 5474 region. Primers that bind up to 50 bp upstream or 300 bp downstream are preferred. Clearly, if the position of the primer is adjusted the size of the restriction enzyme fragment generated from the amplified product will vary accordingly. However, the fragments generated from the normal product and the 5569A product will remain the same relative to each other and the fragments generated from the 5474 mutant allele will differ in size from those generated from the normal and 5569A alleles.

Other experiments using the primers previously described by Feder et al. (supra) at annealing temperatures of 62°C, 60°C, 58°C and 56°C were performed with DNA from five 5474A/5569A compound heterozygotes. Only one of these was correctly identified as heterozygous for 5474A by the Rsa 1 assay at an annealing temperature of 60°C or less. The prevalence of the 5569 G>A polymorphism was determined by DNA sequence analysis in 113 normal individuals who were homozygous for 5474G and 3514C. Of these, 24 (21%) were heterozygous for 5569A, corresponding to an allele frequency

of 0.106 of this sample. Interestingly, among 288 alleles so far sequenced, no instance has been observed of the 5569A polymorphism occurring on an allele that harbours the 5474A mutation.

Hill and Robertson's maximum likelihood estimate of linkage disequilibrium  $D$  was 0.71 ( $P < 0.005$ ), which confirmed the presence of moderate to strong linkage disequilibrium between the two variant sites. If linkage disequilibrium between 5569A and 5474A exists, it could mean that the number of 5474G alleles has been systematically underestimated, and thus the number of 5474A homozygotes has been overestimated. This, in turn, could contribute to the discrepancy between the reported prevalence of 5474A homozygotes, namely 1 in 100 to 300, and the reported phenotypic prevalence of hemochromatosis, namely 1 in 200 to 500 (Adams, P.C., (1999), Hepatology, v. 29, p1324). In the study described herein, the prevalence of hemochromatosis was reduced from 1 in 168 to 1 in 327 by the use of the new primers.

These results have major public health implications regarding the utility of population screening for hemochromatosis. Individuals previously considered to be non-expressing 5474A homozygotes on the basis of a PCR-based assay using the previously described primers require re-screening to rule out polymorphism at nucleotide position 5569. For example, this applies to the commonly used assays involving restriction enzyme digestion with either RsaI or Sna B 1.

The diagnostic method of the invention may be carried out on any suitable genomic DNA-containing sample obtained from a test subject. As is understood by those skilled in the art, an RNA sample may also be employed, with reverse transcription to give the cDNA. Suitable samples from which to obtain genomic DNA include buccal smears, peripheral blood samples and tissue samples.

The DNA is extracted from the sample and PCR amplification is carried out by conventional methods, well known to those skilled in the art.

The invention includes nucleotide sequences complementary to Sequence ID No. 2.

The invention also includes primer pairs for PCR amplification that flank the 5474 nucleotide position of the HFE gene but do not span the 5569 nucleotide position of the gene. The identification of the polymorphism at position 5569 by the present inventors enables those skilled in the art to design suitable primers in addition to those specifically disclosed herein. The polymorphism disclosed herein also permits the design of suitable amplification primers flanking the 5569 position, enabling the identification of the 5569 polymorphism in a subject.

## 10 EXAMPLES

The examples are described for the purposes of illustration and are not intended to limit the scope of the invention.

### 15 Example 1

Peripheral blood samples were obtained from 5211 voluntary blood donors and DNA was extracted as described by Cogswell et al., (1998), Ann. Int. Med., v. 129, pp. 971-979.

A Rsa 1 genotype assay, as described by Jouanelle A-M, et al., (1996), Nature Genetics, v. 14, pp. 251-252, was carried out on the samples. A PCR amplification was performed using the previously described HFE intron 3 sense primer (5'TGGCAAGGGTAAACAGATCC) and HFE intron 4 antisense primer (5'CTCAGGCACTCCTCTCAACC) which includes the 5569 G>A polymorphism site.

25 PCR parameters were: 94°C for 5 min, 85°C for 5 min, then 40 cycles of 94°C for 1 min, 62°C for 1 min, 72°C for 1 min, and finally 72°C for 10min. The PCR product was digested with Rsa 1 and separated on an agarose gel.

The presence of the 5474 G>A mutation was determined for all samples.

### 30 Example 2

Thirty one subjects who were determined to be apparent 5474A

homozygotes by the assay of Example 1 were re-examined by conventional DNA sequence analysis of DNA samples obtained from peripheral blood as described in Example 1.

An HFE exon 4 sense primer (5'-AAGCAGCCAATGGATGCCAAGC-3')  
5 and an HFE exon 5 antisense primer (5'TCCAATGAACAAGATGACGAC-3')  
were used to amplify by PCR a 243 bp fragment. The reaction was  
denatured at 94°C for 5min, 85°C for 5 min, then 40 cycles of 94°C for 1 min,  
62°C for 1 min, 72°C for 1 min and finally 72°C for 10 min. DNA sequencing  
was performed using the ABI Prism™, BigDye™ Terminator Cycle  
10 Sequencing Ready Reaction Kit according to the manufacturer's instructions  
(Perkin Elmer Applied Biosystems, Mississauga, Ontario, Canada) by the J P  
Robarts Research Institute Sequencing Facility (London, Ontario, Canada).  
Linkage disequilibrium was estimated using Hill and Robertson's  
standardized D statistic, as described by Hegele et al., (1990), Genet.  
15 Epidemiol., v. 7, pp. 69-81.

Sixteen of the thirty one were confirmed to be 5474A homozygotes and  
fifteen were 5474A heterozygotes.

### **Example 3**

The fifteen subjects who were determined to be 5474A heterozygotes  
20 by the assay described in Example 2 were re-examined by the  
PCR/restriction enzyme digestion assay described in Example 1, but using as  
antisense primer the new HFE intron 4 antisense primer (5'-  
TACCTCCTCAGGCACTCCTC) which does not overlap the 5569 G>A  
polymorphism site (Figure 3). All fifteen subjects were clearly identified by  
25 this assay as 5474A heterozygotes.

### **Example 4**

A Rsa 1 genotype assay was carried out on the DNA of four  
5474A/5569A compound heterozygotes, using for PCR amplification either

the antisense primer described in Example 1, which includes the 5569 G>A polymorphism site, or the antisense primer described in Example 3, which does not include the 5569 G>A polymorphism site. The results are shown in Figure 4. (a) PCR reaction using intron 3 sense primer and intron 4  
5 antisense primer that spans the 5569 G>A polymorphism site. (b) PCR reaction using intron 3 sense primer and the new intron 4 antisense primer that does not span the 5569 G>A polymorphism site.

Using the previously described primers which span the 5569 G>A site, the PCR product is 387 bp. The PCR product from a normal HFE allele is  
10 cleaved by Rsa 1 to give 247 bp and 140 bp fragments. The product of the 5474A mutant allele is cleaved to give 247 bp, 111 bp and 29 bp fragments. A 5474A homozygote shows the 111 bp band on a gel, whereas a 5474A heterozygote should show both the 140 bp and 111 bp bands, allowing identification of the heterozygote.

15 The presence of the 5569A polymorphism in a heterozygote, however, results in the production of a 111 bp fragment, rather than both a 111 bp fragment and a 140 bp fragment since the normal allele is not amplified. Since only this 111 bp fragment is detected, the subject is presumed falsely to be a 5474A homozygote, as seen in Figure 3 lanes 1a, 2a, 3a and 4a.

20 Using the new antisense primer, which does not span the 5569 site, the PCR product is 393 bp. The product from a normal allele is cleaved by Rsa 1 to give 247 bp and 146 bp fragments. The product of the 5474A mutant allele gives 247 bp, 117 bp and 29 bp fragments. In the presence of the 5569A polymorphism, and the absence of a 5474A mutant allele, 247 bp  
25 and 146 bp fragments are generated. A 5474A mutant homozygote shows only a 117 bp band, which is distinguishable from the 5474A heterozygote with 5569A polymorphism which shows two bands, 146 bp and 117 bp, as seen in Figure 3 (b) lanes. Thus, the method of the present invention has surprising advantages over the method currently available in that 5474A  
30 heterozygotes can be accurately distinguished from 5474A homozygotes.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: London Health Sciences Centre
- (ii) TITLE OF INVENTION: METHOD FOR DIAGNOSIS OF HEREDITARY HEMOCHROMATOSIS
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Sim & McBurney
  - (B) STREET: 6th Floor, 330 University Avenue
  - (C) CITY: Toronto
  - (D) PROVINCE: Ontario
  - (E) COUNTRY: Canada
  - (F) POSTAL CODE: M5G 1R7
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: 2,272,410
  - (B) FILING DATE: May 21, 1999
- (vii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Rae, Patricia A. (Dr.)
  - (B) REFERENCE: 4767-97/PAR
- (viii) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (416) 595-1155
  - (B) TELEFAX: (416) 595-1163

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12146 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

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

```

ggatccttta accgaggaga ttattatagc cggagctctg aagcagcaat ctcagttcctt 60
gtgatagtga gcaaagaact acaaactaac accaaaatgc aagcttaaag caaagtttat 120
tgaagcacia taatacactc tgagggacag cgggcttatt tctgcgaagt gaactcagca 180
cttctttaca gagctcaagg tgcttttatg gggtttggtg ggaggagttg aggtttgggc 240
tgtatctgag tgacaggatg atgttatctg attgaagttt atagctatac aatctaaaat 300
taaactgtgc atggtcttac ctataatttg ttaagaaaag cctcccaggg atgggggggc 360
aaaactgtat gtaaattcta ttataatgat ggcatgatga acttgggggtg aacttgaaga 420
caggcttttg tggttggtgg catgtgccac cttaggggat ttccacctgt accctccttt 480
ctctttctcc aggatatttt ggccacagac tttatcataa actccatccc ttaggggtggc 540
attagggtag tcttgggcct gaatttaggt gggccagtgg ctgtcttagt gacagccttt 600

