

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
11 August 2005 (11.08.2005)

PCT

(10) International Publication Number
WO 2005/073408 A2

(51) International Patent Classification⁷: **C12Q 1/68**

Mitchell; *. FREIJE, Wadiha; 117 Elgin Ave. #1, Forest Park, IL 60130.

(21) International Application Number:
PCT/US2005/002164

(74) Agent: MARTIN, Alice, O.; P.O. Box 2786, Chicago, IL 60690-2786 (US).

(22) International Filing Date: 24 January 2005 (24.01.2005)

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/538,791 23 January 2004 (23.01.2004) US
60/539,728 26 January 2004 (26.01.2004) US

(71) Applicant (for all designated States except US): PYXIS GENOMICS, INC. [US/US]; 2201 W. Campbell Park Drive, Chicago, IL 60612 (US).

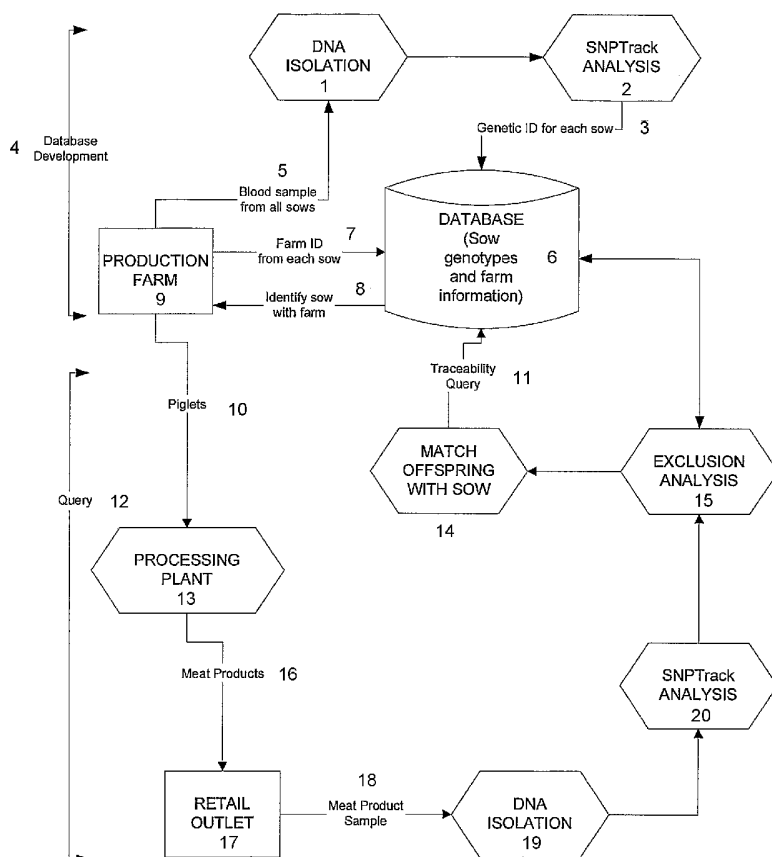
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),

(72) Inventors; and

(75) Inventors/Applicants (for US only): ABRAHAMSON,

[Continued on next page]

(54) Title: SMALL SEGMENTS OF DNA DETERMINE ANIMAL IDENTITY AND SOURCE



(57) Abstract: Methods and compositions to trace an animal to a country or a farm of origin or to identify an animal based on SNPTrack analyses are described. Methods of developing SNPTracks that include a plurality of markers with one or more SNPs and SNPTrack analysis kits are described.

WO 2005/073408 A2



European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, 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).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

- *without international search report and to be republished upon receipt of that report*

SMALL SEGMENTS OF DNA DETERMINE ANIMAL IDENTITY AND SOURCE

[0001] INVENTORS: Mitchell Abrahamsen and Wadiha Freije.

[0002] Short segments of mitochondrial, autosomal, X, and Y chromosomal DNA, are used to identify lineage of individual animals (identity), and to trace their country or farm of origin (traceability).

BACKGROUND

[0003] There are many reasons for determining the identity and source of animals bred as food for humans. The ability to trace back a food sample to the farm or country of origin offers improved quality control, safer food, and can demand a higher price. The ability to rapidly trace lineage back to the sire and dam better localizes the cause of any problem and gives an opportunity to take preventive measures, for example, to minimize the unwanted distribution of contaminated meat. The ability to rapidly trace lineage back to the sire and dam also provides useful information on meat quality in genetic selection decisions to improve meat quality. Human populations may also be monitored with respect to immigration and forensic purposes. Forensic applications include immigration documentation, criminal trials, adoption disputes, and paternity testing.

[0004] Pedigrees are cumbersome to analyze directly and have problems arising from the nature of breeding programs, *e.g.*, because commercial pigs are sired by AI (artificial insemination), the same boar is mated to sows on several different farms. In those situations, the farm can only be tracked via the dam (mother of the slaughtered pig).. Moreover, parents may be dead or unavailable for genetic testing.

[0005] For example, in swine demonstration of the sire line is problematic. A further complication is that, in order to maximize fertility, the semen from two or more boars may be mixed together for use on commercial farms. Piglets from the same litter may therefore have different genetic sires. To overcome these complications, DNA genotyping requires a large number of markers that are developed and characterized specifically in a specific breeding population. Unfortunately, as the numbers of markers increases, costs of testing also generally increase.

[0006] Rapid tracing to the farm of origin of a cattle infected with Mad Cow Disease is necessary to identify and limit the spread of Mad Cow Disease-infected beef in human food chain. Similar investigation is necessary to limit the spread of Scrapie-infected sheep in the human food chain. Current methods of tracing a contaminated meat sample or an infected animal to its country or farm of origin rely on time consuming and laborious procedures, some of which are discussed herein.

[0007] Determining the lineage and source of individual mammals previously relied on the use of restriction fragment length polymorphisms (RFLPs), variable number of tandem repeats (VNTRs), and microsatellite markers. All these known DNA sequence methods have the same disadvantages. They are costly to execute and require highly specialized laboratory settings. In addition, because of their repetitive nature, VNTRs and microsatellite repeats can be affected by errors in the replication process that result in novel alleles that are non-informative.

SUMMARY

[0008] New methods and compositions are provided to determine lineage and trace the source of animals, in particular mammals used for food. For use in forensics and food security, the tests are robust, simple to perform, and have sufficient power to identify closely related individuals. Single nucleotide polymorphisms (SNPs) are abundant and simple to analyze.

[0009] Short segments of mitochondrial, autosomal, X, and Y chromosomal DNA, are used to identify lineage of individual mammals (identity), and to trace their country or farm of origin (traceability). A DNA test for identification uses genetic information to uniquely determine the identity of each animal and to trace the animal to its country or farm of origin.

[00010] To minimize costs and maximize the efficiency of the number of genetic markers needed, short segments of mitochondrial, autosomal, X and Y DNA are used to determine the genetic origin of humans or domestic farm mammals. Genetic origins include family, country, or farm of origin, and lineage. Single nucleotide polymorphisms (SNPs) and/or insertions/deletions in mitochondrial DNA (mtDNA), sex chromosomes, and autosomal loci are useful. Short segments that contain one or more SNPs which map to the mitochondria, the non-recombining portion of the Y chromosome, autosomes, or in any region in the genome are referred to herein as "SNPTracks". Multiple SNPTracks are aspects of the disclosure. The information

content of SNPs and ins/dels which map within 0.2 to 10 kb fragments on mitochondria, autosomes or the sex chromosomes, is combined to generate informative SNPTracks. Each SNPTrack frequency is determined in a reference population of mammals, *e.g.* humans from different ethnic groups or cattle and pigs from various origins. Frequencies of SNPTracks vary in highly inbred populations and therefore should be determined from an experimental population.

[00011] The methods disclosed herein make use of the stability and polymorphisms in short DNA sequences and the minimal amount of DNA needed to reduce the cost and facilitate automated and portable detection of markers, *e.g.*, on a farm instead of in a laboratory.

[00012] A method for identification of an animal includes the steps of obtaining a sample from the animal or from a processed product of the animal; performing single nucleotide polymorphism (SNP) analysis that includes markers, such that the markers include one or more SNPs; generating SNPTracks of the animal such that the SNPTracks contain one or more markers with one or more SNPs; and comparing the SNPTracks of the animal to a database that includes pre-existing SNPTracks to identify the animal. The animal is a farm animal, which includes animals such as cow, pig, sheep, and poultry. The animal can also be a mammal and the mammal may be a human.

[00013] The markers are designed and developed from autosomes, sex chromosomes, and mitochondrial DNA, wherein, if there are more than one SNP per marker, they are present within a nucleotide region of 0.2 to about 10 kb.

[00014] The various swine markers are selected from the group that includes markers designated ACY-STS7; COX2; EG-STS7; GALT; IKBA; LEPR; P450-STS18; PBE3; PBE42; PBE43; PBE 57; PBE59; PBE 64; PBE 73; PBE 84; PBE132; PBE137; PRKAG-STS3; RYRA-STS6; SCAMP; VAN-STS1; WSCR-STS1; BG; AMG; CTSL; PBE112; and MYF5. The SNP positions of these markers are designated as follows: ACY-STS7 (245 G/C 421 C/T); COX2 (368 C/T 533 G/A 939 C/T); EG-STS7 (774 G/A 805 G/A 817 G/A); GALT (478 G/A 758 C/G 866 C/G); IKBA (4476 C/T 4679 T/A 4904 G/T); LEPR (426 A/C/G 810 G/C); P450-STS18 (71 C/T 138 G/A 361 G/A); PBE3 (115 G/A 192 A/T 555 T/G); PBE42 (111 G/A 118 T/G 181 C/T); PBE43 (314 C/T 471 C/A 524 C/T); PBE 57 (75 C/T 109 G/A 197 C/T 268 T/G); PBE59 (276 C/T 494 C/T); PBE 64 (115 G/A 419 C/G 515 C/T); PBE 73

(93 A/C 116 A/G 177 C/T 477 C/T); PBE 84 (14 A/G 97 T/C 428 T/C); PBE132 (102 C/G 127 A/G 193 C/T 371 A/G); PBE137 (121 C/T 278 C/T 409 A/G); PRKAG-STS3 (1845 G/A 1938 G/A 2050 G/A); RYRA-STS6 (402 A/C 408 C/G 567 A/C); SCAMP (184 C/T 389 G/A 516 C/T 582 G/A 939 A/T); VAN-STS1 (889 C/T 950 C/T 1009 G/A 1065 C/T); WSCR-STS1 (411 C/T 599 C/T); BG (1257 C/T 1323 C/T 1425 C/A 1966 C/T); AMG (907 A/C 975 A/G 1467 A/G); LCN (87 C/T 373 G/C 402 C/T); CTSL (252 A/G 272 A/G); PBE112 (61 C/T 87 C/T); and MYF5 (1833 A/G 2204 C/G 2335 A/C).

[00015] The various markers are also selected from the group that includes swine mitochondrial markers designated at positions 15543 (C/T), 15558 (A/T), 15615 (C/T), 15616 (C/T), 15675 (C/T), 15714 (C/T), 15840 (C/T), and 16127 (G/A).

[00016] A system for identifying an animal from which a sample is derived, the system includes the steps of: obtaining a sample from the animal or from a processed product of the animal; performing single nucleotide polymorphism tracking (SNPTrack) analysis of the sample with a plurality of markers, wherein the plurality of markers include one or more SNPs; storing one or more SNPTracks of the sample in a computer system; comparing the one or more SNPTracks of the sample with known SNPTracks in a database; and identifying the animal to a particular location.

[00017] A computer system for identifying a sample, the computer system includes: a software module comprising instructions operative to provide a searchable database that includes SNPTracks obtained from animals, the SNPTracks include one or more markers, wherein the markers include one or more SNPs; a software module that includes instructions operative to provide an algorithm to determine exclusion probabilities; and a software module that includes instructions operative to provide an interface to accept and compare a query SNPTrack with the plurality of SNPTracks in the database.

[00018] A method of developing a database for identifying an animal or a product sample derived from the animal, the method includes the steps of: performing SNPTrack analysis of a plurality of samples obtained from a plurality of animals from one or more sources with a plurality of markers, wherein the plurality of markers includes one or more SNPs; obtaining and storing the SNPTracks of the plurality of samples in a database, wherein the database is searchable; performing SNPTrack analysis of the product sample; comparing one or more SNPTracks of the product

sample with the SNPTracks of the plurality of the samples stored in the database; and identifying the product sample to a location. The database further includes data selected from the group that includes production farm, farm of origin, retail outlet, whole saler, breeding record, animal identification, offspring data, sibling data, and lineage data. The sources are selected from the group that includes farm of origin, production farm, processing center, retail outlet, distribution center, and whole saler. The SNPTracks of the plurality of the samples stored in the database are associated with an animal identification system.

[00019] The SNPTrack analysis is performed on the plurality of samples between birth and slaughter of the plurality of the animals. The database may have limited access to an authorized user.

[00020] A method of obtaining a SNPTrack of a sample, the method includes the steps of: selecting a first marker such that the first marker has one or more SNPs within a 0.2 to about 10 kb region in the genome; selecting a second marker such that the second marker has one or more SNPs within a 0.2 to about 10 kb region in the genome; performing SNPTrack analysis of the sample with the first and second markers; and obtaining the SNPTrack of the sample. A method of obtaining a SNPTrack of a sample further includes the steps of performing a SNPTrack analysis with a plurality of markers and obtaining the SNPTrack for the plurality of markers.

[00021] A nucleotide segment from swine genome that includes a plurality of SNPs, wherein the segment is amplified with primers listed in TABLE 3. The nucleotide segment is selected from the group that includes autosomes, sex chromosomes, and mitochondrial DNA.

[00022] A high-throughput system for tracking an animal and a meat product through a supply chain, the system includes the steps of: performing SNPTrack analysis of a plurality of samples obtained from a plurality of animals; obtaining and storing a plurality of SNPTracks in a database, wherein the database is searchable; obtaining a plurality of samples from meat products; performing SNPTrack analysis of the plurality of samples from meat products; and tracing the plurality of samples from meat products to the farm or the processing plant. The high-throughput system includes samples obtained from the plurality of animals prior to slaughtering in a farm or a processing center. The high-throughput system is capable of identifying and tracing animals and its products from birth to post-slaughter. The high-throughput

system is capable of identifying and tracing a meat product from a consumer to a farm of origin.

[00023] A software module includes instructions operative to provide a database comprising SNPTracks identified in TABLE 7.

[00024] A SNPTrack analysis kit to identify an animal includes: a plurality of oligonucleotides corresponding to a plurality of markers that contain one or more SNPs within a 0.2 to about 10 kb region of each marker; reagents to perform SNPTrack analysis; and access to compare SNPTracks with a plurality of SNPTracks in a database to identify the animal. The kit further includes an instruction manual to identify the animal.

[00025] A method of tracing an infected meat sample to a particular location, the method includes the steps of: performing SNPTrack analysis of a plurality of samples obtained from a plurality of animals with a plurality of markers, wherein the plurality of markers include one or more SNPs; obtaining and storing the SNPTracks of the plurality of samples in a database, wherein the database is searchable; performing SNPTrack analysis of the infected meat sample; comparing one or more SNPTracks of the infected meat sample with the SNPTracks of the plurality of the samples stored in the database; and tracing the infected meat sample to a particular location. The infected meat sample is a beef sample infected with mad cow disease or bovine spongiform encephalopathy (BSE). The infected meat sample is a sheep or goat sample infected with scrapie disease. The infected meat sample is an infected pork sample. The infected meat sample is obtained from a meat product in the market. The particular location can include farm of origin, production farm, processing center, retail outlet, distribution center, and whole saler.

[00026] A method of enhancing food safety and quality assurance of a meat product, the method includes the steps of: performing SNPTrack analysis of a plurality of samples obtained from a plurality of animals prior to slaughtering with a plurality of markers, wherein the plurality of markers include one or more SNPs; obtaining and storing a plurality of SNPTracks of the plurality of samples in a database, wherein the database is searchable; tracing the meat product to its source by performing SNPTrack analysis of the meat product and comparing a SNPTrack of the meat product with the SNPTracks of the plurality of samples stored in the database; and determining if the meat product is safe.

DEFINITIONS

- [00027]** **Allele frequency:** The frequency at which a particular allele (polymorphism) occurs in the members of a population under study.
- [00028]** **Animal identification system:** Information capable of being used to identify or trace a particular animal to pre-existing records. For example, an alpha-numeric code that is associated with a particular SNPTrack to identify a particular animal or a sample derived from that animal.
- [00029]** **Database:** An organized collection of information or data, stored preferably in an electronic computer readable format, that is capable of being updated and queried. The information stored in a database is managed by a database management system, which includes a software mechanism for managing that data. The database and its associated database management system can be accessed over the Internet or by any other electronic means. The database can store information or data including but not limiting to SNPTracks, farm of origin, production farm, processing plant, distribution center, retail outlet, wholesaler, breeding record, commercially valuable trait information, sibling information, offspring information, pedigree analysis, and other animal identification that in an organized and searchable manner.
- [00030]** **Haplotype:** A set of closely linked alleles (genes, genetic loci, or DNA polymorphisms) in a chromosome that is usually inherited as a single recombination unit. Some haplotypes may be in linkage disequilibrium. A haplotype may also be one of a set of single nucleotide polymorphisms along a region of a chromosome.
- [00031]** **High-throughput system:** A technique or a methodology or a platform capable of analyzing a plurality of samples simultaneously or in batches for a specific assay. For example, a plurality of DNA samples derived from blood samples of farm animals can be simulataneously analyzed for SNPs to obtain SNPTracks using a set of pre-defined assays.
- [00032]** **Identity:** A unique genetic identification of an individual by analyzing certain biological characteristics such as the DNA sequence, SNPTracks, and other genetic markers.
- [00033]** **Identification:** A method of determining a genetic identity of a mammal or a sample derived from a mammal, based on a comparison of SNPTracks with pre-existing SNPTracks of that mammal in a database (identity determination). Identification also includes traceability/tracing analysis that involves, for example

determining the farm of origin or country of origin based on a comparison of SNPTracks obtained from a mammal or a sample derived from a mammal with pre-existing SNPTracks in a database obtained from matings pairs, maternal or paternal breeding populations. Identification of a mammal therefore involves determining the identity based on a unique SNPTrack and also tracing the mammal to a source or location based on SNPTracks of its maternal or paternal breeding populations.

[00034] Lineage: Genetic ancestry; Line of descent of the descendants from an original source of parentage.

[00035] Location: A place where a sample can be traced back for identification purposes. A location can include a country, farm, production farm, processing plant, distribution center, retail outlet, wholesaler, or any other geographical territory.

[00036] Marker: A biomolecule that is capable of distinguishing biological samples. A marker can be a sequence of nucleotides.

[00037] Product: A portion of an animal, generally after slaughter, including processed meat sample that is available in a chain of commerce such as for example, a beef product at a grocery store. A processed meat sample includes any meat sample that is obtained post-slaughter.

[00038] Sample: Any material that can be analyzed to determine identity or traceability. Samples include processed meat samples, skin, blood, hair, bodily fluids or any other biological material obtained from a dead or live mammal. Sample also includes DNA or other genetic material that can be used for genotyping, haplotype determination, and SNPTrack analysis.

[00039] SNP: Single Nucleotide Polymorphism in a nucleotide sequence that includes insertions, deletions, and substitutions when compared between two or more members of a population. SNPs may be in the coding, non-coding, introns, exons, and the regulatory regions of DNA or RNA derived from mitochondrial, autosomal and sex-chromosomes.

[00040] SNPTrack: A SNPTrack includes a plurality of markers such that each marker includes a plurality of SNPs within a 0.2 to 10 kb interval in any region of a chromosome or mitochondrial DNA that may be inherited as a single unit during recombination if recombination occurs. The set of SNPs in a SNPTrack may also be in linkage disequilibrium, wherein the SNPs are linked.

- [00041] **System:** An organized assembly or platform of components, resources, materials, tools, equipments, procedures or methods or processes or operations, software, interacting and functioning in a unified way to perform a specific function.
- [00042] **Traceability/Tracing:** A methodology to track a mammal to its farm of origin or country of origin or any other location by a genetic analysis. For example, a test meat sample or a test cattle may be traced to its farm or country of origin by excluding the sows that cannot be the mother of the test sample through an analysis of single nucleotide polymorphisms.

BRIEF DESCRIPTION OF DRAWINGS

- [00043] FIG. 1 is a database development scheme and the traceability of a test meat sample to its farm of origin using the database are shown (query). The schematic illustrations represent an identification scheme based on sow parentage.
- [00044] FIG. 2 is a schematic representation of an identity determination model of a piglet. A database development scheme and the identity assay of a test meat product sample from a piglet using the database (query) are shown.
- [00045] FIG. 3A shows a schematic representation of an identity/traceability determination scheme for a meat sample involving wholesalers and farms.
- [00046] FIG. 3B shows a flowchart of the various steps and instructions for identification of a meat product sample.
- [00047] FIG. 4 is a representation of determining SNPTracks based on a two SNP example.
- [00048] FIG. 5 demonstrates determining SNPTracks based on a three SNP example (FIG. 5A); for an offspring A (FIG. 5B); and for an offspring B (FIG. 5C).

DETAILED DESCRIPTION

- [00049] New methods and products designed to replace and/or complement existing methods to determine the identity (lineage) of an individual animal as well as traceability of the source, *e.g.* farm or country of origin are disclosed. Previous methods relied on the use of RFLPs, VNTRs and microsatellite markers. In contrast, present methods rely on the use of small DNA fragments that include single nucleotide polymorphisms (SNPs) in the mitochondrial genome, the autosomes, and the sex chromosomes. It relies further on the power of combining information from

multiple SNPs which map within 0.2 to 10 kb of each other to identify the various SNPTracks present in the population and to determine their allele frequency. In an embodiment, pig genomic DNA was used to validate the method of using multiple SNPTracks and prove its utility in traceability and in determining identity. A number of swine breeds were used for this analysis including Duroc, Landrace, and Large White. However, the methods are suitable for all animals including mammals.

[00050] Because farms share semen for pig breeding, the farm of origin can only be identified based on DNA from the sows. Therefore, animals may only be traced to their farm of origin using the DNA inherited from their mothers. Identification of genetic markers (lineage specific, X-chromosome specific and autosomal,) with sufficient information content to allow the unambiguous identification of each animal is an aspect of the present invention.

[00051] Lineage markers are markers that are inherited from the mother or the father. For example, they are derived from the mitochondrial genome or the Y chromosome. The mitochondrial genome is maternally inherited, which means that all offspring inherit their mitochondria only from their mothers. There are a number of known polymorphisms in the mitochondrial genome including a region which exhibits a higher than expected level of variation in about 1 kb of DNA (D-loop region). Mitochondrial SNPs may not have the power to identify each animal, but they allow grouping the sows into several genetically related sets based on the polymorphisms of their mitochondrial DNA, and to tailor the development of additional markers to supplement a traceability methodology.

[00052] The Y chromosome is passed on from father to son. Male progeny inherit an exact copy of the Y chromosome from their father. Because it does not undergo recombination outside of the pseudoautosomal regions, it behaves as a locus similar in its inheritance to the mitochondrial genome. The Y has accumulated a number of sequence variations due to errors in DNA replication during evolution. Such variations can be scored as SNPs and help group the boars into several genetically related groups for further marker development.

[00053] SNPs represent single nucleotide variations at specific chromosomal locations. SNPs were determined by direct sequence analysis following amplification from genomic DNA. They range in frequency from very rare (<1%) to the very common (49%). The distribution and the allele frequency of SNPs may vary between breeds.

Therefore, it is important to develop and validate SNP markers in a specific population under study.

[00054] Using SNPs in this type of analysis is to identify sequence variations that are relatively common in the study group. Even “common” SNPs usually have only two alleles limiting the genetic information content of each marker and requiring the genotyping of a large number of SNPs to provide for the level of confidence needed to identify an individual animal. For that reason, it is useful to combine information from multiple SNPs that reside physically in close proximity (within 0.2 to 10 kb in the DNA sequence). Such markers are unlikely to be separated by recombination during mating, and their genetic and physical location information are combined to generate a SNPTrack. The power of each SNP or a combination of SNPs to identify an individual in a population is determined statistically based on the data generated using a population under study.

[00055] Because the traceability of each animal will depend heavily on the sows, it is advantageous to emphasize X-chromosome markers in addition to the mitochondrial SNPs. Male progeny inherit their X chromosome from their mothers, while female progeny inherit one X chromosome from each parent. The mode of inheritance of X specific SNPs allows a more precise allele assignment increasing the power of each marker. Therefore, it is useful to develop a large number of X-chromosomes SNPs.

[00056] Additional markers are developed from autosomal genes to be used when needed to identify an individual or trace its farm of origin. The autosomal SNP markers are identified using the same strategy used for the X specific markers. For example, the mitochondria and the non-recombining portion of the Y represent two useful genetic areas. In an embodiment using pigs, sequencing of the mitochondria was not limited to the D-loop region, but SNP detection was performed in all mitochondrial genes. Additional markers were selected from known X chromosome genes including the amelogenin gene, and the androgen receptor gene. Autosomal markers were selected from mostly unlinked loci on various chromosomes. They included an IL-2 receptor, an obesity gene (leptin), and a fatty acid binding protein.

[00057] In an embodiment, additional DNA sequences for SNP detection were derived from pig DNA contained in bacterial artificial chromosomes (BACs) from the SRY locus on the Y chromosome and three X chromosome-specific genes. The BACs

were subcloned to generate novel DNA sequences that were used to identify SNPs from different breeds of pigs.

[00058] Efforts focused on the identification of short DNA fragments (0.2 to 10 kb) which harbor multiple SNPs. The polymorphic information of multiple SNPs was combined to generate SNPTracks. The SNPs on individual chromosomes was determined using a number of methods. For example, the SNPTracks could be determined from homozygous individuals, subcloning and sequencing of individual haplotypes, pedigree analysis, and amplification using allele-specific nucleotides followed by sequencing. X chromosome markers can be determined by analyzing male animals because they are haploid for the non-recombining regions of the sex chromosomes.

