

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
20 November 2008 (20.11.2008)

PCT

(10) International Publication Number
WO 2008/140467 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

NISSING, Nicholas, J. [US/US]; Monsanto Company,
800 North Lindbergh Blvd., St. Louis, MO 63167 (US).

(21) International Application Number:
PCT/US2007/021187

(74) Agent: MADSEN, Matthew, L.; Monsanto Company,
800 North Lindbergh Blvd., St. Louis, MO 63167 (US).

(22) International Filing Date:
28 September 2007 (28.09.2007)

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, 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, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/848,541 29 September 2006 (29.09.2006) US
60/919,099 20 March 2007 (20.03.2007) US
60/931,680 24 May 2007 (24.05.2007) US

(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, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, 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): MONSANTO TECHNOLOGY LLC [US/US]; 800 North Lindbergh Blvd., St. Louis, MO 63167 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): DU, Fengxing [CN/US]; Monsanto Company, 800 North Lindbergh Blvd., St. Louis, MO 63167 (US). GROSZ, Michael, D. [US/US]; Monsanto Company, 800 North Lindbergh Blvd., St. Louis, MO 63167 (US). LOHUIS, Michael, M. [US/US]; Monsanto Company, 800 North Lindbergh Blvd. (US). TAO, Nengbing [US/US]; Monsanto Company, 800 North Lindbergh Blvd., St. Louis, MO 63167 (US).

Published:

- without international search report and to be republished upon receipt of that report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

(54) Title: GENETIC MARKERS AND METHODS FOR IMPROVING DAIRY PRODUCTIVITY AND FITNESS TRAITS

(57) Abstract: The present invention provides methods for improving desirable animal traits including improved fitness and productivity in dairy animals. Also provided are methods for determining a dairy animal's genotype with respect to multiple markers associated with milk production, somatic cell score, daughter pregnancy rate, productive life, and/or net merit. The invention also provides methods for selecting or allocating animals for predetermined uses such as progeny testing or nucleus herd breeding, for picking potential parent animals for breeding, and for producing improved progeny animals. Also provided are methods for identifying genes and/or genetic markers for traits including fitness and productivity traits; including identifying markers that are in allelic association with the SNPs disclosed herein and for identifying the genotypes of a causative mutation underlying a quantitative trait locus (QTL). The invention further provides methods for genotyping animals for multiple SNPs associated with the traits described above.



WO 2008/140467 A2

GENETIC MARKERS AND METHODS FOR IMPROVING DAIRY PRODUCTIVITY AND FITNESS TRAITS

PRIORITY CLAIM

[0001] This application claims the benefit of U.S. Provisional Application Serial No. 60/848,541, filed September 29, 2006. This application also claims the benefit of U.S. Provisional Application Serial No. 60/919,099, filed March 20, 2007. This application also claims the benefit of U.S. Provisional Application Serial No. 60/931,680, filed May 24, 2007. Each of the above-referenced applications is hereby incorporated by reference in their entirety.

INCORPORATION OF SEQUENCE LISTING

[0002] A sequence listing is contained in the file named "Dairy_ComboApp_st25_FINAL.txt" which is 211,087,360 bytes (201 megabytes) (measured in MS-Windows XP) and was created on September 27, 2007 and is located in computer readable form on a compact disc (in accordance with 37 C.F.R. §1. 52(e) and 37 C.F.R. §1. 1.821) enclosed herewith and incorporated herein by reference.

FIELD OF THE INVENTION

[0003] The invention relates to the enhancement of desirable characteristics in dairy cattle. Embodiments relate to genes, gene expression, and genetic markers used in methods for improving dairy cattle fitness and/or productivity. More specifically, it relates to the use of genetic markers in methods for improving dairy cattle, including improvements with respect to Milk production (MP), Somatic Cell Score (SCS), Daughter Pregnancy Rate (DPR), Productive Life (PL), and Net Merit (NM).

BACKGROUND OF THE INVENTION

[0004] The future viability and competitiveness of the dairy industry depends on continual improvement in productivity and fitness traits. Examples of key traits include milk productivity (e.g. milk, fat, protein yield, fat%, protein % and

persistence of lactation), health (e.g. Somatic Cell Count, mastitis incidence), fertility (e.g. pregnancy rate, display of estrus, calving interval and non-return rates in bulls), calving ease (e.g. direct and maternal calving ease), longevity (e.g. productive life), and functional conformation (e.g. udder support, proper foot and leg shape, proper rump angle, etc.). Unfortunately efficiency traits are often unfavorably correlated with fitness traits. Although fitness traits all have some degree of underlying genetic variation in commercial cattle populations, the accuracy of selecting breeding animals with superior genetic merit for many of them is low due to low heritability or the inability to measure the trait cost effectively on the candidate animal. In addition, many productivity and fitness traits can only be measured on females. Thus, the accuracy of conventional selection for these traits is moderate to low and ability to make genetic change through selection is limited, particularly for fitness traits.

[0005] Genomics offers the potential for greater improvement in productivity and fitness traits through the discovery of genes, or genetic markers linked to genes, that account for genetic variation and can be used for more direct and accurate selection. Close to 1000 markers with associations with productivity and fitness traits have been reported (see www.bovineqtl.tamu.edu/ for a searchable database of reported QTL), however, the resolution of QTL location is still quite low which makes it difficult to utilize these QTL in marker-assisted selection (MAS) on an industrial scale. Only a few QTL have been fully characterized with a strong putative or well-confirmed causal mutation: *DGATI* on chromosome 14 (Grisard et al., 2002; Winter et al, 2002; Kuhn et al., 2004) *GHR* on chromosome 20 (Blott et al., 2003), *ABCG2* (Cohen-Zinder et al., 2005) or *SPP1* on chromosome 6 (Schnabel et al., 2005). However, these discoveries are rare and only explain a small portion of the genetic variance for productivity traits and no genes controlling fitness traits have been fully characterized. A more successful strategy employs the use of whole-genome high-density scans of the entire bovine genome in which QTL are mapped with sufficient resolution to explain the majority of genetic variation around the traits of interest.

[0006] Cattle herds used for milk production around the world originate predominantly from the Holstein or Holstein-Friesian breeds which are known for high levels of production. However, the high production levels in Holsteins have also been linked to greater calving difficulty and reduced levels of fertility. It is unclear

whether these unfavorable correlations are due to pleiotropic gene effects or simply due to linked genes. If the latter is true, with marker knowledge, it may be possible to select for favorable recombinants that contain the favorable alleles from several linked genes that are normally at frequencies too low to allow much progress with traditional selection. Since Holstein germplasm has been sold and transported globally for several decades, the Holstein breed has effectively become one large global population held to relatively moderate inbreeding rates. Also, the outbred nature of such a large population selected for several generations has allowed linkage disequilibrium to break down except within relatively short distances (i.e. less than a few centimorgans) (Hayes et al., 2006). Given this pattern of linkage disequilibrium, dense marker coverage is required to refine QTL locations with sufficient precision to find markers that are in tight linkage disequilibrium with them. Therefore, markers that are in tight linkage disequilibrium with the QTL are essential for effective population-wide MAS or whole-genome selection (WGS).

[0007] The presence of linkage disequilibrium between very closely linked markers provides one skilled in the art an opportunity to substitute one marker for the other in the context of a marker-assisted breeding/management/allocation program. Linkage disequilibrium can be defined as the observance of alleles at two distinctive loci occurring in gametes more frequently than expected given the known allele frequencies and recombination fraction between the two loci (Source: NHBLI/NCBI Glossary). Given the knowledge that allele 1 at marker A occurs in gametes disproportionately with allele 1 at marker B confers the equivalent practical and economic value of marker B upon marker A, such that they can be considered as equivalent.

[0008] It is well documented that linkage disequilibrium between two markers increases as physical distance decreases. As such, there is value in identifying markers that are located physically close to markers that have proven value when used in marker-assisted breeding/management/allocation programs. Because the nearby marker will likely display linkage disequilibrium with the marker of proven value, the nearby marker can be used as a proxy as it will show a similar association with the phenotypic trait. Thus the proven economic value of any given marker can be replicated by identifying and substituting very closely linked (nearby) markers.

[0009] Furthermore, it is well documented that DNA sequences within (intronic, 5' UTR, 3' UTR) and outside of (promoter sequence motifs, enhancer sequence motifs) the formal boundaries of a gene can significantly influence the expression of the gene. For example, McDowell and Dean (1999) showed that an enhancer element more than 2 kb from the translational start codon directly influenced transcription of the gene. More recently, Markstein et al (2004) described the effects of an enhancer element residing 15 kb from the translational start site. Given these observations, it is clear that sequence variation at locations outside of the formal boundaries of a gene (typically from the translational start codon to the translational termination codon), can alter the expression patterns of a gene.

[0010] Most traits are quantitative in nature and hence are governed by a large number of QTL of small to moderately sized effects. Therefore, to characterize enough QTL to explain a majority of genetic variation for these traits, the whole genome needs to be scanned with a set of markers mapped to the genome at high resolution (i.e. greater than 1 marker/cM); otherwise known as whole-genome analysis. Furthermore, a sufficient number of marked QTL must be used in MAS in order to accurately predict the breeding value of an animal without phenotyping records on relatives or the animal itself. The application of such a high-density whole-genome marker map to discover and finely-map QTL explaining variation in productivity and fitness traits is described herein. The large number of resulting linked markers can be used in several methods of marker selection or marker-assisted selection, including whole-genome selection (WGS) (Meuwissen et al., Genetics 2001) to improve the genetic merit of the population for these traits and create value in the dairy industry.

SUMMARY OF THE INVENTION

[0011] This section provides a non-exhaustive summary of the present invention.

[0012] Various embodiments of the invention provide methods for evaluating an animal's genotype at 10 or more positions in the animal's genome. In various aspects of these embodiments the animal's genotype is evaluated at positions within a segment of DNA (an allele) that contains at least one SNP selected from the SNPs described in the Tables and Sequence Listing of the present application.

[0013] Other embodiments of the invention provide methods for allocating animals for use according to their predicted marker breeding value for productivity and/or fitness. Various aspects of this embodiment of the invention provide methods that comprise: a) analyzing the animal's genomic sequence at one or more polymorphisms (where the alleles analyzed each comprise at least one SNP) to determine the animal's genotype at each of those polymorphisms; b) analyzing the genotype determined for each polymorphisms to determine which allele of the SNP is present; c) allocating the animal for use based on its genotype at one or more of the polymorphisms analyzed. Various aspects of this embodiment of the invention provide methods for allocating animals for use based on a favorable association between the animal's genotype, at one or more polymorphisms disclosed in the present application, and a desired phenotype. Alternatively, the methods provide for not allocating an animal for a certain use because it has one or more SNP alleles that are either associated with undesirable phenotypes or are not associated with desirable phenotypes.

[0014] Other embodiments of the invention provide methods for allocating animals for use according to their predicted marker breeding value for Milk production, Somatic Cell Score, Daughter Pregnancy Rate, Productive Life, and/or Net Merit. Various aspects of this embodiment of the invention provide methods that comprise: a) analyzing the animal's genomic sequence at one or more polymorphisms (where the alleles analyzed each comprise at least one SNP) to determine the animal's genotype at each of those polymorphisms; b) analyzing the genotype determined for each polymorphisms to determine which allele of the SNP is present; c) allocating the animal for use based on its genotype at one or more of the polymorphisms analyzed. Various aspects of this embodiment of the invention provide methods for allocating animals for use based on a favorable association between the animal's genotype, at one or more polymorphisms disclosed in the present application, and a desired phenotype. Alternatively, the methods provide for not allocating an animal for a certain use because it has one or more SNP alleles that are either associated with undesirable phenotypes or are not associated with desirable phenotypes.

[0015] Other embodiments of the invention provide methods for selecting animals for use in breeding to produce progeny. Various aspects of these methods comprise: A) determining the genotype of at least one potential parent animal at one or more

locus/loci, where at least one of the loci analyzed contains an allele of a SNP selected from the group of SNPs described in Table 3, 4, 5 and 6 and the Sequence Listing. B) Analyzing the determined genotype at one or more positions for at least one animal to determine which of the SNP alleles is present. C) Correlating the analyzed allele(s) with one or more phenotypes. D) Allocating at least one animal for use to produce progeny.

[0016] Other embodiments of the invention provide methods for selecting animals for use in breeding to produce progeny. Various aspects of these methods comprise: A) determining the genotype of at least one potential parent animal at one or more locus/loci, where at least one of the loci analyzed contains an allele of a SNP selected from the group of SNPs described in Table 4 and the Sequence Listing and/or other SNPs within one or more genes described in Table 4. B) Analyzing the determined genotype at one or more positions for at least one animal to determine which of the SNP alleles is present. C) Correlating the analyzed allele(s) with one or more phenotypes. D) Allocating at least one animal for use to produce progeny.

[0017] Other embodiments of the invention provide methods for producing offspring animals (progeny animals). Aspects of this embodiment of the invention provide methods that comprise: breeding an animal that has been selected for breeding by methods described herein to produce offspring. The offspring may be produced by purely natural methods or through the use of any appropriate technical means, including but not limited to: artificial insemination; embryo transfer (ET), multiple ovulation embryo transfer (MOET), *in vitro* fertilization (IVF), or any combination thereof.

[0018] Other embodiments of the invention provide for databases or groups of databases, each database comprising lists of the nucleic acid sequences, which lists include a plurality of the SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing. Preferred aspects of this embodiment of the invention provide for databases comprising the sequences for 50 or more SNPs. Other aspects of these embodiments comprise methods for using a computer algorithm or algorithms that use one or more database(s), each database comprising a plurality of the SNPs described in Table 3, 5 and 6 and the Sequence Listing to identify phenotypic traits associated with the

inheritance of one or more alleles of the SNPs, and/or using such a database to aid in animal allocation.

[0019] Other embodiments of the invention provide for databases or groups of databases, each database comprising lists of the nucleic acid sequences, which lists include a plurality of the SNPs described in Table 4 and the Sequence Listing and/or other SNPs located with the genes listed in Table 4. Preferred aspects of this embodiment of the invention provide for databases comprising the sequences for 50 or more SNPs. Other aspects of these embodiments comprise methods for using a computer algorithm or algorithms that use one or more database(s), each database comprising a plurality of the SNPs described in Table 4 and the Sequence Listing to identify phenotypic traits associated with the inheritance of one or more alleles of the SNPs, and/or using such a database to aid in animal allocation.

[0020] The expression of genes identified in this application are likely responsible (either through quantitative or qualitative variation of the protein) for a significant proportion of the observed genetic variation for the trait. As such, other embodiments of the invention may include modulating the presence of the protein products of genes described in Table 4 in the animal through various established methods, which would therefore likely modulate the phenotype of the animal in a predictable fashion.

[0021] Additional embodiments of the invention provide methods for identifying other genetic markers that are in allelic association with one or more of the SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing.

DEFINITIONS

[0022] The following definitions are provided to aid those skilled in the art to more readily understand and appreciate the full scope of the present invention. Nevertheless, as indicated in the definitions provided below, the definitions provided are not intended to be exclusive, unless so indicated. Rather, they are preferred definitions, provided to focus the skilled artisan on various illustrative embodiments of the invention.

[0023] As used herein the term "allelic association" preferably means: nonrandom deviation of $f(A_i B_j)$ from the product of $f(A_i)$ and $f(B_j)$, which is specifically defined

by $r^2 > 0.2$, where r^2 is measured from a reasonably large animal sample (e.g., ≥ 100) and defined as

$$r^2 = \frac{[f(A_1B_1) - f(A_1)f(B_1)]^2}{f(A_1)(1-f(A_1))(f(B_1)(1-f(B_1)))} \quad \text{[Equation 1]}$$

where A_1 represents an allele at one locus, B_1 represents an allele at another locus; $f(A_1B_1)$ denotes frequency of gametes having both A_1 and B_1 , $f(A_1)$ is the frequency of A_1 , $f(B_1)$ is the frequency of B_1 in a population.

