



(22) Date de dépôt/Filing Date: 2013/08/09

(41) Mise à la disp. pub./Open to Public Insp.: 2014/02/09

(30) Priorité/Priority: 2012/08/09 (US61/681,152)

(51) Cl.Int./Int.Cl. *C12N 15/12* (2006.01),
A01K 67/00 (2006.01), *C07H 21/04* (2006.01),
C12Q 1/68 (2006.01), *C40B 30/00* (2006.01),
C40B 40/06 (2006.01)

(71) Demandeur/Applicant:
WISCONSIN ALUMNI RESEARCH FOUNDATION, US

(72) Inventeur/Inventor:
KHATIB, HASAN, US

(74) Agent: MCCARTHY TETRAULT LLP

(54) Titre : POLYMORPHISMES DE NUCLEOTIDE SIMPLE ASSOCIES A LA FERTILITE DES TAUREAUX

(54) Title: SINGLE NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH BULL FERTILITY

```
1081 gagcgcgctc tgcagagcgc gaataagcgtc tgcgtgggaa accctgggag gggcggtaac
1141 ggtccagcgg ggcgcgactc ggcgcgcggg accgacccgc ccgcagccc caccggggt
1201 gccgcagtcg cccgcgcgca tgcacccgcg cccacacgct cggcgccccc acgtggcggg
1261 gccggcgggg cgtgcgttcc gccggtgctt ggggggtgct tggcccccct cgtctgcctc
1321 aaccggggtt tgttgcgctc gaagtccccc ggtgggcagc tctgtccctt gctttccctt
1381 tccccccggg cctgcgcgac ctcgccctct cctcgacgca ggtccggggt cctccacaaa
1441 ccccgagccc cgtctctatc tctcggggtg gttgtgaagc ggtctggccc cagacacccc
1501 gctgggtggg gtggtgcagt ccgcacgtgc gtctctacgg cagcccgagg cggacaaagg
1561 ggtttcagcg agccctcggt cccacgccc gcccgcacac ggcacgagaa gacgcggagaa
1621 aacccctgtt tagatgaaaa aaaattaaac tctagtggct tgcctctgcn ttgaaaacg
1681 gtcgctgtg cccagacaaa aaggctgcag ggtggagact cgggttgagc acctgggtct
1741 ttgttttaac tttaagcacc tgggtttact ttctggctta aaaaagaaa aaatgtgag
1801 cagaaagcag aattgtttat ttattcgga aaaaanaatg gtagcattag ctaggctaat
1861 agtgaatgag ctcatttgag catctaatga ccttccattt gtaaacat agtctcttga
1921 atttgatcag taactaatga gttttatttg acattttctc cttgcaaat ctccacgta
1981 ctgactttag gtgtgcctgc tgtatgaac tatgcctttg gattttttaa aaatatat
2041 ttgtcagcac ttgtgctggg aagtaaatat aaacaggtg tgtttctatg taataaagg
2101 gcataagcac ccgatgtggt gtagggagag ggaaccccat cccatttttg gagatggggg
2161 cccagagggg ccggaagcat agggagacaa gctcaggttc aacacaccca taagtgcacg
2221 aatgcggggt ctgcagcacc acgtgttttt ggcgcgtcat gatgttagtg caggccagagc
2281 gaacatccca ggaagatcaa agagcaattt ctacagcttg cctttataca aaggcagtag
2341 tcatccactt gtggcgagca cccatgcag gtaggtgggt acattataca aaggcagtag
2401 attatcttcc tgcctttgtt cctcgtctgt ccccttttca gtagcattac ctccatctga
2461 ttatctgaag ctgcttctgt ttgaggtata actagatcca gtagcattac ctccatctga
2521 gacgtggtct agcatatttt tccagctact ggaacttagt gtagcattac ctccatctga
2581 tcttgccccc ttcttttttt aaaaattatt gatgtattat gtagcattac ctccatctga
2641 atcctaaagc agggagcttg tgaattttta caaaagaca taactgtatc cgtatgagag
2701 accgacattc ctgggacccc cagaaacccc ctaggggttc ctctttcagc agcatgcctt
2761 cctgattttt tagtgccttg atttagtcag ctctttcggg ggaactctga aatagcagta
2821 ctgctcccat taagaatcaa ccaagtgaat gagaaagacc atctctcttg gacagctaac

2881 agatgctgct ttctctctgg agatcaactc tcttggtccc tggggctctg gatgcagctc
2941 aaaccacagg gctctccagt gctaaagggg aatactgctc ctgtagcttt tctgacagag
3001 atggtttagc aacagggcag gctgtgctgt gctctggcct gatctgatgc ttctccagct
3061 ctccgctctg agcagggcct cctctccag aggtcactct gataccttgg ctccgggggg

3121 tgggtccctt gggccagctt gctggaagca agatggagat tgaactccgggg ctccggcctg
SNP2
3181 cttgtgtgac gacgcctttc cgcagggaaa cgggttctac ttctgcccct ctgtgtgtgc
3241 taaccttcca tccggatgag gctctctgag tctggagaca ctagggaact gattttttgc
SNP3
3301 ctttgtatga tttaaggctc taagctgggag ctagaataa aatctgagg aatagggcgg
3361 aatgcaagt acctcaatag gctgaatca gccaccatg gaggctgagt acatgtctt
3421 tgaacactag tttacttga tcaactctg gcgcctcacc tgcctgagct aaagcctatt
3481 tactgacccc aggtgtgtgt gaaggaatgt gcagcattat tgaagatgc tgaatgaagg
3541 agcaggggga gctgtgtgtt aagaccccca aaatccccc aggtttttag caaagcattt
3601 taaagggcag tgaaggggag ggtgtcccaa agtgtgtgat cagctctatc acaattctct
3661 aattggaatc tagatgagt atgaggtaac aagactactg tgggtctctt tctgtgttta
3721 gatgtgaag aatctgcta caatgcagga gactctggt ttatccctgg gttgagatcc
3781 cttgagagag ggcagacaaa caatttccag tattcttgcc tgggaatatc ccatgggacg
3841 aggaagcttg tgggtacagt ccatgggggt gcaagagttt ggcctgactt gattgactag
3901 gcaagagaaa agcctatggt aatgaggaat atgtgggagt ggtcactcag tagttaattt
3961 cttccatttg atgtgtgttt tagabotga acaataactc aggaatata cttcagatag
4021 taatatctag gatttcaag gaggaaacta agcagagggg ataggggagg ggtctgccc
4081 aggaatctcc ctaaggttcc tgccttggtta ccaagagaca gaagacatc attgctttct
4141 tcaagcttct tcaaatgcca taatgtaggg tctgttattt tcttaagtct gaacacagc
4201 tcttgttcca gactgttcta gctttattta gttagtattg caggactcca tgcacaccca
4261 ttgtcagctc ttggaaccca ctttccccc ttctaatca gtttccaca tagaactaac
4321 agtctctccc gtttttagca ggtctctccc tctccagttt ttttttggg tttgagctac
4381 ataacctaat gttttgcaaa acgaattcta gacgtctgtt aaaagtttct attgcttctt
4441 tttatttttt ttttttaaat aaacatgcta aaaaacacta agtgcagtt tgggttccat
4501 tctgtttaga gtttaacttt cctctggtta acattgggaa tcaagcaggt tttgttctca
4561 gaaagataga tgtgtgtaac atctactgac ctggattttt agatgtaagt ttgtctaat
4621 acattgtcat ttattgata cttcagctct acatttggtta agagatttaa gaaacacgga
4681 aaggttagg gctttgttca cttttgttt taacaggaac acaacggggt gatgggaatg
4741 aagagaaatg ttgggaatg atattttac tatatccatc catcttcta cgtttcaaaa
4801 tcaaaagcca gttgggttcc tgcctgctct gcttaatttg tttatctaac tcaagagctt
```

Partial genomic sequence of MAP1B showing the location of SNPs 1, 2 and 3 on the MAP1B gene and the locations of the primers (1F and 1R, corresponding to positions 2859-2876, and positions 3346-3362, respectively) used to amplify the region comprising the SNP sites. The number of the sequences is per that of GenBank Accession No. (514739)

(57) Abrégé/Abstract:

Single nucleotide polymorphic sites of the bovine MAP1B, PPP1R11, and DDX4 genes are associated with improved bull fertility as measured by e.g. sire conception rates. Nucleic acid molecules, arrays, kits, methods of genotyping and marker-assisted bovine breeding methods based on these SNPs are disclosed.



ABSTRACT OF THE DISCLOSURE

Single nucleotide polymorphic sites of the bovine MAP1B, PPP1R11, and DDX4 genes are associated with improved bull fertility as measured by e.g. sire conception rates. Nucleic acid molecules, arrays, kits, methods of genotyping and marker-assisted bovine breeding methods based on these SNPs are disclosed.

SINGLE NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH BULL FERTILITY

FIELD OF THE INVENTION

[0001] The present invention relates to compositions and methods for testing and selecting cattle using molecular genetic methods by assaying for the presence of at least one genetic marker indicative of increased bull fertility. Specifically, genetic variations in the *MAP1B* and *PPP1R11* genes are tested and used for selecting cattle animals with improved blastocyst or fertilization rates, or both.

BACKGROUND OF THE INVENTION

[0002] The dairy cattle genome has been significantly restructured over the past 30 years due to intensive breeding effort selecting for production traits, including high quality milk and high and sustained productivity. However, while those efforts led to dramatic improvement of productivity, there has been significant reproductive deterioration in high-producing dairy cows, which in turn has caused substantial economic loss in the dairy cattle industry (Lucy, 2007, Fertility in high-producing dairy cows: reasons for decline and corrective strategies for sustainable improvement. *Society of Reproduction and Fertility Supplement*. 64 237-254). Key factors contributing to decreasing fertility of dairy cow are low fertilization rates and decreased embryonic survival.

[0003] Fertility is a complex trait that comprises developmental stages such as combining sperm and egg to form a zygote, compaction of embryo cells to form a morula, establishment of the blastocyst, attachment of the embryo to the uterus, and fetal development (Amann and DeJarnette, 2012). This complexity makes accurate prediction of successful pregnancy difficult, as aberrant development of sperm, oocyte, embryo, or fetus all would lead to conception failure. Conception rate in dairy cattle is about 40%, and only 50% of the fertilized eggs produce viable embryos (Santos et al., 2004). The decline in reproductive performance in cattle over the past few decades (Dobson et al., 2007) has been ascribed primarily to fertilization failure and early embryonic loss (Santos et al., 2004).

[0004] Previous studies have shown that genetic makeup of an individual plays crucial roles in embryonic development and reproductive success (Weigel, 2006; Shook, 2006). Although a male and female parent each contributes half of its genetic material to the new zygote and both are necessary for embryo development, it is not obvious whether or not this

contribution is equally important to pregnancy success. For example, it is well established that the paternal genome supports growth of extra-embryonic tissues while the maternal genome fosters development of the embryo proper (Barton et al., 1984). After fertilization, the development of an embryo is controlled by maternal genomic information that is accumulated during oogenesis (Telford et al., 1990). It is only at the 8-cell stage in the bovine embryo that the embryonic genome activates and the embryo switches to transcribing its own RNA (Memili and First, 2000).

[0005] Despite that most breeding schemes in cattle are focused on the selection of elite bulls using progeny testing or genomic selection, and that some semen traits (e.g., sperm motility and percentage of abnormal sperm) show moderate to high heritabilities (Druet et al., 2009), most fertility studies in cattle have focused on the maternal contribution, and the paternal contribution to reproductive performance has not been thoroughly investigated, and only a few studies have been reported in the literature (Feugang et al., 2009; Khatib et al., 2010; Peñagaricano et al., 2012). Therefore, characterization of bull fertility markers is both feasible and highly desirable, and the deployment of these markers in cattle breeding would lead to improved reproductive performance in cattle.

[0006] A recent comparative genomics study has characterized many genes involved in the control of spermatogenesis that were highly conserved from fly to human (Bonilla and Xu, 2008). Some of these genes were reported to be crucial for human fertility. However, it is not known whether or not these spermatogenesis genes play important roles in the fertility of bulls.

SUMMARY OF THE INVENTION

[0007] The present inventor carried out an association analysis between highly conserved spermatogenesis genes and sire conception rate (SCR) as a measure of bull fertility, with the objective that significant polymorphisms associated with bull fertility can be used as genetic markers in breeding programs aimed at improving reproductive performance in cattle.

[0008] Specifically, an association analysis is performed between highly conserved spermatogenesis genes and SCR in US Holstein populations as a measure of bull fertility. Sequence analysis revealed 24 single nucleotide polymorphisms (SNPs) in 9 genes in the bull population using the pooled DNA sequencing approach. These 9 genes were selected for their high level of sequence conservation between flies and humans. Overall, the 24 SNPs were

tested for association with SCR in a population of 1,988 bulls. Three SNPs located in the MAP1B gene, one SNP in the PPP1R11 gene and one SNP in the DDX4 gene showed significant associations with SCR. Nucleotide probes based upon these SNPs are found to be useful for genetic testing of bull animals for improved fertilization rate.

[0009] Accordingly, in one embodiment, the present invention provides an isolated oligo- or poly-nucleotide molecule consisting of

- 1) a nucleotide of Position 3066 of SEQ ID NO:1 and at least 12, but not more than 200 contiguous nucleotides of Fig. 1 adjacent to position 3066;
- 2) Position 3323 of SEQ ID NO:1 and at least 12, but not more than 200 contiguous nucleotides of Fig. 1 adjacent to position 3323;
- 3) Position 87071 of SEQ ID NO:2 and at least 12, but not more than 200 contiguous nucleotides of Fig. 3 adjacent to position 87071,
- (4) Position 112 of SEQ ID NO:3 and at least 12, but not more than 200 contiguous nucleotides of SEQ ID NO:3 adjacent to position 112,
- (5) Position 61646 of SEQ ID NO:4 and at least 12, but not more than 200 contiguous nucleotides of Fig. 6 adjacent to position 61646, and
- 6) Position 34239 of SEQ ID NO:5 and at least 12, but not more than 200 contiguous nucleotides of Fig. 5 adjacent to position 61646.

[0010] In one embodiment, the nucleotide molecule of the present invention comprises at least about 15 contiguous nucleotides adjacent to its respective position (hereinafter the "SNP position") of the respective figure. In one embodiment, the nucleic acid molecule of the present invention comprises at least about 20 contiguous nucleotides adjacent to the respective SNP position. In one embodiment, the oligonucleotide molecule of the present invention consists of not more than about 100 nucleotides. In one embodiment, the oligonucleotide molecule of the present invention consists of not more than about 50 nucleotides. In one embodiment, the SNP position of the nucleotide molecule of the present invention near or at the center of the molecule; alternatively, the SNP position is at the 3'-end of the oligonucleotide molecule.

[0011] Also provided herein is an array of nucleic acid molecules, comprising the isolated oligonucleotide molecule of the present invention, supported on a substrate. The substrate may be any suitable medium, known and readily available to one of ordinary skills in the art, and the array may be addressable.

[0012] The present invention further provides a kit comprising an isolated oligonucleotide molecule of the present invention, and a suitable container.

[0013] In another embodiment, the present invention provides a method for detecting single nucleotide polymorphism (SNP) in a gene listed in Table 1 below in a bovine cell, the method comprising optionally isolating an DNA from the bovine cell, determining the identity of a nucleotide on the gene of the cell at a SNP position identified in Table 1 below, and comparing the identity to the preferred nucleotide identity at a corresponding position in Table 1.

