

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 May 2009 (14.05.2009)

PCT

(10) International Publication Number
WO 2009/062042 A2

- (51) International Patent Classification:
C12Q 1/68 (2006.01)
- (21) International Application Number:
PCT/US2008/082811
- (22) International Filing Date:
7 November 2008 (07.11.2008)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/986,241 7 November 2007 (07.11.2007) US
- (74) Agent: **LI, Kening**; PINSENT MASONS, c/o, Baker Donelson Bearman, Caldwell & Berkowitz, PC, 555 Eleventh Street, NW, 6th Floor, Washington, District Of Columbia 20004 (US).
- (81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (*for all designated States except US*): **WISCONSIN ALUMNI RESEARCH FOUNDATION** [US/US]; 614 Walnut Street, Madison, Wisconsin 53726 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **KHATIB, Hasan** [IL/US]; 31D University Houses, Madison, Wisconsin 53705-1470 (US). **WEN, Huang** [CN/US]; 2302 University Avenue, Apt. 318, Madison, Wisconsin 53726 (US).

Published:
— *without international search report and to be republished upon receipt of that report*

(54) Title: METHODS AND COMPOSITIONS FOR IMPROVED CATTLE LONGEVITY AND MILK PRODUCTION

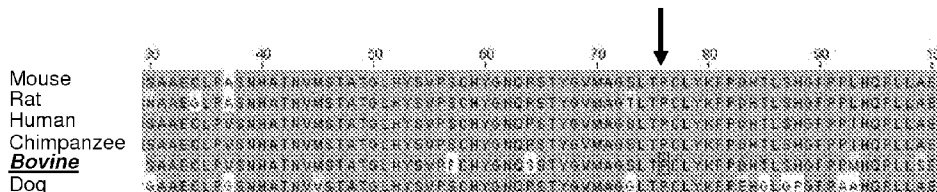


Figure 2

(57) Abstract: A single nucleotide polymorphic site at position 10793 of the bovine POU1F1 gene is associated with improved longevity and milk product traits. Disclosed are nucleic acid molecules, kits, methods of genotyping and marker assisted bovine breeding methods.



WO 2009/062042 A2

METHODS AND COMPOSITIONS FOR
IMPROVED CATTLE LONGEVITY AND MILK PRODUCTION

CROSS-REFERENCE TO RELATED APPLICATION: 60/986,241

FILED DATE: Nov. 7 2007

GOVERNMENT INTEREST

- 5 [0001] This invention was made with United States government support awarded by the following agencies: USDA/CSREES 05-CRHF-0-6055. The United States government has certain rights in this invention.

FIELD OF THE INVENTION

- 10 [0002] The present invention relates to a method of cattle progeny testing using molecular genetic methods by assaying for the presence of at least one genetic marker which is indicative of longevity and improved milk production.

BACKGROUND OF THE INVENTION

- 15 [0003] Dairy cows are significant investments for dairy farmers, and enormous efforts, such as animal breeding and artificial insemination, have been and continue to be invested in ensuring that the animals have high and sustained productivity, and that the milk produced are of high quality. Traditional breeding techniques involve the studying of sire progenies, and evaluating their traits including milk production ratings (transmitting abilities) to guide further breeding. This standard technique is time consuming and costly, requiring years to evaluate the true genetic value by progeny testing each bull. Many cows
20 must be bred and give birth to offspring. The females must be raised, bred, allowed to give birth and finally milked for a length of time to measure their phenotypic traits.

- [0004] Furthermore, selection based purely on phenotypic characteristics does not efficiently take into account genetic variability caused by complex gene action and interactions, and the effect of the environmental and developmental variants. There is thus a

5 need for a method of genetically evaluating cattle to enable breeders to more accurately select animals at both the phenotypic and the genetic level.

[0005] Marker-assisted selection can lower the high cost of progeny testing currently used to improve sires, since young bull progeny could be evaluated immediately after birth, and young bulls that are determined by genetic testing to have undesirable markers would
10 never be progeny tested. Testing may even be conducted prior to birth, for the presence/absence of the marker. Therefore, there is also a need for genetic markers for improved milk production traits.

[0006] POU1F1 is a member of the tissue specific POU (Pit, Oct, Unc) homeobox transcription factor DNAbinding protein family that is found in all mammals studied so far
15 (Bastos et al., 2006; Ingraham et al., 1988; Ingraham et al., 1990). The pituitary specific expression of POU1F1 is required for the activation of growth hormone (GH), prolactin (PRL), and thyroid stimulating hormone (TSH) (Li et al., 1990). These genes are involved in a variety of signaling pathways that are important for many developmental and physiological processes, including pituitary gland development (Li et al., 1990, Mullis,
20 2007), mammary gland development and growth (Svennersten-Sjaunja and Olsson, 2005), milk protein expression (Akers, 2006), and milk production and secretion (Svennersten-Sjaunja and Olsson, 2005). Moreover, binding of GH and PRL to their receptors on the cell membrane triggers a cascade of signaling events including the JAK/STAT pathway, which has been shown to be required for adult mammary gland development and lactogenesis (Liu
25 et al., 1997).

[0007] Mutations in *POU1F1* often result in severe GH deficiency as well as defects in development (Mullis, 2007). In a dwarf mouse model, mutations in *POU1F1* lead to the loss of three pituitary cell types— somatotropes, lactotropes and thyrotropes — (Li et al., 1990). Lactotropes produce prolactin, which is necessary for mammary gland development and
30 lactation.

[0008] Several genes in the same pathway of POU1F1 have been reported to be associated with different milk production and health traits. For example, growth hormone

5 receptor (*GHR*) and prolactin receptor (*PRLR*) have shown associations with milk yield and composition (Viitala et al., 2006). Also, the signal transducer and activator of transcription 1 (*STAT1*) and osteopontin (*OPN*) genes have been shown to have significant effects on milk yield and milk protein and fat yields in Holstein dairy cattle (Cobanoglu et al., 2006; Leonard et al., 2005; Schnabel et al., 2005). The uterine milk protein (*UTMP*) is another
10 gene in the pathway of *POUIF1* that has been found to be associated with productive life in dairy cattle (Khatib et al., 2007b).

[0009] *POUIF1* is located on bovine chromosome region BTA1q21-22 (Woollard et al., 2000), where multiple quantitative trait loci (QTL) affecting milk production traits have been identified (Georges et al., 1995; Nadesalingam et al., 2001). In previous studies,
15 *POUIF1* variants have been reported to be associated with milk yield and conformation traits (Renaville et al., 1997; Tuggle and Freeman, 1994). Taken together, the biological functions of *POUIF1* and associations with production traits of genes in the same pathway of *POUIF1* suggest that this gene could be functionally involved in milk yield and composition traits.

20 SUMMARY OF THE INVENTION

[0010] The present inventors investigated the effects of *POUIF1* on health and milk composition traits in two independent North American Holstein cattle populations. A pooled DNA sequencing approach was used to identify single nucleotide polymorphisms (SNP) in the gene. A SNP (C/A) in exon 3 of *POUIF1* that changes a proline to a histidine was
25 identified. A total of 2141 individuals from two independent North American Holstein cattle populations were genotyped for this SNP using a modified PCR-RFLP method. The frequencies of allele A were 14.9% and 16.8% in the two examined populations respectively. Statistical analysis revealed significant association of *POUIF1* variants with milk yield and productive life, which makes *POUIF1* a strong candidate for marker assisted
30 selection in dairy cattle breeding programs.

[0011] Based on the results, the present invention provides an isolated nucleic acid molecule comprising a polymorphic site of position 10793 ("SNP 10793") of SEQ ID NO: 1

5 and at least 17 contiguous nucleotides or bases of SEQ ID NO: 1 adjacent to the polymorphic site, wherein the nucleic acid molecule comprises an adenine base at position 10793 of SEQ ID NO: 1. It is recognized that SEQ ID NO: 1 is already known, and the nucleic acid molecule therefore does not encompass one that consists of SEQ ID NO: 1.

10 [0012] Preferably, the nucleic acid molecule which comprises at least 15, more preferably at least 20, still more preferably at least 25, contiguous bases of SEQ ID NO: 1 adjacent to the polymorphic site. In one embodiment, the isolated nucleic acid molecule comprises not more than 1,500 nt, preferably not more than 1000 nt, more preferably not more than 900 nt, more preferably not more than 800 nt, more preferably not more than 700 nt, preferably not more than 600 nt, more preferably not more than 500 nt, preferably not
15 more than 400 nt, more preferably not more than 300 nt, more preferably not more than 150 nt., preferably not more than 100 nt., still more preferably not more than 50 nt.