[0024] As used herein the terms “allocating animals for use” and “allocation for use” preferably mean deciding how an animal will be used within a herd or that it will be removed from the herd to achieve desired herd management goals. For example, an animal might be allocated for use as a breeding animal or allocated for sale as a non-breeding animal (e.g. allocated to animals intended to be sold for meat). In certain aspects of the invention, animals may be allocated for use in sub-groups within the breeding programs that have very specific goals (e.g. productivity or fitness). Accordingly, even within the group of animals allocated for breeding purposes, there may be more specific allocation for use to achieve more specific and/or specialized breeding goals.

[0025] As used here in the term “anchor SNP” and “anchor Marker” preferably refer to a SNP located in a region determined to be in genetic association/linkage with one or more traits. In Table 3 anchor markers are identified as those SNP/Markers that have a SEQ ID NO: identical to region number of the region containing them. For example, SEQ ID NO:1 provides the nucleic acid sequence for the anchor marker for region number 1. In Table 4, anchor markers are identified as those SNP/Markers that have a SEQ ID NO listed in column 1. For example, SEQ ID NO:173931 provides the nucleic acid sequence for the anchor marker for the first region, corresponding nearby genes are listed in column 2, and SEQ ID numbers of markers within the gene are listed in column 3.

[0026] As used herein the terms “animal” or “animals” preferably refer to dairy cattle.

[0027] As used herein “fitness” preferably refers to traits that include, but are not limited to: pregnancy rate (PR), daughter pregnancy rate (DPR), productive life (PL), somatic cell count (SCC) and somatic cell score (SCS).

[0028] As used herein, PR and DPR refer to the percentage of non-pregnant animals that become pregnant during each 21-day period.

[0029] As used herein, PL is calculated as months in milk in each lactation, summed across all lactations until removal of the cow from the herd (by culling or death).

[0030] As used herein, somatic cell score can be calculated using the following relationship: $SCS = \log_2(SCC/100,000)+3$, where SCC is somatic cells per milliliter of milk.

[0031] As used herein the term “growth” refers to the measurement of various parameters associated with an increase in an animal’s size and/or weight.

[0032] As used herein the term “linkage disequilibrium” preferably means allelic association wherein A_1 and B_1 (as used in the above definition of allelic association) are present on the same chromosome.

[0033] As used herein the term “marker-assisted selection (MAS) preferably refers to the selection of animals on the basis of marker information in possible combination with pedigree and phenotypic data.

[0034] As used herein the terms “marker breeding value (MBV)” and “predicted marker breeding value (PMBV)” refer to an estimate of an animal's genetic transmitting ability with respect to specific traits and is based on its genotype.

[0035] As used herein the term “natural breeding” preferably refers to mating animals without human intervention in the fertilization process. That is, without the use of mechanical or technical methods such as artificial insemination or embryo transfer. The term does not refer to selection of the parent animals.

[0036] As used herein the terms “neighboring markers” or “neighboring SNPs” preferably refer to SNPs in close proximity to an anchor SNP, most preferably these terms refer to markers/SNPs located within about 70 kilobases of the anchor SNP.

[0037] As used herein the term “net merit” preferably refers to a composite index that includes several commonly measured traits weighted according to relative economic value in a typical production setting and expressed as lifetime economic worth per

cow relative to an industry base. Examples of a net merit indexes include, but are not limited to, \$NM or TPI in the USA, LPI in Canada, etc (formulae for calculating these indices are well known in the art (e.g. \$NM can be found on the USDA/AIPL website: www.aipl.arsusda.gov/reference.htm)

[0038] As used herein, the term “milk production” preferably refers to phenotypic traits related to the productivity of a dairy animal including milk fluid volume, fat percent, protein percent, fat yield, and protein yield.

[0039] As used herein the term “predicted value” preferably refers to an estimate of an animal’s breeding value or transmitting ability based on its genotype and pedigree.

[0040] As used herein “productivity” and “production” preferably refers to yield traits that include, but are not limited to: total milk yield, milk fat percentage, milk fat yield, milk protein percentage, milk protein yield, total lifetime production, milking speed and lactation persistency.

[0041] As used herein the term “quantitative trait” is used to denote a trait that is controlled by multiple (two or more, and often many) genes each of which contributes small to moderate effect on the trait. The observations on quantitative traits often follow a normal distribution.

[0042] As used herein the term “quantitative trait locus (QTL)” is used to describe a locus that contains polymorphism that has an effect on a quantitative trait.

[0043] As used herein the term “reproductive material” includes, but is not limited to semen, spermatozoa, ova, and zygote(s).

[0044] As used herein the term “single nucleotide polymorphism” or “SNP” refer to a location in an animal’s genome that is polymorphic within the population. That is, within the population some individual animals have one type of base at that position, while others have a different base. For example, a SNP might refer to a location in the genome where some animals have a “G” in their DNA sequence, while others have a “T”.

[0045] As used herein the terms “hybridization under stringent conditions” and “stringent hybridization conditions” preferably mean conditions under which a “probe” will hybridize to its target sequence to a detectably greater degree than to other sequences (e.g., at least 5-fold over background). Stringent conditions are target-sequence-dependent and will differ depending on the structure of the

polynucleotide. By controlling the stringency of the hybridization and/or washing conditions, target sequences that are 100% complementary to the probe can be identified (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing).

[0046] Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30 °C for short probes (e.g., 10 to 50 nucleotides) and at least about 60 °C for long probes (e.g., greater than 50 nucleotides). Stringency may also be adjusted with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37 °C, and a wash in 1X to 2X SSC (20X SSC=3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55 °C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37 °C, and a wash in 0.5X to 1X SSC at 55 to 60 °C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65 °C. The duration of hybridization is generally less than about 24 hours, usually about 4 to about 12 hours.

[0047] Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the thermal melting point (T_m) can be approximated from the equation of Meinkoth and Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m=81.5^\circ\text{C}+16.6 (\log M)+0.41 (\% \text{ GC})-0.61 (\% \text{ form})-500/L$; where M is the molarity of monovalent cations, % GC is the percentage of guanine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. T_m is reduced by about 1°C for each 1% of mismatching; thus, T_m , hybridization, and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with 90% identity are sought, the T_m can be decreased 10°C. Generally, stringent conditions are selected to be about 5°C lower than the T_m for the specific sequence and its complement at a defined ionic strength

and pH. However, highly stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T_m , those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45°C (aqueous solution) or 32°C (formamide solution), it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 (Elsevier, N.Y.); and Ausubel et al., eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See also Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.).

[0048] As used herein the terms “marker breeding value (MBV)” and “predicted marker breeding value (PMBV)” respectively refer to an estimate of an animal’s genetic transmitting ability with respect to either productivity traits or fitness traits that is based on its genotype.

[0049] As used herein the term “whole-genome analysis” preferably refers to the process of QTL mapping of the entire genome at high marker density (i.e. at least about one marker per cM) and detection of markers that are in population-wide linkage disequilibrium with QTL.

[0050] As used herein the term “whole-genome selection (WGS)” preferably refers to the process of marker-assisted selection (MAS) on a genome-wide basis in which markers spanning the entire genome at moderate to high density (e.g. at least about one marker per 1-5 cM), or at moderate to high density in QTL regions, or directly neighboring or flanking QTL that explain a significant portion of the genetic variation controlling one or more traits.

ILLUSTRATIVE EMBODIMENTS OF THE INVENTION

[0051] Various embodiments of the present invention provide methods for evaluating an animal's (especially a dairy animal's) genotype. In preferred embodiments of the invention, the animal's genotype is evaluated at 10 or more positions (i.e. with respect to 10 or more genetic markers). Aspects of these embodiments of the invention provide methods that comprise determining the animal's genomic sequence at 10 or more locations (loci) that contain single nucleotide polymorphisms (SNPs). Embodiments of the invention provide methods for evaluating an animal's genotype by determining which of two or more alleles for the SNP are present for each of 10 or more SNPs selected from the group consisting of the SNPs described in Tables 5 and 6 of the instant application. Furthermore, embodiments of the invention provide methods for evaluating an animal's genotype by determining which of two or more alleles for the SNP are present for each of 10 or more SNPs selected from the group consisting of the SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing of the instant application and/or SNPs located on the same chromosome and within approximately 70 kilobases of one or more of the SNP anchor markers described in Tables 3 and 4 and the Sequence Listing. That is, within approximately 70,000 base pairs of the anchor SNP marker, in either the 5' or 3' direction from an anchor SNP.

[0052] Various embodiments of the present invention provide methods for evaluating an animal's (especially a dairy animal's) genotype. In preferred embodiments of the invention, the animal's genotype is evaluated at 10 or more positions (i.e. with respect to 10 or more genetic markers). Aspects of these embodiments of the invention provide methods that comprise determining the animal's genomic sequence at 10 or more locations (loci) that contain single nucleotide polymorphisms (SNPs). Specifically, the invention provides methods for evaluating an animal's genotype by determining which of two or more alleles for the SNP are present for each of 10 or more SNPs selected from the group consisting of the SNPs described in Table 4 and the Sequence Listing of the instant application and/or SNPs located on the same chromosome and within a gene, wherein the gene has a portion of its sequence within approximately 70 kilobases of one or more of the SNP anchor markers described in Table 4 and the Sequence Listing. That is, within a gene residing within approximately 70,000 base pairs of the anchor SNP marker, in either the 5' or 3' direction from the anchor SNP.

[0053] In preferred aspects of these embodiments the animal's genotype is evaluated to determine which allele is present for 25 or more SNPs selected from the group consisting of the SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing and SNPs on the same chromosome and within about 70 kilobases of one or more of the SNP anchor markers, described in Tables 3, 4, 5, and 6 and the Sequence Listing. More, preferably the animal's genotype is determined for positions corresponding with 50, 100, 200, 500, or 1000, or more of the SNPs selected from this group.

[0054] In preferred aspects of these embodiments the animal's genotype is evaluated to determine which allele is present for 25 or more SNPs selected from the group consisting of the SNPs described in Table 4 and the Sequence Listing and SNPs on the same chromosome and within a gene, wherein the gene has a portion of its sequence within about 70 kilobases of one or more of the SNP anchor markers, described in Table 4 and the Sequence Listing. More, preferably the animal's genotype is determined for positions corresponding with 50, 100, 200, 500, or 1000, or more of the SNPs selected from this group.

[0055] In preferred aspects of these embodiments the animal's genotype is evaluated to determine which allele is present for 25 or more SNPs selected from the group of SNPs described in Tables 5 and 6. More, preferably the animal's genotype is determined for positions corresponding with 50, 100, 200, 500, or 1000, or more of the SNPs described in Tables 5 and 6.

[0056] In other aspects of this embodiment, the animal's genotype is analyzed with respect to at least 10, 25, 50, 100, 200, 500, or more SNPs that have been shown to be associated with productivity and/or fitness (*see* Table 5 for a list of the SNPs associated with these traits). For example, embodiments of the invention provides a method for genotyping 10 or more, 25 or more, 50 or more, 100 or more, 200 or more, or 500 or more, or 1000 or more SNPs that have been determined to be significantly associated with productivity, as described in Tables 5 and 6.

[0057] In other aspects of this embodiment, the animal's genotype is analyzed with respect to at least 10, 25, 50, 100, 200, 500, or more SNPs that have been shown to be

associated with one or more traits selected from the group consisting of Milk Production, Somatic Cell Score, Daughter Pregnancy Rate, Productive Life, and Net Merit (*see* Tables 3 and 4 for a list of the SEQ ID NOs of the SNPs associated with these traits, including associated anchor SNPs). For example, embodiments of the invention provides a method for genotyping 10 or more, 25 or more, 50 or more, 100 or more, 200 or more, or 500 or more, or 1000 or more SNPs that have been determined to be significantly associated with one or more of these traits. These SNPs are preferably selected from the group consisting of the SNPs described in Table 3 and 4 and the Sequence Listing and the respective neighboring SNPs.

[0058] Aspects of the present invention also provides for both whole-genome analysis and whole genome-selection (WGS) (that is marker-assisted selection (MAS) on a genome-wide basis). Various aspects of this embodiment of the invention provide for either whole-genome analysis or WGS wherein the makers analyzed for an animal span the animal's entire genome at moderate to high density. That is, the animal's genome is analyzed with markers that on average occur, at least, approximately every 1 to 5 centimorgans in the genome. Moreover the invention provides that of the markers used to carry out the whole-genome analysis or WGS, 10 or more, 25, or more, 50 or more, 100 or more, 200 or more, 500 or more, or 1000 or more are selected from the markers described in Tables 5 and 6. In preferred aspects of this embodiment the markers may be associated with fitness or associated with productivity, or may be associated with both fitness and productivity.

[0059] Aspects of the present invention also provides for both whole-genome analysis and whole genome-selection (WGS) (i.e. marker-assisted selection (MAS) on a genome-wide basis). Various aspects of this embodiment of the invention provide for either whole-genome analysis or WGS wherein the makers analyzed for an animal span the animal's entire genome at moderate to high density. That is, the animal's genome is analyzed with markers that on average occur, at least, approximately every 1 to 5 centimorgans in the genome. Moreover the invention provides that of the markers used to carry out the whole-genome analysis or WGS, 10 or more, 25, or more, 50 or more, 100 or more, 200 or more, 500 or more, or 1000 or more are selected from the group consisting of the markers described in Table 3 and 4 and the Sequence Listing and SNPs on the same chromosome and within about 70 kilobases

of one or more of the SNP anchor markers in Table 3. In preferred aspects of this embodiment the markers may be associated with fitness or associated with productivity, or may be associated with both fitness and productivity. Furthermore, the invention provides that of the markers used to carry out the whole-genome analysis or WGS, 10 or more, 25, or more, 50 or more, 100 or more, 200 or more, 500 or more, or 1000 or more are selected from the group consisting of the markers described in Tables 3 and 4 and the Sequence Listing and SNPs on the same chromosome and within a gene, wherein the gene has a portion of its sequence within about 70 kilobases of one or more of the SNP anchor markers in Table 4. In preferred aspects of this embodiment the markers may be associated with fitness or associated with productivity, or may be associated with both fitness and productivity.

[0060] In any embodiment of the invention the genomic sequence at the SNP locus may be determined by any means compatible with the present invention. Suitable means are well known to those skilled in the art and include, but are not limited to direct sequencing, sequencing by synthesis, primer extension, Matrix Assisted Laser Desorption /Ionization-Time Of Flight (MALDI-TOF) mass spectrometry, polymerase chain reaction-restriction fragment length polymorphism, microarray/multiplex array systems (e.g. those available from Affymetrix, Santa Clara, California), and allele-specific hybridization.