[0014] In one embodiment, the bovine cell may be an adult cell, an embryo cell, a sperm, an egg, a fertilized egg, or a zygote. The identity of the nucleotide may be determined by many methods known and readily available to those ordinarily skilled in the art, such as but not limited to sequencing a nucleic acid molecule comprising a suitable portion of the gene of the cell comprising a respective SNP position, or by hybridizing a suitable probe to a nucleic acid preparation from the cell, which probe may be suitably labeled e.g. fluorescently or radioactively.

[0015] The nucleic acid molecule may be isolated from the cell via a large variety of methods, known and readily available to an ordinarily skilled artisan, such as amplification by the polymerase chain reaction (PCR) of genomic DNA of the cell, or when appropriate, by RT-PCR of the mRNA of the cell.

[0016] In preferred embodiment, both copies of the gene in a diploid genome are genotyped according to the method of the present invention.

[0017] The identity of the nucleotide may be determined based on genotypes of the parent of the cell, genotypes of the daughter of the cell, or both, through genetic analysis methods well-known to those skilled in the art.

[0018] A method is further provided for determining whether an individual bovine animal is suitable as a gamete donor for natural mating, artificial insemination or *in vitro*

fertilization, the method comprising detecting the SNP according to the above method of the present invention, and excluding as gamete donor an individual which does not have the preferred allele identity at the respective SNP position as described in Table 1.

[0019] In one embodiment, the individual is excluded as a gamete donor if the individual, whose genotype is not homozygous of the preferred allele with regard to the respective SNP position.

[0020] The present invention additionally provides a method of selecting a bovine embryo for planting in a uterus, the method comprising genotyping the embryo according to the present invention, while preserving the viability of the embryo, and excluding from planting an embryo which does not have the preferred allele identity at the respective SNP position as described in Table 1.

[0021] In another embodiment, the present invention further provides a method for selectively breeding cattle using a multiple ovulation and embryo transfer procedure (MOET), the method comprising superovulating a female animal, collecting eggs from said superovulated female, in vitro fertilizing said eggs from a male animal which has at least one preferred allele identity at the respective SNP position as described in Table 1, implanting said fertilized eggs into other females allowing for an embryo to develop.

[0022] In another embodiment, the method for selectively breeding cattle using MOET may further comprise a step of genotyping the male animal according to the method of the present invention.

[0023] In another embodiment, the method for selectively breeding cattle using MOET may further comprise a step of genotyping the developing embryo, and allowing pregnancy to proceed only if the genotype of the embryo comprises at least one preferred allele identity at the respective SNP position as described in Table 1.

DESCRIPTION OF THE DRAWINGS

[0024] FIG. 1 depicts a partial genomic sequence of MAP1B, showing the locations of SNPs 1, 2 and 3 on the MAP1B gene and the locations of the primers (1F and 1R, corresponding to positions 2859-2876, and positions 3346-3362, respectively) used to amplify the region comprising the SNP sites. The numbering of the sequences is according to

that of GenBank Accession No. (Gene ID: 514739, updated on 12-Jul-2012), incorporated herein by reference in its entirety.

[0025] FIG. 2 is the partial genomic sequence of the MAP1B gene showing the locations of SNPs 4 and 6 on the MAP1B gene and the locations of the primers 2F and 2R (corresponding to positions 51447-51466, and positions 51931-51948, respectively used to amplify the region comprising the SNP sites. The numbering of the sequence is according to that of GenBank Accession No. (Gene ID: 514739, updated on 12-Jul-2012), incorporated herein by reference in its entirety.

[0026] FIG. 3 is the partial genomic sequence of the MAP1B gene showing the location of SNP 5 on the MAP1B gene and the locations of the primers 3F and 3R (corresponding to positions 86634-86651, and positions 87260-87278, respectively) used to amplify the region comprising the SNP sites. The numbering of the sequence is per that of GenBank Accession No. (Gene ID: 514739), incorporated herein by reference in its entirety.

[0027] FIG. 4. is the partial genomic sequence of the PPP1R11 gene showing the locations of SNPs 1-4 on the PPP1R11 gene and the locations of the primers 1F and 1R (corresponding to positions -102 to -85, and positions 470-489, respectively) used to amplify the region comprising the SNP sites. The numbering of the sequences is according to GenBank Accession No. (Gene ID: 504846), incorporated herein by reference in its entirety.

[0028] FIG. 5 is the partial genomic sequence of the DDX4 gene, showing the location of SNP 2 on the DDX4 gene and the locations of the primers 1F and 1R (corresponding to positions 34014-34031, and positions 34398-34417, respectively) used to amplify the region comprising the SNP site. The numbering of the sequences is according to that of GenBank Accession No. (Gene ID: 493725), incorporated herein by reference in its entirety.

[0029] FIG. 6 is the partial genomic sequence of the DDX4 showing the location of SNP 1 on the DDX4 gene and the locations of the primers 2F and 2R (corresponding to positions 61531-61549 , and positions 61867-61884, respectively) used to amplify the region comprising the SNP site. The number of the sequences is per that of GenBank Accession No. (Gene ID: 493725), incorporated herein by reference in its entirety.

DETAILED DESCRIPTION OF THE INVENTION

[0030] The present inventor has found that the spermatogenesis genes *MAP1B*, *PPP1R11* and *DDX4* showed significant associations with SCR. Table 1 below summarizes the SNPs demonstrated to be significantly associated with sire conception rates according to the present invention.

Table 1. Genetic markers significantly associated with sire conception rate

SNP ID	Location/Position	Nucleotide Identity of Preferred Allele	Nucleotide Identity of Polymorph in GenBank
SNP1.MAP1B	Position 3066 of Fig. 1	A	A
SNP3.MAP1B	Position 3323 of Fig. 1	T	T
SNP5.MAP1B	Position 87071 of Fig. 3	C	C
SNP1.PPP1R11	Position 112 of Fig. 4	G	T
SNP1.DDX4	Position 61646 of Fig. 6	A	A
SNP2.DDX4	Position 34239 of Fig. 5	G	G

[0031] Three SNPs in *MAP1B*, in low to moderate linkage disequilibrium (LD), were significantly associated with SCR. After correction for multiple testing, only one SNP in intron 5, SNP 5, showed the most significant association with SCR.

[0032] The *MAP1B* gene belongs to the microtubule-associated protein family and is known to affect neuronal development such as axon growth (Tymanskyj et al., 2012), development of dendritic spine and synaptic maturation (Tortosa et al., 2011), and regulation of the interaction between microtubules and actin microfilaments for axonal development (Montenegro-Venegas et al., 2010). Recent reports on the expression of *MAP1B* in the male reproductive tract in both rat and human (Queiróz et al., 2006) and in testis of fruit fly and mouse (Bonilla and Xu, 2008) suggest important functions of this gene in the regulation of male fertility. The finding by the present inventor, disclosed herein, that the *MAP1B* gene is associated with SCR, supports the conclusion that *MAP1B* plays a role in male fertility across a wide range of species.

[0033] The present inventor also found that a SNP in the 5'UTR of *PPP1R11* was associated with SCR in the bull population examined in this study. This is the first report of association between male fertility in cattle and *PPP1R11*, which is consistent with previous reports on the roles of this gene in spermatogenesis in mouse and human. For example, the different isoforms of *PPP1R11* (also known as *TCTEX5*) were found to be expressed in most mouse tissues with high expression in testis, epididymis, and in the head and tail regions of spermatozoa (Han et al., 2007). In a subsequent study, it was shown that mutations in the long transcript of *PPP1R11* were associated with normal sperm function (Han et al., 2008).

The authors concluded that *PPP1R11* plays important roles in sperm motility and spermatogenesis. A recent study reported that an isoform of protein phosphatase 1 (PP1 γ 2), which has an essential role in spermatogenesis, forms a complex with PPP1R11 in the testis (Cheng et al., 2009). Given that PP1 γ 2 is regulated by *PPP1R11*, these results further support the idea that *PPP1R11* has important functions in spermatogenesis.

[0034] The spermatogenesis genes investigated in this study were selected from a pool of genes whose expression is highly conserved in testis of both fruit fly and mouse (Bonilla and Xu, 2008). The protein sequence identities between cattle and human and between cattle and fly are 91% and 32%, respectively for *MAP1B* and 99% and 47%, respectively for *PPP1R11*. As such, the association of these genes with bull fertility testifies to the usefulness of the comparative genomics approach in selecting candidate male fertility genes.

[0035] To further explore involvement of male fertility genes identified in this study in female fertility, we tested the association of the SNPs in *MAP1B*, with fertilization and embryo survival rates using data from the IVF system. *MAP1B* genotypes of the cows, from which oocytes were extracted and used for fertilization and embryo culture, were significantly associated with differential fertilization rate and embryo survival rate. Recently, the expression of *MAP1B* was found to be downregulated in follicular cystic follicles compared to normal follicles, suggesting that alteration in *MAP1B* expression may be involved in reproduction failure in cattle (Choe et al., 2010). The instant disclosure again demonstrates the significance of both parental genomes to embryonic development and fertility.

[0036] Accordingly, the present invention provides nucleic acid-based genetic markers for identifying bovine animals, especially bulls, with superior fertility, specifically, sire conception rate as a measure of male fertility. In general, for use as markers, isolated oligonucleotide or polynucleotide molecules, or isolated nucleic acid fragments, preferably DNA fragments, as used. Such markers will be of at least 10 nucleotides (nt), preferably at least 11, 12, or 15 nt, usually at least 20 nt, often at least 50 nt. Such small DNA fragments are useful as primers for the polymerase chain reaction (PCR), and probes for hybridization screening, etc.

[0037] In one embodiment, the isolated nucleic acid molecule comprises not more than 1,500 nt, or not more than 1000 nt, or not more than 900 nt, or not more than 800 nt, or not

more than 700 nt, or not more than 600 nt, or not more than 500 nt, or not more than 400 nt, or not more than 300 nt, or not more than 200 nt., or not more than 150 nt., or not more than 100 nt., or not more than 75 nt.

[0038] In the context of the present invention, the term "isolated" refers to a nucleic acid molecule purified to some degree from endogenous materials with which the nucleic acid molecule may naturally occur or exist. At the least, the term "isolated" refers to a nucleic acid molecule separated from chromatin or other protein or components of the genomic DNA. Preferably, the isolated oligonucleic acid molecule or polynucleic acid molecule of the present invention comprises a fragment that is shorter than that which is naturally occurring.

[0039] In the context of the present invention, the provided sequences also encompass the complementary sequence corresponding to any of the provided polymorphisms. Where appropriate, and in order to provide an unambiguous identification of the specific site of a polymorphism, the numbering of the original nucleic sequences in the GenBank may be used; alternatively, the numbering may simply refer to the specific sequence in the Sequence Listing accompanying this disclosure.

[0040] The term primer refers to a single-stranded oligonucleotide capable of acting as a point of initiation of template-directed DNA synthesis under appropriate conditions (i.e., in the presence of four different nucleoside triphosphates and an agent for polymerization, such as, DNA or RNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature. The appropriate length of a primer depends on the intended use of the primer but typically ranges from 15 to 30 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template but must be sufficiently complementary to hybridize with a template. The term primer site, or priming site, refers to the area of the target DNA to which a primer hybridizes. The term primer pair means a set of primers including a 5' upstream primer that hybridizes with the 5' end of the DNA sequence to be amplified and a 3', downstream primer that hybridizes with the complement of the 3' end of the sequence to be amplified. One of these two primers is often referred to as the "forward primer," while the other the "reverse primer."

[0041] The term "probe" or "hybridization probe" denotes a defined nucleic acid segment (or nucleotide analog segment) which can be used to identify by hybridization a specific

polynucleotide sequence present in samples, said nucleic acid segment comprising a nucleotide sequence complementary of the specific polynucleotide sequence to be identified. "Probes" or "hybridization probes" are nucleic acids capable of binding in a base-specific manner to a complementary strand of nucleic acid.

[0042] An objective of the present invention is to determine which embodiment of the polymorphisms a specific sample of DNA has. For example, it is desirable to determine whether the nucleotide at a particular position is A or G. An oligonucleotide probe can be used for such purpose. Preferably, the oligonucleotide probe will have a detectable label, and contains an A at the corresponding position. Experimental conditions can be chosen such that if the sample DNA contains an A, they hybridization signal can be detected because the probe hybridizes to the corresponding complementary DNA strand in the sample, while if the sample DNA contains a G, no hybridization signal is detected.

[0043] Similarly, PCR primers and conditions can be devised, whereby the oligonucleotide is used as one of the PCR primers, for analyzing nucleic acids for the presence of a specific sequence. These may be direct amplification of the genomic DNA, or RT-PCR amplification of the mRNA transcript of the gene of interest. The use of the polymerase chain reaction is described in Saiki et al. (1985) Science 230:1350-1354. Amplification may be used to determine whether a polymorphism is present, by using a primer that is specific for the polymorphism. Alternatively, various methods are known in the art that utilize oligonucleotide ligation as a means of detecting polymorphisms, for examples see Riley et al (1990) Nucleic Acids Res. 18:2887-2890; and Delahunty et al (1996) Am. J. Hum. Genet. 58:1239-1246. The detection method may also be based on direct DNA sequencing, or hybridization, or a combination thereof. Where large amounts of DNA are available, genomic DNA is used directly. Alternatively, the region of interest is cloned into a suitable vector and grown in sufficient quantity for analysis. The nucleic acid may be amplified by PCR, to provide sufficient amounts for analysis.

[0044] Hybridization may be performed in solution, or such hybridization may be performed when either the oligonucleotide probe or the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent

to synthesis. Solid-supports suitable for use in detection methods of the invention include substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide or target nucleic acid. For screening purposes, hybridization probes of the polymorphic sequences may be used where both forms are present, either in separate reactions, spatially separated on a solid phase matrix, or labeled such that they can be distinguished from each other.

[0045] Hybridization may also be performed with nucleic acid arrays and subarrays such as described in WO 95/11995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites. One or both polymorphic forms may be present in the array. Usually such an array will include at least 2 different polymorphic sequences, i.e. polymorphisms located at unique positions within the locus, and may include all of the provided polymorphisms. Arrays of interest may further comprise sequences, including polymorphisms, of other genetic sequences, particularly other sequences of interest. The oligonucleotide sequence on the array will usually be at least about 12 nt in length, may be the length of the provided polymorphic sequences, or may extend into the flanking regions to generate fragments of 100 to 200 nt in length. For examples of arrays, see Ramsay (1998) Nat. Biotech. 16:4044; Hacia et al. (1996) Nature Genetics 14:441-447; Lockhart et al. (1996) Nature Biotechnol. 14:1675-1680; and De Risi et al. (1996) Nature Genetics 14:457-460.

[0046] The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using riboprobes (Winter et al., Proc. Natl. Acad. Sci. USA 82:7575, 1985; Meyers et al., Science 230:1242, 1985) and proteins which recognize nucleotide mismatches, such as the *E. coli* mutS protein (Modrich, P. Ann. Rev. Genet. 25:229-253, 1991). Alternatively, variant alleles can be identified by single strand conformation polymorphism (SSCP) analysis (Orita et al., Genomics 5:874-879, 1989; Humphries et al., in Molecular Diagnosis of Genetic Diseases, R. Elles, ed., pp. 321-340, 1996) or denaturing gradient gel electrophoresis (DGGE) (Wartell et al., Nucl. Acids Res. 18:2699-2706, 1990; Sheffield et al., Proc. Natl. Acad. Sci. USA 86:232-236, 1989).