[0013] The nucleic acid molecule preferably contains the polymorphic site which is within 4 nucleotides of the center of the nucleic acid molecule. Preferably, the polymorphic site is at the center of the nucleic acid molecule.

20 [0014] In another embodiment, the nucleic acid molecule contains the polymorphic site which is at the 3'-end of the nucleic acid molecule.

[0015] The present invention also provides an array of nucleic acid molecules comprising at least two nucleic acid molecules described above.

25 [0016] The present invention further provides a kit comprising a nucleic acid molecule described above, and a suitable container.

[0017] Also provided is a method for detecting single nucleotide polymorphism (SNP) in bovine POU1F1 gene, wherein the POU1F1 gene has a nucleic acid sequence of SEQ ID NO: 1, the method comprising determining the identity of a nucleotide at position 10793, and comparing the identity to the nucleotide identity at a corresponding position of SEQ ID
30 NO: 1.

5 [0018] In another embodiment, the present invention provides a method for genotyping a bovine cell, using the method above. Suitable 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 sequencing the POU1F1 gene, or a relevant fragment thereof, isolated from the cell. The POU1F1 gene or a relevant fragment thereof is isolated from the cell via
10 amplification by the polymerase chain reaction (PCR) of genomic DNA of the cell, or by RT-PCR of the mRNA of the cell. Preferably, the PCR or RT-PCR is conducted with a pair of primers having the following sequences:

CAAATGGTCCTTTTCTTGTTGTTACAGGGAGCTTAAGGC (SEQ ID NO: 2)

CTTTAACTCATTGGCAAACCTTTTC (SEQ ID NO: 3).

15 [0019] In a further embodiment, the present invention provides a method for progeny testing of cattle, the method comprising collecting a nucleic acid sample from the progeny, and genotyping said nucleic sample as described above.

[0020] Further provided is a method for selectively breeding cattle using a multiple ovulation and embryo transfer procedure (MOET), the method comprising superovulating a
20 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, genotyping the developing embryo, and terminating pregnancy if the developing embryo does not have adenine (A) at position 10793. Preferably, pregnancy is terminated if the embryo is homozygously A at position 10793.

25 [0021] In a preferred embodiment, the method is used for selectively breeding dairy cattles, comprising selecting a bull that is hemizyously or homozygously A at position 10793 of its POU1F1 gene, and using its semen for fertilizing a female animal. Preferably the bull is homozygously A at position 10793. More preferably, the female animal is also hemizyously or homozygously A at position 107931, preferably homozygously A. MOET
30 procedure may be preferably used for the selective breeding.

5 [0022] The present invention also provides a method for testing a dairy cattle for longevity or its milk production trait, or both, comprising genotyping its cells, wherein a cattle being homozygously A at position 107931 indicates that the cattle has desirable longevity or milk production trait.

DESCRIPTION OF THE DRAWINGS

10 [0023] **Figure 1** shows the POU1F1 gene sequence (SEQ ID NO: 1) where the relevant polymorphic site is shown.

[0024] **Figure 2** shows the protein sequence alignment of POU1F1 from mammalian species. Protein sequences of POU1F1 from mouse, rat, human, chimpanzee, bovine, and dog were aligned using the multiple alignment algorithm ClustalW and visualized with
15 Jalview(EBI). Numbers on the top are the relative positions of amino acids. The position of the Pro76His mutation is indicated by the arrow.

DETAILED DESCRIPTION OF THE INVENTION

[0025] It has been found that a specific site, i.e. position 10793 (see fig.1), in the POU1F1 gene sequence is polymorphic. The term “polymorphism” as used herein refers to
20 the occurrence of two or more alternative genomic sequences or alleles between or among different genomes or individuals. “Polymorphic” refers to the condition in which two or more variants of a specific genomic sequence can be found in a population. A “polymorphic site” is the locus at which the variation occurs. Polymorphisms generally have at least two alleles, each occurring at a significant frequency in a selected population. A
25 polymorphic locus may be as small as one base pair. The first identified allelic form is arbitrarily designated as the reference form, and other allelic forms are designated as alternative or variant alleles. The allelic form occurring most frequently in a selected population is sometimes referred to as the wild type form. Diploid organisms may be homozygous or heterozygous for allelic forms. A biallelic polymorphism has two forms,
30 and a triallelic polymorphism has three forms, and so on.

5 [0026] Polymorphisms may result in functional differences, through changes in the encoded polypeptide, changes in mRNA stability, binding of transcriptional and translation factors to the DNA or RNA, and the like. Polymorphisms are often used to detect genetic linkage to phenotypic variation.

10 [0027] One type of polymorphism, single nucleotide polymorphisms (SNPs), has gained wide use for the detection of genetic linkage recently. SNPs are generally biallelic systems, that is, there are two alleles that an individual may have for any particular SNP marker. In the instant case, SNPs are used for determining the genotypes of the POU1F1 gene, which are found to have strong correlation to longevity and milk production traits.

15 [0028] In the context of the present specification, the provided sequences also encompass the complementary sequence, including those corresponding to the provided polymorphisms. In order to provide an unambiguous identification of the specific polymorphic site the numbering of the original POU1F1 sequence in the GenBank is shown in Figure 1 and is used.

20 [0029] The present invention provides nucleic acid based genetic markers for identifying bovine animals with superior longevity and milk production traits. In general, for use as markers, nucleic acid fragments, preferably DNA fragments, will be of at least 12 nucleotides (nt), preferably at least 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.

25 [0030] 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
30 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

5 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.

10 [0031] 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
15 base-specific manner to a complementary strand of nucleic acid.

[0032] 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 C. An oligonucleotide probe can be used for such purpose. Preferably, the oligonucleotide probe will have a detectable label,
20 and contains an A at the corresponding position. Experimental conditions can be chosen such that if the sample DNA contains an A at the polymorphic site, a 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.

25 [0033] 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 POU1F1 gene. The use of the polymerase chain reaction is described in Saiki et al. (1985) Science 230:1350-1354.
30 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

5 (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.

10 [0034] 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
15 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
20 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.

[0035] Hybridization may also be performed with nucleic acid arrays and subarrays such
25 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, for example the polymorphism of position 10793 may be represented by either, or both, of the listed nucleotides. Usually such an array will include at least 2 different polymorphic sequences, i.e. polymorphisms located at unique positions
30 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

5 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.

[0036] The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using
10 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
15 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).

[0037] 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
20 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
25 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).

30 [0038] 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-

5 2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-
tetramethyl-6-carboxyrhodamine (TAMRA), radioactive labels, e.g. ^{32}P , ^{35}S , ^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
10 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.

[0039] 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
15 has a maximum effect 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
20 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
25 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
30 nucleotides of the center," and so on.

[0040] 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

5 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.

[0041] Alternatively, the relevant portion of the POU1F1 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 Figure 1. It is readily recognized that, other than those
10 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.

[0042] DNA markers have several advantages; segregation is easy to measure and is unambiguous, and DNA markers are co-dominant, i.e., heterozygous and homozygous
15 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
20 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
25 stage may be assayed for presence of the marker, and selection decisions made accordingly.

[0043] In one embodiment of the invention an assay is provided for detection of presence of a desirable genotype using the markers.

[0044] 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
30 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

5 the specific allele or the specific nucleotide carried by an individual at a polymorphic marker.

[0045] 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
10 desired genotypes of the present invention, some of which are indicative of improved milk production traits.

[0046] Further provided is a method for genotyping the bovine POU1F1 gene, comprising determining for the two copies of the POU1F1 gene present the identity of the nucleotide pair at position 10793.

15 [0047] One embodiment of a genotyping method of the invention involves examining both copies of the POU1F1 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
20 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.

[0048] The present invention further provides a kit for genotyping a bovine sample, the
25 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
30 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.

5 [0049] In one embodiment the present invention provides a breeding method whereby genotyping as described above is conducted on bovine embryos, and based on the results, certain cattle are either selected or dropped out of the breeding program.

[0050] Through use of the linked marker loci, procedures termed “marker assisted selection” (MAS) may be used for genetic improvement within a breeding nucleus; or
10 “marker assisted introgression” for transferring useful alleles from a resource population to a breeding nucleus (Soller 1990; Soller 1994).

