



(86) Date de dépôt PCT/PCT Filing Date: 2001/10/31
(87) Date publication PCT/PCT Publication Date: 2002/05/10
(45) Date de délivrance/Issue Date: 2013/05/14
(85) Entrée phase nationale/National Entry: 2003/04/28
(86) N° demande PCT/PCT Application No.: NZ 2001/000245
(87) N° publication PCT/PCT Publication No.: 2002/036824
(30) Priorités/Priorities: 2000/10/31 (NZ507888);
2000/12/06 (NZ508662)

(51) Cl.Int./Int.Cl. *C12Q 1/68* (2006.01),
A01K 67/027 (2006.01), *C07K 16/40* (2006.01),
C12N 15/12 (2006.01), *C12N 15/52* (2006.01),
C12N 15/85 (2006.01), *C12N 9/10* (2006.01)
(72) Inventeurs/Inventors:
GEORGES, MICHEL ALPHONSE JULIEN, BE;
COPPIETERS, WOUTER HERMAN ROBERT, BE;
GRISART, BERNARD MARIE-JOSEE JEAN, BE;
...
(73) Propriétaires/Owners:
GEORGES, MICHEL ALPHONSE JULIEN, BE;
COPPIETERS, WOUTER HERMAN ROBERT, BE;
...
(74) Agent: SMART & BIGGAR

(54) Titre : SELECTION ASSISTEE PAR MARQUEURS DE BOVINS A PRODUCTION LAITIERE AMELIOREE FAISANT
APPEL AU GENE DIACYLGLYCEROL ACYLTRANSFERASE DGAT1
(54) Title: MARKER ASSISTED SELECTION OF BOVINE FOR IMPROVED MILK PRODUCTION USING
DIACYLGLYCEROL ACYLTRANSFERASE GENE DGAT1

(57) **Abrégé/Abstract:**

The present invention provides a method of genotyping bovine for improved milk production traits by determining the DGAT1 genotypic state of said bovine, wherein the DGAT1 gene and polymorphisms within said gene have been found to be associated with such improved milk production traits.



(72) **Inventeurs(suite)/Inventors(continued)**: SNELL, RUSSELL GRANT, NZ; REID, SUZANNE JEAN, NZ;
FORD, CHRISTINE ANN, NZ; SPELMAN, RICHARD JOHN, NZ

(73) **Propriétaires(suite)/Owners(continued)**: GRISART, BERNARD MARIE-JOSEE JEAN, BE; SNELL, RUSSELL GRANT,
NZ; REID, SUZANNE JEAN, NZ; FORD, CHRISTINE ANN, NZ; SPELMAN, RICHARD JOHN, NZ

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 May 2002 (10.05.2002)

PCT

(10) International Publication Number
WO 02/36824 A1

(51) International Patent Classification⁷: **C12Q 1/68**,
C12N 15/12, A01K 67/027, C07K 16/40

(74) Agents: **CALHOUN, Douglas, C.** et al.; A J Park, Hud-
dart Parker Building, 6th Floor, Post Office Square, P.O.
Box 949, 6015 Wellington (NZ).

(21) International Application Number: PCT/NZ01/00245

(22) International Filing Date: 31 October 2001 (31.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
507888 31 October 2000 (31.10.2000) NZ
508662 6 December 2000 (06.12.2000) NZ

(71) Applicants and
(72) Inventors: **GEORGES, Michel, Alphonse, Julien**
[BE/BE]; 24 rue Vieux Tige, B-4161 Villers-aux-Tours
(BE). **COPPIETERS, Wouter, Herman, Robert**
[BE/BE]; Beekstraat 2, B-3401 Landen (BE). **GRISART,**
Bernard, Marie-Josée, Jean [BE/BE]; 2 Rue Bordia,
B-4218 Couthuin (BE). **SNELL, Russell, Grant** [NZ/NZ];
3/28 Tennyson Street, Balmoral, Auckland (NZ). **REID,**
Suzanne, Jean [NZ/NZ]; ViaLactia Biosciences (NZ)
Limited, P O Box 109-185, Newmarket, Auckland (NZ).
FORD, Christine, Ann [NZ/NZ]; ViaLactia Biosciences
(NZ) Limited, P O Box 109/185, Newmarket, Auckland
(NZ). **SPELMAN, Richard, John** [NZ/NZ]; 340 Horotiu
Road, R.D. 8, Hamilton (NZ).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN,
YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

Published:
— *with international search report*

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: MARKER ASSISTED SELECTION OF BOVINE FOR IMPROVED MILK PRODUCTION USING DIACYLGLYC-
EROL ACYLTRANSFERASE GENE DGAT1

(57) Abstract: The present invention provides a method of genotyping bovine for improved milk production traits by determining the DGAT1 genotypic state of said bovine, wherein the DGAT1 gene and polymorphisms within said gene have been found to be associated with such improved milk production traits.



WO 02/36824 A1

**Marker Assisted Selection of Bovine for Improved Milk Production Using Diacylglycerol
Acyltransferase gene DGAT1**

FIELD OF THE INVENTION

5 This invention relates to an application of marker assisted selection of bovine for a
quantitative trait loci (QTL) associated with milk production, particularly although by
no means exclusively, by assaying for the presence of at least one allele which is
associated with increased milk volume as well as improved milk composition. The
present invention also relates to the gene associated with the QTL, various
10 polymorphisms within the gene sequence, proteins encoded by these sequences as well
as to the application of all of these in the farming industry.

BACKGROUND

15 The genetic basis of bovine milk production is of immense significance to the dairy
industry. An ability to modulate milk volumes and content has the potential to alter
farming practices and to produce products which are tailored to meet a range of
requirements. In particular, a method of genetically evaluating bovine to select those
which express desirable traits, such as increased milk production and improved milk
20 composition, would be desirable.

To date, bovine genomics are poorly understood and little is known regarding the genes
which are critical to milk production. While there have been reports of quantitative trait
loci (QTLs) on bovine chromosome 14 postulated to be associated with milk production
25 (Coppieters *et al* (1998)), the specific genes involved have not to date been identified.

Marker assisted selection, which provides the ability to follow a specific favourable
genetic allele, involves the identification of a DNA molecular marker or markers that
segregate with a gene or group of genes associated with a QTL. DNA markers have
30 several advantages. They are relatively easy to measure and are unambiguous, and as
DNA markers are co-dominant, heterozygous and homozygous animals can be
distinctively identified. Once a marker system is established, selection decisions are
able to be made very easily as DNA markers can be assayed at any time after a DNA
containing sample has been collected from an individual infant or adult animal, or even
35 earlier as it is possible to test embryos in vitro if such embryos are collected.

The applicants have now identified a gene responsible for the QTL effect on bovine chromosome 14 as well as a number of polymorphisms which are associated with distinct genetic merits of animals for milk composition and volume.

- 5 It is an object of the present invention to provide an application method for marker assisted selection of this bovine gene, and in particular, of the polymorphisms in the bovine gene which are associated with increased milk volume and altered milk composition; and/or to provide genetic markers for use in such a method; and/or to provide the nucleic acid and amino acid sequences of this gene and encoded
10 polypeptide; and/or to provide animals selected using the method of the invention as well as milk produced by the selected animals; and/or to provide the public with a useful choice.

SUMMARY OF THE INVENTION

15

This invention relates to the discovery of the bovine Diacylglycerol-o-acyltransferase (DGAT1) gene and polymorphisms within the bovine DGAT1 gene which are associated with increased milk yield and altered milk composition.

- 20 More specifically, several polymorphisms in the bovine DGAT1 gene have been identified distinguishing multiple DGAT1 alleles in different cattle breeds. These polymorphisms include: K232A (Bases 6829/30 AA-CG nucleic acid change and K-A amino acid change); Nt984+8(Base 7438 A-G nucleic acid change); Nt984+26(Base 7456 C-T nucleic acid change); Nt1470+85(Base 8402 C-T nucleic acid change); Nt191+435 (Base
25 626 T-G nucleic acid change); Nt191-3321 (Base 3512 T-G nucleic acid change); Nt279+144 (Base 4040 T-C nucleic acid change); Nt279+1067 (Base 4963 A-G nucleic acid change); Nt279+1107 (Base 5003 G-A nucleic acid change); Nt358 (Base 5997 C-T nucleic acid change); Nt754+3 (Base 6892 G-A nucleic acid change); Nt897+32 (Base 7224/5 GG-AC nucleic acid change); Nt1251+42 (Base 7987 G-A nucleic acid change)
30 as summarised in Table 1. In particular, DGAT1 alleles characterized by the K232A mutation have been identified as being associated with an increased milk volume and altered milk composition in animals dependent upon whether they are homozygous with or without the mutation or heterozygous carrying one mutated allele. More specifically, the presence of the K232A mutation results in a decrease in milkfat percentage, milkfat
35 yield, solid fat content and milk protein percentage, while increasing milk volume and milk protein yield.

00345-11

3

The present invention thus relates to the use of the polymorphisms in a method of identification and selection of a bovine having at least one of said polymorphisms as well as to providing markers specific for such identification. Kits comprising said markers for use in marker selection also form part of the present invention as do
5 animals so selected, the milk produced by such selected animals and products produced from such milk, particularly as such milk and milk products affect processing and/or health characteristics for consumers.

In particular, the present invention is directed to a method of genotyping cows or bulls
10 for one or more of the polymorphisms disclosed herein, selected cows or bulls so genotyped and milk and semen from said selected cows and bulls respectively.

According to a further aspect the present invention is directed to the isolated DGAT1 nucleic acid and allelic nucleic acid molecules comprising polymorphisms as well as to
15 the proteins encoded thereby and their polypeptide sequences. Antibodies raised against said proteins are also contemplated, as are vectors comprising the nucleic acid molecules, host cells comprising the vectors; and protein molecules expressed in said host cells; and the application of all of them in the farming industry.

20 In particular, such applications include methods for modulating milk production and/or composition in a lactating bovine by affecting DGAT1 activity, by reducing the activity of DGAT1 (e.g. by use of specific ribozymes, antisense sequences and/or antibodies, or by transgenic technology to produce a "knock out" bovine and/or bovine with introduced transgenes containing the DGAT1 gene and/or variations of this gene
25 driven by various promoters).

68348-71

3a

In one aspect, the invention relates to a method of identifying a bovine, or a sample derived from said bovine, with a genotype indicative of increased milk volume production or increased milk protein to fat ratio in said bovine, the method comprising determining the presence of alanine (A) at amino acid position 232 in the diacylglycerol acyltransferase1 (DGAT1) protein of SEQ ID NO:2, wherein the presence of alanine is
5 determined with respect to DNA, mRNA or protein from the bovine.

In another aspect, the invention relates to a method of selecting a bovine that possesses a genotype indicative of increased milk volume production or increased milk protein to fat ratio in said bovine, the method comprising the step of selecting a
10 bovine identified by the method as described above.

In another aspect, the invention relates to a method of identifying a bovine, or a sample derived from said bovine, with a genotype indicative of decreased milk volume production or decreased milk protein to fat ratio in said bovine, the method comprising determining the presence of lysine (K) at amino acid position 232 in the diacylglycerol
15 acyltransferase1 (DGAT1) protein of SEQ ID NO:2, wherein the presence of lysine is determined with respect to DNA, mRNA or protein from the bovine.

In another aspect, the invention relates to a method of selecting a bovine that possesses a genotype indicative of decreased milk volume production or decreased milk protein to fat ratio in said bovine, the method comprising the step of selecting a
20 bovine identified by the method as described above.

In another aspect, the invention relates to a primer comprising a nucleotide sequence having at least 12 contiguous bases of SEQ ID NO: 1 or 4 and further comprising the K232A (AA to GC) polymorphism of SEQ ID NO: 13 associated with improved milk production traits.

25 In another aspect, the invention relates to use of the DGAT1 gene sequence of SEQ ID NO: 1 or 4, comprising the K232A (AA to GC) polymorphism of SEQ ID NO: 13 in the identification of one or more molecular DNA markers useful in the method as described above.

68348-71

3b

In another aspect, the invention relates to use of one or more polymorphic sequences selected from the group consisting of SEQ ID NOs: 7 to 19 in the method as described above.

In another aspect, the invention relates to a kit, comprising the primer as
5 described above or a probe selected from any 5 or more contiguous nucleotides of the DGAT1 sequence of SEQ ID NO: 1 or 4 comprising the K232A (AA to GC) polymorphism of SEQ ID NO: 13 which is sufficiently complementary with said nucleic acid sequence so as to bind thereto in 6x sodium citrate/sodium chloride at 45°C, and instructions for using the primer or the probe for genotyping a bovine with respect to milk composition and
10 volume associated with DGAT1.

In another aspect, the invention relates to an isolated nucleic acid molecule having the sequence of SEQ ID NOs: 1 or 4 comprising the K232A (AA to GC) polymorphism of SEQ ID NO: 13 wherein said polymorphism is associated with improved milk production traits.

15 In another aspect, the invention relates to an isolated nucleic acid molecule comprising the whole or at least 20 nucleotides of the nucleotide sequence of SEQ ID NO: 1 or 4 or which varies from the sequence due to the degeneracy of the genetic code, or a nucleic acid strand capable of hybridising with said nucleic acid molecule in 6x sodium citrate/sodium chloride at 45°C, wherein said nucleic acid molecule further
20 comprises the K232A (AA to GC) polymorphism of SEQ ID NO: 13 which is associated with improved milk production traits.

In another aspect, the invention relates to an isolated mRNA transcribed from DNA having a sequence which corresponds to the nucleic acid molecule as described above.

25 In another aspect, the invention relates to a recombinant cloning vector comprising the nucleic acid molecule as described above.

In another aspect, the invention relates to a prokaryotic or eukaryotic cell containing the cloning vector as described above.

06348-11

3c

In another aspect, the invention relates to a transfected cell line which expresses a protein encoded by the nucleic acid molecule as described above.

In another aspect, the invention relates to a diagnostic kit useful in detecting the DNA molecule as described above, comprising a first and second
5 primer for amplifying the DNA, the primers being complementary to nucleotide sequences of the DNA upstream and downstream, respectively, of the K232A (AA to GC) polymorphism of SEQ ID NO: 13 in the portion of the DNA encoding DGAT1 protein of SEQ ID NO: 2 which results in altered relative milk lipid and protein
10 production and milk volume, wherein at least one of the nucleotide sequences is selected to be from a non-coding region of the DGAT1 gene of SEQ ID NO: 1.

In another aspect, the invention relates to a process for producing a protein encoded by the nucleic acid molecule as described above, comprising the steps: a) preparing a DNA fragment including a nucleotide sequence which encodes the protein; b) incorporating the DNA fragment into an expression vector to obtain a
15 recombinant DNA molecule which includes the DNA fragment and is capable of undergoing replication; c) transforming a host cell with the recombinant DNA molecule to produce a transformant which can express the protein; d) culturing the transformant to produce the protein; and e) recovering the protein from resulting cultured mixture.

20 In another aspect, the invention relates to a purified protein having an amino acid sequence encoded by the isolated nucleic acid molecule as described above, having diacylglycerol acyltransferase activity.

In another aspect, the invention relates to an antibody raised against the protein as described above.

25 In another aspect, the invention relates to use of a transgenic non-human animal having cells that contain a transgene, which transgene is introduced into the animal or an ancestor of the animal at a prenatal stage, wherein said transgene comprises a bovine cDNA comprising: a) the nucleotide sequence of SEQ ID NO: 4 including the K232A (AA to GC) polymorphism of SEQ ID NO: 13;
30 b) a sequence capable of hybridising to SEQ ID NO: 4 in 6x sodium chloride/sodium

citrate (SSC) at 45°C, which includes the K232A (AA to GC) polymorphism of SEQ ID NO: 13; or c) a sequence which varies from SEQ ID NO: 4 due to the degeneracy of the genetic code and which includes the K232A polymorphism; to produce milk.

The invention will now be described with reference to the Figures of the accompanying drawings in which:

Figure 1: Shows a BAC contig spanning the *BULGE13-BULGE09* interval relative to a schematic diagram of bovine chromosome 14 and a schematic diagram showing the location of the genetic markers. The most likely position of the QTL is shown as a bar on the FISH-anchored linkage map proximal to BTA14q. The BACs composing the contigs spanning the *BULGE13-BULGE09* interval are shown as a series of horizontal lines. The symbols on each BAC indicate their individual STS content: solid circles

correspond to STS derived from BAC ends, open boxes to microsatellite markers, and solid triangles to gene-specific Comparative Anchored Tagged Sequences. The arrow heads mark the BACs from which the respective BAC end STS were derived. The length of the lines do not reflect the actual insert size of the corresponding BACs. The BAC contig was aligned with the orthologous human HSA8q24.3 genomic "golden path" sequence represented according to the Ensembl Human Genome Server: individual sequence contigs are shown in alternating light and dark; a horizontal line indicates a gap in the sequence assembly; genetic markers are indicated under the contig map; the lines and boxes above the contig map represent "curated", "predicted known" or "predicted novel" genes.

Figures 2a and 2b: Show the genomic sequence of the bovine DGAT1 gene. Figure 2a is the 31 base pair sequence upstream but adjacent to the ATG or translation start site and is 5'UTR. Figure 2b is the genomic sequence in the bovine DGAT1 gene from the ATG translation start site (base 1) through to genomic sequence flanking the gene at the 3' end. The significant features including intron/exon boundaries, polymorphic sites, polyadenylation signal, and alternate splicing site and some of the primer sequences used in the assays described herein, are indicated;

Figure 3: Shows the genomic organization, four polymorphisms and haplotypes found in the bovine *DGAT1* gene. Leader and trailer sequences are shown in light grey, coding sequences in dark grey and intronic sequences as a hollow line. The positions of four of the identified polymorphisms are marked as shown on the gene, and detailed in the underlying boxes including the corresponding sequence traces. All the sequence variations are summarised in Table 1. The four *DGAT1* haplotypes which were found in the Dutch and New-Zealand Holstein-Friesian population as defined by these polymorphisms are shown and referred to as "SHQ-D", "SHQ-NZ", "SHQ-III" for the fat increasing haplotypes and "shs" for the fat decreasing haplotype;

Figure 4a: Shows the corresponding full length amino acid sequence for DGAT1 sequence of Figure 2b including annotation of the amino acid substitution;

Figure 4b: Shows the amino acid sequence predicted as a result of alternate splicing with exon VIII;

Figure 5: Shows the multiple peptide alignment of a portion of the *DGAT1* protein flanking the K232A substitution from *Bos taurus*, *Bison bison*, *Ovis aries*, *Sus scrofa*,

Homo sapiens, *Cercopithecus aethiops*, *Mus musculus domesticus* and *Rattus norvegicus* showing the evolutionary conservation of the lysine mutated in the bovine K232A polymorphism;

- 5 **Figure 6:** A. Shows the frequency distribution of observed *DGAT1* SNP haplotypes in the Dutch and New Zealand Holstein-Friesian dairy cattle populations. B-D. Shows the frequency distribution of the combined microsatellite (*BULGE09-BULGE11*) and SNP *DGAT1* haplotypes. The H^{Q-D} and H^{Q-NZ} haplotypes are shown; and
- 10 **Figure 7:** Shows the lod score due to LD when including (+) or excluding (-) the four *DGAT1* polymorphisms shown in Figure 3 in a combined linkage and LD multipoint maximum likelihood mapping method. The lod score corresponds to the $\log_{10}$ of the ratio between the likelihood of the data assuming LD and linkage between the markers and the likelihood of the data assuming linkage in the absence of LD. The positions of
- 15 the microsatellites and SNP markers utilized in the analysis are shown on the X-axis, while the position of the *DGAT1* SNPs is marked by a red arrow at the top of the figure.

DETAILED DESCRIPTION OF THE INVENTION

- 20 It has been discovered for the first time that the *DGAT1* gene in bovine is associated with the QTL on chromosome 14 which is linked with improved milk production traits. More particularly, a number of novel polymorphisms on the *DGAT1* gene have been discovered. It is thought that one or more of these polymorphisms is responsible for these traits.

25

- The method used for isolating genes which cause specific phenotypes is known as positional candidate cloning. It involves: (i) the chromosomal localisation of the gene which causes the specific phenotype using genetic markers in a linkage analysis; and (ii) the identification of the gene which causes the specific phenotype amongst the
- 30 "candidate" genes known to be located in the corresponding region. Most of the time these candidate genes are selected from available mapping information in humans and mice.

- The tools required to perform the initial localisation (step (i) above) are microsatellite
- 35 marker maps, which are available for livestock species and are found in the public domain (Bishop *et al.*, 1994; Barendse *et al.*, 1994; Georges *et al.*, 1995; and Kappes, 1997). The tools required for the positional candidate cloning, particularly the BAC libraries, (step (ii) above) are partially available from the public domain. Genomic

libraries with large inserts constructed with Bacterial Artificial Chromosomes (BAC) are available in the public domain for most livestock species including cattle. For general principles of positional candidate cloning, see Collins, 1995 and Georges and Anderson, 1996.

5

Recently, a quantitative trait locus (QTL) with major effect on milk solids composition, located at the centromeric end of bovine chromosome 14, has been reported (Coppieters *et al.*, (1998)). This QTL was shown to effect milk fat content and in particular to significantly affect protein %, volume, protein yield and fat yield of milk. The linkage study as well as subsequent marker assisted segregation analyses allowed for the identification of thirteen Holstein-Friesian sires predicted to be heterozygous "Qq" for the corresponding QTL (Coppieters *et al.*, (1998); Riquet *et al.*, (1999)).

10

Linkage disequilibrium methods were applied to refine the map position of the QTL to a ≈ 5 cM interval bounded by microsatellite markers BULGE09 and BULGE30.

15

A bovine DGAT1 nucleotide sequence was determined by the applicants and is shown in Figures 2a and 2b with the corresponding amino acid sequences (long and short forms) being shown in Figures 4a and 4b respectively. Table 1 sets out all the polymorphisms located to date with reference to the sequence in Figure 2b. Some of the genetic polymorphisms identified in the bovine DGAT1 gene are reported in Figure 3. The nucleic acid and protein sequences of the DGAT1 alleles including the K232A mutation are shown in Figures 2a and 2b (SEQ ID NOs: 3 and 1), annotated to show the alternatively spliced forms. The cDNA sequence is also set out in SEQ ID NO: 4.

20

25

The sequence information in the Figures gives rise to numerous, and separate, aspects of the invention.

30

35

In one aspect, the invention provides a method of determining genetic merit of a bovine with respect to milk composition and volume which comprises the step of determining the bovine DGAT1 genotypic state of said bovine. In particular, this method is useful for genotyping and selecting cows and bulls having the desired genotypic state so that milk and semen may be collected from said cows and bulls respectively. Such semen would be useful for breeding purposes to produce bovine having the desired genotypic and, as a result, phenotypic state. In addition, cows genotyped by the methods of the present invention are also useful for breeding purposes, particularly for breeding with the selected bulls and/or to be artificially inseminated with the semen from selected

bulls. The embryos and offspring produced by such cows also form part of the present invention.

In one embodiment, the genotypic state is determined with respect to DNA obtained
5 from said bovine.

Alternatively, said genotypic state is determined with reference to mRNA obtained from said bovine.

10 In yet a further embodiment, the genotypic state is determined with reference to the amino acid sequence of expressed bovine DGAT1 protein obtained from said bovine.

Conveniently, in said method, the genotypic state of DNA encoding bovine DGAT1 is determined, directly or indirectly.

15

Alternatively, in said method the genotypic state of at least one nucleotide difference from the nucleotide sequence encoding bovine DGAT1 is determined, directly or indirectly.

20 More specifically, in said method the genotypic state of bovine DGAT1 allele(s) characterised by one or more of the polymorphisms shown in Table 1 below, is determined, directly or indirectly.

68348-71

8

Table 1:

Table of polymorphisms in the bovine DGAT1 gene

Start codon (atg); the a residue is denoted as position 1

5

Base number relative to exonic sequence ¹	Nucleotide distance from start of SEQ ID NO:1	substitution	SEQ ID NO:	Intron/exon #
Nt 191 + 435	626	T-G CAGTGCTAGGGG CAGTGCAGGGG	22 23	Intron 1
Nt 191 + 3321	3512	T-G GCATTGCGCT GCATGGCGCT	24 25	Intron 1
Nt 279 + 144	4040	T-C TACCCTGGGAC TACCCCGGGAC	26 27	Intron 2
Nt 279 +1067	4963	A-G CTCTTAGCAGC CTCTTGGCAGC	28 29	Intron 2
Nt 279 +1107	5003	G-A ACAGGCAACT ACAGACAACT	30 31	Intron 2
Nt 358	5997	C-T TGTCTCTGTTC TGTCTTTGTTC	32 33	Exon IV
Nt 692	6829	AA-GC GGTAAGAAGGCCAA (Q) GGTAAGGCGGCCAA (q)	34 35	K232A Exon VIII *
Nt 754 +3	6892	G-A GCGGTGAGGAT GCGGTAAGGAT	36 37	Intron VIII
Nt 897 +32	7224	GG-AC GGGGGGGGGGGACTCT GGGGGACGGGGACTCT	38 39	Intron X
Nt 984 +8	7438	A-G GAGTGACCTGC GAGTGGCCTGC	40 41	Intron XII *
Nt 984 +26	7456	C-T GGACGGGTGGG GGACGTGTGGG	42 43	Intron XII *
Nt1251 +42	7987	G-A GGTGGGGGTGG GGTGGAGGTGG	44 45	Intron XV
Nt 1470 +85	8402	C-T CTGGGCGCAGC CTGGGTGCAGC	46 47	3' flanking region *

The numbers given are for the actual nucleotide or in the case of two nucleotide substitutions to the first nucleotide in the variation (counting 5' to 3')

*More detail of these polymorphisms is given in Figure 2b.

¹ e.g. Nt 191 represents nucleotide number 191 from the start site of the coding sequence, + 435 represents number of nucleotides from and including base 192 in the genomic sequence (intron 1) to the polymorphic nucleotide. The polymorphic nucleotides are shaded

10

Preferably, the invention is directed to a method of determining the genotypic state of bovine DGAT1 allele(s) by determining the presence of the K232A polymorphism, either directly or indirectly.

15

There are numerous art standard methods known for determining whether a particular DNA sequence is present in a sample. An example is the Polymerase Chain Reaction

(PCR). A preferred aspect of the invention thus includes a step in which ascertaining whether a polymorphism(s) in the sequence of DGAT1 DNA is present, includes amplifying the DNA in the presence of primers based on the nucleotide sequence of the DGAT1 gene and flanking sequence, and/or in the presence of a primer containing at least a portion of a polymorphism as disclosed herein and which when present results in altered relative milk lipid and protein production, and milk volume.

A primer of the present invention, used in PCR for example, is a nucleic acid molecule sufficiently complementary to the sequence on which it is based and of sufficient length to selectively hybridise to the corresponding portion of a nucleic acid molecule intended to be amplified and to prime synthesis thereof under *in vitro* conditions commonly used in PCR. Likewise, a probe of the present invention, is a molecule, for example a nucleic acid molecule of sufficient length and sufficiently complementary to the nucleic acid molecule of interest, which selectively binds under high or low stringency conditions with the nucleic acid sequence of interest for detection thereof in the presence of nucleic acid molecules having differing sequences.

In another aspect, the invention provides a method for determining the genetic merit of bovine with respect to milk content and volume with reference to a sample of material containing mRNA obtained from the bovine. This method includes ascertaining whether a polymorphism(s) in the sequence of the mRNA encoding DGAT1 is present. The presence of such polymorphisms again indicates an association with altered relative milk lipid and protein production and milk volume.

Again, if an amplification method such as PCR is used in ascertaining whether a polymorphism(s) in the sequence of the mRNA encoding (DGAT1) is present, the method includes reverse transcribing the mRNA using a reverse transcriptase to generate a cDNA and then amplifying the cDNA in the presence of a pair of primers complementary to a nucleotide sequence encoding a protein having biological activity of wild type DGAT1.

In a further aspect, the invention includes the use of a probe in the methods of genotyping according to the invention wherein the probe is selected from any 5 or more contiguous nucleotides of the DGAT1 sequence as shown in Figure 2b, which is therefore sufficiently complementary with a nucleic acid sequence encoding such bovine DGAT1, or its complement, so as to bind thereto under stringent conditions. Diagnostic kits containing such a probe are also included. Such probes may be selected from

ForAA (FAM): CGTTGGCCTTCTTA or DgatADGC (VIC): TTGGCCGCCTTACC. (SEQ ID NOs: 20 and 21 respectively.)

5 The invention further includes isolated nucleic acid molecules encoding the DGAT1 variant proteins i.e. those proteins encoded by SEQ ID NOs: 1 and 4 (Figure 2b), comprising one or more polymorphisms of SEQ ID NOs: 7 to 19 (Table 1), or a fragment or variant thereof. Particularly, the invention includes an isolated nucleic acid molecule comprising a DNA molecule having in whole or in part the nucleotide sequence identified in Figure 2b or which varies from the sequence due to the degeneracy of the
10 genetic code, or a nucleic acid strand capable of hybridising with said nucleic acid molecule under stringent hybridisation conditions.

15 The invention includes isolated mRNA transcribed from DNA having a sequence which corresponds to a nucleic acid molecule of the invention.

The invention includes isolated DNA in a recombinant cloning vector and a prokaryotic or eukaryotic cell containing and expressing heterologous DNA of the invention.

20 The invention includes a transfected cell line which expresses a protein encoded by the nucleic acid molecules of the invention.

The invention also includes a primer composition useful for detection of the presence of one or more polymorphisms associated with improved milk production traits in bovine DNA encoding DGAT1 and/or the presence of DNA encoding a variant protein. In one
25 form, the composition can include a nucleic acid primer substantially complementary to a nucleic acid sequence encoding DGAT1. The nucleic acid sequence can in whole or in part be that identified in Figure 2b. Diagnostic kits including such a composition are also included.

30 The invention further provides a diagnostic kit useful in detecting DNA encoding a variant DGAT1 protein in bovine which includes first and second primers for amplifying the DNA, the primers being complementary to nucleotide sequences of the DNA upstream and downstream, respectively, of a polymorphism in the portion of the DNA encoding DGAT1 which results in altered relative milk lipid, solid fat content and
35 protein production and milk volume, wherein at least one of the nucleotide sequences is selected to be from a non-coding region of the DGAT1 gene. The kit can also include a third primer complementary to a polymorphism, disclosed herein, located on the DGAT1 gene.