[0047] A polymerase-mediated primer extension method may also be used to identify the polymorphism(s). Several such methods have been described in the patent and scientific literature and include the "Genetic Bit Analysis" method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Pat. No. 5,679,524). Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Pat. Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism may be detected by mass spectrometry as described in U.S. Pat. No. 5,605,798. Another primer extension method is allele-specific PCR (Ruao et al., Nucl. Acids Res. 17:8392, 1989; Ruao et al., Nucl. Acids Res. 19, 6877-6882, 1991; WO 93/22456; Turki et al., J. Clin. Invest. 95:1635-1641, 1995). In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al. (WO 89/10414).

[0048] A detectable label may be included in an amplification reaction. Suitable labels include fluorochromes, e.g. fluorescein isothiocyanate (FITC), rhodamine, Texas Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein (JOE), 6-carboxy-X-rhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA), radioactive labels, e.g. ³²P, ³⁵S, ³H; etc. The label may be a two stage system, where the amplified DNA is conjugated to biotin, haptens, etc. having a high affinity binding partner, e.g. avidin, specific antibodies, etc., where the binding partner is conjugated to a detectable label. The label may be conjugated to one or both of the primers. Alternatively, the pool of nucleotides used in the amplification is labeled, so as to incorporate the label into the amplification product.

[0049] It is readily recognized by those ordinarily skilled in the art that in order to maximize the signal to noise ratio, in probe hybridization detection procedure, the polymorphic site should be at the center of the probe fragment used, whereby a mismatch has a maximum effect on destabilizing the hybrid molecule; and in a PCR detection procedure, the polymorphic site should be placed at the very 3'-end of the primer, whereby a mismatch has the maximum effect on preventing a chain elongation reaction by the DNA polymerase. The location of nucleotides in a polynucleotide with respect to the center of the polynucleotide are described herein in the following manner. When a polynucleotide has an odd number of nucleotides, the nucleotide at an equal distance from the 3' and 5' ends of the polynucleotide

is considered to be “at the center” of the polynucleotide, and any nucleotide immediately adjacent to the nucleotide at the center, or the nucleotide at the center itself is considered to be “within 1 nucleotide of the center.” With an odd number of nucleotides in a polynucleotide any of the five nucleotides positions in the middle of the polynucleotide would be considered to be within 2 nucleotides of the center, and so on. When a polynucleotide has an even number of nucleotides, there would be a bond and not a nucleotide at the center of the polynucleotide. Thus, either of the two central nucleotides would be considered to be “within 1 nucleotide of the center” and any of the four nucleotides in the middle of the polynucleotide would be considered to be “within 2 nucleotides of the center,” and so on.

[0050] In some embodiments, a composition contains two or more differently labeled oligonucleotides for simultaneously probing the identity of nucleotides or nucleotide pairs at two or more polymorphic sites. It is also contemplated that primer compositions may contain two or more sets of allele-specific primer pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic site.

[0051] Alternatively, the relevant portion of the gene of the sample of interest may be amplified via PCR and directly sequenced, and the sequence be compared to the wild type sequence shown in the figures. It is readily recognized that, other than those disclosed specifically herein, numerous primers can be devised to achieve the objectives. PCR and sequencing techniques are well known in the art and reagents and equipments are readily available commercially.

[0052] Alternatively, an invasive signal amplification assay, as described in e.g. U.S. Pat. No. 5,422,253 and Lyamichev et al., 2000, Biochemistry 39:9523-9532, both incorporated herein by reference in their entirety, may be used for detecting the SNP of interest. This assay takes advantage of enzymes such as the 5' nuclease activity of a DNA polymerase or the gene 6 product from bacteriophage T7 in their ability to cleave polynucleotide molecules by recognizing specific structures instead of specific sequences. A single-stranded target molecule is annealed to a pilot oligonucleotide such that the 5' end of the pilot forms a duplex with the target molecule. If the 3' end of the pilot oligonucleotide does not pair with the target, a 3' arm is formed. When exposed to a cleavage agent such as a DNA polymerase having a 5' nuclease activity or the gene 6 product from bacteriophage T7, the target molecule is cleaved in the 5' region, one nucleotide into the duplex adjacent to the unpaired region of the target. If a cut in a double-stranded molecule is required, the double-stranded

molecule is denatured. Because this unpaired 3' arm can be as short as one nucleotide, this assay can be used for detecting a single-nucleotide difference, e.g. in the context of SNP detection. The pilot oligonucleotide is designed such that it pairs perfectly with one allele, but has a 3', single nucleotide mismatch with another allele. Cleavage only occurs if there is a mismatch between the target molecule and the pilot. To achieve signal amplification, the above invasive reaction is modified such that cleavage occurs on the pilot oligonucleotide. Two oligonucleotides are annealed in an adjacent manner to the target molecule. The resulting adjacent duplexes overlaps by at least one nucleotide to create an efficient substrate, called the overlapping substrate, for the 5' nucleases. The 5' end of the downstream oligonucleotide, also called the probe, contains an unpaired region termed the 5' arm (Lyamichev et al., 1993, *Science* 260:778-783.) or flap (Harrington and Lieber, 1994, *EMBO J* 13: 1235-1246) that is not required for the enzyme activity; however, very long arms can inhibit cleavage (Lyamichev et al., 1993, *Science* 260:778-783). Specific cleavage of the probe, termed invasive cleavage (Lyamichev et al., 1999, *Nat. Biotechnol.* 17 292-296; Kwiatkowski et al., 1999, *Mol. Diagn.* 4, 353-364.), occurs at the position defined by the 3' end of the upstream oligonucleotide, which displaces or "invades" the probe. If the overlap between the adjacent oligonucleotides is only one nucleotide, cleavage takes place between the first two base pairs of the probe, thus releasing its 5' arm and one nucleotide of the base paired region (Lyamichev et al., 1999, *Proc. Natl. Acad. Sci. USA.* 96: 6143-6148, and Kaiser et al., 1999, *J Biol. Chem.* 274:21387-21394). If the upstream oligonucleotide and the probe are present in large molar excess over the target nucleic acid, and invasive cleavage is carried out near the melting temperature of the probe, a cut probe can rapidly dissociate, and an intact probe will anneal to the target more frequently than will a cut probe, thus initiating a new cycle of cleavage. This allows multiple probes to be cut for each target molecule under isothermal conditions, resulting in linear signal amplification with respect to target concentration and time (Lyamichev et al., 1999, *Nat. Biotechnol.* 17: 292-296).

[0053] DNA markers have several advantages; segregation is easy to measure and is unambiguous, and DNA markers are co-dominant, i.e., heterozygous and homozygous animals can be distinctively identified. Once a marker system is established selection decisions could be made very easily, since DNA markers can be assayed any time after a blood sample can be collected from the individual infant animal, or even earlier by testing embryos *in vitro* if very early embryos are collected. The use of marker assisted genetic selection will greatly facilitate and speed up cattle breeding problems. For example, a

modification of the multiple ovulation and embryo transfer (MOET) procedure can be used with genetic marker technology. Specifically, females are superovulated, eggs are collected, *in vitro* fertilized using semen from superior males and implanted into other females allowing for use of the superior genetics of the female (as well as the male) without having to wait for her to give birth to one calf at a time. Developing blastomeres at the 4-8 cell stage may be assayed for presence of the marker, and selection decisions made accordingly.

[0054] In one embodiment of the invention an assay is provided for detection of presence of a desirable genotype using the markers.

[0055] The term “genotype” as used herein refers to the identity of the alleles present in an individual or a sample. In the context of the present invention a genotype preferably refers to the description of the polymorphic alleles present in an individual or a sample. The term “genotyping” a sample or an individual for a polymorphic marker refers to determining the specific allele or the specific nucleotide carried by an individual at a polymorphic marker.

[0056] The present invention is suitable for identifying a bovine, including a young or adult bovine animal, an embryo, a semen sample, an egg, a fertilized egg, or a zygote, or other cell or tissue sample therefrom, to determine whether said bovine possesses the desired genotypes of the present invention, some of which are indicative of improved reproduction traits.

[0057] Further provided is a method for genotyping the bovine genes identified in Table 1, comprising determining for the two copies of the gene in a diploid genome present the identity of the nucleotide pair at the relevant SNP position (see below).

[0058] One embodiment of a genotyping method of the invention involves examining both copies of the gene, or a fragment thereof, to identify the nucleotide pair at the polymorphic site in the two copies to assign a genotype to the individual. In some embodiments, “examining a gene” may include examining one or more of: DNA containing the gene, mRNA transcripts thereof, or cDNA copies thereof. As will be readily understood by the skilled artisan, the two “copies” of a gene, mRNA or cDNA, or fragment thereof in an individual may be the same allele or may be different alleles. In another embodiment, a genotyping method of the invention comprises determining the identity of the nucleotide pair at the polymorphic site.

[0059] The present invention further provides a kit for genotyping a bovine sample, the kit comprising in a container a nucleic acid molecule, as described above, designed for detecting the polymorphism, and optionally at least another component for carrying out such detection. Preferably, a kit comprises at least two oligonucleotides packaged in the same or separate containers. The kit may also contain other components such as hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may contain, preferably packaged in separate containers, a polymerase and a reaction buffer optimized for primer extension mediated by the polymerase, such as PCR.

[0060] In one embodiment the present invention provides a breeding method whereby genotyping as described above is conducted on a bovine animal, especially a bull, and based on the results, certain cattle are either selected or dropped out of the breeding program.

[0061] Through use of the linked marker loci, procedures termed “marker assisted selection” (MAS) may be used for genetic improvement within a breeding nucleus; or “marker assisted introgression” for transferring useful alleles from a resource population to a breeding nucleus (Soller 1990; Soller 1994).

[0062] A method is further provided for determining whether an individual bovine animal is suitable as a gamete donor for natural mating, artificial insemination or in vitro fertilization, the method comprising determining the identity of one or more SNPs according to the present invention using a method of the present invention, and excluding as gamete donor an individual which does not have the preferred allele identity at the respective SNP position as described in Table 1.

[0063] Specifically, an individual bovine animal, especially a bull, is excluded as a gamete donor, if its genome does not have at least:

- 1) adenine (A) at the position of its MAP1B gene corresponding to position 3066 of Fig. 1;

- 2) thymine (T) at the position of its MAP1B gene corresponding to position 3323 of Fig. 1;

- 3) cytosine (C) at the position of its MAP1B gene corresponding to position 87071 of Fig. 3;

4) guanine (G) at the position of its PPP1R11 gene corresponding to position 112 of Fig. 4;

5) adenine (A) at the position of its DDX4 gene corresponding to position 61646 of Fig. 6, or

6) G at the position of its DDX4 gene corresponding to position 34239 of Fig. 5.

[0064] In one embodiment, the individual is excluded as a gamete donor if the genotype of the individual is not homozygous of the preferred allele with regard to the respective SNP position.

[0065] The present invention additionally provides a method of selecting a bovine embryo for planting in a uterus, the method comprising genotyping the embryo according to the present invention, while preserving the viability of the embryo, and excluding from planting an embryo which does not have the preferred allele identity at the respective SNP position as described in Table 1.

[0066] In another embodiment, the present invention further provides a method for selectively breeding cattle using a multiple ovulation and embryo transfer procedure (MOET), the method comprising superovulating a female animal, collecting eggs from said superovulated female, in vitro fertilizing said eggs from a male animal which has at least one preferred allele identity at the respective SNP position as described in Table 1 and above, implanting said fertilized eggs into other females allowing for an embryo to develop.

[0067] In another embodiment, the method for selectively breeding cattle using MOET may further comprise a step of genotyping the male animal according to the method of the present invention.

[0068] In another embodiment, the method for selectively breeding cattle using MOET may further comprise a step of genotyping the developing embryo, and allowing pregnancy to proceed only if the genotype of the embryo comprises at least one preferred allele identity at the respective SNP position as described in Table 1.

[0069] The following examples are intended to illustrate preferred embodiments of the invention and should not be interpreted to limit the scope of the invention as defined in the claims.

EXAMPLES

[0070] Materials and methods

[0071] Associations of candidate fertility genes examined were carried out in two experiments. In the first part, single nucleotide polymorphisms (SNPs) in the spermatogenesis genes were tested for associations with sire conception rate (SCR) in a large bull population. In the second part, genes found significantly associated with SCR were tested for association with female fertility traits (fertilization and blastocyst rates). Male fertility genes that play roles in female fertility can be used to improve reproductive performance in cattle using genetic information from both males and females in breeding schemes.

[0072] Gene Selection, SNP Identification, and Genotyping for Bull Fertility

[0073] A total of 58 spermatogenesis genes, with conserved testicular expression from fly to human, were reported in Bonilla and Xu (2008). Of those genes, only 22 were annotated in the bovine genome. For SNP identification and genotyping, genomic DNA was extracted from semen samples of 268 Holstein bulls (Genex Cooperative/CRI, Shawano, WI) using standard phenol/chloroform protocols. One DNA pool was constructed from 20 random semen samples with equal amounts of DNA. The DNA pool was amplified using primers designed in the 22 candidate genes to amplify 5' untranslated regions (UTRs), exons, introns, and 3' UTRs. The PCR products were sequenced, and SNPs were identified by visually inspecting sequence traces. PCR amplification and sequencing were performed as described in Khatib et al. (2008). Table 2 shows the primer sets used to amplify the nine candidate spermatogenesis genes found to be polymorphic in the bull population.

Table 2. Primers used to amplify the nine spermatogenic genes

Gene	Forward Primer	SEQ ID NO:	Reverse Primer	SEQ ID NO:	Product Size
<i>DCUN1D1</i>	ATACCCTTAGGCAGTTAG	7	AATTGTAAACCCTGAGAC	8	536
<i>DDX4(1)</i>	AAACACGGAACAGAGGGT	9	AGGCAGGATTAGCAAGTATG	10	404
<i>DDX4(2)</i>	AACCAAGTGGCTGGGATG	11	CAGACTCAAATGCGACAA	12	354
<i>DNAI1(1)</i>	CGGTAAGTGAGCAGCATC	13	ACTGAAGCCTTTGCCCTA	14	495
<i>DNAI1(2)</i>	CCCAGTGCTCCAAATCCT	15	ATGGCTCATCTTGTCTTCAGTA	16	413
<i>DNAI1(3)</i>	CGTGACTGGGTTTAGGAT	17	CTGGTGGCTGCTGTCTAT	18	602
<i>GAPDHS(1)</i>	CCAGGAAACGGCATCACC	19	ACACGCAGCAGGGCAACT	20	414
<i>GAPDHS(2)</i>	GTGAAGGCCAGGGACTATGA	21	ACATGAACAAGAGGGCTGCT	22	541
<i>GSTM3</i>	TTCTCTCCCTGCAAGTCGT	23	TGAGAACAGCTGCCATCATC	24	664
<i>MAP1B(1)</i>	CCATTTCCTAAGGCACAG	25	TTCCGCCATCTTCCTACA	26	504
<i>MAP1B(2)</i>	CTTATGGTCGTGATTATGAA	27	AAGGCTAACACTGCTGGT	28	502
<i>MAP1B(3)</i>	GGCTGTGACATACCTACC	29	CAGACCTTCCCTACTTATT	30	645
<i>PPP1R1(1)</i>	CACATTACGGCGGAACATA	31	ATCCCAAGCAGTATCACCTA	32	591
<i>PPP1R1(2)</i>	ACCTGTTCTATCTCCTCCCA	33	GTCACCTACCCACCTTGC	34	543
<i>SPATA20</i>	TTGGAGAAGAAACCCACCAG	35	CCTCACAAGCAAGGCTAAGG	36	459
<i>UBC</i>	TCGCTCAGTCGTGTCTTAC	37	TCAACCAACGCCTAATGT	38	420

[0074] A total of 24 SNPs located in the nine spermatogenic genes were genotyped in the Genex population (268 animals) by MALDI-TOF MS (GeneSeek Inc., Lincoln, NE).