```

|            |             |             |             |             |             |      |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| ccgctctctt | ctgtcatccc  | ctcccaactg  | ctaattgtcta | actacctaac  | aattacccat  | 660  |
| taaatcagtg | tgtctggggg  | taggagcagg  | cctcaatatg  | tttaatcatt  | ctccagataa  | 720  |
| tcccaatact | gtaaagtttg  | tgaaacactt  | gtcagataat  | tcaattatga  | aggctgtgga  | 780  |
| acgtgtttca | gtaggatcta  | attggttaat  | gttatgactt  | aattaatttg  | aatcaaaaaa  | 840  |
| caaaatgaaa | aagctttata  | tttctaagtc  | aaataagaca  | taagttggtc  | taaggttgag  | 900  |
| ataaaatttt | taaatgtatg  | attgaatttt  | gaaaatcata  | aatattttaa  | tatctaaagt  | 960  |
| tcagatcaga | acattgcgaa  | gctactttcc  | ccaatcaaca  | acaccccttc  | aggattttaa  | 1020 |
| aaccaagggg | gacactggat  | cacctagtgt  | ttcacaagca  | ggtaccttct  | gctgtaggag  | 1080 |
| agagagaact | aaagttctga  | aagacctgtt  | gcttttcacc  | aggaagtttt  | actgggcatc  | 1140 |
| tcctgagcct | aggcaatagc  | tgtaggggtga | cttctggagc  | catccccgtt  | tccccgcccc  | 1200 |
| ccaaaagaag | cggagattta  | acggggacgt  | gcggccagag  | ctgggggaaat | ggggccgcga  | 1260 |
| gccaggccgg | cgcttctcct  | cctgatgctt  | ttgcagaccg  | cggtcctgca  | ggggcgcttg  | 1320 |
| ctgcgtgagt | ccgagggctg  | cgggcgaact  | aggggcgcgg  | cgggggtgga  | aaaatcgaaa  | 1380 |
| ctagcttttt | ctttgcgctt  | gggagtttgc  | taactttgga  | ggacctgctc  | aaccaatcc   | 1440 |
| gcaagcccct | ctccctactt  | tctgcgtcca  | gaccccgctga | gggagtgcct  | accactgaac  | 1500 |
| tgcagatagg | ggtccctcgc  | cccaggacct  | gccccctccc  | ccggctgtcc  | cggctctgcg  | 1560 |
| gagtgacttt | tggaaccgcc  | cactcccttc  | ccccaaactag | aatgctttta  | aataaatctc  | 1620 |
| gtagttcctc | acttgagctg  | agctaagcct  | ggggctcctt  | gaacctggaa  | ctcgggttta  | 1680 |
| tttccaatgt | cagctgtgca  | gttttttccc  | cagtcactctc | caaacaggaa  | gttcttccct  | 1740 |
| gagtgccttg | cgagaaggct  | gagcaaacc   | acagcaggat  | ccgcacgggg  | tttccacctc  | 1800 |
| agaacgaatg | cgttgggcgg  | tgggggcgcg  | aaagagtggc  | gttggggatc  | tgaattcttc  | 1860 |
| accattccac | ccacttttgg  | tgagacctgg  | ggtggagggtc | tctaggggtg  | gaggctcctg  | 1920 |
| agagaggcct | acctcggggc  | tttccccact  | cttggcaatt  | gttcttttgc  | ctggaaaatt  | 1980 |
| aagtatatgt | tagttttgaa  | cgtttgaact  | gaacaattct  | cttttcggct  | aggctttatt  | 2040 |
| gatttgcaat | gtgctgtgta  | attaagaggc  | ctctctacaa  | agtactgata  | atgaacatgt  | 2100 |
| aagcaatgca | ctcacttcta  | agttacattc  | atatctgatc  | ttatttgatt  | ttcactaggc  | 2160 |
| atagggaggt | aggagcta    | aatacgttta  | ttttactaga  | agttaactgg  | aattcagatt  | 2220 |
| atataactct | tttcagggtta | caaagaacat  | aaataatctg  | gttttctgat  | gttatttcaa  | 2280 |
| gtactacagc | tgcttcta    | cttagttgac  | agtgattttg  | ccctgtagtg  | tagcacagtg  | 2340 |
| ttctgtgggt | cacacgccgg  | cctcagcaca  | gcactttgag  | ttttgggtact | acgtgtatcc  | 2400 |
| acattttaca | catgacaaga  | atgaggcatg  | gcacggcctg  | cttcctggca  | aattttattca | 2460 |
| atggtacacg | gggctttggg  | ggcagagctc  | atgtctccac  | ttcatagcta  | tgattcttaa  | 2520 |
| acatcacact | gcattagagg  | ttgaataata  | aaatttcatg  | ttgagcagaa  | atattcattg  | 2580 |
| tttacaagtg | taaatgagtc  | ccagccatgt  | gttgcactgt  | tcaagcccca  | agggagagag  | 2640 |
| cagggaacaa | agtctttacc  | ctttgatatt  | ttgcattcta  | gtgggagaga  | tgacaataag  | 2700 |
| caaatgagca | gaaagatata  | caacatcagg  | aaatcatggg  | tgttgtgaga  | agcagagaag  | 2760 |
| tcagggcaag | tcactctggg  | gctgacactt  | gagcagagac  | atgaaggaaa  | taagaatgat  | 2820 |
| attgactggg | agcagtat    | cccaggcaaa  | ctgagtgggc  | ctggcaagtt  | ggattaaaaa  | 2880 |
| gcgggttttc | tcagcactac  | tcatgtgtgt  | gtgtgtgggg  | gggggggcgg  | cgtgggggtg  | 2940 |
| ggaaggggga | ctaccatctg  | catgtaggat  | gtctagcagt  | atcctgtcct  | ccctactcac  | 3000 |
| taggtgctag | gagcactccc  | ccagtcttga  | caacccaaaa  | tgtctctaaa  | ctttgccaca  | 3060 |
| tgtcacctag | tagacaaact  | cctgggttaag | aagctcgggt  | tgaaaaaaat  | aaacaagtag  | 3120 |
| tgctggggag | tagaggccaa  | gaagtaggta  | atgggctcag  | aagaggagcc  | acaacaagg   | 3180 |
| ttgtgcaggc | gcctgtaggc  | tgtgggtgtga | attctagcca  | aggagtaaca  | gtgatctgtc  | 3240 |
| acaggctttt | aaaagattgc  | tctggctgct  | atgtggaaag  | cagaatgaag  | ggagcaacag  | 3300 |
| taaaagcagg | gagcccagcc  | aggaagctgt  | tacacagtcc  | aggcaagagg  | tagtgagagt  | 3360 |
| ggctgggtgg | gaacagaaaa  | gggagtgaca  | aaccattgtc  | tcctgaatat  | attctgaagg  | 3420 |
| aagttgctga | aggattctat  | gttgtgtgag  | agaaagagaa  | gaattggctg  | ggtgtagtag  | 3480 |
| ctcatgccaa | ggaggaggcc  | aaggagagca  | gattcctgag  | ctcaggagtt  | caagaccagc  | 3540 |
| ctgggcaaca | cagcaaaacc  | ccttctctac  | aaaaaataca  | aaaattagct  | gggtgtgggt  | 3600 |
| gcatgcacct | gtgatcctag  | ctactcggga  | ggctgaggtg  | gagggtattg  | cttgagccca  | 3660 |
| ggaagttgag | gctgcagtga  | gccatgactg  | tgccactgta  | cttcagccta  | ggtgacagag  | 3720 |
| caagaccctg | tctcccctga  | ccccctgaaa  | aagagaagag  | ttaaagttga  | ctttgttctt  | 3780 |
| tattttaatt | ttattggcct  | gagcagtggg  | gtaattggca  | atgccatttc  | tgagatgggt  | 3840 |
| aaggcagagg | aaagagcagt  | ttggggtaaa  | tcaaggatct  | gcatttggac  | atgttaagtt  | 3900 |
| tgagattcca | gtcaggcttc  | caagtgggtga | ggccacatag  | gcagttcagt  | gtaagaattc  | 3960 |
| aggaccaagg | cagggcacgg  | tggctcactt  | ctgtaatccc  | agcactttgg  | tggctgaggg  | 4020 |
| aggtagatca | tttgagggtca | ggagtgttag  | acaagcttgg  | ccaacatggg  | gaaaccccat  | 4080 |
| gtctactaaa | aatacaaaaa  | ttagcctggg  | gtgggtggcg  | acgcctatag  | tcccagggtt  | 4140 |
| tcaggaggct | taggtaggag  | aatcccttga  | accaggagg   | tgcaggttgc  | agtgagctga  | 4200 |
| gattgtgcca | ctgcactcca  | gcctgggtga  | tagagtgaga  | ctctgtctca  | aaaaaaaaaa  | 4260 |