[0061] Other embodiments of the invention provide methods for allocating animals for subsequent use (e.g. to be used as sires or dams or to be sold for meat or dairy purposes) according to their predicted value for productivity or fitness. Various aspects of this embodiment of the invention comprise determining at least one animal's genotype for at least one SNP selected from the group of SNPs consisting of the SNPs described in Table 3 and the sequence listing and SNPs on the same chromosome and within about 70 kilobases of one or more of the SNP anchor markers, (methods for determining animals' genotypes for one or more SNPs are described *supra*). Thus, the animal's allocation for use may be determined based on its genotype at one or more, 5 or more, 10 or more, 25 or more, 50 or more, 100 or more, 300 or more, or 500 or more, or 1000 or more of the SNPs in this group. The instant invention also provides embodiments where analysis of the genotypes of the SNPs described in Tables 3 and the Sequence Listing is the only analysis done. Other

embodiments provide methods where analysis of the SNPs disclosed herein is combined with any other desired type of genomic or phenotypic analysis (e.g. analysis of any genetic markers beyond those disclosed in the instant invention). Moreover, the SNPs analyzed may be selected from those SNPs associated only with milk production (MP), only with somatic cell score (SCS), only with daughter pregnancy rate (DPR), only with reproductive life (PL), or only with net merit (NM). Alternatively, the analysis may be done for SNPs selected from any desired combination of MP, SCS, DPR, PL, and NM. SNPs associated with various traits include those listed in Tables 3 and others that are located on the same chromosome and within about 70 kb of an anchor marker.

[0062] Other embodiments of the invention provide methods for allocating animals for subsequent use (e.g. to be used as sires or dams or to be sold for meat or dairy purposes) according to their predicted value for productivity or fitness. Various aspects of this embodiment of the invention comprise determining at least one animal's genotype for at least one SNP selected from the group of SNPs consisting of the SNPs described in Table 4 and the sequence listing and SNPs on the same chromosome and within a gene, wherein the gene has a portion of its sequence within about 70 kilobases of one or more of the SNP anchor markers, (methods for determining animals' genotypes for one or more SNPs are described *supra*). Thus, the animal's allocation for use may be determined based on its genotype at one or more, 5 or more, 10 or more, 25 or more, 50 or more, 100 or more, 300 or more, or 500 or more, or 1000 or more of the SNPs in this group. The instant invention also provides embodiments where analysis of the genotypes of the SNPs described in Table 4 and the Sequence Listing is the only analysis done. Other embodiments provide methods where analysis of the SNPs disclosed herein is combined with any other desired type of genomic or phenotypic analysis (e.g. analysis of any genetic markers beyond those disclosed in the instant invention). Moreover, the SNPs analyzed may be selected from those SNPs associated only with milk production (MP), only with somatic cell score (SCS), only with daughter pregnancy rate (DPR), only with reproductive life (PL), or only with net merit (NM). Alternatively, the analysis may be done for SNPs selected from any desired combination of MP, SCS, DPR, PL, and NM. SNPs associated with various traits include those SNPs listed in Table 4 and others that are located on the same chromosome and within a gene,

wherein the gene has a portion of its sequence within about 70 kb of an anchor marker.

[0063] According to various aspects of these embodiments of the invention, once the animal's genetic sequence for the selected SNP(s) have been determined, this information is evaluated to determine which allele of the SNP is present for at least one of the selected SNPs. Preferably the animal's allelic complement for all of the determined SNPs is evaluated. Finally, the animal is allocated for use based on its genotype for one or more of the SNP positions evaluated. Preferably, the allocation is made taking into account the animal's genotype at each of the SNPs evaluated, but its allocation may be based on any subset or subsets of the SNPs evaluated.

[0064] The allocation may be made based on any suitable criteria. For any SNP, a determination may be made as to whether one of the allele(s) is associated/correlated with desirable characteristics or associated with undesirable characteristics. This determination will often depend on breeding or herd management goals. Determination of which alleles are associated with desirable phenotypic characteristics can be made by any suitable means. Methods for determining these associations are well known in the art; moreover, aspects of the use of these methods are generally described in the EXAMPLES, below.

[0065] Phenotypic traits that may be associated with the SNPs of the current invention include, but are not limited to; fitness traits and productivity traits (including for example, MP, SCS, DPR, PL, and NM).

[0066] According to various aspects of this embodiment of the invention allocation for use of the animal may entail either positive selection for the animals having the desired genotype(s) (e.g. the animals with the desired genotypes are selected for productivity traits), negative selection of animals having undesirable genotypes (e.g. animals with an undesirable genotypes are culled from the herd), or any combination of these methods. According to preferred aspects of this embodiment of the invention animals identified as having SNP alleles associated with desirable phenotypes are allocated for use consistent with that phenotype (e.g. allocated for breeding based on phenotypes positively associated with fitness). Alternatively, animals that do not have SNP genotypes that are positively correlated with the desired phenotype (or

possess SNP alleles that are negatively correlated with that phenotype) are not allocated for the same use as those with a positive correlation for the trait.

[0067] Other embodiments of the invention provide methods for selecting potential parent animals (*i.e.*, allocation for breeding) to improve fitness and/or productivity in potential offspring. Various aspects of this embodiment of the invention comprise determining at least one animal's genotype for at least one SNP selected from the group of SNPs consisting of the SNPs described in Table 3, 5 and 6 and the Sequence Listing and SNPs on the same chromosome and within about 70 kilobases of one or more of the SNP anchor markers. Furthermore, determination of whether and how an animal will be used as a potential parent animal may be based on its genotype at one or more, 10 or more, 25 or more, 50 or more, 100 or more, 300 or more, or 500 or more of the SNPs from that group. Moreover, as with other types of allocation for use, various aspects of these embodiments of the invention provide methods where the only analysis done is to genotype the animal for one or more of the SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing. Other aspects of these embodiments provide methods where analysis of one or more SNPs disclosed herein is combined with any other desired genomic or phenotypic analysis (e.g. analysis of any genetic markers beyond those disclosed in the instant invention). Moreover, the SNP(s) analyzed may all be selected from those associated only with MP, only with SCS, only with DPR, only with PL, or only with NM. Conversely, the analysis may be done for SNPs selected from any desired combination of these or other traits.

[0068] Other embodiments of the invention provide methods for selecting potential parent animals (*i.e.*, allocation for breeding) to improve fitness and/or productivity in potential offspring. Various aspects of this embodiment of the invention comprise determining at least one animal's genotype for at least one SNP selected from the group of SNPs consisting of the SNPs described in Table 4 and the Sequence Listing and SNPs on the same chromosome and within a gene, wherein the gene has a portion of its sequence within about 70 kilobases of one or more of the SNP anchor markers. Furthermore, determination of whether and how an animal will be used as a potential parent animal may be based on its genotype at one or more, 10 or more, 25 or more, 50 or more, 100 or more, 300 or more, or 500 or more of the SNPs from that group. Moreover, as with other types of allocation for use, various aspects of these embodiments of the invention provide methods where the only analysis done is to

genotype the animal for one or more of the SNPs described in Table 4 and the Sequence Listing. Other aspects of these embodiments provide methods where analysis of one or more SNPs disclosed herein is combined with any other desired genomic or phenotypic analysis (e.g. analysis of any genetic markers beyond those disclosed in the instant invention). Moreover, the SNP(s) analyzed may all be selected from those associated only with MP, only with SCS, only with DPR, only with PL, or only with NM. Conversely, the analysis may be done for SNPs selected from any desired combination of these or other traits.

[0069] According to various aspects of these embodiments of the invention, once the animal's genetic sequence at the site of the selected SNP(s) have been determined, this information is evaluated to determine which allele of the SNP is present for at least one of the selected SNPs. Preferably the animal's allelic complement for all of the sequenced SNPs is evaluated. Additionally, the animal's allelic complement is analyzed and evaluated to estimate the probability that the animal's progeny will express one or more phenotypic traits or to predict the animal's progeny's genetic merit or phenotypic value of traits of interest. Finally, the animal is allocated for breeding use based on its genotype for one or more of the SNP positions evaluated and the probability that it will pass the desired genotype(s)/allele(s) to its progeny. Preferably, the breeding allocation is made taking into account the animal's genotype at each of the SNPs evaluated. However, its breeding allocation may be based on any subset or subsets of the SNPs evaluated.

[0070] The breeding allocation may be made based on any suitable criteria. For example, breeding allocation may be made so as to increase the probability of enhancing a single certain desirable characteristic in a population, in preference to other characteristics, (e.g. increased fitness, or even specifically lowering somatic cell score (SCS) as part of fitness); alternatively, the selection may be made so as to generally maximize overall production based on a combination of traits. The allocations chosen are dependent on the breeding goals. Sub-categories falling within fitness, include, *inter alia*: daughter pregnancy rate (DPR), productive life (PL), and somatic cell score. Sub-categories falling within productivity include, *inter alia*: milk fat percentage, milk fat yield, total milk yield, milk protein percentage, and total milk protein.

[0071] Other embodiments of the instant invention provide methods for producing progeny animals. According to various aspects of this embodiment of the invention, the animals used to produce the progeny are those that have been allocated for breeding according to any of the embodiments of the current invention. Those using the animals to produce progeny may perform the necessary analysis or, alternatively, those producing the progeny may obtain animals that have been analyzed by another. The progeny may be produced by any appropriate means, including, but not limited to using: (i) natural breeding, (ii) artificial insemination, (iii) *in vitro* fertilization (IVF) or (iv) collecting semen/spermatozoa and/or at least one ovum from the animal and contacting it, respectively with ova/ovum or semen/spermatozoa from a second animal to produce a conceptus by any means.

[0072] According to preferred aspects of this embodiment of the invention the progeny are produced by a process comprising natural breeding. In other aspects of this embodiment the progeny are produced through a process comprising the use of standard artificial insemination (AI), *in vitro* fertilization, multiple ovulation embryo transfer (MOET), or any combination thereof.

[0073] Other embodiments of the invention provide for methods that comprise allocating an animal for breeding purposes and collecting/isolating genetic material from that animal: wherein genetic material includes but is not limited to: semen, spermatozoa, ovum, zygotes, blood, tissue, serum, DNA, and RNA.

[0074] It is understood that most efficient and effective use of the methods and information provided by the instant invention employ computer programs and/or electronically accessible databases that comprise all or a portion of the sequences disclosed in the instant application. Accordingly, the various embodiments of the instant invention provide for databases comprising all or a portion of the sequences corresponding to at least 10 SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing. In preferred aspect of these embodiments the databases comprise sequences for 25 or more, 50 or more, 100 or more, 200 or more, 500 or more, or 1000 or more, or substantially all of the SNPs described in Table 3, 5 and 6 and the Sequence Listing.

[0075] It is further understood that efficient analysis and use of the methods and information provided by the instant invention will employ the use of automated

genotyping; particularly when large numbers (e.g. 100s) of markers are evaluated. Any suitable method known in the art may be used to perform such genotyping, including, but not limited to the use of micro-arrays.

[0076] Other embodiments of the invention provide methods wherein one or more of the SNP sequence databases described herein are accessed by one or more computer-executable programs. Such methods include, but are not limited to, use of the databases by programs to analyze for an association between the SNP and a phenotypic trait, or other user-defined trait (e.g. traits measured using one or more metrics such as gene expression levels, protein expression levels, or chemical profiles), and programs used to allocate animals for breeding or market.

[0077] Other embodiments of the invention provide methods comprising collecting genetic material from an animal that has been allocated for breeding. Wherein the animal has been allocated for breeding by any of the methods disclosed as part of the instant invention.

[0078] Other embodiments of the invention provide for diagnostic kits or other diagnostic devices for determining which allele of a SNP is present in a sample; wherein the SNP(s) are selected from the group of SNPs described in Tables 5 and 6. In various aspects of this embodiment of the invention, the kit or device provides reagents/instruments to facilitate a determination as to whether nucleic acid corresponding to the SNP is present. Such kit/or device may further facilitate a determination as to which allele of the SNP is present. In certain aspects of this embodiment of the invention the kit or device comprises at least one nucleic acid oligonucleotide suitable for DNA amplification (e.g. through polymerase chain reaction). In other aspects of the invention the kit or device comprises a purified nucleic acid fragment capable of specifically hybridizing, under stringent conditions, with at least one allele of at least one SNP described in Tables 5 and 6.

[0079] Other embodiments of the invention provide for diagnostic kits or other diagnostic devices for determining which allele of one or more SNP(s) is/are present in a sample; wherein the SNP(s) are selected from the group of SNPs consisting of the SNPs described in Table 3 and the sequence listing and SNPs on the same chromosome and within about 70 kilobases of one or more of the SNP anchor markers. In various aspects of this embodiment of the invention, the kit or device

provides reagents/instruments to facilitate a determination as to whether nucleic acid corresponding to the SNP is present. Such kit/or device may further facilitate a determination as to which allele of the SNP is present. In certain aspects of this embodiment of the invention the kit or device comprises at least one nucleic acid oligonucleotide suitable for DNA amplification (e.g. through polymerase chain reaction). In other aspects of the invention the kit or device comprises a purified nucleic acid fragment capable of specifically hybridizing, under stringent conditions, with at least one allele of at least one of the SNPs described above.

[0080] In particularly preferred aspects of this embodiment of the invention the kit or device comprises at least one nucleic acid array (e.g. DNA micro-arrays) capable of determining which allele of one or more of the SNPs are present in a sample; where the SNPs are selected from the group of SNPs consisting of the SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing and SNPs on the same chromosome and within about 70 kilobases of one or more of the SNP anchor markers. Preferred aspects of this embodiment of the invention provide DNA micro-arrays capable of simultaneously determining which allele is present in a sample for 10 or more SNPs. Preferably, the DNA micro-array is capable of determining which SNP allele is present in a sample for 25 or more, 50 or more, 100 or more, 200 or more, 500 or more, or 1000 or more SNPs. Methods for making such arrays are known to those skilled in the art and such arrays are commercially available (e.g. from Affymetrix, Santa Clara, California).

[0081] Other embodiments of the instant invention provide methods of identifying genetic markers associated with MP, SCS, DPR, PL, and/or NM (and/or any other fitness or productivity trait) that are in allelic association with one or more of the SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing. According to various aspects of this embodiment of the invention the method comprises: (a) identifying a marker, B_1 , that is suspected of being in allelic association with at least one marker, A_1 , wherein A_1 is selected from the group of SNPs described in Tables 3, 4, 5, and 6 and the Sequence Listing; (b) determining whether A_1 and B_1 are in allelic association; wherein allelic association is determined to exist if $r^2 > 0.2$ for Equation 1 for a population sample of at least 100 animals and wherein Equation 1 is:

$$r^2 = \frac{[f(A_1B_1) - f(A_1)f(B_1)]^2}{f(A_1)(1 - f(A_1))f(B_1)(1 - f(B_1))} \quad \text{[Equation 1]}$$

and wherein for Equation 1 A_1 represents an allele at one locus (e.g., a SNP described in Table 3); B_1 represents a genetic marker at another locus; $f(A_1B_1)$ denotes frequency of having both A_1 and B_1 ; $f(A_1)$ is the frequency of A_1 in the population, $f(B_1)$ is the frequency of B_1 in the population. In preferred aspects of this embodiment of the invention B_1 is a SNP. Even more preferred aspects of this embodiment of the invention provide methods for identifying genetic markers in linkage disequilibrium with one or more SNPs selected from the group of SNPs described in Tables 3, 4, 5, and 6. In other aspects of these embodiments of the invention the value of r^2 is greater than 0.5, greater than 0.7, or greater than 0.9.

[0082] In other aspects of this embodiment of the invention the method comprises finding polymorphisms within about 70 kb of one of the anchor SNPs listed in Table 3 or for finding a polymorphism in a gene, wherein the gene has a portion of its sequence within about 70 kb of one of the anchor SNPs listed in Table 4.

[0083] Additional embodiments of the invention provide for genetic markers for fitness and/or productivity that are in allelic association with one or more of the SNPs described in Tables 3, 4, 5, and 6. Markers provided as part of this embodiment of the invention may be identified by any suitable means known to those of ordinary skill in the art. A marker falls within this embodiment of the invention if it is determined to be in allelic association with one or more of the SNPs described in Tables 3, 4, 5, and 6 as defined by Equation 1, *supra*, where r^2 is greater than 0.2, greater than 0.5, greater than 0.7, or greater than 0.9. In preferred aspects of these embodiments of the invention the markers are in linkage disequilibrium with one or more of the SNPs described in Table 3.