[00059] One of the differences between traceability and identity depends on which animals are genotyped and entered into a database. For example, traceability relates to the sows (mothers) and identity relates to the actual animals that are slaughtered (*e.g.*, piglets or offsprings). In a traceability approach, a genetic identification of the meat sample is used to locate the mother by excluding who cannot be the mother (FIG. 1). In identity, a matching of genetic identification of the meat sample with those in the database is performed (provided SNPTrack analysis was performed for the animal before slaughter) to find that same genetic identification in the database.

[00060] Reduced heterozygosity in highly inbred populations limits a polymorphism's value. However, the methods that are used currently suffer from similar disadvantages. This limitation can be overcome by analyzing additional loci to increase the possibility of identifying polymorphic markers in the population under study. In farm animals, it is difficult to determine the haplotype frequencies with absolute certainty because of the number of breeds involved and the breeding practices of farm animals. This advantage can be overcome by selecting highly informative markers in most breeds and determining the haplotype frequencies in a relevant sample of the population to be analyzed.

[00061] A method of identifying an animal includes the steps of (1) obtaining a sample from the animal (for example blood sample from a cow prior to slaughter) or from a processed product (for example, a beef product in the market) of the animal; (2) performing single nucleotide polymorphism (SNP) analysis that includes a one or more markers, such that the markers include one or more SNPs—SNP analysis can

be performed, for example, by isolating DNA from the sample, followed by PCR amplification using marker specific primers (*e.g.*, listed in Table 3) and sequencing to determine the base pairs; (3) generating SNPTracks of the animal such that the SNPTracks contain one or more markers with one or more SNPs (SNPTracks combine the SNP data from multiple markers from (2)); and comparing the SNPTracks of the animal to a database that includes pre-existing SNPTracks to identify the animal (the database has been previously created using similar markers and performing SNPTrack analysis, for example, with samples obtained from animals in a farm). The SNPTracks can be generated by combining the individual SNP data from a plurality of markers in to one or more track that contains a string of SNPs from different regions of the genome (for example, a sample spectrum of genetic regions analyzed for developing swine markers is listed in Table 4).

[00062] The markers are designed and developed from autosomes, sex chromosomes, and mitochondrial DNA, wherein the SNPs in a marker are present within a nucleotide region of 0.2 to about 10 kb. Other genetic segments that span regions larger than 10 kb and shorter than 0.2 kb are also within the scope of this disclosure.

[00063] A method of genotyping or performing SNPTrack analysis to identify and/or trace a meat sample to a farm of origin is illustrated in FIG. 3A. In an illustration, a company, PYXIS performs a SNPTrack analysis to determine the genotype of a meat sample provided by a customer or obtained through any other source. The meat sample can include fresh or processed samples from pig, cows, and sheep. DNA is isolated from the sample using standard DNA isolation procedures and SNPTrack analysis is performed with a set of markers, such as, for example markers from Table 7.

[00064] The results of the SNPTrack analysis are queried against databases developed and maintained by wholesalers. For example, the genotype data obtained through SNPTrack analysis by PYXIS, are compared against pre-existing SNPTrack data in the databases developed by Wholesaler 1, 2 and 3. In the illustrated model in FIG. 3A, PYXIS has limited access to search and compare for matches in the Wholesaler databases. If the Wholesaler 1 database has a match (or no match) to a query by PYXIS, the Wholesaler 1 database will return an appropriate search result, *e.g.*, "Match" or "No Match".

[00065] The Wholesaler databases will have identification records to trace a particular meat sample to a farm of origin or to another upstream source such as another wholesaler. In the illustrated model in FIG. 3A, a Wholesaler can track a meat sample (if there is a “Match”) to a particular farm of origin. The wholesaler may inform the farm of origin and take appropriate measures to insure safety of the meat products that may contain an infected or defective meat sample that was tested.

[00066] In the model illustrated in FIG. 3A, the Wholesaler databases were developed by performing SNPTrack analysis of meat samples (sample collected from either slaughtered or prior to slaughtering) and the mating population from various farms. Thus, the databases may contain haplotype data (SNPTracks) of slaughtered meat samples, meat samples prior to slaughtering or culling, and mating population (breeding animals). PYXIS performs SNP analysis for the samples; develops the SNPTracks; and provides the software module to search and compare Wholesaler databases in a limited way. PYXIS also provides integrated product development solutions wherein a Wholesaler or a farm develops and independently maintains genetic identification databases based on SNPTracks for animals used in the food chain. Thus, PYXIS assists in the development of searchable databases of SNPTracks that are independently maintained by a larger user such as a Wholesaler and also provides an integration platform wherein a smaller user can have its sample analyzed and traced/identified to a farm of origin. PYXIS acts as an intermediate service provider to enable meat samples to be genotyped (or analyzed using SNPTracks) and identified or traced to a particular farm of origin. The larger user such as a Wholesaler independently maintains the database, thus insuring confidentiality of the breeding records and other valuable information.

[00067] In an embodiment, the PYXIS develops and maintains the Wholesaler databases in a single database management system. PYXIS maintains confidentiality among multiple Wholesaler databases and provides SNPTrack analysis to identify a meat sample. Thus, confidentiality is maintained among the multiple Wholesaler databases, especially for commercially valuable and proprietary information.

[00068] The method shown in FIG. 3A may be performed in connection with a software module as generally depicted in FIG. 3B. The term “computer module” or “software module” referenced in this disclosure is meant to be broadly interpreted and cover various types of software code including but not limited to routines, functions,

objects, libraries, classes, members, packages, procedures, methods, or lines of code together performing similar functionality to these types of coding. The components of the present disclosure are described herein in terms of functional block components, flow charts and various processing steps. As such, it should be appreciated that such functional blocks may be realized by any number of hardware and/or software components configured to perform the specified functions. For example, the present disclosure may employ various integrated circuit components, *e.g.*, memory elements, processing elements, logic elements, look-up tables, and the like, which may carry out a variety of functions under the control of one or more microprocessors or other control devices. Similarly, the software elements of the present invention may be implemented with any programming or scripting language such as C, C++, SQL, Java, COBOL, assembler, PERL, or the like, with the various algorithms being implemented with any combination of data structures, objects, processes, routines or other programming elements. Further, it should be noted that the present disclosure may employ any number of conventional techniques for data transmission, signaling, data processing, network control, and the like as well as those yet to be conceived.

EXAMPLES

Example 1: Determining Traceability and Identity and of Pigs

[00069] An objective of this example was to develop a genetic test to trace fresh and processed pork products back to the farm of origin, and to verify that the product was indeed from the farm stated to be the origin. A second objective was to trace the product back to the parent boar and sow, and thus from parentage records to the grandparents in the pure nucleus herd populations. The ability to determine parentage provides breeders with power to combine the information from genetic lineage and physical attributes to select animals with preferred traits for breeding programs and to eliminate animals responsible for poor quality.

a) Identification of SNPTracks

[00070] SNPTracks were determined by manually comparing the DNA sequence (0.2 - 10 kb) from the same genetic region (locus) across 60 different animals representing some of the major breeds used in pork production such as, for example, Duroc,

Landrace, and Large White. A set of 100 to 200 SNP markers were identified based on the differences in the nucleotide sequence within the 0.2 to 10 kb region of DNA in either autosomal, sex chromosomes, or mitochondrial DNA. The sequences were available either in a proprietary database or were obtained by direct sequencing of desired regions. Differences (insertions/deletions/substitutions) among the DNA sequence were identified as SNPs. Approximately 100 genetic regions (markers) were evaluated for the presence of SNPs by comparing the sequences of each region from the 60 different animals. The selection of which genetic regions (markers) to be included in the test was accomplished by determining which markers were actually polymorphic (*i.e.* contained SNPs) in the target production population. Some of the markers that were polymorphic in the original 60 animals, were not polymorphic in the target population and thus were excluded. A set of 20 markers, each with a SNPTrack composed of 2 or more SNPs (a total of 60 SNPs), is based on the exclusion power predicted by theoretical calculations on the chance of miss identifying an unrelated animal based on chance (see TABLE 6).

b) Determining allele frequencies and minimizing the number of SNPTracks needed for identity or traceability studies

[00071] The most useful SNPs are those that are frequently represented in a population. A single SNP has two alleles. A SNP is most useful if the two alleles are present at equal frequency. Most SNPs have two alleles with frequencies between 20 and 40%. Combining multiple SNPs spanning 0.2 to 10 kb facilitates segregation as a single locus with 5 or 6 alleles (since it is unlikely that they will ever be separated by recombination). For a hypothetical sequence GGGAATATTTATTACCTAT(G/C)TTATATTGGA, allele 1 is GGGAATATTTATTACCTAT(G)TTATATTGGA (50%) and allele 2 is GGGAATATTTATTACCTAT(C)TTATATTGGA (50%). An ideal situation is where allele 1 and allele 2 are present at equal frequencies (*i.e.* 50%) A second SNP is identified (that occurred by random mutation). Since this arose after the first SNP, it is only present in one of the alleles.

[00072] The second SNP is denoted as GGGAATATTTATTACCTAT(C)TTATA(T/C)TGGA. Allele 1 remains the same GGGAATATTTATTACCTAT(G)TTATA(T)TGGA (50%). However, the original allele 2 is now either GGGAATATTTATTACCTAT(C)TTATA(T)TGGA (allele 2;

25%) or GGGAATATTTATTACCTAT(C)TTATA(C)TGGA (allele 3; 25%). Together (allele 2 and allele 3), the frequency is 50%. If the second SNP is present at equal frequencies, then the overall frequency of the 3 alleles is 50%, 25%, and 25%. A value of 50%, 30%, 20% reflects empirically determined data.

[00073] In the example above, haplotypes are assembled based on the combination of SNPs (SNPTrack) at each genetic region (locus/marker). The three SNPTracks above are GT, CT, and CC. If a different genetic region had the SNPTracks CT, AT, AG the nine possible haplotypes would be GT/CT, GT/AT, GT/AG, CT/CT, CT/AT, CT/AG, CC/CT, CC/AT, CC/AG, wherein the first SNPTrack represents one genetic region and the second SNP track represent a different genetic region. The SNPTracks and the approximate frequencies were determined by comparing the DNA sequences of the same genetic region from 60 different animals. The actual allele frequencies may be different for any given population and may change over time. The value of the markers in the prediction of parentage depends on their frequency, which then governs the total number of markers required for analysis.

[00074] An informative list of the characteristics of short amplified fragments (amplicons) with SNPs is shown in Table 3 and a description of genetic regions examined for SNPs is shown in Table 4. A marker is informative if there are multiple alleles present. In the example above, the informative marker was the one in which the 3 alleles were present in the boar and sow population. It could also be the case that in the target boar and sow population there was only a single allele represented. This is determined by directly determining the DNA sequence at the SNP positions in the DNA isolated from the target boar and sow animals. (*i.e.* empirically determined) SNPTracks are composed of 2 or more SNPs that are identified within a genetic region of approximately 0.2 to 10 kb. The allele frequencies of some of the short amplified fragments (amplicons) of Table 3 are shown in Table 4. Table 5 shows SNPTracks and allele frequency distribution for some of the amplicons generated by the primers listed in Table 3. A segregation analysis of the markers across the study population was done.

c) Developing a database with SNPTrack data for sows and sires from various farms

[00075] Using the set of markers identified Table 3, SNPTrack analysis was performed. The DNA obtained from each of the sow and sire was subject to SNP analysis using

the oligonucleotide primers described in Table 3 and subsequent SNPTracks were identified. The data obtained from this analysis were used to develop a SNPTrack database that contained unique SNPTracks for each of the sows and sires that were genotyped with the set of markers identified in Table 5 (see FIG. 1).

d) Validation of the SNPTrack data in a sample population

[00076] A validation assay for a sample of piglets was performed. For example, a sample of 2000 piglets representing 200 piglets per farm may be used for validation. A minimum of 10 commercial farms with 200 piglets per farm and a minimum of 30 dams per farm may be used. A minimum of 8 different sires per farm (same sire may be used on more than one farm) and on each farm, a minimum of 6 sows to be mated with an un-mixed semen from a single sire was used. The sire and dam of each litter in the study were recorded, along with the date of birth and farm. For mixed semen, the identities of all the contributing boars were listed. An ear tagging system was used to identify all study animals (piglets) with a unique number and an ear tissue sample for DNA extraction and for subsequent SNPTrack analysis was provided. All records included the unique identification number of the ear tag or any other suitable identification system. SNPTrack data from the piglets derived from the various farms were queried into the SNPTrack database with no prior knowledge of the farm of origin. The SNPTracks from the sample population were used to validate the SNPTrack database.

[00077] For a field study, however, not all animals need to be identified using the same set of markers. The exact marker set used is tailored for each animal tested to minimize the number of markers needed and to reduce the cost of testing 100 to 200 markers. A final outcome may be a set of markers that will be placed in groups of 5 to 8 on a branched tree. The markers used to identify or trace each animal may depend on the results of the first set of markers analyzed and so on. The grouping of the markers was done statistically based on the data generated to minimize the number of markers needed to trace each animal.

e) Testing a meat sample to trace the farm of origin

[00078] The DNA from a meat sample to be tested was extracted and SNP analysis was performed for markers identified in TABLE 5. The resulting SNPTracks were

queried in the SNPTrack database to trace the farm of origin. Based on the exclusion probabilities, the meat sample is traced to its farm of origin (see FIG. 1)

Example 2: Kits to determine identity or farm of origin

[00079] Kits to determine identity or farm of origin includes oligonucleotide primers for a set of SNP markers, suitable buffers, enzymes and any other biochemical components necessary to perform SNP analysis. A database enriched with SNP marker analysis of breeding animals from various farms is useful in determining the results obtained using the kits disclosed herein. For example, oligonucleotides, whose sequences are described in TABLE 3, is provided in a multi-well high-throughput format for SNP analysis along with suitable buffers and enzymes. PCR amplification followed by direct sequencing or any other form of SNP detection are implemented to develop SNPTracks for any given sample. The SNPTracks are then used to identify the sample or trace the sample's farm of origin.

Example 3: SNPTrack Application in Humans

[00080] The human SNP database contains over a million SNPs. Current validation has focused on sequence variation within genes. These could be within coding sequences or in the 5' and 3' untranslated region. SNPs within human genes also help identify SNPs in the pig homologs because they identify regions within genes that tolerate sequence variations.

[00081] Current SNP cataloging in humans have focused on disease association. The methods disclosed herein are helpful in tracing humans to their country of origin for immigration- and for developing SNPTrack databases for security-related purposes. These methods are also helpful in reconstructing genealogical trees.

Example 4: Three-tier Searching Approach for Pork Traceability assay.

[00082] Some of the DNA matching procedures include 1) mitochondrial matching test ; 2) mating-sample DNA matching test using available mating information; 3) parent-sample DNA matching test, independent of mating information. These activities are independent procedures that are conducted simultaneously during the matching process.

[00083] The 'mating-sample DNA matching test' is a novel design to trace the sample to a specific location based on the mating-pair information. The 'parent-sample DNA matching test' is a paternity test.

[00084] The mitochondrial DNA matching test (MT test) involves a simple matching of mitochondrial genotypes to identify sows with the same mitochondrial genotype as the query sample.

[00085] The mating-sample DNA matching test (MS test) involves exhaustive DNA matching against each known mating pair. SNPTracks of various markers obtained for a particular sample are compared against a database populated with SNPTracks obtained from various mating pairs (breeding population). This test attempts to answer the question whether a particular sample came from an offspring of the mating pair. The sample is excluded if its DNA profile (SNPTrack) is incompatible with that of any mating pair. The mating-sample DNA matching test (MS test) possess higher exclusion power than paternity testing ($E = 0.28125$, $Q_1 = 3/16 = 0.1875$, $Q_0 = 1/8 = 0.125$). The MS test requires about 63% of markers to achieve same exclusion power as paternity testing with a known sire and requires about 41% of markers to achieve same power as paternity testing without known sire. Implementation of the MS test requires a database of breeding records and marker genotypes (SNPTracks).

[00086] Derivation of exclusion probability under the MS test is shown in Table 1. Table 1 shows the derivation of exclusion probability (E) of the MS test assuming bi-allelic markers. Assuming equal allele frequency, $E = 0.28125$. In comparison, the exclusion probability for paternity testing is $Q_1 = 0.1875$ if the sire is known and $Q_0 = 0.125$ if the sire is unknown. For the MS test, heterozygous offspring and parents all contribute to the exclusion power. For paternity test, heterozygous disputed parentage does not contribute to exclusion power.

[00087] Table 1. Exclusion probability (E) for mating-sample DNA matching test.

Dam	Sire	Mating freq	Excluded sample		Exclusion probability
			Genotype	Frequency	
AA	AA	p^4	Aa, aa	$1 - p^2$	$p^4(1 - p^2)$
AA	Aa	$(2pq)p^2$	aa	q^2	$(2pq)p^2 q^2$
AA	aa	$p^2 q^2$	AA, aa	$1 - 2pq$	$p^2 q^2(1 - 2pq)$

Aa	AA	$(2pq)p^2$	aa	q^2	$(2pq)p^2q^2$
Aa	Aa	$(2pq)^2$	-	0	0
Aa	aa	$(2pq)q^2$	AA	p^2	$(2pq)p^2q^2$
aa	AA	p^2q^2	AA, aa	$1-2pq$	$p^2q^2(1-2pq)$
aa	Aa	$(2pq)q^2$	AA ,	p^2	$(2pq)p^2q^2$
aa	aa	q^4	AA, Aa	$1-q^2$	$q^4(1-q^2)$
Sum		1			E

[00088] $E = p^4(1-p^4) + q^4(1-q^2) + 4p^2q^2(1-pq)$
 $= 9/32 = 0.28125$ if $p=q=1/2$. 3)

[00089] Possible genotypes for two bi-allelic markers based on direct matching of meat sample to all possible mating progeny (Mating test) is shown in Tables 2A and 2B.

Table 2C shows the genotypes of Sow 1 and Boar 1 for Markers 1 and 2.

Table 2A:

Sow	Marker 1			Marker 2		
1	C/C	G/G	C/A	G/C	T/A	A/A

Table 2A:

Boar	Marker 1			Marker 2		
1	C/A	G/A	C/A	C/C	A/A	G/A

[00090] Potential alleles from mating between Sow 1 and Boar 1 (mitochondrial markers are excluded) is shown in Table 2C. One allele comes from Sow 1 and one allele comes from Boar 1.

Table 2C:

Marker 1			Marker 2		
C/C	G/G	C/C	G/C	T/A	A/A
C/A	G/A	C/A	C/C	A/A	A/G
		A/A			

[00091] For Marker 1 there are 12 possible genotypes of the offspring

- | | | | |
|-----|-----|-----|-----|
| 1) | C/C | G/G | C/C |
| 2) | C/C | G/G | C/A |
| 3) | C/C | G/G | A/A |
| 4) | C/C | G/A | C/C |
| 5) | C/C | G/A | C/A |
| 6) | C/C | G/A | A/A |
| 7) | C/A | G/G | C/C |
| 8) | C/A | G/G | C/A |
| 9) | C/A | G/G | A/A |
| 10) | C/A | G/A | C/C |
| 11) | C/A | G/A | C/A |
| 12) | C/A | G/A | A/A |

[00092] For Marker 2 there are 8 possible genotypes of the offspring

- | | | | |
|----|-----|-----|-----|
| 1) | G/C | T/A | A/A |
| 2) | G/C | T/A | A/G |
| 3) | G/C | A/A | A/A |
| 4) | G/C | A/A | A/G |
| 5) | C/C | T/A | A/A |
| 6) | C/C | T/A | A/G |
| 7) | C/C | A/A | A/A |
| 8) | C/C | A/A | A/G |

[00093] Considering only Markers 1 and 2, there are 96 possible (8x12) genotypes for their offspring. The meat sample genotype would be compared against these 96 possibilities to identify an exact match.

[00094] The parent-sample DNA matching test (PS test) involves exhaustive DNA matching against each potential parent in the absence of mating information. This test answers the question whether a disputed parent is the true parent of the known offspring. This test is implemented in the absence of any mating information. Therefore, knowing the sire significantly improves the power in identifying the dam.

[00095] These three tests may be performed sequentially or simultaneously. But care should be taken not to exclude the right mother because of a genotyping error.

Example 5: Determining SNPTracks—2 and 3 SNP Haplotypes

[00096] Determination of SNPTracks based on 2 or 3 SNP haplotype examples is illustrated in FIGS. 4, 5A-C. In FIG. 4, the population has 2 SNP allele at positions 1 and 2 of the SNPTrack designated “TG” respectively. The male and female symbols refer to the respective copy inherited from the father and the mother respectively.

Each copy is shown as complementary double stranded DNA. For example, the original allele as indicated by reading the top strand is “TG” and is “AC” as indicated by reading the bottom strand. Therefore, depending upon which strand is read during the haplotype determination, the SNPTracks may vary because of base complementarity. The “X” denotes any intervening base between the SNPs. The “population” refers to a representative sample from the general population of a specific group of animals. The “founder animal” refers to an original animal that has a specific SNPTrack. As DNA is doubled-stranded, assays can be designed to detect the SNP on either strand. Therefore a SNP that is identified as a T/C, could also be detected as a A/G on the complementary DNA strand. Due to technical issues related to SNP detection technologies, an SNP assay may be designed to detect the complementary SNP rather than the indicated SNP.

[00097] In the illustrated example in FIG. 4, the “founder animal 1” has a mutant allele “GG” inherited from the father and the original allele “TG” inherited from the mother. The alleles denoted herein, unless specified otherwise, refer to the top strand in a double stranded DNA sequence. Founder animal 1 has two alleles—the original allele “TG” and the mutant allele “GG”. There can be another “founder animal”, such as, for example “founder animal 2” that has a second mutant allele “GA” inherited from the father and the original allele “TG” inherited from the mother. Therefore, the two founder animals have three alleles designated “TG/GG/GA”. These three alleles give rise to six possible genotypes as illustrated in FIG. 4. The genotype data for these three alleles are also illustrated in FIG. 4. For example, for the TG/GA heterozygous allele, the genotype data will be T/G and G/A at positions 1 and 2 of the SNPTrack respectively. The genotype data or the SNPTrack determination is unique for a specific allele pair. Thus, two founder animals with a total of three alleles for a SNPTrack that includes 2 SNPs, there are six possible genotypes that can be determined by a SNPTrack analysis. Offsprings generated between these founder animals, assuming the founder animals 1 and 2 are of opposite sex, will have one of the six possible genotypes. Because the possible genotypes of a sample can be predicted if the SNPTracks of the father and the mother are known, genotyping errors including sequencing errors can be corrected or filtered off from affecting the SNPTrack analysis. General assumptions in the haplotype or SNPTrack determination model discussed above include that the mutation events are independent; SNPs are

close enough that recombination does not happen; the SNP at position 1 is always linked to the SNP at position 2. SNPs that are within 0.2 to 10 kb are assumed to segregate together and are considered linked.

[00098] A three SNP example is illustrated in FIGS. 5A-5C. In FIG. 5A, the three SNPs at positions 1, 2 and 3 are designated as “TGA”—the original allele in the population. The general descriptions for “founder animal”, “population” and other notations and nomenclature are the same as described for the 2 SNP example in FIG. 4. The founder animal 1 has an original allele TGA inherited from the mother and a mutant allele GGA inherited from the father. The founder animal 2 has 2nd mutant allele GAA inherited from the father and the original allele TGA inherited from the mother. The three alleles from these founder animals 1 and 2 are designated “TGA/GGA/GAA”. The six possible genotypes derived from these three alleles are designated in FIG. 5A. These include three homozygous and three heterozygous genotypes.

[00099] In FIGS. 5A and 5B, founder animal 3 has a 3rd mutant allele GAC inherited from the father and an original allele TGA inherited from the mother. Therefore, among the founder animals 1, 2, and 3, there are four different alleles—one original allele and three mutant alleles. There are ten possible genotypes derived from these four alleles as illustrated in FIG. 5B. These include 4 homozygous and 6 heterozygous genotypes. In FIG. 5B, alleles TAA/TAC/TGC/GGC are shown not to exist as an illustration to demonstrate the predictive power of SNPTrack analysis. Because these alleles do not exist among the ten genotypes derived from the founder animals 1, 2, and 3, the offsprings from these founder animals also cannot have any of those alleles. The genotype data can therefore predict the exact allele combination to trace an offspring to a parent or to a particular location depending upon the database records. FIG. 5C is an illustration to demonstrate the power of SNPTrack analysis to identify parent genotypes based on the genotype data of sample offsprings (A) and (B). In FIG. 5C, the ten possible parent genotypes based on the four alleles (1. CAG; 2. GTG; 3. GAG; 4. GAC) are designated as A through J respectively. For example, for an offspring with the SNP designated by a genotype CAG, there are 4 possible matching parent genotypes (A, E, F, G) under both the SNP/SNP Match column (comparison done by a SNP match, that is if there is one matching SNP) and the SNPTrack Match column (comparison done by identifying a matching SNPTrack, that is all three SNPs

must match). However, for an offspring with a genotype GAG, under the SNP Match column, there are 7 possible parent genotypes (C, E, F, G, H, I, J), whereas under the SNPTrack Match column, there are only 4 possible parent genotypes (C, F, H, J). Thus, SNPTrack analysis and exclusion is more powerful than convention SNP/SNP Match analysis. SNPTrack analysis in this example reduces the number of potential parents for further exclusion analysis.

[000100] Some of the general assumptions and observations in determining the SNPTrack analysis based on the three SNP example include independent mutations; SNPs are close enough that recombination does not occur; and new mutations occur only on a single previous allele.