The invention includes a process for producing a protein of the invention, including preparing a DNA fragment including a nucleotide sequence which encodes the protein; incorporating the DNA fragment into an expression vector to obtain a recombinant DNA molecule which includes the DNA fragment and is capable of undergoing replication; transforming a host cell with the recombinant DNA molecule to produce a transformant which can express the protein; culturing the transformant to produce the protein; and recovering the protein from resulting cultured mixture.

Thus in a further aspect, the invention provides a purified protein encoded by the nucleic acid molecule of the invention and having biological activity of DGAT1. The terms "isolated" and "purified" as used herein, each refer to a protein substantially free of cellular material or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesised. In certain preferred embodiments, the protein having biological activity of DGAT1 comprises an amino acid sequence and variants shown in Figures 4a and 4b (SEQ ID NOs: 2, 5 and 6). Furthermore, proteins having biological activity of DGAT1 that are encoded by nucleic acids which hybridise under stringent conditions to a nucleic acid comprising a nucleotide sequence shown in Figure 2b (SEQ ID NOs: 1 and 4) are encompassed by the invention.

Proteins of the invention having DGAT1 activity can be obtained by expression of a nucleic acid coding sequence in a suitable host cell using techniques known in the art. Suitable host cells include prokaryotic or eukaryotic organisms or cell lines, for example, yeast, *E. coli*, insect cells and COS1 cells. The recombinant expression vectors of the invention can be used to express a protein having DGAT1 activity in a host cell in order to isolate the protein. The invention provides a method of preparing a purified protein of the invention comprising introducing into a host cell a recombinant nucleic acid encoding the protein, allowing the protein to be expressed in the host cell and isolating and purifying the protein. Preferably, the recombinant nucleic acid is a recombinant expression vector. Proteins can be isolated from a host cell expressing the protein and purified according to standard procedures of the art, including ammonium sulfate precipitation, column chromatography (eg. ion exchange, gel filtration, affinity chromatography, etc.) electrophoresis, and ultimately, crystallisation (see generally "Enzyme Purification and Related Techniques". *Methods in Enzymology*, 22, 233-577 (1971)).

Alternatively, the protein or parts thereof can be prepared by chemical synthesis using techniques well known in the chemistry of proteins such as solid phase synthesis (Merrifield, 1964), or synthesis in homogeneous solution (Houbenwycyl, 1987).

5 It will of course be understood that a variety of substitutions of amino acids is possible while preserving the structure responsible for activity of the DGAT1 proteins disclosed herein. Conservative substitutions are described in the patent literature, as for example, in United States Patent No 5,264,558 or 5,487,983. It is thus expected, for example, that interchange among non-polar aliphatic neutral amino acids, glycine,
10 alanine, proline, valine and isoleucine, would be possible. Likewise, substitutions among the polar aliphatic neutral amino acids, serine, threonine, methionine, asparagine and glutamine could possibly be made. Substitutions among the charged acidic amino acids, aspartic acid and glutamic acid, could probably be made, as could substitutions among the charged basic amino acids, lysine and arginine. Substitutions
15 among the aromatic amino acids, including phenylalanine, histidine, tryptophan and tyrosine would also likely be possible. These sorts of substitutions and interchanges are well known to those skilled in the art. Other substitutions might well be possible. Of course, it would also be expected that the greater percentage of homology ie. sequence similarity, of a variant protein with a naturally occurring protein, the greater
20 the retention of activity.

A further advantage may be obtained through chimeric forms of the proteins, as known in the art. A DNA sequence encoding each entire protein, or a portion of the protein, could be linked, for example, with a sequence coding for the C-terminal portion of *E. coli*
25 β -galactosidase to produce a fusion protein.

The proteins of the invention, or portions thereof, have numerous applications in turn. By way of example, each protein can be used to prepare antibodies which bind to a distinct epitope in an unconserved region of the protein. An unconserved region of the
30 protein is one which does not have substantial sequence homology to other proteins.

Still further, the invention includes an antibody to a bovine DGAT1 variant protein encoded by a nucleotide sequence of the present invention as well as a diagnostic kit containing such an antibody.

35

Conventional methods can be used to prepare the antibodies. For example, by using a DGAT1 peptide, polyclonal antisera or monoclonal antibodies can be made using

standard methods. A mammal, (eg. a mouse, hamster, or rabbit) can be immunised with an immunogenic form of the peptide which elicits an antibody response in the mammal. Techniques for conferring immunogenicity on a peptide include conjugation to carriers or other techniques well known in the art. For example, the peptide can be administered in the presence of adjuvant. The progress of immunisation can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassay can be used to assess the levels of antibodies. Following immunisation, antisera can be obtained and, if desired, polyclonal antibodies isolated from the sera.

To produce monoclonal antibodies, antibody producing cells (lymphocytes) can be harvested from an immunised animal and fused with myeloma cells by standard somatic cell fusion procedures, thus immortalising these cells and yielding hybridoma cells. Such techniques are well known in the art. For example, the hybridoma technique originally developed by Kohler and Milstein (Kohler, 1975) as well as other techniques such as the human B-cell hybridoma technique (Kozbor, 1983) and screening of combinatorial antibody libraries (Huse, 1989). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with the peptide, and monoclonal antibodies isolated.

The term "antibody" as used herein is intended to include fragments thereof which are also specifically reactive with the target protein. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab')₂ fragments can be generated by treating antibody with pepsin. The resulting F(ab')₂ fragment can be treated to reduce disulfide bridges to produce Fab' fragments.

Another method of generating specific antibodies, or antibody fragments, reactive against the target proteins is to screen expression libraries encoding immunoglobulin genes, or portions thereof, expressed in bacteria, with peptides produced from the nucleic acid molecules of the present invention. For example, complete Fab fragments, VH regions and FV regions can be expressed in bacteria using phage expression libraries. See for example Ward *et al.*, Huse *et al.*, and McCafferty *et al.* (Ward, 1989); Huse 1989; McCafferty, 1990). Screening such libraries with, for example, a DGAT1 protein can identify immunoglobulin fragments reactive with that DGAT1. Alternatively, the SCID-hu mouse developed by Genpharm can be used to produce antibodies, or fragments thereof.

The polyclonal, monoclonal or chimeric monoclonal antibodies can be used to detect the proteins of the invention, portions thereof or closely related isoforms in various biological materials. For example, they can be used in an ELISA, radioimmunoassay or histochemical tests. Thus, the antibodies can be used to quantify the amount and
5 location of a DGAT1 protein of the invention, portions thereof or closely related isoforms in a sample in order to determine the role of DGAT1 proteins. Using methods described hereinbefore, polyclonal, monoclonal antibodies, or chimeric monoclonal antibodies can be raised to non-conserved regions of DGAT1 and used to distinguish a particular DGAT1 from other proteins.

10

The polyclonal or monoclonal antibodies can be coupled to a detectable substance or reporter system. The term "coupled" is used to mean that the detectable substance is physically linked to the antibody. Suitable detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials and
15 radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, and acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride and
20 phycoerythrin; an example of a luminescent material includes luminol; and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S and ^3H . In a preferred embodiment, the reporter system allows quantitation of the amount of protein (antigen) present.

Such an antibody-linked reported system could be used in a method for determining
25 whether a fluid or tissue sample of a bovine contains a deficient amount or an excessive amount of the relevant DGAT1 protein. Given a normal threshold concentration of such a protein, test kits can be developed.

The availability of such antibodies gives rise to further applications. One is a diagnostic
30 kit for identifying cells comprising an antibody (such as a monoclonal antibody) which binds to a protein comprising an amino acid sequence shown in Figure 4a and 4b; means for detecting the antibody when bound to the protein, unreacted protein or unbound antibody; means for determining the amount of protein in the sample; and means for comparing the amount of protein in the sample with a standard. In some
35 embodiments of the invention, the detectability of the antibody which binds to a specific DGAT1 protein is activated by the binding (eg. change in fluorescence spectrum, loss of

radioisotopic label). The diagnostic kit can also contain an instruction manual for use of the kit.

Antibody-based diagnostics are of course not the only possibility. A further diagnostic
5 kit comprises a nucleotide probe complementary to the sequence, or an oligonucleotide
fragment thereof, shown in Figure 2a and 2b, for example, for hybridisation with mRNA
from a sample of cells; means for detecting the nucleotide probe bound to mRNA in the
sample with a standard. In a particular aspect, the kit of this aspect of the invention
includes a probe having a nucleic acid molecule sufficiently complementary with a
10 sequence identified in Figure 2a and 2b, or its complement, so as to bind thereto under
stringent conditions. "Stringent hybridisation conditions" takes on its common
meaning to a person skilled in the art. Appropriate stringency conditions which
promote nucleic acid hybridisation, for example, 6x sodium chloride/sodium citrate
(SSC) at about 45°C are known to those skilled in the art, including in Current
15 Protocols in Molecular Biology, John Wiley & Sons, NY (1989). Appropriate wash
stringency depends on degree of homology and length of probe. If homology is 100%, a
high temperature (65°C to 75°C) may be used. If homology is low, lower wash
temperatures must be used. However, if the probe is very short (<100bp), lower
temperatures must be used even with 100% homology. In general, one starts washing
20 at low temperatures (37°C to 40°C), and raises the temperature by 3-5°C intervals until
background is low enough not to be a major factor in autoradiography. The diagnostic
kit can also contain an instruction manual for use of the kit.

One of the major applications of the present invention is in the marker assisted
25 selection of bovines having a polymorphism in the DGAT1 gene and which are
associated with improved milk production traits. The invention therefore provides a
diagnostic kit which can be used to determine the DGAT1 genotype of bovine genetic
material, for example. One kit includes a set of primers used for amplifying the genetic
material. A kit can contain a primer including a nucleotide sequence for amplifying a
30 region of the genetic material containing one of the polymorphisms described herein.
Such a kit could also include a primer for amplifying the corresponding region of the
normal DGAT1 gene, i.e. the sequence without polymorphisms. Usually, such a kit
would also include another primer upstream or downstream of the region of interest
complementary to a coding and/or non-coding portion of the gene. These primers are
35 used to amplify the segment containing the mutation, i.e. polymorphism, of interest.

In particular, the invention is directed to the use of the polymorphisms in the DGAT1 gene in the genotyping of cows and bulls as well as to cows and bulls selected by such genotyping which have one or more of said polymorphisms in the DGAT1 gene. Such bulls so selected are of valuable breeding stock and the invention is also directed to the semen produced by such selected bulls for breeding purposes. Cows so selected are also useful as breeding stock as are their offspring. In addition, such cows may produce valuable dairy herds as the milk produced by such cows is produced in greater volumes than equivalent non-selected cows, and/or has an altered composition in that it comprises less milkfat and more milk protein. Such milk and products made therefrom also form part of the invention. It is also noted that the milk from these selected cows will be valuable as the fat content is not only decreased but is also characterised by being softer. Without being bound by theory, it is thought that this increased fat softness is due to the fatty acid composition being such that there is less saturated and more unsaturated fat in the milk of selected cows. Thus it is anticipated that products made from such milk will have processing advantages, such as in the production of more spreadable butter, as well as having a health benefit on consumers, as generally unsaturated fats are considered to be more "healthy" than saturated fats. The protein composition of milk produced by such selected cows is also altered. In particular, such milk comprises an altered protein yield compared to milk for non-selected cows and the casein:whey ratio is also altered which makes such milk valuable for cheese production.

Thus, the present invention involves genotyping bovine, both cows and bulls, for the DGAT1 polymorphisms disclosed herein, selected cows and bulls so genotyped, milk and semen produced by the selected cows and bulls so genotyped, offspring produced by the selected bovine, including embryos and cells (including cell lines) useful for cloning said selected bovine.

The actual genotyping is carried out using primers that target specific polymorphisms as described herein and that could function as allele-specific oligonucleotides in conventional hybridisation, Taqman assays, OLA assays, etc. Alternatively, primers can be designed to permit genotyping by microsequencing.

One kit of primers can include first, second and third primers, (a), (b) and (c), respectively. Primer (a) is based on a region containing a DGAT1 mutation such as described above. Primer (b) encodes a region upstream or downstream of the region to be amplified by primer (a) so that genetic material containing the mutation is amplified,

by PCR, for example, in the presence of the two primers. Primer (c) is based on the region corresponding to that on which primer (a) is based, but lacking the mutation. Thus, genetic material containing the non-mutated region will be amplified in the presence of primers (b) and (c). Genetic material homozygous for the DGAT1 gene will thus provide amplified products in the presence of primers (b) and (c). Genetic material
5 homozygous for the mutated gene will thus provide amplified products in the presence of primers (a) and (b). Heterozygous genetic material will provide amplified products in both cases.

- 10 The present invention also contemplates the modulation of milk production and content in non-human animals by modulating the activity of the DGAT1 protein. In particular, this aspect of the invention includes a method of modulating milk production and/or milk content in a lactating bovine, the method comprising administering to the bovine an effective amount of a nucleic acid molecule substantially complementary to at least a
15 portion of mRNA encoding the bovine DGAT1 variant proteins and being of sufficient length to sufficiently reduce expression of said DGAT1, i.e. by use of antisense nucleic acids.

- Antisense nucleic acids or oligonucleotides (RNA or preferably DNA) can be used to
20 inhibit DGAT1 production in a bovine if this is considered desirable e.g. in order to produce a bovine capable of improved milk production, i.e. increased milk volume and decreased milkfat content. Antisense oligonucleotides, typically 15 to 20 bases long, bind to the sense mRNA or pre mRNA region coding for the protein of interest, which can inhibit translation of the bound mRNA to protein. The cDNA sequence encoding
25 DGAT1 can thus be used to design a series of oligonucleotides which together span a large portion, or even the entire cDNA sequence. These oligonucleotides can be tested to determine which provides the greatest inhibitory effect on the expression of the protein (Stewart 1996). The most suitable mRNA target sites include 5'- and 3'-untranslated regions as well as the initiation codon. Other regions might be found to be
30 more or less effective.

Alternatively, an antisense nucleic acid or oligonucleotide may bind to DGAT1 coding sequences.

- 35 In yet another embodiment, the invention provides a method of modulating milk production and/or milk content in a lactating bovine, including administering to the bovine an effective amount of a nucleic acid molecule having ribozyme activity and a

nucleotide sequence substantially complementary to at least a portion of mRNA encoding a bovine DGAT1 and being of sufficient length to bind selectively thereto to sufficiently reduce expression of said DGAT1.

- 5 Rather than reducing DGAT1 activity in the bovine by inhibiting gene expression at the nucleic acid level, activity of the relevant DGAT1 protein may be directly inhibited by binding to an agent, such as, for example, a suitable small molecule or a monoclonal antibody.
- 10 Thus, the invention also includes a method of inhibiting the activity of bovine DGAT1 in a lactating bovine so as to modulate milk production and/or milk solids content, comprising administering an effective amount of an antibody to the relevant DGAT1.

The invention still further includes a method of modulating milk production and/or
15 milk solids content by raising an autoantibody to a bovine DGAT1 in the bovine. Raising the autoantibody can include administering a protein having DGAT1 activity to the bovine.

In still a further embodiment, nucleic acids which encode DGAT1 proteins can be used
20 to generate transgenic animals. A transgenic animal (eg. a mouse) is an animal having cells that contain a transgene, which transgene is introduced into the animal or an ancestor of the animal at a prenatal, eg. an embryonic stage. A transgene is DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, a bovine cDNA, comprising the nucleotide sequence shown in Figure 2b,
25 or an appropriate variant or subsequence thereof, can be used to generate transgenic animals that contain cells which express the relevant DGAT1. Likewise, variants can be used to generate transgenic animals. "Knock out" animals can also be generated.

Methods for generating transgenic animals, particularly animals such as mice, have
30 become conventional in the art are described, for example, in US Patent Nos. 4,736,866 and 4,870,009. In such methods, plasmids containing recombinant molecules are microinjected into mouse embryos. In particular, the plasmids can be microinjected into the male pronuclei of fertilised one-cell mouse eggs; the injected eggs transferred to pseudo-pregnant foster females; and the eggs in the foster females allowed to develop to
35 term. (Hogan, 1986). Alternatively, an embryonal stem cell can be transfected with an expression vector comprising nucleic acid encoding a DGAT1 protein, and cells containing the nucleic acid can be used to form aggregation chimeras with embryos

from a suitable recipient mouse strain. The chimeric embryos can then be implanted into a suitable pseudopregnant female mouse of the appropriate strain and the embryo brought to term. Progeny harbouring the transfected DNA in their germ cells can be used to breed uniformly transgenic mice.

5

Such animals could be used to determine whether a sequence related to an intact DGAT1 gene retains biological activity of the encoded DGAT1. Thus, for example, mice in which the murine DGAT1 gene has been knocked out and containing the nucleic acid sequence identified in Figure 2b or fragment or variant thereof could be generated. The animals could be examined with reference to milk production and content.

10

The pattern and extent of expression of a recombinant molecule of the invention in a transgenic mouse is facilitated by fusing a reporter gene to the recombinant molecule such that both genes are co-transcribed to form a polycistronic mRNA. The reporter gene can be introduced into the recombinant molecule using conventional methods such as those described in Sambrook *et al.*, (Sambrook, 1989). Efficient expression of both cistrons of the polycistronic mRNA encoding the protein of the invention and the reporter protein can be achieved by inclusion of a known internal translational initiation sequence such as that present in poliovirus mRNA. The reported gene should be under the control of the regulatory sequence of the recombinant molecule of the invention and the pattern and extent of expression of the gene encoding a protein of the invention can accordingly be determined by assaying for the phenotype of the reporter gene. Preferably the reporter gene codes for a phenotype not displayed by the host cell and the phenotype can be assayed quantitatively. Examples of suitable reporter genes include lacZ (β -galactosidase), neo (neomycin phosphotransferase), CAT (chloramphenicol acetyltransferase) dhfr (dihydrofolate reductase), aphIV (hygromycin phosphotransferase), lux (luciferase), uidA (β -glucuronidase). Preferably, the reporter gene is lacZ which codes for β -galactosidase. β -galactosidase can be assayed using the lactose analogue X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside) which is broken down by β -galactosidase to a product that is blue in colour.

15

20

25

30

35

Still further transgenic applications of the invention arise from knocking out the endogenous gene encoding DGAT1 in non-human mammals and replacing this with a bovine transgene, in order to obtain a desired effect. This is particularly true in cattle raised for milk production. For example, additional copies of the bovine gene encoding DGAT1 can be inserted as a transgene, or the endogenous gene associated with a high level expression promoter in a transgene. It may also prove advantageous to substitute

a defective gene rather than delete the entire sequence of DNA encoding for a protein having DGAT1 activity. A method of producing a transgenic bovine or transgenic bovine embryo is described in United States Patent No. 5,633,076, issued 27 May 1997, for example.

5

These transgenic animals of the invention can again be used to investigate the molecular basis of DGAT1 action. For example, it is expected that mutants in which one or more of the conserved cysteine residues has been deleted would have diminished activity in relation to a DGAT1 protein in which all such residues are retained. Further,
10 deletion of a proteolytic cleavage site would likely result in a mutant lacking biological activity of DGAT1.

Transgenic animals of the invention can also be used to test substances for the ability to prevent, slow or enhance DGAT1 activity. A transgenic animal can be treated with
15 the substance in parallel with an untreated control transgenic animal. Substances which could be tested in this way include proteins extracted from foods ingested by the animal. For example, proteins extracted from pastoral grasses and other fodder can be tested to determine their effect on DGAT1 activity, including to determine whether breed-specific effects can be induced.

20

Thus, in further aspects, the invention provides transgenic non-human animals. These include by way of example only a transgenic bovine having a genome lacking a gene encoding a protein having biological activity of DGAT1 (or indeed any DGAT1 activity at all); a transgenic mouse having a genome containing a gene encoding a bovine protein
25 having biological activity of any DGAT1; and a transgenic bovine having a gene encoding a bovine protein having biological activity of a bovine DGAT1 and heterologous nucleotide sequence antisense to the gene. The transgenic bovine can include a gene encoding a nucleic acid sequence having ribozyme activity and in transcriptional association with the nucleotide sequence antisense to the gene.

30

The invention further provides a transgenic bovine having a genome which includes additional copies of a gene encoding a protein having biological activity of DGAT1 or copies of a gene encoding a protein having biological activity of DGAT1 under control of a high expression promoter.

35

These are but a selection of the applications of this invention. Others will be apparent to those persons skilled in this art and are in no way excluded. To the contrary, the

invention extends to cover not only the specific teaching provided but also all variations and modifications which are within the skill and contemplation of the addressee.

The invention will now be defined by specific examples which are illustrative only and are not intended to limit the invention in any way.

EXPERIMENTAL

1. Location of the gene responsible for the observed QTL

Construction of a BAC contig spanning the BULGE9-BULGE30 interval.

In order to clone the gene(s) responsible for the observed QTL effect, a BAC contig spanning the corresponding marker interval was constructed. This was accomplished by screening a BAC library by filter hybridisation with the microsatellite markers available for proximal BTA14q, as well as with human cDNA clones mapping to the orthologous chromosome segment on the human RH transcript map: 8q23.3-ter (Riquet *et al.*, (1999)). The ends of the isolated BACs were sequenced, sequence tagged sites (STS) developed from the corresponding sequences, and mapped onto a bovine x hamster whole genome radiation hybrid panel. This STS content mapping approach lead to the construction of the BAC contig shown in Figure 1.

DGAT1 maps to the BULGE9-BULGE30 interval and is a strong positional candidate for the QTL.

A murine gene encoding a protein with Diacylglycerol-o-acyltransferase (DGAT1) activity was identified (Cases *et al.*, (1998)) and shown to completely inhibit lactation when knocked out in the mouse (Smith *et al.*, (2000)). This gene was reported in the human to map to HSA8qter (Cases *et al.*, (1998)), ie. in the region orthologous to that containing the bovine QTL. Screening the publicly available databases with the published murine and human DGAT1 cDNA sequences allowed identification of (i) a human BAC clone containing the human DGAT1 gene (AF205589), and (ii) three bovine Expressed Sequence Tags (AW446908; AW446985; AW652329) jointly covering approximately two thirds of the bovine gene. Aligning the human DGAT1 genomic sequences with the human and bovine cDNA sequences allowed the corresponding intron-exon boundaries to be identified. Primers were developed to PCR amplify a portion of the bovine DGAT1 gene. Screening the BACs composing the BULGE9-BULGE30 contig clearly indicated that the bovine DGAT1 gene was contained in a

subset of the BACs allowing us to accurately position the DGAT1 gene in the contig of Figure 1.

These results demonstrated that the map position of DGAT1 coincided with the most likely position of the chromosome 14 QTL as determined by linkage and linkage disequilibrium analyses. Knowing that the QTL primarily affects fat content, knowing the enzymatic activity of DGAT1 and the effect of a DGAT1 knock-out on lactation, this gene was considered to be a very strong positional candidate for the corresponding QTL.

10 Organisation of the bovine DGAT1 gene

The organisation of the bovine DGAT1 gene was determined by sequence analysis of one of the DGAT1 containing BACs. Primers were designed based on the available bovine, murine and human cDNA sequences which were either used for direct sequencing of the BAC clone or to generate PCR products corresponding to different parts of the bovine DGAT1 gene from this BAC which were then subjected to cycle-sequencing. All available sequences were then merged using the Phred / Phrap software (Ewing *et al.*, (1998); Ewing & Green, (1998); Gordon *et al.*, (1998)) to yield the consensus sequence shown in Figures 2a and b.

RT-PCR, 5' and 3' RACE experiments were performed on mRNA isolated from bovine mammary gland and the obtained PCR products subjected to cycle sequencing. Comparison of the genomic and cDNA sequences showed that the bovine *DGAT1* gene spans 8.6 Kb and comprises 17 exons measuring 121.8 bp on average (range: 42 – 436 bp) and allowed intron-exon boundaries to be identified (Figures 2a, 2b and 3). The cDNA sequence is also set out in SEQ ID NO: 4. While the first two introns are respectively 3.6 and 1.9 Kb long, the remaining 14 introns are only 92.4 bp long on average (range: 70 – 215 bp). All introns conform to the GT-AG rule and are strictly conserved between human and bovine. The bovine *DGAT1* gene is transcribed in a mRNA comprising >31 bp of 5' UTR sequence (Figure 2a), 1470 bp coding for a protein of 489 amino-acids, and 275 bp of 3' UTR sequence including a canonical AATAAA polyadenylation signal. The human and bovine *DGAT1* nucleotide (coding) and protein sequences are respectively 89.5% and 92.5% identical (Figures 2a, 2b, 4a and 4b). In addition, an alternative splicing variant is predicted in the bovine for exon VIII (Figure 2b). The corresponding bovine cDNAs are predicted to encode proteins comprising respectively 489 and 467 (alternative splicing variant) amino-acid residues (Figures 4a and 4b).

The predicted "Q" and "q" QTL alleles differ by a non conservative lysine to alanine amino-acid substitution in the DGAT1 gene.

Assuming that DGAT1 is indeed the QTL, it is predicted that the identified "Q" and "q" QTL alleles will correspond to functionally distinct DGAT1 alleles, ie. will differ at one or more mutations causing these alleles to be functionally different. To test this hypothesis, the structure of the DGAT1 gene in individuals predicted to be of different QTL genotypes: "QQ", "Qq" and "qq" was examined. More specifically, the DGAT1 gene from:

- 10 (i) two sires with " H^{Q-D}/h^q " genotype as well as two of their " H^{Q-D}/H^{Q-D} " offspring, two of their " h^q/h^q " offspring and one " H^{Q-D}/h^q " offspring, and
- (ii) one " H^{Q-NZ}/h^q " sire with one of its " H^{Q-NZ}/H^{Q-NZ} " offspring

15 was analysed wherein H^{Q-D} corresponds to the Dutch Q haplotype and H^{Q-NZ} corresponds to the New Zealand Q haplotype, and primer pairs were designed that allowed for the amplification from genomic DNA of (i) the coding portion of exon I, (ii) exon II, and (iii) the chromosome regions spanning exons III to XVII. The corresponding PCR products from the selected individuals were cycle-sequenced and the resulting
20 sequences examined with the *Polyphred* software.

Additional sequencing analysis, as described above, on DNA from a range of breeds revealed additional polymorphisms included in Table 1 (see Methods section for breeds). Four such polymorphisms were investigated further:

25

(i) K232A: a substitution of a *ApA* by a *GpC* dinucleotide in exon VIII (respectively positions 694 and 695 counting from the start codon in the cDNA). The substitution of these two adjacent nucleotides results in a non conservative lysine (hydrophylic basic amino acid) to alanine (hydrophobic amino acid) substitution in the DGAT1 protein. The
30 *lysine* residue affected by this polymorphism is conserved in the human and murine DGAT1 sequences. Together with the resulting change in the electrical charge of the protein, this strongly suggests that this amino-acid substitution is likely to result in a functional difference between the two corresponding alleles and to be at least partly responsible for the observed QTL effect.

35

(ii) Nt984+8(Base 7438 A-G): A A to G substitution in intron 12, eight base pairs downstream of exon XII. Following standard nomenclature, this polymorphism will be

referred to as *Nt984+8(A-G)*. This polymorphism cannot be predicted as such to modify the functionality of the corresponding alleles although an effect on the splicing mechanism cannot be excluded given its proximity to the intron-exon boundary.

5 ***(iii) Nt984+26(Base 7456 C-T):*** A C to T substitution in intron 12, 26 base pairs downstream of exon XII. Following standard nomenclature, this polymorphism will be referred to as *Nt984+26(Base 7456 C-T)*. Again, this polymorphism cannot be predicted as such to modify the functionality of the corresponding alleles although an effect on the splicing mechanism cannot be excluded given its proximity to the intron-exon
10 boundary.

(iv) Nt1470+85(Base 8402 C-T): A C to T substitution in the 3' UTR. Following standard nomenclature, this polymorphism will be referred to as *Nt1470+85(Base 8402 C-T)*. Again, this polymorphism cannot be predicted as such to modify the functionality
15 of the corresponding alleles although an effect on polyadenylation or mRNA stability cannot be excluded.

Conclusion

20 These four polymorphisms were shown to assort into three distinct *SNP* haplotypes referred to as *sH^{Q-D}*, *sH^{Q-NZ}* and *sh^g* because in the sequenced samples they coincided respectively with *microsatellite* haplotypes *μH^{Q-D}*, *μH^{Q-NZ}* and *μh^g*. The base pair compositions of these three SNP haplotypes are shown in Figure 3.

25 Because the *sH^{Q-NZ}* and *sh^g* marker haplotypes share the G residue at the *DGAT1* *Nt984+8(Base 7438 A-G)* site, the causality of this polymorphism in the determinism of the QTL could be excluded. For the three remaining polymorphic sites, however, the *DGAT1* haplotypes associated with marker haplotypes *sH^{Q-D}* and *sH^{Q-NZ}* proved identical to each other while different from the *sh^g* *DGAT1* haplotype. Either of these three
30 polymorphisms could therefore be responsible for the observed QTL effect. The *Nt984+26(Base 7456 C-T)* and *Nt1470+85(Base 8402 C-T)* polymorphisms are *a priori* more likely to be neutral with respect to *DGAT1* activity because of their respective location in an intron and the 3' UTR and likewise the other non coding or neutral polymorphism shown in Table 1. A direct effect of the *K232A* mutation on *DGAT1*
35 activity, however, is very plausible. Indeed, the corresponding lysine residue is conserved amongst all examined mammals (i.e. human, mouse, rat, pig, sheep, bison) demonstrating its functional importance (Figure 5). The evolutionary conservation of

this lysine residue also demonstrates that the *K* residue characterizing the *sH^{Q-D}* and *sH^{Q-NZ}* marker haplotypes is more than likely the ancestral state and that it is the *A* residue characterizing the *sh^a* haplotypes that corresponds to a more recently evolved state.

5

2. Genotype Testing and Analysis I

This summarises the genotype testing and subsequent analysis of Holstein-Friesian animals sourced from New Zealand and Holland which were tested for the presence of the K232A polymorphism. Reference to allele "Q" corresponds to the *K* residue and allele "q" to the *A* residue (as shown in Figure 3 and Table 1).