[0051] The present invention discloses the association between *POUIF1* and milk production and longevity in a total of 2141 individuals from two independent Holstein dairy cattle populations. SNP10793 allele A was associated with a significant increase in PTA for
15 milk yield in the CDDR granddaughter design population but not in the daughter design UW resource population. Although the granddaughter design has more power than the daughter design for detecting QTL (Weller et al., 1990), the use of PTA values in the CDDR population may limit the detection of epistasis and dominance effects. Thus, to test whether there is any genetic interaction between the A and C alleles of SNP10793, genotypic effects
20 were estimated in the UW population using the YD data. Genotype AA was found to be associated with an increase in milk yield and productive life compared to the CC and AC genotypes. This is an indication for complete dominance of the C allele over the A allele in determining the phenotypic value of productive life and milk yield. Also, this could explain the lack of significant association between *POUIF1* and productive life when PTAs were
25 used in the allele substitution model.

[0052] A different SNP located in exon 6 of *POUIF1* has been reported to be associated with milk yield (Tuggle and Freeman, 1994; Renaville et al., 1997). The population size used in those studies was relatively small (115 and 98, respectively) compared to a total of
30 2141 individuals from two independent populations investigated in the current study. In addition, the SNP reported in the current study is a missense mutation compared to a synonymous mutation SNP reported in Tuggle and Freeman (1994) and Renaville et al. (1997).

5 [0053] The candidate gene approach has been widely and successfully used in medical and agricultural studies to identify underlying genes responsible for complex traits such as susceptibility to diseases and production traits. We have used this approach and identified a number of genes including *OLRI* (Khatib et al., 2006, Khatib et al., 2007a), *PI* (Khatib et al., 2005), *OPN* (Leonard et al., 2005), *STAT1* (Cobanoglu et al., 2006), and *UTMP* (Khatib et al., 2007b) that are associated with milk production and health traits. In addition to the functional candidate gene approach, positional information about the investigated gene is usually incorporated into this approach to identify candidate genes. However, production traits are very complex by nature and determined by multiple factors including single gene effects, interaction between genes, and environmental factors. Therefore, in addition to positional and functional information of the single gene, functional information of the signaling pathway and regulatory network in which the candidate gene is involved should be incorporated to aid the identification of candidate genes. In light of this notion, *POU1F1* was chosen as a candidate gene for milk production traits. First, *POU1F1* is a transcription factor that controls the expression of *GH* and *PRL*, two important genes in mammary gland development and milk production and secretion. Second, genes that are downstream of the *POU1F1* signaling pathway (e.g. *STAT1*, *OPN*, *UTMP*) have been reported to be associated milk production and health traits. Third, the amino acid proline is highly conserved among mammalian species; as such mutations at this position could change the function of the protein.

25 [0054] In summary, based on the positional, functional, and regulatory information, *POU1F1* was chosen as a candidate gene for investigation of association with milk production and health traits. We identified SNP10793, a C to A nucleotide change that changes a proline to a histidine in the protein. The rarer AA genotype was associated with a significant increase in productive life and milk yield. These results suggest that *POU1F1* could be used in marker assisted selection programs in dairy cattle.

30 [0055] 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.

5 EXAMPLES

MATERIALS AND METHODS

[0056] *Cattle Population and Phenotypic Data*

[0057] Semen samples from 31 Holstein sires and their 1299 sons were obtained from the Cooperative Dairy DNA Repository (CDDR) of the USDA Bovine Functional Genomics
10 Laboratory (Beltsville, MD). Blood samples (n = 842) were obtained from the University of Wisconsin (UW) resource population (Gonda et al., 2006; Cobanoglu et al., 2006; Khatib et al., 2007b). Phenotypic data including predicted transmitting abilities (PTAs) and yield deviations (YDs) for milk yield, fat yield, protein yield, productive life, and SCS score were obtained from the USDA Animal Improvement Program Laboratory (Beltsville, MD).

15 [0058] *Single Nucleotide Polymorphism (SNP) Identification*

[0059] SNP were identified in *POUIF1* using the pooled DNA sequencing approach as described in Leonard et al. (2005). Briefly, genomic DNA was extracted from 30 individuals, quantified using a spectrophotometer, then equal amounts of DNA from each individual were pooled together and subjected to PCR amplification using different pairs of
20 primers designed in *POUIF1*. PCR products were sequenced using forward and reverse primers and SNP were identified by visual inspection of the chromatograms. To validate SNP identified in the pools, individuals composing these pools were also sequenced.

[0060] *SNP Genotyping*

[0061] Genotyping of the identified SNP was done by a PCR-restriction fragment length
25 polymorphism (PCR-RFLP) based method. To genotype SNP3699, primers (forward: atactcatcagagaactgcc and reverse: cattaaccctgttggtatgg) were used to amplify a 771 bp genomic fragment of *POUIF1*. The PCR products were digested with the restriction enzyme *TaqI*. Depending on the availability of restriction enzymes and suitability of the sequence, a PCR primer can be designed to change the nucleotide sequence near the SNP to
30 create a restriction site. For SNP10793, primers (forward: caaatggtccttttctgtgttacagggagcttaaggc and reverse: ctttaaactcattggcaaacttttc) were designed

5 to amplify a PCR product of 234 bp. Two Cs were mutated to Gs at positions 2 and 3 nucleotides upstream of the SNP in order to create a recognition site for the restriction enzyme *StuI*.

[0062] A touchdown PCR program was used as follows: initial denaturing at 94°C for 5 min, followed by 33 cycles of 94°C for 45s, touchdown annealing for 45s (from 63 °C to 50
10 °C, stay at 50 °C for 25 cycles), and 72 °C for 45s, and final extension at 72 °C for 7 min. The PCR products were subjected to *TaqI* or *StuI* (Promega, Madison, WI) digestion according to manufacturer's instructions, followed by 2% agarose gel electrophoresis. The A allele of SNP3699 was indicated by two bands of 338 and 433 bp, and the G allele was indicated by three bands of 84, 254, and 433 bp. For SNP10793, the C allele was indicated
15 by two bands of 38 and 198 bp, while the A allele was indicated by a single band of 234 bp.

[0063] *Statistical Analysis*

[0064] For the CDDR population, association analysis between number of alleles at the *POUIF1* locus and productive traits was carried out through a weighted least square allele substitution model of the following form:

$$20 \quad y_{ij} = \mu + \text{sire}_i + \beta x_{ij} + \varepsilon_{ij}$$

where y_{ij} is the PTAs of the trait considered, μ represents a general constant, sire_i is the fixed effect of the i^{th} sire, β represents half of the allele substitution effect ($\alpha/2$), x_{ij} is the number of A alleles (0, 1, 2) at SNP3699 or SNP10793, and ε_{ij} is the residual term.

[0065] For the UW resource population, association of *POUIF1* polymorphism with
25 production traits was evaluated with the following mixed effect model:

$$[0066] \quad y_{ijklm} = \mu + h_i + s_j + mgs_k + d_{ijkl}\tau + p_m + \varepsilon_{ijklm}$$

[0067] where y_{ijklm} represents in turn the yield deviation for milk protein and fat or productive life of daughter l of sire j and maternal grandsire k ; τ represents an effect associated with *M. Paratuberculosis* infectious status; d_{ijkl} is an indicator variable assuming
30 values 0 or 1 for non infected and infected cows respectively; p_m represents the effect of

5 POU1F1 (1 = AA, AG, GG). Herd h , sire s and maternal grand sire mgs effects were fitted in the model as random. In the analysis, correlation between individuals was not accounted for and therefore variance structure for sire and maternal grand sire effect had form, $I \otimes \sigma_s^2$ and $I \otimes \sigma_{mgs}^2$ respectively. Variance structure for herd effect was $I \otimes \sigma_h^2$. Standard assumptions were made for the residual term ε_{ijklm} . Additive genetic effect was estimated as half of the difference between the two homozygotes groups and dominant genetic effect was computed as the difference between heterozygote and the average of two homozygotes. Degree of dominance was estimated as the ratio of dominant effect over additive effect (Falconer and Mackay, 1996) with values approaching 1 indicating complete dominance (same effect for heterozygote and homozygote). All statistical analyses procedures were implemented using “lm” and “lme” of the freely and publicly available R software v. 2.5.1.

[0068] EXAMPLE 1: Identification of SNP

[0069] Sequencing of 30 grandsires from the CDDR populations and of the pooled DNA samples revealed four SNP identified in protein coding exons of *POU1F1*. An A/G SNP at position 3699 (GenBank accession number NW_001501776) was identified in exon 2, and 3 SNPs were identified in exon 3: SNP A/C at position 10793, SNP C/T at position 10822, and SNP A/G at position 10863. Importantly, SNP3699 (exon 2), SNP10822 (exon 3), and SNP10863 (exon 3) were found to be in complete linkage disequilibrium (LD), therefore only SNP3699 was used for genotyping. SNP10793 (exon 3) was not in LD with other SNPs, therefore it was genotyped independently. SNP3699, SNP10822, and SNP10863 are synonymous mutations whereas SNP10793 is a missense mutation in which the change from a C to an A (minor allele) changes amino acid 76 of the POU1F1 protein from proline (Pro) to histidine (His). Alignment of protein sequences of POU1F1 from mouse, rat, human, chimpanzee, bovine, and dog using the multiple alignment algorithm ClustalW, revealed that proline is highly conserved among these species (Figure 2).