[0084] Additional embodiments of the invention provide for genetic markers for fitness and/or productivity that are in allelic association with one or more of the SNPs described in Tables 5 and 6. Markers provided as part of this embodiment of the invention may be identified by any suitable means known to those of ordinary skill in the art. A marker falls within this embodiment of the invention if it is determined to be in allelic association with one or more of the SNPs described in Tables 5 and 6 as defined by Equation 1, *supra*, where r^2 is greater than 0.2, greater than 0.5, greater than 0.7, or greater than 0.9. In preferred aspects of these embodiments of the invention the markers are in linkage disequilibrium with one or more of the SNPs described in Tables 5 and 6.

[0085] Genetic markers that are in allelic association with any of the SNPs described in the Tables may be identified by any suitable means known to those skilled in the art. For example, a genomic library may be screened using a probe specific for any of the sequences of the SNPs described in the Tables. In this way clones comprising at least a portion of that sequence can be identified and then up to 300 kilobases of 3' and/or 5' flanking chromosomal sequence can be determined. Preferably up to about 70 kilobases of 3' and/or 5' flanking chromosomal sequences are evaluated. By this means, genetic markers in allelic association with the SNPs described in the Tables will be identified.

[0086] Other embodiments of the present invention provide methods for identifying genes that may be associated with phenotypic variation. According to various aspects of these embodiments, the chromosomal location of a SNP associated with a particular phenotypic variation can be determined, by means well known to those skilled in the art. Once the chromosomal location is determined genes suspected to be involved with determination of the phenotype can be analyzed. Such genes may be identified by sequencing adjacent portions of the chromosome or by comparison with analogous section of the human genetic map (or known genetic maps for other species). An early example of the existence of clusters of conserved genes is reviewed in Womack (1987), where genes mapping to the same chromosome in one species were observed to map to the same chromosome in other, closely related, species. As mapping resolution improved, reports of the conservation of gene structure and order within conserved chromosomal regions were published (for example, Grosz et al, 1992). More recently, large scale radiation hybrid mapping and BAC sequence have yielded chromosome-scale comparative mapping predictions between human and bovine genomes (Everts-van der Wind et al., 2005), between human and porcine genomes (Yasue et al., 2006) and among vertebrate genomes (Demars et al., 2006)

[0087] Other embodiments of the invention provide methods for identifying causal mutations that underlie one or more quantitative trait loci (QTL). Various aspects of this embodiment of the invention provide for the identification of SNPs that are in allelic association with one or more of the SNPs described in Table 3. Once these SNPs are identified, it is within the ability of skilled artisans to identify mutations

located proximal to such SNP(s). Further, one skilled in the art can identify genes located proximate to the identified SNP(s) and evaluate these genes to select those likely to contain the causal mutation. Once identified, these genes and the surrounding sequence can be analyzed for the presence of mutations, in order to identify the causal mutation.

[0088] Other embodiments of the invention provide methods for identifying causal mutations that underlie one or more quantitative trait loci (QTL). Various aspects of this embodiment of the invention provide for the identification QTL that are in allelic association with one or more of the SNPs described in Tables 5 and 6. Once these SNPs are identified, it is within the ability of skilled artisans to identify mutations located proximal to such SNP(s). Further, one skilled in the art can identify genes located proximate to the identified SNP(s) and evaluate these genes to select those likely to contain the causal mutation. Once identified, these genes and the surrounding sequence can be analyzed for the presence of mutations, in order to identify the causal mutation.

[0089] Still other embodiments of the invention provide methods to modulate the expression and/or concentration of a gene product from the genes described in Table 4 with the intent of manipulating the performance and/or product quality of an animal or animal product. This can be performed through any of a number of technologies designed to positively or negatively influence gene expression, including but not limited to transgenesis, RNA interference, and anti-sense protocols.

EXAMPLES

[0090] The following examples are included to demonstrate general embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventors to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the invention.

[0091] All of the compositions and methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied without departing from the concept and scope of the invention.

Example 1: *Determining Associations between Genetic Markers and Phenotypic Traits*

[0092] Simultaneous discovery and fine-mapping on a genome-wide basis of genes underlying quantitative traits (Quantitative Trait Loci: QTL) requires genetic markers densely covering the entire genome. As described in this example, a whole-genome, dense-coverage marker map was constructed from microsatellite and single nucleotide polymorphism (SNP) markers with previous estimates of location in the bovine genome, and from SNP markers with putative locations in the bovine genome based on homology with human sequence and the human/cow comparative map. A new linkage-mapping software package was developed, as an extension of the CRIMAP software (Green et al., Washington University School of Medicine, St. Louis, 1990), to allow more efficient mapping of densely-spaced markers genome-wide in a pedigreed livestock population (Liu and Grosz Abstract C014; Grapes et al. Abstract W244; 2006 Proceedings of the XIV Plant and Animal Genome Conference, www.intl-pag.org). The new linkage mapping tools build on the basic mapping principles programmed in CRIMAP to improve efficiency through partitioning of large pedigrees, automation of chromosomal assignment and two-point linkage analysis, and merging of sub-maps into complete chromosomes. The resulting whole-genome discovery map (WGDM) included 6,966 markers and a map length of 3,290 cM for an average map density of 2.18 markers/cM. The average gap between markers was 0.47 cM and the largest gap was 7.8 cM. This map provided the basis for whole-genome analysis and fine-mapping of QTL contributing to variation in productivity and fitness in dairy cattle.

Discovery and Mapping Populations

[0093] Systems for discovery and mapping populations can take many forms. The most effective strategies for determining population-wide marker/QTL associations

include a large and genetically diverse sample of individuals with phenotypic measurements of interest collected in a design that allows accounting for non-genetic effects and includes information regarding the pedigree of the individuals measured. In the present example, an outbred population following the grand-daughter design (Weller *et al.*, 1990) was used to discover and map QTL: the population, from the Holstein breed, had 529 sires each with an average of 6.1 genotyped sons, and each son has an average of 4216 daughters with milk data. DNA samples were collected from approximately 3,200 Holstein bulls and about 350 bulls from other dairy breeds; representing multiple sire and grandsire families.

Phenotypic Analyses

[0094] Dairy traits under evaluation include fitness and productivity traits such as milk yield (“MILK”) (pounds), fat yield (“FAT”) (pounds), fat percentage (“FATPCT”) (percent), productive life (“PL”) (months), somatic cell score (“SCS”) (Log), daughter pregnancy rate (“DPR”) (percent), protein yield (“PROT”) (pounds), protein percentage (“PROTPCT”) (percent), and net merit (“NM”) (dollar). These traits are sex-limited, as no individual phenotypes can be measured on male animals. Instead, genetic merits of these traits defined as PTA (predicted transmitting ability) were estimated using phenotypes of all relatives. Most dairy bulls were progeny tested with a reasonably larger number of daughters (e.g., >50), and their PTA estimation is generally more or considerably more accurate than individual cow phenotype data. The genetic evaluation for traditional dairy traits of the US Holstein population is performed quarterly by USDA. Detailed descriptions of traits, genetic evaluation procedures, and genetic parameters used in the evaluation can be found at the USDA AIPL web site (www.aipl.arsusda.gov). It is meaningful to note that the dairy traits evaluated in this example are not independent: FAT and PROT are composite traits of MILK and FATPCT, and MILK and PROTPCT, respectively. NM is an index trait calculated based on protein yield, fat yield, production life, somatic cell score, daughter pregnancy, calving difficulty, and several type traits. Protein yield and fat yield together account for >50% of NM, and the value of milk yield, fat content, and protein content is accounted for via protein yield and fat yield.

[0095] PTA data of all bulls with progeny testing data were downloaded from the USDA evaluation published at the AIPL site in November 2005. The PTA data were analyzed using the following two models:

$$y_{ij} = s_i + \text{PTAd}_{ij} \quad [1]$$

$$y_i = \mu + \beta_1(\text{SPTA})_i + \text{PTAd}_i \quad [2]$$

where y_i (y_{ij}) is the PTA of the i^{th} bull (PTA of the j^{th} son of the i^{th} sire); s_i is the effect of the i^{th} sire; $(\text{SPTA})_i$ is the sire's PTA of the i^{th} bull of the whole sample; μ is the population mean; PTAd_i (PTAd_{ij}) is the residual bull PTA.

[0096] Equation [1] is referred to as the sire model, in which sires were fitted as fixed factors. Among all USA Holstein progeny tested bulls, a considerably large number of sires only have a very small number of progeny tested sons (e.g., some have one son), and it is clearly undesirable to fit sires as fixed factors in these cases. It is well known the USA Holstein herds have been making steady and rapid genetic progress in traditional dairy traits in the last several decades, implying that the sire's effect can be partially accounted for by fitting the birth year of a bull. For sires with <10 progeny tested sons, sires were replaced with son's birth year in Eq. [1]. Eq. [2] is referred to as the SPTA model, in which sire's PTA are fitted as a covariate. Residual PTA (PTAd_i or PTAd_{ij}) were estimated using linear regression.

SNP-trait Association Analyses and Identification of Anchor Markers

[0097] In the present example, linkage disequilibrium (LD) mapping was performed in the aforementioned discovery population using statistical analyses based on probabilities of individual ordered genotypes estimated conditional on observed marker genotypes. The first step was to estimate sire's ordered genotype probabilities at all linked markers conditional on grandsire's and offspring marker genotype data. The exact calculation quickly becomes computationally infeasible as the size and complexity of the pedigree and number of linked markers increases. For example, there are, in total 2^k ordered genotypes for all linked loci when a sire has k linked heterozygous loci. A stepwise procedure developed based on a likelihood ratio test was used for estimating probabilities of sire's ordered genotypes at all linked markers.

[0098] The probabilities of ordered genotypes at loci of interest were estimated conditional on flanking informative markers as follows:

$$P(H_{sik}H_{dlk} | \mathbf{M}) = \sum_a \sum_b P(\mathbf{H}_{sa}\mathbf{H}_{db} | \mathbf{M}) * P(H_{sik}H_{dlk} | \mathbf{H}_{sa}\mathbf{H}_{db}, \mathbf{M}) \quad [3]$$

Where $P(\mathbf{H}_{sa}\mathbf{H}_{db} | \mathbf{M})$ is the probability of sire having a pair of haplotypes (or order genotype) $\mathbf{H}_{sa}\mathbf{H}_{db}$ at all linked loci conditional on the observed genotype data \mathbf{M} , and $P(H_{sik}H_{dlk} | \mathbf{H}_{sa}\mathbf{H}_{db}, \mathbf{M})$ is the probability of a son having ordered genotype $H_{sik}H_{dlk}$ at loci of interest conditional on sire's ordered genotype $\mathbf{H}_{sa}\mathbf{H}_{db}$ at all linked loci and the observed genotype data \mathbf{M} .

[0099] To determine associations between haplotypes probabilities and trait phenotypes, haplotypes of neighboring (and/or non-neighboring) markers across each chromosome were defined by setting the maximum length of a chromosomal interval and minimum and maximum number of markers to be included. Clearly, one needs to set similar parameters to form or define groups of marker loci for haplotype evaluation. The association between pre-adjusted trait phenotypes and haplotype (or pair of haplotype that is alternatively termed as ordered genotypes) was evaluated via a regression approach with the following models:

$$PTAd_k = \sum_i \beta_{si} P(H_{sik}) + e_k \quad [4]$$

$$PTAd_k = \sum_i \beta_{di} P(H_{dik}) + e_k \quad [5]$$

$$PTAd_k = \sum_i \beta_{si} [P(H_{sik}) + P(H_{dik})] + e_k \quad [6]$$

$$PTAd_k = \sum_i \beta_{si} [P(H_{sik}H_{djk}) + P(H_{sjk}H_{dik})] + e_k \quad [7]$$

where $PTAd_k$ is the preadjusted PTA of the k^{th} bull as defined in Eq. [1] under the sire model and can be replaced with $PTAd_i$ as defined in Eq [2] under the SPTA model, and e_k is the residual; $P(H_{sik})$ and $P(H_{dik})$ are the probability of paternal and maternal haplotype of individual k being haplotype i ; $P(H_{sik}H_{dik})$ is the probability of individual k having paternal haplotype i and maternal haplotype j that can be estimated using Eq.

[3]; all β are corresponding regression coefficients. Equations [4], [5], [6], and [7] are designed to model paternal haplotype, maternal haplotype, additive haplotype, and genotype effects, respectively.

[0100] Least-squares methods were used to estimate the effect of a haplotype or haplotype pair on a phenotypic trait and the regular F-test used to test the significance of the effect. Permutation tests were performed based on phenotype permutation (20,000) within each paternal half-sib family to estimate Type I error rate (p value)

Linkage Disequilibrium:

[0101] In the Holstein bulls that were sampled in our study, high level of LD was found between tightly linked SNPs (Table 2). As linkage distance between two SNPs increased, the level of LD decreased. Furthermore, it is reasonable to believe that the extent and range of LD in other dairy cattle populations is similar to that in Holstein population.

Table 1. LD as a function of linkage distance

| Linkage Distance (cM) | Average r^2 |
|-----------------------|---------------|
| <0.1 | 0.465 |
| 0.10 | 0.203 |
| 0.29 | 0.164 |
| 0.50 | 0.127 |
| 0.70 | 0.109 |
| 0.88 | 0.085 |
| 1.07 | 0.081 |

[0102] These results suggested that, for any two loci in very tight linkage, the probability of having high level of LD is high. Therefore, an anonymous SNP that is in very tight linkage with a SNP that was found to be statistically associated with a dairy trait can be found, with high probability, to be statistically associated with the same dairy trait, and the probability is generally an increasing function of LD. As a result, it is an efficient way to exploit similar SNP-trait association by discovering

SNP in the neighboring regions. The threshold for minimum level of LD for this approach is dependent on various factors including the nature of an application, effect size, and sample size. Although not limited by this theory, a minimum of $r^2 = 0.1$ is believed to be a conservative estimate of useful LD for breeding purposes. More preferably, $r^2 = 0.2$ is expected to be a useful LD (Zondervan et al, 2004).

Mapping of SNPs to Bovine Sequence Assembly

[0103] The mapping of SNPs to bovine genomic sequence assembly was done by comparing SNP sequences with the assembled bovine genomic sequences obtained from <ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/fasta/Btau20060815-freeze>. The SNPs' sequences were constructed by concatenating the left flank sequence, one of the SNP alleles (1st character of ALLELES), and right flank sequence. The sequences were blasted (linux megablast 2.2.15 was used for sequence comparison on a computer farm, against the bovine sequence assembly. The matches were further filtered to remove matches that have match length $< 0.90 * seq_length$. Furthermore, those SNPs that match to ambiguous locations and unknown chromosomes are ignored in this study.