An oligonucleotide ligation assay (OLA) was developed as described in the method section below that allows for efficient genotyping of the four *DGAT1* polymorphisms simultaneously. This OLA-test was used to genotype a previously described (Farnir et al., 2000) "grand-daughter design" (i.e. series of paternal half-brother pedigrees) comprising 1,818 Dutch Holstein-Friesian sires as well as a "daughter design" (i.e. series of paternal half-sister pedigrees) comprising 529 New Zealand Holstein-Friesian cows selected according to phenotype as described below. The marker linkage phase for each individual was determined as described below.

Fig. 6 summarizes the frequency distribution of *DGAT1* haplotypes encountered in the Dutch and New Zealand populations respectively. Four distinct *SNP* haplotypes were identified. Three of these correspond to the *sH^{Q-D}*, *sH^{Q-NZ}* and *sh^a* that were previously identified by sequencing, and jointly account for 99% and 98% of the chromosomes in the Dutch and New-Zealand populations respectively. A fourth minor haplotype was found accounting for the remaining 1% and 2% of the chromosomes. As this haplotype codes for a *K* residue at position 232 it was assumed to correspond to a fat increasing "Q" allele and was therefore referred to as *sH^{Q-III}* (Fig. 3). The observation that the *K* residue is found on three distinct *DGAT1* haplotypes while the *A* residue is found on a unique *DGAT1* haplotype is in agreement with *K* being the more ancient state.

The *sH^{Q-D}* and *sH^{Q-NZ}* *SNP* haplotypes (coding for a *K* residue at position 232) appear to be in strong linkage disequilibrium (LD) with the flanking microsatellite markers *BULGE09* and *BULGE11*, as they are in essence associated with unique microsatellite haplotypes corresponding respectively to the previously defined *μH^{Q-D}* and *μH^{Q-NZ}* haplotypes (Fig. 6C&D). In sharp contrast, the *sh^a* haplotype (coding for an *A* residue at

35

position 232) is nearly evenly distributed across more than ten distinct microsatellite haplotypes (Fig. 6B).

These observations are in excellent agreement with the results of the combined linkage
 5 and LD analysis (Fernier et. al., 2000). These studies indeed predicted (i) that in the Dutch population the vast majority (estimates ranging from 81% to 92%) of “Q” allele (= K) would reside on the μH^{Q-D} microsatellite haplotype, (ii) that in the New Zealand population a large fraction (estimates ranging from 36% to 51%) of “Q” alleles would reside on haplotype μH^{Q-NZ} (we now see that the remainder correspond mainly to the
 10 μH^{Q-D} microsatellite haplotype) and (iii) that in both populations the “q” alleles (= A) would correspond to multiple marker haplotypes, corresponding to h^q .

Figure 7 illustrates the gain in LD signal that could be obtained in the Dutch Holstein-Friesian grand-daughter design when adding the *DGAT1* polymorphisms to the
 15 previously available markers for proximal BTA14q and performing a joint linkage and LD multipoint analysis (Fernier et. al., 2000) using the sires “daughter yield deviations” (DYD (Van Raden and Wiggans, 1991) corresponding to half breeding values) for milk fat percentage as phenotype. It can be seen that the lod score attributable to LD essentially doubles (from 3.7 to 7.8), and maximizes exactly at the position of the
 20 *DGAT1* gene. This result strongly supports the causal involvement of the *DGAT1* gene in the QTL effect. The corresponding ML estimates of the “Q” to “q” allele substitution effect ($\alpha/2$) (as defined in Falconer and Mackay, 1996), residual standard deviation (σ), population frequency of the “Q” allele (f_Q), number of generations to coalescence (g) and heterogeneity parameter (ρ) were respectively 0.11% ($\alpha/2$), 0.06% (σ), 0.20 (f_Q), 5 (g) and
 25 0.84 (ρ).

Using the same Dutch Holstein-Friesian population, the additive effect of the *DGAT1* K232A polymorphism on milk yield and composition was examined. The sons DYDs for milk yield (kgs), protein yield (kgs), fat yield (kgs), protein percentage and fat
 30 percentage, were analysed using a mixed model including (i) a regression on the number of K alleles in the genotype (0, 1 or 2), and (ii) a random polygenic component estimated using an individual animal model and accounting for all known pedigree relationships. Table 2 below, reports the obtained results. It can be seen that the K232A mutation has an extremely significant effect on the five analysed dairy traits.
 35 The proportion of the trait variance explained by this polymorphism in this population ranges from 8% (protein yield) to 51% (fat percentage), corresponding to between 10% (protein yield) and 64% (fat percentage) of the genetic variance (= QTL + polygenic).

Note that the proportion of the variance explained by the full model ($1-r^2_{\text{error}}$) is of the order of 70% for the yield traits and 80% for the percentage traits, which is in agreement with the known reliabilities of the corresponding DYDs (Van Raden and Wiggans, 1991). An interesting feature of this QTL effect is that the “q” to “Q” substitution increases fat yield, while decreasing milk and protein yield, despite the strong overall positive correlation characterizing the three yield traits.

Table 2:

Effect of the *DGAT1* K232A mutation on sire's daughter yield deviations (DYDs) for milk yield and composition.

Trait	$\alpha/2$	r^2_{QTL}	p-value _{QTL}	$r^2_{\text{polygenic}}$	r^2_{error}
Milk yield (Kgs)	-158 Kgs	0.18	5.00E-35	0.49	0.32
Fat yield (Kgs)	5.23 Kgs	0.15	1.57E-29	0.55	0.30
Protein yield (Kgs)	-2.82 Kgs	0.08	1.70E-15	0.65	0.26
Fat %	0.17 %	0.51	4.33E-122	0.29	0.19
Protein %	0.04 %	0.14	5.05E-28	0.66	0.20

(i) $\alpha/2$: QTL allele substitution effect on DYD (half breeding value), corresponding in the mixed model to the regression coefficient on the number of K alleles in the *DGAT1* K232A genotype, and to $\alpha/2$, where α is defined according to ref. Falconer and Mackay, 1996. (ii) r^2_{QTL} : proportion of the trait variance explained by the *DGAT1* K232A polymorphism. (iii) p-value_{QTL}: statistical significance of the *DGAT1* K232A effect. (iv) $r^2_{\text{polygenic}}$: proportion of the trait variance explained by the random, polygenic effect in the mixed model. (v) r^2_{error} : proportion of the trait variance unexplained by the model.

The two previous analyses examined the effect of the *DGAT1* polymorphism on estimated breeding values. By definition, this phenotype will only account for the additive component of the *DGAT1* effect, and justifies the use of a regression on the number of K alleles in the mixed model. To evaluate the dominance relationship between the *DGAT1* alleles, the effect of the K232A genotype on the lactation values (first yield deviations (Van Raden and Wiggans, 1991)) of the cows composing the New Zealand daughter design were analysed. This was achieved by using a mixed model including (i) a fixed effect corresponding to the K232A genotype, and (ii) a random polygenic component accounting for all known pedigree relationships (“animal model”). Very significant effects of K232A genotype on all examined yield and composition traits were found in this population as well (Table 3, below), accounting for between 1% (protein yield) and 31% (fat percentage) of the trait variance. The observed dominance deviations, d , corresponding to the difference between the genotypic value of the KA

genotype and the midpoint between the *AA* and *KK* genotypic values (Falconer and Mackey, 1996) are shown in Table 3 below. Genotypic values of the heterozygous genotype are systematically in between alternate homozygotes. None of the *d*-values proved to be significantly different from zero, indicating an absence of dominance.

5 Average *K* to *A* QTL allele substitution effects, α (Falconer and Mackey, 1996), were computed from the estimates of *a*- and *d*-values, as well as the population frequencies of the *K* and *A* alleles (Table 3). The predicted substitution effects are in good agreement with those computed from the grand-daughter design: the *K* allele increases fat yield, fat % and protein %, while decreasing milk and protein yield. The absolute
10 values of α estimated from the grand-daughter and daughter design are in perfect agreement for fat and protein %, while for the yield traits estimates are larger in the grand-daughter design when compared to the daughter design. The exact reasons for this are being explored. It could be due to the fact that the sire population in the grand-daughter design is not representative of the cow population in general, or to
15 intrinsic differences between the Dutch and New-Zealand populations and/or environment.

Table 3:

**Effect of the *DGAT1* K232A mutation on cows' lactation values
for milk yield and composition.**

Trait	a	d	α	r^2_{QTL}	p-val QTL	$r^2_{\text{polygenic}}$	r^2_{error}
Milk yield (Kgs)	-144 Kgs	-42 Kgs	-161 Kgs	0.03	1.05E-8	0.54	0.43
Fat yield (Kgs)	7.82 Kgs	-0.89 Kgs	7.46 Kgs	0.09	1.77E-20	0.46	0.45
Protein yield (Kgs)	-2.34 Kgs	-0.76 Kgs	-2.64 Kgs	0.01	4.35E-2	0.37	0.42
Fat %	0.41 %	0.03 %	0.42%	0.31	2.5E-108	0.49	0.20
Protein %	0.08%	0.03 %	0.08%	0.04	1.60E-20	0.72	0.24

(i) **a**: half the difference between the genotypic values of the *KK* and *AA* genotypes (Falconer and Mackey, 1996). (ii) **d**: dominance deviation (Falconer and Mackey, 1996): deviation of the *KA* genotypic value from the midpoint between the *AA* and *KK* genotypic values; none of these proved to be significantly different from zero. (iii) **α** : average *K* to *A* substitution effect, computed as " $a + d(q-p)$ " (Falconer and Mackey, 1996), where *q* is the allelic frequency of *K* (=0.7) and *p* of *A* (=0.3) (iv) **r^2_{QTL}** : proportion of the trait variance explained by the *DGAT1* K232A polymorphism. (v) **p-val QTL**: statistical significance of the *DGAT1* K232A effect. (vi) **$r^2_{\text{polygenic}}$** : proportion of the trait variance explained by the random, polygenic effect in the mixed model. (vii) **r^2_{error}** : proportion of the trait variance unexplained by the model.

Pedigree material and phenotypes. The pedigree material used for the association studies comprised a “grand-daughter design” (Weller et. al., 1990) counting 1,818 Holstein-Friesian bulls sampled in the Netherlands, as well as a “daughter-design” (Weller et. al., 1990) counting 529 Holstein-Friesian cows sampled in New Zealand. The phenotypes of the sires were “daughter yield deviations”(DYD: unregressed weighted averages of the daughter’s lactation performances adjusted for systematic environmental effects and breeding values of the daughter’s dams and expressed as deviations from the population mean (Van Raden and Wiggans, 1991)) obtained directly from CR-Delta (Arnhem – The Netherlands). The phenotypes of the cows were “lactation values” (first lactation yield deviations (YD), i.e. weighted average lactation performances expressed as deviations from the population mean, adjusted for management group, permanent environmental effects and herd-sire interaction effects (Van Raden and Wiggans, 1991)) obtained directly from Livestock Improvement Corporation (Hamilton – New Zealand).

Combined linkage and linkage disequilibrium analysis and association studies. The maximum likelihood procedure for combined linkage and linkage disequilibrium analysis is described in detail in Farnir, 2000. The association study in the grand-daughter design was performed using the following model:

$$y_i = \mu + \beta x_i + a_i + e_i$$

where y_i is the DYD of son i , μ is the overall population mean, β is a fixed regression coefficient estimating the A to K allele substitution effect, x_i is an indicator variable corresponding to the number of K alleles in the $K232A$ genotype, a_i is a random polygenic component accounting for all known pedigree relationships (“animal model” Lynch and Walsh, 1997) and e_i is a random residual. The association study in the daughter design was performed using the model:

$$y_i = \mu + g_i + a_i + e_i$$

where y_i is the lactation value of cow i , g_i is a fixed effect corresponding to the $DGAT1$ genotype (KK , KA , or AA), a_i is a random polygenic component accounting for all known pedigree relationships (“animal model” Lynch and Walsh, 1997) and e_i is a random residual. In both instances, maximum likelihood solutions for β , g_i , a_i , e_i , σ^2_a , σ^2_e were obtained using the MTDFREML program (Boldman et al, 1997).

3. Genotype Testing and Analysis II

This summarises the genotype testing and subsequent analysis of Holstein-Friesian,
 5 Jersey and Ayrshire animals in a separate population from those presented in genotype testing and analysis I, above.

Progeny tested sires

Each year Livestock Improvement Corporation (New Zealand) progeny test some 200-
 10 300 bulls per year. This entails the bulls being genetically evaluated on the basis of 50-85 daughters per sire. The sires are evaluated for milk fat, milk protein, milk volume and 20 non-production traits. Semen has been retained from all progeny tested sires since the early 1970s. DNA was extracted from the semen and genotyped for the K232A DGAT1 polymorphism using the 7900 Taqman system (see Methods section below).

15 Statistical analysis was undertaken on this dataset using Restricted Maximum Likelihood (REML) and the average information algorithm (Johnson and Thompson, 1995). The linear model included the fixed effects of DGAT1 (3 classes; 0, 1 and 2 copies of the Q allele i.e. the K residue) and a covariate corresponding to the proportion
 20 of overseas genetics. The random effect was animal with a relationship matrix based on all known relationships. Daughter yield deviations (DYDs), weighted averages of a sire's daughter's lactation performances expressed as deviations from the population mean (van Raden and Wiggans 1991) were used as the phenotypic measurement. The phenotypes were weighted by a weighting factor based on the variance of the DYD for a
 25 son being:

$$\text{Var DYD} = \left[\frac{1 + (n-1) \frac{1}{4} h^2}{n} \right] \sigma_p^2$$

where Var DYD is the variance of son's DYD; n is the number of daughters contributing
 30 to the DYD; h^2 is the heritability, which was taken as 0.35 for yield traits.

The dataset was analysed separately for the 3 major breeds; Holstein-Friesian, Jersey and Ayrshire.

Seventeen hundred and thirteen Holstein-Friesian sires were included in the analysis. The effect of the DGAT1 polymorphism was extremely significant for the three milk production traits (Table 4). With each additional Q allele the level of milk fat production increases by approximately 6 kg per lactation, milk protein production decreases by approximately 2.5 kg per lactation and milk volume decreases by approximately 125 litres per lactation.

Table 4:

Effect of the DGAT1 polymorphism on milk production in the Holstein-Friesian bull population (kilograms per lactation).

	Fat	Protein	Milk
qq	0	0	0
Qq	6.86	-2.13	-128
QQ	11.83	-4.80	-266
st. error	0.87	0.68	24

The effects for the Jersey and Ayrshire breeds were less significant than those of the Holstein-Friesian breed but were consistent in direction of effects.

Daughters for milk components

Data collection was integrated with LIC's herd testing service using a sample of 102 herds involved in LIC's Sire Proving Scheme (SPS) in 1995. In addition to milk volume from herd testing, the concentrations of fat, crude protein (total nitrogen), casein, whey and lactose were determined. The data was collected from over 3,000 cows born in 1996 and first calving in the 1998 spring season, these being predominantly the daughters of approximately 220 SPS bulls. The milk characteristics were measured at three herd tests on each cow, with each herd having a herd test in each of the Sept/Oct, Nov/Dec and Jan/Feb periods. The Milkoscan FT120, which employs Fourier transform infrared spectrophotometry with enhanced milk calibrations (Foss Electric Application Note Nos. 95, P/N 492280 and 102, P/N 578377), was used to determine the milk component concentrations.

Nine hundred and twelve daughters were genotyped for the DGAT1 polymorphism using the OLA system. Analysis was undertaken using SAS (Statistics, Version 5, 1985) fitting a general linear model. The model included sire and maternal grandsire as fixed effects, DGAT1 polymorphism (3 classes; 0, 1 and 2 copies of the Q allele), covariates including 16ths of Holstein-Friesian, Jersey, Ayrshire and other, proportion of overseas

genetics within the Holstein-Friesian, Jersey and Ayrshire breeds. Yield deviations that were pre-adjusted for herd, stage of lactation among other fixed effects were used (Johnson et al 2000).

- 5 The DGAT1 polymorphism is statistically significant for Lactose, casein, beta-casein and whey yield and also for casein and beta-casein percent as outlined in Table 5.

Table 5:
Effect of the DGAT1 polymorphism on milk components.

10

Trait	qq	Qq	QQ	p-value
Lactose yield*	48	23	0	<0.0001
Casein yield*	11.0	5.8	0	0.01
Casein %	-0.13	-0.06	0	<0.0001
Whey yield*	6.86	2.31	0	<0.0001
β -casein yield*	3.98	2.19	0	0.05
β -casein %	-0.43	-0.23	0	0.0001

* Units = g/day for lactose, casein and whey yield and g/litre for β -casein yield

Daughters for solid fat content

- 15 Six hundred and ninety-two daughters were phenotyped for solid fat content. Solid fat content of the milkfat is a characteristic which has a major influence on the functionality of milkfat products, and in particular has a significant effect on the hardness of butter (MacGibbon & McLennan, 1987). The solid fat content at 10°C (SFC 10) was used for comparison of the properties of the milkfat as it relates well to the
- 20 sectility hardness measurement of butter, a major functional property. Thus the performance of milkfat products may be predicted from the characteristics of the milk produced. The solid fat content (SFC) of the extracted fat was determined by pulsed nuclear magnetic resonance (NMR) and expressed as percentage solid fat (MacGibbon & McLennan, 1987). As the milkfat was melted to remove any thermal history, prior to
- 25 recrystallization under standard conditions, the SFC simply reflects the chemical composition of the milkfat.

The 692 daughters were a subset of the 912 daughters that were phenotyped and genotyped for the results presented in Table 2. The solid fat content measures were

collected over 2 lactations. Breeding values were calculated using an animal model similar to that of Johnson et al 2000.

5 The same statistical model was fitted for solid fat content as was for the milk component analysis. The DGAT1 polymorphism has a statistically significant effect (p-value <0.0001) on solid fat content, increasing it by approximately by 1% for each addition of the Q allele.

10 This effect was further confirmed in 50 daughters (predominantly Holstein-Friesian) that were farmed at one location and measured for SFC on the same day. The estimated effect for of the DGAT1 polymorphism on SFC was to increase it by approximately 2% per addition of each Q allele. This finding was significant at the five percent threshold level.

15 The genetic standard deviation for SFC is 2.25 (D Johnson personal communication) and thus the effect of DGAT1 is approximately 0.5 of a genetic standard deviation.

4. Relative Transcript Levels of the Splice Variant

20 Real time PCR experiments were conducted using reverse transcribed mRNA isolated from lactating bovine mammary gland(s) (see experimental methods). These experiments revealed that the alternatively spliced transcript as shown on Figure 2b, was approximately 100 fold less abundant than the full length transcript.

METHODS SECTION

In order to identify other polymorphisms within the bovine DGAT1 gene, DNA was isolated from sperm, PCR amplified and then using primers designed from the sequence shown in Figures 2a and 2b and/or the cDNA sequence (SEQ ID NO: 4) direct sequenced on an ABI 3100. The breeds examined were:

Ayrshire, Angler, Belgian Blue, Blond D'Aquitaine, Brown Swiss, Charolais, Red Devon, Devon, Dexter, Friesian, Guernsey, Belted Galloway, Gelbvieh, Hereford, Jersey, Limousin, Longhorn, Maine Anjou, MRI (Meuse-rhine-yssel), Murray Grey, Piedmontese, Romangola, Sahiwal, Santa Gertrudis, Scottish Highland, Shorthorn, South Devon, Sussex, Swedish Red, Simmental, Wagyu, Welsh Black, Angus, and Zebu.

All the polymorphisms discovered are listed in Table 1, above.

The majority of the primers are also listed in Figure 2b or contained in the cDNA sequence (SEQ ID NO: 4).

Experimental method for the OLA analysis of four SNP's in DGAT1

PCR amplification of the regions containing the polymorphisms

Protocol for the PCR amplification of exon VIII, intron XII and 3' UTR, the regions containing the four polymorphisms that were initially described in the DGAT1 gene.

Component	For 1 sample	Final concentration
HotStar Qiagen Buffer (10 x)	1.5 µl	0.7 µM
Primer 17F at 100µM	0.07 µl	0.7µM
18R at 10 µM	0.07 µl	0.5 µM
Primer 6F at 100µM	0.05µl	0.5 µM
AW 446985dn1 at 100µM	0.05µl	0.5 µM
Primer InsUp1	0.05µl	0.5 µM
14R2	0.05µl	10%
DMSO	1 µl	300 µM
dNTP 10mM	0.3 µl	0.1 U/µl
HotStarQiagen Taq (CatNr 203205:5U/µl)	0.2 µl	
H2O	1.66 µl	
DNA (5 ng/µl)	5 µl	
Total	10 µl	

Primer sequences are given in the following table as well as the genomic region targeted by them.

68348-71

SNP targeted	Primer name	Primer sequence	SEQ ID No:	5 ' base position
Exon VIII SNP (DG 1)	17F	CCTGAGCTTGCCTCTCCCACAGT	48	6579
	18R	CCAGGAGTCGCCGCAGCAGGAAG	49	7058
Exon XII SNPs (DG 2 and DG3)	6F	CCGGCCATCCAGAACTCCATGAAG	50	7280
	AW446985 dn1	TAGAACTCGCGGTCTCCAAAC	51	7605
	InsUp1	TGGCTGTCACTCATCATCGGGCA	52	8222
3'UTR SNP (DG4)	14R2	TTGCACAGCACTTTATTGACACA	53	8566

PCR amplification was performed on MJ PTC100 or PTCT200 cyclers using the following steps:

Step	Temperature	Time	Comment
1° Activation of the enzyme	94 °C	12 minutes	One times
2° Denaturation	92 °C	1 minute	Repeat step 2 to 4, 35 times
3° Hybridisation	60°C	1 minute 30 seconds	
4° Elongation	72°C	1 minute 30 seconds	
5° Inactivation of the enzyme	99°C	45 minutes	

5

Oligonucleotide Ligation Assay (OLA)

The oligonucleotides used in the OLA multiplex reaction are given in the table below. The detection of each mutation relies on the use of two fluorescent-labelled oligonucleotide (SNPx_FAM and SNPx_HEX) and one common 3' and 5' phosphorylated, non-labelled oligonucleotide (SNPx_2P)

10

Locus	Oligo	Sequence	SEQ ID NO:	5' base position	Number of spacer phosphoramidites	Size of the ligation product ^a
DG1	SNP1_FAM	AGC TTT GGC AGG TAA GGC	54	6813	0	32
	SNP1_HEX	AGC TTT GGC AGG TAA GAA	55	6813		
	SNP1_2P	GGC CAA CGG GGG AG	56	6831		
DG2	SNP2_FAM	GCT GGC GGT GAG TGA	57	7424	3	39
	SNP2_HEX	GCT GGC GGT GAG TGG	58	7424		
	SNP2_2P	CCT GCT GGG TGG GGA	59	7439		
DG3	SNP3_FAM	GCT GGG TGG GGA CGC	60	7442	0	29
	SNP3_HEX	GCT GGG TGG GGA CGT	61	7442		
	SNP3_P	GTG GGG GCG GGT GG	62	7457		
DG4	SNP4_FAM	TGC CCC AAC CTG GGT	63	8388	2	36
	SNP4_HEX	TGC CCC AAC CTG GGC	64	8388		
	SNP4_2P	GCA GCA GGA GGA GGC	65	8403		

^a The size of the ligation products is the sum of the number of nucleotides of the two ligated oligonucleotides plus 3 bases equivalents per spacer phosphoramidites molecule, present at the 5' end of the common oligonucleotide.

15

For each SNP a mixture of the three oligonucleotides was prepared first, following the dilution guidelines in the table below.

SNP mixture	Oligonucleotide to mix	Quantity	Final concentration
DG1 (oligo. mixture)	SNP1_FAM 10 μ M SNP1_HEX 10 μ M SNP1_2P 10 μ M H ₂ O	10 μ l 20 μ l 20 μ l 50 μ l	1 μ M 2 μ M 2 μ M
DG2 (oligo. mixture)	SNP2_FAM 10 μ M SNP2_HEX 10 μ M SNP2_2P 10 μ M H ₂ O	10 μ l 20 μ l 20 μ l 50 μ l	1 μ M 2 μ M 2 μ M
DG3 (oligo. mixture)	SNP3_FAM 10 μ M SNP3_HEX 10 μ M SNP3_2P 10 μ M H ₂ O	10 μ l 20 μ l 20 μ l 50 μ l	1 μ M 2 μ M 2 μ M
DG4 (oligo. mixture)	SNP4_FAM 10 μ M SNP4_HEX 10 μ M SNP4_2P 10 μ M H ₂ O	10 μ l 30 μ l 20 μ l 40 μ l	1 μ M 3 μ M 2 μ M

The ligation reaction for one sample was performed as follow:

Component	Quantity persample
DG1 oligonucleotide mixture ^a (35, 70 and 70 nM)	0.7 μ l
DG2 oligonucleotide mixture ^a (12.5, 25 and 25 nM)	0.25 μ l
	0.25 μ l
DG3 oligonucleotide mixture ^a (12.5, 25 and 25 nM)	0.25 μ l
	2 μ l
	2 μ l
DG4 oligonucleotide mixture ^a (12.5, 37.5 and 25 nM)	
DMSO	1 μ l
Incubation buffer of the Tsc DNA ligase (Roche, Cat Nr 1 939 807 or 1 939 815)	8.55 μ l
Tsc DNA ligase	5 μ l
H ₂ O	
Multiplex PCR (see above)	
Total	20 μ l

5

^a The final concentration of the oligonucleotides in the ligation reaction is given between parenthesis (SNPx_FAM, SNP_x_HEX and SNP_x_2P respectively)

10 The sample was submitted to the following temperature cycling program in a MJ PTC100 or PTC 200 PCR machine.

Step	Temperature	Time	Comment
1° Initial denaturation step	98 °C	2 minutes	One times
2° Dénaturation	94 °C	30 seconds	Repeat step 2 to 3, 30 times
3° Hybridisation and ligation	45°C	3 minutes	
5° Inactivation of the enzyme	99°C	45 minutes	

15 Following the LCR, 20 μ l of H₂O was added to the ligation reaction. To 0.5 μ l of the diluted ligation reaction, either 2 μ l of loading buffer was added, or 2 μ l loading buffer containing TAMRA350 internal line size standard.

The loading buffer was composed as follows: 1 part of blue dextran (50mg/ml)/ EDTA (25mM) and 6 parts of formamide

68348-71

37

The TAMRA350 containing loading buffer was composed as follows: 3 parts TAMRA350 (Applied Biosystems 401736 ; 8nM), 10 parts of Blue dextran (50 mg/ml)/EDTA (25 mM) and 60 parts of formamide.

5 TAMRA containing samples was placed alternately with TAMRA free samples when loaded onto the sequencing gel, in order to ease the identification of the lanes on the gel image.

The samples may require further dilution in order to avoid a too intense fluorescent signal on the sequencer. It is also very likely that from one primer batch to another, oligonucleotides concentrations will need adjustment.

10 The samples were denatured for 5 minutes at 95°C before loading. The samples were then loaded onto a 6% denaturing acrylamide gel on sequencer ABI 373 or a 4% gel on sequencer ABI 377.

In addition to the OLA assays referred to above, genotyping of the DGAT1 polymorphism was carried out by utilizing two different techniques for detection of
15 PCR products.

Gel-based Genotyping Assay

Primer sequences 5' to 3', genomic sequence position within SEQ ID NO:1 in brackets:

DGAT1 21: GTAGCTTTGGCAGGTAAGAA (6811) (SEQ ID NO:66)

20 DGAT1 22: GGGGCGAAGAGGAAGTAGTA (6984) (SEQ ID NO:67)

DGAT1 23: TGGCCCTGATGGTCTACACC (6613) (SEQ ID NO:68)

DGAT1 24B: GGGCAGCTCCCCCGTTGGCCGC (6850) (SEQ ID NO:69)

The final reaction conditions were 1X Gold PCR buffer, 2.5mM MgCl₂ (Applied Biosystems), 200µM each dNTP (Roche), 600nM DGAT1 21 and DGAT1 22, 400nM
25 DGAT1 23 and DGAT1 24B (Invitrogen), 10% dimethylsulphoxide (Sigma), 3µl DNA template and 2.5 units AmpliTaq Gold DNA polymerase (Applied Biosystems) in a total volume of 50µl.

68348-71

38

Cycling conditions were a 94°C initial denaturation for 5 minutes, then 35 cycles of denaturation at 94°C for 30 seconds, annealing at 56°C for 30 seconds, extension at 72°C for 20 seconds followed by one extension cycle of 72°C for 2 minutes.

Primer positions around polymorphism (in bold) on genomic sequence from 6587 to 6986.

```

                                DGAT123    TGGC  CCTGATGGTC TACACC
                                                →
TGCCTCTCCC ACAGTGGGCT CCGTGCTGGC CCTGATGGTC TACACCATCC
TCTTCCTCAA GCTGTTCTCC TACCGGGACG TCAACCTCTG GTGCCGAGAG
CGCAGGGGCTG GGGCCAAGGC CAAGGCTGGT GAGGGCTGCC TCGGGCTGGG
10  GCCACTGGGC TGCCACTTGC CTCGGGACCG GCAGGGGCTC GGCTCACCCC
                                DGAT1 21 GTAGCT TTGGCAGGTA AGAA
                                                →
CGACCCGCCC CCTGCCGCTT GCTCGTAGCT TTGGCAGGTA AGAAAGGCCAA
                                ←
                                CGCCGGTT
CGGGGGGAGCT GCCCAGCGCA CCGTGAGCTA CCCCAGACAAC CTGACCTACC
15  GCCCCCTCGA CGGG DGAT1 24B
GCGGTGAGGA TCCTGCCGGG GGCTGGGGGG ACTGCCCGGC GGCCTGGCCT
GCTAGCCCCG CCCTCCCTTC CAGATCTCTA CTACTTCCTC TTCGCCCCCA
(SEQ ID NO:70)

```

```

                                ← AT GATGAAGGAG AAGCGGGG DGAT1 22
20  The Q allele has polymorphic sequence AA and is detected by the DGAT1 21 + 22
    primers, producing a band of 174bp. The q allele has polymorphic sequence GC and
    is detected by the DGAT123 + 24 primers, producing a band of 238bp.