30

5 **EXAMPLE 2:** *Association of POU1F1 with Milk Production and Health Traits*

[0070] The allele and genotype frequencies of SNP3699 and SNP10793 and the corresponding chi square test of Hardy-Weinberg equilibrium (HWE) are listed in Table 1. The genotype frequencies were consistent with those expected of a population in Hardy-Weinberg equilibrium. Association testing of SNP3699 in the CDDR population did not
10 show significance with any of the examined traits (data not shown), so this SNP was not investigated in the UW resource population.

[0071] In contrast to SNP3699, SNP10793 was found to be significantly associated ($P = 0.027$) with milk yield in the CDDR population using the allele substitution model (Table 2). PTAs analysis in the UW population did not detect association of milk yield with
15 *POU1F1* locus. However, in the YD analysis, AA genotype was found to be associated with higher milk yield (Table 3). The AA genotype was also found to be positively associated with productive life. Because of PTAs additivity assumptions, dominance effects were estimated in the UW population using yield deviation data (Table 3). For both, yield and
20 productive life, the ratio of dominant effect over additive was close to 1, suggesting complete dominance. Frequencies of the allele A of SNP10793 were 15% and 17% in the CDDR and UW populations, respectively (Table 1).

Table 1. Allele and genotype frequencies and tests HWE of the identified SNPs of *POU1F1* in the CDDR and UW resource Holstein cattle populations

| Population | MAF ^a | Genotype | χ^2 (DF = 1) ^b | |
|------------|---------------------------------|----------|-----------------------------------|-------|
| CDDR | SNP3699 (n=480) ^c | 0.18 | n(AA)=12, n(AG)=149, n(GG)=319 | 1.228 |
| | SNP10793 (n=1299) | 0.15 | n(AA)=31, n(AC)=325, n(CC)=943 | 0.227 |
| UW | SNP10793 (n=842) | 0.17 | n(AA)=26, n(AC)=231, n(CC)=585 | 0.300 |

25 ^a MAF: minor allele frequency, A allele in SNP3699 and A allele in SNP10793.

^b For degree of freedom of 1, the 5% significance level for χ^2 is 3.84.

^c n = number of individuals genotyped

5 **Table 2.** Estimates of the allele substitution effects and standard errors (SE) of SNP10793^a for production trait PTA values in the CDDR and UW Holstein cattle populations

| Traits | CDDR $\alpha/2 \pm \text{SE}$ | UW $\alpha/2 \pm \text{SE}$ |
|--------------------|----------------------------------|--------------------------------|
| Milk yield | 72.24 ± 32.64* | 37.11 ± 35.02 |
| Fat yield | 1.69 ± 1.23 | 1.59 ± 1.34 |
| Fat percentage | -0.362 ± 0.500 | 0.052 ± 0.482 |
| Protein yield | 1.22 ± 0.835 | 0.792 ± 0.936 |
| Protein percentage | -0.353 ± 0.229 | -0.101 ± 0.231 |
| Productive life | -0.520 ± 0.684 | 0.028 ± 0.047 |

^a The effect of substituting allele C with allele A.

* $P < 0.05$.

10

Table 3. Estimates of the genotypic effects, standard errors (SE), and additive and dominance effects of SNP10793 in the UW resource Holstein cattle population

| Effect | Milk yield | P value | Productive life | P value |
|----------------------------------|------------------|-----------|-----------------|-----------|
| Genotype effect ^a | | | | |
| AC | -13.26 ± 150.88 | 0.9301 | 0.46 ± 0.87 | 0.5973 |
| AA | 722.55 ± 378.80 | 0.0592 | 5.17 ± 2.25 | 0.0240 |
| Dominance effect ^a | -374.54 ± 220.66 | 0.0926 | -2.12 ± 1.30 | 0.1065 |
| Additive effect | 361.28 ± 189.40 | 0.0592 | 2.58 ± 1.12 | 0.0240 |
| Degree of dominance ^b | 1.04 | | 0.82 | |

15 ^a Estimates of yield deviations from the UW population, the effect of genotype CC was arbitrarily set to zero

^b Degree of dominance was estimated as the ratio of dominance effect over additive effect

5 REFERENCES

- Akers, R. M. 2006. Major advances associated with hormone and growth factor regulation of mammary growth and lactation in dairy cows. *J. Dairy Sci.* 89(4):1222-1234.
- Bastos, E., I. Santos, I. Parmentier, J. L. Castrillo, A. Cravador, H. Guedes-Pinto, and R. Renaville. 2006. *Ovis aries* POU1F1 gene: cloning, characterization and polymorphism analysis. *Genetica* 126(3):303-314.
- 10 Cobanoglu, O., I. Zaitoun, Y. M. Chang, G. E. Shook, and H. Khatib. 2006. Effects of the signal transducer and activator of transcription 1 (STAT1) gene on milk production traits in Holstein dairy cattle. *J. Dairy Sci.* 89(11):4433-4437.
- Falconer, D.S. and T.F.C. Mackay. 1996. *Introduction to Quantitative Genetics*. 4th ed. Addison Wesley Longman Limited, England.
- 15 Georges, M., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto, A. T. Pasquino, L. S. Sargeant, A. Sorensen, M. R. Steele, X. Zhao, and et al. 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* 139(2):907-920.
- 20 Gonda, M. G., Y. M. Chang, G. E. Shook, M. T. Collins, and B. W. Kirkpatrick. 2006. Genetic variation of *Mycobacterium avium* ssp. *paratuberculosis* infection in US Holsteins. *J. Dairy Sci* 89(5):1804-1812.
- Ingraham, H. A., R. P. Chen, H. J. Mangalam, H. P. Elsholtz, S. E. Flynn, C. R. Lin, D. M. Simmons, L. Swanson, and M. G. Rosenfeld. 1988. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. *Cell* 55(3):519-529.
- 25 Ingraham, H. A., S. E. Flynn, J. W. Voss, V. R. Albert, M. S. Kapiloff, L. Wilson, and M. G. Rosenfeld. 1990. The POU-specific domain of Pit-1 is essential for sequence-specific, high affinity DNA binding and DNA-dependent Pit-1-Pit-1 interactions. *Cell* 61(6):1021-1033.

- 5 Khatib, H., E. Heifetz, and J. C. Dekkers. 2005. Association of the protease inhibitor gene with production traits in Holstein dairy cattle. *J. Dairy Sci* 88(3):1208-1213.
- Khatib, H., S. D. Leonard, V. Schutzkus, W. Luo, and Y. M. Chang. 2006. Association of the OLR1 gene with milk composition in Holstein dairy cattle. *J. Dairy Sci.* 89(5):1753-1760.
- 10 Khatib, H., G. J. Rosa, K. Weigel, F. Schiavini, E. Santus, and A. Bagnato. 2007a. Additional support for an association between OLR1 and milk fat traits in cattle. *Anim. Genet.* 38(3):308-310.
- Khatib, H., V. Schutzkus, Y. M. Chang, and G. J. Rosa. 2007b. Pattern of expression of the uterine milk protein gene and its association with productive life in dairy cattle. *J. Dairy Sci.* 90(5):2427-2433.
- 15 Leonard, S., H. Khatib, V. Schutzkus, Y. M. Chang, and C. Maltecca. 2005. Effects of the osteopontin gene variants on milk production traits in dairy cattle. *J. Dairy Sci.* 88(11):4083-4086.
- Li, S., E. B. Crenshaw, 3rd, E. J. Rawson, D. M. Simmons, L. W. Swanson, and M. G. Rosenfeld. 1990. Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene pit-1. *Nature* 347(6293):528-533.
- 20 Liu, X., G. W. Robinson, K. U. Wagner, L. Garrett, A. Wynshaw-Boris, and L. Hennighausen. 1997. Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes Dev.* 11(2):179-186.
- 25 Mullis, P. E. 2007. Genetics of growth hormone deficiency. *Endocrinol. Metab. Clin. North Am.* 36(1):17-36.
- Nadesalingam, J., Y. Plante, and J. P. Gibson. 2001. Detection of QTL for milk production on Chromosomes 1 and 6 of Holstein cattle. *Mamm. Genome* 12(1):27-31.