Bovine Physical Map Distance vs. Genetic Map Distance

[0104] To calculate the average physical map distance for 1 cM of genetic distance, we used those SNP markers on bovine genetic map that can be placed on bovine sequence assembly, and calculated the distance between SNPs that were placed on the two extreme ends of each chromosome as the physical map distance and genetic map distance for physical map and genetic map respectively (ignoring the unmapped assembly sequences for the purpose of this study)

Table 2. Bovine Physical Map Distance vs. Genetic Map Distance

| CHROMOSOME | GENETIC DISTANCE | PHYSICAL DISTANCE | Chromosome-wide P/G | Gemome-wide P/G |
|------------|------------------|-------------------|---------------------|-----------------|
| 1 | 182.4 | 145459514 | 797475.406 | 702659.104 |
| 2 | 143.1 | 125272635 | 875420.231 | 702659.104 |

| | | | | |
|----|-------|-----------|------------|------------|
| 3 | 150.7 | 116288861 | 771658.003 | 702659.104 |
| 4 | 147.9 | 110696246 | 748453.32 | 702659.104 |
| 5 | 149.7 | 117884278 | 787470.127 | 702659.104 |
| 6 | 139.9 | 111654569 | 798102.709 | 702659.104 |
| 7 | 144.6 | 100731928 | 696624.675 | 702659.104 |
| 8 | 154.8 | 145881479 | 942386.815 | 702659.104 |
| 9 | 126.9 | 94564311 | 745187.636 | 702659.104 |
| 10 | 146.2 | 95751528 | 654935.212 | 702659.104 |
| 11 | 131 | 100314855 | 765762.252 | 702659.104 |
| 12 | 113.5 | 77411399 | 682038.758 | 702659.104 |
| 13 | 121.3 | 82744210 | 682145.177 | 702659.104 |
| 14 | 123 | 81898107 | 665838.268 | 702659.104 |
| 15 | 108.1 | 73391295 | 678920.398 | 702659.104 |
| 16 | 101.4 | 71894317 | 709016.933 | 702659.104 |
| 17 | 106.5 | 98165715 | 921743.803 | 702659.104 |
| 18 | 90.3 | 54044521 | 598499.679 | 702659.104 |
| 19 | 109 | 62440234 | 572846.183 | 702659.104 |
| 20 | 94.1 | 67289824 | 715088.459 | 702659.104 |
| 21 | 98.3 | 62906616 | 639945.229 | 702659.104 |
| 22 | 93.8 | 56761946 | 605138.017 | 702659.104 |
| 23 | 78.3 | 48265813 | 616421.622 | 702659.104 |
| 24 | 80.1 | 60029443 | 749431.248 | 702659.104 |
| 25 | 66.9 | 41927437 | 626718.042 | 702659.104 |
| 26 | 74.7 | 47491137 | 635758.193 | 702659.104 |
| 27 | 68.4 | 43195094 | 631507.222 | 702659.104 |
| 28 | 73.9 | 37681315 | 509896.008 | 702659.104 |
| 29 | 81.2 | 44877972 | 552684.384 | 702659.104 |

[0105] In Table 2, “*P/G*” is calculated as the physical distance over the genetic distance for each chromosome (Chromosome-wide *P/G*), and an average of physical distance over the genetic distance all chromosomes yields *Genome-wide P/G*. Based on the data presented in Table 2, a genetic distance of 1 cM is assumed to be equivalent to about 702660 bp in physical map distance.

Markers in close Proximity to Anchor Markers

[0106] Public SNPs that are in close proximity to the SNPs that we have identified

to be significantly associated with traits include Daughter Pregnancy Rate, Milk Yield and Composition, Net Merit, Productive Life, and Somatic Cell Score, were identified by their physical map location. If an Anchor Marker is located at position z , those public SNPs that are located between $z-70266$ and $z+70266$ are deemed as in close proximity of the Anchor Marker. These markers, which are closely associated with the anchor markers, are called “neighboring markers” or “neighboring SNPs”.

Genes in close Proximity to Anchor Markers

[0107] Bovine genes were retrieved from the file “Bos_taurus.Btau_3.1.43.pep.known.fa.gz” from ensemble ftp site at “ftp.ensembl.org/pub/current_bos_taurus/data/fasta/pep/”. In total 18654 entries were included in the file with protein sizes ranging from 8 to 23992. The Ensembl genes are annotated by the Ensembl automatic analysis pipeline using either a GeneWise model from a species-specific or vertebrate protein, a set of aligned species-specific cDNAs followed by GenomeWise for ORF prediction or from GENSCAN exons supported by protein, cDNA and EST evidence. GeneWise models are further combined with available aligned cDNAs to annotate UTRs (www.ensembl.org/Homo_sapiens/helpview?se=1;ref=;kw=geneview). The positions where these genes locate on the bovine genome were retrieved from the FASTA sequence header lines. For example, the genome mapping information can be obtained from the header line “>ENSP00000328693 pep:novel chromosome:NCBI35:1:904515:910768:1 gene:ENSG00000158815:transcript:ENST00000328693” as from 904,515 bp to 910,768 bp on chromosome 1 of the bovine genome (build NCBI35, the same as Btau_3.1). The stable transcript identifier is used as gene name in this application.

[0108] Genes that are in close proximity to “anchor markers” (the SNPs that we have identified to be significantly associated with traits, including Daughter Pregnancy Rate, Milk Yield and Composition, Net Merit, Productive Life, and Somatic Cell Score), were identified by their physical map location. If an anchor marker is located at position z , those bovine genes with any part located between $z-70266$ and $z+70266$ are deemed as in close proximity of the anchor marker. These genes, which are closely associated with the anchor markers are called “neighboring genes”. Markers within these genes, such as for example SNPs, which are closely

associated with the anchor markers, are called “neighboring markers” or “neighboring SNPs”.

Example 2: Use of single nucleotide polymorphisms to improve offspring traits

[0109] To improve the average genetic merit of a population for a chosen trait, one or more of the markers with significant association to that trait can be used in selection of breeding animals. In the case of each discovered locus, use of animals possessing a marker allele (or a haplotype of multiple marker alleles) in population-wide LD with a favorable QTL allele will increase the breeding value of animals used in breeding, increase the frequency of that QTL allele in the population over time and thereby increase the average genetic merit of the population for that trait. This increased genetic merit can be disseminated to commercial populations for full realization of value.

[0110] For example, a progeny-testing scheme could greatly improve its rate of genetic progress or graduation success rate via the use of markers for screening juvenile bulls. Typically, a progeny testing program would use pedigree information and performance of relatives to select juvenile bulls as candidates for entry into the program with an accuracy of approx 0.5. However, by adding marker information, young bulls could be screened and selected with much higher accuracy. In this example, DNA samples from potential bull mothers and their male offspring could be screened with a genome-wide set of markers in linkage disequilibrium with QTL, and the bull-mother candidates with the best marker profile could be contracted for matings to specific bulls. If superovulation and embryo transfer (ET) is employed, a set of 5-10 offspring could be produced per bull mother per flush procedure. Then the marker set could again be used to select the best male offspring as a candidate for the progeny test program. If genome-wide markers are used, it was estimated that accuracies of marker selection could reach as high as 0.85 (Meuwissen et al., 2001). This additional accuracy could be used to greatly improve the genetic merit of candidates entering the progeny test program and thereby increasing the probability of successfully graduating a marketable progeny-tested bulls. This information could also be used to reduce program costs by decreasing the number of juvenile bull candidates tested while maintaining the same number of successful graduates. In the

extreme, very accurate marker breeding values (MBV) could be used to directly market semen from juvenile sires without the need of progeny-testing at all. Due to the fact that juveniles could now be marketed starting at puberty instead of 4.5 to 5 years, generation interval could be reduced by more than half and rates of gain could increase as much as 68.3% (Schrooten et al., 2004). With the elimination of the need for progeny testing, the cost of genetic improvement for the artificial insemination industry would be vastly improved (Schaeffer, 2006).

[0111] In an alternate example, a centralized or dispersed genetic nucleus (GN) population of cattle could be maintained to produce juvenile bulls for use in progeny testing or direct sale on the basis of MBVs. A GN herd of 1000 cows could be expected to produce roughly 3000 offspring per year, assuming the top 10-15% of females were used as ET donors in a multiple-ovulation and embryo-transfer (MOET) scheme. However, markers could change the effectiveness MOET schemes and in vitro embryo production. Previously, MOET nucleus schemes have proven to be promising from the standpoint of extra genetic gain, but the costs of operating a nucleus herd together with the limited information on juvenile animals has limited widespread adoption. However, with marker information, juveniles can be selected much more accurately than before resulting in greatly reduced generation intervals and boosted rates of genetic response. This is especially true in MOET nucleus herd schemes because, previously, breeding values of full-sibs would be identical, but with marker information the best full-sib can be identified early in life. The marker information would also help limit inbreeding because less selection pressure would be placed on pedigree information and more on individual marker information. An early study (Meuwissen and van Arendonk, 1992) found advantages of up to 26% additional genetic gain when markers were employed in nucleus herd scenarios; whereas, the benefit in regular progeny testing was much less.

[0112] Together with MAS, female selection could also become an important source of genetic improvement particularly if markers explain substantial amounts of genetic variation. Further efficiencies could be gained by marker testing of embryos prior to implantation (Bredbacka, 2001). This would allow considerable selection to occur on embryos such that embryos with inferior marker profiles could be discarded prior to implantation and recipient costs. This would again increase the cost effectiveness of nucleus herds because embryo pre-selection would allow equal progress to be made

with a smaller nucleus herd. Alternatively, this presents further opportunities for pre-selection prior to bulls entering progeny test and rates of genetic response predicted to be up to 31% faster than conventional progeny testing (Schrooten et al., 2004).

[0113] The first step in using a SNP for estimation of breeding value and selection in the GN is collection of DNA from all offspring that will be candidates for selection as breeders in the GN or as breeders in other commercial populations (in the present example, the 3,000 offspring produced in the GN each year). One method is to capture shortly after birth a small bit of ear tissue, hair sample, or blood from each calf into a labeled (bar-coded) tube. The DNA extracted from this tissue can be used to assay an essentially unlimited number of SNP markers and the results can be included in selection decisions before the animal reaches breeding age.

[0114] One method for incorporating into selection decisions the markers (or marker haplotypes) determined to be in population-wide LD with valuable QTL alleles (see Example 1) is based on classical quantitative genetics and selection index theory (Falconer and Mackay, 1996; Dekkers and Chakraborty, 2001). To estimate the effect of the marker in the population targeted for selection, a random sample of animals with phenotypic measurements for the trait of interest can be analyzed with a mixed animal model with the marker fitted as a fixed effect or as a covariate (regression of phenotype on number of allele copies). Results from either method of fitting marker effects can be used to derive the allele substitution effects, and in turn the breeding value of the marker:

$$\alpha_1 = q[a + d(q - p)] \quad \text{[Equation 3]}$$

$$\alpha_2 = -p[a + d(q - p)] \quad \text{[Equation 4]}$$

$$\alpha = a + d(q - p) \quad \text{[Equation 5]}$$

$$g_{A_1A_1} = 2(\alpha_1) \quad \text{[Equation 6]}$$

$$g_{A_1A_2} = (\alpha_1) + (\alpha_2) \quad \text{[Equation 7]}$$

$$\mathbf{g}_{A_2A_2} = 2(\alpha_2) \quad \text{[Equation 8]}$$

where α_1 and α_2 are the average effects of alleles 1 and 2, respectively; α is the average effect of allele substitution; p and q are the frequencies in the population of alleles 1 and 2, respectively; a and d are additive and dominance effects, respectively; $\mathbf{g}_{A_1A_1}$, $\mathbf{g}_{A_1A_2}$ and $\mathbf{g}_{A_2A_2}$ are the (marker) breeding values for animals with marker genotypes A1A1, A1A2 and A2A2, respectively. The total trait breeding value for an animal is the sum of breeding values for each marker (or haplotype) considered and the residual polygenic breeding value:

$$EBV_{ij} = \sum \hat{\mathbf{g}}_j + \hat{U}_i \quad \text{[Equation 9]}$$

where EBV_{ij} is the Estimated Trait Breeding Value for the i^{th} animal, $\sum \hat{\mathbf{g}}_j$ is the marker breeding value summed from $j = 1$ to n where n is the total number of markers (haplotypes) under consideration, and \hat{U}_i is the polygenic breeding value for the i^{th} animal after fitting the marker genotype(s).

[0115] These methods can readily be extended to estimate breeding values for selection candidates for multiple traits, the breeding value for each trait including information from multiple markers (haplotypes), all within the context of selection index theory and specific breeding objectives that set the relative importance of each trait. Other methods also exist for optimizing marker information in estimation of breeding values for multiple traits, including random models that account for recombination between markers and QTL (e.g., Fernando and Grossman, 1989), and the potential inclusion of all discovered marker information in whole-genome selection (Meuwissen et al., Genetics 2001). Through any of these methods, the markers reported herein that have been determined to be in population-wide LD with valuable QTL alleles may be used to provide greater accuracy of selection, greater rate of genetic improvement, and greater value accumulation in the dairy industry.

Example 3: Identification of SNPs

[0116] A nucleic acid sequence contains a SNP of the present invention if it comprises at least 20 consecutive nucleotides that include and/or are adjacent to a polymorphism described in Tables 3, 4, 5, and 6 and the Sequence Listing. Alternatively, a SNP of the present invention may be identified by a shorter stretch of consecutive nucleotides which include or are adjacent to a polymorphism which is described in Table 3, 4, 5, and 6 and the Sequence Listing in instances where the shorter sequence of consecutive nucleotides is unique in the bovine genome. A SNP site is usually characterized by the consensus sequence in which the polymorphic site is contained, the position of the polymorphic site, and the various alleles at the polymorphic site. "Consensus sequence" means DNA sequence constructed as the consensus at each nucleotide position of a cluster of aligned sequences.

[0117] Consensus sequence can be based on either strand of DNA at the locus, and states the nucleotide base of either one of each SNP allele in the locus and the nucleotide bases of all Indels in the locus, or both SNP alleles using degenerate code (IUPAC code: M for A or C; R for A or G; W for A or T; S for C or G; Y for C or T; K for G or T; V for A or C or G; H for A or C or T; D for A or G or T; B for C or G or T; N for A or C or G or T; Additional code that we use include I for "-" or A; O for "-" or C; E for "-" or G; L for "-" or T; where "-" means a deletion). Thus, although a consensus sequence may not be a copy of an actual DNA sequence, a consensus sequence is useful for precisely designing primers and probes for actual polymorphisms in the locus.

[0118] Such SNP have a nucleic acid sequence having at least 90% sequence identity, more preferably at least 95% or even more preferably for some alleles at least 98% and in many cases at least 99% sequence identity, to the sequence of the same number of nucleotides in either strand of a segment of animal DNA which includes or is adjacent to the polymorphism. The nucleotide sequence of one strand of such a segment of animal DNA may be found in a sequence in the group consisting of SEQ ID NO:1 through SEQ ID NO:262,149. It is understood by the very nature of polymorphisms that for at least some alleles there will be no identity at the polymorphic site itself. Thus, sequence identity can be determined for sequence that is exclusive of the polymorphism sequence. The polymorphisms in each locus are described in the tables and the sequence listing.