[000101] FIGS. 4-5C illustrate SNPTrack determination and genotype analysis for two SNP and three SNP models. SNPTracks that have more than 3 SNPs and SNPTracks that include a plurality of 2 or 3 SNP haplotypes can also be designed and developed. For example, the dataset in Table 8 demonstrate the power of combining a plurality of 2 or 3 SNP haplotypes in developing SNPTracks based on approximately 15 markers from the pig genome.

[000102] Data shown in Table 8 illustrate a SNPTrack analysis performed with samples derived from a group of pigs that included mothers and their offsprings. The results of the SNPTrack analysis is shown in Table 8. Under the Animal I.D. column, for example, MLV1 represents the mother and the four following designations MLV1-P1...P4 represent the 4 different offsprings. SNPTrack analysis was performed with a set of about 15 markers listed in Table 8. Positions indicated with "F" represent assay failures and were not included in the exclusion analysis. By comparing the SNPTracks of the offsprings against the SNPTracks of their mothers illustrate the power of SNPTrack analysis to identify the correct mother and eliminate the incorrect mothers. When the SNPTrack of a mother, such as, for example, MLV1 is known, an offspring such as MLV1-P1 can be identified and traced through its mother by comparing the SNPTrack obtained from MLV1-P1 with the SNPTracks stored in a database that also includes the SNPTrack obtained from MLV1, the mother. If the database includes the SNPTrack of MLV1-P1 itself (obtained and stored previously), then the offspring can be uniquely identified and traced to a particular location such as farm of origin.

[000103] A matching and a non-matching example wherein non-matching mothers are excluded is shown in Table 9. In the first example shown in Table 9, MLAC1 is an offspring and PGG1-5 are 5 possible mothers. SNPTrack analysis was performed and the SNPTracks were compared. In this assay, failures are indicated as N/A and some SNPs are indicated as D/I for deletion and insertion. None of the 5 mothers could be the parent of the same MLAC1 because they are excluded by the SNPTrack analysis. The excluded positions are highlighted in gray.

[000104] In the second example shown in Table 9, among the five possible mothers PGG1-5, only PGG4 could be the parent of the offspring MLAC2 because PGG1-3 and PGG5 are excluded, with the excluded positions shown in grey highlights.

Example 6: Developing SNPTracks to identify and/or trace a beef sample to a particular location or a farm of origin.

[000105] SNPTrack analysis disclosed herein can be adapted to traceability and identity assays to track beef products to a specific location or farm of origin. Based on the disclosure provided herein, SNPTracks that include a plurality of segments of SNPs in cow genome can be obtained from SNP sources such as the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/genome/guide/cow/>).

[000106] Another source to obtain bovine SNPs is <http://www.livestockgenomics.csiro.au/ibiss/> discussed in Hawken *et al.*, (2004), An interactive bovine in silico database (IBISS). Mammalian Genome 15, 819-827.

TABLE 3: Characteristics of amplicons with SNPs from pig genome.

Amplicon Name	Length of the oligo (bp)	Oligo Sequence	Amplicon size (bp)	Annealing temperature (°C)
PBE3F	23	CTGACAGTTAAAGACTGCCCAAC	658	63
PBE3R	24	AGCTCCATGTCTTACTCATCCACC		
ACY-F7	24	GCAGCAATACTGTGCCTTAGAAAC	976	65
ACY-R7	24	AAGTAATAGGGAGTGAGAGTCTCC		
BG-F2	23	ATCTCGACAACCTCAAGGGCACC	1145	63
BG-R2	24	CCACTCACATGCGTGCTTTACAAC		
COX2-F4	23	TGGTGCCTGGTCTGATGATGTAC	1148	63
COX2-R4	24	AGCAGAAAGCGCTTGCGGTATTCA		
EG-F7	24	CACCCTGCAGCATCTTCTTAGCTG	1074	63
EG-R7	24	CCAAACTGAGGCCGGGTGTGCTC		
GALT-F4	22	TGCAAGCATCCAGGGCTGCTTT	1394	63
GALT-R4	24	CGAAACAAGTTCTAGTGAGCTCTG		
IKBA-F5	23	ACACGGAGTCAGAGTTCACAGAG	1291	65
IKBA-R5	24	AGCAGAAGTAAGGTCTGGCTGAA		
LepR-F7	22	AAACCGCTGCCTCCATCCAGTG	1330	63
LepR-R7	24	AGGCTGGAGTACTCCAATTACTCC		
LPL-F2	24	TCCATGATAGGCTGCATCCTAGAA	1244	63
LPL-R2	24	CCAGAGGTACGCGGAACAGAACTG		
P450-F18	24	TCCCTGTTCTCTATGGCCTGCTTC	896	63
P450-R18	24	GATGTGGTGGTGCTGAACTCCAAG		
PBE42F	24	TCTGAGCCTCATATTCTAATGGAC	534	60
PBE42R	23	CCTTTCAGTGCAGAAGTTCCAGG		
PBE43F	25	GATTGCAGTATTTTTTGTCTTGGAG	704	63
PBE43R	23	CTGAAAGATCCATCCATTGTTC		

PBE57F	24	ATGCTGCGATTTCTCTGGAGTTCC	602	63
PBE57R	24	TGTCCAGACGTTCCCTTTGGCTCC		
PBE59F	26	TTGAAATCCCTACTTAAGTCCTGCTC	717	63
PBE59R	27	CAAGAGTAAATCATGCAAGGAAATGTG		
PBE64F	25	TAGGACAGGAGAAGGATAACAAACC	760	63
PBE64R	24	AGTCCCTGGACATATGCAGGTTAG		
PBE73F	25	AGATGATCATCGAGCTGTAGGATAG	590	63
PBE73R	25	TTCCAATCCTTTGTGAATATCTGGC		
PBE77F	24	GCAGGGACAGCTCTGCCAGGGAAC	529	63
PBE77R	27	GTTACCTTACTTGAACCCTTTCTTTCC		
PBE84F	25	AGAGAACCCTCAACTCTCAGCTGTG	663	63
PBE84R	26	TTCAGCCTTATTGAGGTATAGTTATC		
PRKAG-F3	23	CAAGAAGCAGAGCTTCGTGGGTG	1043	63
PRKAG-R3	24	CGATGAGTCCATGAGCTTAGAACC		
RYRA-F6	25	GTGTAATCTGTTGGAGTATTTCTGT	1339	63
RYRA-R6	24	TCGGTAAGATTATCATCTGACTTC		
SCAMP1-F3	23	TGAGCAGCAGGTCTGGAATCTAG	1175	63
SCAMP1-R3	24	GAGTAGCCACAAGAATATACCAAG		
VAN-F1	23	CCACAATGCTTGCTTTGCAGAG	1272	63
VAN-R1	24	CGCATCACACAAATGTGTTTCATGG		
WSCR-F1	23	GAGAGGCAGTGGGTCCAGACCAA	1249	60
WSCR-R1	24	TCCAAGGTGGTCTGAGCTGACCTG		

TABLE 4: Description of Genetic Regions Examined for SNPs

Name	Length (bp)	Description
AB049327S1	1470	Sus scrofa IL7 gene for interleukin 7, exon 1, partial sequence.
SSIGFBPII2	1515	Sus scrofa IGF binding protein-2 (IGF) gene, exons 2 and 3.
SSAJ3734	1551	Sus scrofa SCAMP1 gene, exon 1 and joined CDS.
AF252874	1638	Sus scrofa bactericidal permeability increasing protein (BPI) gene, partial cds.
PIGMEF2A2	1768	Sus scrofa domestica myocyte enhancer factor 2A (MEF2A) gene, exon and partial cds.
AF329087	1835	Sus scrofa Niemann-Pick type C1 protein gene, promoter region and partial cds.
AF415201S2	1839	Sus scrofa alpha-1,3-galactosyltransferase gene, exons 2 and 3.
SSIGFBPII1	1861	Sus scrofa IGF binding protein-2 (IGF) gene, exon 1.
SSC309827	1933	Sus scrofa partial ABCD3 gene for peroxisomal membrane protein 1, exons 13-14.
SSC430415	2053	Sus scrofa partial GLUL gene for glutamate-ammonia ligase, exons 3-4.
AF019044	2132	Sus scrofa DAX-1 gene, partial cds.
SSBETAG	2392	Sus scrofa beta-globin gene.
SSC344137	2395	Sus scrofa partial ATP1A1 gene for Na ⁺ /K ⁺ ATPase alpha 1 subunit, exons 13-16.
AF331845	2443	Sus scrofa androgen receptor gene, promoter region and 5' untranslated region, partial sequence.
SSC2BTXNS	2472	S.scrofa C2 gene (exons 13-18) and BF gene (exons 1-2).
AF415201S1	2695	Sus scrofa alpha-1,3-galactosyltransferase gene, exon 1.
AF202775	2798	Sus scrofa androgen receptor mRNA, complete cds.
SSC404884	3032	Sus scrofa partial cdkn3 gene, exons 7-8.
AF247680	3101	Sus scrofa immunoreceptor DAP12 gene, complete cds.
SSAJ3742	3265	Sus scrofa SCAMP1 gene, exon 9.
SSDNAMYO	3506	S.scrofa myogenin gene.
AY028583	3621	Sus scrofa prostaglandin G/H synthase-2 (PGHS-2) mRNA, complete cds.
SSMYF5G	3680	Sus scrofa myf-5 gene and 3 microsatellite sequences.
AY044189	3733	Sus scrofa uroplakin II gene, complete cds.
AF430245	3823	Sus scrofa vanin-1 gene, promoter region, exon 1, intron 1 and partial cds.
SSC249746-EPO	3874	Sus scrofa epo gene for erythropoietin, exons 1-5.
AF492499	4161	Sus scrofa obese (ob) gene, intron 1.

SSU96150	4631	Sus scrofa tear lipocalin/von Ebner's lingual gland protein (LCN1) gene, complete cds.
SSC6076	4911	Sus scrofa HSL gene, exons 6 to 9 and 3' UTR.
AY237828	5121	Sus scrofa vanin-1 gene, promoter region, 5'UTR, and partial cds.
E15380	5418	Porcine MCP promoter.
SSIKBAGE	5764	S.scrofa IkBa gene.
AF214521	5888	Sus scrofa AMPK gamma subunit (PRKAG3) gene, complete cds.
AF328419	6239	Sus scrofa amelogenin gene, exons 3, 4a, 4b, 5, 6, and 7a.
AF458070	6337	Sus scrofa bone morphogenetic protein 15 (BMP15) gene, exons 1 and 2 and partial cds.
SSU14331	6511	Sus scrofa myogenin gene, complete cds.
AY116585	6727	Sus scrofa inhibin beta B precursor subunit (INHBB) gene, exons 1 and 2, complete cds.
SSTNFAB	7218	Porcine TNF-alpha and TNF-beta genes for tumour necrosis factors alpha and beta, respectively.
AC090553trunk	8043	
AY112657	8053	Sus scrofa fibrinogen-like protein 2 (FGL2) gene, complete cds.
SSY16039	8144	Sus scrofa A-FABP gene for fatty acid-binding protein, exons 1-5.
AC091506trunk	8212	
AF036005	8480	Sus scrofa interleukin-2 receptor alpha chain gene, partial cds.
AF535216	8660	Sus scrofa endothelial nitric oxide synthase (NOS) gene, exons 3 through 14 and partial cds.
AB017196	9361	Sus scrofa ACY-1 and rpl29/HIP genes, complete cds.
SSC404883	9923	Sus scrofa partial cdkn3 gene, exons 1-6 and join CDS.
SSC296176	10281	Sus scrofa LIF gene for leukemia inhibitory factor.
SSC315771	12715	Sus scrofa CTSL gene for cathepsin L, exons 1-8.
SSC7302	15604	Sus scrofa triadin gene.
AL773560P	17040	Pig DNA sequence from clone containing cytochrome P450-21-hydroxylase
SSPPK	19298	S.scrofa ppk98 gene.
SSRYRA	17808	S.scrofa gene for skeletal muscle ryanodine receptor.

TABLE 5: Allele frequency distribution

Marker	No. of SNP s	SNPTrac k 1	Frequency	SNPTrac k 2	Frequency	SNPTrac k 3	Frequenc y	SNPTrac k 4	Frequency
EG	3	CGC	46%	CAC	22%	TAC	17%	CAT	15%
RYRA	3	ACA	36%	ACC	22%	AGC	22%	CCC	22%
VAN	3	CGC	60%	CAC	19%	TGT	13%	CGT	8%
WSCR	3	GCG	44%	AAA	22%	GAG	18%	GAA	16%
BG	3	CCC	35%	CAC	30%	TCC	24%	CCT	11%
LEPR	2	AG	25%	AC	25%	CC	25%	GC	25%
IKBA	3	CAG	38%	CTG	26%	TAG	26%	CAT	10%
COX2	3	TAT	54%	TGT	17%	CAC	17%	CAT	12%
GALT	3	ACC	47%	GGG	24%	GGC	17%	GCC	17%
PRKA G	3	GAG	43%	GGA	27%	AAG	20%	GAA	11%
SCAM P	3	TAT	33%	TGA	26%	TGT	22%	CGT	19%
P450	3	CGG	40%	CAG	23%	TGG	21%	CGA	16%
PBE42	3	AGC	48%	AGT	27%	GGT	17%	ATC	8%
PBE43	3	-CT	41%	+CC	21%	-AT	25%	-CC	14%
PBE3	3	G-G	34%	A+G	33%	G-(gg)	18%	G+G	15%
PBE64	3	GGC	39%	GCT	34%	GGT	27%	AGT	6%
PBE84	3	GTC	40%	GCC	28%	TCC	22%	GTT	10%
PBE57	3	CAT	50%	TAC	22%	CGC	15.50%	CAC	12.50%
PBE59	2	TC	50%	CC	30%	CT	20%	na	na
ACY	2	GC	52%	GT	30%	CC	18%	na	na

TABLE 6: SNPTrack Exclusion Probabilities

Marker type	Exclusion probability	n ₂	n ₁
Autosomal markers only	$Q = 1 - 10^{-6}$	12.4	16.8
	$Q = 1 - 10^{-7}$	14.5	19.5
	$Q = 1 - 10^{-8}$	16.6	22.3
	$Q = 1 - 10^{-9}$	18.6	25.1
Autosomal markers and mitochondrial DNA typing	$Q = 1 - 10^{-6}$	10.4	14
	$Q = 1 - 10^{-7}$	12.4	16.8
	$Q = 1 - 10^{-8}$	14.5	19.5
	$Q = 1 - 10^{-9}$	16.6	22.3

Number of autosomal markers required to achieve a given exclusion probability (Q): n₂ = number of autosomal markers required when both the alleged dam and sire have marker genotypes; n₁ = number of autosomal markers required with the alleged dam has marker genotypes by the alleged sire does not have marker genotypes.

TABLE 7: Marker Assignment/Position

Marker	No. SNPs	Old Marker	SNP positions							
P1SMIT	8	Mitochondrial	15542 C/T	15558 A/T	15615 C/T	15616 C/T	15675 C/T	15714 C/T	15840 C/T	16127 G/A
P1S001	2	ACY-STS7	245 G/C	421 C/T						
P1S002	3	Cox2	368 C/T	533 G/A	939 C/T					
P1S003	3	EG-STS7	774 G/A	805 G/A	817 G/A					
P1S004	3	GALT	478 G/A	758 C/G	866 C/G					
P1S005	3	IKBA	4476 C/T	4679 T/A	4904 G/T					
P1S006	2	LepR	426 A/C/G	810 G/C						
P1S007	3	P450-STS18	71 C/T	138 G/A	361 G/A					
P1S008	3	PBE3	115 G/A	192 A/T	555 T/G					
P1S009	3	PBE42	111 G/A	118 T/G	181 C/T					
P1S010	3	PBE43	314 C/T	471 C/A	524 C/T					
P1S011	4	PBE 57	75 C/T	109 G/A	197 C/T	268 T/G				
P1S012	2	PBE59	276 C/T	494 C/T						
P1S013	3	PBE 64	115 G/A	419 C/G	515 C/T					
P1S014	4	PBE 73	93 A/C	116 A/G	177 C/T	477 C/T				
P1S015	3	PBE 84	14 A/G	97 T/C	428 T/C					
P1S016	4	PBE132	102 C/G	127 A/G	193 C/T	371 A/G				
P1S017	3	PBE137	121 C/T	278 C/T	409 A/G					
P1S018	3	PRKAG-STS3	1845 G/A	1938 G/A	2050 G/A					
P1S019	3	RYRA-STS6	402 A/C	408 C/G	567 A/C					
P1S020	5	SCAMP	184 C/T	389 G/A	516 C/T	582 G/A		939 A/T		
P1S021	4	VAN-STS1	889 C/T	950 C/T	1009 G/A	1065 C/T				
P1S022	2	WSCR-STS1	411 C/T	599 C/T						
P1S023	4	BG	1257 C/T	1323 C/T	1425 C/A			1966 C/T		
P1S024	3	AMG	907 A/C	975 A/G	1467 A/G					

-34-

P1S025	3	LCN	87 C/T	373 G/C	402 C/T					
P1S026	2	CTSL	252 A/G	272 A/G						
P1S027	2	PBE112	61 C/T	87 C/T						
P1S028	3	MYF5	1833 A/G	2204 C/G	2335 A/C					

TABLE 8: SNPTrack Analysis of Mothers and Offsprings

TABLE 8: GNT Track Analysis of Injuries and Crisps																											
	PBE3			ACY-STS7			EG-STS7			RYRA-STS6			PBE59			PBE43			GALT			VAN-STS1			IKBA		
	G/A	A/T	T/G	G/C	C/T	G/A	G/A	G/A	A/C	C/G	A/C	C/T	C/T	C/T	C/T	C/A	C/T	G/A	C/G	C/G	C/G	C/T	G/A	C/T	C/T	T/A	G/T
Animal I.D.	115	193	IN 555	150	326	773	804	817	402	408	567	276	494	494	4INDI	471	524	478	758	866	950	1009	1065	4476	4679	4904	
MLV1	G	A/T	T	G	C/T	G	G/A	G/A	A	G	C	C	T	T	C	C/A	C/T	G/A	C/G	C/G	C/T	G	C/T	C/T	T/A	G	
MLV1-P1	G/A	A/T	T	G	C/T	G	A	G/A	A/C	C/G	A/C	C	C/T	C	C	C/A	C/T	G/A	C/G	C	C	G/A	C	C/T	A	G	
MLV1-P2	G/A	T	T	G	C/T	G	A	G/A	A/C	C/G	A/C	C/T	C	C	C	C/A	C/T	G/A	C/G	C/G	F	F	F	C/T	A	G	
MLV1-P3	G/A	T	T	G	C/T	G	A	G/A	A/C	C/G	C	C/T	T	T	C	C/A	C/T	G	C/G	C/G	F	F	F	C/T	A	G	
MLV1-P4	G/A	A/T	T	G	C/T	G	G/A	G	A/C	C/G	C	C	T	T	C	C/A	T	G/A	C/G	C/G	C	G/A	C	C/T	A	G	
MLV2	G/A	A/T	T	G/C	C/T	G	G	G	A	C	A/C	C	T	T	C	C/A	T	G	C/G	C/G	C	G/A	C	C/T	T/A	G	
MLV2-P1	A	T	T	G/C	C/T	G	G/A	G	A/C	C	A/C	C	C/T	C	C	C/A	T	G	G	C/G	C	G/A	C	C	T/A	G	
MLV2-P2	A	T	T	C	C	G	G/A	G	A	C	A/C	C/T	T	T	C/T	C	C/T	G	C	C	C	G/A	C	C/T	A	G	
MLV2-P3	G/A	A/T	T	G/C	C/T	G	G/A	G	A/C	C	A/C	C	C/T	C	C	A	T	G	G	C/G	C	G	C	C/T	A	G	
MLV2-P4	G	A	T	G	T	G	G/A	G	A/C	C	A/C	C	C/T	C	C	C/A	C/T	G	C/G	C	C	G/A	C	C/T	A	G	
MLV3	G/A	T	T	G	T	G	G	G	A	C	A	C	C/T	C	C	C	C/T	G/A	C/G	C/G	C	G	C	C	A	G/T	
MLV3-P1	A	T	T	G	T	G	G/A	G	A/C	C	A/C	C	C/T	C	C	C/A	T	G/A	C/G	C	C	G	C	C	A	G/T	
MLV3-P2	A	T	T	G/C	C/T	G	G/A	G	A/C	C	A/C	C	T	T	C	C	T	G/A	C/G	C/G	C/T	G	C	C	T/A	G/T	
MLV3-P3	G	A/T	T	G	C/T	G	G/A	G	A/C	C	A/C	C	C/T	C	C	C	C	G/A	C/G	C	C	G/A	C	C	A	G/T	
MLV3-P4	A	T	T	G/C	C/T	G	G/A	G	A/C	C	A/C	C	T	T	C	C	T	G/A	C/G	C/G	C	G/A	C	C	A	G	
MLV4	G/A	T	T	C	C	G	A	G/A	A/C	C	A/C	C	C	C	C	A	T	A	C	C	C	G/A	C	C	A	G	
MLV4-P1	A	T	T	G	C	G	A	G/A	C	C	C	C	C	C	C	C	T	A	C	C	C	G/A	C	C	A	G	
MLV4-P2	A	T	T	C	C	G	G/A	G	A	C	A/C	C	C	C	C	C	T	A	C	C	C	G	C	C	A	G	
MLV4-P3	A	T	T	G	C	G/A	G/A	G	A	C	A/C	C	C	C	C	C/A	T	G/A	C/G	C/G	C/T	G	C	C	T/A	G	
MLV4-P4	G/A	T	T	G/C	C	G	G/A	G/A	A/C	C	C	C	C/T	C	C	C/A	T	A	C	C	C	G/A	C	C	T/A	G	
MLV5	G	A/T	T	G/C	C	G	G/A	G/A	A	C	A/C	C	C/T	C	C	C/A	C/T	G/A	C/G	C	C	G	C	C	T/A	G	
MLV5-P1	G	A	T	G	C	G	G/A	G	A/C	C	A/C	C	C/T	C	C	C/A	C/T	G/A	C/G	C	C	G/A	C	C	A	G	
MLV5-P2	G/A	A/T	T	G/C	C/T	G	A	G/A	A/C	C	A/C	C	C	C	C	A	T	G/A	C/G	C	C	G	C	C	A	G	
MLV5-P3	G	A	T	G/C	C	G	G/A	G	A	C	A	C	C/T	C	C	C/A	C/T	G	G	C	C	G	C	C	A	G	
MLV5-P4	G/A	A/T	T	C	C	G	A	G/A	A/C	C	A/C	C	T	T	C	C/A	T	G/A	C/G	C/G	C	G/A	C	C	T/A	G	
MLV6	G/A	T	T	G	C/T	G	A	G/A	A/C	C	C	C	C/T	C	C	C	T	G/A	C/G	C/G	C	G/A	C	C/T	T/A	G	
MLV6-P1	G/A	A/T	T	G	C	G	A	G/A	A/C	C	C	C	C	C	C	C	T	G/A	C/G	C/G	C	A	C	C	T	G	
MLV6-P2	A	T	T	G	C/T	G	A	G/A	A/C	C	C	C	C	C	C	C	T	G	G	C	C	G/A	C	C	T/A	G	
MLV6-P3	G	A/T	T	G	C/T	G	A	G/A	A/C	C	C	C	C	C	C	C	T	G/A	C/G	C/G	C	G	C	C	T	G	
MLV6-P4	G/A	A/T	T	G	C/T	G	A	G/A	A	C	C	C	C	C	C	C	T	G/A	C/G	C/G	C	G/A	C	C	A	G	
MLV7	G/A	A/T	T	G/C	C/T	G	G/A	G/A	A	C/G	A/C	C	T	T	C	C	C/A	C/T	G/A	C/G	C/G	C	G	C	T/A	G/T	