- 5 Renaville, R., N. Gengler, E. Vrech, A. Prandi, S. Massart, C. Corradini, C. Bertozzi, F. Mortiaux, A. Burny, and D. Portetelle. 1997. Pit-1 gene polymorphism, milk yield, and conformation traits for Italian Holstein-Friesian bulls. *J. Dairy Sci.* 80(12):3431-3438.
- Schnabel, R. D., J. J. Kim, M. S. Ashwell, T. S. Sonstegard, C. P. Van Tassell, E. E. Connor, and J. F. Taylor. 2005. Fine-mapping milk production quantitative trait loci on
10 BTA6: analysis of the bovine osteopontin gene. *Proc. Natl. Acad. Sci. U S A* 102(19):6896-6901.
- Svennersten-Sjaunja, K. and K. Olsson. 2005. Endocrinology of milk production. *Domest Anim. Endocrinol* 29(2):241-258.
- Tuggle, C. K. and A. E. Freeman, Inventors. 1994. Genetic marker for improved milk
15 production traits in cattle. Iowa State University Research Foundation, Inc., assignee. U.S. Pat. No. 5,614,364.
- Viitala, S., J. Szyda, S. Blott, N. Schulman, M. Lidauer, A. Maki-Tanila, M. Georges, and J. Vilkki. 2006. The role of the bovine growth hormone receptor and prolactin receptor genes in milk, fat and protein production in Finnish Ayrshire dairy cattle. *Genetics* 173(4):2151-
20 2164.
- Weller, J. I., Y. Kashi, and M. Soller. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci* 73(9):2525-2537.
- Woollard, J., C. K. Tuggle, and F. A. Ponce de Leon. 2000. Rapid communication:
25 localization of POU1F1 to bovine, ovine, and caprine 1q21-22. *J Anim. Sci* 78(1):242-243.

5 WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a polymorphic site at position 10793 and at least 15 contiguous bases of SEQ ID NO: 1 adjacent to the polymorphic site, wherein the nucleic acid molecule comprises an adenine base at position 10793, or a nucleic acid molecule that is fully complementary to the nucleic acid molecule.

10 2. A nucleic acid molecule according to Claim 1, which comprises at least 17 contiguous bases of SEQ ID NO: 1 adjacent to the polymorphic site.

3. A nucleic acid molecule according to Claim 1, which comprises at least 20 contiguous bases of SEQ ID NO: 1 adjacent to the polymorphic site.

15 4. An isolated nucleic acid molecule according to Claim 1, which comprises not more than 150 nt.

5. An isolated nucleic acid molecule according to Claim 1, which comprises not more than 100 nt.

6. An isolated nucleic acid molecule according to Claim 1, which comprises not more than 50 nt.

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

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

25 9. A nucleic acid molecule according to Claim 1, wherein the polymorphic site 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 8.

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

 12. A method for detecting single nucleotide polymorphism (SNP) in bovine POU1F1 gene, wherein the PI gene have a nucleic acid sequence of SEQ ID NO: 1, the method comprising determining the identity of a nucleotide at position 10793, and
10 comparing the identity to the nucleotide identity at a corresponding position of SEQ ID NO: 1.

 13. A method for genotyping a bovine cell, comprising obtaining a nucleic acid sample from said cell and determining the identity of the nucleotide of a position of 10793 of the bovine POU1F1 gene according to claim 12.

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

 16. A method according to Claim 16, wherein the gene or a relevant fragment thereof is isolated from the cell via amplification by the polymerase chain reaction (PCR) of
20 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
25 nucleic acid sample from said progeny, and genotyping said nucleic sample according to Claim 13.

 19. A method for selectively breeding of 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
30 suitable male animal, implanting said fertilized eggs into other females allowing for an

5 embryo to develop, and genotyping said developing embryo according to Claim 15, and terminating pregnancy if said developing embryo does not have A at position 10793 of the POU1F1 gene.

20. A method according to Claim 19, wherein pregnancy is terminated if the embryo is not homozygously A at position 10793 of the POU1F1 gene.

10 21. A method for selectively breeding dairy cattles, comprising selecting a bull that is homozygously A at position 10793 of the POU1F1 gene and using its semen for fertilizing a female animal.

22. A method according to Claim 21, wherein the female animal is in vitro fertilized.

15 23. A method according to Claim 21, wherein MOET procedure is used.

24. A method according to Claim 21, wherein said female animal is also homozygously A at position 10793 of the POU1F1 gene.

20 25. A method for testing a dairy cattle for its longevity or milk production trait, or both, comprising genotyping its cells according to Claim 13, wherein a cattle homozygously having A at position 10793 of the POU1F1 gene indicates that the cattle has desirable milk production trait.

Figure 1

```

1  ttctccgttt ctattctttt gtgggaatga gttgcccaacc ttttacttcg actgatacct
61  ttatacctct gaattctgag tcttctgcaa ctctgcctct gataatgcat cccagtgctg
121 cggagtgctt accggtctcc aaccacgcc aacacgtgat gtccacaggt actaacttca
181 ataacagtct acatgctgct ggccttaag atatcaggat gggtgctttg aatcttgcta
241 gtttagaatc tcattttaaa aaaaatttta atgtgtatta agttaaatat tcaggatatt
301 aaagaagaca aaatgcctaa gaaaatattc aagaaatatt agtaagtga gtaatttcaa
361 gtattggtga ttttattata aaacctgtct tctatacaag taaacagtga gtctgaaaac
421 cactctatgc aaacatgat gcataaagag taaaactaaa aaaatagtgt aaaataattt
481 aaaaggtgat ttttttccct cttagagatt tatttggtt ttagattaca tttactacca
541 tgtctaaaaa atgccgtaca cctctgagta aacacaactg aatgattcct gataccaaca
601 acagtgggtt tcaaagaata aaataattag atgtttaaat acactttatt cagatggtaa
661 agaactctcc tacaatgcag gagaccagc tctgatccct gggtcaggaa gatcccagtg
721 agaagggaat ggcaaccac tccactatth ttgcctggcg aagtccaagg acagaggagc
781 ctggcaggtt acagtccatg gggttgcaa gagtcggaca caactgaacg actaacactt
841 tcactttaaa tacaactaac ttctttagaa taattaaaaa ctcagaaacc aagtctagag
901 gcaatataat atttatttgt ttatttttaa aactttttat tttatattgg agtatagcca
961 gttaacaatg ttgtgatatt ttcagatgga cagcagagga attcagccat acatacacta
1021 tacccttttt tacccaaact cccctccc ataggctaga ggcaatataa ttttagaat
1081 ccatgaactg atttgaatct tttcaagatt tagttctgaa gagatttcaa ctttaaatca
1141 agtttttata ttttcttata taactttctt agatgaaaat taaaattaat tcaaatatat
1201 tgtctcaaac acaaccaagt aaataaatat gttttgttaa tgtgtctaatttttactcag
1261 taacaaaggg ttttgattaa taaacaatgt cttgggtaaa tgtctgttta atagcacttt
1321 tcacttgtaa agcaatctaa atatgagcta aattttataa atttaaaaat caatttgagt
1381 tgacacatth atatgccaat ttaaaatatt tttagccttat attttgtaa aattaagcag
1441 ctttgaaagt tatatgtaag tttggtatta atcttttaaca gaattttcaa atactgaag
1501 atcatattat ctagatttgg ggaaaattat taatgtagag atgtttttta tctactgaag
1561 agtctcagaa ttttaattaaa catatgctga tgcaaatatc ctcaaacatg gttttaaaaa
1621 aaagtaataa agtattttta tgaagaggtg tggataaat aaattttaa ac aattatcaca
1681 tattgttgta ttgttcagtc gctaagctg aactctgcaa ctctttgcaa ccccatggac
1741 tgcagcatgc cagacttccc tgtccttcac tatctcctgg actcttccca tacaattatc
1801 acatagctaa ttgtgtatth taaataatgg acattttaat agattgagaa tgagaaggat
1861 agtgcattth tgcataaat actgtatgtg gatagtatth ggctagaata aaaattggaa
1921 ttgaggttaa aattaagcag gtttaagctc aaaaccctat ttatttaaaa aaaaataaaa
1981 gtaaataaat aatgtatcta gtaaattagc ttgtcattta tatgacttcc ccaaatcagc
2041 aaaacaaaag aatttaaacc ttggaggtaa aaaaattcat aattagtaac aatgtcctaa
2101 atggctattg ttattgtttt aaatagcaat agcattgatt ctataatttc tccatcaaat
2161 tgtaaccgta tatatctaath tgtctaggta aagaattaga aatgattata atgagagggg
2221 aatggcaacc cactccagtg tttttgctg gagaatcca gggacggggg agcctgggtg
2281 gctgctgtct atggggctac acagagttgg acacgactga agcacttag caacagcagc
2341 agcataatga gaggagcagt tggatattca atgcgtttca tgtttcttta attttggat
2401 gcatagaaat tattataaca caaaattatt tcagttctat gtgtcagttg taaaatgaaa
2461 atttatacag atactttaaa atattaaaaa tgagttatag tttagaaaaa tgttttactt
2521 tatgtcttta attgatagac tacatattth caaaaaggac tataaatctc ttctaataatc
2581 tcatattatt caagaagata caattatatt tttgaaaatc caccactcta ttcaggacat
2641 aatgggaaaa agtataaaat aagatataaa agaaaacatt gattagtgat tttattgtta
2701 gtagttcatt gtgattcatg ttgaaagctt gatagaagtt aagtaattca aatttaattt
2761 caaaatacac cttaattcaa tgatacttgt agctatttht aatgaagtt aagttatatg
2821 ctggactcaa cttttataaa cctctagtta ctcagtgatc aatgaattta ttttttgacc
2881 agcttcatca ttagaattgt tcaattaatt gtgtgcatgc agaaaatcta caagctcatg
2941 gtttcaatac cagcctaaat acattgcaca ttgtcacaga gtttttgtaa gcctctgaat
3001 ttaatggagt tttatgatca tacccaataa aaggctagtg gctctgccac ttctctttg
3061 atggaagaaa tttgttaggt gaatgagcag aagtaaaagt aaagacccat gaagagttga
3121 tttctgtatt cattgctcca gcaaaagaca attttatggt accatgttat acgagaattt