[0119] Shown below are examples of public bovine SNPs that match each other:

SNP ss38333809 was determined to be the same as ss38333810 because 41 bases (with the polymorphic site at the middle) from each sequence match one another perfectly (match length=41, identity=100%).

```

ss38333809:  tcttacacatcaggagatagytccgaggtggatttctacaa
                |||
ss38333810:  tcttacacatcaggagatagytccgaggtggatttctacaa

ss38333809 is SEQ ID NO:262146
ss38333810 is SEQ ID NO:262147
    
```

[0120] SNP ss38333809 was determined to be the same as ss38334335 because 41 bases (with the polymorphic site at the middle) from each sequence match one another at all bases except for one base (match length=41, identity=97%).

```

ss38333809:  tcttacacatcaggagatagytccgaggtggatttctacaa
                |||
ss38334335:  tcttacacatcaggagatggytccgaggtggatttctacaa

ss38333809 is SEQ ID NO:262146
ss38334335 is SEQ ID NO:262149
    
```

Example 4: Quantification of and genetic evaluation for production traits

[0121] Quantifying production traits can be accomplished by measuring milk of a cow and milk composition at each milking, or in certain time intervals only. In the USDA yield evaluation the milk production data are collected by Dairy Herd Improvement Associations (DHIA) using ICAR approved methods. Genetic evaluation includes all cows with the known sire and the first calving in 1960 and later and pedigree from birth year 1950 on. Lactations shorter than 305 days are extended to 305 days. All records are preadjusted for effects of age at calving, month of calving, times milked per day, previous days open, and heterogeneous variance. Genetic evaluation is conducted using the single-trait BLUP repeatability model. The model includes fixed effects of management group (herd x year x season plus register

status), parity x age, and inbreeding, and random effects of permanent environment and herd by sire interaction. PTAs are estimated and published four times a year (February, May, August, and November). PTAs are calculated relative to a five year stepwise base i.e., as a difference from the average of all cows born in the current year, minus five (5) years. Bull PTAs are published estimating daughter performance for bulls having at least 10 daughters with valid lactation records.

Example 5: Quantification of reproductive traits in daughters (cows) and sires' PTAs

[0122] Quantification of and genetic evaluation of the reproductive capability such as calving ease (CE), occurrence of stillbirths (SB) and daughter pregnancy rate (DPR). Calving ease measures the ability of a particular cow (daughter) to calve easily. CE is scored by the owner on a scale of 1 to 5, 1 meaning no problems encountered or unobserved birth and 5 meaning extreme difficulty. The CE PTAs for sires are expressed as percent difficult births in primiparous daughter heifers (%DBH), where difficult births are those scored as requiring considerable force or being extremely difficult (4 or 5 on a five point scale). SB is scored by the owner on a scale of 1 to 3, 1 meaning the calf was born alive and was alive 48 h postpartum, 2 meaning the calf was born dead, and 3 indicating the calf was born alive but died within 48 h postpartum. SB scores of 2 and 3 are combined into a single category *for* evaluation. The SB PTAs for sires are expressed as percent stillbirths in daughter heifers (%SBH), where stillborn calves are those scored as dead at birth or born alive but died within 48 h of birth (2 or 3 on a three point scale). Pregnancy rate is a function of the number of days open, which is the number of days between calving and a successful breeding. DPR is defined as the percentage of nonpregnant cows (daughters) that become pregnant during each 21-day period. A DPR PTA of "1" implies that daughters from this bull are 1% more likely to become pregnant during that estrus cycle than a bull with a DPR PTA of zero.

Example 6: Quantification of and genetic evaluation for productive life (PL)

[0123] Productive life (PL) is defined as the length of time a cow remains in a milking herd before removal by voluntary or involuntary culling (due to health or fertility problems), or death. PL is usually measured as the number of days, months,

or days in milk (DIM) from the first calving to the day the cow exits the herd (due to death, culling, or selling to non-dairy purposes). Because some cows are still alive at the time of data collection, their records are projected (VanRaden, P.M. and E.J.H. Klaaskate. 1993) or treated as censored (Ducrocq, 1987). The USDA genetic evaluation for PL includes all cows with first calving in 1960 and later (born in 1950 and later for the pedigree). Cows born at least 3 years prior to evaluation, with a valid sire ID and first lactation records are considered. PL is considered to be completed at 7 years of age. Records are extended for cows that have not had the opportunity to reach 7 years of age because they are still alive, were sold for dairy purposes, or the herd discontinued testing. Cows sold for dairy purposes or in herds that discontinued testing receive extended records if they had opportunity to reach 3 years of age; otherwise their records are discarded. The method of genetic evaluation is a single trait BLUP animal model. The statistical model includes effects of management group (based on herd of first lactation and birth date) and sire by herd interaction. Sires' PTAs for PL are calculated relative to a five year stepwise base i.e., as a difference from the average PL of all cows born in the current year, minus five (5) years.

Example 7: Quantification of somatic cell score in daughters (cows) and sires' PTAs

[0124] Quantifying somatic cell score (SCS) is accomplished by calculating $\log_2(\text{SCC}/100,000)+3$, where SCC is number of somatic cells per milliliter of milk from a cow (daughter). The SCS PTAs for sires are expressed as a deviation from a SCS PTA of zero.

REFERENCES

[0125] The references cited in this application, both above and below, are specifically incorporated herein by reference.

Non-Patent Literature

- Abdel-Azim, G and Freeman, AE, (2002) *J. Dairy Sci.* 85:1869–1880.
- Blott, S., Kim, J.J., Moiso, S., et al. (2003). *Genetics* **163**: 253-266.
- Ciobanu, DC, Bastiaansen, JWM, Longergan, SM, Thomsen, H, Dekkers, JCM, Plastow, GS, and Rothschild, MF, (2004) *J. Anim. Sci.* 82:2829–39.
- Cohen-Zinder, M. et al. (2005) *Genome Res.* **15**:936-44.
- Davis, GP and DeNise, SK, (1998) *J. Anim. Sci.* 76:2331–39.
- Dekkers, JCM, and Chakraborty, R, (2001) *J. Anim. Sci.* 79:2975-90.
- Demars J, Riquet J, Feve K, Gautier M, Morisson M, Demeure O, Renard C, Chardon P, Milan D. (2006), *BMC Genomics*, 24:7-13
- Du and Hoeschele, (2000) *Genetics* 156:2051–62.
- Ducrocq, V. 1987. An analysis of length of productive life in dairy cattle. Ph.D. Diss.. Cornell Univ., Ithaca, NY; Univ. Microfilms Int., Ann Arbor, MI.
- Everts-van der Wind A, Larkin DM, Green CA, Elliott JS, Olmstead CA, Chiu R, Schein JE, Marra MA, Womack JE, Lewin HA. (2005) *Proc Natl Acad Sci U S A*, 20;102(51):18526-31.
- Falconer, DS, and Mackay, TFC, (1996) *Introduction to Quantitative Genetics*. Harlow, UK: Longman.
- Fernando, R, and Grossman, M, (1989) Marker assisted selection using best linear unbiased prediction. *Genetics Selection Evolution* 21:467-77.
- Franco, MM, Antunes, RC, Silva, HD, and Goulart, LR (2005) *J. Appl. Genet.* 46(2):195–200.
- Grisart, B. et al. (2002) *Genome Res.* **12**:222-231
- Grosz, MD, Womack, JE, and Skow, LC (1992) *Genomics*, 14(4):863-868.
- Hayes, B, and Goddard, ME, (2001) *Genet. Sel. Evol.* 33:209–229.
- Hayes, B, Chamberlain, A.J., Goddard, M.E. (2006) *Proc. 8th WCGALP* **22**:(16).
- Kaminski, S, Ahman, A, Rusc, A, Wojcik, E, and Malewski, T (2005) *J. Appl. Genet.* 46(1):45–58.

- Kuhn, C. et al. (2004). *Genetics* **167**:1873-81.
- Kwok PY, Methods for genotyping single nucleotide polymorphisms, (2001), *Annu. Rev. Genomics Hum. Genet.*, 2:235–258.
- Markstein M, Zinzen R, Markstein P, Yee KP, Erives A, Stathopoulos A, and Levine M. 2004. A regulatory code for neurogenic gene expression in the *Drosophila* embryo. *Development* 131, 2387-2394
- McDowell JC and Dean A. 1999. Structural and functional cross-talk between a distant enhancer and the epsilon-globin gene promoter shows interdependence of the two elements in chromatin. *Mol Cell Biol.* 19(11):7600-9.
- Meuwissen, THE, and Van Arendonk, JAM, (1992) *J. Dairy Sci.* 75:1651–1659.
- Meuwissen, THE, Hayes, BJ, and Goddard, ME, (2001) *Genetics.* 157:1819–29.
- Rothschild and Plastow, (1999), *AgBioTechNet* 10:1-8.
- Schaeffer, LR (2006) *J. Anim. Breed. Genet.* 123:218–223.
- Schnabel, R. et al. (2005) *PNAS* **102**:6896-6901.
- Schrooten, C, Bovenhuis, H, van Arendonk, JAM, and Bijma, P (2005) *J. Dairy Sci.* 88:1569–1581.
- Sharma, BS, Jansen, GB, Karrow, NA, Kelton, D, and Jiang, Z, (2006) *J. Dairy Sci.* 89:3653–3663.
- Short, TH, *et al.* (1997) *J. Anim. Sci.* 75:3138–3142.
- Spelman, RJ and Bovenhuis, H, (1998) *Animal Genetics*, 29:77–84.
- Spelman, RJ and Garrick, DJ, (1998) *J. Dairy Sci.* 81:2942–2950.
- Stearns, TM, Beever, JE, Southey, BR, Ellis, M, McKeith, FK and Rodriguez-Zas, SL, (2005) *J. Anim. Sci.* 83:1481–93.
- Syvänen AC, Accessing genetic variation: genotyping single nucleotide polymorphisms, (2001) *Nat. Rev. Genet.* 2:930–942.
- VanRaden, P.M. and E.J.H. Klaaskate. 1993. *J. Dairy Sci.* 76:2758-2764.
- Verrier, E, (2001) *Genet. Sel. Evol.* 33:17–38.
- Villanueva, B, Pong-Wong, R, Fernandez, J, and Toro, MA (2005) *J. Anim. Sci.* 83:1747–52.
- Weller JI, Kashi Y, Soller M. (1990) *J. Dairy Sci.* 73:2525-37

Williams, JL, (2005), *Rev. Sci. Tech. Off. Int. Epiz.* 24(1):379–391.

Windig, JJ, and Meuwissen, THE, (2004) *J. Anim. Breed. Genet.* 121:26-39.

Winter, A. et al. (2002). *PNAS*, **99**:9300-9305.

Womack, J, (1987), *Dev. Genet.* 8(4):281-293.

Yasue H, Kiuchi S, Hiraiwa H, Ozawa A, Hayashi T, (2006), *Cytogenet. Genome Res.*, 112(1-2):121-125.

Youngerman, SM, Saxton, AM, Oliver, SP, and Pighetti, GM, (2004) *J. Dairy Sci.* 87:2442–2448.

Zondervan, K and Cardon, L (2004)_*Nat Rev Genet.* 2004 Feb; 5(2):89-100

Patent Literature (Swine)

| Patent/Publication Number | Title | Inventors | Pub. Date |
|---------------------------|---|---|-----------|
| 6,143,880 | Pig myogenin gene and method to identify polymorphisms related to muscle growth | Soumillion, Ann Elisabeth; te Pas, Marinus Frederik | 11/7/00 |
| 6,458,531 | Leptin receptor gene as a genetic marker for leanness in pigs | Rothschild, Max F.; Vincent, Amy L.; Ernst, Catherine W. | 10/1/02 |
| 6,803,190 | Melanocortin-4 receptor gene and use as a genetic marker for fat content, weight gain, and/or feed consumption of animals | Rothschild, Max F.; Kim, Kwan Suk; Larson, Niels J. | 10/12/04 |
| 6,919,177 | PRKAG3 alleles and use of the same as genetic markers for reproductive and meat quality traits | Rothschild, Max F.; Ciobanu, Daniel C.; Malek, Massoud; Plastow, Graham | 7/19/05 |
| US2003017470A1 | Novel PRKAG3 alleles and use of the same as genetic markers for reproductive and meat quality traits | Rothschild, Max F.; Ciobanu, Daniel C.; Malek, Massoud; Plastow, Graham | 1/23/03 |
| US2003104392A1 | Detection of growth performance of pigs | Lin, En-Chung; Chen, Ming-Yu; Huang, San-Yuan; Tsou, Hui-Liang; Lee, Wen-Chuan | 6/5/03 |
| US2003129610A1 | New QTL's on chromosomes X, 2, 6 and 7 of pigs | de Koning, Dirk-Jan; Groenen, Martinus Antonius, M.; van Arendonk, Johannes Antonius, M. | 7/10/03 |
| US2003211512A1 | Novel Ghrelin alleles and use of the same for genetically typing animals | Rothschild, Max F.; Kim, Kwan-Suk; Anderson, Lloyd L. | 11/13/03 |
| US2004029145A1 | Novel HMGA alleles and use of the same as genetic markers for growth, fatness, meat quality, and feed efficiency traits | Rothschild, Max F.; Kim, Kwan-Suk; Nguyen, Nguyet Thu | 2/12/04 |
| US2004101842A1 | Method to screen for improved meat characteristics in pigs | Haley, Christopher Simon; Matlin, Charlotte Anne | 5/27/04 |
| US2004235030A1 | Melanocortin-4 receptor gene and use as a genetic marker for fat content, weight gain, and/or feed consumption of animals | Rothschild, Max F.; Kim, Kwan Suk; Larsen, Niels J. | 11/25/04 |
| US2004253594A1 | Genetic markers for improved disease resistance in animals (nramp1) | Tuggle, Christopher K; Marklund, Lena; Stabel, Thomas J; Mellencamp, Martha A; Stumbaugh, Ambar | 12/16/04 |

| Patent/Publication Number | Title | Inventors | Pub. Date |
|---------------------------|---|---|-----------|
| US2004261138A1 | Genetic markers for improved meat characteristics in animals (MC4R) | Rothschild, Max F.; Kim, Kwan Suk; Emmett, Rebecca S. | 12/23/04 |
| US2005112648A1 | Genetic markers for pig backfat thickness | Lin, En-Chung; Chen, Ming-Yu; Huang, San-Yuan; Tsou, Hui-Liang; Lee, Wen-Chuan | 5/26/05 |
| US2005142560A1 | Method for identifying and managing livestock by genotype | Marquess, Foley Leigh-Shaw | 6/30/05 |
| US2005153317A1 | Methods and systems for inferring traits to breed and manage non-beef livestock | DeNise, Sue; Rosenfeld, David; Kerr, Richard; Bates, Stephen; Holm, Tom | 7/14/05 |
| US2005208551A1 | Novel PRKAG3 alleles and use of the same as genetic markers for reproductive and meat quality traits | Rothschild, Max F.; Ciobanu, Daniel C.; Malek, Massoud; Plastow, Graham | 9/22/05 |
| WO9723644A1 | The pig myogenin gene and method to identify polymorphisms related to muscle growth | SOUMILLION, Ann, Elizabeth, Maria; TE PAS, Marinus, Frederik, Willem | 7/3/97 |
| WO9846792A2 | Genetic marker based pig selection | Chang, Kin, Chow | 10/22/98 |
| WO0006777A2 | Melanocortin-4 Receptor Gene And Use As A Genetic Marker For Fat Content, Weight Gain, And/Or Feed Consumption Of Animals | ROTHSCHILD, MAX, F.; LARSON, NIELS, J.; KIM, KWAN, SUK | 2/10/00 |
| WO0036143A2 | Selecting animals for parentally imprinted traits | ANDERSSON, Leif; GEORGES, Michel; SPINCEMAILLE, Geert | 6/22/00 |
| WO0069882A1 | Genetic marker for meat quality, growth, carcass and reproductive traits in livestock | GREGGER, Douglas, L. | 11/23/00 |
| WO0146406A2 | New QTLs On Chromosomes X, 2, 6 And 7 Of PigS | DE KONING, Dirk-Jan; GROENEN, Martinus, Antonius, Mathilda; VAN ARENDONK, Johannus, Antonius, Maria | 6/28/01 |
| WO0175161A2 | Genetic markers for improved meat characteristics in animals | ROTHSCHILD, Max, F.; EMNETT, Rebecca; KIM, Kwan, Suk | 10/11/01 |
| WO0192570A2 | Method to screen for improved meat characteristics in pigs | HALEY, Christopher, Simon; MALTIN, Charlotte, Anne | 12/6/01 |

| Patent/Publication Number | Title | Inventors | Pub. Date |
|---------------------------|---|---|-----------|
| WO0220850A2 | Novel PRKAG3 alleles and use of the same as genetic markers for reproductive and meat quality traits | ROTHSCHILD, Max, F.; CIOBANU, Daniel, C.; MALEK, Massoud; PLASTOW, Graham | 3/14/02 |
| WO05001032A2 | Fine mapping of chromosome 17 quantitative trait loci and use of same for marker assisted selection | ROTHSCHILD, Max F.,; RAMOS, Antonio; KIM, Kwan Suk | 1/6/05 |
| WO05017204B1 | Use single nucleotide polymorphsm in the coding region of the porcine leptin receptor gene to enhance pork production | KOJIMA, Cheryl J., 1 DU, Fengxing, GROSZ, Michael, D., BYATT, John, C., | 9/29/05 |
| WO05040400A2 | Methods and systems for inferring traits to manage non-beef livestock | DENISE, Sue, K.; ROSENFELD, David; KERR, Richard; BATES, Stephen; HOLM, Tom | 5/6/05 |
| WO05078133A2 | Marker Assisted Best Linear Unbiased Predicted (MA-BLUP): software adaptations for practical applications for large breeding populations in farm animal species | WANG, Tianlin; LOHUIS, Michael, M.; KOJIMA, Cheryl, J.; DU, Fengxing; BYATT, John | 8/25/05 |