```

The primers DGAT123 and DGAT122 also successfully PCR the DGAT1 gene producing a product of 372bp in all reactions. Therefore, a QQ homozygote would have bands at 372bp and 174bp, a qq homozygote would have bands at 372bp and 238bp and a Qq heterozygote would have all 3 bands at 372bp, 238bp and 174bp.

18µl of PCR product was separated on a 1.2% agarose TAE gel, stained with ethidium bromide and scored independently by two investigators on the basis of the number and size of bands present.

68348-71

39

TaqMan Allelic Discrimination Genotyping Assay

Primer sequences 5' to 3', genomic sequence position in brackets:

DGAT1forAD: TTCTCCTACCGGGACGTCAA (6651) (SEQ ID NO:71)

ReverseNZ: CCGCGGTAGGTCAGGTTGTC (6890) (SEQ ID NO:72)

5 Probe sequences 5' to 3', genomic sequence position in brackets:

ForAA (FAM): CGTTGGCCTTCTTA (6838) (SEQ ID NO:73)

DGAT1ADGC (VIC): TTGGCCGCCTTACC (6836) (SEQ ID NO:74)

Both probes use MGB (minor groove binder) as a non-fluorescent quencher.

10 The final reaction conditions are 1x Universal PCR Mastermix (Applied Biosystems),
500nM each primer (Invitrogen), 70nM ForAA (FAM) probe, 300nM DGAT1ADGC
(VIC) probe (Applied Biosystems) and 2µl of a 1/20 dilution of DNA template in a total
volume of 10µl.

15 Cycling conditions were 50°C for 2 minutes, 95°C initial denaturation for 10 minutes,
then 37 cycles of denaturation at 94°C for 15 seconds, annealing and extension 60°C
for 1 minute.

Primer positions around polymorphism (in bold) on genomic sequence from 6587 to
6986 of SEQ ID NO:1.

TGCCTCTCCC ACAGTGGGCT CCGTGCTGGC CCTGATGGTC TACACCATCC

DGAT1forAD TTCTCC TACCGGGACG TCAA →

20 TCTTCCTCAA GCTGTTCTCC TACCGGGACG TCAACCTCTG GTGCCGAGAG

CGCAGGGCTG GGGCCAAGGC CAAGGCTGGT GAGGGCTGCC TCGGGCTGGG

GCCACTGGGC TGCCACTTGC CTCGGGACCG GCAGGGGCTC GGCTCACCCC

ForAA(FAM) A T TCTTCCGGTTGC

CGACCCGCCC CCTGCCGCTT GCTCGTAGCT TTGGCAGGTA AGAAGGCCAA

25 **DGAT1ADGC (VIC) CCAT TCCGCCGGTT**

68348-71

40

CGGGGGAGCT GCCCAGCGCA CCGTGAGCTA CCCCAGACAAC CTGACCTACC

←

CTGTTG GACTGGATGG

GCGGTGAGGA TCCTGCCGGG GGCTGGGGGG ACTGCCCGGC GGCCTGGCCT

CGCC ReverseNZ

- 5 GCTAGCCCCG CCCTCCCTTC CAGATCTCTA CTACTTCCTC TTCGCCCCCA
(SEQ ID NO:75)

A 240bp product is produced in this reaction. When the Q allele (AA) is present the FAM-labelled probe binds and fluoresces at 518nm. When the q allele (GC) is present the VIC-labelled probe binds and fluoresces at 554nm. After cycling is
10 complete, the plate is scanned on the ABI7900 Sequence Detection System, the fluorescence from each well detected, and a scattergraph is drawn. The scattergraph separates out into 3 clumps with Q homozygotes in the upper left hand corner, q homozygotes in the lower right hand corner and Qq heterozygotes in between. Each clump is circled and the software automatically determines the genotype for each
15 sample. On each plate there are controls with 8 wells each of known Q homozygotes, q homozygotes, Qq heterozygotes and no template controls.

Splice Variant Gene Expression

To determine the relative gene expression of the splice variants created by insertion/deletion of 66bp around the polymorphic site by alternate exon usage, RNA
20 was extracted from mammary tissue and reverse transcribed using oligodT primer using a first strand cDNA synthesis kit (Invitrogen). Real time PCR to determine relative quantities of each variant was then carried out.

Primer sequences 5' to 3', genomic sequence position in brackets:

DGAT1forRT66: TCTCCTACCGGGACGTCAAC (6652) (SEQ ID NO:76)

- 25 DGAT1revRT66: GAGATCGCGGTAGGTCAGGTT (6964) (SEQ ID NO:77)

DGAT1forRTless66: GCTGCTTTGGCAGATCTCTACTACTT (6711) (SEQ ID NO:78)

DGAT1revRTless66: AAGCGCTTTCGGATGCG (7038) (SEQ ID NO:79)

Probe sequences 5' to 3', genomic sequence position in brackets:

68348-71

41

DGAT1with66 (FAM): CCGTGAGCTACCC (6857) (SEQ ID NO:80)

DGAT1less66 (VIC): CTTCGCCCCCACCCT (6976) (SEQ ID NO:81)

Both probes use MGB (minor groove binder) as a non-fluorescent quencher.

Final reaction conditions were 1X Universal PCR Mastermix (Applied Biosystems),
5 60nM each primer (Invitrogen), 60nM each probe (Applied Biosystems) and 1µl of
template cDNA in a total volume of 10µl.

Cycling conditions were 50°C for 2 minutes, 95°C initial denaturation for 10 minutes,
then 37 cycles of denaturation at 94°C for 15 seconds, annealing and extension 60°C
for 1 minute.

10 Primer positions around 66bp insertion (in italics) on cDNA sequence. The start of
the cDNA sequence is equivalent to position 6479 on the genomic sequence, with the
last base of the cDNA equivalent to position 7428 of the genomic sequence.

CCGTGGCCTT TCTCCTCGAG TCTATCACTC CAGTGGGCTC CGTGCTGGCC

DGAT1forRT66 TCTCCT ACCGGGACGT

15 CTGATGGTCT ACACCATCCT CTTCCTCAAG CTGTTCTCCT ACCGGGACGT

CAAC → DGAT1forRTless66 GCTGCTT

CAACCTCTGG TGCCGAGAGC GCAGGGCTGG GGCCAAGGCC AAGGCTGCTT

TGGCAG DGAT1with66(FAM) C CCGTGAGCTAC

TGGCAGGTAA *GAAGGCCAAC* GGGGGAGCTG CCCAGCGCAC CGTGAGCTAC

20 CC ATCTCTAC TACTT →

CCCGACAACC TGACCTACCG CGATCTCTAC TACTTCCTCT TCGCCCCCAC

← TTGG ACTGGATGGC GCTAGAGDGAT1revRT66 CT TCGCCCCCAC

CCTGTGCTAC GAGCTCAACT TCCCCCGCTC CCCCCGCATC CGAAAGCGCT

CCT DGAT1less66 (VIC) ← GCGTAG GCTTTCGCGA

68348-71

42

TCCTGCTGCG GCGACTCCTG GAGATGCTGT TCCTCACCCA GCTCCAGGTG

A DGAT1revRTless66

GGGCTGATCC AGCAGTGGAT GGTCCCGGCC ATCCAGAACT CCATGAAGCC

CTTCAAGGAC ATGGACTACT CCCGCATCGT GGAGCGCCTC CTGAAGCTGG

5 (SEQ ID NO:82)

This reaction detects the presence of the insertion splice variant by creating a 145bp product which binds the FAM probe only. The deletion splice variant is detected by a 92bp product that binds the VIC probe only.

The cDNA for each alternate splice variant was cloned into pGemT (Promega). A
10 dilution series of the same, known amount, of each variant plasmid DNA was used to create a standard curve that established the linearity of the PCR reaction over a range of DNA concentrations. The threshold cycle number of the sample variants was converted back to a DNA amount by linear regression and the amounts of each variant present compared.

15 The presence of an alternate spice variant raises the possibility of an alternate function that is at this stage unknown.

It will be appreciated that it is not intended to limit the invention to the above
examples only, many variations, which may readily occur to a person skilled in the art,
being possible without departing from the scope thereof as defined in the
20 accompanying claims.

INDUSTRIAL APPLICATION

The present invention is directed to a method of genotyping bovine for improved milk
production traits. In particular, such traits include increased milk volume and milk
protein content and decreased milkfat content and solid fat content. It is anticipated
25 that herds of bovine selected for such a trait will produce milk which will be more
easily processed and such milk and products made therefrom may provide health
benefits to consumers, as well as producing an increased milk yield.

REFERENCES

- CASES, S.; SMITH, S.J.; ZHENG, Y.-W.; MYERS, H.M.; LEAR, S.R.; SANDE, E.; NOVAK, S.; COLLINS, C.; WELCH, C.B.; LUSIS, A.J.; ERICKSON, S.K.; FARESE, R.V. JR (1998). Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. *Proceedings of the National Academy of Sciences, USA*. 95(22):13018-23.
- 10 COPPIETERS, W.; RIQUET, J.; ARRANZ, J.-J.; BERZI, P.; CAMBISANO, N.; GRISART, B.; KARIM, L.; MARCQ, F.; SIMON, P.; VANMANSHOVEN, P.; WAGENAAR, D.; GEORGES, M. (1998) A QTL with major effect on milk yield and composition maps to bovine chromosome 14. *Mammalian Genome* 9: 540-544.
- 15 DELOUKAS et al (1998) A physical map of the 30,000 human genes. *Science* 282, 744-746
- EWING, B.; HILLIER, L.; WENDL, M.C.; GREEN, P. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Research* 8(3):175-185.
- 20 EWING, B.; GREEN, P. (1998). Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Research* 8(3):186-194.
- GORDON, D.; ABAJIAN, C.; GREEN, P. (1998). Consed: a graphical tool for sequence finishing. *Genome Research* 8(3): 195-202.
- 25 RIQUET, J.; COPPIETERS, W.; CAMBISANO, N.; ARRANZ, J.-J.; BERZI, P.; DAVIS, S.; GRISART, B.; FARNIR, F.; KARIM, L.; MNI, M.; SIMON, P.; TAYLOR, J.; VANMANSHOVEN, P.; WAGENAAR, D.; WOMACK, J.E.; GEORGES, M. (1999). Identity-by-descent fine-mapping of QTL in outbred populations: application to milk production in dairy cattle. *Proceedings of the National Academy of Sciences, USA* 96: 9252-9257.
- 30 FARNIR, F.; GRISART, B.; COPPIETERS, W.; RIQUET, J.; BERZI, P.; CAMBISANO, N.; KARIM, L.; MNI, M.; SIMON, P.; WAGENAAR, D.; GEORGES, M. (2000). Simultaneous mining of linkage and linkage disequilibrium to fine-map QTL in outbred half-sib pedigrees: revisiting the location of a QTL with major effect on milk production on bovine chromosome 14. *Ph.D Thesis, University of Liege* 2000.
- 35

- SMITH, S.J.; CASES, S.; JENSEN, D.R.; CHEN, H.C.; SANDE, E.; TOW, B.; SANAN, D.A.; RABER, J; ECKEL, R.H.; FARESE, R.V. JR (2000). Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking DGAT. *Nature Genetics* 5 25(1):87-90.
- BARENDSE, W.; ARMITAGE, S.M.; KOSSAREK, L.M.; SHALOM, A.; KIRKPATRICK, B.W.; RYAN, A.M.; CLAYTON, D.; LI, L.; NEIBERGS, H.L.; ZHANG, N.; GROSSE, W.M.; WEISS, J.; CREIGHTON, P.; MCCARTHY, F.; RON, M.; TEALE, A.J.; FRIES, R.; 10 McGRAW, R.A.; MOORE, S.S.; GEORGES, M.; SOLLER, M.; WOMACK, J.E.; HETZEL, D.J.S. (1994). A genetic linkage map of the bovine genome. *Nature Genet.* **6**: 227-235.
- BISHOP, M.D.; KAPPES, S.M.; KEELE, J.W.; STONE, R.T.; SUNDEN, S.L.F.; HAWKINS, G.A.; SOLINAS TOLDO, S.; FRIES, R.; GROSZ, M.D.; YOO, J.; BEATTIE, C.W. (1994). A 15 genetic linkage map for cattle. *Genetics* **136**: 619-639.
- COLLINS, F.S. (1995). Positional cloning moves from perditional to traditional. *Nature Genet.* **9**: 347-350.
- 20 GEORGES, M.; ANDERSSON, L. (1996). Livestock genomics comes of age. *Genome Research.* **6**: 907-921.
- GEORGES, M; NIELSEN, D.; MACKINNON, M.; MISHRA, A.; OKIMOTO, R.; PASQUINO, A.T.; SARGEANT, L.S.; SORENSEN, A.; STEELE, M.R.; ZHAO, X.; WOMACK, J.E.; 25 HOESCHELE, I. (1995). Mapping quantitative trait loci controlling milk production by exploiting progeny testing. *Genetics* **139**: 907-920.
- HOUBENWCYL (1987). Methods of Organic Chemistry, ed. E. Wansch. Vol. 15 I and II. Thieme, Stuttgart.
- 30 HUSE *et al.* (1989). *Science* **246**: 1275-1281.
- KAPPES, S.M.; KEELE, J.W.; STONE, R.T.; McGRAW, R.A.; SONSTEGARD, T.S.; SMITH, T.P.L.; LOPEZ-CORRALES, N.L.; BEATTIE, C.W. (1997) A Second-Generation 35 Linkage Map of the Bovine Genome. *Genome Research* **7**: 235-249.
- KOHLER and MILSTEIN (1975). *Nature.* **256**: 495-497.

KOZBOR *et al.* (1983). *Immunol. Today* **4**: 72.

MERRIFIELD (1964). *J. Am. Chem. Assoc.* **85**: 2149-2154.

5

McCAFFERTY *et al.* (1990) *Nature* **348**: 552-554.

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbour Lab Press, Cold Spring Harbour, New York.

10

STEWART, A.J.; CANITROT, Y.; BARACCHINI, E.; DEAN, N.M.; DEELEY, R.G. and COLE, S.P.C. (1996). Reduction of expression of the multidrug resistance protein (MRP) in human tumour cells by antisense phosphorothioate oligonucleotides. *Biochem. Pharmacol.* **51**: 461-469.

15

VAN RADEN, P. M., AND G. R. WIGGANS, 1991 Derivation, calculation, and use of National Animal Model Information. *J. Dairy Sci.* **74**: 2737-2746.

20

JOHNSON, D.L.; THOMPSON, R. 1995. Restricted maximum likelihood estimation of variance components for univariate animal models using sparse matrix techniques and average information. *J. Dairy Sci.* **78**: 449-456.

JOHNSON D.L., S.F. PETCH, A.M. WINKELMAN AND M. BRYANT 2000 Genetics of milk characteristics in New Zealand dairy cattle Proceedings of New Zealand Society of Animal Production **60**:318-319.

25

MACGIBBON A.K.H. & MCLENNAN W. D. 1987. Hardness of New Zealand patted butter: seasonal and regional variations. *New Zealand Journal of Dairy Science and Technology*, **22**: 143-156.

FALCONER D.S. and MACKAY T.F.C. Introduction to Quantitative Genetics, 4th Edition. Longman Scientific and Technical, New York, 1996.

30

WELLER, J.I. *et al.* Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci.* **73**, 2525-2537, (1990).

LYNCH M and WALSH B (1997). Genetics and analysis of quantitative traits. Sinuaer Associates, Inc. Sunderland, Massachusetts.

SAS Institute Inc. 1985 SAS Users Guide: Statistics, Version 5 edition, Cary, N.C. 956p.

- 5 BOLDMAN K.G.; KRIESE L.A., VAN VLECK L.D.; VAN TASSELL C.P.; KACHMAN S.D. A manual for the use of MTDFREML, 1997.
<http://aiipl.arsusda.gov/curtvvt/mtdfreml.html>.

SEQUENCE LISTING IN ELECTRONIC FORM

In accordance with Section 111(1) of the Patent Rules, this description contains a sequence listing in electronic form in ASCII text format (file: 68348-71 Seq 06-05-11 v3.txt).

A copy of the sequence listing in electronic form is available from the Canadian Intellectual Property Office.

The sequences in the sequence listing in electronic form are reproduced in the following table.

SEQUENCE TABLE

<110> REID, SUZANNE J
 FORD, CHRISTINE A
 GEORGES, MICHEL A J
 COPPIETERS, WOUTER H R
 GRISART, BERNARD M J J
 SNELL, RUSSELL G
 SPELMAN, RICHARD J

<120> MARKER ASSISTED SELECTION OF BOVINE FOR IMPROVED MILK COMPOSITION

<130> 68348-71

<140> PCT/NZ01/245
 <141> 2001-10-31

<160> 82

<170> PatentIn version 3.1

<210> 1
 <211> 11771
 <212> DNA
 <213> Bos taurus

<220>
 <221> CDS
 <222> (1)..(191)
 <223> Exon 1 CDS, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
 <221> misc_signal
 <222> (1)..(4)
 <223> these bases correspond to bases 4 to 7 of the Kozak recognition sequence. See DGAT1 cDNA for the complete recognition sequence.

<220>
 <221> CDS
 <222> (3809)..(3896)
 <223> Exon 2, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (5840)..(5880)
<223> Exon 3, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (5960)..(6045)
<223> Exon 4, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (6138)..(6190)
<223> Exon 5, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (6406)..(6511)
<223> Exon 6, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (6601)..(6714)
<223> Exon 7, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (6815)..(6889)
<223> Exon 8, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> variation
<222> (6824)..(6889)
<223> this sequence is deleted from the alternately spliced transcript.

<220>
<221> CDS
<222> (6960)..(7063)
<223> Exon 9, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (7154)..(7192)
<223> Exon 10, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (7271)..(7312)
<223> Exon 11, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (7386)..(7430)
<223> Exon 12, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (7505)..(7617)
<223> Exon 13, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (7705)..(7770)
<223> Exon 14, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (7858)..(7945)
<223> Exon 15, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (8027)..(8089)
<223> Exon 16, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> CDS
<222> (8162)..(8314)
<223> Exon 17 CDS, determined by alignment with an amino acid sequence deduced from the cDNA

<220>
<221> misc_feature
<222> (8315)..(8317)
<223> translation stop codon

<220>
<221> polyA_site
<222> (8572)..(8578)

<220>
<221> variation
<222> (626)..(626)
<223> thymidine (T) to guanine (G) substitution polymorphism

<220>
<221> variation
<222> (3512)..(3512)
<223> thymidine (T) to guanine (G) substitution polymorphism

<220>
<221> variation
<222> (4040)..(4040)
<223> thymidine (T) to cytosine (G) substitution polymorphism

<220>
<221> variation
<222> (4963)..(4963)
<223> adenine (A) to guanine (G) substitution polymorphism

<220>
<221> variation
<222> (5003)..(5003)
<223> guanine (G) to adenine (A) substitution polymorphism

<220>
<221> variation
<222> (5997)..(5997)
<223> cytosine (C) to thymine (T) substitution polymorphism

<220>
<221> variation
<222> (6829)..(6830)
<223> adenine (A)-adenine (A) to guanine(G)-cytosine (C) substitution
polymorphism; AA corresponds to the Q allele
GC corresponds to the q allele

<220>
<221> variation
<222> (6892)..(6892)
<223> guanine (G) to adenine (A) substitution polymorphism

<220>
<221> variation
<222> (7224)..(7225)
<223> guanine (G)-guanine (G) to adenine (A)-cytosine (C) substitution
polymorphism
GG-AC

<220>
<221> variation
<222> (7438)..(7438)
<223> adenine (A) to guanine (G) substitution polymorphism

<220>
<221> variation
<222> (7456)..(7456)
<223> cytosine (C) to thymidine (T) substitution polymorphism

<220>
<221> variation
<222> (7987)..(7987)
<223> guanine (G) to adenine (A) substitution polymorphism

<220>
<221> variation
<222> (8402)..(8402)
<223> ctosine (C) to thymidine (T) substitution polymorphism

<220>
<221> misc_feature
<222> (9434)..(9434)
<223> ambiguous nucleotide

<220>
<221> misc_feature
<222> (9496)..(9496)
<223> ambiguous nucleotide

<220>
<221> misc_feature
<222> (10402)..(10417)
<223> ambiguous nucleotides

<220>
<221> primer_bind
<222> (6579)..(6601)
<223> Primer 17F
CCTGAGCTTGCCTCTCCCACAGT

<220>
<221> primer_bind
<222> (7036)..(7058)
<223> Primer 18R
CCAGGAGTCGCCGCAGCAGGAAG
reverse primer

<220>
<221> primer_bind
<222> (7280)..(7303)
<223> Primer 6F
CCGGCCATCCAGAACTCCATGAAG

<220>
<221> primer_bind
<222> (7585)..(7605)
<223> Primer AW446985 dn1
TAGAACTCGCGGTCTCCAAAC
reverse primer

<220>
<221> primer_bind
<222> (8222)..(8244)
<223> Primer InsUp1
TGGCTGTCACTCATCATCGGGCA

<220>
<221> primer_bind
<222> (8566)..(8589)
<223> Primer 14R2
TTGCACAGCACTTTATTGACACA

<220>
<221> primer_bind
<222> (6813)..(6830)
<223> Primer SNP1_FAM
AGC TTT GGC AGG TAA GGC

<220>
<221> primer_bind
<222> (6813)..(6830)
<223> Primer SNP1_HEX
AGC TTT GGC AGG TAA GAA

<220>
<221> primer_bind
<222> (6831)..(6844)
<223> Primer SNP1_2P
GGC CAA CGG GGG AG

<220>
<221> primer_bind
<222> (7424)..(7438)
<223> Primer SNP2_FAM
GCT GGC GGT GAG TGA

<220>
<221> primer_bind
<222> (7424)..(7438)
<223> Primer SNP2_HEX
GCT GGC GGT GAG TGG

<220>
<221> primer_bind
<222> (7439)..(7453)
<223> Primer SNP2_2P
CCT GCT GGG TGG GGA

<220>
<221> primer_bind
<222> (7442)..(7456)
<223> Primer SNP3_FAM
GCT GGG TGG GGA CGC

<220>
<221> primer_bind
<222> (7442)..(7456)
<223> Primer SNP3_HEX
GCT GGG TGG GGA CGT

<220>
<221> primer_bind
<222> (7457)..(7470)
<223> Primer SNP3_P
GTG GGG GCG GGT GG

<220>
<221> primer_bind
<222> (8388)..(8402)
<223> Primer SNP4_FAM
TGC CCC AAC CTG GGT

<220>
<221> primer_bind
<222> (8388)..(8402)
<223> Primer SNP4_HEX
TGC CCC AAC CTG GGC

<220>
<221> primer_bind
<222> (8403)..(8417)
<223> Primer SNP4_2P
GCA GCA GGA GGA GGC

<220>
<221> primer_bind
<222> (6811)..(6830)
<223> Primer Dgat 21
GTAGCTTTGGCAGGTAAGAA

<220>
<221> primer_bind
<222> (6965)..(6984)
<223> Primer Dgat 22
GGGGCGAAGAGGAAGTAGTA
reverse primer

<220>
<221> primer_bind
<222> (6613)..(6632)
<223> Primer Dgat 23
TGGCCCTGATGGTCTACACC

<220>
<221> primer_bind
<222> (6829)..(6850)
<223> Primer Dgat 24B
GGGCAGCTCCCCCGTTGGCCGC
reverse primer

<220>
<221> primer_bind
<222> (6651)..(6670)
<223> Primer DgatforAD
TTCTCCTACCGGGACGTCAA

<220>
<221> primer_bind
<222> (6871)..(6890)
<223> Primer ReverseNZ
CCGCGGTAGGTCAGGTTGTC
reverse primer

<220>
<221> primer_bind
<222> (6825)..(6838)
<223> Probe ForAA (FAM)
CGTTGGCCTTCTTA

<220>
<221> primer_bind
<222> (6823)..(6836)
<223> Probe DgatADGC (VIC)
TTGGCCGCCTTACC

<220>
<221> primer_bind
<222> (6651)..(6670)
<223> Primer DgatforAD
TTCTCCTACCGGGACGTCAA

<220>
<221> primer_bind

```

<222> (6878)..(6972)
<223> Primer DgatreVAD
      AAGTAGTAGAGATCGCGGTAGGTCA
      reverse primer

<220>
<221> primer_bind
<222> (6825)..(6838)
<223> Probe ForAA (FAM)
      CGTTGGCCTTCTTA

<220>
<221> primer_bind
<222> (6823)..(6836)
<223> Probe DgatADGC (VIC)
      TTGGCCGCCTTACC

<220>
<221> primer_bind
<222> (6652)..(6671)
<223> Primer DgatforRT66
      TCTCCTACCGGGACGTCAAC

<220>
<221> primer_bind
<222> (6874)..(6964)
<223> Primer DgatreVRT66
      GAGATCGCGGTAGGTCAGGTT
      reverse primer

<220>
<221> primer_bind
<222> (6711)..(6972)
<223> Primer DgatforRTless66, GCTGCTTTGGCAGATCTCTACTACTT
      This primer was designed to selectively bind and amplify
      the cDNA splice variant.

<220>
<221> primer_bind
<222> (6711)..(6972)
<223> The corresponding binding site in this genomic sequence comprises
      bases 6711 to 6715, 6815 to 6823 and 6960 to 6972.

<220>
<221> primer_bind
<222> (7022)..(7038)
<223> Primer DgatreVRTless66
      AAGCGCTTTCGGATGCG
      reverse primer

<220>
<221> primer_bind
<222> (6857)..(6870)
<223> Probe Dgatwith66 (FAM)
      CCGTGAGCTACCC

<220>
<221> primer_bind

```

gtagcgggtgc	gcgtgacccc	taacctttga	cccttgatac	ggggcccttg	cgacccaacc	251
tggtggccca	ggcctgtcgg	cggcagctcg	ggctcgagtc	cgagagtcctg	gcgcctggac	311
cttggtgcac	agctgtgccc	ctcgggcctc	cacggggaaa	cttagcggga	ggttgggggc	371
ggaggggtctc	ctgcccggaa	cacccaggta	cggggggccga	ggggagggca	gcggctcaac	431
ttctagacgc	cctccctctg	ccttcctttg	gtgggttctg	aagctttccc	agggtgagcc	491
cactacgcac	agtgtcctct	acctggaagg	agatacaggg	gtccttctctg	agggctatga	551
ggggtgcctt	gtgggttgat	aaagctcccg	ggggaggagg	gtggaccggc	ggagaacaga	611
ggcagggggca	gtgctagggg	atttctcatc	cctcgcagac	cctccagaga	atggctcttca	671
caaaggtccc	tcatacgtca	cccggcgatt	gactggccta	ggatcctgct	tattaccagc	731
acaaatggct	gctctagggg	caaagtgggt	cctgtaatgg	gaccctcacc	cctggtttggg	791
gtacagggga	ggagttggaa	gtgcgcacac	ccacaggtgg	gcgccttgct	tagctgaagg	851
actgatggga	aggagttggg	ggagcaagct	gcggctgaaa	gggaggatct	gaccacgtg	911
ggcatcagct	aagtcttgct	ggctgcctcc	aggcgttccc	tttgccatcc	tccacgcccc	971
tcccccgggg	cctgaccttc	atcctggtca	agggtctctca	ggggctctgg	ttttgggatac	1031
agctccagag	ctagaggtta	tcaaggagga	agtgggcaac	aggtcagtc	gcaaggattt	1091
gctatcttca	ctgggtgctg	tggggagggg	agggacaagg	gcagttgggg	tgcaggcact	1151
gtccctgccc	ttggggggca	cacagttcac	ctgagagata	agatagccgc	agccctgaag	1211
agtgagagca	aaggtcaggc	acagagttca	ggatgacacc	aggggagggg	ggctctgtga	1271
ggggcactgg	cttcctacag	gccccagggtg	gtcctgaggg	ggcggctgca	aaggccagga	1331
ggcccacagg	ccctcttgcc	cactcctggg	gaactggatt	tgggggtcact	ttgtatgagg	1391
tggggggcgg	taccagcttt	gggccaaagct	gtcaccctgg	atggggccatc	acttgccctgc	1451
tctgtatagg	ccagatggcc	agaagctgct	cctgtcctgt	tgatggccca	tcctcgaggt	1511
ctggaccctc	gggaagagga	gcagttggtg	gcagggatgg	gccaccggag	accctcctga	1571
cctccaggac	acgcagctgt	gtgtgcctgt	ccccaaagcca	catgccacat	ggctaggggc	1631
ctcctggggc	agggctgggc	attggtctgg	ctactcttgg	tatcgcctat	ggcttccctg	1691
cctcccagtc	atcatacctc	cacctctgcc	tccctgcctg	ttcctctctt	tctcctcagg	1751
cccttccgga	catttctctgc	tcacctaggt	ctgggcaggc	gggggtcaggt	gccgggtgtg	1811
agctcaactcc	ttccggcagc	aagggtgtagc	tatgtgccgg	aaggaaggcc	gctgctgttg	1871
cctcgcctct	gagtgcatac	cttccaggtc	ctccacactc	ccctgtgccc	cgacacctgg	1931
tgcgtccttc	agccattggg	tcatagtgtcc	tccaggcaca	gctttctagt	ccagagcctc	1991
taggctgggt	gcaggaagtg	ctgaggaagt	ggcagccggg	aggcgagctg	gcaccctgtc	2051
cctccttggt	ctgtccgtcc	ctgcccctgg	accgtatggc	cccgcatagtg	tgatccccac	2111
ttggggctgt	gcctctgggc	aagttgggaa	gcttggtgag	cctcattttc	atgtgcccgc	2171
ctcccagtac	tgatgtgcag	gttgaatgag	gtgccaaactg	taatgagttg	gaatggccct	2231
gctggctggg	tgggactggg	gagcaggtgg	gggccgctgg	ggggcacaga	ggcacaccca	2291
gtgcctcagt	cagggagagg	gtgacagaga	agctctgggt	gaggccccac	ctccactctg	2351
gccatggctg	ctgccctttg	gtccactgca	gtgaactgtg	ccatggggct	ggacctctgt	2411
ggggattggg	gggcagtggg	ctttcttccc	gcttggggcc	tctgacctct	gggggcaggg	2471
cgctgcccgg	gtgggacagt	cggaaaggctg	gtagaggggac	ctgagggggtc	tgtgtgggtgg	2531