```

Figure 1

3181 taagagacta ttctctggtc agttctttgc aaagtctaca ttgggtgaaa cctgcaagaa
 3241 aactgcagct tccagttgtg tgcgtgtgct cagtcactca gtcatgtctg acacttggca
 3301 acctcatgga ctgtagccca ccaaggttaa aactggtcag tcacgccaac ccagaacact
 3361 ggatataaatt aatgaaggct ctagtccact gtgatgattc atgaactcgt tacttgggaa
 3421 aaatgtcaac ccctagtttt agcataactca tcagagaact gcccccaaat gagaacaaat
 3481 tattggcata taactttaag aatagcataa atgtgtacat ttgaaatgaa acgaatgtgt
 3541 cttgaaatcct catacatttt cttaccagtc ccgtctattt tgcctttgat ccaaactcct
 3601 aatgttttgt gcacatgttt tgtggtgaca atgctgggaa acacagcaac aggacttcat
 3661 tattctgttc cttcctgtca ttatggaaac cagtcactca cctatggcgt gatggcagg
 3721 aagaaaaatt gtctttacat gtaagattga gtttggggac gcttggatgc attttctggg
 3781 tcgaagggaa tcttgaccag agtgtatcat gaaattcaga tctcctaacc ttagaaattg
 3841 ctgctaaatc caccacttac tataatggtc cctgatctgt aacgccttca gagatcataa
 3901 tagttatacc tgatcactgc tgttctccac atgacctgaaa tgaactgcta tgcttcttaa
 3961 cgctgtggtt tgctttgtat gatttttatt ctaattctct gttgccaaac tgtaattgt
 4021 cacttgctta tgccattggg ggccttgccct ccagtaataa taaggtagct gctagcctta
 4081 tgtaactatt taaatttaaa agtaaaattaa ctaaaataag gtcagattta aaattcctca
 4141 gttgccaaaca gtgagcaagt cagcaaaaagc tttaaatgac ttcattgccc gcctccatac
 4201 caacagggtt aatgattatg aggtacggaa cttaatagtc atcatttccc ttgagatttt
 4261 ggctgttcct tcagttcctg aattttaaata aaaacaggaa agtgcttaat aatcttttgt
 4321 ggctacaaa gcagatagac taacttatag agcatattcc tcatcacagt atttcttttg
 4381 gccatttggg ctgtaaagat aaataagaca ttgaaggcta agagagattt gcagcaactt
 4441 tagccatagt tgaatccttg gcagccctgt ctgagatgct ctgagcttg gataagattg
 4501 aatagatata agttacatac ttatttccct ctggataaag aaagcatcac tataatcaaga
 4561 taaactatct tgtttgctga cacttaaatg aaagtattgt tagcaacttt cttaggggaa
 4621 agaaaatctg acgtaagaga tccactcatc acaaaattta tgacttaatt ttagaccatg
 4681 gctatacagt aatttttttg tgcacagaaa ttatttagaa atctcaatct tttatattta
 4741 tttctgcagc aagaaataat ttcagtgaca ataaactact aatgttctat gtgaaaaata
 4801 taaacgtgtg tgtgtgcatg ctcgtgtaat agttgttatt caagtaggta attaaaaaat
 4861 caaacataga tttatgccta taaatttgaa attaaattga acaaagaatt tagttcacca
 4921 acctttaatt tctaaaaaatt tctcaaattt acactaaata aaacattaaa actttaaaat
 4981 aatataattct agaacaaatt agcttttctt acattctggt agagagcatt ctgtaaaac
 5041 ttatatcctc acctttatga tgaagtgaaa actgatgttc ttaatggfta cttttaaact
 5101 ttaatgtctc tcttcaataa ttaaaaaata aggaaagtca tatttttctt cctctgtgaa
 5161 aacggaccgt ttatgattcg tgtacttagc tgtaatcag tatccatggt taatgagagg
 5221 atctctcggt gaagatgaac tactttcatt atagaaatag tagaacaatg tgcatttagc
 5281 cactagcggg gccgctcaaa tgctttctgt ggccctccacc acacttggat tacagagaca
 5341 agaaatgatg agtttataga catgtgctac tctttcttgg aggagagagt ccttgcctgt
 5401 ccttcacagt agttcccagg aaccacatg caacttaagt catctctgca cccagcaagg
 5461 ggcagcaatg agtggagggt agtgctcaga ccaaacactg tgagatgatt tttactacta
 5521 cctccctccc accagatcta ggtagaagct gctgagttca cagctggcct gagaaacctg
 5581 acttcttcc tgtttccagg gttgcaggt acaagcccag gttcacactg agagtccaca
 5641 ggcaactcag gttgtgctgg tccaagttc ctgagcaggc cctcagcaaa ctttcatctc
 5701 cattagaagt taaaaaggag aggtctttgg aaaaaccttt taaatgatt caaagcccag
 5761 gttggtgata ataataaaaa aatggatgt cctaagcttt ggaggattg ctgtagtctt
 5821 aaacaaaaat tacatatatt ttgtcgctga aagaaaaatt caccatgtca agcatagaga
 5881 ttatatttga atatttttga gagggagatc attaaaatgt gatgtgcttt gtaatttgtt
 5941 acaatgtcac tattttgaca tatattcagt attagccttt gcatcacggc agatttaatc
 6001 tcagagaatc tgaacctttc tgctgtcatc cccagagaag ctacacaaat tgagtataa
 6061 aacacaggga ttcagaatat tcaggaattc acaaaagggt taaaatacac aagagaacc
 6121 tgataatgtc atttgaattt tctgagaagt ctgagcgtt catcaaatca gccactcca
 6181 cacaatccac agaagcgtct aatcaccaac aacgaagggt cattgtcaca tactccatat
 6241 aaaagtgaag ctgcacaaaag agcaatattt caaagggtgt gaaattgctt tatttgaata
 6301 tggttatttt atccctatac tctgatgctt atatgtagtc tgagcatttc attaaattaa
 6361 taactcactg gaaagcaaga gctacagctc atattttaag ctaaggttgg taggaaggat
 6421 tctaatacctc agatgaatcc tctgtacctc atagtccagg tgtggggata tcaaagagaa
 6481 tgattggtgg ggtgaaatga attagccaca ccaggaacta attacaaacc tcgttcaagt
 6541 tagcatctca aagtgttagt ctcttccat ccaagttgtg atgggtacca tacatagctc