|             |            |             |            |             |             |      |
|-------------|------------|-------------|------------|-------------|-------------|------|
| aaaaaaaaa   | aaaaaaaaa  | aactgaagga  | attattcctc | aggatttggg  | tctaatttgc  | 4320 |
| cctgagcacc  | aactcctgag | ttcaactacc  | atggctagac | acaccttaac  | atttttctaga | 4380 |
| atccaccagc  | tttagtggag | tctgtctaata | catgagtatt | ggaataggat  | ctgggggag   | 4440 |
| tgaggggggtg | gcagccacgt | gtggcagaga  | aaagcacaca | aggaaagagc  | acccaggact  | 4500 |
| gtcatatgga  | agaaagacag | gactgcaact  | cacccttcac | aaaatgagga  | ccagacacag  | 4560 |
| ctgatggtat  | gagttgatgc | aggtgtgtgg  | agcctcaaca | tcctgctccc  | ctcctactac  | 4620 |
| acatgggttaa | ggcctggtgc | tctgtctcca  | ggttcacact | ctctgcacta  | cctcttcatg  | 4680 |
| ggtgcctcag  | agcaggacct | tgggtctttcc | ttgtttgaag | ctttgggcta  | cgtggatgac  | 4740 |
| cagctgttcg  | tgttctatga | tcagagaggt  | cgccgtgtgg | agccccgaac  | tccatggggt  | 4800 |
| tccagtagaa  | tttcaagcca | gatgtggctg  | cagctgagtc | agagtctgaa  | aggggtggat  | 4860 |
| cacatgttca  | ctgttgactt | ctggactatt  | atggaaaatc | acaaccacag  | caagggtatg  | 4920 |
| tggagagggg  | gcctcacctt | cctgaggttg  | tcagagcttt | tcactctttc  | atgcatcttg  | 4980 |
| aaggaaacag  | ctggaagtct | gaggtccttg  | gggagcaggg | aagaggggag  | gaatttgctt  | 5040 |
| cctgagatca  | tttgggtcct | ggggatggtg  | gaaataggga | cctattcctt  | tgggtgcagt  | 5100 |
| taacaaggct  | ggggattttt | ccagagtccc  | acaccctgca | ggtcatcctg  | ggctgtgaaa  | 5160 |
| tgcaagaaga  | caacagtacc | gagggctact  | ggaagtacgg | gtatgatggg  | caggaccacc  | 5220 |
| ttgaattctg  | ccctgacaca | ctggattgga  | gagcagcaga | acccagggcc  | tggcccacca  | 5280 |
| agctggagtg  | ggaaaggcac | aagattcggg  | ccaggcagaa | cagggcctac  | ctggagaggg  | 5340 |
| actgccctgc  | acagctgcag | cagttgctgg  | agctggggag | aggtgttttg  | gaccaacaag  | 5400 |
| gtatggtgga  | aacacacttc | tgcccctata  | ctctagtggc | agagtggagg  | aggttgcagg  | 5460 |
| gcacggaatc  | cctgggttga | gtttcagagg  | tggctgaggg | tgtgtgcctc  | tccaaattct  | 5520 |
| gggaagggac  | tttctcaatc | ctagagtctc  | taccttataa | ttgagatgta  | tgagacagcc  | 5580 |
| acaagtcagt  | ggtttaattt | cttttctcca  | tgcataatgg | tcaaagggaa  | gtgtctatgg  | 5640 |
| cccttgcttt  | ttatttaacc | aataatcttt  | tgtatattta | tacctgttaa  | aaattcagaa  | 5700 |
| atgtcaaggc  | cgggcacggg | ggctcacccc  | tgtaatccca | gcactttggg  | aggccgaggc  | 5760 |
| gggtggtcac  | aaggtcagga | gtttgagacc  | agcctgacca | acatgggtgaa | acccgtctct  | 5820 |
| aaaaaaatac  | aaaaattagc | tgggtcacagt | catgctgacc | tgtagtccca  | gctaatttga  | 5880 |
| aggctgaggg  | aggagcatcg | cttgaacctg  | ggaagcggaa | gttgactgta  | gccaagatcg  | 5940 |
| cgccactgca  | ctccagccta | ggcagcagag  | tgagactcca | tcttaaaaaa  | aaaaaaaaaa  | 6000 |
| aaaaagagaa  | ttcagagatc | tcagctatca  | tatgaatacc | aggacaaaat  | atcaagttag  | 6060 |
| gccacttatc  | agagtagaag | aatcctttag  | gttaaaagtt | tctttcatag  | aacatagcaa  | 6120 |
| taatcactga  | agctacctat | cttacaagtc  | cgcttcttat | aacaatgcct  | cctaggttga  | 6180 |
| cccaggtgaa  | actgaccatc | tgtattcaat  | cattttcaat | gcacataaag  | ggcaatttta  | 6240 |
| tctatcagaa  | caaagaacat | gggtaacaga  | tatgtatat  | tacatgtgag  | gagaacaagc  | 6300 |
| tgatctgact  | gctctccaag | tgacactgtg  | ttagagtcca | atcttaggac  | acaaaatggg  | 6360 |
| gtctctcctg  | tagcttggtt | ttttctgaaa  | agggtatttc | cttcctccaa  | cctatagaag  | 6420 |
| gaagtgaag   | ttccagtctt | cctggcaagg  | gtaaacagat | ccctctcct   | catecttctt  | 6480 |
| ctttcctgtc  | aagtgcctcc | tttgggtgaag | gtgacacatc | atgtgacctc  | ttcagtgacc  | 6540 |
| actctacggg  | gtcgggcctt | gaactactac  | ccccagaaca | tcaccatgaa  | gtggctgaag  | 6600 |
| gataagcagc  | caatggatgc | caaggagttc  | gaacctaaag | acgtattgcc  | caatggggat  | 6660 |
| gggacctacc  | agggttggtg | aaccttggtc  | gtacccctg  | gggaagagca  | gagatatagc  | 6720 |
| tgccaggtgg  | agcaccaggg | cctggatcag  | ccctcatttg | tgatctgggg  | tatgtgactg  | 6780 |
| atgagagcca  | ggagctgaga | aaatctattg  | gggttgtaga | ggagtgcctg  | aggaggtaat  | 6840 |
| tatggcagtg  | agatgaggat | ctgctctttg  | ttaggggatg | ggctgagggg  | ggcaatcaaa  | 6900 |
| ggctttaact  | tgttttttct | gttttagagc  | cctcaccgtc | tggcacccta  | gtcattggag  | 6960 |
| tcactagtg   | aattgctgtt | tttgctgtca  | tcttggtcat | tggaaattttg | ttcataatat  | 7020 |
| taaggaagag  | gcagggttca | agtgagtagg  | aacaaggggg | aagtctctta  | gtacctctgc  | 7080 |
| cccagggcac  | agtgggaaga | ggggcagagg  | ggatctggca | tccatgggaa  | gcatttttct  | 7140 |
| catttatatt  | ctttggggac | accagcagct  | ccctgggaga | cagaaaataa  | tggttctccc  | 7200 |
| cagaatgaaa  | gtctctaatt | caacaaacat  | cttcagagca | cctactattt  | tgcaagagct  | 7260 |
| gtttaaggta  | gtacaggggc | tttgaggttg  | agaagtcact | gtggctattc  | tcagaaccca  | 7320 |
| aatctggtag  | ggaatgaaat | tgatagcaag  | taaatgtagt | taaagaagac  | cccatgaggt  | 7380 |
| cctaaagcag  | gcaggaagca | aatgcttagg  | gtgtcaaagg | aaagaatgat  | cacattcagc  | 7440 |
| tggggatcaa  | gatagccttc | tggatcttga  | aggagaagct | ggattccatt  | aggtgaggtt  | 7500 |
| gaagatgatg  | ggaggtctac | acagacggag  | caacctatgc | aagtaggaga  | gtataaggca  | 7560 |
| tactgggaga  | ttagaaataa | ttactgtacc  | ttaacctga  | gtttgcttag  | ctatcactca  | 7620 |
| ccaattatgc  | atttctaccc | cctgaacatc  | tgtggtgtag | ggaaaagaga  | atcagaaaga  | 7680 |
| agccagctca  | tacagagtcc | aagggtcttt  | tgggatattg | ggttatgatc  | actgggggtg  | 7740 |
| cattgaagga  | tcctaagaaa | ggaggaccac  | gatctccctt | atatgggtgaa | tgtgttggtt  | 7800 |
| agaagttaga  | tgagaggtga | ggagaccagt  | tagaaagcca | ataagcattt  | ccagatgaga  | 7860 |
| gataatgggt  | cttgaaatcc | aatagtgccc  | aggtctaaat | tgagatgggt  | gaatgaggaa  | 7920 |