ctgggggag	gcctcaggaa	tttgacagca	gggatctgga	aaagctttaa	taacattatt	2591
tggtgtcagg	attgggaaat	gctccccctcc	ccccctcccc	tctttcatct	tagagactgc	2651
tgcacatctg	gtcagtgtgg	tcttcttggg	ggcccccaag	gtggcagggg	tcacactgtt	2711
atgaaaccgt	cccctgggta	tgtgggtgcag	acatgcacat	gcagatgggtg	attggcaggt	2771
tgtagcatga	ggtaggctttg	ggacgggtcc	agtgcagctg	agtgaggctgg	atctgggggg	2831
ttctggggcag	gtccatcaag	cggatacccc	cacagactgt	cctcttggga	tagttggggc	2891
tgggagccct	gcttgccctg	ccaaaaggca	ggcgcagagt	catgaagaag	agggcttggg	2951
ggctcagagc	cccactgtgt	gtgcagccca	gggtggacct	ggaggagggtg	cgtgggcagg	3011
ctgggcccggc	cggggcgggg	ggtagggggg	cctgggtgtga	aagggaacca	gggcccagact	3071
gtcagcgctg	cctggctgag	gatgctggca	ccctgtcctc	cccagccgtc	tgtctcctgg	3131
gtgcagccat	ctgagtgtg	accccagccg	cccctggagg	ctggctgttc	tcctgtgccc	3191
tattgctggg	gacatgtgtc	cacaggaggg	aaagggaagc	cccggcctct	ccccttacia	3251
aactggaggc	cttgctcaat	gccctggatg	gcctcctggt	ggcaggggtg	ttgggtgggag	3311
gtggggctgc	tgcttagaac	ccgccagcgg	gcctgggcct	gggctgagct	gcacccctcc	3371
acctctgcct	ccagctgagg	gttggcttcc	atctccacca	ggcccagcac	tgggcacagg	3431
gctctcagag	gcaggctctg	aaagtccctc	gctggcttct	gcagtggact	ccaggcgccg	3491
agcccccagg	gggctcgcct	tgctctcacc	ctgcgaagcc	acgtgaaggc	tgggtcctcc	3551
cctccggaag	ggccaaatgc	agggcatggg	tggtttgaat	ggtagggcct	gggctccccg	3611
gagggaccag	ctgctgtgag	ggccgcccc	tccccacttc	cgtcttgcat	caccagctcc	3671
tgtggcactc	cccacgcccc	gtccccagct	gggagcggca	ggcccccggt	ggctctgccc	3731
gcggaggggg	atgtgtgggc	ggcgggggtg	ccttgctgcc	agatgctctg	ccccgagtgt	3791

ccgtctccgc	tctccag	g	tgt	cac	cgc	ctg	cag	gat	tcc	ctg	ttc	agt	tct	3842
			Cys	His	Arg	Leu	Gln	Asp	Ser	Leu	Phe	Ser	Ser	
			65					70					75	

gac	agt	ggc	ttc	agc	aac	tac	cgt	ggc	atc	ctg	aat	tgg	tgt	gtg	gtg	3890
Asp	Ser	Gly	Phe	Ser	Asn	Tyr	Arg	Gly	Ile	Leu	Asn	Trp	Cys	Val	Val	
			80						85					90		

atg	ctg	gtacgtagag	tgacaccttg	gagcaagggt	cctgacggcc	ggggggccat	3946
Met	Leu						

gggctcttct	ccaggggtag	gtgtctgtac	ttgtgtagct	gtgggtgaatg	gagctctgtg	4006
ctggcggtgg	gggtccctgg	agcagccgta	ccctgggacc	ctaccgggag	catgctcatg	4066
ccgtccctgc	tgaatcccag	gagatgcctg	cagagggcag	cctgggagcc	tctgagctgg	4126
gggtctgcgc	ccagggggca	ctggagtctc	cccagggggc	gagagagagt	aggcagggat	4186
gggtctgggtg	ccctgggtgg	gggatggctg	ctccgtgggc	ccaggccctc	cctggcagca	4246
caggtgagtg	gtcttggggg	tccacgtaga	acttctctct	ctgttccaaa	ttgccctcat	4306
gggtgcggca	tgcttgggtg	aacctggggg	agcaggggtga	ggacatgctt	ctcagcccag	4366
cccacagctc	caggccacac	tctgcaggac	tctggcccc	ccctcagccc	tggaggggagc	4426
aggactggag	tcctgtgtcc	gccttgctct	gacctggccg	aggccactgc	tgtggggccc	4486
cagcaggcct	gcccagcaga	aggtggagtg	cagggacccc	aggggcagcc	ttcagggtgg	4546
ggcaggggtga	ggcccagactg	ggcccagccc	caccgctcag	tgctgatgtg	gcgcgaggcc	4606
ttcgccccctc	cagctgacgt	gtctgcctgc	cctgggtgtg	gctccagagg	ctgcctgtgt	4666
accagggggcc	cccacgcttc	tgtttgtggg	tctgggcagt	cccctgggga	gcgggtggggg	4726
ctgtgtgcca	gtccagaccc	agtagtccac	gcgtcctggt	ctctggaggc	cgtggctggt	4786
ccaggactgt	ggcaagggtg	tcgtgcaggg	caggccctca	gcagcctgtc	tgttctcctg	4846
cagccccccag	cctcctggcc	ctttgggtgca	cccacaaagc	tccccctcc	cccaggagct	4906
ggggccgcct	gctgcgtcct	ctcggcagcc	tgggcttcca	ggtaggctggg	cctcttagca	4966
gctccaactc	ttgcctgtgg	tgggctctca	ggacaggcaa	ctgccagtcg	gcagacattg	5026
caggaccacg	tgtgtcctgg	taagctggct	ggtaggtgtg	ttagctgggg	gatgggtgtg	5086
caggtggccc	ctgcatctct	gagcctgtca	cctcctcggg	aagccttctg	ggtagggggac	5146
tccacccatg	tcgcctggag	aagcatcact	tttccacaga	gccttctgca	acccccgtgg	5206
ggcctgagcc	tgggggtggg	gaggtgggtg	cccctgctcc	tgcagaggcc	agccaggcat	5266
ctggccccag	gccactggca	agagctcggt	gtgttggggg	atctgtcctt	tgctgctgct	5326
gcaggagcgg	ccgaggcagg	cgggggcgtg	agtaggggtg	gagacccagg	cccagcttcc	5386
ccagccccctc	aggaccggcc	tgctctttcc	caccacccca	ccaagtgcgt	gggcacaccc	5446
cgctgtgag	gatgggcccc	gtagggcagg	cggagccctg	ggagggtggc	agtgcgcggg	5506
gcaggccttg	acttcactgg	ggcttggggg	tgtcgtgtgt	gccagggggc	ctgacccgct	5566

tggtgggacg gacggccgct gggcagcagg tttcttctgc cacggtggca caggcacctg	5626
gggttgtggt tggctccagg cgggcggggg ctgcgtgccc ctgcgcaggc acataggccg	5686
tgggtgggga gtctcagagc ttggcgtgag gtcccacagg gctgggcctg caggatggag	5746
gccactgtcc tgagctgcag gtgctggcag gagctggggg gggcgttctg gggccgtggc	5806
tgacagcgtt atgtccctct ctctctatcg cag atc tta agc aac gca cgg tta	5860
Ile Leu Ser Asn Ala Arg Leu	
95 100	
ttt cta gag aac ctc atc aa gtgagtgggc cccggcctgc cccagccct	5910
Phe Leu Glu Asn Leu Ile Lys	
105	
gccacctcac ccctcgcta cacagaccct caccacactg cgtctgcag g tat ggc	5966
Tyr Gly	
atc ctg gtg gac ccc atc cag gtg gtg tct ctg ttc ctg aag gac ccc	6014
Ile Leu Val Asp Pro Ile Gln Val Val Ser Leu Phe Leu Lys Asp Pro	
110 115 120 125	
tac agc tgg cca gct ctg tgc ctg gtc att g gtgagctggg tgcccaggag	6065
Tyr Ser Trp Pro Ala Leu Cys Leu Val Ile	
130 135	
gcctcaggcc ggcggtgggt gggacagggc tgatctgggc ctgaacctgc cctggggtgc	6125
ttctgtctc a gtg gcc aat atc ttt gcc gtg gct gcg ttc cag gtg gag	6175
Val Ala Asn Ile Phe Ala Val Ala Ala Phe Gln Val Glu	
140 145	
aag cgc ctg gcc gtg gtaagcagt cctcacgcc ctcccctgac ttgcctcaag	6230
Lys Arg Leu Ala Val	
150	
gtccttacca gtcgggctta gggcgggcca ccagctggtc ccaactgtgct tcagggtttt	6290
gggcctttcg tggccttct gagaggggct gcacctcagg cctgggtggct ctctctcagg	6350
gaggtcctct gaccagggag gggggtcctt ggctgacgct ctgctccac cccag gga	6408
Gly	
gct ctg acg gag cag gcg ggg ctg ctg ctg cac ggg gtc aac ctg gcc	6456
Ala Leu Thr Glu Gln Ala Gly Leu Leu Leu His Gly Val Asn Leu Ala	
155 160 165 170	
acc att ctc tgc ttc cca gcg gcc gtg gcc ttt ctc ctc gag tct atc	6504
Thr Ile Leu Cys Phe Pro Ala Ala Val Ala Phe Leu Leu Glu Ser Ile	
175 180 185	
act cca g gtgggcccc cccccgccc cgccccgcc cacgctgtct cggccacggg	6561
Thr Pro	
cagcgcgggg ggcgtggcct gagcttgcct ctcccacag tg ggc tcc gtg ctg	6614
Val Gly Ser Val Leu	
190	
gcc ctg atg gtc tac acc atc ctc ttc ctc aag ctg ttc tcc tac cgg	6662
Ala Leu Met Val Tyr Thr Ile Leu Phe Leu Lys Leu Phe Ser Tyr Arg	
195 200 205	

gac gtc aac ctc tgg tgc cga gag cgc agg gct ggg gcc aag gcc aag 6710
 Asp Val Asn Leu Trp Cys Arg Glu Arg Arg Ala Gly Ala Lys Ala Lys
 210 215 220 225

gct g gtgagggctg cctcgggctg gggccactgg gctgccactt gcctcgggac 6764
 Ala

cggcagggggc tcggctcacc cccgacccgc cccctgccgc ttgctcgtag ct ttg 6819
 Ala Leu

gca ggt aag aag gcc aac ggg gga gct gcc cag cgc acc gtg agc tac 6867
 Ala Gly Lys Lys Ala Asn Gly Gly Ala Ala Gln Arg Thr Val Ser Tyr
 230 235 240

ccc gac aac ctg acc tac cgc g gtgaggatcc tgccgggggc tggggggact 6919
 Pro Asp Asn Leu Thr Tyr Arg
 245 250

gcccggcggc ctggcctgct agccccgccc tcccttcag at ctc tac tac ttc 6973
 Asp Leu Tyr Tyr Phe
 255

ctc ttc gcc ccc acc ctg tgc tac gag ctc aac ttc ccc cgc tcc ccc 7021
 Leu Phe Ala Pro Thr Leu Cys Tyr Glu Leu Asn Phe Pro Arg Ser Pro
 260 265 270

cgc atc cga aag cgc ttc ctg ctg cgg cga ctc ctg gag atg 7063
 Arg Ile Arg Lys Arg Phe Leu Leu Arg Arg Leu Leu Glu Met
 275 280 285

gtgagggcggg gcctcgcggg ccaggggtggg cgggcctgcc ggcacccggc accggggctc 7123

agctcactgt ccgcttgctt ccttccccag ctg ttc ctc acc cag ctc cag gtg 7177
 Leu Phe Leu Thr Gln Leu Gln Val
 290

ggg ctg atc cag cag gtacgtgccc ggggggggggg ggggggggggg gggggggact 7232
 Gly Leu Ile Gln Gln
 295

ctggggccgt tggggagctg actctgcgct ttttgcag tgg atg gtc ccg gcc atc 7288
 Trp Met Val Pro Ala Ile
 300 305

cag aac tcc atg aag ccc ttc aag gtgagcaggc aggcctggca ggggtgggttc 7342
 Gln Asn Ser Met Lys Pro Phe Lys
 310

cgggggtcagg gctgagggag ccagctgtgc cctgtgccca cag gac atg gac tac 7397
 Asp Met Asp Tyr
 315

tcc cgc atc gtg gag cgc ctc ctg aag ctg gcg gtgagtgacc tgctgggtgg 7450
 Ser Arg Ile Val Glu Arg Leu Leu Lys Leu Ala
 320 325

ggacgcgtgg gggcgggtgg ggctgttctg gcacctggca cccactcccc acag gtc 7507
 Val

ccc aac cac ctc atc tgg ctc atc ttc ttc tac tgg ctc ttc cac tcc	7555
Pro Asn His Leu Ile Trp Leu Ile Phe Phe Tyr Trp Leu Phe His Ser	
330 335 340 345	
tgc ctg aac gcc gtg gct gag ctc atg cag ttt gga gac cgc gag ttc	7603
Cys Leu Asn Ala Val Ala Glu Leu Met Gln Phe Gly Asp Arg Glu Phe	
350 355 360	
tac cgg gac tgg tg gtgggtggcc ttgccggggc gggggtggtg ggggcccccg	7657
Tyr Arg Asp Trp Trp	
365	
ctggggctgg ggccggagcc cctgcccact ctgccccgcc cccgcag g aac tcc gag	7714
Asn Ser Glu	
tcc atc acc tac ttc tgg cag aac tgg aac atc cct gtt cac aag tgg	7762
Ser Ile Thr Tyr Phe Trp Gln Asn Trp Asn Ile Pro Val His Lys Trp	
370 375 380 385	
tgc atc ag gtgggtgtgc gcctgggggc ggggggttgg ggggtgggac	7810
Cys Ile Arg	
ggggtcgcgt ggcccgggcg cccagcccac tgccgcctcc cccgcag a cac ttc tac	7867
His Phe Tyr	
390	
aag ccc atg ctc cgg cgg ggc agc agc aag tgg gca gcc agg acg gca	7915
Lys Pro Met Leu Arg Arg Gly Ser Ser Lys Trp Ala Ala Arg Thr Ala	
395 400 405	
gtg ttt ctg gcc tcc gcc ttc ttc cac gag gtcagtgcac tgagggcgcg	7965
Val Phe Leu Ala Ser Ala Phe Phe His Glu	
410 415	
ccctgcccct ggtgggggtg ggggtggggg tgggggctcg ctgacgcccc tctcccctca	8025
g tac ctg gtg agc atc ccc ctg cgc atg ttc cgc ctc tgg gcc ttc acc	8074
Tyr Leu Val Ser Ile Pro Leu Arg Met Phe Arg Leu Trp Ala Phe Thr	
420 425 430	
ggc atg atg gcg cag gtgagcagcc ctggaccccc gctccgcccc gccccgcgag	8129
Gly Met Met Ala Gln	
435	
cgcagaggct cactcccgtc ctgtgtcccc ag atc ccg ctg gcc tgg ata gtg	8182
Ile Pro Leu Ala Trp Ile Val	
440 445	
ggc cgc ttc ttc cgc ggc aac tac ggc aac gcg gcc gtg tgg ctg tca	8230
Gly Arg Phe Phe Arg Gly Asn Tyr Gly Asn Ala Ala Val Trp Leu Ser	
450 455 460	
ctc atc atc ggg cag ccg gtg gcc gtc ctg atg tac gtc cac gac tac	8278
Leu Ile Ile Gly Gln Pro Val Ala Val Leu Met Tyr Val His Asp Tyr	
465 470 475	
tac gtg ctc aac cgt gag gcg ccg gca gcc ggc acc tgagcgcctc	8324
Tyr Val Leu Asn Arg Glu Ala Pro Ala Ala Gly Thr	
480 485	

caggctggcc	ccctcgtggg	tgttggactg	ctttgccgcg	ctgcctgcgg	ctggactaga	8384
gcctgcccc	acctgggccc	agcaggagga	ggcctggctg	gtggaagctg	cctcctggcc	8444
tccaccaggc	ctctgcctga	agggcttcc	cctgccaggg	gagagcaggc	ccgacgcagt	8504
tctggccct	gggaggtgcc	catgctctgg	aaaccctaca	gatctcgccc	aagggctctga	8564
atgtgtcaat	aaagtgtgtg	gcacagttag	ctccctcagc	ctccagggca	cagggctggc	8624
aggagggggc	ggccctccca	cgtggggcca	tgtgtggga	aggaggcccc	agcgcttggg	8684
gaggagctgg	ggctgtggtg	accctccctg	cctcacaggg	ctctgtggtc	agacgtcttg	8744
ccctgcaagg	tggagactcc	atgctccaag	gccccctgtg	cctgaggtct	gcacacaagt	8804
ggattcaact	tgggtcaggc	cagaggctaa	ggtgtggaag	aaggttgaga	atcaggctga	8864
cttgaacggc	agcaaagact	ccaaggcaag	gctgcagaag	tctcaaaagc	tatgcgcaca	8924
gtccctgct	ggggtgctca	cctgggctgg	gctctgggct	gcttggaaca	agcaggtggc	8984
ctggctcagc	cctcaccgag	ggccttcctt	gggggcagag	gcttgccctga	tgccaggggc	9044
tccccgtttt	tccaggccct	cagcaggtag	ttgggtgtgg	ccctcatgat	accttgggtcc	9104
cagagctccg	ccactcaaaa	agcttggcag	tgaggcaagg	gcaaccccgg	gctgttcccc	9164
cctctactgg	ctctgccgcc	tgggttggaa	accctgaggc	tgtgccaggc	aggtgtacct	9224
tgacagccag	ccatggccca	gtaagatggg	tgcccagaggt	ggtacctggg	cagcggacct	9284
agctgtgctg	ccccgcgcc	aaccagaagc	cgctctagcc	catggtgggtc	gtctgggcga	9344
gacaggctgg	ttggctaggg	actgtttggt	ctacagcagg	tgtaggcagc	gtctccctga	9404
cccctgcctt	ctaggaagcc	accaccctgn	gcctactca	tcagcaagga	cagcgagcag	9464
ggctgagctg	ggggtgcgtg	ggctgctacg	gnccccgcac	cttcatcaca	tgacacctctg	9524
cacccctgct	tgccctgactc	aagagtgggg	gggggggtcc	tgtgcttcct	tcattccaga	9584
cccacgggtg	tgaccagtg	cacccacctg	gtccgctagt	gctgacctgg	ccacagggct	9644
cctgtggggc	cacgctgatc	ccgccctggt	cccttcataa	agaactcttg	agcacatgca	9704
gcccagggga	gccaggaggc	tccagtgtgc	tgtgtccatc	tgccctccctt	cagcccttcc	9764
cgagacactg	cgcctcatgc	ccccctccac	ccccaccac	actggcagga	ggaacacaca	9824
gggagaccac	acacagagct	cgttgtttat	aaatctctgc	ctggctcatc	ggtctgtttg	9884
tccatgtata	tatctgcata	tctctatgga	aggggaaagg	gggactcgtg	taaaaatcca	9944
aaatacaatt	ctatgaacac	ctgcctcctg	gtcagctctga	gtgtggccgt	gaagcccagg	10004
tgagctgtgg	ctcacagggc	taggccctcg	gtgctggccg	ggggccactc	cccacccctt	10064
ctccccccct	ccgccagcca	ggggaccagg	ctcctggaca	ccaggcctgc	ccaaggcctg	10124
ctctcctcct	ggggcttcta	cgagacagtg	gggtccttgg	ctttgggggg	ttctgagccc	10184
gtcagcaggg	agatgggtgg	gtcatctctt	atatttcgtc	tccctcggag	aagtaggagc	10244
cctccccccag	ctcgaagagc	accggcagggt	cgctgctccc	cacgtccacg	gagcccggtg	10304
ccaggagcag	caggggctgg	gcgggtgtagt	gcaccagctg	cttccctagg	ggtgcgactg	10364
ggtcagggtg	ccggtggggc	cgggggggcg	ggtgggggnnn	nnnnnnnnnn	nnnnnnnnnn	10424
ccccggcccc	agccaccccg	cctacgcacg	ctggccaggc	tgctgtccag	gtcgggcagg	10484
ctcatgtcgg	gcaccgtaac	cgaggggctg	aacagctgca	gggaagaggg	gcgggtcaga	10544
ctgccctgga	agcggggcg	gctgggcgcg	gggcgggcag	ggttagactc	acatccagca	10604
gggtgctggt	gtccacgctg	aagccatggc	ttgtcagcat	ggtctgcagg	ttgtccagggt	10664
tggagtccat	ggcgtccaag	tggctcgtga	gctcggctct	ggccagagga	aggggagcag	10724
gtgacgtggc	atccaggcac	ccccagggtg	agtcctgccg	gcccttaagc	ccagctgcca	10784
gcctgccctg	cgcaggctca	cgggaggccc	tgagggtggg	ccctgggcct	gggacacttg	10844
gccaccgggtg	gtggctcagg	ctccctcatc	accgtgaggc	cccgttaaggc	agcacgtcag	10904
gctggggccg	ctaagacatc	agggggctga	gttcagggtc	cagaagggcc	cggggcccca	10964
gcttctcagg	ttcagcaacg	cctgcccttc	acactggagg	acagcaggag	ggtggagctg	11024
tgctcggcca	ggctgggcca	gcctacccca	gcctgcgcca	ggatagatta	ggcctgcccc	11084
ctcgaggagg	ccaggcagta	cctcgggggtg	ctgcctcagg	ccccagagggt	gggtggggag	11144
cctgttcaca	ggctgggcag	gggtctgctc	agaggctggg	gaggggcccc	agctctggcg	11204
gagccctcct	ccctccccag	tgggacagcg	ctaaccctgg	ctggactcgg	ccatgcagag	11264
aggaggaggg	gcaggggaag	aggcgggcga	ccccagacat	ctttggagtg	cgagccaaac	11324
tgcaagatag	aaagacaaga	gccccctctt	gcgggtgggg	cccagcgagt	ccaagcccc	11384
caccccgctt	acccacctct	ggggtggtag	gacagacaga	ccaggcgggc	cggggagggc	11444
ggctcagggg	gcaggggtgc	ccgccacgtg	cctgggtccc	acgggtgcagg	ggctgcggga	11504
ccccctggct	ggcatgtctc	cctccccagc	tccatgctct	gggccagggtc	ggccagtctt	11564
ccccagcaaa	tcccatgccg	agggcctgag	ggccatgcgt	gtccagcccc	gcactgtcct	11624
tgctcggagg	ccctgcgctt	ccagagaccg	aagggcgcct	ggaaggcact	cacttgtcta	11684
ggcaggcgac	gctgaggcac	ttctcgggag	ccgaggcggg	gagggtcgac	ggatagcggc	11744
cccagtgat	ccgatagaag	cttcgta				11771

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

Arg Ile Arg Lys Arg Phe Leu Leu Arg Arg Leu Leu Glu Met Leu Phe
 275 280 285
 Leu Thr Gln Leu Gln Val Gly Leu Ile Gln Gln Trp Met Val Pro Ala
 290 295 300
 Ile Gln Asn Ser Met Lys Pro Phe Lys Asp Met Asp Tyr Ser Arg Ile
 305 310 315 320
 Val Glu Arg Leu Leu Lys Leu Ala Val Pro Asn His Leu Ile Trp Leu
 325 330 335
 Ile Phe Phe Tyr Trp Leu Phe His Ser Cys Leu Asn Ala Val Ala Glu
 340 345 350
 Leu Met Gln Phe Gly Asp Arg Glu Phe Tyr Arg Asp Trp Trp Asn Ser
 355 360 365
 Glu Ser Ile Thr Tyr Phe Trp Gln Asn Trp Asn Ile Pro Val His Lys
 370 375 380
 Trp Cys Ile Arg His Phe Tyr Lys Pro Met Leu Arg Arg Gly Ser Ser
 385 390 395 400
 Lys Trp Ala Ala Arg Thr Ala Val Phe Leu Ala Ser Ala Phe Phe His
 405 410 415
 Glu Tyr Leu Val Ser Ile Pro Leu Arg Met Phe Arg Leu Trp Ala Phe
 420 425 430
 Thr Gly Met Met Ala Gln Ile Pro Leu Ala Trp Ile Val Gly Arg Phe
 435 440 445
 Phe Arg Gly Asn Tyr Gly Asn Ala Ala Val Trp Leu Ser Leu Ile Ile
 450 455 460
 Gly Gln Pro Val Ala Val Leu Met Tyr Val His Asp Tyr Tyr Val Leu
 465 470 475 480
 Asn Arg Glu Ala Pro Ala Ala Gly Thr
 485

<210> 3

<211> 31

<212> DNA

<213> Bos taurus

<220>

<221> misc_signal

<222> (29)..(31)

<223> bases 1 to 3 of the Kozak recognition sequence. See genomic sequence from start codon for bases 4 to 7 of the Kozak recognition sequence or the DGAT1 cDNA for the complete recognition sequence.

<400> 3

acttggccgc ggcggggtgc gaactaaggc c

31

<210> 4

<211> 1732

<212> DNA

<213> Bos taurus

<220>

<221> misc_feature

<222> (1)..(3)

<223> translation start codon

<220>

<221> misc_feature

<222> (1468)..(1470)

<223> translation stop codon

<220>
<221> variation
<222> (689)..(755)
<223> this sequence is deleted in the alternately spliced transcript

<220>
<221> variation
<222> (358)..(358)
<223> cytosine (C) to thymine (T) substitution polymorphism

<220>
<221> variation
<222> (694)..(695)
<223> adenine (A)-adenine (A) to guanine(G)-cytosine (C) substitution
polymorphism; AA corresponds to the Q allele
GC corresponds to the q allele

<220>
<221> primer_bind
<222> (616)..(635)
<223> Primer DgatforAD
TTCTCCTACCGGGACGTCAA

<220>
<221> primer_bind
<222> (743)..(767)
<223> Primer DgatrevAD
AAGTAGTAGAGATCGCGGTAGGTCA
reverse primer

<220>
<221> primer_bind
<222> (690)..(703)
<223> Primer ForAA (FAM)
CGTTGGCCTTCTTA
reverse primer

<220>
<221> primer_bind
<222> (688)..(709)
<223> Primer DgatADGC (VIC)
TTGGCCGCCTTACC
reverse primer

<220>
<221> primer_bind
<222> (617)..(636)
<223> Primer DgatforRT66
TCTCCTACCGGGACGTCAAC

<220>
<221> primer_bind
<222> (739)..(760)
<223> Primer DgatrevRT66
GAGATCGCGGTAGGTCAGGTT
reverse primer

<220>
<221> primer_bind

```

<222> (676)..(767)
<223> Primer DgatforRTless66 GCTGCTTTGGCAGATCTCTACTACTT

<220>
<221> primer_bind
<222> (676)..(767)
<223> This primer selectively binds and amplifies a contiguous sequence
characteristic of the splice variant & generated by the deletion
bases 689 to 755 of this sequence.

<220>
<221> primer_bind
<222> (817)..(833)
<223> Primer DgatrevRTless66
AAGCGCTTTCGGATGCG

<220>
<221> primer_bind
<222> (722)..(734)
<223> Primer Dgatwith66 (FAM)
CCGTGAGCTACCC

<220>
<221> primer_bind
<222> (771)..(785)
<223> Primer Dgatless66 (VIC)
CTTCGCCCCCACCCT

<400> 4
atgggcgacc gcggcgggcgc gggcgggctcc cggcgccgga ggacgggggtc gcggccttcg 60
atccagggcg gcagtgggcc cgcggcagcg gaagaggagg tgccgggatgt gggcgccgga 120
ggggacgcgc cggtcggga cacagacaag gacggagacg tagacgtggg cagcggccac 180
tgggacctga ggtgtcaccg cctgcaggat tccctgttca gttctgacag tggcttcagc 240
aactaccgtg gcatacctgaa ttggtgtgtg gtgatgctga tcttaagcaa cgcacgggta 300
tttctagaga acctcatcaa gtatggcatc ctggtggacc ccatccagggt ggtgtctctg 360
ttcctgaagg acccctacag ctggccagct ctgtgcctgg tcattgtggc caatatcttt 420
gccgtggctg cgttccagggt ggagaagcgc ctggccgtgg gagctctgac ggagcaggcg 480
gggctgctgc tgcacgggggt caacctggcc accattctct gcttcccagc ggccgtggcc 540
tttctcctcg agtctatcac tccagtgggc tccgtgctgg ccctgatggg ctacaccatc 600
ctcttcctca agctgttctc ctaccgggac gtcaacctct ggtgccgaga gcgcagggct 660
ggggccaagg ccaaggctgc tttggcagggt aagaaggcca acggggggagc tgcccagcgc 720
accgtgagct accccgacaa cctgacctac cgcgatctct actacttcct cttcgcccc 780
accctgtgct acgagctcaa cttccccgcg tcccccgca tccgaaagcg cttcctgctg 840
cggcgactcc tggagatgct gttcctcacc cagctccagg tggggctgat ccagcagtgg 900
atggtcccg gcatccagaa ctccatgaag cccttcaagg acatggacta ctcccgcatc 960
gtggagcgcc tcctgaagct ggcgggtccc aaccacctca tctggctcat cttcttctac 1020
tggtctctcc actcctgcct gaacgccgtg gctgagctca tgcagtgttg agaccgcgag 1080
ttctaccggg actggtggaa ctccgagtc atcacctact tctggcagaa ctggaacatc 1140
cctgttcaca agtgggtgcat cagacacttc tacaagccca tgctccggcg gggcagcagc 1200
aagtgggcag ccaggacggc agtgtttctg gcctccgcct tcttccacga gtacctgggtg 1260
agcatcccc tgcgcatggt ccgcctctgg gccttcaccg gcatgatggc gcagatcccg 1320
ctggcctgga tagtgggccg cttcttccgc ggcaactacg gcaacgcggc cgtgtggctg 1380
tactcatca tcgggcagcc ggtggccgct ctgatgtacg tccacgacta ctacgtgctc 1440
aaccgtgagg cgccggcagc cggcacctga ggcctccag gctggcccc tcgtgggtgt 1500
tggaactgct tgccgcgctg cctgcggctg gactagagcc tgccccaacc tgggcgcagc 1560
aggaggaggc ctggctggtg gaagctgcct cctggcctcc accaggcctc tgcctgaagg 1620
gcttcctcct gccaggggag agcaggcccg acgcagttct ggccccctgg aggtgccccat 1680
gctctgga aa ccctacagat ctgcaccaag ggtctgaatg tgtcaataaa ga 1732

```

<210> 5
 <211> 489
 <212> PRT
 <213> Bos taurus

<220>
 <221> VARIANT
 <222> (232)..(232)
 <223> an amino acid substitution (K -> A) caused by a polymorphism at bases 6289 to 6830 of the genomic sequence (measured from the adenine residue of the translation start codon).