Patent Literature (Dairy)

| Patent/Publication Number | Title | Inventors | Pub. Date |
|---------------------------|--|--|-----------|
| US5041371 | Genetic marker for superior milk products in dairy cattle | Cowan, Charles M.; Dentine, Margaret R.; Ax, Roy L.; Schuler, Linda A. | 8/20/91 |
| US5374523 | Allelic variants of bovine somatotropin gene:genetic marker for superior milk production in bovine | Collier, Robert J.; Hauser, Scott D.; Krivi, Gwen G.; Lucy, Matthew C. | 12/20/94 |
| US5582987 | Methods for testing bovine for resistance or susceptibility to persistent lymphocytosis by detecting polymorphism in bola-dr3 exon 2 | Lewin, Harris A.; van Eijk, Michiel J. T. | 12/10/96 |
| US5614364 | Genetic marker for improved milk production traits in cattle | Tuggle, Christopher K.; Freeman, Albert E. | 3/25/97 |
| US2003039737A1 | Population of dairy cows producing milk with desirable characteristics and methods of making and using same | Cooper, Garth J. S. | 2/27/03 |
| US2004076977A1 | Marker assisted selection of bovine for improved milk production using diacylglycerol acyltransferase gene dgat1 | Georges, Michel Alphonse Julien; Coppieters, Wouter Herman Robert; Grisart, Bernard Marie-Josée Jean; Shell, Russell Grant; Jean Reid, Suzanne; Ford, Christine Ann; Spelman, Richard John | 4/22/04 |
| US2004234986A1 | Method of testing a mammal for its predisposition for fat content of milk and/ or its predisposition for meat marbling | Fries, Hans-Rudolf; Winter, Andreas | 11/25/04 |
| US2004241723A1 | Systems and methods for improving protein and milk production of dairy herds | Marquess, Foley Leigh Shaw; Laarveld, Bernard; Cleverly Buchanan, Fiona; Van Kessel, Andrew Gerald; Schmutz, Sheila Marie; Waldner, Cheryl; Christensen, David | 12/2/04 |

| | | | |
|----------------|--|--|----------|
| US2004254104A1 | Marker assisted selection of bovine for improved milk composition | Blott, Sarah; Kim, Jong-Joo; Schmidt-Kuntzel, Anne; Cornet, Anne; Berzi, Paulette; Cambisano, Nadine; Grisart, Bernard; Karim, Latifa; Simon, Patricia; Georges, Michel; Farnir, Frederic; Coppieters, Wouter; Moisiso, Sirja; Vilkki, Johanna; Spelman, Richard; Johnson, Dave; Ford, Christine; Snell, Russell | 12/16/04 |
| US2005123929A1 | Methods and compositions for genetically detecting improved milk production traits in cattle | Khatib, Hasan | 6/9/05 |
| US2005136440A1 | Method for identifying animals for milk production qualities by analysing the polymorphism of the pit-1 and kappa-casein genes | Renaville, Robert; Gengler, Nicolas | 6/23/05 |
| US2005137805A1 | Gene expression profiles that identify genetically elite ungulate mammals | Lewin, Harris A.; Liu, Zonglin; Rodriguez-Zas, Sandra; Everts, Robin E. | 6/23/05 |
| US2005153317A1 | Methods and systems for inferring traits to breed and manage non-beef livestock | DeNise, Sue; Rosenfeld, David; Kerr, Richard; Bates, Stephen; Holm, Tom | 7/14/05 |
| US2006037090A1 | Selecting animals for desired genotypic or potential phenotypic properties | Andersson, Leif; Andersson, Goran; Georges, Michel; Buys, Nadine | 2/16/06 |
| US2006094011A1 | Method for altering fatty acid composition of milk | Morris, Christopher Anthony; Tate, Michael Lewis | 5/4/06 |
| US2006121472A1 | Method for determining the allelic state of the 5'-end of the κ (a)s1-casein gene | Prinzenberg, Eva-Maria; Erhardt, George | 6/8/06 |
| US2006166244A1 | DNA markers for increased milk production in cattle | Schnabel, Robert D.; Sonstegard, Tad S.; Van Tassell, Curtis P.; Ashwell, Melissa S.; Taylor, Jeremy F. | 7/27/06 |

| | | | |
|--------------|--|---|----------|
| WO0236824A1 | Marker assisted selection of bovine for improved milk production using diacylglycerol acyltransferase gene dgat1 | GEORGES, MICHEL, ALPHONSE, JULIEN; COPPIETERS, WOUTER, HERMAN, ROBERT; GRISART, BERNARD, MARIE-JOSEE, JEAN; SNELL, RUSSELL, GRANT; REID, SUZANNE, JEAN; FORD, CHRISTINE, ANN; SPELMAN, RICHARD, JOHN | 5/10/02 |
| WO03104492A1 | Marker assisted selection of bovine for improved milk composition | BLOTT, SARAH; KIM, JONG-JOO; SCHMIDT-KUNTZEL, ANNE; CORNET, ANNE; BERZI, PAULETTE; CAMBISANO, NADINE; GRISART, BERNARD; KARIM, LATIFA; SIMON, PATRICIA; GEORGES, MICHEL; FARNIR, FREDERIC; COPPIETERS, WOUTER; MOISIO, SIRJA; VILKKI, JOHANNA; JOHNSON, DAVE; SPELMAN, RICHARD; FORD, CHRISTINE; SNELL, RUSSELL | 12/18/03 |
| WO04004450A1 | Method for altering fatty acid composition of milk | MORRIS, Christopher Anthony; TATE, Michael Lewis | 1/15/04 |
| WO04048609A2 | Methods and kits for the selection of animals having certain milk production capabilities, based on the analysis of a polymorphism in the somatotropin receptor gene | RENAVILLE, Robert; PARMENTIER, Isabelle | 6/10/04 |

| | | | |
|--------------|--|--|---------|
| WO04083456A1 | Systems and methods for improving protein and milk production of dairy herds | MARQUESS, Foley, Leigh, Shaw; LAARVELD, Bernard; CLEVERLY BUCHANAN, Fiona; VAN KESSEL, Andrew, Gerald; SCHMUTZ, Sheila, Marie; WALDNER, Cheryl; CHRISTENSEN, David | 9/30/04 |
| WO05007881A2 | Improving production characteristics of cattle | SCHMUTZ, SHEILA MARIE; GOODALL, JULIE JANINE | 1/27/05 |
| WO05030789A1 | Adrenergic receptor snp for improved milking characteristics | COLLIER, Robert, J.; LOHIUS, Michael; GROSZ, Michael | 4/7/05 |
| WO05040400A2 | Methods and systems for inferring traits to manage non-beef livestock | DENISE, Sue, K.; ROSENFELD, David; KERR, Richard; BATES, Stephen; HOLM, Tom | 5/6/05 |
| WO05056758A2 | Methods and compositions for genetically detecting improved milk production traits in cattle | KHATIB, Hasan | 6/23/05 |
| WO05089122A2 | Animals with reduced body fat and increased bone density | JOHNSON, Geoffrey, B.; PLATT, Jeffrey, L.; JOHNSON, Joel, W. | 9/29/05 |
| WO06076563A2 | Dna markers for increased milk production in cattle | SCHNABEL, Robert, D.; SONSTEGARD, Tad, S.; VAN TASSELL, Curtis, P.; ASHWELL, Melissa, S.; TAYLOR, Jeremy, F. | 7/20/06 |
| WO9213102A1 | Polymorphic dna markers in bovidae | Georges, Michel; MASSEY, Joseph, M. | 8/6/92 |
| WO9319204A1 | Bovine alleles and genetic markers and methods of testing of and using same | LEWIN, Harris, A.; VAN EIJK, Michiel, J., T. | 9/30/93 |
| WO9403641A1 | Genetic marker for dairy cattle production superiority | COLLIER, Robert, Joseph; HAUSER, Scott, David; KRIVI, Gwen, Grabowski; LUCY, Matthew, Christian | 2/17/94 |

DESCRIPTION OF THE TABLES

[0126] Table 3: provides a list of regions in the first column, SEQ ID numbers of anchor markers in the second column, SEQ ID numbers for markers within the region in the third column, and trait association abbreviations in the fourth column. Columns five through eight, nine through twelve, and thirteen through sixteen contain similarly arranged information. The abbreviations used for the trait associations are as follows:

M = Milk production: MILK, FAT, FATPCT, PROT, PROTPCT

S = Somatic Cell Score: SCS

D = Daughter Pregnancy Rate: DPR

P = Productive Life: PL

N = Net Merit: NM

[0127] Table 4: provides a list of anchor marker SEQ ID numbers in the first column, gene names in the second column, SEQ ID numbers of SNPs within the gene in the third column, and trait association abbreviations in the fourth column. Columns five through eight, nine through twelve, and thirteen through sixteen contain similarly arranged information. The abbreviations used for the trait associations are as follows:

M = Milk production: MILK, FAT, FATPCT, PROT, PROTPCT

S = Somatic Cell Score: SCS

D = Daughter Pregnancy Rate: DPR

P = Productive Life: PL

N = Net Merit: NM

[0128] In order to reduce redundancy, the prefix of each gene name has been replaced with a single letter. For example, ENSBTAT00000033728 is represented as Z033728; where “Z” in gene name stands for “ENSBTAT00000”.

[0129] **Table 5:** provides a list of phenotypic traits and the assigned identification numbers of SNPs found to be associated with each trait. The left column provides a counter to allow easier reading of the table. The “Trait” column lists the following traits: “FITNESS”, rows 1–2397; and “PRODUCTIVITY”, rows 2398–5117.

[0130] **Table 6:** provides the SEQ ID NO of the sequence associated with each of the SNPs listed in Table 5. The “SNP POSITION” column provides the position (nucleotide number) of the SNP within the associated sequence (SEQ ID NO) and the “SNP ALLELE 1” and “SNP ALLELE 2” columns provide the identity of the two nucleotides that occur most frequently at the SNP POSITION within the population analyzed.

[0131] **Table 7:** provides a list of the SNP ID NOs and SEQ ID NOs, listed in Tables 5 and 6, sorted numerically according to SNP ID NO.

What is claimed is:

1. A method for allocating an animal for use according to the animal's predicted marker breeding value for one or more traits selected from the group consisting of milk production, somatic cell score, daughter pregnancy rate, productive life, and/or net merit, the method comprising:
 - a. determining the animal's genotype at one or more locus/loci; wherein at least one locus contains a single nucleotide polymorphism (SNP), having at least two allelic variants; and wherein at least one SNP is:
 - i) selected from the SNPs described in Table 3 and the Sequence Listing; and/or
 - ii) SNPs located on the same chromosome and within 70 kilobases of one or more of the SNP anchor markers described in Table 3 and the Sequence Listing";
 - b. analyzing the determined genotype of at least one evaluated animal at one or more SNPs selected from:
 - i) the SNPs described in Table 3 and the Sequence Listing; and/or
 - ii) SNPs located on the same chromosome and within 70 kilobases of one or more of the SNP anchor markers described in Table 3 and the Sequence Listing,
 - c. correlating the analyzed genotype with one or more phenotypes using Table 3 and the Sequence Listing
 - d. allocating the animal for use according to its determined genotype.

2. The method of claim 1 wherein the animal's genotype is evaluated at 10 or more loci that contain SNPs selected from:
 - a. the SNPs described in Table 3 and the Sequence Listing;
 - b. and/or SNPs located on the same chromosome and within 70 kilobases of one or more of the SNP anchor markers described in Table 3 and the Sequence Listing.

3. The method of claim 2 wherein the animal's genotype is evaluated at 50 or more loci .
4. The method of claim 2 wherein the animal's genotype is evaluated at 100 or more loci.
5. The method of claim 2 wherein the animal's genotype is evaluated at 200 or more loci.
6. The method of 2 wherein the animal's genotype is evaluated at 500 or more loci.
7. The method of any of claims 1 to 6 that comprises whole-genome analysis.
8. A method for selecting a potential parent animal for improvement of one or more traits selected from the group consisting of milk production, somatic cell score, daughter pregnancy rate, productive life, and net merit :
 - a. determining at least one potential parent animal's genotype at one or more genomic locus/loci; wherein at least one locus contains a single nucleotide polymorphism (SNP) that has at least two allelic variants, and wherein at least one SNP is selected from:
 - i) the SNPs described in Table 3 and the Sequence Listing;
 - ii) and/or SNPs located on the same chromosome and within 70 kilobases of one or more of the SNP anchor markers described in Table 3 and the Sequence Listing;
 - b. analyzing the determined genotype of at least one evaluated animal for one or more SNPs selected from:
 - i) the SNPs described in Table 3;
 - ii) and/or SNPs located on the same chromosome and within 70 kilobases of one or more of the SNP anchor markers described in Table 3 and the Sequence Listing;
 - c. correlating the identified allelic variants with a milk production phenotype, somatic cell score phenotype, daughter pregnancy rate

- phenotype, productive life phenotype, and/or net merit phenotype; using the information provided in Table 3 and the sequence listing.
- d. allocating at least one animal for breeding use based on its genotype.
9. The method of claim 8 wherein the potential parent animal's genotype is evaluated at 10 or more loci that contain SNPs selected from:
- a. the SNPs described in Table 3; and/or
 - b. SNPs located on the same chromosome and within 70 kilobases of one or more of the SNP anchor markers described in Table 3 and the Sequence Listing.
10. The method of claim 8 wherein the potential parent animal's genotype is evaluated at 50 or more loci.
11. The method of claim 8 wherein the potential parent animal's genotype is evaluated at 100 or more loci.
12. The method of claim 8 wherein the potential parent animal's genotype is evaluated at 200 or more loci.
13. The method of claim 8 wherein the potential parent animal's genotype is evaluated at 500 or more loci.
14. The method of any of claims 8 to 13 that comprises whole-genome analysis.
15. A method of producing progeny animals comprising:
- a) identifying at least one potential parent animal that has been allocated for breeding in accordance with the method of any of claims 1 to 7;
 - b) producing progeny from the allocated animal through a process selected from the group consisting of:
 - i) natural breeding;
 - ii) artificial insemination;
 - iii) *in vitro* fertilization; and

- iv) collecting semen/spermatozoa or at least one ovum from the animal and contacting it, respectively, with ovum/ova or semen/spermatozoa from a second animal to produce a conceptus by any means.
16. The method of claim 15 comprising producing progeny through natural breeding.
17. The method of claim 15 comprising producing offspring through artificial insemination, embryo transfer, and/or *in vitro* fertilization.
18. A nucleic acid array for determining which allele of at least 100 SNPs associated with one or more traits selected from the group consisting of milk production, somatic cell score, daughter pregnancy rate, productive life, and net merit are present in a sample; wherein the array comprises 100 or more nucleic acid sequences capable of hybridizing, under stringent conditions, with at 100 or more SNPs selected from:
- a. the SNPs described in Table 3 and the Sequence Listing; and/or
 - b. SNP(s) located on the same chromosome and within 70 kilobases of one or more of the SNP anchor markers described in Table 3 and the Sequence Listing.
19. A method of identifying quantitative trait locus associated with milk production, somatic cell score, daughter pregnancy rate, productive life, and/or net merit by identifying a genetic marker located on the same chromosome and within 70 kilobases of at least one SNP selected from the group of SNP anchor markers described in Table 3 and the Sequence Listing, the method comprising:
- a) selecting a SNP from the group of anchor SNPs described in Table 3 and the Sequence Listing; wherein said SNP is associated with at least one trait as described in Table 3.
 - b) identifying a genetic polymorphism within 70 kilobases of this SNP.