[0075] Imputation of SNPs for validation of significant SNPs found in Genex population in a larger bull population obtained from the USDA

[0076] Solely for the purpose of validating the conclusions drawn from the results using the Genex population, and to increase the sample size and improve the statistical power of the study, the 24 SNPs identified in the 9 candidate genes and genotyped in 268 bulls (from here forward, the reference population) were imputed in a total of 1,720 bulls (from here forward, the imputed population) so that a final dataset of 1,988 bulls with genotypic data was generated for subsequent statistical analyses. Bulls in the reference and imputed populations have been previously genotyped with the Illumina BovineSNP50 Bead Chip, and hence shared SNP were used to infer the genotypes of the unshared SNP in the imputed population. Genotypes of the reference ($n = 268$) and the imputed populations ($n = 1,720$) for the 50K SNP Chip were provided by Genex Cooperative/CRI (Shawano, WI) and the Animal Improvement Programs Laboratory of the United States Department of Agriculture (AIPL, USDA; Beltsville, MD), respectively. SNP with minor allele frequencies below 5% were removed. After data editing, 38,265 SNP spanning the entire bovine genome were available in both populations for the imputation process.

[0077] Imputation of SNPs was carried out for each candidate gene separately. In each case, a total of 100 SNP on each side of the gene were used to infer the genotypes of the ungenotyped SNP. Imputation was performed using the population-based haplotype clustering algorithm of Scheet and Stephens (2006), which was implemented via the fastPHASE version 1.2 software using the default settings for all parameters (University of Washington TechTransfer Digital Ventures Program, Seattle, WA).

[0078] Phenotypic Data for Bull Fertility

[0079] The 1,988 bulls genotyped with the 50K SNP Chip were evaluated for sire conception rate (SCR), a phenotypic evaluation of bull fertility provided to dairy producers by AIPL-USDA as described in Peñagaricano et al. (2012). Briefly, SCR is the expected difference in conception rate of a sire compared with the mean of all other evaluated sires (Kuhn and Hutchison, 2008; Kuhn et al., 2008).

[0080] In this study, SCR values ranged from -10.66% to +6.80%, and the number of breedings per bull ranged from 303 to 111,402. SCR data were obtained from seven consecutive evaluations provided by AIPL-USDA between August 2008 and December 2010. For bulls with multiple evaluations, the most recent SCR evaluation was used in the analysis.

[0081] **Statistical Analysis for Bull Fertility**

[0082] The association between each SNP and SCR was evaluated using the following mixed linear model,

$$SCR_{ijkl} = \mu + EVAL_j + \beta SNP_k + sire_l + e_{ijkl}$$

where μ is the general mean, $EVAL_j$ is the fixed effect of the j^{th} AIPL-USDA SCR evaluation ($j = 1, 2, \dots, 7$), SNP_k is the number of copies of one allele of the SNP (corresponding to 0, 1 or 2 copies) carried by the i^{th} animal ($i = 1, 2, \dots, 1988$), β is the regression coefficient for the SNP considered (also known as the allele substitution effect), $sire_l$ represents the random additive genetic effect of the l^{th} sire ($l = 1, 2, \dots, 246$) of the i^{th} animal, and e_{ijkl} represents the random residual for each observation. To detect possible deviations from the additive model, associations between genotype and SCR were evaluated using SNP as a categorical variable.

[0083] Random effects were assumed to follow the multivariate normal distribution,

$$\begin{pmatrix} \mathbf{s} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim N \left[\mathbf{0}, \begin{pmatrix} \mathbf{A}\sigma_s^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{W}^{-1}\sigma_\varepsilon^2 \end{pmatrix} \right]$$

where $\mathbf{s}$ and $\boldsymbol{\varepsilon}$ are the vectors of sire and residual effects, respectively; σ_s^2 and σ_ε^2 are the sire and residual effect variances, respectively; $\mathbf{A}$ represents the matrix of additive relationships between sires in the pedigree ($1,558 \times 1,558$) and $\mathbf{W}$ is a diagonal matrix of order 1,988 with its elements representing reliabilities of SCR values. The $\mathbf{A}$ matrix was calculated based on a five-generation pedigree of sires downloaded from AIPL-USDA. The association between each SNP and SCR was tested using a likelihood ratio test by comparing to a reduced model without the SNP effect.

[0084] **Phenotypic and Genotypic Data for Cow Fertility**

[0085] The most significant SNP for SCR (rs109423562 located in *MAP1B*) was further investigated for association analysis with fertilization and blastocyst rates—the main cow fertility traits— using an IVF system. The procedures of in vitro fertilization and subsequent embryo culture were described in Khatib et al. (2008). To generate fertilization and blastocyst rate data, a total of 6,282 in vitro fertilizations were performed, and a total of 4,207 embryos were produced using oocytes from 359 ovaries collected from 359 Holstein cows and semen samples from 12 Holstein bulls. For 74 ovaries, oocytes were fertilized by two different bulls each. Fertilization rate was calculated as the number of cleaved embryos at Day 2 post-fertilization divided by the total number of fertilized oocytes collected from one ovary. Blastocyst rate was calculated as the number of embryos that reached the blastocyst stage (Day 8) and appeared normal out of the total number of embryos produced.

[0086] The 359 ovaries were genotyped for SNP rs109423562 (G/A) using PCR – RFLP. A 171 bp fragment was amplified using the primers 5'-GCAGCTCTTTTAGGAGTGTTAGCGTCTGAT-3' (SEQ ID NO: 39) (forward) and 5'-CTCACAGAGGGCATTGACA-3' (SEQ ID NO: 40) (reverse). The PCR product was then digested by the restriction enzyme *HinfI* and electrophoresed on a 2.0% agarose gel. Allele G was cut while allele A was uncut.

[0087] Statistical Analysis for Cow Fertility

[0088] Association between SNP rs109423562 (G/A) in *MAP1B* and fertilization and blastocyst rates were analyzed using the following mixed linear model,

$$y_{ijk} = \mu + \text{ovary}_i + \text{sire}_j + \text{SNP}_{ijk} + e_{ijk}$$

where y_{ijk} represents the fertilization or blastocyst rate of oocyte k from ovary i fertilized with semen from bull j , μ represents a general mean for the trait considered, ovary_i represents the random effect of the individual ovary from which oocytes were harvested, sire_j represents the random effect of the sire used in the fertilization, SNP_{ijk} represents the fixed effect of the ovary genotype for the SNP considered, and e_{ijk} represents the residuals, assumed normal, independent and identically distributed with mean 0 and variance $\text{I}\sigma_e^2$. Ovaries and bulls were assumed uncorrelated with variance structures $\text{I}\sigma_o^2$ and $\text{I}\sigma_s^2$, respectively. Association between the SNP and fertilization or blastocyst rate was tested again using a likelihood ratio test by comparing with a reduced model without the SNP effect. All the statistical analyses

were performed using the *pedigreemm* package (Vazquez et al. 2010) of the R language/environment (R Development Core Team 2009).

RESULTS

[0089] SNP Identification and Association of Candidate Genes with SCR

[0090] Sequencing analysis revealed 24 SNPs in 9 spermatogenesis genes (*DCUN1D1*, *DNAI1*, *DDX4*, *GAPDHS*, *GSTM3*, *MAP1B*, *PPP1R11*, *SPATA20* and *UBC*) in the bull population using the pooled DNA sequencing approach. All 24 SNPs located in 9 candidate genes were tested for association with sire conception rate first in Genex population and then in a larger population of 1,988 bulls for validation.

[0091] Association analysis in Genex population

[0092] SNPs in the genes *MAP1B*, *PPP1R11*, and *DDX4* are associated with SCR

[0093] For *MAP1B* gene, a SNP C/T in intron 5 at position 9331992 (University of Maryland bovine version 3.1; UMD3.1) showed significant association with sire conception rate. Primers used to amplify the gene and SNP location are shown in Figure 1. Frequency of allele T was 25% and frequency of allele C was 75%. The allele substitution effect was -0.15 ± 0.01 (p-value = 0.01) so that allele C is favorable for SCR.

[0094] For *PPP1R11*, one SNP (T/G) in the 5'UTR region at position 28710268 (UMD3.1) was significantly associated with SCR in the Genex population, with allele G associated with increased SCR. SEQ ID NO:6 shows SNP location and primers used to identify the SNP.

[0095] For *DDX4*, two SNPs located at positions 23382814 (SNP1; A/G) and 23410221 (SNP2; G/A) were identified in the gene. For SNP locations see Figure 3. The two SNPs were in almost complete linkage disequilibrium, so they have the same allele frequencies in the bull population examined. Genotype AA of SNP1 has 0.783 units of SCR versus -0.214 SCR units for GG genotype (P-value = 0.05). Similarly, GG genotype of SNP2 has 0.749 units of SCR versus -0.494 for AA genotype.

[0096] Association of spermatogenesis genes in the combined Genex and USDA populations

[0097] Three SNPs located in *MAP1B* and one SNP in *PPP1R11* showed significant associations with SCR (Table 3). The SNP with the most significant association with SCR is SNP5.MAP1B located in intron 5 with an allele substitution effect of -0.24 and a *P*-value of 0.001. The other two significant SNPs were SNP1.MAP1B and SNP3.MAP1B, both located in intron 1 with an allele substitution effect of 0.15 and *P*-values of 0.025 and 0.039, respectively. After Bonferroni correction for multiple testing, only SNP5.MAP1B remained significant (*P*-value = 0.024). Pair-wise linkage disequilibrium (LD) tests of *MAP1B* SNPs showed a moderate LD ($r^2 = 0.38$) between SNP1.MAP1B and SNP3.MAP1B and SNP5.MAP1B. The LD between SNP3.MAP1B and SNP5.MAP1B was relatively low ($r^2 = 0.14$).

[0098] SNP1.PPP1R11 located in the 5'UTR region of *PPP1R11* showed significant association with SCR with an allele substitution effect of 0.15 and a *P*-value of 0.046.

Table 3. Genetic markers significantly associated with sire conception rate

SNP ID	Gene	Genotype (N)	Allele substitution effect ± SE	<i>P</i> -value
SNP1.MAP1B	<i>MAP1B</i>	GG (463)	0.15 ± 0.01	0.025
		GA (987)		
		AA (538)		
SNP3.MAP1B	<i>MAP1B</i>	CC (972)	0.15 ± 0.01	0.039
		CT (847)		
		TT (169)		
SNP5.MAP1B	<i>MAP1B</i>	CC (1081)	-0.24 ± 0.01	0.001
		TC (763)		
		TT (144)		
SNP1.PPP1R11	<i>PPP1R11</i>	TT (1033)	0.15 ± 0.01	0.046
		TG (821)		
		GG (134)		

[0099] Association of MAP1B with Fertilization Rate and Embryo Survival Rate

[00100] SNP5.MAP1B located in intron 5 of *MAP1B* showed the most significant association with sire conception rate in the bull population analyzed above. To characterize its impact on female fertility, we tested the association of this SNP with fertilization rate and blastocyst rate in the IVF system. SNP5.MAP1B showed significant associations with both fertilization rate (*P*-value = 0.027) and blastocyst rate (*P*-value = 0.029) (Table 4). Oocytes collected from genotype CT cows showed the lowest fertilization rate (59.9%) compared with that from CC (66.4%) and TT (66.3%) cows (Table 4). For blastocyst rate, the CT genotype

again showed the lowest rate (27.1%) while homozygous CC and TT individuals showed blastocyst rates of 31.0% and 41.8%, respectively (Table 4).

Table 4. Association between SNP5 of *MAP1B* gene and fertilization rate and blastocyst rate

Genotype (N)	Fertilization rate		Blastocyst rate	
	Estimate $\pm$ SE	<i>P</i> -value	Estimate $\pm$ SE	<i>P</i> -value
CC (321)	0.664 $\pm$ 0.03	0.027	0.310 $\pm$ 0.02	0.029
CT (84)	0.599 $\pm$ 0.03		0.271 $\pm$ 0.03	
TT (28)	0.663 $\pm$ 0.04		0.418 $\pm$ 0.05	

REFERENCES

- Amann, R.P., J. M. Dejarnette. 2012. Impact of genomic selection of AI dairy sires on their likely utilization and methods to estimate fertility: A paradigm shift. *Theriogenology* 77:795-817.
- Barton, S. C., M. A. Surani, and M. L. Norris. 1984. Role of paternal and maternal genomes in mouse development. *Nature*. 311:374–376.
- Bonilla, E., and E. Y. Xu. 2008. Identification and characterization of novel mammalian spermatogenic genes conserved from fly to human. *Mol. Hum. Reprod.* 14:137–142.
- Cheng, L., S. Pilder, A. C. Nairn, S. Ramdas, and S. Vijayaraghavan. 2009. PP1gamma2 and PPP1R11 are parts of a multimeric complex in developing testicular germ cells in which their steady state levels are reciprocally related. *PLoS One*. 4:e4861.
- Choe, C., Y. W. Cho, C. W. Kim, D. S. Son, J. Han, and D. Kang. 2010. Identification of differentially expressed genes in bovine follicular cystic ovaries. *Korean J. Physiol. Pharmacol.* 14:265–272.
- Dobson, H., R. Smith, M. Royal, Ch. Knight, and I. Sheldon. 2007. The high-producing dairy cow and its reproductive performance. *Reprod Domest Anim.* 42 Suppl 2:17-23.
- Druet, T., S. Fritz, E. Sellem, B. Basso, O. Gérard, L. Salas-Cortes, P. Humblot, X. Druart, and A. Eggen. 2009. Estimation of genetic parameters and genome scan for 15 semen characteristics traits of Holstein bulls. *J. Anim. Breed. Genet.* 126:269–277.
- Feugang, J. M., A. Kaya, G. P. Page, L. Chen, T. Mehta, K. Hirani, L. Nazareth, E. Topper, R. Gibbs, and E. Memili. 2009. Two-stage genome-wide association study identifies integrin beta 5 as having potential role in bull fertility. *BMC Genomics*. 10:176.
- Han, Y. B., H. L. Feng, C. K. Cheung, P. M. Lam, C. C. Wang, and C. J. Haines. 2007. Expression of a novel T-complex testis expressed 5 (Tctex5) in mouse testis, epididymis, and spermatozoa. *Mol. Reprod. Dev.* 74:1132–1140.
- Han, Y., X. X. Song, H. L. Feng, C. K. Cheung, P. M. Lam, C. C. Wang, and C. J. Haines. 2008. Mutations of t-complex testis expressed gene 5 transcripts in the testis of sterile t-haplotype mutant mouse. *Asian J Androl.* 10:219-226.
- Imumorin, I. G., E. H. Kim, Y. M. Lee, D. J. De Koning, J. A. van Arendonk, M. De Donato, J. F. Taylor, and J. J. Kim. 2011. Genome Scan for Parent-of-Origin QTL Effects on Bovine Growth and Carcass Traits. *Front. Genet.* 2:44.
- Khatib, H., R. L. Monson, W. Huang, R. Khatib, V. Schutzkus, H. Khateeb, and J. J. Parrish. 2010. Short communication: Validation of in vitro fertility genes in a Holstein bull population. *J. Dairy. Sci.* 93:2244–2249.