<220>
 <221> VARIANT
 <222> (232)..(232)
 <223> Lysine (K) corresponds to the Q allele, alanine (A) corresponds to the q allele.

<400> 5
 Met Gly Asp Arg Gly Gly Ala Gly Gly Ser Arg Arg Arg Arg Thr Gly
 1 5 10 15
 Ser Arg Pro Ser Ile Gln Gly Gly Ser Gly Pro Ala Ala Ala Glu Glu
 20 25 30
 Glu Val Arg Asp Val Gly Ala Gly Gly Asp Ala Pro Val Arg Asp Thr
 35 40 45
 Asp Lys Asp Gly Asp Val Asp Val Gly Ser Gly His Trp Asp Leu Arg
 50 55 60
 Cys His Arg Leu Gln Asp Ser Leu Phe Ser Ser Asp Ser Gly Phe Ser
 65 70 75 80
 Asn Tyr Arg Gly Ile Leu Asn Trp Cys Val Val Met Leu Ile Leu Ser
 85 90 95
 Asn Ala Arg Leu Phe Leu Glu Asn Leu Ile Lys Tyr Gly Ile Leu Val
 100 105 110
 Asp Pro Ile Gln Val Val Ser Leu Phe Leu Lys Asp Pro Tyr Ser Trp
 115 120 125
 Pro Ala Leu Cys Leu Val Ile Val Ala Asn Ile Phe Ala Val Ala Ala
 130 135 140
 Phe Gln Val Glu Lys Arg Leu Ala Val Gly Ala Leu Thr Glu Gln Ala
 145 150 155 160
 Gly Leu Leu Leu His Gly Val Asn Leu Ala Thr Ile Leu Cys Phe Pro
 165 170 175
 Ala Ala Val Ala Phe Leu Leu Glu Ser Ile Thr Pro Val Gly Ser Val
 180 185 190
 Leu Ala Leu Met Val Tyr Thr Ile Leu Phe Leu Lys Leu Phe Ser Tyr
 195 200 205
 Arg Asp Val Asn Leu Trp Cys Arg Glu Arg Arg Ala Gly Ala Lys Ala
 210 215 220
 Lys Ala Ala Leu Ala Gly Lys Lys Ala Asn Gly Gly Ala Ala Gln Arg
 225 230 235 240
 Thr Val Ser Tyr Pro Asp Asn Leu Thr Tyr Arg Asp Leu Tyr Tyr Phe
 245 250 255
 Leu Phe Ala Pro Thr Leu Cys Tyr Glu Leu Asn Phe Pro Arg Ser Pro
 260 265 270
 Arg Ile Arg Lys Arg Phe Leu Leu Arg Arg Leu Leu Glu Met Leu Phe
 275 280 285
 Leu Thr Gln Leu Gln Val Gly Leu Ile Gln Gln Trp Met Val Pro Ala
 290 295 300
 Ile Gln Asn Ser Met Lys Pro Phe Lys Asp Met Asp Tyr Ser Arg Ile
 305 310 315 320

66

Val	Glu	Arg	Leu	Leu	Lys	Leu	Ala	Val	Pro	Asn	His	Leu	Ile	Trp	Leu
				325					330					335	
Ile	Phe	Phe	Tyr	Trp	Leu	Phe	His	Ser	Cys	Leu	Asn	Ala	Val	Ala	Glu
			340					345					350		
Leu	Met	Gln	Phe	Gly	Asp	Arg	Glu	Phe	Tyr	Arg	Asp	Trp	Trp	Asn	Ser
		355					360					365			
Glu	Ser	Ile	Thr	Tyr	Phe	Trp	Gln	Asn	Trp	Asn	Ile	Pro	Val	His	Lys
	370					375					380				
Trp	Cys	Ile	Arg	His	Phe	Tyr	Lys	Pro	Met	Leu	Arg	Arg	Gly	Ser	Ser
385					390					395					400
Lys	Trp	Ala	Ala	Arg	Thr	Ala	Val	Phe	Leu	Ala	Ser	Ala	Phe	Phe	His
				405				410						415	
Glu	Tyr	Leu	Val	Ser	Ile	Pro	Leu	Arg	Met	Phe	Arg	Leu	Trp	Ala	Phe
			420					425					430		
Thr	Gly	Met	Met	Ala	Gln	Ile	Pro	Leu	Ala	Trp	Ile	Val	Gly	Arg	Phe
		435					440					445			
Phe	Arg	Gly	Asn	Tyr	Gly	Asn	Ala	Ala	Val	Trp	Leu	Ser	Leu	Ile	Ile
	450					455					460				
Gly	Gln	Pro	Val	Ala	Val	Leu	Met	Tyr	Val	His	Asp	Tyr	Tyr	Val	Leu
465					470					475					480
Asn	Arg	Glu	Ala	Pro	Ala	Ala	Gly	Thr							
				485											

<210> 6
 <211> 467
 <212> PRT
 <213> Bos taurus

<400> 6

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

67

Cys Tyr Glu Leu Asn Phe Pro Arg Ser Pro Arg Ile Arg Lys Arg Phe
 245 250 255
 Leu Leu Arg Arg Leu Leu Glu Met Leu Phe Leu Thr Gln Leu Gln Val
 260 265 270
 Gly Leu Ile Gln Gln Trp Met Val Pro Ala Ile Gln Asn Ser Met Lys
 275 280 285
 Pro Phe Lys Asp Met Asp Tyr Ser Arg Ile Val Glu Arg Leu Leu Lys
 290 295 300
 Leu Ala Val Pro Asn His Leu Ile Trp Leu Ile Phe Phe Tyr Trp Leu
 305 310 315 320
 Phe His Ser Cys Leu Asn Ala Val Ala Glu Leu Met Gln Phe Gly Asp
 325 330 335
 Arg Glu Phe Tyr Arg Asp Trp Trp Asn Ser Glu Ser Ile Thr Tyr Phe
 340 345 350
 Trp Gln Asn Trp Asn Ile Pro Val His Lys Trp Cys Ile Arg His Phe
 355 360 365
 Tyr Lys Pro Met Leu Arg Arg Gly Ser Ser Lys Trp Ala Ala Arg Thr
 370 375 380
 Ala Val Phe Leu Ala Ser Ala Phe Phe His Glu Tyr Leu Val Ser Ile
 385 390 395 400
 Pro Leu Arg Met Phe Arg Leu Trp Ala Phe Thr Gly Met Met Ala Gln
 405 410 415
 Ile Pro Leu Ala Trp Ile Val Gly Arg Phe Phe Arg Gly Asn Tyr Gly
 420 425 430
 Asn Ala Ala Val Trp Leu Ser Leu Ile Ile Gly Gln Pro Val Ala Val
 435 440 445
 Leu Met Tyr Val His Asp Tyr Tyr Val Leu Asn Arg Glu Ala Pro Ala
 450 455 460
 Ala Gly Thr
 465

<210> 7
 <211> 12
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (7)..(7)
 <223> thymidine (T) to guanine (G) substitution polymorphism

<400> 7
 cagtgctagg gg

12

<210> 8
 <211> 10
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (6)..(6)
 <223> thymidine (T) to guanine (G) substitution polymorphism

<400> 8
 gcattgcgct

10

68

<210> 9
 <211> 11
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (6)..(6)
 <223> thymidine (T) to cytosine (G) substitution polymorphism

<400> 9
 taccctggga c 11

<210> 10
 <211> 11
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (6)..(6)
 <223> adenine (A) to guanine (G) substitution polymorphism

<220>
 <221> variation
 <222> (6)..(6)
 <223> adenine (A) to guanine (G) substitution polymorphism

<400> 10
 ctcttagcag c 11

<210> 11
 <211> 10
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (5)..(5)
 <223> guanine (G) to adenine (A) substitution polymorphism

<400> 11
 acaggcaact 10

<210> 12
 <211> 11
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (6)..(6)
 <223> cytosine (C) to thymine (T) substitution polymorphism

<400> 12
 tgtctctgtt c 11

<210> 13
 <211> 14
 <212> DNA
 <213> Bos taurus

 <220>
 <221> variation
 <222> (7)..(8)
 <223> adenine (A)-adenine (A) to guanine(G)-cytosine (C) substitution
 polymorphism; AA corresponds to the Q allele
 GC corresponds to the q allele

 <400> 13
 ggtaagaagg ccaa 14

 <210> 14
 <211> 11
 <212> DNA
 <213> Bos taurus

 <220>
 <221> variation
 <222> (6)..(6)
 <223> guanine (G) to adenine (A) substitution polymorphism

 <400> 14
 gcggtgagga t 11

 <210> 15
 <211> 16
 <212> DNA
 <213> Bos taurus

 <220>
 <221> variation
 <222> (6)..(7)
 <223> guanine (G)-guanine (G) to adenine (A)-cytosine (C) substitution
 polymorphism
 GG-AC

 <400> 15
 gggggggggg gactct 16

 <210> 16
 <211> 11
 <212> DNA
 <213> Bos taurus

 <220>
 <221> variation
 <222> (6)..(6)
 <223> adenine (A) to guanine (G) substitution polymorphism

 <400> 16
 gagtgacctg c 11

<210> 17
 <211> 11
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (6)..(6)
 <223> cytosine (C) to thymidine (T) substitution polymorphism

<400> 17
 ggacgcgtgg g 11

<210> 18'
 <211> 11
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (6)..(6)
 <223> guanine (G) to adenine (A) substitution polymorphism

<400> 18
 ggtgggggtg g 11

<210> 19
 <211> 11
 <212> DNA
 <213> Bos taurus

<220>
 <221> variation
 <222> (6)..(6)
 <223> ctosine (C) to thymidine (T) substitution polymorphism

<400> 19
 ctgggcgcag c 11

<210> 20
 <211> 14
 <212> DNA
 <213> Bos taurus

<400> 20
 cgttggcctt cttta 14

<210> 21
 <211> 14
 <212> DNA
 <213> Bos taurus

<400> 21
 ttggccgcct tacc 14

<210> 22	
<211> 12	
<212> DNA	
<213> Bos taurus	
<400> 22	
cagtgctagg gg	12
<210> 23	
<211> 12	
<212> DNA	
<213> Bos taurus	
<400> 23	
cagtgcgagg gg	12
<210> 24	
<211> 10	
<212> DNA	
<213> Bos taurus	
<400> 24	
gcattgcgct	10
<210> 25	
<211> 10	
<212> DNA	
<213> Bos taurus	
<400> 25	
gcatggcgct	10
<210> 26	
<211> 11	
<212> DNA	
<213> Bos taurus	
<400> 26	
taccctggga c	11
<210> 27	
<211> 11	
<212> DNA	
<213> Bos taurus	
<400> 27	
taccccggga c	11
<210> 28	
<211> 11	
<212> DNA	
<213> Bos taurus	

<400> 28
ctcttagcag c 11

<210> 29
<211> 11
<212> DNA
<213> Bos taurus

<400> 29
ctcttggcag c 11

<210> 30
<211> 10
<212> DNA
<213> Bos taurus

<400> 30
acaggcaact 10

<210> 31
<211> 10
<212> DNA
<213> Bos taurus

<400> 31
acagacaact 10

<210> 32
<211> 11
<212> DNA
<213> Bos taurus

<400> 32
tgtctctgtt c 11

<210> 33
<211> 11
<212> DNA
<213> Bos taurus

<400> 33
tgtctttgtt c 11

<210> 34
<211> 14
<212> DNA
<213> Bos taurus

<400> 34
ggtaagaagg ccaa 14

<210> 35
<211> 14

<212> DNA	
<213> Bos taurus	
<400> 35	
ggtaaggcgg ccaa	14
<210> 36	
<211> 11	
<212> DNA	
<213> Bos taurus	
<400> 36	
gcggtgagga t	11
<210> 37	
<211> 11	
<212> DNA	
<213> Bos taurus	
<400> 37	
gcggtgagga t	11
<210> 38	
<211> 16	
<212> DNA	
<213> Bos taurus	
<400> 38	
ggggggggggg gactct	16
<210> 39	
<211> 16	
<212> DNA	
<213> Bos taurus	
<400> 39	
gggggacggg gactct	16
<210> 40	
<211> 11	
<212> DNA	
<213> Bos taurus	
<400> 40	
gagtgacctg c	11
<210> 41	
<211> 11	
<212> DNA	
<213> Bos taurus	
<400> 41	
gagtgacctg c	11

<210> 42
<211> 11
<212> DNA
<213> Bos taurus

<400> 42
ggacgcgtgg g

11

<210> 43
<211> 11
<212> DNA
<213> Bos taurus

<400> 43
ggacgtgtgg g

11

<210> 44
<211> 11
<212> DNA
<213> Bos taurus

<400> 44
ggtgggggtg g

11

<210> 45
<211> 11
<212> DNA
<213> Bos taurus

<400> 45
ggtggaggtg g

11

<210> 46
<211> 11
<212> DNA
<213> Bos taurus

<400> 46
ctgggcgcag c

11

<210> 47
<211> 11
<212> DNA
<213> Bos taurus

<400> 47
ctgggtgcag c

11

<210> 48
<211> 23
<212> DNA
<213> Bos taurus

<400> 48
cctgagcttg cctctccac agt 23

<210> 49
<211> 23
<212> DNA
<213> Bos taurus

<400> 49
ccaggagtcg ccgcagcagg aag 23

<210> 50
<211> 24
<212> DNA
<213> Bos taurus

<400> 50
ccggccatcc agaactccat gaag 24

<210> 51
<211> 21
<212> DNA
<213> Bos taurus

<400> 51
tagaactcgc ggtctccaaa c 21

<210> 52
<211> 23
<212> DNA
<213> Bos taurus

<400> 52
tggctgtcac tcatcatcgg gca 23

<210> 53
<211> 23
<212> DNA
<213> Bos taurus

<400> 53
ttgcacagca ctttattgac aca 23

<210> 54
<211> 18
<212> DNA
<213> Bos taurus

<400> 54
agctttggca ggtaaggc 18

<210> 55
<211> 18

76

<212> DNA
<213> Bos taurus

<400> 55
agcttttggca ggtaagaa

18

<210> 56
<211> 14
<212> DNA
<213> Bos taurus

<400> 56
ggccaacggg ggag

14

<210> 57
<211> 15
<212> DNA
<213> Bos taurus

<400> 57
gctggcggtg agtga

15

<210> 58
<211> 15
<212> DNA
<213> Bos taurus

<400> 58
gctggcggtg agtgg

15

<210> 59
<211> 15
<212> DNA
<213> Bos taurus

<400> 59
cctgctgggt gggga

15

<210> 60
<211> 15
<212> DNA
<213> Bos taurus

<400> 60
gctgggtggg gacgc

15

<210> 61
<211> 15
<212> DNA
<213> Bos taurus

<400> 61
gctgggtggg gacgt

15

<210> 62
<211> 14
<212> DNA
<213> Bos taurus

<400> 62
gtgggggcggtgg

14

<210> 63
<211> 15
<212> DNA
<213> Bos taurus

<400> 63
tgccccaacc tgggt

15

<210> 64
<211> 15
<212> DNA
<213> Bos taurus

<400> 64
tgccccaacc tgggc

15

<210> 65
<211> 15
<212> DNA
<213> Bos taurus

<400> 65
gcagcaggag gaggc

15

<210> 66
<211> 20
<212> DNA
<213> Bos taurus

<400> 66
gtagctttgg caggtaagaa

20

<210> 67
<211> 20
<212> DNA
<213> Bos taurus

<400> 67
ggggcgaaga ggaagtagta

20

<210> 68
<211> 20
<212> DNA
<213> Bos taurus

78

<400> 68
tggccctgat ggtctacacc

20

<210> 69
<211> 22
<212> DNA
<213> Bos taurus

<400> 69
gggcagctcc cccgttggcc gc

22

<210> 70
<211> 400
<212> DNA
<213> Bos taurus

<400> 70
tgcctctccc acagtgggct ccgtgctggc cctgatggtc tacaccatcc tcttcctcaa 60
gctgttctcc taccgggacg tcaacctctg gtgccgagag cgcagggctg gggccaaggc 120
caaggctggt gagggctgcc tcgggctggg gccactgggc tgccacttgc ctggggaccg 180
gcaggggctc ggctcacccc cgaccgcgcc cctgccgctt gctcgtagct ttggcaggta 240
agaaggccaa cgggggagct gccacgcgca ccgtgagcta ccccgacaac ctgacctacc 300
gcggtgagga tcctgccggg ggctgggggg actgcccggc ggcttggcct gctagccccg 360
ccctcccttc cagatctcta ctacttcctc ttcgccccca 400

<210> 71
<211> 20
<212> DNA
<213> Bos taurus

<400> 71
ttctcctacc gggacgtcaa

20

<210> 72
<211> 20
<212> DNA
<213> Bos taurus

<400> 72
ccgcggtagg tcaggttgct

20

<210> 73
<211> 14
<212> DNA
<213> Bos taurus

<400> 73
cgttggcctt ctta

14

<210> 74
<211> 14
<212> DNA
<213> Bos taurus

79

<400> 74
ttggccgcct tacc

14

<210> 75
<211> 400
<212> DNA
<213> Bos taurus

<400> 75
tgcctctccc acagtgggct ccgtgctggc cctgatggtc tacaccatcc tcttcctcaa 60
gctgttctcc taccgggacg tcaacctctg gtgccgagag cgcagggctg gggccaaggc 120
caaggctggg gagggctgcc tcgggctggg gccactgggc tgccacttgc ctcgggaccg 180
gcaggggctc ggctcacccc cgacccgccc cctgccgctt gctcgtagct ttggcaggta 240
agaaggccaa cgggggagct gcccagcgca ccgtgagcta ccccgacaac ctgacctacc 300
gcggtgagga tcctgccggg ggctgggggg actgcccggc ggccctggcct gctagccccg 360
ccctcccttc cagatctcta ctacttctc ttcgccccca 400

<210> 76
<211> 20
<212> DNA
<213> Bos taurus

<400> 76
tctcctaccg ggacgtcaac

20

<210> 77
<211> 21
<212> DNA
<213> Bos taurus

<400> 77
gagatcgcgg taggtcagggt t

21

<210> 78
<211> 26
<212> DNA
<213> Bos taurus

<400> 78
gctgctttgg cagatctcta ctactt

26

<210> 79
<211> 17
<212> DNA
<213> Bos taurus

<400> 79
aagcgctttc ggatgcg

17

<210> 80
<211> 13
<212> DNA
<213> Bos taurus

80

<400> 80
ccgtgagcta ccc

13

<210> 81
<211> 15
<212> DNA
<213> Bos taurus

<400> 81
cttcgcccc accct

15

<210> 82
<211> 50
<212> DNA
<213> Bos taurus

<400> 82
cttcaaggac atggactact cccgcatcgt ggagcgctc ctgaagctgg

50

68348-71

81

CLAIMS:

1. A method of identifying a bovine, or a sample derived from said bovine, with a genotype indicative of increased milk volume production or increased milk protein to fat ratio in said bovine, the method comprising determining the presence of alanine (A) at amino acid position 232 in the diacylglycerol acyltransferase1 (DGAT1) protein of SEQ ID NO:2, wherein the presence of alanine is determined with respect to DNA, mRNA or protein from the bovine.
2. The method of claim 1, wherein the presence of alanine (A) is determined by detecting the presence of nucleotides G and C at positions 6829 and 6830 respectively in DGAT1 gene of SEQ ID NO:1.
3. The method of claim 2, wherein detection of the presence of the nucleotides G and C is made via amplification of a nucleic acid sequence comprising the nucleotides.
4. The method of claim 3, wherein a primer consisting of SEQ ID NO: 21 is used in the amplification.
5. The method of claim 2, wherein detection of the presence of the nucleotides G and C is made using an oligonucleotide ligation assay (OLA).
6. The method of claim 5, wherein the OLA is performed using a primer comprising the sequence of any one of SEQ ID NOs: 54 and 56.
7. The method of claim 2 in which detection of the nucleotide G and C is made via hybridisation of a probe comprising a sequence complementary to at least 5 contiguous nucleotides of the sequence or complement of SEQ ID NO:1 that comprises said G and C, wherein the probe is capable of hybridisation to said sequence or complement of SEQ ID NO:1 in 6x sodium citrate/sodium chloride (SSC) at 45°C.
8. A method of selecting a bovine that possesses a genotype indicative of increased milk volume production or increased milk protein to fat ratio in said bovine, the method comprising the step of selecting a bovine identified by the method of any one of claims 1 to 7.

68348-71

82

9. A method of identifying a bovine, or a sample derived from said bovine, with a genotype indicative of decreased milk volume production or decreased milk protein to fat ratio in said bovine, the method comprising determining the presence of lysine (K) at amino acid position 232 in the diacylglycerol acyltransferase1 (DGAT1) protein of SEQ ID NO:2, wherein the presence of lysine is determined with respect to DNA, mRNA or protein from the bovine.
10. The method of claim 9, wherein the presence of lysine (K) is determined by detecting the presence of nucleotides A and A at positions 6829 and 6830 respectively in DGAT1 gene of SEQ ID NO:1.
11. The method of claim 10, wherein detection of the presence of the nucleotides A and A is made via amplification of a nucleic acid sequence comprising the nucleotides.
12. The method of claim 11, wherein primers consisting of SEQ ID NOs: 20 and 21 are used in the amplification.
13. The method of claim 10, wherein detection of the presence of the nucleotides A and A is made using an oligonucleotide ligation assay (OLA).
14. The method of claim 13, wherein the OLA is performed using a primer comprising the sequence of any one of SEQ ID NOs: 55 and 56.
15. The method of claim 10 in which detection of the nucleotide A and A is made via hybridisation of a probe comprising a sequence complementary to at least 5 contiguous nucleotides of the sequence or complement of SEQ ID NO:1 that comprises said A and A, wherein the probe is capable of hybridisation to said sequence or complement of SEQ ID NO:1 in 6x sodium citrate/sodium chloride (SSC) at 45°C.
16. A method of selecting a bovine that possesses a genotype indicative of decreased milk volume production or decreased milk protein to fat ratio in said bovine, the method comprising the step of selecting a bovine identified by the method of any one of claims 10 to 15.

68348-71

83

17. A primer comprising a nucleotide sequence having at least 12 contiguous bases of SEQ ID NO: 1 or 4 and further comprising the K232A (AA to GC) polymorphism of SEQ ID NO: 13 associated with improved milk production traits.
18. Use of the DGAT1 gene sequence of SEQ ID NO: 1 or 4, comprising the
5 K232A (AA to GC) polymorphism of SEQ ID NO: 13 in the identification of one or more molecular DNA markers useful in the method of any one of claims 1 or 9.
19. Use of one or more polymorphic sequences selected from the group consisting of SEQ ID NOs: 7 to 19 in the method of any one of claims 1 to 9.
20. A kit, comprising the primer of claim 17 or a probe selected from any
10 5 or more contiguous nucleotides of the DGAT1 sequence of SEQ ID NO: 1 or 4 comprising the K232A (AA to GC) polymorphism of SEQ ID NO: 13 which is sufficiently complementary with said nucleic acid sequence so as to bind thereto in 6x sodium citrate/sodium chloride at 45°C, and instructions for using the primer or the probe for genotyping a bovine with respect to milk composition and volume associated with
15 DGAT1.
21. An isolated nucleic acid molecule having the sequence of SEQ ID NOs: 1 or 4 comprising the K232A (AA to GC) polymorphism of SEQ ID NO: 13 wherein said polymorphism is associated with improved milk production traits.
- 20 22. An isolated nucleic acid molecule comprising the whole or at least 20 nucleotides of the nucleotide sequence of SEQ ID NO: 1 or 4 or which varies from the sequence due to the degeneracy of the genetic code, or a nucleic acid strand capable of hybridising with said nucleic acid molecule in 6x sodium citrate/sodium chloride at 45°C, wherein said nucleic acid molecule further comprises the K232A (AA to GC)
25 polymorphism of SEQ ID NO: 13 which is associated with improved milk production traits.
23. An isolated mRNA transcribed from DNA having a sequence which corresponds to the nucleic acid molecule as claimed in claim 21 or 22.

68348-71

84

24. A recombinant cloning vector comprising the nucleic acid molecule of claim 21 or 22.

25. A prokaryotic or eukaryotic cell containing the cloning vector of claim 24.

5 26. A transfected cell line which expresses a protein encoded by the nucleic acid molecule of claim 21 or 22.

27. A diagnostic kit useful in detecting the DNA molecule as claimed in claim 21 or 22, comprising a first and second primer for amplifying the DNA, the primers being complementary to nucleotide sequences of the DNA upstream and
10 downstream, respectively, of the K232A (AA to GC) polymorphism of SEQ ID NO: 13 in the portion of the DNA encoding DGAT1 protein of SEQ ID NO: 2 which results in altered relative milk lipid and protein production and milk volume, wherein at least one of the nucleotide sequences is selected to be from a non-coding region of the DGAT1 gene of SEQ ID NO: 1.

15 28. The kit according to claim 27, further comprising a third primer complementary to a polymorphism selected from the group consisting of SEQ ID NO: 7 to 19 located on the DGAT1 gene of SEQ ID NO: 1.

29. A process for producing a protein encoded by the nucleic acid molecule of claim 21 or 22, comprising the steps:

20 a) preparing a DNA fragment including a nucleotide sequence which encodes the protein;

b) incorporating the DNA fragment into an expression vector to obtain a recombinant DNA molecule which includes the DNA fragment and is capable of undergoing replication;

25 c) transforming a host cell with the recombinant DNA molecule to produce a transformant which can express the protein;

d) culturing the transformant to produce the protein; and

e) recovering the protein from resulting cultured mixture.

68348-71

85

30. A purified protein having an amino acid sequence encoded by the isolated nucleic acid molecule of claims 21 or 22, having diacylglycerol acyltransferase activity.

31. An antibody raised against the protein as claimed in claim 30.

5 32. Use of a transgenic non-human animal having cells that contain a transgene, which transgene is introduced into the animal or an ancestor of the animal at a prenatal stage, wherein said transgene comprises a bovine cDNA comprising:

10 a) the nucleotide sequence of SEQ ID NO: 4 including the K232A (AA to GC) polymorphism of SEQ ID NO: 13;

b) a sequence capable of hybridising to SEQ ID NO: 4 in 6x sodium chloride/sodium citrate (SSC) at 45°C, which includes the K232A (AA to GC) polymorphism of SEQ ID NO: 13; or

15 c) a sequence which varies from SEQ ID NO: 4 due to the degeneracy of the genetic code and which includes the K232A polymorphism;

to produce milk.

33. The use of claim 32, wherein the endogenous DGAT1 gene has been knocked out and replaced with the transgene.

1/34

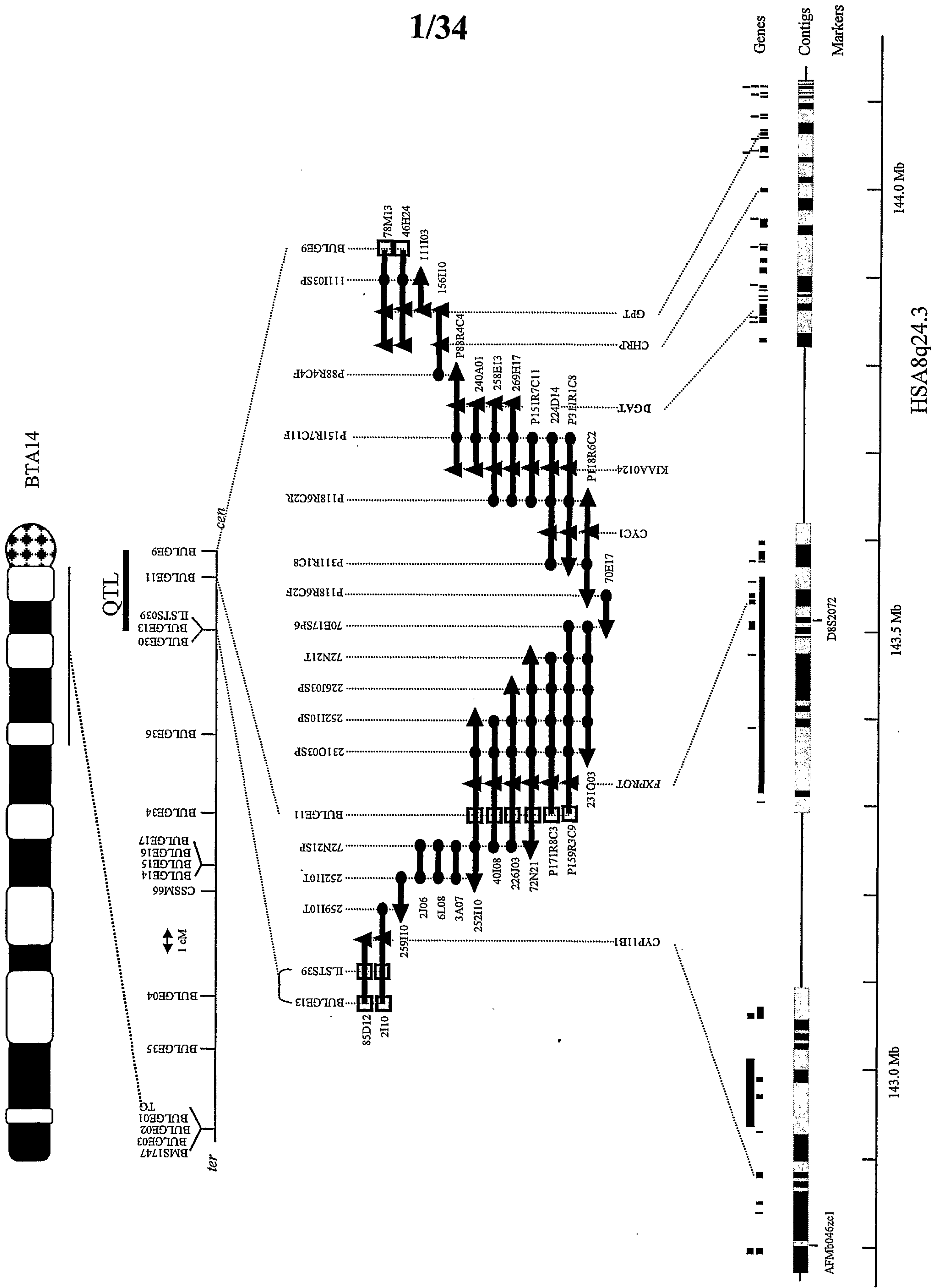


FIGURE 1

2/34

<222> (29)..(31)

<223> bases 1 to 3 of the Kozak recognition sequence. See the genomic sequence from the start codon for bases 4 to 7 of the Kozak recognition sequence or the DGAT1 cDNA for the complete recognition sequence.

acttggccgc gccggggtgc gaactaaggc c

FIGURE 2a

3/34

<210> 1

<211> 11771

<212> DNA

<213> Bos taurus

<220>

<221> CDS

<222> (1)..(191)

<223> Exon 1 CDS, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> misc_signal

<222> (1)..(4)

<223> these bases correspond to bases 4 to 7 of the Kozak recognition sequence. See DGAT1 cDNA for the complete recognition sequence.