20. The method of claim 19 wherein the genetic polymorphism is a SNP.
21. The method of claim 19 wherein the genetic polymorphism is a causal mutation underlying a quantitative trait locus
22. A method comprising evaluating an animal's genotype at 100 or more genomic locus/loci; wherein at least 100 loci each comprise a single nucleotide polymorphism (SNP) selected from:
 - a. the SNPs described in Table 3 and the Sequence Listing; and/or
 - b. SNPs located on the same chromosome and within 70 kilobases of one or more of the SNP anchor markers described in Table 3 and the Sequence Listing.
23. The method of claim 22 wherein said SNP is associated with at least one trait as described in table 3.
24. The method of claim 22 wherein the animal's genotype is evaluated at 200 or more loci.
25. The method of claim 22 wherein the animal's genotype is evaluated at 500 or more loci
26. The method of any of claim 22 that comprises whole-genome analysis.
27. A method for allocating an animal for use according to the animal's predicted marker breeding value for one or more traits selected from the group consisting of milk production, somatic cell score, daughter pregnancy rate, productive life, and/or net merit, the method comprising:
 - a. determining the animal's genotype at one or more locus/loci; wherein at least one locus contains a single nucleotide polymorphism (SNP), having at least two allelic variants; and wherein at least one SNP is selected from:

- i) the SNPs described in Table 4 and the Sequence Listing; and/or
 - ii) a SNP located in a gene described in Table 4;
 - b. analyzing the determined genotype of at least one evaluated animal at one or more SNPs selected from:
 - i) the SNPs described in Table 4 and the Sequence Listing; and/or
 - ii) a SNP located in a gene described in Table 4;
 - c. correlating the analyzed genotype with one or more phenotypes using Table 4 and the Sequence Listing
 - d. allocating the animal for use according to its determined genotype.
28. The method of claim 27 wherein the animal's genotype is evaluated at 10 or more loci that contain SNPs selected from:
 - a. the SNPs described in Table 4 and the Sequence Listing; and/or
 - b. a SNP located in a gene described in Table 4;
29. The method of claim 28 wherein the animal's genotype is evaluated at 50 or more loci .
30. The method of claim 28 wherein the animal's genotype is evaluated at 100 or more loci.
31. The method of claim 28 wherein the animal's genotype is evaluated at 200 or more loci.
32. The method of 28 wherein the animal's genotype is evaluated at 500 or more loci.
33. The method of any of claims 27 to 32 that comprises whole-genome analysis.

34. A method for selecting a potential parent animal for improvement of one or more traits selected from the group consisting of milk production, somatic cell score, daughter pregnancy rate, productive life, and net merit :
- a. determining at least one potential parent animal's genotype at one or more genomic locus/loci; wherein at least one locus contains a single nucleotide polymorphism (SNP) that has at least two allelic variants, and wherein at least one SNP is selected from:
 - i) the SNPs described in Table 4 and the Sequence Listing; and/or
 - ii) a SNP located in a gene described in Table 4;
 - b. analyzing the determined genotype of at least one evaluated animal for one or more SNPs selected from:
 - i) the SNPs described in Table 4; and/or
 - ii) a SNP located in a gene described in Table 4;
 - c. correlating the identified allelic variants with a milk production phenotype, somatic cell score phenotype, daughter pregnancy rate phenotype, productive life phenotype, and/or net merit phenotype; using the information provided in Table 4 and the sequence listing.
 - d. allocating at least one animal for breeding use based on its genotype.
35. The method of claim 34 wherein the potential parent animal's genotype is evaluated at 10 or more loci that contain SNPs selected from:
- a. the SNPs described in Table 4; and/or
 - b. a SNP located in a gene described in Table 4;
36. The method of claim 34 wherein the potential parent animal's genotype is evaluated at 50 or more loci.
37. The method of claim 34 wherein the potential parent animal's genotype is evaluated at 100 or more loci.
38. The method of claim 34 wherein the potential parent animal's genotype is evaluated at 200 or more loci.

39. The method of claim 34 wherein the potential parent animal's genotype is evaluated at 500 or more loci.
40. The method of any of claims 34 to 39 that comprises whole-genome analysis.
41. A method of producing progeny animals comprising:
- a) identifying at least one potential parent animal that has been allocated for breeding in accordance with the method of any of claims 27 to 34;
 - b) producing progeny from the allocated animal through a process selected from the group consisting of:
 - i) natural breeding;
 - ii) artificial insemination;
 - iii) *in vitro* fertilization; and
 - iv) collecting semen/spermatozoa or at least one ovum from the animal and contacting it, respectively, with ovum/ova or semen/spermatozoa from a second animal to produce a conceptus by any means.
42. The method of claim 41 comprising producing progeny through natural breeding.
43. The method of claim 41 comprising producing offspring through artificial insemination, embryo transfer, and/or *in vitro* fertilization.
44. A nucleic acid array for determining which allele of at least 100 SNPs associated with one or more traits selected from the group consisting of milk production, somatic cell score, daughter pregnancy rate, productive life, and net merit are present in a sample; wherein the array comprises 100 or more nucleic acid sequences capable of hybridizing, under stringent conditions, with at 100 or more SNPs selected from:
- a. the SNPs described in Table 4 and the Sequence Listing; and/or
 - b. a SNP located in a gene described in Table 4;

45. A method of identifying quantitative trait locus associated with milk production, somatic cell score, daughter pregnancy rate, productive life, and/or net merit by identifying a genetic marker located on the same chromosome and within 70 kilobases of at least one SNP selected from the group of SNP anchor markers described in Table 4 and the Sequence Listing, the method comprising:
- a) selecting a SNP from the group of anchor SNPs described in Table 4 and the Sequence Listing; wherein said SNP is associated with at least one trait as described in Table 4.
 - b) identifying a gene within 70 kilobases of this SNP.
 - c) associating said gene with at least one trait described in Table 4
46. The method of claim 45 further comprising identifying a genetic polymorphism within said gene.
47. The method of claim 46 wherein the genetic polymorphism is a causal mutation underlying a quantitative trait locus
48. A method comprising evaluating an animal's genotype at 100 or more genomic locus/loci; wherein at least 100 loci each comprise a single nucleotide polymorphism (SNP) selected from:
- a. the SNPs described in Table 4 and the Sequence Listing; and/or
 - b. a SNP located in a gene described in Table 4;
49. The method of claim 48 wherein said SNP is associated with at least one trait as described in Table 4.
50. The method of claim 48 wherein the animal's genotype is evaluated at 200 or more loci.

51. The method of claim 48 wherein the animal's genotype is evaluated at 500 or more loci.
52. The method of any of claim 48 that comprises whole-genome analysis.
53. A method for allocating an animal for use according to the animal's predicted marker breeding value for productivity and/or fitness, the method comprising:
 - a. determining the animal's genotype at one or more locus/loci; wherein at least one locus contains a single nucleotide polymorphism (SNP), having at least two allelic variants; and wherein at least one SNP is selected from the SNPs described in Tables 5 and 6;
 - b. analyzing the determined genotype of at least one evaluated animal at one or more SNPs selected from the SNPs described in Tables 5 and 6 to determine which allelic variant is present;
 - c. allocating the animal for use according to its determined genotype.
54. The method of claim 53 wherein the animal's genotype is evaluated at 10 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
55. The method of claim 53 wherein the animal's genotype is evaluated at 50 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
56. The method of claim 53 wherein the animal's genotype is evaluated at 100 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
57. The method of claim 53 wherein the animal's genotype is evaluated at 200 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
58. The method of any of claims 53 to 57 wherein SNPs evaluated are associated with fitness.

59. The method of any of claims 53 to 57 wherein SNPs evaluated are associated with productivity.
60. The method of any of claims 53 to 57 that comprises whole-genome analysis.
61. A method for selecting a potential parent animal for breeding to improve fitness and/or productivity in potential offspring:
 - a. determining at least one potential parent animal's genotype at one or more genomic locus/loci; wherein at least one locus contains a single nucleotide polymorphism (SNP) that has at least two allelic variants, and wherein at least one SNP is selected from the SNPs described in Tables 5 and 6;
 - b. analyzing the determined genotype of at least one evaluated animal for one or more SNPs selected from the SNPs described in Tables 5 and 6 to determine which allele is present;
 - c. correlating the identified allele with a fitness and/or productivity phenotype;
 - d. allocating at least one animal for breeding use based on its genotype.
62. The method of claim 61 wherein the potential parent animal's genotype is evaluated at 10 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
63. The method of claim 61 wherein the potential parent animal's genotype is evaluated at 50 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
64. The method of claim 61 wherein the potential parent animal's genotype is evaluated at 100 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.

65. The method of claim 61 wherein the potential parent animal's genotype is evaluated at 200 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
66. The method of any of claims 61 to 65 wherein the potential parent animal is selected to improve fitness in the potential offspring.
67. The method of any of claims 61 to 65 wherein the potential parent animal is selected to improve productivity in the potential offspring.
68. The method of any of claims 61 to 65 that comprises whole-genome analysis.
69. A method of producing progeny animals comprising:
 - a) identifying at least one potential parent animal that has been allocated for breeding in accordance with the method of claim 1;
 - b) producing progeny from the allocated animal through a process comprising:
 - i) natural breeding;
 - ii) artificial insemination;
 - iii) *in vitro* fertilization; and/or
 - ii) collecting semen/spermatozoa or at least one ovum from the animal and contacting it, respectively, with ovum/ova or semen/spermatozoa from a second animal to produce a conceptus by any means.
70. The method of claim 69 comprising producing progeny through natural breeding.
71. The method of claim 69 comprising producing offspring through artificial insemination, embryo transfer, and/or *in vitro* fertilization.
72. The method of claim 69 wherein the potential parent animal's genotype is evaluated at 10 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.

73. The method of claim 69 wherein the potential parent animal's genotype is evaluated at 50 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
74. The method of claim 69 wherein the potential parent animal's genotype is evaluated at 100 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
75. The method of claim 69 wherein the potential parent animal's genotype is evaluated at 200 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
76. The method of any of claims 69 to 75 wherein the potential parent animal is selected to improve fitness in the offspring.
77. The method of any of claims 69 to 75 wherein the potential parent animal is selected to improve productivity in the offspring.
78. The method of any of claims 69 to 75 that comprises whole-genome analysis.
79. A nucleic acid array for determining which allele of at least 10 of the SNPs described by Tables 5 and 6 are present in a sample; wherein the array comprises 10 or more nucleic acid sequences capable of hybridizing, under stringent conditions, with at 10 or more SNPs selected from the group consisting of the SNPs described by Tables 5 and 6.
80. The array of claim 79 wherein the array is capable of determining which allele is present for each of 25 or more SNPs selected from the group consisting of the SNPs described by Tables 5 and 6.
81. The array of claim 79 wherein the array is capable of determining which allele is present for each of 50 or more SNPs selected from the group consisting of the SNPs described by Tables 5 and 6.

82. The array of claim 79 wherein the array is capable of determining which allele is present for each of 100 or more SNPs selected from the group consisting of the SNPs described by Tables 5 and 6.
83. The array of any of claims 79 to 82 for determining which alleles are present for SNPs associated with fitness.
84. The array of any of claims 79 to 82 for determining which alleles are present for SNPs associated with productivity.
85. A method of identifying quantitative trait locus associated with production or fitness by identifying a genetic marker in allelic association with at least one SNP selected from the group of SNPs described in Tables 5 and 6, the method comprising:
- a) identifying a genetic marker B_1 suspected of being in allelic association with a marker A_1 selected from the group of SNPs described in Tables 5 and 6;
 - b) determining whether A_1 and B_1 are in allelic association;
- wherein allelic association exists if $r^2 > 0.2$ for Equation 1 for a population sample of at least 100 animals and wherein Equation 1 is:

$$r^2 = \frac{[f(A_1B_1) - f(A_1)f(B_1)]^2}{f(A_1)(1 - f(A_1))(f(B_1)(1 - f(B_1)))} \quad \text{[Equation 1]}$$

and wherein A_1 represents an allele of a SNP described in Tables 5 and 6; B_1 represents a genetic marker at another locus; $f(A_1B_1)$ denotes frequency of having both A_1 and B_1 ; $f(A_1)$ is the frequency of A_1 in the population; and $f(B_1)$ is the frequency of B_1 in a population.

86. The method of claim 85 wherein the genetic marker B_1 is a SNP.
87. The method of claim 85 wherein the genetic marker identified is in linkage disequilibrium with at least one SNP selected from the group of SNPs described in Tables 5 and 6.

88. The method of claim 85 wherein B_1 is a causal mutation underlying a quantitative trait locus
89. The method of claim 85 wherein $r^2 > 0.5$.
90. The method of claim 85 wherein $r^2 > 0.9$.
91. The method of any of claims 85 to 90 for identifying a genetic marker in allelic association with a SNP associated with fitness.
92. The method of any of claims 85 to 90 for identifying a genetic marker in allelic association with a SNP associated with productivity.
93. The method of claim 91 wherein the genetic marker is in linkage disequilibrium with a SNP associated with fitness.
94. The method of claim 92 wherein the genetic marker is in linkage disequilibrium with a SNP associated with productivity.
95. The method of claim 91 wherein the identified genetic marker is a causative mutation.
96. The method of claim 92 wherein the identified genetic marker is a causative mutation.
97. A method comprising evaluating an animal's genotype at 10 or more genomic locus/loci; wherein at least 10 loci each comprise a single nucleotide polymorphism (SNP) selected from the SNPs described in Tables 5 and 6.
98. The method of claim 97 wherein the animal's genotype is evaluated at 50 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.

99. The method of claim 98 wherein the animal's genotype is evaluated at 100 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
100. The method of claim 97 wherein the animal's genotype is evaluated at 200 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
101. The method of claim 97 wherein the animal's genotype is evaluated at 500 or more loci that contain SNPs selected from the SNPs described in Tables 5 and 6.
102. The method of any of claims 97 to 101 where the SNPs are associated with fitness.
103. The method of any of claims 97 to 101 wherein the SNPs are associated with productivity.
104. The method of any of claims 97 to 101 that comprises whole-genome analysis.