- Khatib, H., W. Huang, X. Wang, A. H. Tran, A. B. Bindrim, V. Schutzkus, R. L. Monson, and B. S. Yandell. 2009. Single gene and gene interaction effects on fertilization and embryonic survival rates in cattle. *J. Dairy Sci.* 92:2238–2247.
- Khatib, H., R. L. Monson, V. Schutzkus, D. M. Kohl, G. J. M. Rosa, and J. J. Rutledge. 2008. Mutations in the STAT5A gene are associated with embryonic survival and milk composition in cattle. *J. Dairy Sci.* 91:784–793.
- Kuhn, M. T., and J. L. Hutchison. 2008. Prediction of dairy bull fertility from field data: use of multiple services and identification and utilization of factors affecting bull fertility. *J. Dairy Sci.* 91:2481–2492.
- Kuhn, M. T., J. L. Hutchison, and H. D. Norman. 2008. Modeling nuisance variables for prediction of service sire fertility. *J. Dairy Sci.* 91:2823–2835.
- Magee, D. A., D. P. Berry, E. W. Berkowicz, K. M. Sikora, D. J. Howard, M. P. Mullen, R. D. Evans, C. Spillane, and D. E. MacHugh. 2011. Single nucleotide polymorphisms within the bovine DLK1-DIO3 imprinted domain are associated with economically important production traits in cattle. *J. Hered.* 102:94–101.
- Magee, D. A., K. M. Sikora, E. W. Berkowicz, D. P. Berry, D. J. Howard, M. P. Mullen, R. D. Evans, C. Spillane, and D. E. MacHugh. 2010. DNA sequence polymorphisms in a panel of eight candidate bovine imprinted genes and their association with performance traits in Irish Holstein-Friesian cattle. *BMC Genet.* 11:93.
- Memili, E., and N. L. First. 2000. Zygotic and embryonic gene expression in cow: A review of timing and mechanisms of early gene expression as compared with other species. *Zygote.* 8:87–96.
- Montenegro-Venegas, C., E. Tortosa, S. Rosso, D. Peretti, F. Bollati, M. Bisbal, I. Jausoro, J. Avila, A. Cáceres, and C. Gonzalez-Billault. 2010. MAP1B regulates axonal development by modulating Rho-GTPase Rac1 activity. *Mol. Biol. Cell.* 21:3518–3528.
- Peñagaricano, F., K. A. Weigel, and K. Khatib. 2012. Genome-wide association study identifies candidate markers for bull fertility in Holstein dairy cattle. *Anim. Genet.* (In press).
- Queiróz, D. B., A. M. Silva, G. Gutiérrez-Ospina, C. S. Porto, G. Grossman, P. Petrusz, and M. C. Avellar. 2006. Cells positive for microtubule-associated protein 1B (MAP 1B) are present along rat and human efferent ductules and epididymis. *Cell Tissue Res.* 325:125–133.
- R Development Core Team. 2009. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Santos, J. E. P., W. W. Thatcher, R. C. Chebel, R. L. A. Cerri, and K. N. Galvao. 2004. The effect of embryonic death rates in cattle on the efficacy of estrus synchronization programs. *Anim. Reprod. Sci.* 83:513–535.
- Scheet, P., and M. Stephens. 2006. A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* 78:629–644.

Shook G. E. 2006. Major advances in determining appropriate selection goals. *J. Dairy Sci.* 89:1349-1361.

Telford, N. A., A. J. Watson, and G. A. Schultz. 1990. Transition from maternal to embryonic control in early mammalian development: A comparison of several species. *Mol. Reprod. Dev.* 26:90–100.

Tortosa, E., C. Montenegro-Venegas, M. Benoist, S. Härtel, C. González-Billault, J. A. Esteban, and J. Avila. 2011. Microtubule-associated protein 1B (MAP1B) is required for dendritic spine development and synaptic maturation. *J. Biol. Chem.* 286:40638–40648.

Tymanskyj, S. R., T. M. Scales, and P. R. Gordon-Weeks. 2012. MAP1B enhances microtubule assembly rates and axon extension rates in developing neurons. *Mol. Cell Neurosci.* 49:110–119.

Vazquez, A. I., D. M. Bates, G. J. M. Rosa, D. Gianola, and K. A. Weigel. 2010. Technical Note: An R package for fitting generalized linear mixed models in animal breeding. *J. Anim. Sci.* 88:497–504.

Weigel, K. A. 2006. Prospects for improving reproductive performance through genetic selection. *Anim. Reprod. Sci.* 96: 323-330.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide selected from the group consisting of:
 - i) a nucleotide at position 3066 or position 3323 of SEQ ID NO:1, and at least 10 but not more than 200 contiguous nucleotides of SEQ ID NO:1 adjacent to position 3066 or position 3323, wherein position 3066 is adenine, and position 3323 is thymine;
 - ii) a nucleotide at position 87071 of SEQ ID NO:2, and at least 10 but not more than 200 contiguous nucleotides of SEQ ID NO:2 adjacent to position 87071, wherein position 87071 is cytosine;
 - iii) a nucleotide at position 112 of SEQ ID NO:3, and at least 10 but not more than 200 contiguous nucleotides of SEQ ID NO:3 adjacent to position 112, wherein position 112 is guanine;
 - iv) a nucleotide at position 34239 of SEQ ID NO:5, and at least 10 but not more than 200 contiguous nucleotides of SEQ ID NO:5 adjacent to position 34239, wherein position 34239 is guanine; and
 - v) a nucleotide at Position 61646 of SEQ ID NO:4, and at least 10 but not more than 200 contiguous nucleotides of SEQ ID NO:4 adjacent to position 61646, wherein position 61646 is adenine.
2. A nucleic acid molecule according to Claim 1, which comprises at least 13, at least 14, or at least 15 contiguous nucleotides of:
 - i) SEQ ID NO:1 adjacent to position 3066 or position 3323;
 - ii) SEQ ID NO:2 adjacent to position 87071;
 - iii) SEQ ID NO:3 adjacent to position 112;
 - iv) SEQ ID NO:5 adjacent to position 34239, or
 - v) SEQ ID NO:4 adjacent to position 61646.
3. A nucleic acid molecule according to Claim 1, which comprises at least 20 contiguous nucleotides of:
 - i) SEQ ID NO:1 adjacent to position 3066 or position 3323;
 - ii) SEQ ID NO:2 adjacent to position 87071;
 - iii) SEQ ID NO:3 adjacent to position 112;
 - iv) SEQ ID NO:5 adjacent to position 34239, or
 - v) SEQ ID NO:4 adjacent to position 61646.
4. An isolated nucleic acid molecule according to Claim 1, which comprises not more than 150 contiguous nucleotides of:

- i) SEQ ID NO:1 adjacent to position 3066 or position 3323;
- ii) SEQ ID NO:2 adjacent to position 87071;
- iii) SEQ ID NO:3 adjacent to position 112;
- iv) SEQ ID NO:5 adjacent to position 34239, or
- v) SEQ ID NO:4 adjacent to position 61646.

5. An isolated nucleic acid molecule according to Claim 1, which comprises not more than 100 contiguous nucleotides of:

- i) SEQ ID NO:1 adjacent to position 3066 or position 3323;
- ii) SEQ ID NO:2 adjacent to position 87071;
- iii) SEQ ID NO:3 adjacent to position 112;
- iv) SEQ ID NO:5 adjacent to position 34239, or
- v) SEQ ID NO:4 adjacent to position 61646.

6. An isolated nucleic acid molecule according to Claim 1, which comprises not more than 50 contiguous nucleotides of:

- i) SEQ ID NO:1 adjacent to position 3066 or position 3323;
- ii) SEQ ID NO:2 adjacent to position 87071;
- iii) SEQ ID NO:3 adjacent to position 112;
- iv) SEQ ID NO:5 adjacent to position 34239, or
- v) SEQ ID NO:4 adjacent to position 61646.

7. A nucleic acid molecule according to Claim 1, wherein the nucleotide is within 4 nucleotides of the center of the nucleic acid molecule.

8. A nucleic acid molecule according to Claim 7, wherein the nucleotide is at the center of the nucleic acid molecule.

9. A nucleic acid molecule according to Claim 1, wherein the nucleotide is at the 3'-end of the nucleic acid molecule.

10. An array of nucleic acid molecules comprising at least two nucleic acid molecules according to Claim 1, wherein the nucleic acid molecules are affixed on a medium.

11. A kit comprising a nucleic acid molecule of Claim 1, and a suitable container.

12. A method for detecting single nucleotide polymorphism in a bovine MAP1B gene whose partial nucleotide sequence is shown in SEQ ID NO:1 and SEQ ID NO:2, or in a bovine PPP1R11 gene whose partial nucleotide sequence is shown in SEQ ID NO:3, or a bovine DDX4 gene whose partial nucleotide sequence is shown in SEQ ID NO:5 and SEQ ID NO:4, the method comprising obtaining a sample of the respective bovine gene from a bovine animal, determining the identity of a nucleotide of the gene at a position corresponding to

i) position 3066 or position 3323 of SEQ ID NO:1;

ii) position 87071 of SEQ ID NO:2;

iii) position 112 of SEQ ID NO:3;

iv) position 34239 of SEQ ID NO:5, or

v) position 61646 of of SEQ ID NO:4,

and comparing the nucleotide identity of the position respectively to the nucleotide identity at a corresponding position of the sequences shown in the respective SEQ ID NO:.

13. A method for genotyping a bovine cell, comprising obtaining a nucleic acid sample from said cell and determining the single nucleotide polymorphism of the MAP1B, PPP1R11, or DDX4 genes according to claim 12.

14. A method according to Claim 13, wherein the bovine cell is an adult cell, an embryo cell, a sperm, an egg, a fertilized egg, or a zygote.

15. A method according to Claim 13, wherein the identity of the nucleotide is determined by sequencing the MAP1B, PPP1R11, or DDX4 gene, or a relevant fragment thereof, isolated from the cell.

16. A method according to Claim 15, wherein the gene or a relevant fragment thereof is isolated from the cell via amplification by the polymerase chain reaction (PCR) of genomic DNA of the cell, or by RT-PCR of the mRNA of the cell.

17. A method according to Claim 15, wherein both copies of the gene in the cell are genotyped.

18. A method for progeny testing of cattle, the method comprising collecting a nucleic acid sample from said progeny, and genotyping said nucleic sample according to Claim 13.

19. A method for selectively breeding cattle using a multiple ovulation and embryo transfer procedure (MOET), the method comprising superovulating a female animal, collecting eggs from said superovulated female, *in vitro* fertilizing said eggs from a suitable male animal, implanting said fertilized eggs into other females allowing for an embryo to develop, and genotyping said developing embryo according to Claim 13, and terminating pregnancy if the developing embryo does not have at least one polymorph selected from the group consisting of:

i) adenine at a position of the MAP1B gene corresponding to position 3066 of SEQ ID NO:1, or thymine at a position of the MAP1B gene corresponding to position 3323 of SEQ ID NO:6, or cytosine at a position of the MAP1B gene corresponding to position 87071 of SEQ ID NO:2, or guanine at a position of the PPP1R11 gene corresponding to position 112 of SEQ ID NO:3, or guanine at a position of the DDX4 gene corresponding to position 34239 of

SEQ ID NO:5, or adenine at a position of the DDX4 gene corresponding to position 61646 of SEQ ID NO:4.

20. The method according to Claim 19, wherein pregnancy is terminated if the developing embryo does not have at least two of the polymorphs.

21. A method for selecting a bull bovine animal as a breeder, wherein the animal is genotyped according to Claim 13, and is only selected for breeding purpose if it comprises at least one polymorph selected from the group consisting of

i) adenine at a position of the MAP1B gene corresponding to position 3066 of SEQ ID NO:1, or thymine at a position of the MAP1B gene corresponding to position 3323 of SEQ ID NO:6, or cytosine at a position of the MAP1B gene corresponding to position 87071 of SEQ ID NO:2, or guanine at a position of the PPP1R11 gene corresponding to position 112 of SEQ ID NO:3, or guanine at a position of the DDX4 gene corresponding to position 34239 of SEQ ID NO:5, or adenine at a position of the DDX4 gene corresponding to position 61646 of SEQ ID NO:4.

22. The method according to claim 21, wherein the cattle animal is selected only if it comprises at least two of the polymorphs.

1081 gagcgcgctc tgcagaggcc gaatagcggg tgcgtgggaa accctgggcg gggccgtaag
 1141 ggtccagcgg gcgcagactc gggcgcgcgg accgacccgc cccgcagccc caccgggct
 1201 gccgcagtgc ccccgcccgga tgcacctgcc cccacactg cggcgcccc actgggcgcg
 1261 gccggccggg cgtgcgttcc gccgggtgctt ggggggtgctg tgcgccctct cgtctgcctc
 1321 aacccggctt tgttgcgctc gaagtccccg ggtgggcagc tctgtcctct gcctttccct
 1381 tccccccgg catcgcagac ctccccctt cctgaccga ggtcgcgggt ccctccacaa
 1441 cccagcccg cgtctattc tctcggggtg gtgctgaagc gcgtctgcc cgagacaccg
 1501 gctggtgggc gtggtgcagt ccgcactgcg gtctctacgg cagcccgagg cggacaaagg
 1561 gcgttcacgc agccctcggt cccacgcgc gcccccact ccccatgaaa gacgcgagaa
 1621 aaccttggtt tagatgaaaa aaaattaaac tctagtgggtc tgcctctgca tttgaaaacg
 1681 gtcgcctgtg cccagaacaa aaggctgcag ggtggagact cgagttgcag acctggttct
 1741 tttgtttaac tttaaagcac tgggtgttact ttttggctta aaaaaagaaa aaaatgtgag
 1801 cagaaagcag acttggtatt ttattcggaa aaaaaaatg gaagaatagg ctaggtcaat
 1861 agtgaaatgc ctcatctgag catctaataa cctttcattt gtcaaactat agtcctttga
 1921 atttgatcag tactaattta gtttatttgc acattttctc cttggaaaat ttcacacgta
 1981 ctgactttgg gtgtggctgc tgtatgaatc tatgactttg gattttttaa aaaatattat
 2041 ttgtcagcac ttttgctggg aagtaataat aaagcagggtg tgtttctatg tataaaagg
 2101 gcataagcac ccgatgtggt gtgaggagag ggaccctcat ccccattttg gagatggggg
 2161 cccagagggg caggaaggat agggagacca gctcaagggtc acaccactca taagtgcag
 2221 aatgccggct ctgcagcatc actgggtttt ggcccgtcat gatgttagtg caggccaggc
 2281 gaacatcaca ggaagatcaa agagcaattt ctagagcttg cttttataca aaggcagtag
 2341 tcatccactt gtgggcagca ccccatgcag ggagggtggt cagcaatcac cttccatgta
 2401 attatcttcc tgctttgtta ctctgctgat caccctttta gtgccctatt tctcagggg
 2461 tttattgaag ctgcttctgt ttgaggataa actagattca cccaagttat cggtcactgt
 2521 gacgtggtct cgactatttt tcacgctact ggaacttagt gatcgagttc aattttgtct
 2581 tcttgccct tttctttttt aaaatttatt gatgtattat gtagtgagg aaaggctcaa
 2641 atcataaacg aggagcttgg tgaattttca caaaaagaca tacctgatcc cgatcaagaa
 2701 acggacattc ctgggacccc cagaaaccct ctcagggtgc cttttcagtc agcatgcctt
 2761 cctgattttt tagtgccctg atttagtcag ctttttgcgg ggaactctga aatagcagta
 2821 ttgatcccat taagaatcaa ccaagtgaat gagcaaagcc **atttcctaag gcacagtaac**