<220>

<221> CDS

<222> (3809)..(3896)

<223> Exon 2, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (5840)..(5880)

<223> Exon 3, determined by alignment with an amino acid sequence deduced from the cDNA

FIGURE 2b

4/34

<220>

<221> CDS

<222> (5960)..(6045)

<223> Exon 4, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (6138)..(6190)

<223> Exon 5, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (6406)..(6511)

<223> Exon 6, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (6601)..(6714)

<223> Exon 7, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (6815)..(6889)

<223> Exon 8, determined by alignment with an amino acid sequence deduced from the cDNA

FIGURE 2b continued

5/34

<220>

<221> variation

<222> (6824)..(6889)

<223> this sequence is deleted from the alternately spliced transcript.

<220>

<221> CDS

<222> (6960)..(7063)

<223> Exon 9, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (7154)..(7192)

<223> Exon 10, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (7271)..(7312)

<223> Exon 11, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (7386)..(7430)

<223> Exon 12, determined by alignment with an amino acid sequence deduced from the cDNA

FIGURE 2b continued

6/34

<220>

<221> CDS

<222> (7505)..(7617)

<223> Exon 13, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (7705)..(7770)

<223> Exon 14, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (7858)..(7945)

<223> Exon 15, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (8027)..(8089)

<223> Exon 16, determined by alignment with an amino acid sequence deduced from the cDNA

<220>

<221> CDS

<222> (8162)..(8314)

<223> Exon 17 CDS, determined by alignment with an amino acid sequence deduced from the cDNA

FIGURE 2b continued

7/34

<220>

<221> misc_feature

<222> (8315)..(8317)

<223> translation stop codon

<220>

<221> polyA_site

<222> (8572)..(8578)

<223>

<220>

<221> variation

<222> (626)..(626)

<223> thymidine (T) to guanine (G) substitution polymorphism

<220>

<221> variation

<222> (3512)..(3512)

<223> thymidine (T) to guanine (G) substitution polymorphism

<220>

<221> variation

<222> (4040)..(4040)

<223> thymidine (T) to cytosine (G) substitution polymorphism

<220>

<221> variation

<222> (4963)..(4963)

<223> adenine (A) to guanine (G) substitution polymorphism

FIGURE 2b continued

8/34

<220>

<221> variation

<222> (5003)..(5003)

<223> guanine (G) to adenine (A) substitution polymorphism

<220>

<221> variation

<222> (5997)..(5997)

<223> cytosine (C) to thymine (T) substitution polymorphism

<220>

<221> variation

<222> (6829)..(6830)

<223> adenine (A)-adenine (A) to guanine(G)-cytosine (C) substitution polymorphism

AA corresponds to the Q allele

GC corresponds to the q allele

<220>

<221> variation

<222> (6892)..(6892)

<223> guanine (G) to adenine (A) substitution polymorphism

<220>

<221> variation

<222> (7224)..(7225)

<223> guanine (G)-guanine (G) to adenine (A)-cytosine (C) substitution polymorphism

GG-AC

FIGURE 2b continued

9/34

<220>

<221> variation

<222> (7438)..(7438)

<223> adenine (A) to guanine (G) substitution polymorphism

<220>

<221> variation

<222> (7456)..(7456)

<223> cytosine (C) to thymidine (T) substitution polymorphism

<220>

<221> variation

<222> (7987)..(7987)

<223> guanine (G) to adenine (A) substitution polymorphism

<220>

<221> variation

<222> (8402)..(8402)

<223> ctosine (C) to thymidine (T) substitution polymorphism

<220>

<221> misc_feature

<222> (9434)..(9434)

<223> ambiguous nucleotide

<220>

<221> misc_feature

<222> (9496)..(9496)

<223> ambiguous nucleotide

FIGURE 2b continued

10/34

<220>

<221> misc_feature

<222> (10402)..(10417)

<223> ambiguous nucleotides

<220>

<221> primer_bind

<222> (6579)..(6601)

<223> Primer 17F
CCTGAGCTTGCCTCTCCCACAGT

<220>

<221> primer_bind

<222> (7036)..(7058)

<223> Primer 18R
CCAGGAGTCGCCGAGCAGGAAG
reverse primer

<220>

<221> primer_bind

<222> (7280)..(7303)

<223> Primer 6F
CCGGCCATCCAGAACTCCATGAAG

<220>

<221> primer_bind

<222> (7585)..(7605)

<223> Primer AW446985 dn1
TAGAACTCGCGGTCTCCAAAC
reverse primer

FIGURE 2b continued

11/34

<220>

<221> primer_bind

<222> (8222)..(8244)

<223> Primer InsUp1
TGGCTGTCACTCATCATCGGGCA

<220>

<221> primer_bind

<222> (8566)..(8589)

<223> Primer 14R2
TTGCACAGCACTTTATTGACACA

<220>

<221> primer_bind

<222> (6813)..(6830)

<223> Primer SNP1_FAM
AGC TTT GGC AGG TAA GGC

<220>

<221> primer_bind

<222> (6813)..(6830)

<223> Primer SNP1_HEX
AGC TTT GGC AGG TAA GAA

<220>

<221> primer_bind

<222> (6831)..(6844)

<223> Primer SNP1_2P
GGC CAA CGG GGG AG**FIGURE 2b continued**

12/34

<220>

<221> primer_bind

<222> (7424)..(7438)

<223> Primer SNP2_FAM
GCT GGC GGT GAG TGA

<220>

<221> primer_bind

<222> (7424)..(7438)

<223> Primer SNP2_HEX
GCT GGC GGT GAG TGG

<220>

<221> primer_bind

<222> (7439)..(7453)

<223> Primer SNP2_2P
CCT GCT GGG TGG GGA

<220>

<221> primer_bind

<222> (7442)..(7456)

<223> Primer SNP3_FAM
GCT GGG TGG GGA CGC

<220>

<221> primer_bind

<222> (7442)..(7456)

<223> Primer SNP3_HEX
GCT GGG TGG GGA CGT

FIGURE 2b continued

13/34

<220>

<221> primer_bind

<222> (7457)..(7470)

<223> Primer SNP3_P
GTG GGG GCG GGT GG

<220>

<221> primer_bind

<222> (8388)..(8402)

<223> Primer SNP4_FAM
TGC CCC AAC CTG GGT

<220>

<221> primer_bind

<222> (8388)..(8402)

<223> Primer SNP4_HEX
TGC CCC AAC CTG GGC

<220>

<221> primer_bind

<222> (8403)..(8417)

<223> Primer SNP4_2P
GCA GCA GGA GGA GGC

<220>

<221> primer_bind

<222> (6811)..(6830)

<223> Primer Dgat 21
GTAGCTTTGGCAGGTAAGAA**FIGURE 2b continued**

14/34

<220>

<221> primer_bind

<222> (6965)..(6984)

<223> Primer Dgat 22
GGGGCGAAGAGGAAGTAGTA
reverse primer

<220>

<221> primer_bind

<222> (6613)..(6632)

<223> Primer Dgat 23
TGGCCCTGATGGTCTACACC

<220>

<221> primer_bind

<222> (6829)..(6850)

<223> Primer Dgat 24B
GGGCAGCTCCCCCGTTGGCCGC
reverse primer

<220>

<221> primer_bind

<222> (6651)..(6670)

<223> Primer DgatforAD
TTCTCCTACCGGGACGTCAA

<220>

<221> primer_bind

<222> (6871)..(6890)

<223> Primer ReverseNZ
CCGCGGTAGGTCAGGTTGTC
reverse primer

FIGURE 2b continued

15/34

<220>

<221> primer_bind

<222> (6825)..(6838)

<223> Probe ForAA (FAM)
CGTTGGCCTTCTTA

<220>

<221> primer_bind

<222> (6823)..(6836)

<223> Probe DgatADGC (VIC)
TTGGCCGCCTTACC

<220>

<221> primer_bind

<222> (6651)..(6670)

<223> Primer DgatforAD
TTCTCCTACCGGGACGTCAA

<220>

<221> primer_bind

<222> (6878)..(6972)

<223> Primer DgatrevAD
AAGTAGTAGAGATCGCGGTAGGTCA
reverse primer

<220>

<221> primer_bind

<222> (6825)..(6838)

<223> Probe ForAA (FAM)
CGTTGGCCTTCTTA**FIGURE 2b continued**

16/34

<220>

<221> primer_bind

<222> (6823)..(6836)

<223> Probe DgatADGC (VIC)
TTGGCCGCCTTACC

<220>

<221> primer_bind

<222> (6652)..(6671)

<223> Primer DgatforRT66
TCTCCTACCGGGACGTCAAC

<220>

<221> primer_bind

<222> (6874)..(6964)

<223> Primer DgatrevRT66
GAGATCGCGGTAGGTCAGGTT
reverse primer

<220>

<221> primer_bind

<222> (6711)..(6972)

<223> Primer DgatforRTless66, GCTGCTTTGGCAGATCTCTACTACTT
This primer was designed to selectively bind and amplify the cDNA
splice variant. The corresponding binding site in this genomic s
equence comprises bases 6711 to 6715, 6815 to 6823 and 6960 to 69
72.

<220>

<221> primer_bind

<222> (7022)..(7038)

<223> Primer DgatrevRTless66
AAGCGCTTTCGGATGCG
reverse primer**FIGURE 2b continued**

17/34

<220>

<221> primer_bind

<222> (6857)..(6870)

<223> Probe Dgatwith66 (FAM)
CCGTGAGCTACCC

<220>

<221> primer_bind

<222> (6976)..(6990)

<223> Probe Dgatless66 (VIC)
CTTCGCCCCCACCCT

<400> 1

atg ggc gac cgc ggc ggc gcg ggc ggc tcc cgg cgc cgg agg acg ggg	48
Met Gly Asp Arg Gly Gly Ala Gly Gly Ser Arg Arg Arg Arg Thr Gly	
1 5 10 15	

tcg cgg cct tcg atc cag ggc ggc agt ggg ccc gcg gca gcg gaa gag	96
Ser Arg Pro Ser Ile Gln Gly Gly Ser Gly Pro Ala Ala Ala Glu Glu	
20 25 30	

gag gtg cgg gat gtg ggc gcc gga ggg gac gcg ccg gtc cgg gac aca	144
Glu Val Arg Asp Val Gly Ala Gly Gly Asp Ala Pro Val Arg Asp Thr	
35 40 45	

gac aag gac gga gac gta gac gtg ggc agc ggc cac tgg gac ctg ag	191
Asp Lys Asp Gly Asp Val Asp Val Gly Ser Gly His Trp Asp Leu Arg	
50 55 60	

gtagcgggtgc gcgtgacccc taacctttga cccctgatac ggggcccctg cgacccaacc	251
--	-----

tggtggccca ggcctgtcgg cggcagctcg ggctcgagtc cgagagtctg gcgcctggac	311
---	-----

cttggtgcac agctgtgccc ctcgggcctc cacggggaaa cttagcggga gggtgggggc	371
---	-----

ggaggggtctc ctgcccggaa caccaggtta cggggggccga ggggagggca gcggctcaac	431
---	-----

ttctagacgc cctccctctg ccttcctttg gtgggttctg aagctttccc aggggtgagcc	491
--	-----

cactacgcac agtgtcctct acctggaagg agatacaggg gtccttcctg agggctatga	551
---	-----

ggggtgcctt gtgggttgat aaagctcccg ggggaggagg gtggaccggc ggagaacaga	611
---	-----

ggcaggggca gtgctagggg atttctcatc cctcgagac cctccagaga atggtcttca	671
--	-----

caaaggtccc tcatccgtca ccggcgatt gactggccta ggatcctgct tattaccagc	731
--	-----

acaaatggct gctctagggt caaagtgggt cctgtaatgg gaccctcacc cctggttggg	791
---	-----

FIGURE 2b continued

18/34

gtacagggga ggagttggaa gtgcgcacac ccacaggtgg gcgccctgct tagctgaagg 851
 actgatggga aggagttggg ggagcaagct gcggctgaaa gggaggatct gaccacgtg 911
 ggcatcagct aagtcctgct ggctgcctcc aggcgttccc ttgcatcc tccacgcccc 971
 tccccgggg cctgaccttc atcctggtca agggctctca ggggctctgg ttttgggatc 1031
 agctccagag ctagaggtta tcaaggagga agtgggcaac aggtcagtca gcaaggattt 1091
 gctatcttca ctgggtgctg tggggagggg agggacaagg gcagttgggg tgcaggcact 1151
 gtccctgccc ttggggggca cacagttcac ctgagagata agatagccgc agccctgaag 1211
 agtgagagca aaggtcaggc acagagttca ggatgacacc aggggagggg ggctctgtga 1271
 ggggcactgg ctccctacag gcccaggtg gtcctgaggg ggcggctgca aaggccagga 1331
 gggccacagg cccctctgcc cactcctggg gaactggatt tggggtcact ttgtatgagg 1391
 tgggggcggg taccagcttt gggccaagct gtcaccctgg atgggccatc acttgccctgc 1451
 tctgtatagg ccagatggcc agaagctgct cctgtcctgt tgatggcca tcctcgaggt 1511
 ctggaccctc gggaagagga gcagttggtg gcagggatgg gccaccggag accctcctga 1571
 cctccaggac acgcagctgt gtgtgcctgt cccaagcca catgccacat ggctaggggc 1631
 ctccctggggc agggctgggc attggtctgg ctactcttgg tatcgccat ggcttccctg 1691
 cctcccagtc atcatcctcc cacctctgcc tccctgcctg ttcctctctt tctcctcagg 1751
 cccttccgga catttccctgc tcacctaggt ctgggcaggc ggggtcaggt gccgggtgtg 1811
 agctcactcc ttccggcagc aagggtgtagc tatgtgccgg aaggaaggcc gctgctgttg 1871
 cctgcctct gagtgcctcc ctccaggtc ctccacactc ccctgtgccc cgacacctgg 1931
 tgcgctcctc agccattggt tcatgtgtcc tccaggcaca gctttctagt ccagagcctc 1991
 taggctgggt gcaggaagtg ctgaggaagt ggcagccggg aggcgagctg gcaccctgtc 2051
 cctccttgtt ctgtccgtcc ctgcccctgg accgtatggc cccgcatgtg tgatccccac 2111
 ttggggctgt gcctctgggc aagtgggaa gcttgggtgag cctcattttc atgtgcccgc 2171
 ctcccagtac tgatgtgcag gttgaatgag gtgccaaactg taatgagttg gaatggccct 2231
 gctggctggg tgggactggg gagcaggtgg gggccgctgg ggggcacaga ggcacacca 2291
 gtgcctcagt caggagaggg gtgacagaga agctctgggt gagggccac ctccactctg 2351
 gccatggctg ctgccctttg gtccactgca gtgaactgtg ccatggggct ggacctctgt 2411
 ggggattggt gggcagtggg ctttcttccc gcttggggcc tctgacctct gggggcaggg 2471
 cgctgcccg gtgggacagt cggaaggctg gtagagggac ctgaggggtc tgtgtggtgg 2531
 ctgggggcag gcctcaggaa ttgacagca gggatctgga aaagctttaa taacattatt 2591

FIGURE 2b continued

19/34

tggtgtcagg attgggaaat gctccccctcc cccctccccc tctttcatct tagagactgc	2651
tgcacatctg gtcagtgtgg tcttcttgggt ggcccccaag gtggcagggg tcacactgtt	2711
atgaaaccgt cccctgggta tgtggtgcag acatgcacat gcagatgggtg attggcaggt	2771
tgtagcatga ggtggctttg ggacgggttcc agtgacagtg agtgggctgg atctgggggg	2831
ttctgggcag gtccatcaag cggatacccc cacagactgt cctcttggga tagttgggcc	2891
tgggagccct gcttgccttg ccaaaaggca ggcgagagt catgaagaag agggcttggg	2951
ggctcagagc cccactgtgt gtgcagccca ggggtggacct ggaggagggtg cgtgggcagg	3011
ctgggccggc cggggcgggg ggtggggggg cctgggtgtga aagggacca gggccagact	3071
gtcagcgtg cctggctgag gatgctggca ccctgtctc cccagccgtc tgtctcctgg	3131
gtgcagccat ctgagtgtg accccagccg cccctggagg ctggctgttc tcctgtgccc	3191
tattgctggg gacatgtgtc cacaggagg aaagggagc cccggcctct ccccttaca	3251
aactggaggc cttgctcaat gccctggatg gcctcctggt ggcagggtgg ttggtgggag	3311
gtggggctgc tgcttagaac ccgccagcgg gcctgggcct gggctgagct gcacccctcc	3371
acctctgcct ccagctgagg gttggcttcc atctccacca ggcccagcac tgggcacagg	3431
gctctcagag gcaggctctg aaagtcccct gctggcttct gcagtggact ccaggcgccg	3491
agccccagg gggctcgcct tgcgctcacc ctgcgaagcc acgtgaaggc tgggtcctcc	3551
cctccggaag ggccaaatgc agggcatggg tggtttgaat ggtggcccct gggtcccccg	3611
gagggaccag ctgctgtgag ggccgcccc tccccacttc cgtcttgcct caccagctcc	3671
tgtggcactc cccacgcccc gtccccaggt gggagcggca ggcccccggt ggctctgccc	3731
gcggaggggg atgtgtgggc ggcgggggtg ccttgcctgc agatgctctg ccccgagtgt	3791
ccgtctccgc tctccag g tgt cac cgc ctg cag gat tcc ctg ttc agt tct	3842
Cys His Arg Leu Gln Asp Ser Leu Phe Ser Ser	
65 70 75	
gac agt ggc ttc agc aac tac cgt ggc atc ctg aat tgg tgt gtg gtg	3890
Asp Ser Gly Phe Ser Asn Tyr Arg Gly Ile Leu Asn Trp Cys Val Val	
80 85 90	
atg ctg gtacgtagag tgacaccttg gagcaagggt cctgacggcc ggggggcat	3946
Met Leu	
gggtcttctt ccaggggtag gtgtctgtac ttgtgtagct gtggtgaatg gagctctgtg	4006
ctggcggtgg gggctccctg agcagccgta ccctgggacc ctaccgggag catgctcatg	4066
ccgtccctgc tgaatcccag gagatgcctg cagagggcag cctgggagcc tctgagctgg	4126
ggtctgcgcc ccagggggca ctggagtctc cccagggggc gagagagagt aggcagggat	4186

FIGURE 2b continued

20/34

```

ggctctgggtgg ccctgggtgg gggatggctg ctccgtgggc ccaggccctc cctggcagca 4246
caggtgagtg gtcttggggg tccacgtaga acttcctctt ctgttccaaa ttgccctcat 4306
gggtgcggca tgcctgggtg aacctggggg agcaggggtga ggacatgctt ctcagcccag 4366
cccacagctc caggccacac tctgcaggac tctggcccct ccctcagccc tggagggagc 4426
aggactggag tcctgtgtcc gccttgctct gacctggccg aggccactgc tgtggggccc 4486
cagcaggcct gccagcaga aggtggagtg cagggacccc aggggcagcc ttcaggggtg 4546
ggcaggggtga ggcccgactg ggcccagccc caccgctcag tgctgatgtg gcgcgaggcc 4606
ttcgccctc cagctgacgt gtctgcctgc cctgggtgtg gctccagagg ctgcctgtgt 4666
accaggggccc ccacgcttc tgtttgtggt tctgggcagt cccctgggga gcggtggggg 4726
ctgtgtgcca gtccagaccc agtagtcac gcgtcctggt ctctggaggc cgtggctggt 4786
ccaggactgt ggcaaggtgg tcgtgcaggg caggccctca gcagcctgtc tgttctcctg 4846
cagccccag cctcctggcc ctttgggtgca ccacaaagc tccccctcc ccaggagct 4906
ggggccgcct gctgcgtcct ctgggcagcc tgggcttcca ggtggctggg cctcttagca 4966
gctccaactc ttgcctgtgg tgggctctca ggacaggcaa ctgccagtcg gcagacattg 5026
caggaccacg tgtgtcctgg taagctggct ggtaggtgt ttagctgggg gatggtgtgg 5086
caggtggccc ctgcatctct gagcctgtca cctcctcggg aagccttctg ggtgggggac 5146
tccacccatg tcgcctggag aagcatcact tttccacaga gccttctgca acccccgtgg 5206
ggcctgagcc tggggtgggg gaggtgggtg cccctgctcc tgcagaggcc agccaggcat 5266
ctggccccag gccactggca agagctcgtt gtgttggggg atctgtcctt tgctgctgct 5326
gcaggagcgg ccgaggcagg cgggggctg agtaggggtg gagaccagc ccagcttcc 5386
ccagcccctc aggaccggcc tgctctttcc caccaccca ccaagtgcgt gggcacaccc 5446
cgctgtgag gatgggcccg gttggcaggg cggagccctg ggaggggtggc agtgcgccgg 5506
gcaggcttgg acttcactgg ggcttggggg tgctgctgtg gccaggggcg ctgaccgct 5566
tggtgggacg gacggccgct gggcagcagg tttcttctgc cacggtggca caggcacctg 5626
gggttgtggt tggctccagg cgggcggggg ctgcgtgccc ctgcgcaggc acataggccg 5686
tgggtgggga gtctcagagc ttggcgtgag gtcccacagg gctgggcctg caggatggag 5746
gccactgtcc tgagctgcag gtgctggcag gagctggggg gggcgttctg gggccgtggc 5806
tgacagcgtt atgtccctct ctctctatcg cag atc tta agc aac gca cgg tta 5860
                                Ile Leu Ser Asn Ala Arg Leu
                                95                                100

ttt cta gag aac ctc atc aa gtgagtgggc cccggcctgc ccagcccct 5910
Phe Leu Glu Asn Leu Ile Lys
                                105

```

FIGURE 2b continued

21/34

gccacctcac ccctcgccta cacagaccct caccacacctg cgtctgcag g tat ggc 5966
Tyr Gly

atc ctg gtg gac ccc atc cag gtg gtg tct ctg ttc ctg aag gac ccc 6014
Ile Leu Val Asp Pro Ile Gln Val Val Ser Leu Phe Leu Lys Asp Pro
110 115 120 125

tac agc tgg cca gct ctg tgc ctg gtc att g gtgagctggg tgcccaggag 6065
Tyr Ser Trp Pro Ala Leu Cys Leu Val Ile
130 135

gcctcaggcc ggcggtgggt gggacagggc tgatctgggc ctgaacctgc cctgggttgc 6125

ttctgtcctc ag tg gcc aat atc ttt gcc gtg gct gcg ttc cag gtg gag 6175
Val Ala Asn Ile Phe Ala Val Ala Ala Phe Gln Val Glu
140 145

aag cgc ctg gcc gtg gtaagcagtg ccctcacgcc ctcccctgac ttgcctcaag 6230
Lys Arg Leu Ala Val
150

gtccttacca gtcgggctta gggcggggcca ccagctggtc ccactgtgct tcagggtttt 6290

gggccttttcg tggccttcct gagaggggct gcacctcagg cctggtggct cttcctcagg 6350

gaggtcctct gaccagggag gggggtcctt ggctgacgct ctgctccac cccag gga 6408
Gly

gct ctg acg gag cag gcg ggg ctg ctg ctg cac ggg gtc aac ctg gcc 6456
Ala Leu Thr Glu Gln Ala Gly Leu Leu Leu His Gly Val Asn Leu Ala
155 160 165 170

acc att ctc tgc ttc cca gcg gcc gtg gcc ttt ctc ctc gag tct atc 6504
Thr Ile Leu Cys Phe Pro Ala Ala Val Ala Phe Leu Leu Glu Ser Ile
175 180 185

act cca g gtgggccccca ccccgcccc cgccccgcc cacgctgtct cggccacggg 6561
Thr Pro

cagcgcgggg ggcgtggcct gagcttgctt ctcccacag tg ggc tcc gtg ctg 6614
Val Gly Ser Val Leu
190

gcc ctg atg gtc tac acc atc ctc ttc ctc aag ctg ttc tcc tac cgg 6662
Ala Leu Met Val Tyr Thr Ile Leu Phe Leu Lys Leu Phe Ser Tyr Arg
195 200 205

gac gtc aac ctc tgg tgc cga gag cgc agg gct ggg gcc aag gcc aag 6710
Asp Val Asn Leu Trp Cys Arg Glu Arg Arg Ala Gly Ala Lys Ala Lys
210 215 220 225

gct g gtgagggctg cctcgggctg gggccactgg gctgccactt gcctcgggac 6764
Ala

cggcaggggc tcggctcacc cccgaccgc cccctgcgc ttgctcgtag ct ttg 6819
Ala Leu

FIGURE 2b continued

22/34

gca ggt aag aag gcc aac ggg gga gct gcc cag cgc acc gtg agc tac Ala Gly Lys Lys Ala Asn Gly Gly Ala Ala Gln Arg Thr Val Ser Tyr 230 235 240	6867
ccc gac aac ctg acc tac cgc g gtgaggatcc tgccgggggc tggggggact Pro Asp Asn Leu Thr Tyr Arg 245 250	6919
gcccggcggc ctggcctgct agccccgccc tcccttcag at ctc tac tac ttc Asp Leu Tyr Tyr Phe 255	6973
ctc ttc gcc ccc acc ctg tgc tac gag ctc aac ttc ccc cgc tcc ccc Leu Phe Ala Pro Thr Leu Cys Tyr Glu Leu Asn Phe Pro Arg Ser Pro 260 265 270	7021
cgc atc cga aag cgc ttc ctg ctg cgg cga ctc ctg gag atg Arg Ile Arg Lys Arg Phe Leu Leu Arg Arg Leu Leu Glu Met 275 280 285	7063
gtgaggcggg gcctcgcggg ccagggtggg cgggcctgcc ggcacccggc accggggctc agctcactgt ccgcttgctt ccttccccag ctg ttc ctc acc cag ctc cag gtg Leu Phe Leu Thr Gln Leu Gln Val 290	7123 7177
ggg ctg atc cag cag gtacgtgccc gggggggggg gggggggggg gggggggact Gly Leu Ile Gln Gln 295	7232
ctggggccgt tggggagctg actctgcgct ttttgcag tgg atg gtc ccg gcc atc Trp Met Val Pro Ala Ile 300 305	7288
cag aac tcc atg aag ccc ttc aag gtgagcaggc aggcctggca ggggtgggttc Gln Asn Ser Met Lys Pro Phe Lys 310	7342
cggggtcagg gctgaggag ccagctgtgc cctgtgccca cag gac atg gac tac Asp Met Asp Tyr 315	7397
tcc cgc atc gtg gag cgc ctc ctg aag ctg gcg gtgagtgacc tgctgggtgg Ser Arg Ile Val Glu Arg Leu Leu Lys Leu Ala 320 325	7450
ggacgcgtgg gggcgggtgg ggctgttctg gcacctggca cccactcccc acag gtc Val	7507
ccc aac cac ctc atc tgg ctc atc ttc ttc tac tgg ctc ttc cac tcc Pro Asn His Leu Ile Trp Leu Ile Phe Phe Tyr Trp Leu Phe His Ser 330 335 340 345	7555
tgc ctg aac gcc gtg gct gag ctc atg cag ttt gga gac cgc gag ttc Cys Leu Asn Ala Val Ala Glu Leu Met Gln Phe Gly Asp Arg Glu Phe 350 355 360	7603