|             |             |             |             |             |            |       |
|-------------|-------------|-------------|-------------|-------------|------------|-------|
| aataaggaag  | agagaagagg  | caagatggtg  | cctaggtttg  | tgatgcctct  | ttcctgggtc | 7980  |
| tcttgtctcc  | acaggaggag  | ccatggggca  | ctacgtctta  | gctgaacgtg  | agtgcacgc  | 8040  |
| agcctgcaga  | ctcactgtgg  | gaaggagaca  | aaactagaga  | ctcaaagagg  | gagtgcattt | 8100  |
| atgagctctt  | catgtttcag  | gagagagttg  | aacctaaaca  | tagaaattgc  | ctgacgaact | 8160  |
| ccttgatttt  | agccttctct  | gttcatttcc  | tcaaaaagat  | ttccccattt  | aggtttctga | 8220  |
| gttcctgcat  | gccggtgatc  | cctagctgtg  | acctctcccc  | tggaactgtc  | tctcatgaac | 8280  |
| ctcaagctgc  | atctagaggc  | ttccttcatt  | tcctccgtca  | cctcagagac  | atacacctat | 8340  |
| gtcattttcat | ttcctatttt  | tggaagagga  | ctccttaa    | ttgggggact  | tacatgattc | 8400  |
| attttaacat  | ctgagaaaag  | ccttgaaccc  | tgggacgtgg  | ctagtcataa  | ccttaccaga | 8460  |
| tttttacaca  | tgtatctatg  | cattttctgg  | accggttcaa  | cttttccttt  | gaatcctctc | 8520  |
| tctgtgttac  | ccagtaactc  | atctgtcacc  | aagccttggg  | gattcttcca  | tctgattgtg | 8580  |
| atgtgagttg  | cacagctatg  | aaggctgtac  | actgcacgaa  | tggaagaggc  | acctgtccca | 8640  |
| gaaaaagcat  | catggctatc  | tgtgggtagt  | atgatgggtg  | tttttagcag  | gtaggaggca | 8700  |
| aatatcttga  | aaggggttgt  | gaagaggtgt  | tttttcta    | tggtcatgaag | gtgtcataca | 8760  |
| gatttgcaaa  | gtttaatggt  | gccttcattt  | gggatgctac  | tctagtattc  | cagacctgaa | 8820  |
| gaatcacat   | aattttctac  | ctggtctctc  | cttggtctga  | taatgaaaat  | tatgataagg | 8880  |
| atgataaaa   | cacttacttc  | gtgtccgact  | cttctgagca  | cctacttaca  | tgcattactg | 8940  |
| catgcacttc  | ttacaataat  | tctatgagat  | aggtactatt  | atccccattt  | ctttttttaa | 9000  |
| tgaagaaagt  | gaagtaggcc  | gggcacggtg  | gctcacgcct  | gtaatcccag  | cactttggga | 9060  |
| ggccaaagcg  | ggtggatcac  | gaggtcagga  | gatcgagacc  | atcctggcta  | acatggtgaa | 9120  |
| accccatctc  | taataaaaat  | acaaaaaatt  | agctgggcgt  | ggtggcagac  | gcctgtagtc | 9180  |
| ccagctactc  | ggaaggctga  | ggcaggagaa  | tggtcatgaac | ccaggaggca  | gagcttgca  | 9240  |
| tgagccgagt  | ttgcgccact  | gcactccagc  | ctaggtgaca  | gagtgcagac  | ccatctcaaa | 9300  |
| aaaataaaaa  | taaaaataaa  | aaaatgaaaa  | aaaaaagaaa  | gtgaagtata  | gagtatctca | 9360  |
| tagtttgtca  | gtgatagaaa  | caggtttcaa  | actcagtcaa  | tctgaccgtt  | tgatacatct | 9420  |
| cagacaccac  | tacattcagt  | agtttagatg  | cctagaataa  | atagagaagg  | aaggagatgg | 9480  |
| ctcttctctt  | gtctcattgt  | gtttcttctg  | aatgagcttg  | aatcacatga  | aggggaacag | 9540  |
| cagaaaacaa  | ccaactgac   | ctcagctgtc  | atgtttcctt  | taaaagtccc  | tgaaggaagg | 9600  |
| tcctggaatg  | tgactccctt  | gctcctctgt  | tgctctcttt  | ggcattcatt  | tctttggacc | 9660  |
| ctacgcaagg  | actgtaattg  | gtggggacag  | ctagtggccc  | tgctgggctt  | cacacacggt | 9720  |
| gtcctcccta  | ggccagtgcc  | tctggagtca  | gaactctggt  | ggtatttccc  | tcaatgaagt | 9780  |
| ggagtaagct  | ctctcatttt  | gagatgggat  | aatggaagcc  | accaagtggc  | ttagaggatg | 9840  |
| cccaggtcct  | tccatggagc  | cactgggggt  | ccggtgcaca  | ttaaaaaaa   | aatctaacca | 9900  |
| ggacattcag  | gaattgctag  | attctgggaa  | atcagttcac  | catgttcaaa  | agagtctttt | 9960  |
| tttttttttt  | gagactctat  | tgcccaggct  | ggagtgcatt  | ggcatgatct  | cggctcactg | 10020 |
| taacctctgc  | ctcccagggt  | caagcgattc  | tcctgtctca  | gcctcccaag  | tagctgggat | 10080 |
| tacaggcgtg  | caccaccatg  | cccggcta    | ttttgtattt  | ttagtagaga  | cagggtttca | 10140 |
| ccatgttggc  | caggctgggtc | tcgaactctc  | ctgacctcgt  | gatccgcctg  | cctcggcctc | 10200 |
| ccaaagtgct  | gagattacag  | gtgtgagcca  | ccctgccag   | ccgtcaaaag  | agtcttaata | 10260 |
| tatatatcca  | gatggcatgt  | gtttacttta  | tggtactaca  | tgcaactggc  | tgcataaatg | 10320 |
| tggtacaagc  | attctgtctt  | gaagggcagg  | tgcttcagga  | taccatatac  | agctcagaag | 10380 |
| tttcttcttt  | aggcattaaa  | tttttagcaa  | gatatctcat  | ctcttctttt  | aaaccatttt | 10440 |
| ctttttttgt  | ggttagaaaa  | gttatgtaga  | aaaaagtaaa  | tgtgatttac  | gctcattgta | 10500 |
| gaaaagctat  | aaaatgaata  | caattaaagc  | tggtatttaa  | ttagccagtg  | aaaaactatt | 10560 |
| aacaacttgt  | ctattacctg  | ttagtattat  | tggtgcatta  | aaaatgcata  | tactttaata | 10620 |
| aatgtacatt  | gtattgtata  | ctgcatgatt  | ttattgaagt  | tcttgttcat  | cttgtgtata | 10680 |
| tacttaatcg  | ctttgtcatt  | ttggagacat  | ttattttgct  | tctaatttct  | ttacattttg | 10740 |
| tcttacggaa  | tattttcatt  | caactgtggt  | agccgaatta  | atcgtgtttc  | ttcactctag | 10800 |
| ggacattgtc  | gtctaagttg  | taagacattg  | gttattttac  | cagcaaacca  | ttctgaaagc | 10860 |
| atatgacaaa  | ttatttctct  | cttaatatct  | tactatactg  | aaagcagact  | gctataaggc | 10920 |
| ttcacttact  | cttctacctc  | ataaggaata  | tggtacaatt  | aatttattag  | gtaagcattt | 10980 |
| gttttatatt  | ggttttat    | cacctgggct  | gagatttcaa  | gaaacacccc  | agtcttcaca | 11040 |
| gtaacacatt  | tcactaacac  | atttactaaa  | catcagcaac  | tgtggcctgt  | taattttttt | 11100 |
| aatagaaatt  | ttaagtcctc  | attttctttc  | ggtgtttttt  | aagcttaatt  | tttctggctt | 11160 |
| tattcataaa  | ttcttaaggt  | caactacatt  | tgaaaaatca  | aagacctgca  | ttttaaattc | 11220 |
| ttattcacct  | ctggcaaaac  | cattcacaaa  | ccatggtagt  | aaagagaagg  | gtgacacctg | 11280 |
| gtggccatag  | gtaaatgtac  | cacggtgggtc | cggtgaccag  | agatgcagcg  | ctgagggttt | 11340 |
| tcctgaagg   | aaaggaataa  | agaatgggtg  | gaggggcgtg  | cactggaaat  | cacttgtaga | 11400 |
| gaaaagcccc  | tgaaaatttg  | agaaaacaaa  | caagaaacta  | cttaccagct  | atttgaattg | 11460 |
| ctggaatcac  | aggccattgc  | tgagctgcct  | gaactgggaa  | cacaacagaa  | ggaaaacaaa | 11520 |
| ccactctgat  | aatcattgag  | tcaagtacag  | caggtgattg  | aggactgctg  | agaggtagag | 11580 |

```

gccaaaattc ttatgttgta ttataataat gtcattcttat aatactgtca gtattttata 11640
aaacattctt cacaaactca cacacattta aaaacaaaac actgtctcta aaatcccca 11700
atTTTTcata aactcagttt taaactaact ttttttcaaa ccacaatctg atttaacaat 11760
gactatcatt taaatatattc tgactttcaa attaaagatt ttcacatgca ggctgatatt 11820
tgtaattgtg attctctctg taggctttgg gtataatgtg ttcttttcct tttttgcatc 11880
agcgattaac ttctacactc taacatgtag aatgttacta caatattaaa gtattttgta 11940
tgacaatttt atttgaaagc ctaggatgag ttgacatcct gcatgcattt attacttgat 12000
atgcatgcat tctggatatc caagcattct atttctgagt aattgtttta ggtgtagaag 12060
agatagatat ggtggatttg gagttgatac ttatatattt tctatttctt ggatggatga 12120
attgtacat taaaagtttt ccatgg                                     12146

```

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 401 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

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

```

aagcagccaa tggatgccaa ggagttcgaa cctaaagacg tattgccccaa tggggatggg 60
acctaccagg gctggataac cttggctgta cccctggggg aagagcagag atatacgtgc 120
caggtggagc acccaggcct ggatcagccc ctcatgtgta tctgggggat gtgactgatg 180
agagccagga gctgagaaaa tctattgggg gttgagagga gtgcctgagg aggtaattat 240
ggcagtgaga tgaggatctg ctctttgtta ggggatgggc tgagggtggc aatcaaaggc 300
tttaacttgc tttttctggt ttagagccct caccgtctgg caccctagtc attggagtca 360
tcagtggaat tgctgttttt gtcgtcatct tgttcattgg a                               401

```

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 401 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

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

```

aagcagccaa tggatgccaa ggagttcgaa cctaaagacg tattgccccaa tggggatggg 60
acctaccagg gctggataac cttggctgta cccctggggg aagagcagag atatacgtgc 120
caggtggagc acccaggcct ggatcagccc ctcatgtgta tctgggggat gtgactgatg 180
agagccagga gctgagaaaa tctattgggg gttgagagga gtgcctgagg aggtaattat 240
ggcagtgaga tgaggatctg ctctttgtta ggggatgggc tgagggtggc aatcaaaggc 300
tttaacttgc tttttctggt ttagagccct caccgtctgg caccctagtc attggagtca 360
tcagtggaat tgctgttttt gtcgtcatct tgttcattgg a                               401

```

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

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

tggcaagggt aaacagatcc

20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

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

ctcaggcact cctctcaacc

20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

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

tacctcctca ggcactcctc

20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: DNA

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

aagcagccaa tggatgccaa gc

22

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

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

tccaatgaac aagatgacga c

21

We claim:

1. A method to determine whether a subject is homozygous for the hereditary hemochromatosis (HFE) gene mutation 5474 G>A, the method comprising:
  - (a) obtaining a genomic DNA sample from the subject;
  - (b) amplifying the DNA sample by polymerase chain reaction (PCR), using suitable oligonucleotide primers to amplify a segment of the gene including nucleotide position 5474 to give an amplified product, wherein said primers do not span nucleotide position 5569 of the HFE gene; and
  - (c) subjecting the amplified product from step (b) to a suitable assay to detect a component thereof indicative of mutant allele amplification and a component thereof indicative of normal allele amplification,wherein the detection of a component indicative of mutant allele amplification but no detection of a component indicative of normal allele amplification indicates that the subject is homozygous for the mutation 5474 G>A.
2. The method of claim 1 wherein step (c) comprises subjecting the amplified product from step (b) to digestion with a restriction enzyme which recognizes the mutant 5474A allele sequence and analysing the digestion product for at least one nucleotide fragment indicative of mutant allele amplification and at least one fragment indicative of normal allele amplification, wherein the detection of the at least one nucleotide fragment indicative of mutant allele amplification but no detection of the at least one fragment indicative of normal allele amplification indicates that the subject is homozygous for the mutation 5474 G>A.
3. The method of claim 2 wherein the restriction enzyme is Rsa 1, and wherein the digestion product is analysed for nucleotide fragments of 146 bp and 117 bp,

wherein the presence of a single nucleotide fragment of 117 bp indicates that the subject is homozygous for the mutation 5474 G>A.

4. The method of any one of claims 1 to 3 wherein the genomic DNA is  
5 obtained from a blood sample or buccal smear obtained from the subject.
5. The method of claim 2 wherein the restriction enzyme is Rsa 1 or Sna  
B 1.
- 10 6. The method of claim 1 wherein the assay of step (a) is an  
oligonucleotide ligation assay.
7. A method for diagnosing hereditary hemochromatosis (HH) in a  
subject, the method comprising determining whether the genome of the  
15 subject has 0, 1 or 2 alleles of a mutant form of the HFE gene.
8. The method of claim 7 wherein the mutant form of the HFE gene is a  
5474 G>A mutation.
- 20 9. The method of claim 8 wherein the method comprises amplification of a  
nucleotide sequence wherein the primers hybridize to sites other than  
nucleotide position 5569.
10. The method of claim 9 wherein the antisense primer has the nucleotide  
25 sequence TACCTCCTCAGGCACTCCTC.
11. The method of claim 9 further comprising reacting the amplified  
sequence with a restriction enzyme to produce reaction products.
- 30 12. The method of claim 11 wherein alternate versions of the gene are  
distinguished one from another by the presence of an Rsa 1 cleavage site.

13. The method of claim 11 wherein said reaction products are separated by gel electrophoresis and visualized.

14. The method of claim 12 further comprising determining the presence of  
5 a 146 bp band and a 117 bp band wherein the presence of only a 117 band is indicative of two alleles having a 5474 G>A mutation, the presence of a 146 bp band and a 117 bp band is indicative of one allele having the mutation and the presence of only a 146 bp band is indicative of neither allele having the mutation.