	PBE64		SCAMP		LEPR		COX2		P450		MITOCHONDRIAL												
	C/G	C/T	C/T	G/A	C/T	G/A	C/G	G/A/C	C/T	G/A	C/T	G/A	C/T	AT	15543	15558	15615	15616	C/T	C/T	15714	C/T	G/A
Animal I.D.	419	515	184	389	516	582	426	810	368	533	939	71	138	361	15543	15558	15615	15616	C/T <td>C/T<td>15714</td><td>C/T<td>G/A</td></td></td>	C/T <td>15714</td> <td>C/T<td>G/A</td></td>	15714	C/T <td>G/A</td>	G/A
MLV1	C/G	C/T	C	G/A	T	G/A	C/G	G/A	T	G/A	T	C	A	G	T	T	C	T	T	C	T	T	G
MLV1-P1	C/G	C/T	C/T	G/A	T	G	C/G	G/A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV1-P2	C/G	C/T	C/T	G/A	T	G	C/G	G/A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV1-P3	C/G	T	C/T	G/A	T	G	C/G	A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV1-P4	C	T	C	G	T	G	C/G	G/A	T	A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV2	G	C	C/T	A	T	G	C/G	A	T	G/A	T	C	A	G	T	T	C	T	T	C	T	T	G
MLV2-P1	G	C	T	A	T	G/A	C/G	A	T	A	T	C	G/A	G/A	T	T	C	T	T	C	T	T	G
MLV2-P2	G	C/T	C/T	A	T	G	C/G	A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV2-P3	G	C/T	C	A	T	G	G	A	T	A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV2-P4	G	C	T	A	T	G	G	A	T	G	T	C	G/A	G/A	T	T	C	T	T	C	T	T	G
MLV3	G	C	C/T	A	T	G/A	C/G	G/A	C/T	A	T	C	G/A	G/A	C	A	C	T	C	C	C	C	G
MLV3-P1	G	C/T	C/T	A	T	G	C/G	G/A	T	G/A	T	C	G	G/A	C	A	C	T	C	C	C	C	G
MLV3-P2	G	C/T	C	A	T	A	C	G/A	T	A	T	C	G/A	G	C	A	C	T	C	C	C	C	G
MLV3-P3	G	C	C/T	A	T	G/A	C/G	G/A	C/T	A	T	C	G/A	G	C	A	C	T	C	C	C	C	G
MLV3-P4	G	C	C/T	A	T	G/A	C/G	A	C/T	G/A	T	C	G	G/A	C	A	C	T	C	C	C	C	G
MLV4	G	T	C/T	A	T	G	C	A	T	A	T	C	A	G	T	T	C	T	T	C	T	T	G
MLV4-P1	G	C/T	T	A	T	G	C/G	A	T	A	T	C	G/A	G/A	T	T	C	T	T	C	T	T	G
MLV4-P2	G	C/T	T	A	T	G	G	A	T	A	T	C	A	G	T	T	C	T	T	C	T	T	G
MLV4-P3	G	T	C/T	A	T	G	G	A	T	A	T	C	A	G	T	T	C	T	T	C	T	T	G
MLV4-P4	G	C/T	C/T	A	T	G	C/G	A	T	A	T	C/T	G/A	G	T	T	C	T	T	C	T	T	G
MLV5	C/G	C/T	C/T	A	T	G/A	C/G	A	T	A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV5-P1	C/G	C/T	C/T	A	T	G/A	C/G	A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV5-P2	C/G	T	C	A	T	G/A	G	A	T	A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV5-P3	G	C/T	T	A	T	G	C/G	A	T	A	T	C	G/A	G	T	T	C	T	T	C	T	T	G
MLV5-P4	C/G	C/T	C/T	G/A	T	G	C/G	A	T	A	T	C	A	G	T	T	C	T	T	C	T	T	G
MLV6	C/G	C/T	C/T	G/A	T	G	C	G/A	C/T	A	C/T	C/T	G/A	G	C	A	C	T	T	C	T	T	G
MLV6-P1	C/G	T	C	G/A	T	G	C	G/A	T	A	T	C	A	G	C	A	C	T	T	T	T	T	G
MLV6-P2	G	C/T	T	A	T	G	C	G/A	C/T	A	C/T	C/T	G/A	G	C	A	C	T	T	T	T	T	G
MLV6-P3	G	C/T	T	A	T	G	C/G	A	T	A	T	C	A	G	C	A	C	T	T	T	T	T	G
MLV6-P4	G	C/T	C/T	A	T	G	C	G	C/T	A	C/T	C	A	G	C	A	C	T	T	T	T	T	G
MLV7	C/G	T	C	G/A	T	G	C/G	A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	T	G

Animal I.D.	115	193IN	555	150	326	773	804	817	402	408	567	276	494	4INDI	471	524	478	758	866	950	1009	1065	4476	4679	4904
MLV7-P1	G/A	A/T	T	G	T	G	G/A	G	A/C	C/G	C	C	C/T	C	A	T	G	G	C/G	C	G	C	C	A	G/T
MLV7-P2	G/A	A/T	T	G/C	C/T	G	G/A	G	A	C/G	A/C	C	C/T	C	A	T	G/A	C/G	C	C	G/A	C	C	T/A	G
MLV7-P3	A	T	T	G	C/T	G	G/A	G/A	A/C	C	A/C	C	C/T	C/T	C/A	C/T	G	G	C/G	C	G	C	C	T/A	G
MLV7-P4	A	T	T	G/C	C/T	G	A	G/A	A	C/G	A/C	C	C/T	C/T	C/A	C/T	G	G	C/G	C	G	C	C	T/A	G
MLV8	A	T	T	G/C	C	G	G/A	G/A	A	C/G	C	C	T	C/T	C/A	C/T	G/A	C/G	C/G	C	G/A	C	C/T	T/A	G
MLV8-P1	A	T	T	G/C	C	G	G/A	G	A/C	C/G	C	C	T	C	C/A	T	G/A	C/G	C/G	C	G/A	C	C/T	A	G
MLV8-P2	A	T	T	G/C	C	G	A	G/A	A/C	C	C	C	C/T	C/T	C/A	C/T	G/A	C/G	C	C	G/A	C	C/T	A	G
MLV8-P3	A	T	T	G/C	C/T	G	G/A	G	A/C	C/G	C	C	C/T	C	A	T	G/A	C/G	C	C	G	C	C	T/A	G
MLV8-P4	G/A	A/T	T	G	C	G	A	G/A	A/C	C	C	C	C/T	C/T	C/A	C/T	G/A	C/G	C	C	G/A	C	C/T	A	G
MLV9	G/A	T	T	G/C	C/T	G/A	G/A	G	A	C	A/C	C	T	C/T	C	C/T	G/A	C/G	C/G	C	G/A	C	C	T/A	G
MLV9-P1	G	A/T	T	G/C	C	G	G/A	G	A	C	A/C	C	C/T	C	C/A	T	G/A	C/G	C	C	G/A	C	C	T/A	G
MLV9-P2	G/A	T	T	G/C	C/T	G/A	A	G	A/C	C	C	C	T	C	C	T	G/A	C/G	C/G	C	G	C	C	T	G
MLV9-P3	G	A/T	T	G	C/T	G/A	A	G	A	C	A	C	C/T	C/T	C	C/T	G/A	C/G	C	C	G/A	C	C	A	G
MLV9-P4	G	A/T	T	G	C/T	G/A	A	G	A	C	A	C	C/T	C/T	C	C/T	G/A	C/G	C	C	G	C	C	T/A	G
MLV10	G/A	A/T	T	F	F	G	G	G	A	C/G	C	C	C/T	C	C/A	T	G/A	C/G	C/G	C	G	C	C/T	T/A	G
MLV10-P1	G/A	A/T	T	C	C	G	G/A	G	A	C	A/C	C	T	C	C/A	T	G/A	C/G	C/G	C	G/A	C	C	T/A	G
MLV10-P2	G	A	T	G/C	C	G	G/A	G	A/C	C	C	C	C	C/T	C/A	C/T	G/A	C/G	C	C	G	C	C/T	A	G
MLV10-P3	A	T	T	G	C/T	G	G/A	G	A/C	C	C	C	C/T	C	C	T	G	G	G	C	G	C	C/T	A	G
MLV10-P4	A	T	T	G/C	C	G	G/A	G	A/C	C	C	C	C/T	C	C	T	G	G	G	C	G	C	C/T	A	G
MLV11	G/A	A/T	T	G/C	C/T	G	A	G/A	A	C	A/C	C	C/T	C	C/A	T	G/A	C/G	C/G	C	G/A	C	C	T/A	G
MLV11-P1	A	T	T	G/C	C	G	A	G	A	C/G	A/C	C	C/T	C	C/A	T	G	G	C/G	C	G/A	C	C	T/A	G
MLV11-P2	A	T	T	G	C/T	G	G/A	G	A	C	A/C	C	T	C/T	C/A	C/T	G/A	C/G	C/G	C	A	C	C	T/A	G
MLV11-P3	G/A	A/T	T	G	C/T	G	A	G/A	A	C/G	A/C	C	T	C	A	T	A	C	C	C	G	C	C	A	G
MLV11-P4	G/A	A/T	T	G/C	C	G	G/A	G	A	C/G	C	C	T	C	A	T	G/A	C/G	C	C	G	C	C	A	G
MLV12	G/A	A/T	T	G/C	C/T	G	G/A	G/A	A	C	A/C	C	T	C/T	C/A	C/T	G/A	C/G	C/G	C	G	C	C/T	T/A	G
MLV12-P1	G/A	A/T	T	G/C	C	G	A	G/A	A	C/G	C	C	T	C	A	T	G/A	C/G	C/G	C	G	C	C	T	G
MLV12-P2	A	T	T	G	C/T	G	A	G/A	A	C	C	C	C/T	C/T	C/A	C/T	G	G	C/G	C	G/A	C	C/T	A	G
MLV12-P3	A	T	T	G/C	C	G	A	G/A	A	C/G	C	C	C/T	C/T	C/A	C/T	G	G	C/G	C	G/A	C	C	T/A	G
MLV12-P4	G/A	A/T	T	G/C	C	G	G/A	G	A	C/G	A/C	C	T	C	A	T	A	C	C	C	G	C	C	T/A	G
MLV13	G/A	A/T	T	G	C/T	G	G/A	G/A	A/C	C	C	C	T	C	C/A	T	G/A	C/G	C	C	G/A	C	C	T/A	G/T
MLV13-P1	A	T	T	G	C	G	A	G/A	A	C/G	C	C	C/T	C	C/A	T	G	G	C	C	G	C	C	A	G/T
MLV13-P2	G/A	A/T	T	G	C/T	G	A	G/A	A/C	C	C	C	T	C	A	T	G/A	C/G	C	C	G/A	C	C	A	G/T
MLV13-P3	A	T	T	G	C	G	G/A	G	A	C	C	C	T	C	C/A	T	G/A	C/G	C	C	A	C	C	A	G/T
MLV13-P4	A	T	T	G	C	G	A	G/A	A	C/G	C	C	C/T	C	C/A	T	G	G	C	C	A	C	C	A	G/T

Animal I.D.	419	515	184	389	516	582	426	810	368	533	939	71	138	361	15543	15558	15615	15616	15675	15714	15840	16127
MLV7-P1	G	C/T	C/T	A	T	G	G	A	T	A	T	C	G	G/A	T	T	C	T	T	C	T	G
MLV7-P2	G	T	C	A	T	G	C	G/A	T	G	T	C	G	G	T	T	C	T	T	C	T	G
MLV7-P3	C/G	T	C	G/A	T	G	C/G	A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV7-P4	G	T	C	A	T	G	C	G/A	T	G	T	C	G/A	G/A	T	T	C	T	T	C	T	G
MLV8	C/G	T	C	G/A	T	G	C	A	T	A	T	T	G	G	T	T	C	T	T	C	T	G
MLV8-P1	C/G	C/T	C	G/A	T	G/A	C/G	A	T	G/A	T	C/T	G	G	T	T	C	T	T	C	T	G
MLV8-P2	G	C/T	C/T	A	T	G	C/G	A	T	A	T	C/T	G	G/A	T	T	C	T	T	C	T	G
MLV8-P3	C/G	C/T	C/T	G/A	T	G	C/G	A	T	A	T	C/T	G	G/A	T	T	C	T	T	C	T	G
MLV8-P4	C/G	T	C	G/A	T	G	C/G	A	T	G/A	T	C/T	G	G/A	T	T	C	T	T	C	T	G
MLV9	G	C	C/T	A	T	G/A	C/G	A	T	A	T	T	G	G	C	A	C	T	T	T	T	G
MLV9-P1	G	C	C/T	A	T	G/A	C/G	A	T	A	T	C/T	G	G	C	A	C	T	T	T	T	G
MLV9-P2	G	C	C/T	A	T	G/A	G	A	T	A	T	C/T	G	G	C	A	C	T	T	T	T	G
MLV9-P3	G	C/T	C/T	A	T	G	C	G/A	T	G/A	T	C/T	G	G/A	C	A	C	T	T	T	T	G
MLV9-P4	G	C	T	A	T	G	C/G	A	T	A	T	C/T	G	G/A	C	A	C	T	T	T	T	G
MLV10	C	T	C	A	T	A	C/G	A	T	A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV10-P1	C/G	C/T	C	A	T	A	C	A	T	G/A	T	C	G	G	T	T	C	T	T	C	T	G
MLV10-P2	C/G	C/T	C/T	A	T	G/A	G	A	T	G/A	T	C	G/A	G/A	T	T	C	T	T	C	T	G
MLV10-P3	C/G	C/T	C	A	T	A	C	A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV10-P4	C/G	C/T	C	A	T	A	C/G	A	T	A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV11	C/G	T	C	G/A	T	G	C/G	A	C/T	A	T	C	G	G/A	T	T	C	T	T	C	T	G
MLV11-P1	G	T	C	A	T	G/A	G	A	T	G/A	T	C	G	G	T	T	C	T	T	C	T	G
MLV11-P2	G	T	C	A	T	A	G	A	T	A	T	C	G/A	G	T	A	C	T	C	C	T	G
MLV11-P3	G	T	C	A	T	G/A	C	C/A	C/T	G/A	T	C	G/A	G	T	T	C	T	C	C	T	G
MLV11-P4	G	T	C	A	T	A	C/G	A	T	A	T	C	A	G	T	A	C	T	C	C	T	G
MLV12	G	C/T	C/T	A	T	G	C/G	A	C/T	A	T	C	G/A	G	C	A	C	T	C	T	T	A
MLV12-P1	G	T	C	A	T	G/A	C/G	A	C/T	G/A	T	C	A	G	C	A	C	T	C	T	T	A
MLV12-P2	G	C/T	C/T	A	T	G/A	G	A	T	G/A	T	C	G/A	G	C	A	C	T	C	T	T	A
MLV12-P3	G	C/T	C/T	A	T	G/A	C/G	C/A	T	G/A	T	C	A	G	C	A	C	T	C	T	T	A
MLV12-P4	G	T	C	A	T	A	C	C/A	T	G/A	T	C	A	G	C	A	C	T	C	C	T	G
MLV13	C/G	T	C	G/A	T	G	C/G	A	T	A	T	C	A	G	C	A	C	T	T	T	T	G
MLV13-P1	C/G	T	C	G/A	T	G/A	C/G	C/A	T	G/A	T	C	A	G	C	A	C	T	T	T	T	G
MLV13-P2	C/G	T	C	G/A	T	G/A	C/G	A	T	G/A	T	C	A	G	C	A	C	T	T	T	T	G
MLV13-P3	G	T	C	A	T	G/A	C/G	C/A	T	G/A	T	C	A	G	C	A	C	T	T	T	T	G
MLV13-P4	C/G	T	C	G/A	T	G/A	C	C/A	T	G/A	T	C	A	G	C	A	C	T	T	T	T	G

Animal I.D.	115	193IN	555	150	326	773	804	817	402	408	567	276	494	4INDI	471	524	478	758	856	950	1009	1065	4476	4679	4904
MLV14	G	AT	T	G	C	G	G	G	A	C/G	AC	C	T	C	C/A	C/T	A	C	C	C/T	G/A	C/T	C	T/A	G
MLV14-P1	G/A	T	T	G	C	G	G/A	G	A	G	C	C	C/T	C	C/A	C/T	G/A	C/G	C	C/T	G	C/T	C	T/A	G
MLV14-P2	F	F	F	G	C	G	G/A	G	A	G	C	C	C/T	C	C/A	C/T	G/A	C/G	C	C/T	G	C/T	C	A	G
MLV14-P3	G/A	AT	T	G	C	G	G/A	G	A	C	AC	C	C/T	C	A	T	G/A	C/G	C	C	A	C	C	A	G
MLV14-P4	G/A	AT	T	G	C	G	G/A	G	AC	C/G	AC	C	C/T	C	A	T	G/A	C/G	C	C	G/A	C	C/T	T/A	G
MLV15	G/A	AT	T	G	C/T	G	G/A	G/A	A	C	AC	C	T	C/T	C	C/T	G/A	C/G	C	C	G/A	C	C	T/A	G
MLV15-P1	A	T	T	G	C	G	A	G/A	A	C	C	C	C/T	C/T	C/A	C/T	G	G	C	C	G	C	C	A	G
MLV15-P2	G/A	AT	T	G	C	G	A	G/A	A	C	C	C	T	C	C/A	T	G	G	C/G	C	A	C	C	T/A	G
MLV15-P3	A	T	T	G	C/T	G	A	G/A	A	C/G	AC	C	C/T	C/T	C/A	C/T	G	G	C/G	C	G	C	C	T/A	G
MLV15-P4	A	T	T	G	C/T	G	A	G/A	A	C/G	C	C	C/T	C/T	C/A	C/T	G	G	C/G	C	G	C	C	T/A	G
MLV16	G	AT	T	G/C	C/T	G	G	G	A	C/G	AC	C	T	C	C/A	T	G/A	C/G	C	C	G	C	C/T	A	G/T
MLV16-P1	G/A	T	T	G/C	C	G	G/A	G	A	C/G	AC	C	T	C	A	T	G/A	C/G	C	C	G/A	C	C	A	G/T
MLV16-P2	G/A	AT	T	G	C/T	G	G/A	G	A	C/G	AC	C	C/T	C	A	T	G/A	C/G	C	C	G/A	C	C/T	A	G
MLV16-P3	G/A	T	T	G	C/T	G	G/A	G	A	C/G	C	C	C/T	C	C/A	T	G	G	C/G	C	G/A	C	C/T	A	G
MLV16-P4	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
MLV17	A	T	T	G	C	G	G/A	G/A	A	C	AC	C	C/T	C	C/A	T	G/A	C/G	C	C	G	C	C	T/A	G
MLV17-P1	A	T	T	G	C	G	A	G/A	A	C/G	AC	C	C/T	C	A	T	G	G	C/G	C	G	C	C	T/A	G
MLV17-P2	A	T	T	G	C	G	G/A	G	A	C	C	C	T	C	C/A	T	G	G	C/G	C	G	C	C	T/A	G
MLV17-P3	G/A	T	T	G	C	G	G/A	G	A	C	C	C	T	C	C/A	T	G	G	C/G	C	G	C	C	A	G
MLV17-P4	A	T	T	G	C	G	A	G/A	A	C/G	C	C	T	C	A	T	G/A	C/G	C	C	G/A	C	C	A	G
MLV18	G	AT	T	G/C	C	G	G/A	G	A	C/G	C	C	T	C	C	C/T	G/A	C/G	C	C/T	G/A	C/T	C/T	T/A	G
MLV18-P1	G/A	T	T	G/C	C	G	A	G	A	C	C	C	T	C	C/A	C/T	G/A	C/G	C	C/T	G/A	C/T	C	T/A	G
MLV18-P2	G/A	T	T	G/C	C	G	G/A	G	A	C/G	C	C	C/T	C	C/A	C/T	G	G	C	C/T	G/A	C/T	C/T	A	G
MLV18-P3	G/A	T	T	G	C	G	G/A	G	A	C/G	C	C	C/T	C	C/A	C/T	G	G	C	C/T	G/A	C/T	C/T	A	G
MLV18-P4	G/A	T	T	G/C	C	G	A	G	A	G	C	C	C/T	C	C/A	C/T	G/A	C/G	C	C/T	G/A	C/T	C/T	A	G
MLV19	G/A	AT	T	G	C	G	G	G	AC	C	C	C	T	C	C/A	T	A	C	C	C	G	C	C/T	T/A	G
MLV19-P1	G/A	AT	T	G	C	G	G/A	G	A	C	C	C	T	C	A	T	A	C	C	C	G	C	C	T	G
MLV19-P2	G/A	AT	T	G	C	G	G/A	G	AC	C/G	C	C	C/T	C	A	T	G/A	C/G	C	C	G	C	C/T	A	G
MLV19-P3	G/A	AT	T	G	C	G	G/A	G	AC	C/G	C	C	T	C	A	T	G/A	C/G	C	C	G/A	C	C	T/A	G
MLV19-P4	A	T	T	G	C	G	G/A	G	AC	C	C	C	T	C	C/A	T	A	C	C	C	G	C	C/T	A	G
MLV20	A	T	T	G/C	C	G	G	G	A	C	AC	C	T	C	C	C/T	G	C/G	C	C	G/A	C	C	T/A	G
MLV20-P1	A	T	T	G/C	C	G	G/A	G	A	C	C	C	C/T	C	C/A	C/T	G	G	C/G	C	G/A	C	C	A	G
MLV20-P2	A	T	T	G/C	C	G	G/A	G	A	C/G	AC	C	T	C	C/A	T	G/A	C/G	C	C	A	C	C	T/A	G

Animal I.D.	419	515	184	389	516	582	426	810	368	533	939	71	138	361	15543	15558	15615	15616	15675	15714	15840	16127
MLV14	G	C/T	C/T	A	T	G	C/G	A	C/T	A	T	T	G	G	T	T	C	T	T	C	T	G
MLV14-P1	G	T	C	A	T	G/A	C/G	C/A	T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV14-P2	G	C/T	C/T	A	T	G/A	F	F	F	F	F	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV14-P3	G	T	C	A	T	G/A	C/G	C/A	C/T	G/A	T	F	G/A	G	T	T	C	T	T	C	T	G
MLV14-P4	G	C/T	C/T	A	T	G/A	C	C/A	C/T	G/A	T	F	F	F	T	T	C	T	T	C	T	G
MLV15	C/G	C/T	C	A	T	A	C/G	G/A	C/T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV15-P1	C/G	T	C	A	T	A	G	A	T	G	T	C	A	G	T	T	C	T	T	C	T	G
MLV15-P2	G	C/T	C	A	T	A	G	A	T	G	T	C	A	G	T	T	C	T	T	C	T	G
MLV15-P3	C/G	T	C	A	T	A	C/G	G/A	T	G	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV15-P4	C/G	T	C	A	T	A	C/G	C/A	T	G	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV16	C/G	T	C	A	T	G	C/G	A	C/T	G/A	T	C	G/A	G	C	A	C	T	C	C	C	G
MLV16-P1	C/G	T	C	A	T	A	C	C/A	C/T	G/A	T	C	G/A	G	C	A	C	T	C	C	C	G
MLV16-P2	C/G	T	C	A	T	A	C	C/A	C/T	G/A	T	C	A	G	C	A	C	T	C	C	C	G
MLV16-P3	G	T	C	A	T	G/A	G	A	T	G	T	C	G/A	G	C	A	C	T	C	C	C	G
MLV16-P4	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
MLV17	G	T	C/T	A	T	G	C/G	A	C/T	A	T	C/T	G	G	C	A	C	T	C	C	C	G
MLV17-P1	G	T	C/T	A	T	G/A	C/G	A	C/T	G/A	T	C/T	G/A	G	C	A	C	T	C	C	C	G
MLV17-P2	G	T	C	A	T	G/A	C/G	C/A	C/T	G/A	T	C	G/A	G	C	A	C	T	C	C	C	G
MLV17-P3	C/G	T	C	A	C/T	G/A	F	F	F	F	F	C/T	G/A	G	C	A	C	T	C	C	C	G
MLV17-P4	G	T	C	A	T	G/A	C/G	C/A	T	G/A	T	C/T	G/A	G	C	A	C	T	C	C	C	G
MLV18	G	C/T	C/T	A	T	G	C/G	G/A	C/T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV18-P1	G	T	C	A	T	G/A	C/G	G/A	C/T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV18-P2	G	C/T	C/T	A	T	G/A	C/G	C/A	C/T	G/A	T	C	A	G	T	T	C	T	T	C	T	G
MLV18-P3	G	T	C	A	T	G/A	C/G	G/A	C/T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV18-P4	G	T	C	A	T	G/A	C	C/G	T	G	T	C	A	G	T	T	C	T	T	C	T	G
MLV19	G	C/T	C/T	A	T	G	C/G	A	C/T	A	T	C	G/A	G/A	T	T	C	T	T	C	T	G
MLV19-P1	G	T	C	A	T	G/A	G	A	T	G/A	T	C	G/A	G/A	T	T	C	T	T	C	T	G
MLV19-P2	G	C/T	C/T	A	T	G/A	C	C/A	C/T	G/A	T	C	A	G	T	T	C	T	T	C	T	G
MLV19-P3	G	T	C	A	T	G/A	C/G	A	T	G/A	T	C	G/A	G/A	T	T	C	T	T	C	T	G
MLV19-P4	G	T	C	A	T	G/A	C/G	A	C/T	A	T	C	A	G	T	T	C	T	T	C	T	G
MLV20	G	C/T	C/T	A	T	G	C/G	G/A	C/T	A	T	T	G	G	T	T	C	T	T	C	T	G
MLV20-P1	G	C/T	C/T	A	T	G/A	C/G	C/A	T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV20-P2	G	T	C	A	T	G/A	C/G	G/A	T	A	T	C/T	G/A	G	T	T	C	T	T	C	T	G