 2881 agatgctgct tttcctctgg agatcaactc tcttgggtcc tggggtcttg gatgcagctc
 2941 aaaccacagg gccttcaggt gctaagggga aatactgctc ctgtagcttt tctgacagag
 3001 atgacttagc aacagggcag gctggctgtt gccctggcct gatagatgc tctctcagct
 3061 ctcag[a/g]ttcc agcagggcct cctcctccag aggtcactct gataccatgg ctgcgggtggc
 SNP1
 3121 tgggtgccc at gggaccagtt gtgtgaagca agatggagat [g/a]ctccggggg ctccggcctg
 SNP2
 3181 cttgtgtgcc gcagcctttc cgcagggaaa gcggttcac tcttgccct cttgcttgtc
 3241 tcccctctca tccggatgag gctctctgag tctggagacc ctagggactt ggatttttgc
 SNP3
 3301 catttgatga ttttaaggctc ta[t/c]gggaaac ctagaaaata aatctgtagg **aagatggcgg**
 3361 **aaatgcaagt** acatcaatag gcctgaatca gccaccatg gaggtgagt acatgcttct
 3421 tgaaacctag tgttactgaa tcaaactgag gcagctcacc tgcagcagt aaagccaatt
 3481 tactgacccc aggttgtggt gaaggaaagt gcagcattat tgtaagatgc tgatgtaagg
 3541 agaacgggga gctcgtgctt aagaccccca aaatcccca agggtttcag caaagcattt
 3601 ttaaaggcag tgtaaggag ggtgtcccag agtgtgtgat cagctcatgc accattctct
 3661 aattgagaat tagattgatg atgaggtaac aagactacgt tgggcttct tgcgtgctta
 3721 gatggtaaa aatctgccta caatgcagga gagctcgggt tgcctcctgg gttgagatcc
 3781 cttgaagaag ggcagacaa cacattccag tattcttgc tggaaaatcc ccatggacag
 3841 agggacctgg tgggtacagt ccatggggtt gcaaagagtt ggacatgact gagtgactag
 3901 gcacagcaca agactatgtt aatgaggact atgtgcggat ggtcatcaag tagttaattt
 3961 cttccatttg atggtgattt tagcatctga aaaataactc aggaaatata cttcagatac
 4021 taatatctag gtatttcaga gaggaactaa agcagaggggt ataggggagg ggtctgtccc
 4081 aggaatgtcc cataagggtc tgcttgggtta cacaaggaca gaagaccatc attcgtttct
 4141 tcagcttcat tcaaaatgca taatgtaggg tctgggtcatt tcttaagtct gaaacaacag
 4201 tctttgctca gactggtcta gctttattta gttagttatg caggactcca tgccaaacca
 4261 ttgtcactcc ttggcacctc ctttgccccc tctctaatta gtctccaaca tagacttaac
 4321 agtcatcccc gtttttagca gagtctcccc tctccagtt tattttggga tgtagactac
 4381 atacactcat gttttgcaaa acgaattcta gagccttgtg aaaagttttc attgcttcgt
 4441 tttattttat tttttaaact aaaacatgta aaaaacacta agtgtcagtt tggtttact
 4501 tgctgttaga gtgtaacttt cctctgggtta acattgggaa tcagcaggat ttgttcatca
 4561 gaaagataga tgtgtgtaac atctactgcc ctggattttt agatgtaagt tttgctaatt
 4621 acattgtcat ttattgaata ctttcagtct acattgggtta agagaattaa gaaccagga
 4681 aagagttagg gctttgtttc cattttgttt taacaggaaa acaacggggg gatgggaatg
 4741 aaaagaaatg ttgaggaatg atatatttac tatatccatc catcttcata cgtttcaaaa
 4801 tcaaaaggca gatggatttc tgcttgccgt gcttaattgg tattatctac tcaagagtt

FIGURE 1

FIGURE 1

Partial genomic sequence of Genomic sequence of MAP1B showing the location of SNPs 1, 2 and 3 on the MAP1B gene and the locations of the primers (1F and 1R, corresponding to positions 2859-2876, and positions 3346-3362, respectively) used to amplify the region comprising the SNP sites. The number of the sequences is per that of GenBank Accession No. (514739)

49681 tccagcctgt tctgtggatg tttgaacttg agaagtgggg tctttgtcca gaggaaacac
 49741 tgcttttcgc ctggttagagg atgggctcca tccgaatcat acccagtttg ttcctttgct
 49801 acttcttcat cttcccgtagg tttcatgtcg agtcagaatg taaggactgt ttagcttttg
 49861 tgaggggcaa aaatgtgttt ttgaactgga caaggtaagg tttgaacca ttcctttgtc
 49921 tttcttgat acttccattt tcaacttgag cacaagcag gttggggaag caggaggggg
 49981 gaagatgtta ttgtggatta gagacagagg aaaaggcagg tgggggggtg gaactgaacc
 50041 ccacttcctg cagccgtttc ccagccgggt ttgaaaagac tctgaaagga gaataacgtc
 50101 tttaatcaag agcaatagta ttagctcctt tactataagt aatacttttc tttgagccta
 50161 tatttatatt accgggctag aaatagctga agttattcca gcagccatga ctattgtcta
 50221 ggagttggat gtgggctggc aatagactgg ctgattacac tgtttagaaa taaaccctt
 50281 tgttggcagt ctcttctggt gagaatggt cataaagggt cctgtggctg gttgttccat
 50341 ggtgcttgc tttatatcag ttcagtaccc tcataatgag gtggtcttct agaatatatt
 50401 attataactc tgttgagag ggctagggc tcgtcagtta tgcaggcaac atcacaaagc
 50461 tttggaagaa atctcttaat taagtgttag ggctgggtgc gcagtgaagg gatgaatgga
 50521 ctgaaatgct tccatctctg agcgtctttt caaactaaac gggccctttg ccgcatcata
 50581 gccaggagtc cagcagacgg acacactgag aaagtgggtg tggatgtgat tggatgtgt
 50641 cctgacttcc tctgacctac ccctggggat ttctgtactt cacgtcacac gtggctcctt
 50701 gttgatggta tgggtgaaaac ataggtgttg aaagaaccag aattgacagc agttcagcga
 50761 ccttttgggt cttcaggtct gagaccatac cctaggcagc atcagtcctt ccacggtaga
 50821 tggcactgga ctttctgtgg cgtttaagac ctaacgttct gtgactgaga atgtggcctg
 50881 tcttggccac agctggtacg atgacaagat gactatctga gttaggaaga aaaagtgaag
 50941 tgaaagtcac ttagttgtgt ccgactccgt gacctgcagc tcttctaggc tctctaggc
 51001 tctctgtcc acgggatttt ccaggcaaga ataattggagt ggggttggcat ttcctcctcc
 51061 agatcttccc aaccaggtc tctgctttg caggcagatt cgttatgtgt ctacaatgaa
 51121 agaaagggta ggagcaaata cagaggcaga agtttgttcc ctctaggaa gggtattctt
 51181 gatctggcca ttcaaagacc ttttcatctt ctctcagatc ttttcaaagt gactaacctg
 51241 aaaaatcttt gatgtgtggg ccaggacatg atggaggaag gcatcttttt ctttctttt
 51301 actccttga gaagagatga aactaaaagg gctctaaggg aaaaaaaaaa attcttaaaa
 51361 aaaaaaagtt aataaaaaaa caaaaggtaa taaatagctc tttgtaaaca gcttaccacc
 51421 ttacttcctg tggttacatg cattacctta **tggctcgtgat tatgaaagat** ttctagagaa
 51481 acgttaggat gatcacataa ctccctccta aggcgccagt ggagcccaaa gtctttgctg
 51541 ttcacgtgcc ttgtgagtgg cccaacacag tggggacttt ataaatatca aatcattgtc
 51601 gttaaaaaac acttcccgct tcaactctgag acccttctt ttaaggagtg catgtgggtg
 51661 [g/a]ggaggatta atgacagcac agcagagtgt gcttgaagggt ggtgacatca cccggcttga
 SNP6
 51721 acccttcagt gccgggtgag aggattttca tctcatccat cctcctgagt ttgcca[c/t]gag
 SNP4
 51781 gggctctcaa gaacaggaaa agaggagtct gaggagagga gacttctgga cattctgtga
 51841 tagtcccctg gctctgtgcc gtattgtttt gtaataaagg cagttatggg ttctagtctg
 51901 ttgtttttct acaaaaatgg aggacgtgtg **accagcagt** **ttagccttcg** tgaatgagat
 51961 tctgtgtttc ggccatcact ggttcaagta ggtaacctaa gagctgagct taagttgctt
 52021 ctcttgagc ccatctttgg ctttcagtaa ggaatctgag caacattaga ctgagaatca
 52081 gacaccttta ccatcacttt acgggatgct tccatttgct gtgtgatagg acgcaggtgc
 52141 aggagggagg cgtctggacc ccagagtcgt ggcgtcagga ggtcccgtgg gagcatctc
 52201 agcctctgca gtggtcctac caggagagga aggtgcttgg gtgtcgggat acccatgctc
 52261 aactcctggg tctcccgtaa tcatgccctg tgactctgga caagtcactt agcctctctg
 52321 aaccttaatg tttgtatctg ttaagacaag ggtttggaat agatccgtcc aaattcacat
 52381 ttctgcagac ctgggttact ggctgtaggg ctctaattgg actggatctc ctcatcctg
 52441 ataacctatt tccagggcgt ggggccccct gctgtacagc gtcttgcttc attttcccac
 52501 cgtcttttag ctcccatctg gatgccatct tctgtgtagt atggaagcct tcttaaacca
 52561 tccagtgact cccttgagcc ctctctctcc ctgctgaggt gtgtgtccag gagccagggg
 52621 ccatcctgcc ctttttgctg gccactgcc catggttctg gtctcatgcc agggctcggc
 52681 ctctgtcagg atgtgggtgg ttgagtttat acctgatctt gatgtaaaca catggcctct
 52741 gccagtcac ttgttctgt tcccacactg gcttccagct cttttgtgga ctctgactct
 52801 ctgctctcct ggctcctga tggctggaca tcttttctt tcttctaga atgccacccc
 52861 tttttttgtt gtctaacttg taaaagcccc atagatcatc tcccatttca aaccttaga
 52921 gatgactatc ttgatattgt gataagaggt gaactttctc agaggagttt cttgttacag
 52981 tgtcaaatgt ggttataaat cactggaact taaggatctg tctgccaagc agtagacag
 53041 agtttgatcc ctgggtcagg aagatccgct gaagaaggga atggctatcc atgccagtat
 53101 tcttgctgg agaatcccat ggacagagga acctggaggc tacagtctat ggtgtcacca
 53161 aaaaattgga catgacttag tgactaaaca acaacaataa gagaagcttt aaggaggtc
 53221 agccctctct ccaccccagc actagaacag tccctagagc agagtctccc aaatctgcct
 53281 ggtgagcaga atcatcacc agtgctttta ttattttt aatatttatt aaaaaattt
 53341 tttgttttg ctatattagg tcttagttgt ggcagaggg atcttttagt gcaacatgtg

FIGURE 2

FIGURE 2

Partial genomic sequence of Genomic sequence of MAP1B showing the location of SNPs 4 and 6 on the MAP1B gene and the locations of the primers 2F and 2R (corresponding to positions 51447-51466, and positions 51931-51948, respectively used to amplify the region comprising the SNP sites. The number of the sequences is per that of GenBank Accession No. (514739)

85081 agtgctctgg ttaatcaacc agtgaggtca gctgaccact gggtagccag tacagattga
 85141 gaaaagagca tccaagactc tactttattcc attctgacca cactgcctca ctaatacaga
 85201 ctcaacatct tgttttaggt cgaaattcag cttggcaaga gcaagcatcc ctgacttcat
 85261 gttcttagac ttatcaagtt ctcggcagtg atctgggttt ttgatgaatc tgggtttctg
 85321 atgatccagc ttcattttgt tgcttcaaaa caatcacagg gatttgaatt catatatatt
 85381 atttgcttta catagttact tgaaggtttt agatcacagt tcacaaacat gtaaagcaaa
 85441 aaataagcaa cactttcttg atttattatg gaaaaattca gtatttagta ctttaggaag
 85501 tactagttac aggtacaagt ttttactttt aggcaacgtg aagcagcaat ttcaagactc
 85561 atatcagatt tcctcttttt atttgacat agaaaacaaa ctgaatttgt tcatgcttag
 85621 aatttgtata gagccaccag ataataaatc ttgatctaaa ggacttaaca gtgaccatgc
 85681 acttaggaga aaacatgaaa tcaattcaaa cagataaaaa cccaactgaa atttgctgcc
 85741 aaactcatga aacttacact atagcccaca caattgattt tatcactttt ttttttttgg
 85801 gtcatttaaa gataattttg agggaaagtg agttaatttg aatttacatt gaggatgctt
 85861 tccaacaga tttttttaaa agacaaaact gcactattaa ttaattttta aacaaacctg
 85921 ggtcaacttc cagtggttct atcagtttgg gcttctatga gcaatgttca tttggtgtca
 85981 acgggagtga ttcaaggtgc aagtggaaac tgcaggcatt taaaaatatt agatgatctg
 86041 taactcacia acctctgcta caagtcagaa ttctttggga gatttacaca tgaatatgtt
 86101 taggactttt agcttaggtt cattataatg gctggttaat ctattcatga aatgattcag
 86161 tttatccaaa taccagtttg gctgattcct actacccctt gccctccaaa ataaaaataa
 86221 accagttcat agctgatttt gactgtggga tggcagctct tatacatccc atggagaaag
 86281 gcaagagaat taaatttagg ggatcttgct agtattttta gtggtttcac agcagtggtc
 86341 tcaaaccaga tacacattag cattggctgg gatgctttta aaaagtgatg gtacctggt
 86401 cagtgaagcc ttaccatagc cattgaagcc agggcatctg tattaagcat gctaagtgat
 86461 tctaactcat tggccaggag gaagaaccac tgccttacia tgctagtctt gttaatgttt
 86521 caaccttctg attagaacia atcagaaagc caattctaga aacaaggtag ccagaaactg
 86581 agattaatct gaaccttcat tttgccagg ctttctgact tggggggaat tttggctgtg
 86641 **acataacctac** cccttacctc agtccggtat gttctgattg gctagagaaa gcagagtctt
 86701 tctgaacctt cctgttgcta aagtttggtt tctagtcttg tctaaggaga gacgtctacc
 86761 atttagagga ctgtcctaag gagagaatac agtgttttca tcagtttatg catgaggctg
 86821 aggtgctgag ggtcttgag atcatatgac attaatgatc gactactggc tagatcaaat
 86881 gtgaggggat aatattcagc tgtgggcca actgctttta aatgaaatcc taacatgaat
 86941 tactaagatg gcttaactat gctttacca atgcagatgc tttcctttgt cctttaaaat
 87001 ctatttctta gatcacattt caaattaaaa gacacactag cagctctttt aggagtgtta
 87061 gcgtctagtt [c/t]tatctttgg ggaaagcctt ggcaactctt cttaattgct aatgtgttta
 SNPs
 87121 agggaaacgc ccattcttct atttctctct agatggtaaa cagtcaagtg atgctgtctc
 87181 agactgccag tgtcaaatgc cctctgtgag agaggggagt gccaacaccc actcccatgt
 87241 ccagagcgc cttctgggga **ataagtaggg aaggtctg**ct ggacagatga gtctctttgc
 87301 atttttgtga ccctggcctt ctctttgttt ttatttgttt acaaagggcc aggaaccacc
 87361 aagacgtcca agccctcagc tgtgccccca ggccccctg tgtacctgga cctatgctat
 87421 attcccaacc atagcaatag taagaatgtc gacgttgaat ttttcaagag agtgagatcg
 87481 tctactacg tggtagtggt gaacgacct gctgctgagg agcccagccg ggtgtcttg
 87541 gatgccttgc tggagggaag agcccagtg ggagcaaca tgcaggtaag agttccagga
 87601 cgggtgttgc acaacacgtg gagctgtgtc cagaggcagc aggaagggat cgtgtttaat
 87661 gaggcaccac cgtggatccc catgaggtgc ccacagggcc tgcctgactt ggacaaagt
 87721 gatttcacac acacaagctg gtctaaaagc attcgcgcca ccagccacca tggacttga
 87781 ggaaggccac tttaccacce taaagtataa tctgcagagt gggcccaaga ttacacaccg
 87841 ttcataatcc aagaaaatta accagcgtaa ccaagtgtca tatttccatg tgagatggat
 87901 aaagattagc ctttacttgt ctttcccaag tagacaaaag ctagagatat ggccatttag
 87961 aaaatcagct gtccacatga gattctgcag gagcactgct gaaaatggct ctcagcagga
 88021 cactcccaac acccaaacat cgtaatgagc cacaaccac tcattatttc agttatggat
 88081 tttatctaag ttttacttac ggttttgtat agtgatctag taaactgtat ttgcataacg
 88141 ttaaatagaa atcctggtta tttcattata tgaaatctaa tgcactcagt ggccctttac
 88201 tgaatactag gtagaattta agctagtaat cacttaccac cccactcct ctgtcccaa
 88261 acacacacac aaagacataa atctttgctc tcatgatgaa atgttagtta acatgcaatt
 88321 agaagggttt cggctgcatt aataactaaa gccctttgt tttaaatatg caatatcttt
 88381 aatgtaaaac atcagttgtg ttaaagaaaa tacaagaaat tccaccttaa ctgaagaact
 88441 tctcataatg ctaaagaatt gaaaactgat atagatgaac taactggcta gtcactgatt
 88501 gcttttgggt ctagtcttca actgccccag aaaaactaat tttttagcag ctttattctg
 88561 gttcctagaa aatgtaagtt ggaaagtcct atggattttc taaggacaat agaataattt
 88621 tctctttccc tttctttttc taatggtcta attaatacct tactgctgtt ctatttttcc
 88681 ccaccccat tctggttctg ctcttcagta gctgttttct ctctccctgc aggtgacct
 88741 gatcccgact catgactcag aagtgatgag ggaatggtag caggagacc atgagaaaca
 88801 gcaggacctc aacatcatgg ttttagcaag cagcagcaca gtggttatgc aagatgaatc
 88861 cttccctgca tgcaagatcg aactgtaaca accaaggta gccgcaccac aggatttgaa
 88921 ctttgtttcc agaaattcct cgatttgaaa ccaccttttc taaaaaaaaa gtcaattcat