10

15. An isolated polynucleotide comprising at least 12 to 12146 nucleotides of the human HFE gene (Sequence ID NO:1), wherein the polynucleotide includes nucleotide position 5569 and wherein nucleotide 5569 is A or a complement thereof.

15

16. A DNA primer pair that flanks nucleotide position 5569 of the HFE gene.

20

17. The nucleotide sequence TACCTCCTCAGGCACTCCTC.

18. A set of primers for use in a DNA amplification assay to determine whether a subject is homozygous for the HFE gene mutation 5474 G>A, comprising the nucleotide sequences 5'TGGCAAGGGTAAACAGATCC and 5'TACCTCCTCAGGCACTCCTC.

25

19. A composition for use in diagnosing HH, said composition comprising one or more primers adapted to amplify a segment of a HFE gene wherein at least one of said primers is complementary to a sequence within 100 bp downstream of nucleotide position 5569 of the HFE gene.

30

20. A kit for diagnosis of HH comprising nucleotide primers adapted to distinguish between 5474A homozygotes and 5474A/5569A heterozygotes.

21. A kit according to claim 20 wherein at least one of said primers is 5'TACCTCCTCAGGCACTCCTC.

**ABSTRACT**

A new polymorphism of the HFE gene is described, the presence of which in a subject can lead to misdiagnosis of the subject as a homozygote  
5 for the mutation causing hereditary hemochromatosis. A new more discriminating diagnostic method is provided.

Figure 1

```

ggatccttta accgaggaga ttattatagc cggagctctg aagcagcaat ctcagttctt 61
gtgatagtga gcaaagaact acaaactaac accaaaatgc aagcttaaag caaagtttat 121
tgaagcacia taatacactc tgagggacag cgggcttatt tctgcgaagt gaactcagca 181
cttctttaca gagctcaagg tgcttttatg gggtttgtgg ggaggagttg aggtttgggc 241
tgtatctgag tgacaggatg atgttatattg attgaagttt atagctatac aatctaaaat 301
taaactgtgc atggtcttac ctataatattg ttaagaaaag cctcccaggg atgggggggc 361
aaaactgtat gtaaattcta ttataatgat ggcattgatga acttgggggtg aacttgaaga 421
caggcttttg tgttggtggg catgtgccac cttagggaat ttccacctgt accctccttt 481
ctctttctcc aggatatttt ggccacagac tttatcataa actccatccc ttaggggtggc 541
attagggtag tcttgggcct gaatttaggt gggccagtggt ctgtcttagt gacagccttt 601
ccgctctctt ctgtcatccc ctcccaactg ctaatgtcta actacctaac aattacccat 661
taaactcagt tgtctggggg taggagcagg cctcaatatg tttaatcatt ctccagataa 721
tcccaatact gtaaagtttg tgaaacactt gtcagataat tcaattatga aggtgtgga 781
acgtgtttca gtaggatcta attggttaat gttatgactt aattaatttg aatcaaaaaa 841
caaatgaaa aagctttata tttctaagtc aaataagaca taagttgggtc taagggtgag 901
ataaaatttt taaatgtatg attgaatttt gaaaatcata aatatttaaa tatctaaagt 961
tcagatcaga acattgcgaa gctactttcc ccaatcaaca acaccccttc aggatttaaa 1021
aaccaagggg gacactggat cacctagtgt ttcacaagca ggtaccttct gctgtaggag 1081
agagagaact aaagtctga aagacctgtt gcttttcacc aggaagtttt actgggcctc 1141
tcctgagcct aggcaatagc tgtaggggtg cttctggagc catccccgtt tccccgcccc 1201
ccaaaagaag cggagattta acggggacgt gcggccagag ctggggaaat ggggccgcga 1261
gccaggccgg cgcttctcct cctgatgctt ttgcagaccg cggctctgca ggggcgcttg 1321
ctgcgtgagt ccgagggtg cgggcgaact aggggcgcgg cgggggtgga aaaatcgaaa 1381
ctagcttttt ctttgcgctt gggagtttgc taactttgga ggacctgctc aaccaatcc 1441
gcaagccctt ctccctactt tctgcgtcca gaccccgctg gggagtgcct accactgaac 1501
tgcagatagg ggtccctcgc cccaggacct gccccctccc ccggtgtgct cggctctgct 1561
gagtgaactt tggaaaccgc cactcccttc cccaactag aatgctttta aataaatctc 1621
gtagttcctc acttgagctg agctaagcct ggggctcctt gaacctggaa ctgggtttta 1681
tttccaatgt cagctgtgca gttttttccc cagtcacttc caaacaggaa gttcttccct 1741
gagtgccttg cgagaaggct gagcaaacc acagcaggat ccgcacgggg tttccacctc 1801
agaacgaatg cggtgggcgg tgggggcgcg aaagagtggc gttggggatc tgaattcttc 1861
accattccac ccacttttgg tgagacctgg ggtggaggtc tctaggggtg gaggtcctg 1921
agagaggcct acctcgggcc tttccccact cttggcaatt gttcttttgc ctggaaaatt 1981
aagtatatgt tagttttgaa cgtttgaact gaacaattct cttttcggct aggttttatt 2041
gatttgcaat gtgctgtgta attaagaggc ctctctacaa agtactgata atgaacatgt 2101
aagcaatgca ctcacttcta agttacattc atatctgac ttatttgatt ttcactaggc 2161
atagggaggt aggagcta atacgttta ttttactaga agttaactgg aattcagatt 2221
atataactct tttcaggtta caaagaacat aaataatctg gttttctgat gttatttcaa 2281
gtactacagc tgcttcta at cttagttgac agtgattttg ccctgtagtg tagcacagtg 2341
ttctgtgggt cacacgccgg cctcagcaca gcactttgag ttttggtact acgtgtatcc 2401
acattttaca catgacaaga atgaggcatg gcacggcctg cttcctggca aatttattca 2461
atggtacacg gggcttttgt ggcagagctc atgtctccac ttcatagcta tgattcttaa 2521
acatcacact gcattagagg ttgaataata aaatttcatg ttgagcagaa atattcattg 2581
tttacaagtg taaatgagtc ccagccatgt gttgcactgt tcaagcccca agggagagag 2641
cagggaaca agtctttacc ctttgatatt ttgcattcta gtgggagaga tgacaataag 2701
caaatgagca gaaagatata caacatcagg aaatcatggg tgttggtgaga agcagagaag 2761
tcagggaag tcaactctgg gctgacactt gagcagagac atgaaggaaa taagaatgat 2821
attgactggg agcagtattt cccaggcaaa ctgagtgggc ctggcaagtt ggattaaaaa 2881
gcgggttttc tcagcactac tcatgtgtgt gtgtgtgggg gggggggcgg cgtgggggtg 2941
ggaaggggga ctaccatctg catgtaggat gtctagcagt atcctgtcct ccctactcac 3001
taggtgctag gagcactccc ccagtcttga caacaaaaa tgtctctaaa ctttgccaca 3061