FIGURE 2b continued

23/34

tac cgg gac tgg tg gtgggtggcc ttgccggggc gggggtggtg ggggcccccg 7657
 Tyr Arg Asp Trp Trp
 365

ctggggctgg ggccggagcc cctgcccact ctgccccgcc cccgcag g aac tcc gag 7714
 Asn Ser Glu

tcc atc acc tac ttc tgg cag aac tgg aac atc cct gtt cac aag tgg 7762
 Ser Ile Thr Tyr Phe Trp Gln Asn Trp Asn Ile Pro Val His Lys Trp
 370 375 380 385

tgc atc ag gtgggtgtgc gcctgggggc ggggggttgg ggggtgggac 7810
 Cys Ile Arg

ggggtcgcgt ggcccgggcg cccagcccac tgccgcctcc cccgcag a cac ttc tac 7867
 His Phe Tyr
 390

aag ccc atg ctc cgg cgg ggc agc agc aag tgg gca gcc agg acg gca 7915
 Lys Pro Met Leu Arg Arg Gly Ser Ser Lys Trp Ala Ala Arg Thr Ala
 395 400 405

gtg ttt ctg gcc tcc gcc ttc ttc cac gag gtcagtgcac tgagggcgcg 7965
 Val Phe Leu Ala Ser Ala Phe Phe His Glu
 410 415

ccctgcccct ggtgggggtg ggggtggggg tgggggctcg ctgacgcccc tctcccctca 8025

g tac ctg gtg agc atc ccc ctg cgc atg ttc cgc ctc tgg gcc ttc acc 8074
 Tyr Leu Val Ser Ile Pro Leu Arg Met Phe Arg Leu Trp Ala Phe Thr
 420 425 430

ggc atg atg gcg cag gtgagcagcc ctggaccccc gctccgcccc gccccgcgag 8129
 Gly Met Met Ala Gln
 435

cgcagaggct cactcccgtc ctgtgtcccc ag atc ccg ctg gcc tgg ata gtg 8182
 Ile Pro Leu Ala Trp Ile Val
 440 445

ggc cgc ttc ttc cgc ggc aac tac ggc aac gcg gcc gtg tgg ctg tca 8230
 Gly Arg Phe Phe Arg Gly Asn Tyr Gly Asn Ala Ala Val Trp Leu Ser
 450 455 460

ctc atc atc ggg cag ccg gtg gcc gtc ctg atg tac gtc cac gac tac 8278
 Leu Ile Ile Gly Gln Pro Val Ala Val Leu Met Tyr Val His Asp Tyr
 465 470 475

tac gtg ctc aac cgt gag gcg ccg gca gcc ggc acc tgagcgctc 8324
 Tyr Val Leu Asn Arg Glu Ala Pro Ala Ala Gly Thr
 480 485

caggctggcc ccctcgtggg tggttgactg ctttgccgcg ctgcctgcgg ctggactaga 8384

gcctgccccca acctgggagc agcaggagga ggcctggctg gtggaagctg cctcctggcc 8444

tccaccaggc ctctgcctga agggcttctt cctgccaggg gagagcaggc ccgacgcagt 8504

tctggcccct gggaggtgcc catgctctgg aaaccctaca gatctcgccc aagggtctga 8564

FIGURE 2b continued

24/34

atgtgtcaat	aaagtgtctgt	gcacagtgtgag	ctccctcagc	ctccagggca	cagggctggc	8624
aggagggggc	ggccctccca	cgtggggcca	tgctgtggga	aggaggcccc	agcgcctgga	8684
gaggagctgg	ggctgtgggtg	accctccctg	cctcacaggg	ctctgtggtc	agacgtcttg	8744
ccctgcaagg	tggagactcc	atgctccaag	gccccctgtg	cctgaggtct	gcacacaagt	8804
ggattcaact	tgggtcaggg	cagaggctaa	ggtgtggaag	aagggtgaga	atcaggctga	8864
cttgaacggc	agcaaagact	ccaaggcaag	gctgcagaag	tctcaaaagc	tatgcgcaca	8924
gtccccctgct	ggggtgctca	cctgggctgg	gctctgggct	gcttggacaa	agcaggtggc	8984
ctggctcagc	cctcaccgag	ggccttcctt	gggggcagag	gttggcctga	tgccaggggc	9044
tccccgtttt	tccaggccct	cagcaggtag	ttgggtgtgg	ccctcatgat	accttggtcc	9104
cagagctccg	ccactcaaaa	agcttggcag	tgaggcaagg	gcaaccccgg	gctgttcccc	9164
cctctactgg	ctctgccgcc	tgggttgga	accctgaggg	tgtgccaggg	aggtgtaccc	9224
tgacagccag	ccatggccca	gtaagatggg	tgcccagaggt	ggtacctggg	cagcggaccc	9284
agctgtgctg	ccccgcgcc	aaccagaagc	cgctctagcc	catgggtggc	gtctgggcga	9344
gacaggctgg	ttggctaggg	actgtttgg	ctacagcagg	tgtaggcagc	gtctccctga	9404
cccctgcctt	ctaggaagcc	accaccctgn	gccctactca	tcagcaagga	cagcgagcag	9464
ggctgagctg	ggggtgcgtg	ggctgctacg	gnccccgcac	cttcatcaca	tgcacctctg	9524
caccccctgc	tgccctgactc	aagagtgggg	gggggggtcc	tgtgcttcct	tcattccaga	9584
cccacgggtgc	tgaccagtg	cacccacctg	gtccgctagt	gctgacctgg	ccacagggct	9644
cctgtggggc	cacgctgatc	ccgccctgg	cccttcataa	agaactcttg	agcacatgca	9704
gcccagggga	gccaggaggc	tccagtgtgc	tgtgtccatc	tgccctccctt	cagccccctc	9764
cgagacactg	cgcctcatgc	ccccctccac	ccccaccac	actggcagga	ggaacacaca	9824
gggagaccac	acacagagct	cgctgtttat	aaatctctgc	ctggctcatc	ggtctgtttg	9884
tccatgtata	tatctgcata	tctctatgga	aggggaaagg	gggactcgtg	taaaaatcca	9944
aaatacaatt	ctatgaacac	ctgcatcctg	gtcagtctga	gtgtggccgt	gaagcccagg	10004
tgagctgtgg	ctcacagggc	taggccctcg	gtgctggccg	ggggccactc	cccaccccct	10064
ctccccccct	ccgccagcca	ggggaccagg	ctcctggaca	ccaggcctgc	ccaaggcctg	10124
ctctcctcct	ggggcttcta	cgagacagtg	gggtccttgg	ctttgggggg	ttctgagccc	10184
gtcagcaggg	agatgggtgg	gtcatctctt	atatttcgtc	tccctcggag	aagtaggagc	10244
cctccccccag	ctcgaagagc	accggcaggt	cgctgctccc	cacgtccacg	gagcccgggt	10304
ccaggagcag	caggggctgg	gcggtgtagt	gcaccagctg	cttccctagg	ggtgcgactg	10364

FIGURE 2b continued

25/34

```

ggtcaggggtg ccggtggggc cggggggcgg ggtggggnnn nnnnnnnnnn nnncccccc 10424
ccccggcccc agccaccccg cctacgcacg ctggccaggc tgctgtccag gtcgggcagg 10484
ctcatgtcgg gcaccgtaac cgaggggctg aacagctgca gggaagaggg gcgggtcaga 10544
ctgccctgga agcggggcgg gctgggcgcg gggcgggcag ggtagactc acatccagca 10604
gggtgctggt gtccacgctg aagccatggc ttgtcagcat ggtctgcagg ttgtccaggt 10664
tgaggtccat ggcgtccaag tggtcgctga gctcggctct ggccagagga aggggagcag 10724
gtgacgtggc atccaggcac cccaggtgc agtcctgccg gcccttaagc ccagctgcca 10784
gcctgccctg cgcaggctca cgggaggccc tgaggtgggg ccctgggcct gggacacttg 10844
gccaccggtg gtggctcagg ctccctcctc accgtgaggc cccgtaaggc agcacgtcag 10904
gctgggcccg ctaagacatc agggggctga gttcagggtc cagaagggcc cggggcccca 10964
gcttctcagg ttccagcaacg cctgcccttc aactggagg acagcaggag ggtggagctg 11024
tgctcggcca ggctgggcca gcctaccca gcatgcgcca ggatagatta ggctgcccc 11084
ctcagggagg ccaggcagta cctcgggggtg ctgcctcagg cccagagggt ggggtggggag 11144
cctgttcaca ggctgggcag gggctctgctc agaggctggg gaggggcca agctctggcg 11204
gagccctcct ccctccccag tgggacagcg ctaaccctgg ctggactcgg ccatgcagag 11264
aggaggaggg gcaggggaag aggcgggcga cccagacat ctttggagtg cgagccaaac 11324
tgcaagatag aaagacaaga gccccctct gcgggtgggg ccagcgagt ccaagcccc 11384
cacc'ccgttc acccacctct ggggtggtag gacagacaga ccaggcggcc cggggagggc 11444
ggctcagggg gcaggggtgc ccgccacgtg cctgggtccc acggtgcagg ggctgcggga 11504
ccccctggct ggcattgtct cctccccagc tccatgtctt gggccaggtc ggccagtttc 11564
ccccagcaaa tcccatgccg agggcctgag ggccatgcgt gtccagcccg gcactgtcct 11624
tgtcggaggc ccctgcgctt ccagagaccg aagggcgcct ggaaggcact cacttgtcta 11684
ggcaggcgac gctgaggcac ttctcgggag ccgaggcggg gagggtcgac ggatagcggc 11744
cccgagtgat ccgatagaag cttcgta 11771

```

FIGURE 2b continued

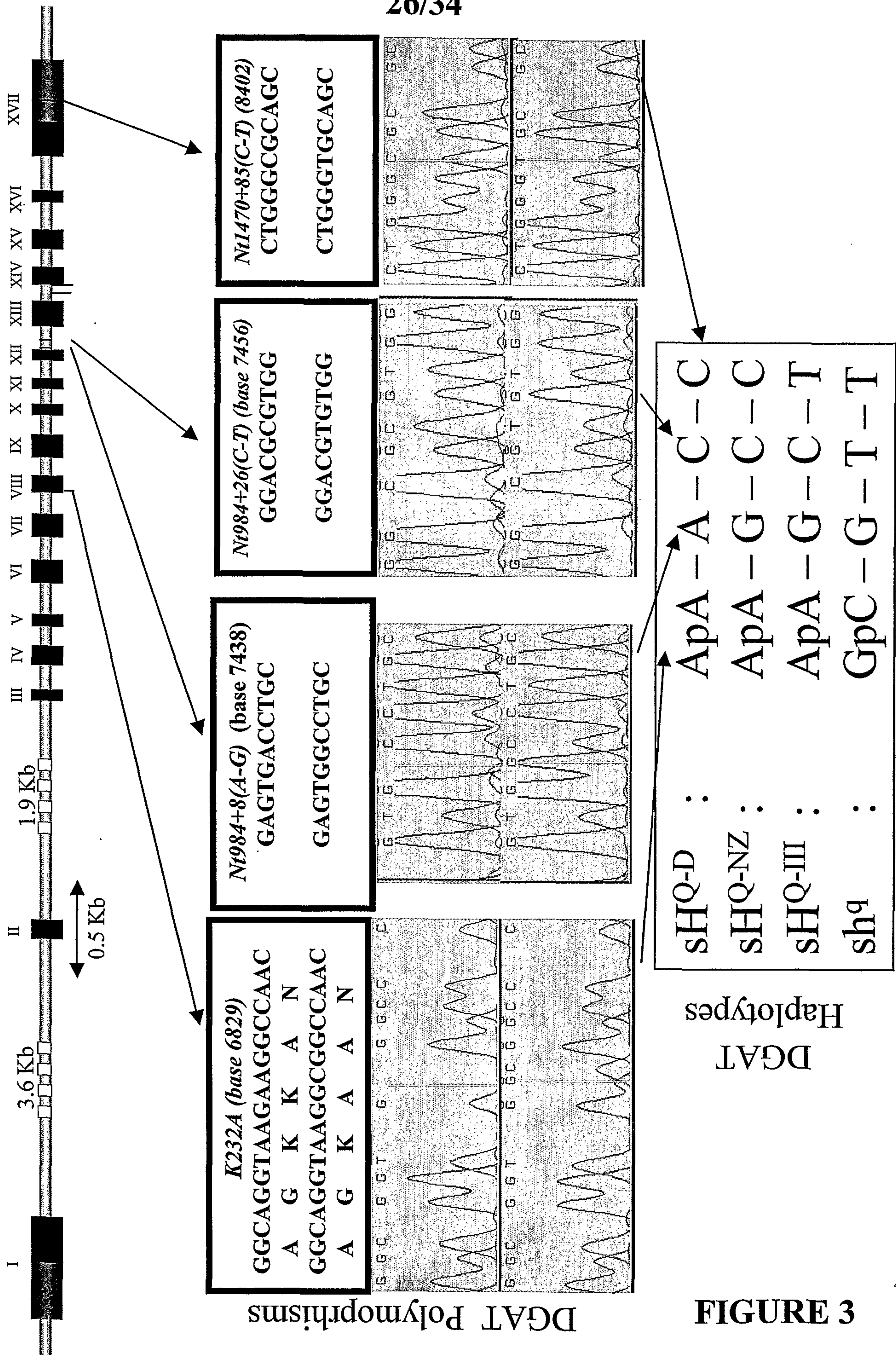


FIGURE 3

27/34

<221> VARIANT

<222> (232)..(232)

<223> an amino acid substitution (K -> A) caused by a polymorphism at bases 7224-7225 of the genomic sequence (measured from the adenine residue of the translation start codon). Lysine (K) corresponds to the Q allele, alanine (A) corresponds to the q allele.

<400> 5

Met Gly Asp Arg Gly Gly Ala Gly Gly Ser Arg Arg Arg Arg Thr Gly
1 5 10 15

Ser Arg Pro Ser Ile Gln Gly Gly Ser Gly Pro Ala Ala Ala Glu Glu
20 25 30

Glu Val Arg Asp Val Gly Ala Gly Gly Asp Ala Pro Val Arg Asp Thr
35 40 45

Asp Lys Asp Gly Asp Val Asp Val Gly Ser Gly His Trp Asp Leu Arg
50 55 60

Cys His Arg Leu Gln Asp Ser Leu Phe Ser Ser Asp Ser Gly Phe Ser
65 70 75 80

Asn Tyr Arg Gly Ile Leu Asn Trp Cys Val Val Met Leu Ile Leu Ser
85 90 95

Asn Ala Arg Leu Phe Leu Glu Asn Leu Ile Lys Tyr Gly Ile Leu Val
100 105 110

Asp Pro Ile Gln Val Val Ser Leu Phe Leu Lys Asp Pro Tyr Ser Trp
115 120 125

Pro Ala Leu Cys Leu Val Ile Val Ala Asn Ile Phe Ala Val Ala Ala
130 135 140

Phe Gln Val Glu Lys Arg Leu Ala Val Gly Ala Leu Thr Glu Gln Ala
145 150 155 160

Gly Leu Leu Leu His Gly Val Asn Leu Ala Thr Ile Leu Cys Phe Pro
165 170 175

FIGURE 4a

28/34

Ala	Ala	Val	Ala	Phe	Leu	Leu	Glu	Ser	Ile	Thr	Pro	Val	Gly	Ser	Val	180	185	190
Leu	Ala	Leu	Met	Val	Tyr	Thr	Ile	Leu	Phe	Leu	Lys	Leu	Phe	Ser	Tyr	195	200	205
Arg	Asp	Val	Asn	Leu	Trp	Cys	Arg	Glu	Arg	Arg	Ala	Gly	Ala	Lys	Ala	210	215	220
Lys	Ala	Ala	Leu	Ala	Gly	Lys	Lys	Ala	Asn	Gly	Gly	Ala	Ala	Gln	Arg	225	230	235
Thr	Val	Ser	Tyr	Pro	Asp	Asn	Leu	Thr	Tyr	Arg	Asp	Leu	Tyr	Tyr	Phe	245	250	255
Leu	Phe	Ala	Pro	Thr	Leu	Cys	Tyr	Glu	Leu	Asn	Phe	Pro	Arg	Ser	Pro	260	265	270
Arg	Ile	Arg	Lys	Arg	Phe	Leu	Leu	Arg	Arg	Leu	Leu	Glu	Met	Leu	Phe	275	280	285
Leu	Thr	Gln	Leu	Gln	Val	Gly	Leu	Ile	Gln	Gln	Trp	Met	Val	Pro	Ala	290	295	300
Ile	Gln	Asn	Ser	Met	Lys	Pro	Phe	Lys	Asp	Met	Asp	Tyr	Ser	Arg	Ile	305	310	315
Val	Glu	Arg	Leu	Leu	Lys	Leu	Ala	Val	Pro	Asn	His	Leu	Ile	Trp	Leu	325	330	335
Ile	Phe	Phe	Tyr	Trp	Leu	Phe	His	Ser	Cys	Leu	Asn	Ala	Val	Ala	Glu	340	345	350
Leu	Met	Gln	Phe	Gly	Asp	Arg	Glu	Phe	Tyr	Arg	Asp	Trp	Trp	Asn	Ser	355	360	365
Glu	Ser	Ile	Thr	Tyr	Phe	Trp	Gln	Asn	Trp	Asn	Ile	Pro	Val	His	Lys	370	375	380
Trp	Cys	Ile	Arg	His	Phe	Tyr	Lys	Pro	Met	Leu	Arg	Arg	Gly	Ser	Ser	385	390	395
Lys	Trp	Ala	Ala	Arg	Thr	Ala	Val	Phe	Leu	Ala	Ser	Ala	Phe	Phe	His	405	410	415

FIGURE 4a continued

29/34

Glu Tyr Leu Val Ser Ile Pro Leu Arg Met Phe Arg Leu Trp Ala Phe
420 425 430

Thr Gly Met Met Ala Gln Ile Pro Leu Ala Trp Ile Val Gly Arg Phe
435 440 445

Phe Arg Gly Asn Tyr Gly Asn Ala Ala Val Trp Leu Ser Leu Ile Ile
450 455 460

Gly Gln Pro Val Ala Val Leu Met Tyr Val His Asp Tyr Tyr Val Leu
465 470 475 480

Asn Arg Glu Ala Pro Ala Ala Gly Thr
485

FIGURE 4a continued

30/34

<400> 6

```

Met Gly Asp Arg Gly Gly Ala Gly Gly Ser Arg Arg Arg Arg Thr Gly
1          5          10          15

Ser Arg Pro Ser Ile Gln Gly Gly Ser Gly Pro Ala Ala Ala Glu Glu
          20          25          30

Glu Val Arg Asp Val Gly Ala Gly Gly Asp Ala Pro Val Arg Asp Thr
          35          40          45

Asp Lys Asp Gly Asp Val Asp Val Gly Ser Gly His Trp Asp Leu Arg
50          55          60

Cys His Arg Leu Gln Asp Ser Leu Phe Ser Ser Asp Ser Gly Phe Ser
65          70          75          80

Asn Tyr Arg Gly Ile Leu Asn Trp Cys Val Val Met Leu Ile Leu Ser
          85          90          95

Asn Ala Arg Leu Phe Leu Glu Asn Leu Ile Lys Tyr Gly Ile Leu Val
          100          105          110

Asp Pro Ile Gln Val Val Ser Leu Phe Leu Lys Asp Pro Tyr Ser Trp
          115          120          125

Pro Ala Leu Cys Leu Val Ile Val Ala Asn Ile Phe Ala Val Ala Ala
130          135          140

Phe Gln Val Glu Lys Arg Leu Ala Val Gly Ala Leu Thr Glu Gln Ala
145          150          155          160

Gly Leu Leu Leu His Gly Val Asn Leu Ala Thr Ile Leu Cys Phe Pro
          165          170          175

Ala Ala Val Ala Phe Leu Leu Glu Ser Ile Thr Pro Val Gly Ser Val
          180          185          190

Leu Ala Leu Met Val Tyr Thr Ile Leu Phe Leu Lys Leu Phe Ser Tyr
          195          200          205

Arg Asp Val Asn Leu Trp Cys Arg Glu Arg Arg Ala Gly Ala Lys Ala
210          215          220

Lys Ala Ala Leu Ala Asp Leu Tyr Tyr Phe Leu Phe Ala Pro Thr Leu
225          230          235          240

```

FIGURE 4b

31/34

Cys Tyr Glu Leu Asn Phe Pro Arg Ser Pro Arg Ile Arg Lys Arg Phe
 245 250 255

Leu Leu Arg Arg Leu Leu Glu Met Leu Phe Leu Thr Gln Leu Gln Val
 260 265 270

Gly Leu Ile Gln Gln Trp Met Val Pro Ala Ile Gln Asn Ser Met Lys
 275 280 285

Pro Phe Lys Asp Met Asp Tyr Ser Arg Ile Val Glu Arg Leu Leu Lys
 290 295 300

Leu Ala Val Pro Asn His Leu Ile Trp Leu Ile Phe Phe Tyr Trp Leu
 305 310 315 320

Phe His Ser Cys Leu Asn Ala Val Ala Glu Leu Met Gln Phe Gly Asp
 325 330 335

Arg Glu Phe Tyr Arg Asp Trp Trp Asn Ser Glu Ser Ile Thr Tyr Phe
 340 345 350

Trp Gln Asn Trp Asn Ile Pro Val His Lys Trp Cys Ile Arg His Phe
 355 360 365

Tyr Lys Pro Met Leu Arg Arg Gly Ser Ser Lys Trp Ala Ala Arg Thr
 370 375 380

Ala Val Phe Leu Ala Ser Ala Phe Phe His Glu Tyr Leu Val Ser Ile
 385 390 395 400

Pro Leu Arg Met Phe Arg Leu Trp Ala Phe Thr Gly Met Met Ala Gln
 405 410 415

Ile Pro Leu Ala Trp Ile Val Gly Arg Phe Phe Arg Gly Asn Tyr Gly
 420 425 430

Asn Ala Ala Val Trp Leu Ser Leu Ile Ile Gly Gln Pro Val Ala Val
 435 440 445

Leu Met Tyr Val His Asp Tyr Tyr Val Leu Asn Arg Glu Ala Pro Ala
 450 455 460

Ala Gly Thr
 465

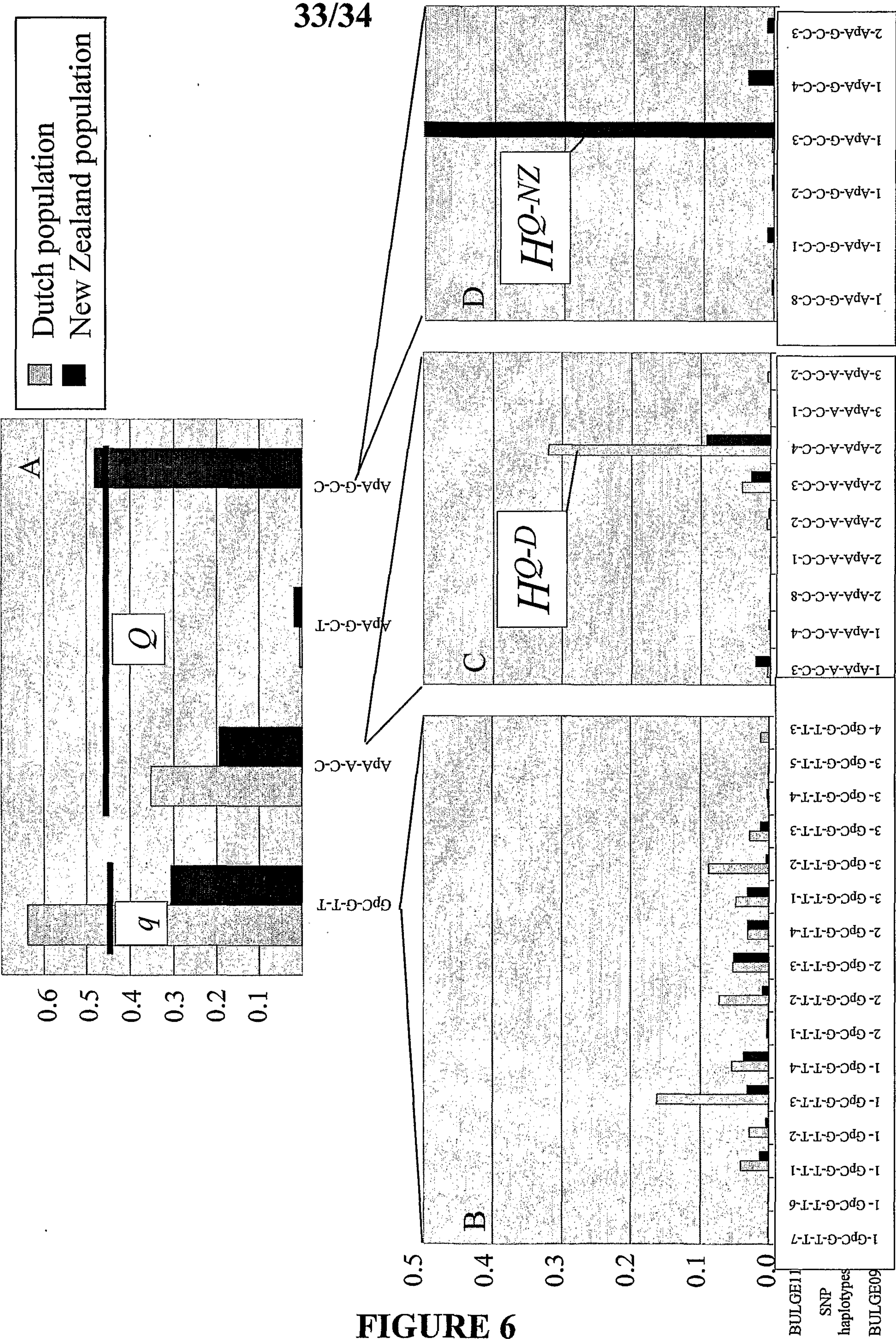
FIGURE 4b continued

K232A



<i>Bos taurus</i> :	LALMVYTIILFLKLFESYRDVNLWCRRERAGAKAKAALAGKKANGGAAQRTVSYPDNLTYRDLYYFELFAPTLCY
<i>Bison bison</i> :	LALMVYTIILFLKLFESYRDVNLWCRRERAGAKAKAALAGKKANGGAAQRTVSYPDNLTYRDLYYFELFAPTLCY
<i>Ovis aries</i> :	LALMVYTIILFLKLFESYRDVNLWCRRERAGAKAKAALAGKKANGGAAQRTVSYPDNLTYRDLYYFELFAPTLCY
<i>Sus scrofa</i> :	LALMVYAIILFLKLFESYRDVNLWCRRERATAKAKAASAGKKANGGAAQHSVSPDNLTYRDLYYFELLAPTLCY
<i>Homo sapiens</i> :	LALMAHTIILFLKLFESYRDVNSWC--RR--ARAKAASAGKKASSAAAPH TVSYPDNLTYRDLYYFELFAPTLCY
<i>Cercopithecus aethiops</i> :	LALMVHTIILFLKLFESYRDVNLWC--RR--ARAKAASAGKRASSAAAPH TVSYPDNLTYRDLYYFELFAPTLCY
<i>M. musculus domesticus</i> :	FALASYSIMFLKLYSYRDVNLWCRQRR--VKAKAVSTGKKVSGAAQQAQAVSYPDNLTYRDLYYFIFAPTLCY
<i>Rattus norvegicus</i> :	FALASYSIIIFLKLSSYRDVNLWCRQRR--VKAKAVSAGKKVSGAAQAQNTVSYPDNLTYRDLYYFIFAPTLCY

FIGURE 5



34/34

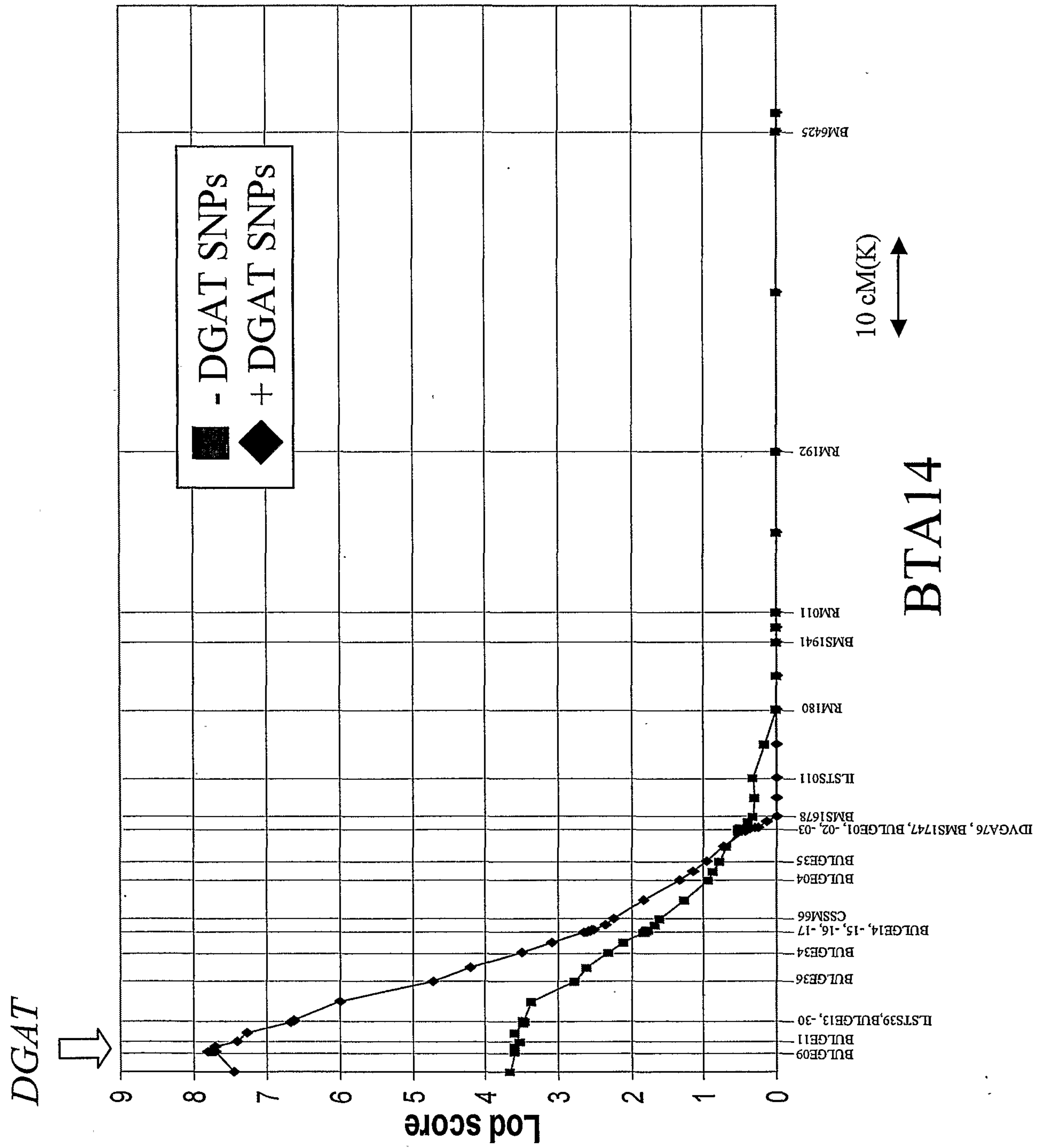


FIGURE 7