FIGURE 3.

FIGURE 3

Partial genomic sequence of Genomic sequence of MAP1B showing the location of SNP 5 on the MAP1B gene and the locations of the primers 3F and 3R (corresponding to positions 86634-86651, and positions 87260-87278, respectively) used to amplify the region comprising the SNP sites. The number of the sequences is per that of GenBank Accession No. (514739), incorporated herein by reference in its entirety.

-120 gcggcgcctg cgcactgtca **cattacggcg gaactaatcc** ggcgacccag cgctttgacg
-60 catttagtac caggaagga aaagggggac cacagaacgc gtcacaccg gaagtagga
1 gccggaactg gggttggaca ggttatcccc aggggtggg cagcggaggc ccaggaggag
61 ggggaaaaaa gaaggtggag gatcctggct gctaactga atcgataccg a[t/g]tctcttag
SNP1
121 acctcagaga cacagaaaag acagaagggt gcctcatccc ctttcctccg cttctctctc
181 tcctcagcct tagccatggc ggaggcaggg gccgggctga gtgagaccgt cactgagaca
241 acggttaccg tgacaacgga gcccgtaga aaggctggg gcggtgctgt ttaggggtct
301 gagagatacc gggagggaag ggataaggct ttggagagtt gctggatggg ctgggcctgg
361 ggatatggga ggaagtgggt ttgggagaat cgcagagtat tagggatttt ttggtgtgtc
421 agagttgggt cagaaggctg gtcaagtga atgcaataga gttaagatgt **aggtgatact**
481 gcttgggatg gtggtgtctg taagtattga aagactggga acttggcgat taatgagcaa
541 gggatgtact gggggaaatg aagggttggt tgagaaagca tgggttgaag ctgcctgtag
601 ggaaacttga cactaagcat gcttatcaat aaatatttct tgaagagatt attgcaaacg
661 gaagcagagg gaatgaggga acaagaaaag ggagatgatg ggagtatttt gaaaaatcag
721 agatgtagag aaaaacagcg tttttgcaa aacattgctt tcaataggag atgttctctg
781 cgggcttaat aaccctttga ttaaggaggt ttagagtaat agttactaga gatgccagga
841 tgctggagaa taggtggata acagattggg agggctgggc ttgaggatga gagatgtgag
901 aacagagtca tttctttaat gggaaaaaga ataggcgttc tgggaaaaga aaggagatc
961 aaagtttagg cattggtgac tgaaaaaata atttcatgt attaatacca ccaaagatga
1021 tttggggagg aagatggagg aacagcgagg attatatatt cctttgaaga tttgctggga
1081 ctttccttag gtttaggaatt gtatcttctc tgtatactag tggttactaa gaatactaag
1141 aacagaattc ctcaaggagc tccttgaggt caaaa**acctg ttctatctcc tcccagcatc**
1201 agctcctctg t[a/t]gctgtgtt tgtgatcctg attgaactgg gaaaggaag aaaggaggcc
SNP4
1261 ccaggaggga cgcaggaaga gttagtagga ggggactagc taggtatgcc tatecttctt
1321 aaccttccag gagaaccgga gcctaaccat caaacttcg aaacggaagc cagagaaaaa
1381 ggtggaatgg acgagtga caatgggaca tgaacacatg ggccgccgct catcaaatg
1441 tgagtaattg ttgccccaca gtaacgctgg agtccctggc cccct[a/c]agca tatcttttgc
SNP2
1501 cttcaggcat tcaactggcct tcccaaagcc cccagatg[c/t]t cacagtccctg tggctgcctt
SNP3
1561 ggtggttctc tggtatcagg gagaggaggt taaagttaga gggaaagagg tagggagggg
1621 cttcaatttc catgtgcaag gcctaaagtc aaaggtatct gaggtgggag aagaggagct
1681 ttggattccc ggctggaaag **gcaaggtggg taggtgacag** agtcccagag ttaggcctg
1741 gggaagctgg atctggaagg tagaaggaga aaatgggtgg aagtaggaat tttgactgag
1801 atccagtggg aatggaactg acactacatc tgaactctc ctctttttc actgggctcc
1861 tccatccaaa tccaggctgc tgtatttatg agaaacctcg ggcttttggc gagagctcca
1921 cagagagtga tgacgaggaa gaggagggtc gtggtcacac acactgtgtg cggggccacc
1981 gcaaaggacg gcgtcatgca accccgggac caagccccac cagccctccc cagcctcctg
2041 accctccca gccctccca gggccaatgc agcactaaat tctcgtctcc cccaccattc
2101 ctgtgtctgt ctggccctga atgtattcat gtggctactc ggggactaaa cccacgattt
2161 gatcccttct ccagccctct cctccctctc cctctgctg acagaggga gagggagagg
2221 aaggtggaca gagatcctgg aattctgact tgctgctatt ccagaacctt ggcttctggg
2281 ttccccagc cctcatttct ccttacaata cccagcctcc tctctccagg gatccaggca
2341 tcttgatccc aatcttttct ctttgttctc actgccaac tgctgtcct gggatccagt
2401 tatcttggcc ccttgactc tctacttgag ttccaaacag cttaattggg tttccagcag
2461 cccagcttt cactgccagg gtcctagtca gattccaggc aatcttgctc cagctatgct
2521 tggtaatcct ggcttagagc tcttccactt atgtatttat gtcaccta ctcttagtgc
2581 ttgcctgtgg gatgtgaggt cttctgtgag acctcagggc tctagccct ttccttctc
2641 tcctgccac ttcccccaag cccttaagag gagttaggag agaggaggt ctttgtcctt
2701 ctcaccttta atgagaaatg gaaaaaagaa atgggcatgt cctctctct caccgttctc
2761 atgtgactag ggtttctgac aaaactggct ccaagactag tcaactagag cccactatct
2821 cctcagcctt tgggtcttcca acttaggaga cagatccgac ccaggggcct gggctcctgg
2881 gagaggatgg aaaaggagg gagccaagag atgcaatctc accccttct tccaaggcct

FIGURE 4

FIGURE 4

Partial genomic sequence of Genomic sequence of PPP1R11 gene showing the location of SNPs 1-4 on the PPPiR11 gene and the locations of the primers 1F and 1R (corresponding to positions -102 to -85, and positions 470-489, respectively) used to amplify the region comprising the SNP sites. The number of the sequences is per that of GenBank Accession No. (504846), incorporated herein by reference in its entirety.

32101 gagtgaggcc ctggatcctg tattttaaag cctccctgct gactctgatg cctctttcat
 32161 ttggttaacca ctgatctatg gagtctttat aacttctctg cacacagatg tatagtgtat
 32221 tttgtggtta tttttcccgat atcttgtgtc ccttagctgg tccagtatag taataaggag
 32281 ctcagattttt ggaatcagac acccgggatt tgagaccag atccttctct tctgtcagt
 32341 tgtattattt taaaaaagca tgtttaactt tgtatttctt tttttctgtt tataaattga
 32401 ggataatacc cacctcataa ggtttttgtg aggattttaa aagttaaaat agaattcatt
 32461 tagaagagtg tcagatggat actgttttat gtgttattat tacattattt ttctataatt
 32521 agtagattat aactgatctt gggatcatta tctcattttt gtttgtgctt aactttattt
 32581 ttaactctac agagattatt ttttaataac tttttatttt gaaatcattt gactcaagaa
 32641 gtttcaaaaa tagtacagaa gatttccttc agcttgcttc taatgtaatt gtactcctca
 32701 cccagtttct cctaattgtc ttagcctatt ttaattccca gtgtgggtcaa aataactgta
 32761 agtttggttg aggagaacag ttggccaaag gttatgtgag gtgggttttc tttcatatat
 32821 gacaataaat gtttagctacc atcatcatta ttccagatga tgatgaatta ttatatcatc
 32881 aaataataat tccattattc cagattattc taagaatctt tgcagaattc tttctttcct
 32941 cttcaacccc cttgtatgaa atctttgctt ctgagaagggt tgtcttgatg ttaaattgatt
 33001 ctttaggaat attgtcaatt gttgatgtca gctcaaatag gagcctgcaa caagagctgt
 33061 gggtcattgt ttattataaa tcaatattaa ttgagtagat tagtactttt gtacataaac
 33121 aactgatagc ttttaactctg tcgagccaca tatgtcatca ctgggacctt gttctctggc
 33181 actaaatgtt agtgatatgt acaaagattc cttgaaacca tcttggtatt ttccaaatat
 33241 gggtttattg gaatcttcta gaaagcttaa agttattact gaaagttata tcaataagggt
 33301 ataatttttt aatttagaaa aattgttcat tcttggtat cactctgcac attcaaaatg
 33361 aatctctcta ggggtgggttc tgataattta tttttaacca actttcctga taatttttag
 33421 tcatccttca gtgtagtgtat cctcatacta ctatgtataa actctggtaa tgttacttaa
 33481 taatgtcact aaggacagaa agccgggatg tccccaaatg cttccattag gatggatagg
 33541 gaaaagtttg catatattaa aaaaacacta taatgcctca ggtttattaa gaaagacaat
 33601 ttacagatta atgatgacat tataaataca atagttatgc atttctgaga tccgtttgac
 33661 tactcaactg ttcagatatt tctgaaactg tttcgtgaca tttatgaaat tcttattttt
 33721 tggctgtgct cagaacttga cagataacat gcttaacatt tagtatttag gtatattagg
 33781 tgatttttaa aaagaattga ctgaataatg tgtttgtatt ttgttgttat ggtgatttta
 33841 aaatttataa ttttggttcat atgttagctt atgaatatat tttttctcag taatttctct
 33901 tgtgatagta ctatttagat actacagtaa tataaatact gcaataatta tttagatgct
 33961 attcacatgt taaattttta ttcaagaatc tagtattgac tgtgaagata atcaaacacg
34021 gaacagaggg ttttccaaga gaggcggtaa ggaccatgtt ttggaacaac ttgtacttaa
 34081 gacagaaatt aaactgaaaa attgattttg gaagagctga aagaaaaatt ctgggtgggtga
 34141 aaacctttca agaaaaatac tttggcatat cttttatgct gttaatattt gagttaatat
 34201 tcagtaggtg tctctccttc tgctttctga tgtcactc[*g/a*]a tttgtctttt cctaagacct
 SNP2
 34261 ccagagtgtt ctatgaacta caaaagggtg gactgtgtga atcttgggtca ttcacagtat
 34321 agataaactg ggatgtcttt gtctctgagt aggaacattg gagatatggg ggaagggaga
 34381 agttgtagat taattaccat **acttgcta** **cctgcct**ctg cttgaggtga gatggtataa
 34441 aaattatagt gctcagttct ggattatcta taggcagaca tgttaaaata gcaacaatat
 34501 ccacgaaaaa ccacagtga cttataaaat tgctacaagt gtgcaaatat atttatgata
 34561 gaactttagt gtttgagct gcactagata catcatagtt tttgctgcaa cttggagata
 34621 tcgttttccc ttgcctatta gatgattggc tcatgataa gatcattgaa tagcaggcct
 34681 tctagtga gctgagactt gctgtggatt tcaactatagc cttggatgag ttgtgagggg
 34741 cgggtgggtg gaatttgggtg gtgaatcagt tcaactatagc agttgtgtct gactctttgc
 34801 gaccccatga attgcagcat gccaggcctc cctgtccatc accgactcct ggagttcatt
 34861 caaactcaag tccatcgagt cgggtgatgc atccaacct ctcactcctt gttgtccctt
 34921 tctcctcctg cccccaatcc ctcccagcat cacagtcttt tccaatgagt cagctctteg
 34981 catgaggtgg ccaaagtact ggagtttcag ctttagcatc attccttcca aagaacaccc
 35041 aggactgatc tcttttagaa tggactagtt ggatctcctt gcagtccaag ggactctcaa
 35101 gagtcttttc caacaccaca gttcaaaagc atcaattctt cggcgctcag ctttcttcac
 35161 agtccaactc tcacatccat atatgaccac tggaaaaacc atagccttga ctatatggac
 35221 ctttggttggc aaagtaattg ctctgctttt cagtatgcta tctaggttgg tcataacttt
 35281 ccttccaagg agtaagcgtc ttttaatttc acagctgcag tccacctctg cagtattttt
 35341 ggagcccaga aaaataaagt ctgccactgt ttccactgtt tccccatcta tttcccatga
 35401 agtgatggga ccagatgcca tgatctttgt tttctgaatg ttgagcttta agccaacttt
 35461 ttcactctcc tctttcactt tcatcaagag gcttttttagt tctctttcac tttctgcata
 35521 agctgagtct ttaattggcaa ttgtaggggg ccttgcaatg atggggcgga catttagtta
 35581 agaaatagac tcgtctttta acattgctct ccttccccct ttaacaagga gttttgacac
 35641 taatgttctt aaaacatagc tcttttgggt tttctgcagaa cagtggctat ctttcttact
 35701 attcagtttt ctttaaatca ttttaattca tatttaagtg cagcaatgaa aagccagttg
 35761 cagctctttg tgcttgatcc tgacttattg actagttagt gttgtctagt aggggtgtccc
 35821 tgattgctat tttctttaga tgtatacatt tgaaggtaga aaacttgtgt gtgacacggt
 35881 ggtcactcag taaatacaga tgtgtgtgta aatagaacct tatctaagtt tatgttgaag
 35941 tatcatgtca tacaggaaca ggttgggtgca tatgtcataa atgtatacag ctcaagtatt

FIGURE 5

Figure 5 shows the partial genomic sequence of DDX4 showing the location of SNP 2 on the DDX4 gene and the locations of the primers 1F and 1R (corresponding to positions 34014-34031, and positions 34398-34417, respectively) used to amplify the region comprising the SNP site. The number of the sequences is per that of GenBank Accession No. (493725), incorporated herein by reference in its entirety.