```

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tgtcacctag  | tagacaaact  | cctgggttaag | aagctcgggt  | tgaaaaaaat  | aaacaagtag  | 3121 |
| tgctggggag  | tagaggccaa  | gaagtaggta  | atgggctcag  | aagaggagcc  | acaaacaagg  | 3181 |
| ttgtgcaggc  | gcctgtaggc  | tgtgggtgtga | attctagcca  | aggagtaaca  | gtgatctgtc  | 3241 |
| acaggctttt  | aaaagattgc  | tctggctgct  | atgtggaaag  | cagaatgaag  | ggagcaacag  | 3301 |
| taaaagcagg  | gagcccagcc  | aggaagctgt  | tacacagtcc  | aggcaagagg  | tagtggagtg  | 3361 |
| ggctgggtgg  | gaacagaaaa  | gggagtgaca  | aaccattgtc  | tcctgaatat  | attctgaagg  | 3421 |
| aagttgctga  | aggattctat  | gttgtgtgag  | agaaagagaa  | gaattggctg  | ggtgtagtag  | 3481 |
| ctcatgccaa  | ggaggaggcc  | aaggagagca  | gattcctgag  | ctcaggagtt  | caagaccagc  | 3541 |
| ctgggcaaca  | cagcaaaacc  | ccttctctac  | aaaaaataca  | aaaattagct  | gggtgtgggtg | 3601 |
| gcatgcacct  | gtgatcctag  | ctactcggga  | ggctgaggtg  | gagggtattg  | cttgagccca  | 3661 |
| ggaagttgag  | gctgcagtga  | gccatgactg  | tgccactgta  | cttcagccta  | ggtgacagag  | 3721 |
| caagaccctg  | tctcccctga  | ccccctgaaa  | aagagaagag  | ttaaagttga  | ctttgttctt  | 3781 |
| tattttaatt  | ttattggcct  | gagcagtggg  | gtaattggca  | atgccatttc  | tgagatggtg  | 3841 |
| aaggcagagg  | aaagagcagt  | ttggggtaaa  | tcaaggatct  | gcatttggac  | atgttaagtt  | 3901 |
| tgagattcca  | gtcaggcttc  | caagtgggtga | ggccacatag  | gcagttcagt  | gtaagaattc  | 3961 |
| aggaccaagg  | cagggcacgg  | tggctcactt  | ctgtaatccc  | agcactttgg  | tggctgaggc  | 4021 |
| aggtagatca  | tttgagggtca | ggagtttgag  | acaagcttgg  | ccaacatggt  | gaaaccccat  | 4081 |
| gtctactaaa  | aatacaaaaa  | ttagcctggt  | gtgggtggcgc | acgcctatag  | tcccaggttt  | 4141 |
| tcaggaggct  | taggtaggag  | aatcccttga  | acccaggagg  | tgcaggttgc  | agtgagctga  | 4201 |
| gattgtgcca  | ctgcactcca  | gcctgggtga  | tagagtgaga  | ctctgtctca  | aaaaaaaaaa  | 4261 |
| aaaaaaaaaa  | aaaaaaaaaa  | aactgaagga  | attattcctc  | aggatttggg  | tctaatttgc  | 4321 |
| cctgagcacc  | aactcctgag  | ttcaactacc  | atggctagac  | acaccttaac  | attttctaga  | 4381 |
| atccaccagc  | tttagtgagg  | tctgtctaat  | catgagtatt  | ggaataggat  | ctggggggcag | 4441 |
| tgaggggggtg | gcagccacgt  | gtggcagaga  | aaagcacaca  | aggaaagagc  | acccaggact  | 4501 |
| gtcatatgga  | agaaagacag  | gactgcaact  | cacccttcac  | aaaatgagga  | ccagacacag  | 4561 |
| ctgatggtat  | gagttgatgc  | aggtgtgtgg  | agcctcaaca  | tcctgctccc  | ctcctactac  | 4621 |
| acatggttaa  | ggcctgttgc  | tctgtctcca  | ggttcacact  | ctctgcacta  | cctcttcatg  | 4681 |
| ggtgcctcag  | agcaggacct  | tggctcttcc  | ttgtttgaag  | ctttgggcta  | cgtggatgac  | 4741 |
| cagctgttcg  | tgttctatga  | tcatgagagt  | cgccgtgtgg  | agccccgaac  | tccatgggtt  | 4801 |
| tccagtagaa  | tttcaagcca  | gatgtggctg  | cagctgagtc  | agagtctgaa  | aggggtgggat | 4861 |
| cacatgttca  | ctgttgactt  | ctggactatt  | atggaaaatc  | acaaccacag  | caagggtatg  | 4921 |
| tggagagggg  | gcctcacctt  | cctgaggttg  | tcagagcttt  | tcattctttc  | atgcatcttg  | 4981 |
| aaggaaacag  | ctggaagtct  | gaggtccttg  | gggagcaggg  | aagaggggaag | gaatttgctt  | 5041 |
| cctgagatca  | tttggtcctt  | ggggatggtg  | gaaataggga  | cctattcctt  | tggttgcagt  | 5101 |
| taacaaggct  | ggggattttt  | ccagagtccc  | acaccctgca  | ggtcatcctg  | ggctgtgaaa  | 5161 |
| tgcaagaaga  | caacagtacc  | gagggctact  | ggaagtacgg  | gtatgatggg  | caggaccacc  | 5221 |
| ttgaattctg  | ccctgacaca  | ctggattgga  | gagcagcaga  | acccagggcc  | tggcccacca  | 5281 |
| agctggagtg  | ggaaaggcac  | aagattcggg  | ccaggcagaa  | cagggcctac  | ctggagaggg  | 5341 |
| actgccctgc  | acagctgcag  | cagttgctgg  | agctggggag  | aggtgttttg  | gaccaacaag  | 5401 |
| gtatgggtgga | aacacacttc  | tgccccata   | ctctagtggc  | agagtggagg  | aggttgcagg  | 5461 |
| gcacggaatc  | cctggttgga  | gtttcagagg  | tggctgaggc  | tgtgtgcctc  | tccaaattct  | 5521 |
| gggaagggac  | tttctcaatc  | ctagagtctc  | taccttataa  | ttgagatgta  | tgagacagcc  | 5581 |
| acaagtcatg  | ggtttaattt  | cttttctcca  | tgcatatggc  | tcaaagggaa  | gtgtctatgg  | 5641 |
| cccttgcttt  | ttatttaacc  | aataatcttt  | tgtatattta  | tacctgttaa  | aaattcagaa  | 5701 |
| atgtcaaggc  | cgggcacggt  | ggctcacccc  | tgtaatccca  | gcactttggg  | aggccgaggc  | 5761 |
| gggtgggtcac | aaggtcagga  | gtttgagacc  | agcctgacca  | acatggtgaa  | acccgtctct  | 5821 |
| aaaaaaaaac  | aaaaattagc  | tggtcacagt  | catgcgcacc  | tgtagtccca  | gctaattgga  | 5881 |
| aggctgaggc  | aggagcatcg  | cttgaacctg  | ggaagcggaa  | gttgactga   | gccaagatcg  | 5941 |
| cgccactgca  | ctccagccta  | ggcagcagag  | tgagactcca  | tcttaaaaaa  | aaaaaaaaaa  | 6001 |
| aaaaagagaa  | ttcagagatc  | tcagctatca  | tatgaatacc  | aggacaaaat  | atcaagttag  | 6061 |
| gccacttatc  | agagtagaag  | aatcctttag  | gttaaaagtt  | tctttcatag  | aacatagcaa  | 6121 |
| taatcactga  | agctacctat  | cttacaagtc  | cgcttcttat  | aacaatgcct  | cctaggttga  | 6181 |
| cccaggtgaa  | actgaccatc  | tgtattcaat  | cattttcaat  | gcacataaag  | ggcaatttta  | 6241 |
| tctatcagaa  | caaagaacat  | gggtaacaga  | tatgtatatt  | tacatgtgag  | gagaacaagc  | 6301 |