Animal I.D.	115	193IN	555	150	326	773	804	817	402	408	567	276	494	4INDI	471	524	478	758	866	950	1009	1065	4476	4679	4904
MLV20-P3	A	T	T	G	C	G	G/A	G	A	C	C	C	T	C	C/A	C/T	G	G	C	C	C	C	C	T/A	G
MLV20-P4	A	T	T	G	C	G	G/A	G	A	C	A/C	C	T	C	C/A	T	G	G	C	C	G/A	C	C	A	G
MLV21	G/A	A/T	T	G	C/T	G	G	G	A/C	C	C	C	T	C/T	C	C/T	G/A	C/G	C/G	C/T	G/A	C/T	C	T/A	G
MLV21-P1	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	
MLV21-P2	G/A	A/T	T	G	C/T	G	G/A	G	A	C	C	C	T	C	C/A	T	G	G	C/G	C/T	G	C/T	C	A	G
MLV21-P3	G/A	A/T	T	G	C/T	G	G/A	G	A	C	A/C	C	T	C/T	C/A	C/T	G	G	C/G	C/T	G	C/T	C	T/A	G
MLV21-P4	A	T	T	G	C/T	G	G/A	G	A	C	C	C	T	C/T	C/A	C/T	G	G	C/G	C/T	G	C/T	C	T/A	G/T
MLV22	G/A	A/T	T	G	C/T	G	A	A	A/C	C/G	C	C	C/T	C/T	C/A	C/T	G	C/G	C/G	C	G	C	C/T	T/A	G
MLV22-P1	G	A	T	G	C/T	G	A	G/A	A	C/G	C	C	C/T	C	C/A	T	G	G	C/G	C	G	C	C/T	A	G
MLV22-P2	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	
MLV22-P3	G/A	A/T	T	G	C/T	G	A	G/A	A/C	C	C	C	C/T	C/T	C	C/T	G	C/G	C	C	G	C	C	T/A	G
MLV22-P4	G	A	T	G	C	G	A	G/A	A	C/G	A/C	C	C/T	C	C/A	T	G	C/G	C	C	G	C	C	T/A	G
MLV24	G	T	T	G	C	G	G	G	A/C	C	C	C	T	C	C/A	T	G/A	C	C	C	G/A	C	C	T/A	G
MLV24-P1	G	A/T	T	G	C	G	G/A	G	A	C	C	C	T	C	C	T	G/A	C/G	C	C	G/A	C	C	T/A	G
MLV24-P2	G/A	T	T	G	C	G	G/A	G	A	C	C	C	T	C	C/A	T	G	C/G	C	C	G	C	C	A	G
MLV24-P3	G/A	T	T	G	C	G	G/A	G	A/C	C	C	C	T	C	A	T	G	C/G	C	C	G/A	C	C	A	G
MLV24-P4	G	A/T	T	G	C	G	G/A	G	A	C	C	C	T	C	C	T	G/A	C/G	C	C	G	C	C	A	G
MLV25	A	T	T	G/C	C	G	G	G	A	C	A/C	C	T	C/T	C/A	C/T	G/A	C/G	C/G	C/T	G	C/T	C	T/A	G
MLV25-P1	A	T	T	G/C	C	G	G/A	G/A	A	C/G	A/C	C	T	C/T	C/A	C/T	G	G	C/G	C/T	G	C/T	C	A	G/T
MLV25-P2	G/A	A/T	T	G	C	G	G/A	G/A	A	C/G	C	C	T	C/T	C/A	C/T	G	G	C/G	C/T	G	C/T	C	A	G/T
MLV25-P3	G/A	A/T	T	G/C	C	G	G/A	G	A	C/G	C	C	T	C	A	T	G/A	C/G	C	C	G	C	C	T/A	G/T
MLV25-P4	G/A	A/T	T	G/C	C	G	G/A	G/A	A	C/G	A/C	C	T	C/T	C/A	C/T	G	G	C/G	C/T	G	C/T	C	A	G
MLV26	G/A	A/T	T	G	C/T	G	G/A	G	A	C	A	C	T	C	A	T	G	C/G	C/G	C	G	C	C	T/A	G/T
MLV26-P1	G/A	A/T	T	G	C/T	G	A	G	A	C	A	C	T	C	A	T	G	G	C/G	C	G	C	C	T/A	G
MLV26-P2	G	A	T	G	C	G	A	G	A	C	A/C	C	T	C	C/A	C/T	G	G	C/G	C	G	C	C	A	G/T
MLV26-P3	G/A	A/T	T	G	C	G	A	G	A	C	A	C	T	C	A	T	G	G	C/G	C	G	C	C	T/A	G
MLV26-P4	G/A	A/T	T	G	C	G	A	G/A	A	C/G	A/C	C	T	C	A	T	G	G	C/G	C	G	C	C	A	G/T
MLV27	G/A	A/T	T	G/C	C	G	G/A	G/A	A/C	C	C	C	T	C/T	C/A	C/T	G/A	C/G	C/G	C	G/A	C	C	T/A	G
MLV27-P1	G	A	T	G	C	G	G/A	G	A	C/G	C	C	T	C	A	T	G/A	C/G	C	C	G/A	C	C	A	G/T
MLV27-P2	A	T	T	G/C	C	G	G/A	G/A	A/C	C/G	C	C	T	C/T	C/A	C/T	G	G	C/G	C	G	C	C	A	G
MLV27-P3	G/A	A/T	T	G/C	C	G	A	A	A	C/G	C	C	T	C/T	C/A	C/T	G	G	C/G	C	G	C	C	A	G
MLV27-P4	G/A	A/T	T	G	C	G	G/A	G	A	C/G	C	C	T	C	A	T	G/A	C/G	C	C	G/A	C	C	A	G
MLV28	G/A	A/T	T	G/C	C/T	G	G	G	A/C	C	C	C	T	C	C/A	T	A	C	C	C	G	C	C/T	T/A	G

Animal I.D.	419	515	184	389	516	582	426	810	368	533	939	71	138	361	15543	15558	15615	15616	15675	15714	15840	16127
MLV20-P3	G	C/T	C/T	A	T	G/A	C	C/G	C/T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV20-P4	G	T	C	A	T	G/A	C/G	G/A	C/T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV21	G	C	C/T	A	T	G/A	C/G	A	C/T	A	T	T	G	G	T	T	C	T	T	C	T	G
MLV21-P1	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	G
MLV21-P2	G	C	C	A	T	A	C/G	C/A	T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV21-P3	G	C/T	C/T	A	T	G/A	G	A	T	G/A	T	C/T	G	G	T	T	C	T	T	C	T	G
MLV21-P4	G	C	C	A	T	A	G	A	C/T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV22	G	T	C/T	A	T	G	C/G	A	T	A	T	C/T	G	G	C	A	C	T	C	T	T	A
MLV22-P1	G	C/T	C	A	T	G/A	C/G	C/A	T	G/A	T	C/T	G/A	G	C	A	C	T	C	T	T	A
MLV22-P2	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	A
MLV22-P3	G	C/T	F	F	F	F	C/G	C/A	T	G/A	T	C	G/A	G	C	A	C	T	C	T	T	A
MLV22-P4	G	C/T	C	A	T	G/A	C	C/A	T	A	T	C	G/A	G	C	A	C	T	C	T	T	A
MLV24	C/G	C/T	C/T	A	T	G/A	C/G	A	T	G	T	T	G	G	C	A	C	T	T	T	T	G
MLV24-P1	G	C	C/T	A	T	G/A	C/G	A	T	G	T	T	G	G	C	A	C	T	T	T	T	G
MLV24-P2	G	C/T	C/T	A	T	G/A	G	A	T	G	T	T	G	G	C	A	C	T	T	T	T	G
MLV24-P3	G	C	C/T	A	T	G/A	C/G	A	T	G	T	T	G	G	C	A	C	T	T	T	T	G
MLV24-P4	G	C	C/T	A	T	G/A	G	A	T	G	T	C/T	G/A	G	C	A	C	T	T	T	T	G
MLV25	G	T	C/T	A	T	G	C/G	A	T	A	T	C	A	G	T	T	C	T	T	C	T	G
MLV25-P1	G	T	C/T	A	T	G/A	C/G	A	T	A	T	C	A	G	T	T	C	T	T	C	T	G
MLV25-P2	G	C/T	C/T	A	T	G/A	C/G	C/A	T	A	T	C	A	G	T	T	C	T	T	C	T	G
MLV25-P3	G	T	C	A	T	G/A	C/G	C/A	T	G/A	T	C	A	G	T	T	C	T	T	C	T	G
MLV25-P4	G	T	C	A	T	G/A	C	C/A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV26	G	C	C/T	A	T	G/A	C/G	A	T	A	T	C	A	G	T	T	C	T	T	C	T	G
MLV26-P1	G	C/T	C	A	T	A	G	A	T	G/A	T	C	A	G	T	T	C	T	T	C	T	G
MLV26-P2	G	C/T	C/T	A	T	G/A	G	A	T	G/A	T	C	A	G	T	T	C	T	T	C	T	G
MLV26-P3	G	C/T	C/T	A	T	G/A	C/G	A	T	G/A	T	C	A	G	T	T	C	T	T	C	T	G
MLV26-P4	G	C/T	C	A	T	A	C/G	A	T	A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV27	G	C/T	C/T	A	T	G	C/G	A	C/T	A	T	C/T	G	G	T	T	C	T	T	C	T	G
MLV27-P1	G	C	C/T	A	T	G/A	C	C/A	C/T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV27-P2	G	C	C/T	A	T	G/A	C/G	A	C/T	G/A	T	C	G	G	T	T	C	T	T	C	T	G
MLV27-P3	G	C	C/T	A	T	G/A	G	A	C/T	A	T	C	G	G	T	T	C	T	T	C	T	G
MLV27-P4	G	C/T	C	A	T	G/A	G	A	C/T	G/A	T	C/T	G	G	T	T	C	T	T	C	T	G
MLV28	G	C	C/T	A	T	G/A	C	C/A	T	G/A	T	T	G	G	C	A	C	T	C	C	C	G

Animal I.D.	115	193IN	555	150	326	773	804	817	402	408	567	276	494	4INDI	471	524	478	758	866	950	1009	1065	4476	4679	4904
MLV28-P1	G/A	A/T	T	G	C/T	G	G/A	G/A	A	C/G	C	C	T	C	A	T	G/A	C/G	C	C	G	C	C	T/A	G/T
MLV28-P2	G/A	A/T	T	G	C/T	G	G/A	G	A/C	C/G	C	C	T	C	C/A	T	G/A	C/G	C	C	G	C	C	T/A	G/T
MLV28-P3	G/A	A/T	T	G/C	C	G	G/A	G/A	A	C/G	C	C	T	C	C/A	T	G/A	C/G	C	C	G	C	C/T	A	G
MLV28-P4	G/A	A/T	T	G	C/T	G	G/A	G/A	A	C/G	C	C	T	C	C/A	T	G/A	C/G	C	C	G	C	C	T/A	G
MLV29	G/A	A/T	T	G/C	C	G	G/A	G/A	A	C/G	A/C	C	T	C/T	C	C/T	G/A	C/G	C/G	C	G	C	C/T	T/A	G
MLV29-P1	G/A	A/T	T	G	C	G	G/A	G	A	C/G	C	C	T	C	C/A	T	A	C	C	C	G	C	C	T/A	G
MLV29-P2	G	A	T	G	C	G	A	G/A	A	C	A	C	T	C	C	C/T	G/A	C/G	C	C	G	C	C	T/A	G
MLV29-P3	G	A	T	G/C	C	G	A	G/A	A	C	A/C	C	T	C	C/A	T	G/A	C/G	C	C	G	C	C/T	A	G/T
MLV29-P4	G/A	A/T	T	G/C	C	G	G/A	G/A	A	C/G	C	C	T	C	C/A	T	G/A	C/G	C	C	G	C	C	T/A	G
MLV30	G/A	T	T	G/C	C/T	G	G/A	G	A/C	C/G	C	C	T	C/T	C/A	C/T	G	G	C/G	C	G/A	C	C/T	T/A	G
MLV30-P1	G/A	A/T	T	G/C	C	G	G/A	G	A	C/G	C	C	T	C/T	C	C	G	G	C/G	C	G	C	C/T	A	G
MLV30-P2	G/A	A/T	T	G	C/T	G	G/A	G	A	C/G	A/C	C	T	C/T	C/A	C/T	G	G	C/G	C	G	C	C	T/A	G
MLV30-P3	G/A	A/T	T	G/C	C	G	A	G	A	C/G	C	C	T	C/T	C	C	G	G	C/G	C	G	C	C/T	A	G
MLV30-P4	G	A/T	T	G	C/T	G	G/A	G	A	C/G	A/C	C	T	C	C/A	C/T	G	G	C/G	C	G/A	C	C/T	A	G

Animal I.D.	419	515	184	389	516	582	426	810	368	533	939	71	138	361	15543	15558	15615	15616	15675	15714	15840	16127
MLV28-P1	G	C/T	C/T	A	T	G/A	C/G	C/A	T	G	T	C/T	G	G	C	A	C	T	C	C	C	G
MLV28-P2	G	C	C/T	A	T	G/A	C	C/A	T	A	T	C/T	G/A	G	C	A	C	T	C	C	C	G
MLV28-P3	G	C/T	C/T	A	T	G/A	C/G	C/A	T	G	T	C/T	G	G	C	A	C	T	C	C	C	G
MLV28-P4	G	C	C	A	T	A	C/G	C/A	T	A	T	C/T	G	G	C	A	C	T	C	C	C	G
MLV29	C/G	T	C	G/A	T	G	C/G	A	T	A	T	C	G	G/A	C	A	C	T	C	C	C	G
MLV29-P1	C/G	C/T	C	G/A	T	G/A	C/G	C/A	T	G/A	T	C	G	G/A	C	A	C	T	C	C	C	G
MLV29-P2	G	C/T	C	A	T	G/A	C/G	A	T	G/A	T	C	G	G	C	A	C	T	C	C	C	G
MLV29-P3	G	C/T	C	A	T	G/A	C	C/A	T	G/A	T	C	G/A	G/A	C	A	C	T	C	C	C	G
MLV29-P4	G	T	C	A	T	G/A	C/G	C/A	T	G/A	T	C	G/A	G	C	A	C	T	C	C	C	G
MLV30	C/G	T	C	G/A	T	G	C/G	A	T	G/A	T	C/T	G	G	T	T	C	T	T	T	T	G
MLV30-P1	G	T	C	A	T	G/A	G	A	T	G	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV30-P2	C/G	T	C	G/A	T	G/A	C/G	A	T	G/A	T	C	G/A	G	T	T	C	T	T	C	T	G
MLV30-P3	G	T	C	A	T	G/A	G	A	T	G	T	C/T	G/A	G	T	T	C	T	T	C	T	G
MLV30-P4	G	T	C	A	T	G/A	G	A	T	G/A	T	C/T	G/A	G	T	T	C	T	T	C	T	G

TABLE 9: Matching and Non-Matching Example Based on SNPTrack Analysis

SNP ID	P1S001- 245	P1S001- 421	P1S001- 48521522	P1S002- 368	P1S002- 533	P1S002- 41855052	P1S002- 41855371	P1S002- 41854995	P1S002- 48521356	P1S003- 774	P1S003- 805	P1S003- 48521502	P1S004- 478	P1S004- 758	P1S004- 866	P1S004- 39180829	P1S005- 4476	P1S005- 41855428	P1S005- 41854889	P1S005- 4904	P1S006- 426
Subject	48521522	41855052	41855371	41854995	48521356	41855003	48521502	41855420	48521532	39180829	41855428	41854889	41855387	48526443							
MLAC1	G	C	T	G	T	G	A	A/G	C/G	C	A	A	G/T	A/G							
PGG1	G	C/T	C	A/G	T	A/G	A	A/G	C	C	A	A	G	A							
PGG2	G	T	N/A	A/G	T	A/G	A	A	C	C	A	A	G/T	A							
PGG3	G	C/T	C/T	A/G	T	A/G	A/G	A	C	C	C/T	A	G	A							
PGG4	G	C/T	C/T	A/G	T	A/G	A	A	C	C	A	A	N/A	A							
PGG5	G	C/T	C/T	A/G	T	G	A/G	A	C	C	A	A	G/T	A							

SNP ID	P1S006- 810	P1S007- 361	P1S007- 48521483	P1S007- 41855060	P1S008- 115	P1S008- 48521463	P1S008- 41854832	P1S008- 41855019	P1S009- 111	P1S009- 118	P1S009- 41855461	P1S010- 471	P1S010- 524	P1S010- 48521329	P1S011- 197	P1S011- 41855068	P1S011- 268	P1S011- 39180952	P1S011- 41855109	P1S012- 276
Subject	41855183	48521483	41855060	41854832	41855019	41855461	41855518	41855519	41855518	41855461	41855518	48521329	41855297	41855297	41855068	39180952	268	39180952	41855109	39180862
MLAC1	C/G	G	C	A/G	A/G	D/I	D	A/G	G	A/G	C/T	C/T	C/T	C/T	C	C	C	C	C	C
PGG1	G	A/G	C	A/G	D/I	D	A/G	G	A	A	A	A	T	T	T	G	G	C	C	C
PGG2	G	A/G	C	A/G	D/I	D	A/G	G	A	A	A	A	T	T	T	G	G	C	C	C
PGG3	G	A/G	C	A/G	D/I	D	A/G	G	A	A	A	A	T	T	T	G	G	C	C	C
PGG4	G	A/G	N/A	A/G	D/I	D	A/G	G	A	A	A	A	T	T	T	G	G	C	C	C
PGG5	G	A	C	G	I	D	A	G	G	G	C	C	T	T	C/T	G	G	C	C	C

SNP ID	P1S012- 494	P1S013- 115	P1S013- 47272028	P1S013- 39180870	P1S014- 116	P1S014- 48521338	P1S014- 39181228	P1S014- 47272037	P1S014- 41855469	P1S015- 428	P1S015- 47272019	P1S016- 102	P1S016- 41855395	P1S016- 41854905	P1S016- 41855305	P1S017- 121	P1S017- 41855240	P1S019- 402
Subject	41855076	47272028	39180870	48521338	48521338	39181228	47272037	41855469	47272019	41855477	41855395	41854905	41855305	41855240	41855362			
MLAC1	N/A	G	C	A	A	C	C/T	A	C/T	C/T	C	A	A	T	C/T	A	A	
PGG1	C/T	A/G	C/T	A	A	C	T	A	C	C	C	A	A	T	C	C	A/C	
PGG2	C/T	G	T	A	A	C	T	A	C	C	C	A	A	T	C	C	A/C	
PGG3	C/T	G	T	A	A	C	T	A/C	C	C	C/G	A	A	T	C	C	A/C	
PGG4	C/T	A/G	T	A	A	C	T	A/C	C	C	C	A	A	T	C	C	A/C	
PGG5	T	A/G	T	A	A	C	T	A/C	N/A	C	C	A	A	T	C	C	A	

None of the mothers (PGG1-5) could be the parent of sample (MLAC1)

SNP ID	P1S001-245	P1S001-421	P1S001-421	P1S002-368	P1S002-533	P1S002-939	P1S002-939	P1S003-774	P1S003-805	P1S004-478	P1S004-758	P1S004-866	P1S005-4476	P1S005-4679	P1S005-4904	P1S006-426
Subject	48521522	41855052	41855052	41855371	41854995	48521356	48521356	41855003	48521502	41855420	48521532	39180829	41855428	41854889	41855387	48526443
MLAC2	G	C	T	A/G	A/G	T	G	A	A/G	A/G	C/G	C	C	A/T	G/T	A/G
PGG1	G	C/T	C	A/G	A/G	T	A/G	A	A/G	A/G	C	C	C	A	G	A
PGG2	G	T	N/A	A/G	A/G	T	A/G	A	A	A	C	C	C	A	G/T	A
PGG3	G	C/T	C/T	A/G	A/G	T	A/G	A/G	A	A	C	C	C/T	A	G	A
PGG4	G	C/T	C/T	A/G	A/G	T	A/G	A/G	A	A	C	C	C	A	N/A	A
PGG5	G	C/T	C/T	A/G	A/G	T	G	A/G	A/G	A	C	C	C	A	G/T	A

SNP ID	P1S006-810	P1S007-361	P1S007-71	P1S007-71	P1S008-115	P1S008-192	P1S008-192	P1S008-555	P1S009-111	P1S009-118	P1S010-471	P1S010-524	P1S011-197	P1S011-268	P1S011-75276	P1S012-276
Subject	41855183	48521483	41855060	41855060	48521463	41854832	41854832	41855019	41855518	41855461	48521329	41855297	41855088	39180952	41855109	39180862
MLAC2	C/G	G	C	C	A/G	D/I	D/I	D	A/G	G	A/C	C/T	C	G	C	C
PGG1	G	A/G	C	A/G	A/G	D/I	D/I	D	A/G	G	A	T	T	G	C	C
PGG2	G	A/G	C	A/G	A/G	D/I	D/I	D	A/G	G	A/C	T	T	G	C	C
PGG3	G	A/G	C	A/G	A/G	D/I	D/I	D	A	G	A	T	T	G	C	C
PGG4	G	A/G	N/A	G	G	D/I	D/I	D	A/G	G	A	T	C/T	G	C	C
PGG5	G	A	C	G	G	I	I	D	A	G	C	T	C/T	G	C	C

SNP ID	P1S012-494	P1S013-115	P1S013-515	P1S013-515	P1S014-116	P1S014-177	P1S014-177	P1S014-477	P1S014-93	P1S015-428	P1S016-471	P1S016-102	P1S016-127	P1S016-193	P1S017-121	P1S019-402
Subject	41855076	47272028	39180870	48521338	48521338	39181228	47272037	47272037	41855469	47272019	41855477	41855395	41854905	41855305	41855240	41855362
MLAC2	T	A/G	C/T	A	A	C	C/T	C/T	A	C	C	C	A	T	C	A/C
PGG1	C/T	A/G	C/T	A	A	C	T	T	A	C	C	C	A	T	C	A/C
PGG2	C/T	G	T	A	A	C	T	T	A	C	C	C	A	T	C	A/C
PGG3	C/T	G	T	A	A	C	T	T	A/C	C	C	C/G	A	T	C	A/C
PGG4	C/T	A/G	T	A	A	C	T	T	A/C	C	C	C	A	T	C	A/C
PGG5	T	A/G	T	A	A	C	T	T	A/C	N/A	C	C	A	T	C	A

Only PGG4 could be from the parent of MLAC2

MATERIALS AND METHODS

A. SNP identification and development of markers to trace or identify an individual animal

- [000107] **DNA Isolation:** Blood samples were collected and DNA was extracted from 30 sows from 10 different farms and from approximately 100 boars.
- [000108] **Bioinformatics:** Approximately 20 to 40 loci or genes from X chromosome and 10 loci or genes from Y chromosome were analyzed. Pig homologs from EST database were identified. PCR primers to amplify and sequence the identified genes and flanking sequences from pig DNA were designed.
- [000109] **Autosomal Markers:** Ligation mediated amplification (LMA) assays from known SNPs in the pig genome were developed. The sequences from flanking regions were expanded to identify additional SNPs. Genotype identification by LMA for about 50 DNA samples was performed to select informative markers.
- [000110] **Mitochondrial Markers:** D-loop region from 50 sows was sequenced to identify informative maternal SNPs. Additional mitochondrial DNA as needed was sequenced to increase the information content of the D-loop SNPs.
- [000111] **Developing X and Y SNPs:** Sequence tagged sites (STSs) from 5 to 10 individuals were amplified. Sequences were analyzed for SNPs. SNPs with minimum heterozygosity of 10% were identified. A mix of markers with heterozygosities between 10 to 40% was chosen.
- [000112] **Validation of autosomal markers:** Allele frequencies in a study population were determined and the power of these markers to trace or identify an individual was estimated. Additional candidate SNPs from databases were identified if necessary.
- [000113] **Mitochondrial LMA SNP assays:** LMA assays were developed and the remaining DNA samples were analyzed from sows to establish allele frequencies. Statistical analyses were performed to estimate the mitochondrial markers' power to distinguish between various maternal lineages.
- [000114] **LMA assays and Validation:** LMA assays were developed for the X and Y chromosome markers. The allele frequencies based on DNA samples from sows and/or boars were determined.
- [000115] **Marker Validation:** SNP assays were performed in a test population blindly to test the mitochondrial markers' ability to identify the maternal lineage.

[000116] **Statistical analysis:** SNP data from sex chromosome markers were combined with the SNP data from mitochondrial and autosomal SNPs. Statistical power of the combined markers was estimated to identify each individual in the study population using the experimental allele frequencies obtained during marker validation. Group markers in sets of 5 to 8 with at least 8 lineage markers were used to genotype all animals. Additional markers may be chosen based on the results obtained with the lineage markers. This process will be repeated until an animal can be identified with certainty. This process will create a SNP marker "tree" with lineage markers at the base, and various X chromosome and autosomal markers defining each branch. It is expected that the actual markers genotyped will differ between animals and will be chosen based on the marker tree generated from the population data collected. The emphasis will be on minimizing the number of markers to be genotyped and as a result of the cost of the test. The marker tree will be validated by blindly genotyping 50 individuals from the population.

B. Statistical analyses to evaluate the power and the number of markers required for traceability or identity studies.

[000117] Statistical analyses were developed to evaluate the power and the number of markers required for traceability and identity testings, and to evaluate potential fluctuations in statistical power in actual experimental situations. Statistical power for traceability testing is measured by exclusion probabilities with and without known sires. The following statistical analyses were conducted.

[000118] The number of autosomal markers required to achieve a given maximum average exclusion probability for the number of alleles varied from 2 to 10, assuming the use of mtDNA testing with a haplotype (SNPTrack) frequency of 1/10. Known sires and unknown sires were considered.

[000119] The effect of mtDNA polymorphisms on the number of autosomal markers required was considered. The results show that the increase in the number of mtDNA haplotypes (SNPTracks) from 10 to 12 results in only a negligible reduction in the number of markers required, *e.g.*, the reduction if 0.2~0.5 markers assuming four alleles with equal allele frequency for each autosome marker, depending on the threshold value of the overall exclusion probability.

[000120] The number of autosomal markers required with mixed semen from known sires was considered. The wrong inclusion rate with genotyping mixed semen was

compared to that without genotyping mixed semen. This analysis shows that genotyping mixed semen from known sires significantly reduces the wrong inclusion rate and hence significantly reduces the number of markers required to achieve a given exclusion power.

[000121] The number of autosomal markers required with different sets of allele frequencies assuming 3 alleles per marker was used to evaluate potential fluctuation in exclusion power due to unequal allele frequency. The results show that the average exclusion power decreases as the differences among allele frequencies increase. The results show that the average exclusion power decreases as the differences among allele frequencies increase.

[000122] The results for identity testing show that identity testing is a much easier task than paternity testing. The number of markers sufficient for paternity testing is more than sufficient for identity testing.

a) Algorithms

[000123] Algorithms used in the statistical analyses included standard mathematical expressions for exclusion probabilities of autosome markers available in the literature, and five other mathematical expressions derived taking into consideration of the genotyping of mtDNA and mixed semen from known sires, and a likelihood ratio test for identity testing that is found in the literature.

[000124] The five other mathematical expressions derived are: 1) Exclusion probability of autosome and mtDNA markers when the sire is known, 2) Exclusion probability of autosome and mtDNA markers when the sire is unknown, 3) Number of markers required to achieve a given exclusion power with equal or unequal allele frequencies and mtDNA markers, 4) Exclusion power of with equal or unequal allele frequencies, mtDNA markers, and mixed semen from known sires, and 5) Number of markers required to achieve a given exclusion power with equal or unequal allele frequencies, mtDNA markers, and mixed semen from known sires.

b) Computer programs

[000125] Six computer programs in SAS computer language (SAS Institute Inc., Cary, NC) were developed to evaluate the power and the number of markers required for paternity and identity testings, and to evaluate potential fluctuations in statistical

power in real situations. Statistical power for paternity testing is measures by exclusion probabilities with and without known sires.