Figure 1

```

6601 tatctgagtt gagttgtggc agctttctac caacttagcc tggttttctc cttgatcttc
6661 tcttcagata cattcatagt taaatttctt cttctgtttc ttctctccta agtacttaaa
6721 taattcaaca gcctgcagga tggataatag gaaacagggg attacaaata ctattcaaaa
6781 atctcttcta gataaacacg aaagaaaact aataacaaaa acctttcaaa attctgattt
6841 ctgggtatac aaggtggtgt ctgttctcat ttgtaaagct gggtgaaagt tggaaaacaa
6901 acttaatgag ctgtgtacct tgcccgccgt cctgtgtgaa ctgggctaat cagctacatg
6961 gtaatgattg ctaaaccag caaggttatg tgttttaatg gccaccaag gtgtgcagta
7021 agagtccttg attaaaaatt gatcttaaga ggaaagcaaa atagctgatg ttggaaaatt
7081 attcaaggag atgaatgctc tttttaaatg agatgtgaat aataagaata aaagatgaat
7141 atagctaaaa agtgtctttt ccacctgggc taggacaaga ggtaccagaa atatgtttca
7201 catttacata ggcaaacagt ctactttgga ggccatctct tcattcttaa attgtctttt
7261 ttctatttcc tttttttctt ttttgtttta ttttagttta ttttgcctc acttattctg
7321 gatgtgttga cactaagggg aggagtaatt ctgaccatct tttgcctctg gatcatgaaa
7381 ggcgcctgca gtagcatgga cactgtgtat tattccttaa attatgtagc atctgtctca
7441 acttcacaac tcaaaagcag ctacaggcaa tctgtaaaca aatggccatg gctgtgctcc
7501 aatcaaacct aatttacaga actaggcaga agacctgttt atatgtggat aggatgtagc
7561 aatgcacaaa agagaaaaaaa aatcacaatt caaatttatg atttgtttca cattgtcatg
7621 catgaacgta aaaaaagaaa ataaatggaa atatttaaat gaactgtgct gtgcatctct
7681 acagcaagggt gggggtattt tactatatcc ttcttcccct gtttgttctt ttaaagcttc
7741 tagctcacia cctcagagat gttgtcagtt agcagctgtc tgagccacgt gctgcagaaa
7801 agcagtgtgc cctgtatcat gggaccttga atcaggtgcc cgtaaggcat gctggacctc
7861 agcttcccat cttctgcttt taaagattta atctctgttc ctctccctcc tctgccgtcc
7921 agaagtccat gcagagcagt tcagagaggc tatggtggat taatctgcag ggtgaaaatc
7981 catcttggac tgtaaagaat gtacaaacct ttccaagcta ttaggtgtg cacttgttct
8041 aggcttatag acaagtttgg tttatcagtg ttttggaaaa tattcatgaa acttctgtgg
8101 cacacaactt cctgcatctt ttctctccgt gctcaaacc cgacttctta atattccagg
8161 ctttagaaaa acctttttaa atacctgtc acatatacca taccaatcaa tccccattht
8221 ctcccaata caggatcata taatagaaat cagtttgtct gaccatgctt gtgactgtcc
8281 caaatttctg tttaacatta gtttcaggaa aactgtgtcc ctgtggataa ttgtggatac
8341 tttcagttaa tgaaaaataa cgcagcacac atcgtgtttt ccttctctca ctagatgcca
8401 aacttttgaa catggtattht cttgtgcact ctttatgaac tcatgttcaa atthtthttc
8461 taaatgtcca tttctccaga gtatttataa agtatacact agctthtaaat tttccataa
8521 atagagatgg cacacttcca atatttttgt aaatatttat aaaactthttg tttgaaagca
8581 tttgtthtaat gtgaccaaaat atatttaaga tgcagaatct ttgagtcatc tgatttccct
8641 gagtacagat tacctgaaaa atgaactgat tattgatgtg gcctattgag gtattcacag
8701 ggcttactt ctagtttcaa ttgtatagtg aattcatctt agcactcaag gaacgttatt
8761 tgtthttata tgaattthta attgtggata aaaaaataca cactthtttc tctgaaata
8821 aaacaaagt ctttagataa atattaagta aatgttaaca taaaggatg ataaatthttt
8881 atataatagt atthtactta actagtagaa atctaaaact aacattaggc caaaaatgag
8941 ctttgtthaa atgtaacaaa cattaaaaaa attgaaattht gtaataaatt aaaaatgat
9001 aaaattgctt thtaaaattg tatctggtct atthcaatgta gcaccacta aagcaaaaag
9061 tggatgatta acaataacaa caaaaaactg gatagthttg atthttgaaa gtaatatgaa
9121 ttgctthtaga cagaaatagt thctthttgta tgtthtttat atccatgaaa cagctatact
9181 taththtatga taththtctt aaththctcac ttgtatgtat thtgtthtcag aacataata
9241 gtctactgga taggggtaca cthcaaaatt ataththtctc atataaataa tatgagcaa
9301 caacttactg acttgcaata thctthtctc attacatatt gatatathaa thtaagattt
9361 atataatctg gcaaaataca ctattactca tgtctgattt ctgactgtat thatagagaa
9421 attataaata gcataaaact gatgaththt thtagaagtt tctaagaatg atthtgcct
9481 atthtctata gthcaattht thtaaththt actcataaga aaaaatgata atthtgaat
9541 atthtaththt tatcaaaatg thcttgttht acccatgtca thtaataata agtgaatthc
9601 thtatgaataa aatgttaata tactgtgtaa gaaaacagaa ataaaatgtg tgggaattht
9661 gcttatacca thctgtcacc aatactccct gactgagatt ctctcttctt thcaaatgaa
9721 gaatgaaatc tagaagcaat thaaacaatta aatcatgtag ththttgaga thtctcattt
9781 aactaactca atggattcag tgcttcatca ththtatgagg atthtattt gtatctcaga
9841 gaaggaaatg gcaactccag ththcttgcc tggagaatcc cagggacaga ggagctgtt
9901 gggctgctgt ccatggggtc gcacagagtc ggacacgact gaagcgactt agcagcagca
9961 gcagcagcat gthcatatc tagaathtga thcttctctg acactthact ththcttctt

```

Figure 1

10021 aaatagtaat caagaaatga ttccatcata gtatattttg caagaatcct gttgctgacc
 10081 ttctaataga tctttaaaag ttactttttg tgcacatcca tcattaccat taaatctatt
 10141 taaaccttat taatgtattt tttgttctca taattttcgt tttacttgcc tttactaaat
 10201 taagcattta ctgataaaac attggatttc ttaatgctgt aaaatcaatt tttcaatgct
 10261 atgaaaatth cccagaatag cacttaacat acaccttgat tataattaag aaaatataca
 10321 ttctggtaga gttgaacctc agattcacia taacaaatga aaagattttt gtttgttttt
 10381 ggggagccat gctgaattcc ataaccagga atcaaacca tgccctctat actggaagga
 10441 tgaagtctta accactggac tgctatggaa gtcttttaag agaaaaacia tgatggaaa
 10501 ctcacacttt tttgaaatth gcttattttct atttaaatta tttgcaagtc cctggttctt
 10561 tcctttggct ttgcattatc tttgtttcca ctgctcccag gaaggtggaa aaaaagatc
 10621 ctatttatac taagctacac tgacttctac ttcagaaaag caaatgtcag gtaacctttt
 10681 agaactgaga ctggctgtca cagaacaatc tgatgggcca aaattttcca tgtatcaaaa
 10741 tgagggataa ttacaaatgg tccttttctt gttgttacag ggagcttaac cctttgtcct
 10801 tataagtttc ctgaccacac gttgagctat ggttttcctc ccatgcatca gcctctcctt
 10861 tcagaggacc ccactgccgc tgatttcaag caggagctca ggcgaaaag caaattgggt
 10921 gaagagccaa tagacatgga ttctccagaa atccgagaac ttgaaaagt tgccaatgag
 10981 tttaaagtga gaagaattaa gctaggtagg tgctgttaa cagctgtggg acacacaact
 11041 ccgtctgcaa agtcttactc tattactgtt taatctctta catgctgctc agaagtctaa
 11101 gacagttctc attctacatc tctactgtgg atgtaagttg aattatgaaa acctatagca
 11161 accttcattt ccttgtaaat tcttagcagc aaaaatata agatttctaa attaatggtc
 11221 tccttttcaa acataagttt agaaatacct ttgttttatt tgaattaata cctttgtgtg
 11281 attcaaaagc taaaagctg tgagaattgt atccctctgc ttatccttc tgtttatcag
 11341 tgttattagt ttctgtgaat ccttctattc tatgcatatt catataagca aattcatgta
 11401 tttcttcata tagacatttt tcatttgact attttgaga gctttcagta ttagtttttt
 11461 agagagctct tctttttaat gcatctctca attccatttt atgaatata taccagtctc
 11521 ctactgatgg aatttagatg gtcttcaatc ttatgttact gaagacaatg atgcaatgat
 11581 taacttcata aatagcattt tgcaccgata aaaatataatg tgcagggaaa gtgttaaaag
 11641 caaaattggt tgattgaaga gcatgtgtat ttgtaatttc agtagatgct aaattggctc
 11701 ctatgtaaat tctaacaaat cttattgcca ccatgagat ctctttgccc actcttcaga
 11761 atatgaattc tgtaaggca gaagtthttt gttttttgtt tatttctttt gctttctcct
 11821 taatggcaca tcctcagcac ttggaacagt gactggcatg tagaataaaa aatagttatt
 11881 gaagggaaaa gatgctatta atcttttggc cctttgcaa tcagttagg taaagtagca
 11941 gtttccattt aaatttccct tatgagtga gagtgaaaga agtttttcat gtgttgaca
 12001 gtcatttgta attccttttc tgtgagcgt ctgttcatag ctttgctca ctttctaggg
 12061 tgaagtgtt ttaatgtaa taatcataat aattaacct gatttaaatg catgaaat
 12121 atttttaatg gattgaactg atagtaaaa cttctgcatt tgtggactga gccattggta
 12181 acttttaatg atatcaattg atgggcatca tatgtaatca ttttatgcat acaggtatat
 12241 agccttgggc cctgaattaa tatgtagtta ctgtttgtca taaacacagc agtaggcatc
 12301 tttatgacat tcattttcaa tttacttttt atatgactgt gaatgtttca aattctacat
 12361 tgatgacatt tgtcaactta tattctgaga atgtttgaga caatctatga aaactttttg
 12421 cctggagata gaagcattac caaatgatat aataaatgct tggatgata cataaaatgt
 12481 tgtgtgacca aagtctcata ataagcatgc ttttagggaa agtaaacaca gttcttagtt
 12541 ttatttggtta acttcaacat gttgaatttt tcaactctac agctgagata aaaatatttg
 12601 tgatatatca ccatatagtt tacatattat attttaatat ctatagattt tgagcatatt
 12661 tcaacagatg cctaataata atgattagag agaattttta aatgtcttct aaagtgtgta
 12721 ttaaggattc cctggtagaa cagctggtaa agaaccacc tgcaatgcag gacaccccag
 12781 ttcaattcct ggtcaggaa gatcagttgg agaaggata ggctacccat tccagtattc
 12841 ttcggtctcc cagtggctca gctgttaaag aatctgcctg ctatgaggga gatctggatt
 12901 cactccctgg gttgggaaga tcccctggag atgggaacag ctaccaccctc cagtattctg
 12961 gcctggagaa ttccatgtac tgtatagtcc atggggttgc aaagaatagt ctgactacac
 13021 ttatattagg ttataaaaat gattcatgta taattactac agtatatagc acagtggcaa
 13081 aaaaataaat ctggatacat taaaaagaat catttacta ctttatacct atgctacatt
 13141 gtctagaaac ttttcttata tatttttgca aagtgtgttt aacacattta tccagtttg
 13201 ctaaataatga atggcagatg ttcctatctg aattcttttg gcttctaaaa tattaactta
 13261 ttaactagaa ggaatttttt aaaatactag acaattctac actgcataac cttactgtta
 13321 ttctaaattg ctaacaaatt tatcgttaa agcaatattt aatagttgac aaaaatacta
 13381 cacaaattta tacaatagt gacccaaact agtgtttctt gcaaaactga agctcatggc