|            |            |             |             |             |             |      |
|------------|------------|-------------|-------------|-------------|-------------|------|
| tgatctgact | gctctccaag | tgacactgtg  | ttagagtcca  | atcttaggac  | acaaaatggg  | 6361 |
| gtctctcctg | tagcttggtt | ttttctgaaa  | aggggtattt  | cttcctccaa  | cctatagaag  | 6421 |
| gaagtgaaag | ttccagtctt | cctggcaagg  | gtaaacagat  | cccctctcct  | catccttcct  | 6481 |
| ctttcctgtc | aagtgcctcc | tttgggtgaag | gtgacacatc  | atgtgacctc  | ttcagtgacc  | 6541 |
| actctacggg | gtcgggcctt | gaactactac  | ccccagaaca  | tcaccatgaa  | gtgggtgaag  | 6601 |
| gataagcagc | caatggatgc | caaggagttc  | gaacctaaag  | acgtattgcc  | caatggggat  | 6661 |
| gggacctacc | agggctggat | aaccttggct  | gtacccccct  | gggaagagca  | gagatatacg  | 6721 |
| tgccaggtgg | agcacccagg | cctggatcag  | cccctcattg  | tgatctgggg  | tatgtgactg  | 6781 |
| atgagagcca | ggagctgaga | aaatctattg  | gggggtgaga  | ggagtgcctg  | aggaggtaat  | 6841 |
| tatggcagtg | agatgaggat | ctgctctttg  | ttaggggatg  | ggctgagggt  | ggcaatcaaa  | 6901 |
| ggctttaact | tgctttttct | gttttagagc  | cctcacctgc  | tggcaccccta | gtcattggag  | 6961 |
| tcacagtggt | aattgctgtt | tttgtcgtca  | tcttggttcat | tgggaattttg | ttcataatat  | 7021 |
| taaggaagag | gcagggttca | agtgagtagg  | aacaaggggg  | aagtctctta  | gtacctctgc  | 7081 |
| cccagggcac | agtgggaaga | ggggcagagg  | ggatctggca  | tccatgggaa  | gcatttttct  | 7141 |
| catttatatt | ctttggggac | accagcagct  | ccctgggaga  | cagaaaataa  | tggttctccc  | 7201 |
| cagaatgaaa | gtctctaatt | caacaaacat  | cttcagagca  | cctactattt  | tgcaagagct  | 7261 |
| gtttaaggta | gtacaggggg | tttgagggtg  | agaagtcact  | gtggctattc  | tcagaaccca  | 7321 |
| aatctggtag | ggaatgaaat | tgatagcaag  | taaatgtagt  | taaagaagac  | cccatgaggt  | 7381 |
| cctaaagcag | gcaggaagca | aatgcttagg  | gtgtcaaagg  | aaagaatgat  | cacattcagc  | 7441 |
| tggggatcaa | gatagccttc | tggatcttga  | aggagaagct  | ggattccatt  | aggtgagggt  | 7501 |
| gaagatgatg | ggaggtctac | acagacggag  | caaccatgcc  | aagtaggaga  | gtataaggca  | 7561 |
| tactgggaga | ttagaaataa | ttactgtacc  | ttaaccctga  | gtttgcttag  | ctatcactca  | 7621 |
| ccaattatgc | atttctaccc | cctgaacatc  | tgtggtgtag  | ggaaaagaga  | atcagaaaga  | 7681 |
| agccagctca | tacagagtc  | aagggtcttt  | tgggatattg  | ggttatgatc  | actgggggtg  | 7741 |
| cattgaagga | tcctaagaaa | ggaggaccac  | gatctccctt  | atatggtgaa  | tgtgttggtta | 7801 |
| agaagttaga | tgagaggtga | ggagaccagt  | tagaaagcca  | ataagcattt  | ccagatgaga  | 7861 |
| gataatgggt | cttgaaatcc | aatagtgtcc  | aggtctaaat  | tgagatgggt  | gaatgaggaa  | 7921 |
| aataaggaag | agagaagagg | caagatgggt  | cctaggtttg  | tgatgcctct  | ttcctgggtc  | 7981 |
| tcttgtctcc | acaggaggag | ccatggggca  | ctacgtctta  | gctgaacgtg  | agtgacacgc  | 8041 |
| agcctgcaga | ctcactgtgg | gaaggagaca  | aaactagaga  | ctcaaagagg  | gagtgcattt  | 8101 |
| atgagctctt | catgtttcag | gagagagttg  | aacctaaaca  | tagaaattgc  | ctgacgaact  | 8161 |
| ccttgatttt | agccttctct | gttcatttcc  | tcaaaaagat  | ttccccattt  | aggtttctga  | 8221 |
| gttctctcat | gccggtgatc | cctagctgtg  | acctctcccc  | tggaaactgtc | tctcatgaac  | 8281 |
| ctcaagctgc | atctagaggg | ttccttcatt  | tctctcgtca  | cctcagagac  | atacacctat  | 8341 |
| gtcatttcat | ttcctatttt | tggaaagagga | ctccttaa    | ttgggggact  | tacatgattc  | 8401 |
| attttaacat | ctgagaaaag | ctttgaaccc  | tgggacgtgg  | ctagtataaa  | ccttaccaga  | 8461 |
| tttttacaca | tgtatctatg | cattttctgg  | accggttcaa  | cttttctctt  | gaatcctctc  | 8521 |
| tctgtgttac | ccagtaactc | atctgtcacc  | aagccttggg  | gattcttcca  | tctgattgtg  | 8581 |
| atgtgagttg | cacagctatg | aaggctgtac  | actgcacgaa  | tggaaagagg  | acctgtccca  | 8641 |
| gaaaaagcat | catggctatc | tgtgggtagt  | atgatgggtg  | tttttagcag  | gtaggaggca  | 8701 |
| aatatcttga | aaggggttgt | gaagaggtgt  | tttttcta    | tggcatgaag  | gtgtcataca  | 8761 |
| gatttgcaaa | gtttaatggg | gccttcattt  | gggatgctac  | tctagtattc  | cagacctgaa  | 8821 |
| gaatcacaat | aattttctac | ctggctctct  | cttgttctga  | taatgaaaat  | tatgataagg  | 8881 |
| atgataaaa  | cacttacttc | gtgtccgact  | cttctgagca  | cctacttaca  | tgcattactg  | 8941 |
| catgcacttc | ttacaataat | tctatgagat  | aggtactatt  | atccccattt  | cttttttaaa  | 9001 |
| tgaagaaagt | gaagtagggc | gggcacgggt  | gctcacgcct  | gtaatcccag  | cactttggga  | 9061 |
| ggccaaagcg | ggtggatcac | gaggtcagga  | gatcgagacc  | atcctggcta  | acatgggtgaa | 9121 |
| accccatctc | taataaaaa  | acaaaaaatt  | agctgggcgt  | ggtggcagac  | gcctgtagtc  | 9181 |
| ccagctactc | ggaaggctga | ggcaggagaa  | tggcatgaac  | ccaggaggca  | gagcttgacg  | 9241 |
| tgagccgagt | ttgcgccact | gcactccagc  | ctaggtgaca  | gagtgagact  | ccatctcaaa  | 9301 |
| aaaataaaaa | taaaaataaa | aaaatgaaaa  | aaaaaagaaa  | gtgaagtata  | gagtatctca  | 9361 |
| tagtttgtca | gtgatagaaa | caggtttcaa  | actcagtc    | tctgaccgtt  | tgatacatct  | 9421 |
| cagacaccac | tacattcagt | agtttagatg  | cctagaataa  | atagagaagg  | aaggagatgg  | 9481 |
| ctcttctctt | gtctcattgt | gtttcttctg  | aatgagcttg  | aatcacatga  | aggggaacag  | 9541 |

```

cagaaaacaa ccaactgata ctcagctgtc atgttttctt taaaagtccc tgaaggaagg 9601
tcctggaatg tgactccctt gctcctctgt tgctctcttt ggcattcatt tctttggacc 9661
ctacgcaagg actgtaattg gtggggacag ctagtggccc tgctgggctt cacacacggt 9721
gtcctcccta ggccagtgcc tctggagtca gaactctggt ggtatttccc tcaatgaagt 9781
ggagtaagct ctctcatttt gagatgggat aatggaagcc accaagtggc ttagaggatg 9841
cccaggctct tccatggagc cactgggggt ccggtgcaca ttaaaaaaaaa aatctaacca 9901
ggacattcag gaattgctag attctgggaa atcagttcac catgttcaaa agagtctttt 9961
tttttttttt gagactctat tgcccagggt ggagtgcagt ggcattgatc cggctcactg 10021
taacctctgc ctcccagggt caagcgattc tcctgtctca gcctcccaag tagctgggat 10081
tacaggcgtg caccaccatg cccggctaata ttttgtattt ttagtagaga cagggtttca 10141
ccatgttggt caggctgggt tcgaactctc ctgacctcgt gatccgcctg cctcggcctc 10201
ccaaagtgtg gagattacag gtgtgagcca cctgcccag cgtcaaaaag agtcttaata 10261
tatatatcca gatggcatgt gtttacttta tggtactaca tgcacttggc tgcataaatg 10321
tggtacaagc attctgtctt gaagggcagg tgcttcagga taccatatac agctcagaag 10381
tttcttcttt aggcattaaa ttttagcaaa gatattctcat ctcttctttt aaaccatttt 10441
ctttttttgt ggttagaaaa gttatgtaga aaaaagtaaa tgtgatttac gctcattgta 10501
gaaaagctat aaaatgaata caattaaagc tgttatttaa ttagccagtg aaaaactatt 10561
aacaacttgt ctattacctg ttagtattat tggtgcatta aaaatgcata tactttaata 10621
aatgtacatt gtattgtata ctgcatgatt ttattgaagt tcttggtcat cttgtgtata 10681
tacttaatcg ctttgtcatt ttggagacat ttattttgct tctaatttct ttacattttg 10741
tcttacggaa tatttttcatt caactgtggt agccgaatta atcgtgtttc ttcactctag 10801
ggacattgtc gtctaagttg taagacattg gttattttac cagcaaacca ttctgaaagc 10861
atatgacaaa ttatttctct cttaatatct tactatactg aaagcagact gctataaggc 10921
ttcacttact cttctacctc ataaggaata tggtacaatt aatttattag gtaagcattt 10981
gttttatatt ggttttattt cacctgggct gagatttcaa gaaacacccc agtcttcaca 11041
gtaacacatt tcactaacac atttactaaa catcagcaac tgtggcctgt taattttttt 11101
aatagaaatt ttaagtcctc attttctttc ggtgtttttt aagcttaatt tttctggctt 11161
tattcataaa ttcttaaggt caactacatt tgaaaaatca aagacctgca ttttaaattc 11221
ttattcacct ctggcaaaac cattcacaaa ccatggtagt aaagagaagg gtgacacctg 11281
gtggccatag gtaaatgtac cacggtgggt cggtgaccag agatgcagcg ctgaggggtt 11341
tcctgaagggt aaaggaataa agaattgggt gaggggcgtg cactggaaat cacttgtaga 11401
gaaaagcccc tgaaaatttg agaaaacaaa caagaaacta cttaccagct atttgaattg 11461
ctggaatcac aggccattgc tgagctgcct gaactgggaa cacaacagaa ggaaaacaaa 11521
ccactctgat aatcattgag tcaagtacag cagggtgatt aggactgctg agaggtacag 11581
gccaaaattc ttatgttgta ttataataat gtcattctat aatactgtca gtattttata 11641
aaacattctt cacaaactca cacacattta aaaacaaaac actgtctcta aaatcccaa 11701
atttttcata aactcagttt taaactaact ttttttcaaa ccacaatctg atttaacaat 11761
gactatcatt taaatatctc tgactttcaa attaaagatt ttcacatgca ggctgatatt 11821
tgtaattgtg attctctctg taggctttgg gtataatgtg ttcttttctt tttttgcac 11881
agcgattaac ttctacactc taacatgtag aatgttacta caatattaaa gtattttgta 11941
tgacaatttt atttgaaagc ctaggatgag ttgacatcct gcatgcattt attacttgat 12001
atgcatgcat tctgggatct caagcattct atttctgagt aattgtttta ggtgtagaag 12061
agatagatat ggtggatttg gagttgatac ttatatattt tctatttctt ggatggatga 12121
atttgtacat taaaagtttt ccatgg//

```

We claim:

1. A method to determine whether a subject is homozygous for the hereditary hemochromatosis (HFE) gene mutation 5474 G>A, the method  
5 comprising:

- (a) obtaining a genomic DNA sample from the subject;
- (b) amplifying the DNA sample by polymerase chain reaction (PCR), using suitable oligonucleotide primers to amplify a segment of the gene including nucleotide position 5474 to give an amplified  
10 product, wherein said primers do not span nucleotide position 5569 of the HFE gene; and
- (c) subjecting the amplified product from step (b) to a suitable assay to detect a component thereof indicative of mutant allele amplification and a component thereof indicative of normal allele  
15 amplification,

wherein the detection of a component indicative of mutant allele amplification but no detection of a component indicative of normal allele amplification indicates that the subject is homozygous for the mutation 5474 G>A.

20 2. The method of claim 1 wherein step (c) comprises subjecting the amplified product from step (b) to digestion with a restriction enzyme which recognizes the mutant 5474A allele sequence and analysing the digestion product for at least one nucleotide fragment indicative of mutant allele amplification and at least one fragment indicative of normal allele  
25 amplification,

wherein the detection of the at least one nucleotide fragment indicative of mutant allele amplification but no detection of the at least one fragment indicative of normal allele amplification indicates that the subject is homozygous for the mutation 5474 G>A.

30

3. The method of claim 2 wherein the restriction enzyme is Rsa 1, and wherein the digestion product is analysed for nucleotide fragments of 146 bp

and 117 bp,

wherein the presence of a single nucleotide fragment of 117 bp indicates that the subject is homozygous for the mutation 5474 G>A.

5     4.     The method of any one of claims 1 to 3 wherein the genomic DNA is obtained from a blood sample or buccal smear obtained from the subject.

5.     The method of claim 2 wherein the restriction enzyme is Rsa 1 or Sna B 1.

10

6.     The method of claim 1 wherein the assay of step (a) is an oligonucleotide ligation assay.

7.     A method for diagnosing hereditary hemochromatosis (HH) in a  
15     subject, the method comprising determining whether the genome of the subject has 0, 1 or 2 alleles of a mutant form of the HFE gene.

8.     The method of claim 7 wherein the mutant form of the HFE gene is a 5474 G>A mutation.

20

9.     The method of claim 8 wherein the method comprises amplification of a nucleotide sequence wherein the primers hybridize to sites other than nucleotide position 5569.

25     10.     The method of claim 9 wherein the antisense primer has the nucleotide sequence TACCTCCTCAGGCACTCCTC.

11.     The method of claim 9 further comprising reacting the amplified sequence with a restriction enzyme to produce reaction products.

30

12.     The method of claim 11 wherein alternate versions of the gene are distinguished one from another by the presence of an Rsa 1 cleavage site.

13. The method of claim 11 wherein said reaction products are separated by gel electrophoresis and visualized.

5 14. The method of claim 12 further comprising determining the presence of a 146 bp band and a 117 bp band wherein the presence of only a 117 band is indicative of two alleles having a 5474 G>A mutation, the presence of a 146 bp band and a 117 bp band is indicative of one allele having the mutation and the presence of only a 146 bp band is indicative of neither allele having  
10 the mutation.

15. An isolated polynucleotide comprising at least 12 to 12146 nucleotides of the human HFE gene (Sequence ID NO:1), wherein the polynucleotide includes nucleotide position 5569 and wherein nucleotide 5569 is A or a  
15 complement thereof.

16. A DNA primer pair that flanks nucleotide position 5569 of the HFE gene.

20 17. The nucleotide sequence TACCTCCTCAGGCACTCCTC.

18. A set of primers for use in a DNA amplification assay to determine whether a subject is homozygous for the HFE gene mutation 5474 G>A, comprising the nucleotide sequences 5'TGGCAAGGGTAAACAGATCC and  
25 5'TACCTCCTCAGGCACTCCTC.

19. A composition for use in diagnosing HH, said composition comprising one or more primers adapted to amplify a segment of a HFE gene wherein at least one of said primers is complementary to a sequence within 100 bp  
30 downstream of nucleotide position 5569 of the HFE gene.

20. A kit for diagnosis of HH comprising nucleotide primers adapted to

distinguish between 5474A homozygotes and 5474A/5569A heterozygotes.