59641 ctgatatctg agattaaata tcacctctaa aaccttcctt ttttgcagtc ctactatctt
 59701 acttaggaaa ttccattcag tatcttgccg ggagaattcc agggacagag gtgcctgggtg
 59761 ggctatagtc catgggggtca caaagaatca atcggacatg actgagcgac tgacactcat
 59821 tccagttgtt cagacacaaa aatatgagge atctttgatt tctttctcca gtgtacatat
 59881 ctgactgttc agcagagtc tcttggttct gtctagcatc aaactatata agaaccatc
 59941 agttgctaag ttaaaacaaa gatcagatct acttcttgcc tccaaaccca atggcttccc
 60001 tcctcaaagt aaagatcttt cagtggattt caaggcactt catgaaatgg ctcccatcat
 60061 cctctccaa atcctgctgc tctccattgc ttacctatt ctactgacac tttgctcatt
 60121 actcatcccc acacgtacca gacatgggtc tacctcattg cctttgttct tgatattccc
 60181 tctgcctaga acgtaaaaac cagatgggga cctcttttct gggttggttg tttgttggtg
 60241 agcacgtaaa gaacacctat acctagaaac ctttataatt gtcttacct ctgacctata
 60301 atattgtttc ttcagatata tgaatagttc acttggttag taatatctgt cctcagatgt
 60361 catctaatta gagaagggtc ataaacccat ttggcccttt attctgctaa attttttct
 60421 tttttagact taatgcatca tatttggtat ttacctctcc acaatgagtg cagatgcttt
 60481 gttttatttg ctgtgttttc agtatctaga atagatactg gcactactg ggtgctaaat
 60541 gtttggtaat aaatggagac agctgatagg gacatggagg agtgggatta caaaaagtaa
 60601 caccctgtta cttctattca ggctatagat tattatcttt cctctctaac attttataga
 60661 aaatttcttg gtgtggaaga gtctttggct catgcttata tttaaaccaa ttaacaaagc
 60721 taacaagtaa atattcttta atgtgttaag accctgagat aatatcttat atatctatat
 60781 atataagata ctttggggat taaggggtt gaattgagat aatgttgaaa catagaatat
 60841 tgtagagttt gtcagccctg ggttgaatgc tattattatg aatgcctctg gcagtcaagg
 60901 taaaataata gattctataa ccatgggaag aacaaaggaa tttatttata gagagggaaa
 60961 tgaagggttg gagccaaatt tggaactgct aagagttggc aggtagatat catttatctt
 61021 tccctaagtc tctacttatt ggttggccta tgaaggcaag gggcaagcgg tcgttttctt
 61081 ttttttttag tcataaagta ggaaaaagtc cttgtctctc tatctttgac atttatccac
 61141 tcaggaaata tttgaattgc tcttacctat caggcctgag tttagggtgct gggaatacag
 61201 tgggtgaataa gacagtttcc tactcaccct atgctttttt ttcttttctt aattttgtag
 61261 tttgaatggt tcattgtctt gtaggtgatg aaagaactat ggtctttgtt gaaactaaga
 61321 aaaaagcaga ttttattgcc acttttcttt gtcaagaaaa aatatcaaca acaagtattc
 61381 atgggtgagt agattattat gatttcttag taaggaggta acttctattt gtcatttgtt
 61441 aagaaatggt ggtatattta actagtaaaa aatcctggaa ttaggatctc aagatctagc
 61501 tattattctt catctttaag tttttataag aaccaagtgg ctgggatgga tgagttagaa
 61561 gtatcttaag ttctcatca tgattatcaa gcagtctcac ttagtctctc taatccttag
 61621 atctaggtgc ttattcagct aatctt/a/gtcca gtgcttggtt tttttttcac tgggcctcaa
 SNF1
 61681 aaatgtcata gtaacacaac ttcattggca ctttactagg agatctaaaa tattaattgg
 61741 tgaatatgta gaattccgag attatacttt taaaaaatca ggaatttttg agaaaggatt
 61801 tgatcaacta gttgtgtatt ttttgtcaaa actagaaaca gtttattagt tggttaagatt
 61861 tgatgtttgt cgcatttgag tctgtatttt ggotgtagggt attagctgtc accttcatct
 61921 gcaaatggag ttaattgtcc taacctcaaa gtattattgt gagaattaaa tgatgactta
 61981 caagaaatta gtataatact ttgtatgtaa gaagtagaca ttaatgttaa ctgttgctat
 62041 ttttggttta ctgttatcca ttgaacttat ccaaaaagaa atctatccat tgaactttaa
 62101 tgggtgtctt taatagtttt cttaagggct attctaatac atgctttgtt ttctttcaag
 62161 tgatcgtgaa cagagagaaa gagagcaagc tcttgagat ttccgctgtg gaaagtgcc
 62221 tgttcttggt gctacttcag tagctgcccg agggctggat attgaaaatg ttcagcatgt
 62281 tattaatttt gatcttctt ctaccattga tgaatatgtt catcgaattg ggcgtactgg
 62341 tcgttggtga aatactggca gagctatttc cttttttgat ctggaatcag atagccagtt
 62401 agcacagcct ctagtgaag tgctatcaga tgtaagtttt taatttttaa aactgaatgg
 62461 atagtgttct tacctgttca ttgaaagcag acattttata tgtatggatt ttcagttcag
 62521 ttcatttgct cagtctgtc tgactcttg caaccccatg gactgcagta tgccaggctt
 62581 cctgtccat caccaactcc cagagcttac tccaactcat gtccatcgag ttggtgatgc
 62641 catccaaccg tctcatcttc tgteatcccc ttctctctct gccttcaatc tttgccagca
 62701 tcagggtctt ttccaatgag tcagttcttt gcactcaggt gccagagtat tggagcttca
 62761 gcttcagcat cagtccttgc gatgaatatt caggactgat ttcttttagg gttgactggg
 62821 ttgatctcct ttagtccag gagactcaag agtcttctcc aacatcacag ttcaaaagca
 62881 tcaattcttc agcactcagt tttctttgaa gtccacttct cacatccata catgactact
 62941 gaagaaacca tagctttgac cagacagacc tttggtggca aaataatgtc tctgggtttt
 63001 aatatgctgt ctagggttgt catagctttt ctgcttttct tccaaggagc aagcatcttt
 63061 taatttcatg gctgtagtca ccactcttag tgatttcgga gcccccaaaa ataaagtctc
 63121 tcaactgttc cattgtttct ccactctgtt gccctgaatt gatgggaccg gatgccatga
 63181 tcttagtttt ctgaatgttg agctttaagc caacttttct actctcttct ttcactttca
 63241 tcaagaggct ctttagttct tcgctttctg ccataagggt ggtgtcgtct gcgtatctca
 63301 ggttattgat atttctccag gcagttctga ttctaccttg tgcttcatcc agcccagcat
 63361 ttcacatgat gtactctaaa atcccacggg cggaggagcc tggtaggctg cagtccatgg
 63421 ggtcgctaag agtcggacac gactgagtga cttcactttc acttttcaat ttcattgcatt
 63481 tgaggaggaa atggcaaccc actccagtgt tcttgcttg agaatcccag ggacggggga
 63541 gcctggtggg ctgcccgtca tgggggttgc cagagtcgga cagactcaa gcgacttagc
 63601 agcagcagca tataaattaa ataagcagg tgacaatata cggccttgac atactcctt
 63661 cctattttgg aaccagtctg ttgttccatg tccagttcta actggttgctt cttgatctgc

FIGURE 6

FIGURE 6 shows the partial genomic sequence of DDX4 showing the location of SNP 1 on the DDX4 gene and the locations of the primers 2F and 2R (corresponding to positions 61531-61549 , and positions 61867-61884, respectively) used to amplify the region comprising the SNP site. The number of the sequences is per that of GenBank Accession No. (493725), incorporated herein by reference in its entirety.

1081 gaggcgctc tgcagagggc gaatacggt tgcgtgggaa accctggggg gggccgtaag
 1141 ggtccagcgg gcgcagactc gggcgcgcg accgaccgc cccgcagccc caccggggt
 1201 gccgcagtgc ccccgcccg tgcacctgc cccacactg cggcgccccc actggggcgg
 1261 gccggcgggg cgtgcgttcc gccgggtgctt ggggggtgctg tgcgcctct cgtctgcctc
 1321 aaccgggctt tgttgcgctc gaagtcctcg ggtgggcagc tctgtcctct gcctttccct
 1381 tccccccgg catcgagagc ctcctctct cctgacoga ggtcgcggtt cctccacaa
 1441 cccagcccg cgtctatctc totcggggtg gtgctgaagc gcgtctgccc cgagacaccg
 1501 gctgggtggg gtggtgcagt ccgcaactgc gtctctacgg cagcccgagg cggacaaagg
 1561 gegtccacgc agccctcgtt cccacgcgc gccccccact ccccatgaaa gacgcgagaa
 1621 aaccttgttt tagatgaaaa aaaattaaac tctagtggtc tgcctctgca tttgaaaacg
 1681 gtcgctgtg occagaacaa aaggctgcag ggtggagact cgagttgcag acctgggtct
 1741 tttgtttaac tttaaagcac tgggtgttact ttttggctta aaaaaagaaa aaaatgtgag
 1801 cagaaagcag acctgttatt ttattcggaa aaaaaaatg gaagaatagg ctagggtcaat
 1861 agtgaatatc ctcatttgag catctaataa ccttcattt gtcaactat agtcctttga
 1921 atttgatcag tactaattta gtttatttgc acattttctc cttggaaant ttcacacgta
 1981 ctgactttgg gtgtggctgc tgtatgaatc tatgactttg gattttttaa aaaatattat
 2041 ttgtcagcac ttttgcgtgg aagtaataat aaagcaggtg tgtttctatg tataaaagg
 2101 gcataagcac ccgatgtggt gtgaggagag ggacctcat cccattttg gagatggggg
 2161 cccagagggg caggaaggat agggagacca gctcaaggte acaccactca taagtgcag
 2221 aatgccggct ctgcagcctc actgggtttt ggcccgctat gatgttagtg caggccaggg
 2281 gaacatcaca ggaagatcaa agagcaattt ctgagcttg cctttataca aaggcagtag
 2341 tcatccactt gtgggcagca cccatgcag ggaggtggct cagcaatcac cttccatgta
 2401 attatcttcc tgccttgtta ctctgctgat caccctttta gtgcctatt tctcagggg
 2461 tttattgaag ctgcttctgt ttgaggataa actagattca cccaagttat cggctactgt
 2521 gacgtggtct cgaactattt tcacgctact ggaacttagt gatcgagttc aattttgtct
 2581 tcttgccctt tttctttttt aaaatttatt gatgtattat gtagtgggg aaaggctcaa
 2641 atcataaacg agggacttgg tgaattttca caaaaagaca tacctgatcc cgaatcagaa
 2701 acggacattc ctgggacccc cagaaacctt ctcaggtgtc ctttccagtc agcatgcctt
 2761 cctgattttt tagtgccctg atttagtcag ctttttgagg ggaactctga aatagcagta
 2821 ttgatcccat taagaatcaa ccaagtgaat gagcaaaagg atttcttaag gcacagtaac

 2881 agatgctgct tttctctggy agatcaactc tottgggtcc tggggctctg gatgcagctc
 2941 aaaccacagg gctttccagt gctaagggga aatactgctc ctgtagcttt tctgacagag
 3001 atgaacttagc aacagggcag gctggctgtt gccctggcct gatatgatgc ttcctcagct
 3061 ctgagta/ttcc agcagggcct cctcctccag aggtcactct gataccatgg ctgcggtggc
 3121 tgggtgccat gggaccagtt gtgtgaagca agatggagat [g/a]ctccggggg ctcgccctg
 3181 cttgtgtgoc gcagcctttc cgcagggaaa gcggttcact ttctgcccct cttgcttgtc
 3241 tccctctca tccggatgag gctctctgag totggagacc ctagggaact ggatttttgc
 3301 catttgatga ttttaaggctc tal[pl]gggaaac ctagaaaata aatctgtagg aagatggcgg
 3361 aatgcaagt acatcaatag gctgaatca gccaccatg gaggtgagt acatgcttct
 3421 tgaacactag tgttactgaa tcaaaactgag gcagctcacc tgcattgcagt aaagccaatt
 3481 tactgacccc aggttgtggt gaaggaaagt gcagcattat tgttaagatgc tgaatgaagg
 3541 agaacgggga gctcgtgctt aagaccccca aaatccccc aagggtttcag caaagcattt
 3601 ttaaaggcag tgttaaggag ggtgtcccag agtgtgtgat cagctcatgc accattctct
 3661 aattgagaat tagattgatg atgaggtaac aagactacgt tgggcttctt tgcgtgctta
 3721 gatggtaag aatctgccta caatgcagga gagctcgggt tcatccctgg gttgagatcc
 3781 cttgaagaag ggcattgaca cacattccag tattcttggc tggaaaatcc ccatggacag
 3841 aggagcctgg tgggtacagt coatgggggt gcaaagagtt ggacatgact gagtgaactg
 3901 gcacagcaca agactatgtt aatgaggact atgtgcggat ggtcatcaag tagttaattt
 3961 cttccatttg atggtgattt tagcatctga aaataaactc aggaatatata cttcagatac
 4021 taatatctag gtatttcaga gaggaactaa agcagaggggt ataggggagg ggtctgtccc
 4081 aggaatgtcc cataaggtec tgcctgggta cacaaggaca gaagaccatc attcgtttct
 4141 tcagcttcat tcaaaatgca taatgtaggg totggtoatt tettaagtct gaaacaacag
 4201 tctttgctca gactggctca gctttattta gttagtattg caggactcca tgccaaacca
 4261 ttgtcactcc ttggcactcc ctttgcctcc tctctaatta gtctcaaca tagacttaac
 4321 agtcactccc gtttttagca gactctccc tctccagtt tattttggga tgtagactac
 4381 atacactcat gttttgcaaa acgaattcta gacettgtg aaaagttttc attgcttctg
 4441 tttattttat tttttaaaact aaaacatgta aaaaacacta agtgtcagtt tggtttctact
 4501 tgotgttaga gtgttaacttt cctctgggta acattgggaa tcagcaggat ttgttcatca
 4561 gaagataga tgtgtgtaac atctactgoc ctggattttt agatgtaagt tttgctaatt
 4621 acattgtcat ttattgaata ctttcagtct acattgggta agagaattaa gaaccagga
 4681 aagagttagg gctttgtttc cattttgttt taacaggaaa acaacggggg gatgggaatg
 4741 aaaagaaatg ttgaggaatg atatatttac tatatccatc catcttcata cgtttcaaaa
 4801 tcaaaaggca gatggatttc tgcctgctgt gcttaattgg tattatctac tcaagagtt

Partial genomic sequence of Genomic sequence of MAP1B showing the location of
 SNPs 1, 2 and 3 on the MAP1B gene and the locations of the primers (1F and 1R,
 corresponding to positions 2859-2876, and positions 3346-3362, respectively)
 used to amplify the region comprising the SNP sites. The number of the
 sequences is per that of GenBank Accession No. (514739)