[000126] The six computer programs used are: 1) `exclud_max.sas`: implements statistical analysis (computer script 1); 2) `exclud_max_mt.sas`: implements statistical analysis (computer script 2); 3) `exclud_err_mix.sas`: implements statistical analysis (computer script 3); 4) `exclud_a3b.sas`: implements statistical analysis (computer script 4); 5) `exclud_a4b.sas`: implements statistical analysis (computer script 5); and 6) `identity.sas`: implements statistical analysis (computer script 6).

[000127] 1) `exclud_max.sas`

/* Computer program 1) for statistical analysis 1) */

```
data exlude;
prohap=1/10;
t100=100000;
mil=1000000;
mil10=10*mil;
mil100=100*mil;
bil=1000*mil;
ext100=1-1/t100;
ex1=1-1/mil;
ex10=1-1/mil10;
ex100=1-1/mil100;
exbil=1-1/bil;
do n=2 to 10 by 1;
  q0=(n-1)*(n**2-3*n+3)/(n**3);
  q1=1-(2*n**3+n**2-5*n+3)/(n**4);
  q1_gm=(n-1)*(n**3-n**2-2*n+3)/(n**4);
  d1=log(1-q1);
  d0=log(1-q0);
  n1_t100=log(1-ext100)/d1;
  n0_t100=log(1-ext100)/d0;
  n1_1=log(1-ex1)/d1;
  n1_10=log(1-ex10)/d1;
  n1_100=log(1-ex100)/d1;
  n1_bil=log(1-exbil)/d1;
  n0_1=log(1-ex1)/d0;
  n0_10=log(1-ex10)/d0;
  n0_100=log(1-ex100)/d0;
  n0_bil=log(1-exbil)/d0;
  n1_t100_mit=log((1-ext100)/prohap)/d1;
  n0_t100_mit=log((1-ext100)/prohap)/d0;
  n1_1_mit=log((1-ex1)/prohap)/d1;
  n1_10_mit=log((1-ex10)/prohap)/d1;
  n1_100_mit=log((1-ex100)/prohap)/d1;
  n1_bil_mit=log((1-exbil)/prohap)/d1;
  n0_1_mit=log((1-ex1)/prohap)/d0;
  n0_10_mit=log((1-ex10)/prohap)/d0;
  n0_100_mit=log((1-ex100)/prohap)/d0;
  n0_bil_mit=log((1-exbil)/prohap)/d0;
output;
```

-51-

```

end;
  keep n q0 q1;
  keep n1_1 n1_10 n1_100 n1_bil n0_1 n0_10 n0_100 n0_bil n1_t100 n0_t100
n1_t100_mit
  n0_t100_mit;
  keep n1_1_mit n1_10_mit n1_100_mit n1_bil_mit n0_1_mit n0_10_mit
n0_100_mit n0_bil_mit;
  proc print;
run;

```

[000128] exclude.max_mt.sas

/* Computer program 2) for statistical analysis 2) */

```

data exlude;
t100=100000;
mil=1000000;
mil10=10*mil;
mil100=100*mil;
bil=1000*mil;
ext100=1-1/t100;
ex1=1-1/mil;
ex10=1-1/mil10;
ex100=1-1/mil100;
exbil=1-1/bil;
do nhap = 10 to 20;
  prohap=1/nhap;
  do n=4 to 4 by 1;
    q0=(n-1)*(n**2-3*n+3)/(n**3);
    q1=1-(2*n**3+n**2-5*n+3)/(n**4);
    q1_gm=(n-1)*(n**3-n**2-2*n+3)/(n**4);
    d1=log(1-q1);
    d0=log(1-q0);
    n1_t100=log(1-ext100)/d1;
    n0_t100=log(1-ext100)/d0;
    n1_1=log(1-ex1)/d1;
    n1_10=log(1-ex10)/d1;
    n1_100=log(1-ex100)/d1;
    n1_bil=log(1-exbil)/d1;
    n0_1=log(1-ex1)/d0;
    n0_10=log(1-ex10)/d0;
    n0_100=log(1-ex100)/d0;
    n0_bil=log(1-exbil)/d0;
    n1_t100_mit=log((1-ext100)/prohap)/d1;
    n0_t100_mit=log((1-ext100)/prohap)/d0;
    n1_1_mit=log((1-ex1)/prohap)/d1;
    n1_10_mit=log((1-ex10)/prohap)/d1;
    n1_100_mit=log((1-ex100)/prohap)/d1;
    n1_bil_mit=log((1-exbil)/prohap)/d1;
    n0_1_mit=log((1-ex1)/prohap)/d0;
    n0_10_mit=log((1-ex10)/prohap)/d0;
    n0_100_mit=log((1-ex100)/prohap)/d0;
    n0_bil_mit=log((1-exbil)/prohap)/d0;
    output;
  end;
end;

```

-52-

```

keep prohap nhap;
* keep n1_1 n1_10 n1_100 n1_bil n0_1 n0_10 n0_100 n0_bil n1_t100 n0_t100;
keep n1_t100_mit n1_1_mit n1_10_mit n1_100_mit n1_bil_mit
    n0_t100_mit n0_1_mit n0_10_mit n0_100_mit n0_bil_mit;
proc print;
run;

```

[000129] exclude_err_mix.sas

```
/* Computer program 3) for statistical analysis 3) */
```

```

data exclude;
mit_prob=0.10;
do m=1 to 36 by 1;
do n=4 to 4 by 1;
q0=(n-1)*(n**2-3*n+3)/(n**3);
q1=1-(2*n**3+n**2-5*n+3)/(n**4);
wrong_inc0=log10((1-q0)**m);
wrong_inc1=log10((1-q1)**m);
q0_m=1-(1-q0)**m;
q1_m=1-(1-q1)**m;
wrong_inc0_mit=(1-q0)**m*mit_prob;
wrong_inc_mix=1-qmix_m_mit;
q0_m_mit=1-(1-q0)**m*mit_prob;
q1_m_mit=1-(1-q1)**m*mit_prob;
qmix_m_mit=q0_m*q1_m_mit + (1-q0_m)*q0_m_mit;
wrong_inc0_mit=(1-q0)**m*mit_prob;
wrong_inc_mix=1-qmix_m_mit;
err_ratio = wrong_inc0_mit/wrong_inc_mix;
output;
end;
end;
data _null_;
set exclude;
* user needs to modify the following statement for file location;
file "file location";
put m 4. n 4. wrong_inc0_mit 15.12 wrong_inc_mix 15.12 q0_m_mit 15.12
q1_m_mit 15.12
qmix_m_mit 15.12 err_ratio 10.2;
cards;
run;

```

[000130] exclude_a3b.sas

```
/* Computer program 4) for statistical analysis 4) */
```

```

data exlude;
input p1 p2 p3;
p_mt_hap=0.10;
mit_ex=1-0.1;
l1=1/prohap;
mil=1000000;
mil10=10*mil;
mil100=100*mil;
bil=1000*mil;
ex1m=log10(1/mil);
ex10m=log10(1/mil10);
ex100m=log10(1/mil100);

```

-53-

```

exbil=log10(1/bil);
do;
  p1m=1-p1;
  p2m=1-p2;
  p3m=1-p3;
  p1ms=p1m**2;
  p2ms=p2m**2;
  p3ms=p3m**2;
  p12=p1*p2;
  p13=p1*p3;
  p23=p2*p3;
  p12m=1-p1-p2;
  p13m=1-p1-p3;
  p23m=1-p2-p3;
  p24m=1-p2-p4;
  p34m=1-p3-p4;
  p12ms=p12m**2;
  p13ms=p13m**2;
  p23ms=p23m**2;
  p1s=p1*p1;
  p2s=p2*p2;
  p3s=p3*p3;
  p4s=p4*p4;
  u12=4-3*(p1+p2);
  u13=4-3*(p1+p3);
  u23=4-3*(p2+p3);
  v12=p12*p12ms;
  v13=p13*p13ms;
  v23=p23*p23ms;

  q0=p1s*p1ms+p2s*p2ms+p3s*p3ms
    +2*(v12 + v13 + v23);
  q1=p1*p1ms+p2*p2ms+p3*p3ms
    -(p1s*p2s*u12+p1s*p3s*u13+p2s*p3s*u23);
  d1=log10(1-q1);
  d0=log10(1-q0);

  n1_1=int(ex1m/d1+0.5);
  n1_10=int(ex10m/d1+0.5);
  n1_100=int(ex100m/d1+0.5);
  n1_bil=int(exbil/d1+0.5);

  n0_1=int(ex1m/d0+0.5);
  n0_10=int(ex10m/d0+0.5);
  n0_100=int(ex100m/d0+0.5);
  n0_bil=int(exbil/d0+0.5);

  n1_1_mt=int((ex1m-log10(p_mt_hap))/d1+0.5);
  n1_10_mt=int((ex10m-log10(p_mt_hap))/d1+0.5);
  n1_100_mt=int((ex100m-log10(p_mt_hap))/d1+0.5);
  n1_bil_mt=int((exbil-log10(p_mt_hap))/d1+0.5);

  n0_1_mt=int((ex1m-log10(p_mt_hap))/d0+0.5);
  n0_10_mt=int((ex10m-log10(p_mt_hap))/d0+0.5);
  n0_100_mt=int((ex100m-log10(p_mt_hap))/d0+0.5);
  n0_bil_mt=int((exbil-log10(p_mt_hap))/d0+0.5);

```

-54-

```

output;
keep ex1m ex10m ex100m exbil n1_1 n1_10 n1_100 n1_bil n0_1 n0_10 n0_100
n0_bil
    n1_1_mt n1_10_mt n1_100_mt n1_bil_mt n0_1_mt n0_10_mt n0_100_mt
n0_bil_mt;
end;
cards;
0.33 0.33 0.34
0.4 0.3 0.3
0.5 0.3 0.2
0.6 0.3 0.1
0.7 0.2 0.1
;

title 'Number of loci required to achieve a given exclusion power';
proc print;
run;

```

[000131] exclude_a4b.sas

```

/* Computer program 5) for statistical analysis 5) */
data exlude;
    input p1 p2 p3 p4;
    p_mt_hap=0.10;
    mit_ex=1-0.1;
    l1=1/prohap;
    mil=1000000;
    mil10=10*mil;
    mil100=100*mil;
    bil=1000*mil;
    ex1m=log10(1/mil);
    ex10m=log10(1/mil10);
    ex100m=log10(1/mil100);
    exbil=log10(1/bil);
do;
    p1m=1-p1;
    p2m=1-p2;
    p3m=1-p3;
    p4m=1-p4;
    p1ms=p1m**2;
    p2ms=p2m**2;
    p3ms=p3m**2;
    p4ms=p4m**2;
    p12=p1*p2;
    p13=p1*p3;
    p14=p1*p4;
    p23=p2*p3;
    p24=p2*p4;
    p34=p3*p4;
    p12m=1-p1-p2;
    p13m=1-p1-p3;
    p14m=1-p1-p4;
    p23m=1-p2-p3;
    p24m=1-p2-p4;
    p34m=1-p3-p4;

```

-55-

```

p12ms=p12m**2;
p13ms=p13m**2;
p14ms=p14m**2;
p23ms=p23m**2;
p24ms=p24m**2;
p34ms=p34m**2;
p1s=p1*p1;
p2s=p2*p2;
p3s=p3*p3;
p4s=p4*p4;
u12=4-3*(p1+p2);
u13=4-3*(p1+p3);
u14=4-3*(p1+p4);
u23=4-3*(p2+p3);
u24=4-3*(p2+p4);
u34=4-3*(p3+p4);
v12=p12*p12ms;
v13=p13*p13ms;
v14=p14*p14ms;
v23=p23*p23ms;
v24=p24*p24ms;
v34=p34*p34ms;

q0=p1s*p1ms+p2s*p2ms+p3s*p3ms+p4s*p4ms
  +2*(v12 + v13 + v14 + v23 + v24 +v34);
q1=p1*p1ms+p2*p2ms+p3*p3ms+p4*p4ms

-(p1s*p2s*u12+p1s*p3s*u13+p1s*p4s*u14+p2s*p3s*u23+p2s*p4s*u24+p3s*p4s*u34);
d1=log10(1-q1);
d0=log10(1-q0);

n1_1=int(ex1m/d1+0.5);
n1_10=int(ex10m/d1+0.5);
n1_100=int(ex100m/d1+0.5);
n1_bil=int(exbil/d1+0.5);

n0_1=int(ex1m/d0+0.5);
n0_10=int(ex10m/d0+0.5);
n0_100=int(ex100m/d0+0.5);
n0_bil=int(exbil/d0+0.5);

n1_1_mt=int((ex1m-log10(p_mt_hap))/d1+0.5);
n1_10_mt=int((ex10m-log10(p_mt_hap))/d1+0.5);
n1_100_mt=int((ex100m-log10(p_mt_hap))/d1+0.5);
n1_bil_mt=int((exbil-log10(p_mt_hap))/d1+0.5);

n0_1_mt=int((ex1m-log10(p_mt_hap))/d0+0.5);
n0_10_mt=int((ex10m-log10(p_mt_hap))/d0+0.5);
n0_100_mt=int((ex100m-log10(p_mt_hap))/d0+0.5);
n0_bil_mt=int((exbil-log10(p_mt_hap))/d0+0.5);

output;
keep ex1m ex10m ex100m exbil n1_1 n1_10 n1_100 n1_bil n0_1 n0_10 n0_100
n0_bil
  n1_1_mt n1_10_mt n1_100_mt n1_bil_mt n0_1_mt n0_10_mt n0_100_mt
n0_bil_mt;

```

-56-

```
end;  
cards;  
0.25 0.25 0.25 0.25  
0.4 0.3 0.2 0.1  
0.5 0.2 0.2 0.1  
0.6 0.2 0.1 0.1  
0.7 0.1 0.1 0.1  
;
```

```
title 'Number of loci required to achieve a given exclusion power';  
proc print;  
run;
```

```
6)identity.sas  
/* Computer program 6) for statistical analysis 6) */
```

```
data exlude;  
prohap=0.10;  
mit_ex=1-0.1;  
l1=1/prohap;  
mil=1000000;  
mil10=10*mil;  
mil100=100*mil;  
bil=1000*mil;  
p_mil=1/mil;  
p_mil10=1/mil10;  
p_mil100=1/mil100;  
p_bil=1/bil;
```

```
do n=2 to 10 by 1;  
  n2=n**2;  
  pii=1/n2;  
  pij=2/n2;  
  kii_mil = -log10(p_mil)/(2*log10(n));  
  kii_mil10 = -log10(p_mil10)/(2*log10(n));  
  kii_mil100 = -log10(p_mil100)/(2*log10(n));  
  kii_bil = -log10(p_bil)/(2*log10(n));  
  
  kij_mil = log10(p_mil)/(log10(2)-2*log10(n));  
  kij_mil10 = log10(p_mil10)/(log10(2)-2*log10(n));  
  kij_mil100 = log10(p_mil100)/(log10(2)-2*log10(n));  
  kij_bil = log10(p_bil)/(log10(2)-2*log10(n));
```

```
output;  
keep kii_mil kii_mil10 kii_mil100 kii_bil kij_mil kij_mil10 kij_mil100  
kij_bil;  
end;  
proc print;  
run;
```

C. Exclusion probabilities with autosome and mtDNA markers

Exclusion probabilities with autosome and mtDNA markers

The overall exclusion probability with or without known sire for a marker set with autosome and mtDNA markers was derived based on previously available exclusion probabilities for autosome markers.

Let

Q_{1k} = exclusion probability of locus k with known genotypes for the sire and offspring,
 Q_{0k} = exclusion probability of locus k with known genotype for the offspring only.

Then,

$$Q_{1k} = \sum_{i=1}^n p_i(1 - p_i)^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i^2 p_j^2 (4 - 3p_i - 3p_j) \quad (1)$$

(Jamieson, 1965, 1995; Garber and Morris, 1983; Weir, 1996)

$$Q_{0k} = \sum_{i=1}^n p_i(1 - p_i)^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i p_j (1 - p_i - p_j)^2 \quad (2)$$

(Garber and Morris, 1983).

Assuming equal allele frequency, equations (1-2) reduces to

$$Q_{1k} = \frac{(n-1)(n^3 - n^2 - 2n + 3)}{n^4} \quad (3)$$

$$Q_{0k} = \frac{(n-1)(n^2 - 3n + 3)}{n^3} \quad (4)$$

[000132]

The overall exclusion probability with or without known sire for a marker set with autosomal and mtDNA markers was derived based on standard exclusion probabilities for autosomal markers.

D. Analysis of paternity testing

[000133]

Equations (1-2) were implemented in computer programs 4 and 5, and equations (3-4) were implemented in computer programs 1, 2, and 3. Note that equations 1 and 2 are general and covers the cases of equations 3 and 4. These two sets of equations were implemented to provide a mutual check for the correctness of the programming code.

[000134]

Let p = the equal haplotype frequency assumed for each mtDNA haplotype

$Q_{i,mit}$ = probability that a random individual is excluded as the parent by at least one autosome locus or the mtDNA haplotype when m autosome markers and one mtDNA haplotype are genotyped

[000135] Then,

$$Q_{i,mit} = 1 - \prod_{k=1}^m (1 - Q_{ik}), \quad i = 1, 0 \quad (5)$$

Equation (5) is used in computer programs 1 through 5.

E. Exclusion probabilities with added genotyping for mixed semen

[000136] A common practice in commercial swine production is the use of mixed semen from a number of sires. If the mixed semen on a dam is genotyped, the exclusion is expected to improve, but none of the above mathematical expressions provide the correct estimate of exclusion probability with added genotyping for mixed semen.

[000137] Let Q_0 = probability that a random individual is excluded as the parent by at least one autosome locus when no known parent is present. Q_{mix} = probability that a random individual is excluded as the parent by at least one autosome locus or the mtDNA haplotype when the sires that were sources of the mixed semen are genotyped for the m autosome markers (potential dam and sow genotyped for autosome and one mtDNA haplotype).

[000138] The probability of Q_0 can be used as the probability that the true sire of the disputed offspring can be determined. If any random individual is included as a potential true sire, two of the sires contributing to the mixed semen will be included. In this case, identifying the true sire is considered failed. The mathematical expressions for Q_0 and Q_{mix} are

$$Q_0 = 1 - \prod_{k=1}^m (1 - Q_{0k}) \quad (6)$$

$$Q_{mix} = Q_0 Q_{1,mit} + (1 - Q_0) Q_{0,mit} \quad (7)$$

[000139]

The equations 6-7 are implemented by computer program 3.

F. Number of autosome markers required

[000140] The number of autosome markers required is derived assuming all autosome markers have the same allele frequencies. Under this assumption, the number of markers required is

$$n = \log[(1 - Q)/p] / \log(1 - Q_i), i = 0, 1 \quad (8)$$

[000141] where Q = the required overall exclusion probability, p = mtDNA haplotype frequency of sow being tested, and Q_i is given by equation (1) or (2). Equation (8) is implemented in computer programs 1 through 5.

[000142] The analysis of paternity testing is implemented by a computer program. The program will conduct an exhaustive allele matching analysis between the offspring, all potential parents, and any known parent (in some cases the sire may be known). The final results of the analysis will identify the true parent, and a likelihood ratio test showing the reliability of test results.

G. Estimation of allele frequencies

[000143] Estimates of allele frequencies affect the reliability statement about the testing results but do not affect the actual testing process. Genotyping the entire population gives the most reliable estimates of allele frequencies but is the most costly and time consuming method. A two-step strategy for estimating allele frequencies is proposed. The first step is to genotype sires and dams and then to predict the population allele frequencies based on the alleles observed in the breeding population (sires and dams) and the relative contribution of each sire and dam to the next generation. This first step is possible because DNA samples from sires and dams are available given that DNA bank for sires and dams is in place and that breeding records are available. The second step is to update the estimates of allele frequencies as more genotyping results become available.

H. Sequences and SNP data of the markers used for SNPTrack analysis in swine.

[000144] **PBE3**

```

1 ctgacagtta aagactgccc aacagtgaag tgaactgcct aaaaaacagt gagttttcta
61 ttttttatgt gttcaaata aggaaaaata aatctgtcca ttatggggat aaggRgatac
121 cagtgttcaa ggggagttaa aacaaaaaga tttctaattgt accttcaaatt tcttaagatt
181 ccatgaaactg (TG) aattttatataggaataa aaaagtgaat ctatttctctt gatattgcaaa
241 gatgaggaaa aaagattttac atgataaaay ttcaaaataa atcgttttgcc attttaagct
301 gtattgttcg agctcaagaa ctttctttta caatattttcc atcttttctaa tttataatta

```

-60-

361 tccaataaaa tatacaatta cctaccactc caaatttttaa agtaaataatg tttgagatca
 421 atgtgcagat gaaaggtytt atttgtataa gaggaaagat agtgctatgt aaatacccct
 481 ttcccatcaa gtatattcct atgcacttcc ataaaggcaa ttcagtgtgt attttcacag
 541 gatccttggttattg (G)ttcatttttaggtctactgacgaagc agaccttcag aaaaatattt
 601 accctgaatt agagsatcaa gatggtggat gagtaagaca tggagcttac cttcttcc

The sequence in () is an in/del. In/del at position 556 (G insertion).

Three SNPs:	Position	115	192/193	556 indel	Estimated Frequency
	G	A	G	18%	
	G	A	T	34%	
	G	T	T	15%	
	A	T	T	33%	

Two SNPs:	Position	115	192/193	Estimated Frequency
	G	A	52%	
	G	T	15%	
	A	T	33%	

[000145] ACY-STS7:

R = A/G; Y=C/T; K=G/T; M=A/C; S=G/C; W=A/T; V=A/C/G

121 cagcaggatt tttgctgagt ttttttggaa accccctcag gaccaacccc cccccccca
 181 aaaagtatta agcaccaaag ttaatagaag agattcacag caaacaaggc agaaccagag
 241 accaSgggta caggggagac aacaaaccag gaccagggt caatctttct gctcccaccc
 301 tacagcctca gtcttctatg ctaatcctga gaaatcccta gcatgggaag ggacactgca
 361 aagcactgta ctgacctagc actggatcag atcaagggtca tatggctggt caatgagcaa
 421 Ygtaaaactt acaaggtact ggggtacacac agccaagggc atcccttccc ttgaaaagct
 481 cttaagccag ggagataaga caacctgccc tcagaaggca gggttacactt gcctaggggg
 541 ttataacctg gccagtaaa ggtcaggcaa agctttacta tgggcctggc agagcatgaa
 601 gtccaggcaa aagctggcta ggcagagaaa aattgtgggc tttggtaggc caagtaagat
 661 gaaggacact taataataat agcactcagg caggggctga agccagcaaa ggctacatga
 721 aagatcctga ctgtgcaa at ggagccaaag agacacactt ctgtgtgatt ccgagcacia
 781 actcaccccc ttaagagatt catatcgttg tgatcggggg ttcttgatgc cgtttctgtg
 841 ccatttttcgg gctgtggaga agtaaagcat taggtcaaca aaatagattt cccccaataa
 901 gaccactacc ccagc

There are 3 expected alleles when combining the information from both positions:

Position	245	421	
	G	C	52%
	C	C	18%
	G	T	30%

[000146] EG-STS7

1 ggtccctgyg ggtycacgtg ggttggtgtc taccgtctt cacaagctgg tactgatttg
 61 gtacgttctc tgcccttatgg gttctgtgct actgatctat atgtcttctt ggttcattca
 121 tgccctgagg gcttttagatt agttggcatt gtttacgggt aataccaaca gtttaacactt
 181 ataccagaa ctcaccacgt cccggggcac agctgcactg cgtgtatata aattccttcg

-61-

```

241 cccctaaggg gaggtacttc tatgaaccct gctttactaa cgaccaaagt gagcccagac
301 gtcagggtcgc tttaRgcac atagtgact tgatcccagg gtggctctgc tgccacttgc
361 cgatctgtct tggttgacat gggctgggct gtcccttaga gtcagacctt tcccaggggc
421 aaaggccact acaagtcagg ggcctaagca gcaaagctga ccatggcctc gccagctcac
481 cagccttccc tggctccctg ttgcctgcag ggtgtggtcc tgctcrggcg cttcctgttt
541 tcctctccaa gacttcttcc ctcaactctgc ccaaaacatc cttcttcccc ttctgcatcc
601 caccagctcc aacgtaggct tcaagatgYc tcctccagga agtcctccca gctgtgctcc
661 tctccacact ccccgctcag ttgatgtctc crccgcacrc acgtccctca tccagcactt
721 cctgtgacag tgcttctccc cctgcatctc ccccggtgag cctcagactg gccRttcccg
781 gaagagcrgt gaYrtggatg agtgRccag agttagcRac cttagagctga ggggccatct
841 cccagtcctg tggcccttac tcccagccg caccccctYg gRcaggggagc acagggaggg
901 ctgctggtgt gttctagcca tggcccgatg accyttgcYg cctcccatg ctgtgttctt
961 gggctgggga aggttctcca caggggaagg agaggttgac aggagagccc cctgcccta
1021 Ytgccctggg gacac

```

Positions to score

Three SNPs:

Position 774*	805	817	Estimated frequency
A	A	G	17%
G	A	A	15%
G	A	G	22%
G	G	G	46%

Same as 773 on previous sheets

Alternate Two SNPs:

Position 793*	805	Estimated frequency
T	A	17%
C	A	37%
C	G	46%

Same as 792 on previous sheets

Alternate SNPS

position 879	
C	73%
T	27%

Position 882	
A	83%
G	17%

[000147]**RYRA-STS6**

```

1 gaccaagagc tgcagcaccg tgtggagtcc ctggcagcct ttgcagaacg ctacgtggac
61 aagctccagg ccaaccagag ggaccgctat ggcatacctca tgaaggcctt caccatgacc
121 gctgccgaga ctgcccagcg tactcgcgag ttccgctccc caccccagga gcaggtcctc
181 taaccccaaa actcagctgg ccttactgtc tcaacctcag cctctccctt tactctgac
241 actgatggca ctcaacctct aaacctgggc ttgacctctg atcctgtggg tatacttctc
301 tcttgctccc ctacctctct ctgaccaga tttcragatc agccagact gaccctaagt
361 cctttcaaac ctttgatctc ccagatatc ctcagtaact cMtgactSca gacaggggct
421 cagttggatc ttagatcctt gacctcagag ttcctgctcc ggggtctctg accctcattc

```

-62-

481 taacctttga ccttccttag atcaacatgc tattgcactt caaagatggc gaggatgagg
 541 aagattgtcc tcttcctgat gagatcMggc aggatttgct ggaattccat caagacctgt
 601 tgactcactg tggtaagaga ggatatcagg gaatcctctt cccagtttt ttctcgagac
 661 ctctctgaaa gtttccttaa gatttcctga tcttgagatt cccgtcgtgg cgcagtgggt
 721 macgaatccg actaggaacc atgaggttgc gggttcggtc cctgcccttg ctcagtgggt
 781 tamygatccg gcgttgccgt gagctgtggt gtaagggtgc aaaccagct caa

Positions to score

Two SNPs:

Position	408	567	Estimated frequency
	C	A	35%
	C	C	43%
	G	C	22%

Three SNPs:

	402	408	567	Estimated frequency
	A	G	C	22%
	A	C	C	22%
	A	C	A	36%
	C	C	C	22%

[000148]

PBE59

1 ttctaaagtt cagcatactt cactagtgat acatgtctta Ytgatacttc cttaagagtt
 61 atgtgcttac ctgcctaggc cctccctcca ctagatggct cagccctggg gatcaggYgt
 121 tatctcctta gctgcatgaa gctragatMg tgtgttggtgc acaccagaat ccacctgcgt
 181 tcaacaccta gctctggagc tcctgctatg gaccaggcat tgtttctgggt gccrctgatg
 241 cagYggagag caaatcagat cccacaaacc catgaYgctc rcatgtgcat Racggaggaa
 301 aaaatatcaa ggaagaagca aatgaaatga gaatacagga ctctgtacct agtccctact
 361 arccagaagt tgtctccaaa grttttcatt ttctatgccy gtggatgatg gtcagaaaga
 421 agatctgcct gtaatcatgt tctcgaaggg atccaagacc tymagcaacc agaaaggaac
 481 tacttcacag taaYactgtt ccaaaccaac agtaagatgc ccgttccctc acttcgcttc
 541 catcttcttt aacrtcaagc agtccttgga gctagctact tcttagtcgt aagaactcga
 601 acgtacataa cgtatttgca gtttccaaag cacatttcc

Two SNPs: Position	276	494	Estimated frequency
	C	T	20%
	C	C	30%
	T	C	50%

[000149]

PBE43

1 cctgaagaat tgatacattha atgtgctatt tcatagcgtg ataagaatgt gcctttgcca
 61 tctcctttga aatgcaaaca tctttattct ttaggggaga ctttgtttac tttgattcaa
 121 cagtgmaaaa aattgggaat tagaaacctt tctgtagttt cccagaagct ggctcttagc
 181 acagatthttt ggtttctctc actggagttg acttgcatcg aaagttggga gcaaccctaa
 241 aaggtatgac ctaaaatcaa gctgggtggca gagtagggga gtctgtacat agcgccgggg
 301 gttctgagcgtgc (tgtagtgt tgtgtastgt akttctsmgt) gtggggaatc ctcaagacag
 361 ggagtccyag gggcctgtag gtattcctct tcctgaaatc atggaatggg tgagccggaa
 421 ggagaacctt ctattctttg ctggccttat ttctttttck ttccctctca Maagttacag
 481 aggtggcttc acagatccag gtcctgctgg agatcttcgg ctgYccctga gaakccagga
 541 agatttttac taaaaaytta ctyttccatc tcctctgcta saaaggcgc cactgtcgct

-63-

601 ttggcctcca cagaggccac aacaccctca gctccagagt cttcactgaa tgtacctgct
 661 ttcacatgaa ca

Three SNPS: Position	314	471	524	Estimated frequency
	T	C	C	21%
	G	A	T	25%
	G	C	T	40%
	G	C	C	14%

[000150] GALT

1 ggatccttaa cccactgagt gaggccaggg atcgaatttg cattctcgta gatactgggc
 61 agattttgtt ctgctgggcc accatgggaa ctccctgggt ttgtctatat atattttttt
 121 ttttttttgt cttttttgcc atttcttggg ccgctcctgc ggcatatgga ggttcccagg
 181 ctaggggtcg aatcggagct gtagccacca gcctacgcca gagccacagc aacgtgggac
 241 ccgagccgag tctgcaacct ataccacagc tcacggcaac gccagatccc ttaaccact
 301 gagcaaggcc agggaccgaa cccgcaacct catggttctt agtcggattc gttaaccact
 361 gcgccacgac gggaaactccc ggttttgtct atttttgaac gttaaataaa tgcaagcatc
 421 cagggctgct ttgactcagt accatgtgtg agatttacc tgttgatgtc agcagctRtg
 481 gctggttcct tctcacggat gtgtgtgacc ctacactgga ccacacctga tctggctgat
 541 gatgggcctt ggggtttttc cagcttttgg tcccakgtca cgtctctgtt tgaacttaaa
 601 tgcacttgct ttcaggtatt aatctggggc ggaatgactg gaacatgagg tgttggttgg
 661 tcagctttag tagatgccag cagggaggat ttcagtagtt tattaagcag atcttgaaga
 721 ctgtggtcaa ctagctcatg cccagagga gggggcgStg aatttcttcc ccagaacagg
 781 agtgagaagc taaattaggc atccatccgc tggaagytag gggggcagtt cttggctcct
 841 ttctgtcagg tttcggcccc ttctcSttag tctggggttt ctaggctcta ctcccaggaa
 901 gwgctctggg ccacttgga agaatgggtg ggggggctyt gagcccctac ttacttcatt
 961 tccctccttc agccaaarcc yctgtgtcc tctgttttac atagtggggt tctgagaatg
 1021 acttywtttt tttttttttt tttttwaaag ctttagctr t kgcgacattt acaaatccmc
 1081 tgctgtgagg tctcttccag ctaggaaatt ctattttggr ascagraggt ggggtgtggr
 1141 agggttaagc attattcagc caaagagttg ggttgggcct cagtacctt ttgaagttct
 1201 tatagcttgg cttgccatgc aggagatctc agaacattct ataaaaatag tgttcaaaca
 1261 gaacaacttc tgaagcctaa aggatgcgaa caagaggctc ggaaggtagc atttcaacgg
 1321 gagttttgag gatgctctcc tttagccacc cctctccatt ttctgcccc ttctttttaa
 1381 attctccatt ggctgtccct gctagtgtgc atttggggtg gtttgggttc agaatgggtc
 1441 tcattttcgc cgaggagtgg gtgatgtggg cggcctgtgt gtctctcca aggggtggg
 1501 ctgtccctcc tccaccacca ggcctagttt ggacctgtag tttcgcttag tgaaggaggc
 1561 cgggccgatc ctgggccgga gagagacgtm tcatgccwtg gcatgcagct ctgagtcac
 1621 aggcctgata aacagcccac ttcccagggc gagcaaggag gaacaaggcc cctggctgct
 1681 gtgggatccg tctgcgtcc tcttcgtgaa accgctgtt attcttttga caggagtgg
 1741 aacgcagcac cttcccttcc tcccagccct gcctccttct gcagagcaga gctcactaga
 1801 acttgtttcg ccttttactc tggggggaga gaagcagagg atgag

Two SNPs: Position	478	758	Estimated frequency
	A	C	47%
	G	C	12%
	G	G	41%

Three SNPs: Position	478	758	866	Estimated frequency
	A	C	C	47%
	G	C	C	12%
	G	G	C	17%
	G	G	G	24%

[000151]

VAN-STSI

1 gaatttgtct cagtarttaa tgactttaaa gctrcaaaag aattaagaaa gaaatagcta
 61 ttaccaggcc aggaagataa aaaccttatac agagacaata tatcagttgt gaagaatcct
 121 ggttctgttt taaagataaa agttagmctt amggcacatt gcttaaatkg ttttacagct
 181 caaccagccc atcaagtact caaccaccag gccaagtgga acctaagaaa ggatgatgcc
 241 agccttggtc gacccttgta actctaatac gctggacctt tgccccagtt ctatgctgaa
 301 ttcttcyctg ctcaagccct ttcattgagta tgaatgtacc ctaryttaa aacttcccca
 361 gttttgctgt ttgagagaca ttgttttggg aactatccct gatactcttc ttactgttaa
 421 gtaaaacaaa tccttcctac tcctgctctt tggcttgact gtgtcttttg gctcaaaacc
 481 caaygagagg tgaaccaggt tttgggtgga cattagggat cagggtggatc aactcaggca
 541 tmgtaggaaa agctactggt gttccaccag ttccagctta agctcatgga tggcattctc
 601 yagcaagaat ygcaattgtc cacctaaaga tgttctctca gcttctgtgg gccaagtagg
 661 ccagaaatgc cagattrgga gttcctgtgg tagctcagca gattaagaac cctactcagt
 721 ktctgygagg atgcagggtt gttcctgaag attargttct tgtctagttg cagaaggaat
 781 tcagaaatga gatggaaatt gagaaagaaa agtgagggtt tatttaagtg agaagtacac
 841 ctcttaag (**gagaggtgggc**) agagaggtggg cagagaggtg agcagctgYc ctgtgttttt
 901 tgggtgctact agttagaagg ggtgtccaat tgtatagatg ggatagtcaY tgagaaagtg
 961 ggggttaggg gtcataattcc ttaattttca ycccagctcc accttcccRa agggaggagg
 1021 gatttttgtc cttagttggt taattggaag tgtcatggca tccaYrtata atgggtactt
 1081 cttatctgca tagctaattg tattakaatt ttattataag gagggcataa tgagcaacag
 1141 kgttccattc agacactgga gattcgkgcc ctcttctacc tttctttgtc tgcagcctgg
 1201 gcacttatca ccccaaaaat gtgtgrtttc ctatcagtct ggtgkttccr gcttttcttt

The sequence (**gagaggtgggc**) in bold is an In/Del.

Positions to score:

Two SNPS:	position	849	889	Estimated frequency
	G	C	54%	
	A	C	15%	
	A	T	31%	

Three SNPs:	position	950	1009	1065	Estimated frequency
	C	G	C	60%	
	C	G	T	8%	
	T	G	T	13%	
	C	A	C	19%	

[000152]

IKBA

4321 gctactcccc gtaccagctc acctggggcc gcccaagcac tcggatacag cagcagctgg
 4381 gccagctgac cctagaaaac ctccagatgc ttccagagag cgaggatgag gagagctatg
 4441 acacggagtc agagttcaca gaggatgagg tgagtYccaa tgacctgtt cacgggtctg
 4501 caaaaagcaa tgctctcgga cccctagagc tcctcctttt cctgaggggtc tcaacataat
 4561 gaggatctca aattagggag cataagcagt gtcctaagag taggtttagg gggaggatta
 4621 tggtytgggg ttttcttttg cttttttgct ctttttgaag gagaggatcc ttaaaggaWa
 4681 acttcagccc aggaagttaa ttcagattcg ggttagaggg aacggagtcc aagaatactt
 4741 gcgttatctc cagtagcagc ccttgccatc accccagcac ctttggcaaa gttctggaag
 4801 tttaacatgc ctttctttcc ctttttagct gccctatgac gactgcgtgc ttggaggcca
 4861 gcgcctgacg ttatgagctt tggaaagtgt ctaaaagacc atgKacttgt acatttgtac
 4921 aaaatcaaga gttttatttt tctaaaaaaa aagaaaaaaa gaaaaaaaagg
 4981 tatacttata accacaccgc acactgcctg gcctgaaaca ttttgctctg gtggattagc
 5041 cccgattttg ttattcttgt gaactttgga aaggcgccaa ggaggatcat cggaatgcag
 5101 agagaacctc ttttaaaccg caccttggtg gggcctgggg gaaaggttat ccctaatttg
 5161 atgggactct tttatttatt gcgcttcttg gttgaaccac catggagtca gtgggtggagc

-65-

5221 ccaggtgtat ctgggaaatg ttagaatcag gtgwggttgtt aaacctgtca gtgggggtggg
 5281 gttaaaagtc acgacctgtc aagggtttgtg ttaccctgct gtaaatactg tacataatgt
 5341 atttttgttg taattatatt ggtacttcta agatgtatat ttattaaatg gattttttaca
 5401 aacagaattc tgatcactgt cttcttcggg cagctgtggg actcctacac tgagagtcac
 5461 tcgaacccca agtggaggtg gaggtggaga attgtgtggg agcatttacc acagccaacc
 5521 acggaactct ttcagagAAC agcttctcac accgtctaca ccagcctccc ggccaggctt
 5581 tgcaggcagc cccaggccca gtgcgtggga ggggaggctg ttgcaagggtg ataggaaaca
 5641 ccagttttcag gcttgggggtg gcagcaagtt ggttggccta cagctggaag gctcttcatt
 5701 gtcgcttgct ttcattcttc tggtttaaatt tcagccagga ccttacttct gctttaggaa
 5761 gctt

Two SNPs: Position 4476 4679 Estimated frequency
 C A 48%
 C T 26%
 T A 26%

Three SNPs: Position 4476 4679 4909 Estimated frequency
 C A T 10%
 C A G 38%
 C T G 26%
 T A G 26%

[000153] PBE64

1 tcaggctgtc accttttatg aaaattttat aaagttttga aaaaagaaga aagaaatcta
 61 tcatgggttg ttgaaagttt tatattcaga attaatgtta taatgtaaat ccaaRataca
 121 taacatttaa aatctaccca tatatagagg gatataagt gaagtaccat agctgtaaac
 181 acttgagtat agataattat ttttaacttaa tttctcccat wcttttttaa gacatgacag
 241 caagtacrar aaacaaacaa acaaaacaaa ycagagtatt gtgcaggat atcaatagcc
 301 ctcaaggaaa gaaacgattc cagcattact acaggatgaa gtctttgcaa caataaacac
 361 aaaaaattga ctgaatgaca aaacagaatt ggattttctg tgtctgacac agaatttcSa
 421 tcttcaaata gatgcctctg ggtatatatt tccaaatgtt gcccaacaat tttattcata
 481 aatatcacac ttgaaaatt cacctgctgt acctYaaaat gataatctaa taaaggaagg
 541 acagaaaaaa tactgcagga tgctcagata gacctcctag gacttaacta aatacatcta
 601 acaaattgaa tcagaattat cattacttga cagctttgta tttgattaca aataattacc
 661 aaaacaccca gtaagatctt gctttttcaa ttatgtaaca ttccrttaca cactaa

Single SNP: Position 515
 C 39%
 T 61%

Two SNPs: Position 419 515 Estimated Frequency
 G C 39%
 C T 27%
 G T 34%

Three SNPs: Position 115 419 515 Estimated Frequency
 G G C 39%
 G C T 27%
 A G T 6%
 G G T 27%

[000154] SCAMP

-66-

```

1  cattgagatg aactgaggag ctgttgataa tgaatgtata gatgaccact taccttctcc
61  cacttttttg tgcctgtagg tccatggact gtatcgcaaca acaggtgcta gttttgagaa
121 ggcccagcaa gagtttgcaa caggcgtgat gtccaacaaa actgtccaga cggcagctgc
181 aaaYgcagct tcaactgcag caactagtgc ggctcagaat gctttcaagg gtaaccagat
241 ttagagagtc ttcaaataat acactgttac cttttgactg tacttttttc tccagttact
301 gtattctata aatatttttt tgttcaaaac acacagtaca cacagcacgt atatttccta
361 atcacttgtg catgggctaa aaccagaaRa acttcgttgt cttattattt acctgacagt
421 ttcttaaatct ttcagtgcc cttgcaggaa aaaaaaatta catgctaaat aaatattctc
481 catatttttt gggggatgaa tgttcagcaa attttYtcgg tggtagacaca ctgaaatcga
541 catggcattt aggattaaaa atgcacttag tacttgctgc aRtcattctt tcaagagtct
601 tagacataag gattacacac tggagcagta aagcaatgct tcattccttt tctttatttg
661 tattgaaaga aataggacat cagaaactta gggactttta aattggcttg ctttttagca
721 gtttcagtca ccagtgaaga gcctatgtgc atttcatagt agataatgta aattttatct
781 ttttattttt tttttctaga gtaattgata ttttgatata aatctctgat cttgcatggg
841 caccatgttt ctaaaaaaaa ytagtatttt gggttatgca ctgcttctgg ttgtaggatt
901 ggggagtttg tagaatcata aaaatgattt tctgtaatWg tttcttttaa ataaaaattt
961 attggagtgc aatatgagga tataatatac agtgcattat ccaaagaaa aagtagataa
1021 ttgatg

```

Two SNPs to score:	Position	389	582	Expected frequency
		G	G	15%
		A	G	50%
		A	A	35%

Alternative two SNPs:	582	939	Expected frequency
	A	T	33%
	G	A	26%
	G	T	41%

Three SNPs to score:

Position:	389	582	939	Expected frequency
	A	A	T	33%
	A	G	A	26%
	A	G	T	26%
	G	G	T	15%

Alternative three SNPs:

Position:	516	582	939	Expected frequency
	T	G	A	26%
	T	G	T	22%
	T	A	T	33%
	C	G	T	19%

SNP at position 184 (c/t) can be included.

[000155] LEPR (Leptin Receptor STS7)

```

121 ttacctggaa atttcttcag cttgcctcac tactaaatat ttatttcctg taactgtctt
181 ttattgcata tgatttgttt tattggcttc aaagcatatc ctctctatt ctgtcgctct
241 tcctgttaaa tagattgaty taattctaac ccctttaagg aatgaaattt ctaaaaattt
301 atcatttccc aaagtgtgtt ttatagaaca ttgatttcat aaaattgttc ttaaaaaaga
361 ttacatgggt aaataaagtt taggaaaccc tacatcactg tatgtccaca gtgtagaatc

```

-67-

```

421 atcttVtata ctaaagggttt ggagaagccc tgaattaaag aaacatttgt gactttgttt
481 catcctatgt tcctcaaact tattttacca aagaaccttt tctcctctaa ccataattctt
541 tagggcgtat gtgttccttt gcatacattt tggaagaagc tgcttttata aatcagaatc
601 atacctatag ttgcaagcat atgtatgatg acttgctgtg tcatttttct gatggcagtc
661 tgcaaagact tacaaatagc agaaactctt aattatgtca ttagatcata atgacttcag
721 ctgaaatgaa tgtgacagtt tacttgctta tagaggagac tatcgagaga ttctctacag
781 caggccctgt ctaaccacag gttaaaattS ttaaaagtct ttgtggatag aggattagtg
841 gacarggatt agcaatgggg ttaagagaaa tgattgggaa gtgacacatt gcagtgagcc
901 agtccaaatc ttgtcatgaa atggaaataa caagatgact aaatggggga aaaatgtaat
961 tgtaatgtat acatgtaagg ataacctgac

```

Two SNPs: Position 426
 A
 G
 C

Three SNPs: Position 426 810
 A G
 A C
 C C
 G C

[000156]

COX2

```

1 aatctggctg cgggaacata atagagtgtg cgatgtgctt aaacaggagc acccggaatg
61 ggacgatgaa cggctgttcc agacgagcag gctgatactg ataggtgcgc aagaacaact
121 cttctcaata acgctcttct ccagggaata cgaaactgtt tctttgcagt ttccagaaat
181 rctgggggta tgtggtgcat gtaaaatcac atgcttcata gtaattcaac ccytgggctt
241 gattaggaat atcaccgacc ttttgttttg atggtaaaaa aggaagacac agaaatcaat
301 agaatatggc aaattaacaa aattgcattt ggggtgcttg aaagtgtgtg agtagaaaga
361 atttgtgYtc taaatygttt aatgttgtgc ccataggaga aacrattaag atttgtatcg
421 aagactaygt acaacacctg agtggctacc acttcaaact gaagtgtgac ccagagctgc
481 ttttcaacca gcaattccaa taccaaaacc gtattgctgc tgagttaaac acRctctacc
541 actggcatcc ccttctgcct gacgccttcc agattgatgg ccacgagtac aactatcaac
601 agtttctcta caataactct atcttactgg aacatggcat caccgaattt gttgaatcat
661 ttagcaggca aattgctggc agggtaagca ttattattat aaacgaaac aaagggctta
721 gtcagtaact ggaatttctg ctgtagaaat gatttttcgt aaacgtatta aaacagtaat
781 tatttgctag tagaattctt cccttaaaat gagaagtcta atatataatt tcggttatag
841 taaatgttat cactataatc tagatgacag aaatattctt gaacagttaa ggtctcagct
901 gggagctgag tcttaccttc tttgtacca agggatgcYt ttaaaataga aatcttaaat
961 atacctaaaa ctcatgttct acaatttcat ttcatttcca caggttgctg gtggtaggaa
1021 tcttccagct gcagtacaaa aagtatcaaa ggcctcaatc gaccagagca gagagatgag
1081 ataccagtct tttaaa

```

Two SNPs: Position 368 533 Estimated frequency
 T A 54%
 T G 17%
 C A 29%

Three SNPs: Position 368 533 939 Estimated frequency
 T A T 54%
 T G T 17%
 C A T 12%
 C A C 17%

[000157]

P450-ST18

```

1 caagagcgtg tcgctgctgg gaaggaaccc tgctctccac cgccaccctc tctctcagga
61 ccctgtgggc YRgggctcca cctcctcacc ctgagaaagg gaaccatgtc caaaatttgg
121 atggaccagt gctcccaRgt tttcatcagg tcctggacac agtcgtgaag ggcattgact
181 aaggtgtcct cctgccggaa atggaggaaa tcctttcaga tcaggacctg gagaaggcca
241 ggcagcggct gaggggtggg tccaggcaca tgtgaaggca agagccttga ccttgtctcc
301 aaaggtgagg caacagatga tgctgcaggg gaggacagag aattccttct ggatgggtcac
361 Rgggggtcccc gcctgggctc tcatgcgctg tggggaggag catgaactca gtagggggcc
421 tgccaggagg gggaagctgg tgcaggatgg ctgagggggg ccagcctcac ctacagaaac
481 tcctgggtca gctgctccac ccggggctcc atggagctgc ggacgccag cagcagggtc
541 gagcgggtga gtttcttgtg agcyttccag aacagrgagt artcccctag cgagatgtcg
601 gggcagtgtc gagacgccag cttgtctgtg agtaagggtt gggggcgggg gttcgaactg
661 accaaaggaa gtcacggacc tgacctccc cgcctcctgc agccctgccc ccttccttca
721 gaaagagccc cacccttac aggatggtat ctgggggtct gccgg

```

Three SNPS: Position	71	72	361	Estimated frequency
	T	G	G	21%
	C	A	G	23%
	C	G	A	16%
	C	G	G	40%

[000158]

AMG

```

1 tgtaagggaa gggtctgcc aagttctctt cttgctggta tttcctgctg gtgtgagggg
61 aacatatgtc tcttcccagc atggagttag tttcatcaag aaataaaaagg aaacaaaaaa
121 aatagaagaa caaagaaaag gagttctctg tggcgcgcca ggttaaggat ctggtgccgt
181 cactgcagcg gctatggcct ctgctgtggg gcaggtttga tccctggccc gggaacgtcc
241 acagactctg ggcacagctg aaagacagac agacagacag aaggaaaatg atgggtgggc
301 tagagtagga ttaactgagc acaggggtgg gaggggtgtc tgaggatgac cggaggataa
361 ctgcatgctg gtttctgctt ccgctgtaag gttacaacct ctgggaaaac catttcgttg
421 ctctggccct ctttaagata agagggctcc tctcctaccc agcatacatg ttccaactaa
481 aagtagacct tcaagatatt ctgcactata tagattttgt aaaagtagct tcggtctctc
541 ttaattgtga aattgcatat tgacttaate tcttcccctc tctctctccc cctccccctt
601 tccctcttcc cttgcacccc ctcaactctt ttcttctcct ttcccccttc ctataaaagc
661 taccacctca tcctgggcac cctgggtata tcaacttcag ctatgaggta atttttctct
721 ttactaattt tgaccattgt ttgacttaac aatgccctgg gctctgtaaa gaatagtggg
781 tggattcttc attcaggatg tttgtcagtc ccatttttct agttctcact gccagcttcc
841 tagtttaagc cctgatgggt cacctcaagc ctgcattgcc ccagaacctt cctacctgcc
901 ccccca (A) cccaacccccgactcagtStctcctccgtatacg gctgtaaaat gaacaccccc
961 tggagggggg acgRcatggt agggcagaaa ctgaactctg gctgaMcaga gttctatccc
1021 ggcctggaaa atatggggac tcaggtaaga tgttatcWac ctaaggctct tKccagccag
1081 accactcctg gttctaagac gtgcacactc tacgtgtctc cctYgctggg ctttcggaaa
1141 gatgagcgac caagggggct gtgtgacatt ctgccgagca agggaaagta tgagatggct
1201 ggaaatcagg tttgaggcgc ttctcatgcc cacacgaacc atgggacctt gggcaaatca
1261 ttgtctctct ctggaacttt gggttcttca tctggaaaag ggaaatgatt ataataccca
1321 acaatttaaa atattgattg gggagcgaaa gagttaagca acataaaagg tgctttgttc
1381 agtttgcctt gagcaaggct gtaattacgg tattgtctat aaatgcttat tactgtctga
1441 aggagtcctt ggacctgagg ttactcGagt ctatacgggt aaggaaaggaa ggaagtgtcg
1501 acttcttttc tcggttcaga tgacaaccca tgggtatggt gactcctaca agctgaggac
1521 aagggttaac aaaaatccga ggaaagattt tctgttaaat ctgaaaagggt tgacatatgt
1661 aaccggcaaa cgcgtttcta ggatgagaaa ctgggttggc ctcttaataa tttttgtgac
1681 atcagatcaa aagaggttac aattcctgtg aggtcacatt aattctctgt tttgttttct
1741 tcttgcaaa aagagcgggc gctggggagc gagacttact gcttttgtaa gctccgtcca