21. A kit according to claim 20 wherein at least one of said primers is 5'TACCTCCTCAGGCACTCCTC.

HFE genomic sequence (GenBank Z92910)

|      |             |            |            |             |             |            |     |         |     |         |     |
|------|-------------|------------|------------|-------------|-------------|------------|-----|---------|-----|---------|-----|
| 5356 | aagcagc     | caatgga    | gc         | caaggag     | ttc         | gaacctaa   | g   | acgtatt | gcc | caatggg | gat |
| 5413 | gggacctacc  | agggctggat | aaccttggct | gtacccccctg | gggaagagca  | gagatat    | acg |         |     |         |     |
| 5473 | tgccagggtgg | agcaccagg  | cctggatcag | ccctcattg   | tgatctgggg  | tatgtgactg |     |         |     |         |     |
| 5533 | atgagagcca  | ggagctgaga | aaatctattg | ggggttgaga  | ggagtgcctg  | aggaggtaat |     |         |     |         |     |
| 5593 | tatggcagtg  | agatgaggat | ctgctctttg | ttaggggatg  | ggctgagggt  | ggcaatcaaa |     |         |     |         |     |
| 5653 | ggctttaact  | tgcttttct  | gttttagagc | cctcacccgtc | tggcaccccta | gtcattggag |     |         |     |         |     |
| 5713 | tcatcagtg   | aattgctgtt | tttgcgtca  | tcttgttcat  | tgga        |            |     |         |     |         |     |

Figure 2A

New HFE genomic sequence

|      |             |            |            |             |             |            |     |         |     |         |     |
|------|-------------|------------|------------|-------------|-------------|------------|-----|---------|-----|---------|-----|
| 5356 | aagcagc     | caatgga    | gc         | caaggag     | ttc         | gaacctaa   | g   | acgtatt | gcc | caatggg | gat |
| 5413 | gggacctacc  | agggctggat | aaccttggct | gtacccccctg | gggaagagca  | gagatat    | acg |         |     |         |     |
| 5473 | tgccagggtgg | agcaccagg  | cctggatcag | ccctcattg   | tgatctgggg  | tatgtgactg |     |         |     |         |     |
| 5533 | atgagagcca  | ggagctgaga | aaatctattg | ggggttaaga  | ggagtgcctg  | aggaggtaat |     |         |     |         |     |
| 5593 | tatggcagtg  | agatgaggat | ctgctctttg | ttaggggatg  | ggctgagggt  | ggcaatcaaa |     |         |     |         |     |
| 5653 | ggctttaact  | tgcttttct  | gttttagagc | cctcacccgtc | tggcaccccta | gtcattggag |     |         |     |         |     |
| 5713 | tcatcagtg   | aattgctgtt | tttgcgtca  | tcttgttcat  | tgga        |            |     |         |     |         |     |

Figure 2B

PRIMER SEQUENCE

Feder sense primer

TGGCAAGGGTAAACAGATCC

Feder antisense primer

CCAACTCTCCTCACGGACTC

New antisense primer

CTCCTCACGGACTCCTCCAT

GENOMIC SEQUENCE (GenBank Z92910)

5192

TCCTGGCAAGGGTAAACAGATCCCCT

5217

5562

5593

GGGGGTT<sup>9</sup><sub>a</sub>AGAGGAGTGCCTGAGGAGGTAATT

5569 polymorphism

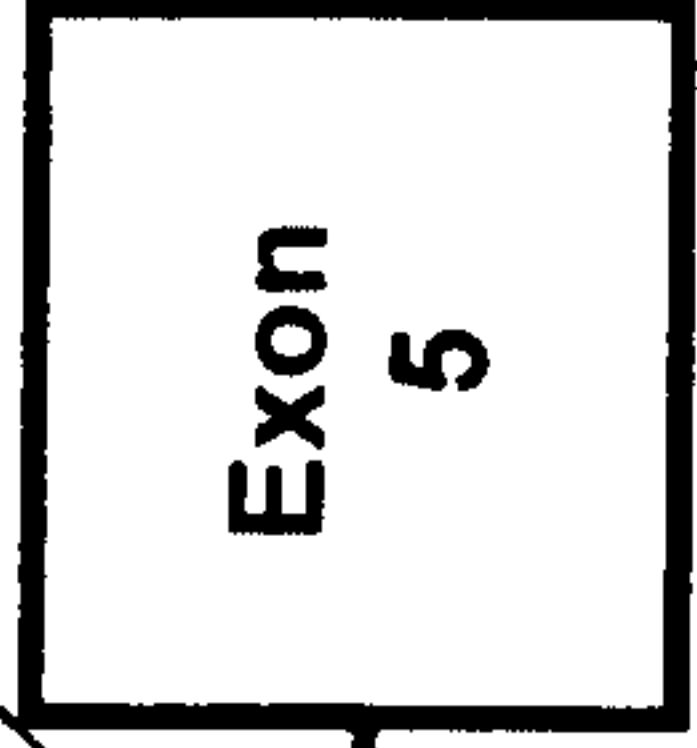
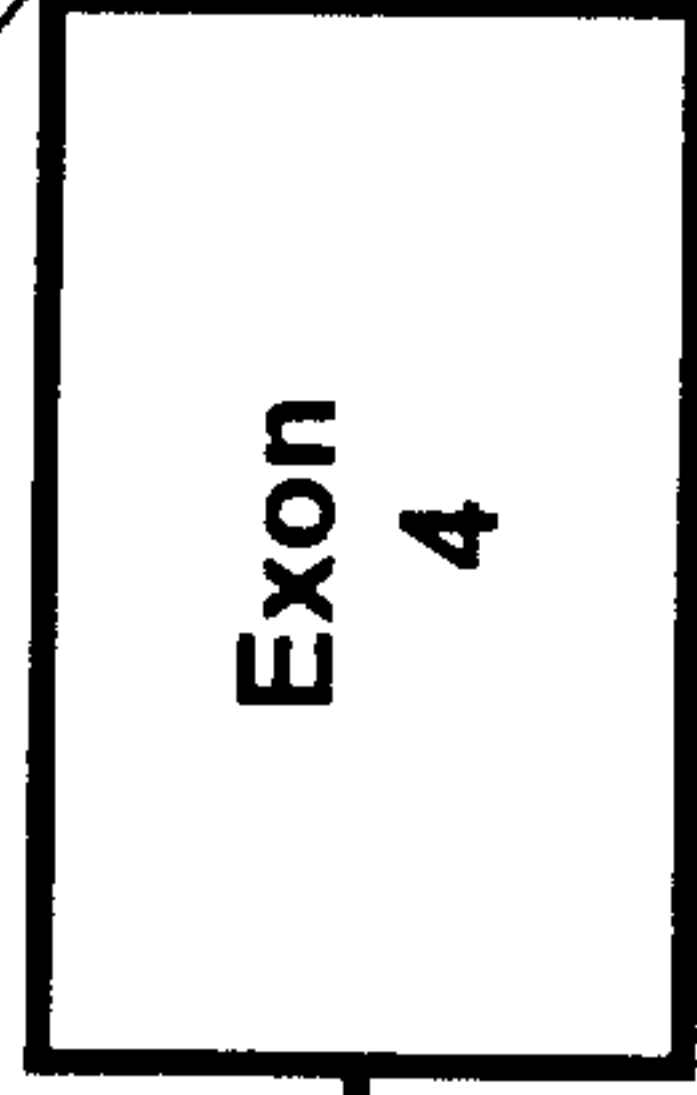
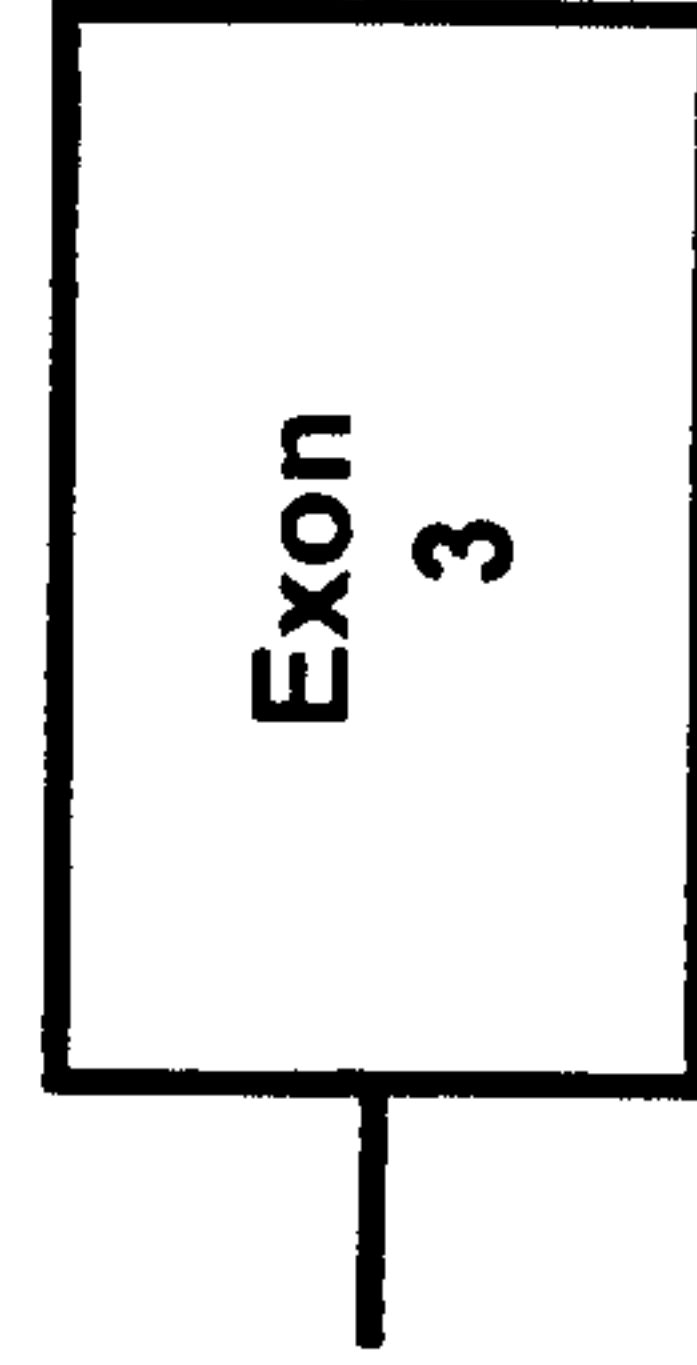


Figure 3