Figure 1

```

13441 ctttgttatt ctttcacagc atacaccag acaaatggtg ggaagctct ggcagctgtg
13501 catggctctg aattcagtca aacaactatc tgccgatttg aaaacctgca gctcagcttc
13561 aaaaatgcat gcaaaactaaa agcaatatta tccaaatggc tggaggaagc cgagcaagta
13621 ggaggtacaa aagctgtggt tctggaacaa gtgatgtttt aacctaaaaa caatggtttc
13681 cctcagttga atttgtacta aagcaagagg tttgaagttt ggtttgattt ttctctttga
13741 catgaaaaat aagtatcttg tttcatcaca ctatgaagaa aagcaaggcc agtgaaagtg
13801 tagaaataaa tttattgaga aggtaaataa tgagagaata aaatatatag ggaaagtttc
13861 tacacaatgt ggcataagggtg tgaagtgggtg aatgattctt ttttaatgta tccagatttt
13921 ttctgtctgt gctatatact gttagtaatta ttcatagaatc atttttacaa cctaataataa
13981 gtgtagccag agcattcgca cacaccgttc tttctagtga atagcaagca attgctagat
14041 aaacaattta atgtgataaa aattatctac ttatattaat gtcaaggctg gctaaagagc
14101 aagatttgat agcattttaga gcaactgttt caacaaagat agggatgatg ttaaagacaa
14161 ctgttaatat ttataaagttt aatatttttt tctgtgttaa cattttaatc tagggcttgt
14221 aatctaaaaat gatgtatact gcaaatattt aaacaaaaaa tgtatggtaa ttctattttg
14281 tattgtttta taaagaaact ttaaatccaa atctgatagt taaaaaaaaa acctggtttg
14341 tttttgtaat tattcattgt tttcgcatca cagctttata caatgagaaa gttggtgcaa
14401 atgaaagaaa aaggaaaacgg agaacaacaa tcaggtatac ttttgagata ttaagagtta
14461 gtgggagaaga aaatgatatt ttacaaatgg aatgaacatt tgagtataat atagtttcaa
14521 tataacataa aaatgaatag agccaattga gaaaataggt gaaaaagcac aacattcaat
14581 aaattacttc tgagaaacag ctggaaattht aaaatttgat ggaaaaatat gtattgtttg
14641 attcaagaac agttttgctc tgcaagtttt ggataaaaaca gaagctgtac aatcacagct
14701 aaaaagaatg actgtttcta ctgtgtcata atgtgttgat ttatgtttag acataaatct
14761 tgctccggga aagaccccat ggactgtagc ctacaggttc ctctgtccat gggattttcc
14821 aggcaagaat aatggagtgg gttgccattt ccttctccag gagatcttcc cgaccagggg
14881 attgaacccg gatctcctac attgtaggca gatgctttac catctgagcc acaagggag
14941 tcatctatct atattatttc aaattaacaa aactggtcac tagtatttta gttgcttaa
15001 gttcaaaaatg acttctagca tttcaagcca gattgttcat ttatcttttt gtagtttccg
15061 tgaggctcat ggaggaattg ctaatatataa ggttttgttt tggttggtta gttgtacact
15121 aacattttca ataacctgag ttctggggga catttagaaa tgcatacaga attattttct
15181 tctcagtaag tcagtgcctt ctgtggcag aaagtggata aacaatgtcg gggttccctc
15241 ctttaatttct tcctgtgact ctggtaaaag gagcctacat gagacaagca tctaaatggt
15301 caaaaaaact tcacatttat tattgttgaa aagctttgaa ggtgttttca gtgtctttag
15361 gtttcttttt tacgttaatg ttagtactaa tatttaggaa atgtaaccta acttgatttt
15421 gatgggccta aaccatcatc tcccttcttt cctgccaact cctcacctcc cagtattgct
15481 gctaaagacg ccctggagag acactttgga gaacagaata agccttctc tcaggagatc
15541 ctgcggatgg ctgaagaact aaacctggag aaagaagtgg tgagggtttg gttttgtaac
15601 cgaaggcaga gagaaaaacg ggtgaagaca agcctgaatc agagtttatt tactatttct
15661 aaggagcatc tcgaatgcag ataggctctc ctattgtgta atagcgagtg tttctacttt
15721 tcattccttt ctcttctcca gccaaaatag aaattagtta tttggttagc ttcaaaaaat
15781 cacatcagta atttttgcag aagtgtttct tttctacttt aaaaataaat acaatttaaa
15841 ttatgttgat gaattattct cagaaggcac atgtacatt ttaagccaaa gactaatagg
15901 attcaaacaa tgattctgtc cctttcacta tatctttccc tctatctctc cc

```

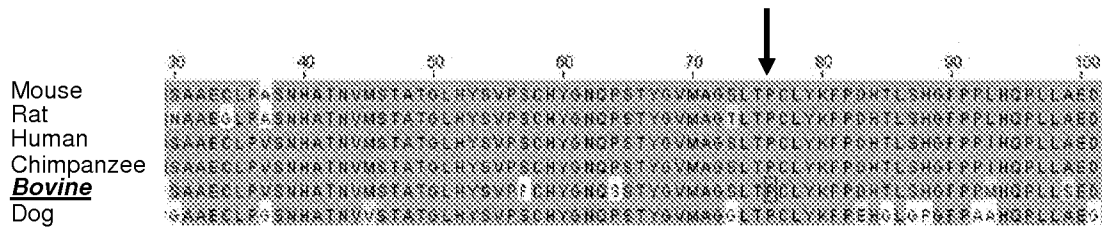



Figure 2