```

-69-

Single SNP: Position 906 Indel

A	37.5%
C	62.5%

Single SNP: Position 1467

G	50%
A	50%

Two SNP assay: Position 906 1467 Estimated frequency

C	A	20%
C	G	37%
A	A	42%
A	G	1%

Single SNP at position 975 could be substituted for the indel at 906

A	56%
G	44%

Single SNP: Position 926

C	73%
G	27%

Single SNP: Position 1006

A	36%
C	64%

Single SNP: Position 1058

A	63%
T	37%

Single SNP: Position 1072 and (1124) co-segregate

G(C)	59%
T(T)	41%

[000159] CTSL

```

1  ttgatgaagc tttcttcatt cacgagagaa agtcacaatt tgataacctc cagaaaccac
61  aggagcccat cagaagacta cccaaagtca gatgatctct agattgaagg aaagcaggcc
121 tgatccttac cagcaaccct gacccttgaa ccttgatggt aagaatcctc agaaatctcc
181 aggttatggt tgtttgtttg ttttggccat gccacagct tgtgaaaatt cctgggccag
241 ggtttgaacc cRgcatcaac agtgacaatg cRggatcctt aacctgttrc accacaaggg
301 aactctaagg acacagggtt ttgagggcat tcgtccactg tgtccctctt tgcctggcaa
361 agcaataaag ctcttctttt ctaccccact caaaactctg tcctcccaga ttcaattcgg
421 crtccatgca cagaggccga gttwcccat cagtattgga gggaattggt aagcggcttc
481 agggkctttt tttttttttt tttttttt

```

Single SNP: Position 252

A	37%
G	63%

Single SNP: Position 272

A	49%
G	51%

Two SNPs: Position 252 272 Estimated frequency

A	G	31%
---	---	-----

-70-

G	G	20%
G	A	50%
A	A	6%

[000160]

LCN

```

1  cccccacggg gtggggcaga gtctggggct gcagagtcgg ggtaggggat cagccggagc
61 ctgatgggag ggcctttctc cagctgYtgg ggagatggta tctgaaggcc atgacctcgg
121 acccggagat tcccgggaag aagcccagat cggtgacccc cctgattctc aaggccctgg
181 aggggggcga cctggaagcc cagataacct ttctgtgagt gtcgcctccc gccttcccct
241 ccccgaccca ggagggcggg ggtctctggg gtgtccttct cagccccttg tgtgacactt
301 agccctggac agctctgggg ggaaccgtcc tagaggggac agaccccgga tgagaccctg
361 tgggtgggag ggScagtgtc gggagaccca ggcaactgcc aYRtgccagc tgatgcctgg
421 cctggagggtg gctgacacgc catcgtccct ccccccctcc ccccgggcta ccacggaccc
481 caggctgcct gtggctgctg ggccaggggg accggagccg gggctgggcc gggctctcaa
541 ggtgggtgac cccagggcag catcacacgt ggcttctgtg ttccaggatt gacggtcagt
601 gccaggacgt gacactggtc ctaaagaaaa ccaaccagcc cttcacgttc acggcctgtg
661 agtctcgggg ccctggccgg gggcaggggt gggggcggcc agcgagtttc tgggacggtc
721 ttgcagcctg aggagcccta ctgctttctg accctattaa atgccaccct ctctcccac
781 tgggtccattt gccttRatga tatgaacccc Ygacgccagg cgtgacggat ttgccctcgg
841 ggggactggt ttgtcccag ctccagctgg gggtcatagc tgtgccaggt caggcccagg
901 agcagagcct ctttgagccc ccctcccccg ctgtcaccgc cggttggggg gtgaccacct
961 ctccagtgtg ggctctccc ggacgtgggg gcccacagc ctgggggggt ggctgggggtg
1021 gaagcccggg gcasgtgcc cgggagggtc ctgggcttac cgggaccggg
cccgtctccg 1081 ca

```

Position of SNP based on this sequence:

```

Position 87 estimated frequency
      C    54%
      T    46%

```

```

Position 373 estimated frequency
      G    66%
      C    34%

```

```

Position 402 estimated frequency
      C    77%
      T    23%

```

Observed three SNP combinations

```

C G C
T C T
C S Y
T C Y
T S C
T S Y
Y G C
Y S C
Y S Y

```

Estimated frequency of all possible combinations

```

TGT    6%
TGC    22%
TCT    13%
TCC    5%
CGT    3%
CGC    43%

```

CCT 2%
CCC 6%

[000161]**MYF5**

1 cgccttccaa aacgggggttt cttagaacac acagcttaga ggtgcagtca gtgggggtcg
61 tccgaccgat ctggggccat ccaaggcctg gcgcacgtgc agagcgggaa cgcgaggggc
121 cggcgtgttc gcggtcgccc ggcccgcctg tgcccaggcc cctcggagga gcagagggag
181 aagtcggagc ttcggagcga gcccggcctt gggggaaaga gcctctccgc tccccccaat
241 aacacagagc ctacaaaacg gcaggctgat gagaacagaa gagactttta aaaaaaaaaa
301 aaatctaaat ccaactccatt actgaaagtg ccatttcaca attttagtg ggcgttatta
361 ggcccactgg gtgaaaaaac aaatttcttc cacaactgaa gtgcatgaaa aagataaaca
421 tctaaaagta aagccatctc tcctgtcccg gattgagaag gcaggtcagc ttgtgtcagc
481 tgaaagtagg gagttctatt tactactttt tttttttttt ttttaatgtg ggagcaatgg
541 caaaacgttg gatttccttc gttccttttg cctgtgacac cctttaagcg tcccttgatt
601 tgctctaaac agcgcaagta agcgggggtg gggagccgtc accccgcccc agcagaaagg
661 cagtaaaacc attagcgtcg aaggggccgt aaacagcca ctgtctctaa aggaaaggcg
721 gaggtttgcc caaacagccc cggcgggggg tgcggtggga tatgctaata gtgccgggcc
781 actggggccg gcctccctcc cccaagaaca tataaagggc cccaaccca gcgcggccga
841 ccagggccgc caggcgtctg ccctgttaa ttagcagagc aaccgagcag ggagttccgc
901 ccgcgacgtg ccgcccgcg gaggcgccag gcccgggct tctccccgat ctgatctatc
961 tcgcagctgc ccaggtgcac cgcccgcctg tccgcagaag atggacctga tggacggctg
1021 ccagttctcg cttctgagt acttctacga tggctcctgc atcccatccccgagggcga
1081 gttcggggac gagtttgagc cagcagtggtc tgctttcggg gcgcacaaag cagacctgccc
1141 cggctcagac gaggaagagc acgtgcgagc acctacgggc caccaccagg ccggccactg
1201 cctcatgtgg gcctgcaaag cgtgcaagag gaaatccacc accatggatc ggcggaaggc
1261 ggccaccatg cgcgagcgga gacgcctgaa gaaggtcaac caggcgtttg agacgtctaa
1321 gaggtgcacc acgactaacc ccaaccagag gctgccccaa gtggagatcc tcaggaatgc
1381 catccgctac attgagagcc tgcaggagct gctgagggaa cagggtgaaa actaatacag
1441 cctgcccagg cagagctgct ctgagcccac cagccccacc tccagctgct ccgacggcat
1501 ggtaagagaa agctcgggac ctctagggcc cttctaattc tttccaaaaa actttacctc
1561 tcgtttaagc caggtgtagc aaccgaatat tctgatagtt ggctgtgggg gtgaggaggc
1621 agttgcccta agagagatgc cccatttaga cagacgccag gaaaccgctg ctgaagagca
1681 taatactttg cctcccaagt tctaggtgaa cgttgccggg ggaggttctc atgtgagacg
1741 ggtggctgtg aatgatcaga ggttttctcc attactcact ttacttccga tatatacccc
1801 ccgtggaccc caccgtacac taacgtttta agRcaactga cggaggctcc ccatagcgca
1861 tggtttctaa ctctaggcag aaatgggatg aaacaccgc atggccccgg ttgctgtctc
1921 gctgaggcct ggctggaaga tgttgatgca ttcttttcag aggggtgtct ctctaacgct
1981 gccaggtttt aatgtgtttt tgccctggga aagtgttctc tctccgaatt agtgtggctt
2041 ccttccaccc caatccatth tgcatggtta acccagtgca cgttgctgcc gaattccacc
2101 ccgcccctct ccttttctcc cagtacagt actgacctca gcgccttgc catttgggga
2161 ggcgagccct tcctaaatca aggcagtga ggtgactgag agtSgtcaac tttcgaagct
2221 ggagggcaag cactgcctca ccctccatca gagcacctt cgccaagacc tgaaaacaaa
2281 tgccttcttt tgtgtctttt attatagcct gaatgcaaca gccctgtctg gtccMgaaag
2341 aacagcagtt ttgacagtat ctactgtccg gatgtacaa atggttaagaa cgaacgcctt
2401 tagaggaggt taaagaccag ttcaacttca cagttcagcc catcaaata gtctgtcctt
2461 ccaggcagtt atctggagga aaagagaatg gtttttacag ggactttttg gcggagcaaa
2521 ataaacatct gctcaaaatt cccctcagag atactcccat gcacacacac aggcacatga
2581 gcttttgctt aaaacattac cgaggggtgt tttccactcc ccacctgcac ccccatgaat
2641 tgctagatat atatgttgca aatttctaca ctgggggtct tgtgaccacc tgacctctgg
2701 gtttcaaagg agctgacctg cagttcaaag ggcaacgtaa gcaagtctac ctattgggtt
2761 tttttttttt taacgttttt ttttcccctt gtatctttag tatatgccac ggataaaaagc
2821 tccttatcca gcctggattg cttatccagc atagtggatc ggatcagcaa ctccgagcaa
2881 cctggactgc ctctccagga cccagcctct ctctctccag ttgccagcac cgattctcag
2941 cctgcaactc caggggcctc tagttccaga cttatctatc acgtgctatg aactaaaaat

-72-

```

3001 ctagtctaga ccatttctgc caggagtgcc tattacacag gaggaaggag gcccaaaagg
3061 cccaaaagca agacaacctg tatataaaca ttttttttca gttgtaaatt tgtaaatact
3121 atcttgccac tttataagaa agtgtattta actaaaaagt cactattgca attaattctt
3181 tatttcttct tcttttcctt tgtcttgga ttaaataatat agttccaatg atattatttc
3241 ttataggggc aattcatcca agggtagctc gttgcaatgc ttaacttata ctttttataa
3301 tattgcttat caaaatatta cctctgttta gagctttatt tttttccctt ttaaaaatat
3361 tagaacaaat actagaactg gaaatcaagt tatagggagt tttaaataata tttaactttt
3421 ttgcttctct ttaatccttt gggttatattg tgtaagtaa aaatataaca tactgcctaa
3481 tggatatatat tttgatctta taagaaatgc atctttttta tgtaagcaca aaatagtact
3541 ttgtggatga tttcaagatg taagagattt tggaaattcc accataaata aaattgttta
3601 aatgaagaat catttgattt atgattttgt taaaagaacc tctaatagca ttggcagtga
3661 ttgatacgta tctttgagct

```

One SNP: Position 2204

C	48%
G	52%

Two SNP Position 1833 2204

G	G	42%
G	C	53%
C	G	5%

All single SNPs added to 2204 to make a 2 SNP combination result in roughly equal results.

Additional SNPs:

Position 2425

A	90%
G	10%

Position 2453

C	98%
G	2%

Position 2513

A	13%
G	87%

Position 2568

C	86%
T	14%

Position 2640

C	12%
T	88%

Position 2678

C	90%
T	10%

[000162]

PBE42

```

1 atgtctccat ttcttttgaa gacaKatgag ayctgagaaa aggtgatgct ccctgatgta
61 gtgggctcag cagtgtctga gcagaactct caggacaacc tgtctRagac rtccaacKtg
121 atagcaaagc tgagacaaaa tgccacctgc tgactcaccg caggagacag caagctggta
181 Ygatatcatt agagaaataa caacacacaa caatagcggg cacttgttgg gcacttcctc
241 tgaaggcagt rtgccaggag cgatcttcac atcttcattc aatccccagg acaacctgtt
301 atcagaggct tcacgatatc cgccttcac cataaaaataa agtatacaac atgctcagca

```

-73-

361 ctgatgggac agagagaagg taggggtccca catgtctagt cttaaattcc acatcatgtg
 421 acaccaggct acaagaaaaa tctgaaggca gaaggcagat ctggagatct agctggccag
 481 gtttcctgga a

There is a complex In/Del/SNP combination at position 7. This region should be avoided when designing primers.

Position	111	118	181	Estimated freq
	A	G	C	48%
	A	T	C	8%
	A	G	T	27%
	G	G	T	17%

Additional combination 2 SNPs

Position	118	181	Estimated freq
	G	C	48%
	T	C	8%
	G	T	44%

Additional combination 3 SNPs

Position	25	118	181	Estimated freq
	G	G	C	48%
	G	T	C	8%
	T	G	T	13%
	G	G	T	31%

[000163]

PBE57

1 cacactcacg cctcagaccc catcgctgtg cagtgcacac ggcattgatg rtggccgggc
 61 acycctgcac cagcYggccc ctcccatttc tccagggaca agawgtagRt ttgaggttcc
 121 tcgtctgact tcagcctcccc(c)trtttcctcactcrtgt cccgtccctt ctccgtccct
 181 gtttctggaa gycgccYtcc atgaccagga cactcagatt cttacttctg aggattatac
 241 tgaaaccctt gccactccct tcctgtgkcc tgttcacac tgcgtcacct gaccacccca
 301 gcaccctctc cctrcgcygtgc(agacacc)ctgttttgcr tgcctccagc agctggcagg
 361 aggaggtctc atcacagagc agccctggcc caccgagtcc ccacctrecc aagaacacgg
 421 ccctcactct gcaactggcc acgccgtgc tcaccacct tgcctygtgc caggaccga
 481 gagmgcactg gtccttgtga gtggygtcct ctgccttcaa gcatcagttc atccatctcc
 541 ccacacacac gag

The sequences in () are In/Dels.

Two SNPs: Position	75	197	Estimated frequency
	C	T	50%
	C	C	28%
	T	C	22%

Three SNPs: Position	75	109	197	Estimated frequency
	C	A	T	50%
	C	A	C	12.5%
	C	G	C	15.5%
	T	A	C	22%

-74-

Additional SNP at position 268

G 80%
T 20%

[000164]**PBE73**

```

1 tttctactac yngscatcat ttcccttccg actgtaggat ttaactttaa catttccttk
61 gtacagggtct gctggtgatg aacgctttct gcMtattttat gcctgaggaa tcttttRtttt
121 gtcttcatta ttgaaagata ttcctgctct gtatagaatt ctacactaat agacttYgct
181 ttcagtaggt acttcggaag atgctgctcc actgtcttct ggcttctttg tttcctatga
241 gaattctgcc atctttatct ttgcttctct gkatggaatg tgtcattttt ttcctcagac
301 tgctaataat atattctctt tatcactggg ttccagaaat ttcattctga tgtgtctkgg
361 tgtaattttc tttacatatt gtgtgcttgg ggatganctg tgattcttag atctatttat
421 agtttttata aaacttttga aaattcagca aattcagcaa ttaatcttct ttttttYtcc
481 tgttggttct tctccccatt ctccttcaga accgcaatta tatgtacatt taggscacs
541 ctgaagstat ccta

```

Positions of possible SNPs based on this consensus:

position 93, frequency

A 80%
C 20%

position 116

A 89%
G 11%

position 177

C 87%
T 13%

Three SNP estimated frequencies

	93	116	177
A	A	C	61%
A	A	T	5%
A	G	C	10%
A	G	T	1%
C	A	C	22%

Other potential SNP positions:

Position 375 Indel

- 86%
+ 14%

position 477

T 96%
C 4%

[000165]**PBE84**

```

1 caaaacttct aacRtgcaga aaaagcggKa gartttacag tgataaacac ttgcattaaa
61 gaaaactaaa gatctgaaat aaacaaccta actttaYacc tcaaataact agaaaaagag
121 gaataaacta agcccaacat tactggaagg aaggaaataa taaagattat ttaagagagM
181 agaaataaat gaaattgaga ttagaaaaca atagaaaata tcagcaaaat taagagctag
241 tttttcaaaa gattaaaaaa aaatgaatct ttagagtaag aaaaagagaa aagatgagta
301 aattaaacta taactgaaag aggagatgtt accaactgac actgcagaaa tacaaaggat
361 cccaagagac tactatgaac aattatgggc caaaaaaatt gggtagcttg gaagaaatgg
421 ataaattYct agaaacatag atcccaccaa gaggc aaata tgacaaaara gaaaatctga

```

-75-

481 acagaccaat aatgagtaag gagattcaaY caattaaaaa aaaaaactcg ctaatgtaaa
 541 gcccaaaaact aaatggtttc actgctgaat tctaccaa Wtttaaagag ttttaa

Two SNPs: Position 29 97 Estimated frequency

G	T	50%
G	C	28%
T	C	22%

Three SNPs: Position 29 97 428 Estimated frequency

G	T	T	10%
G	T	C	40%
G	C	C	28%
T	C	C	22%

Alternate SNP: Position 14

A	48%
G	52%

[000166]**PBE 112**

1 ctgctccctc cacaccgtct ctttctttgc gggtcggtgg gttcctctta gggacccaga
 61 Yaacttgctc ctccctacct ctcatcYctg aaaccggaat ttctgtccct ggcctggcag
 121 ccgtccaccc ccagccctcc Yagctaaggc gtttaacatg cctgcccgca gtatggtggt
 181 ctagaaactg ccagggagcc aggggggttc ctgcaaattc tgtcccaggc tgtggcccca
 241 gctgggtggg ggaaagggac ctgtaagctg caagataagg ggacctcagg tgtggggagt
 301 tgttgatgtt ttcacctaca agtgtacctg tttgctggct tgtgtgtatt cccacctgtc
 361 tgcctgggac tctgaaatag tctggcctgg ggtttctaac cctaccccc aaggtgaggc
 421 ctccagggtg ggggtgggaag tctaatacagg tttccccatt cctgaaaatg gggctgtagt
 481 gagttgcagc catttgctca aggccttctc aggagtggga ggcagcgga atcagaggaa
 541 acagtgaatc aaagagactc caaagagggc ccatctacac tgacagaaga actagcagcc
 601 cgagagtaaa tccaaccctt gttaatcctg tttgatgtct cttggaRtga tgcaagtcac
 661 tgagacactc act

Position of SNP based on this sequence:

Position 61 Estimated frequency

C	34%
T	66%

Alternate SNPs

Position 87

C	94%
T	6%

Position 141

C	90%
T	10%

Position 647

-76-

A 94%
G 6%

87 and 141 cosegregate, either could be used.

Two SNP assay

61 87 (or 141)
C C 34%
T C 60%
T T 6%

[000167]

PBE132

```

1 gcggtcgggg ttgggtagca catcctggcc tcttcctggt ggccccacag gctgagggag
61 atgtcacttc ctgctgttgg gaggtgaccg gcaggatgga Sgggagcagg aaggggtggc
121 cgtgaRccag cactaaccag cctgggccat ttcctcacc cggaaagagg gaaggagaga
181 gagagggaag acYgggtgtt gggagtctcc tgtggccctc acaccagag rcagtgggtg
241 ctcagacccc tgctcccacg ggcactgac tggcacaggc aggcgttcag aaaaaaaaaa
301 ctttcaaaac aatcctatgc actggcccct gcacgtcaca gtttccaatg gccttctcan
361 ccccggcctR agggctttcc tgccatccct ctgactgtag gacagagttg ggggattatt
421 gctgccaccg gaaaatgagg aanctgaggc ccagtgaggt tgcaatganc agtctgcctt
481 ggatygtaca gcaaattant accaggactt g

```

Single SNP:

Position 102

C 53%
G 47%

Alternate SNPs:

Position 127

A 63%
G 37%

Position 194

T 63%
A 37%

Position 371

A 58%
G 42%

4 SNP observed combinations

CATA 30
GGCG 18
GRYG 4
GRYR 3
SATR 6
SRYA 5
SRYR 19

All individuals with A at position 127 have T at position 194.

3 SNP estimated frequencies

CAA 41%

-77-

CAG 4%
 CGA 4%
 CGG 3%
 GAA 1%
 GAG 7%
 GGA 5%
 GGG 27%

[000168]**PBE137**

1 tgggttatng gctgttttga ttcyracgca aaatttgagg gaaacacaaa gatgaagtga
 61 agctcttaat ttactttttc actttctata cttcctttaa aaatgacttc aataccctag
 121 Yattttttcca tatgaacggt tctatttttc atgattatta aagtgggtaca aaattggaac
 181 aactgcaaaa tccattggta agtgattgat tgtgatatag taatgcaatc tttatgaaga
 241 ttaaaatgct ttattacat aaggagatgt tcataatYta ttttgaagtg aaaaaagcag
 301 gtggcacaaat agaatatatt tgatgtgaaa ataaccatta acagtgggtg catgagtga
 361 aggaaggttg tagagggact ttccaaggga cctaactctc tatgcttgRa aaatctgtag
 421 acaccttaag tgcaygtaac taacacctgg aggttgatct ttacactgrg gsnssaaara
 481 aaaa

SNPs to score 121, 278, 409, estimated frequencies:

CCA 22%
 CCG 17%
 CTA 5%
 CTG 5%
 TCA 5%
 TCG 4%
 TTA 5%
 TTG 36%

Alternate SNP:

position 435	Estimated frequency
C	77%
T	23%

[000169]**PRKAG_STS3**

1501 ttaacaaaat cgatactatg tgtgtctgta cacttatgac tttggagtag aaacactggg
 1561 ttgggtttccc acaccttgga gtgcttgggg aggggtcacc tcagyacctc tggccaccag
 1621 cagccttaga tctggaacaa atgtgcagac aaggatctcg tggagggcat gccaggacgt
 1681 gggagaggca gacagcagge tcatgtagag gcaggcccgg gaggcgcccg gtggaagaac
 1741 cctggctggc aggggacctc tgaggcgagc ggaacgattc accctcaact gttctctccg
 1801 gcgctcagat caagaaggcc ttctttgctt tgggtggccaa cggcRtccga gcggcacctt
 1861 tgtgggacag caagaagcag agcttcgtgg gtgaggaggg gctggggagg cagaggtggy
 1921 ggggaaggga ataggggRac cttgtggggg gattctaggg ccgagctctg acayaccaca
 1981 ggcttsaacc aagcaggggc ctggcctgg [rgrrscgga]gggagcatytga ccccggtctc
 2041 cygggtggccR gctgggagat ctcaactgta ggagagctgt gaccagctga cccctccagc
 2101 tctactaccc caaggtccct gtcgcaggtg ctaagtaaga agaggacagg cggaggaagg
 2161 aagtcagaaa atagaagaag cagggcagga aggagagaaa tgacagggga agcataagag
 2221 ggacaacccc atttstcagg cacgggaggg gctgccctcc tgtcctcttt tggccaccct
 2281 cagtaaaagg atgtgggcag ggtgggggga ggggcccggg ctgaccccca ttgctcccc
 2341 cyyttgcccc ccsyacaggg atgctgacca tcacagactt catcttggtg ctgcaccgct
 2401 attacaggtc ccccctggtg aggagtggtc tgggggtcct ggaacacca tctgggctgg
 2461 ggtggaagga gttcagggga ccctcgctg actttgggag ttccgttgct gtctttaggt

-78-

```

2521 ccagatctac gagattgaag aacataagat tgagacctgg aggggtgagc aggcgagggg
2581 acsgkcgaaag gggctgaggg tgtgtgggtg aggatgrggc caaggacctc agggagagca
2641 tgcgcagtgg aggtttcctg gaggaagcgg gaggagggtg atcgggagcc caggggatct
2701 aagggagggg gacagtctgg ggggtggccac gtgaggcggg gtggtcggcc cctttgtgct
2761 gattctggct tttcctgcag agatctacct tcaaggctgc ttcaagcctc tgggtctccat
2821 ctctcccaat gacaggtgag cttccccagc crccactcg agcctccttg ccccgcacag
2881 accccttctc cagctcatcg gttctaagct catggactca tcgtccgtgg actgcagatg
2941 cggcagcttt gacaccctgt cctcctctcc aggggggctg ggatgaaggg gctctctttc

```

Three SNPs to score:

The sequence within [] is a complex in/del. The SNPs at position 1920, 1974, 1886 and the indel are carried on the same haplotype and can be used interchangeably.

Position	1845	1938	2050	Estimated frequency
	G	G	A	27%
	G	A	A	10.5%
	G	A	G	43%
	A	A	G	19.5%

Position 1845 splits the AG haplotype.

[000170]

WSCR STS1

```

1 acggacattc ctgacctcca ctctttggcc tgagggtgtg gaccaagggg cagYRggcca
61 ccMgggtgaa cYacaacagc cccagaYctc ccctgcgaag ggagtccagg tcctgggggc
121 ctaagggaacccc (C) tgccttgagcagcca atcaggccac gtgcacacgg ctaacctggg
181 gctcctccct tcccctctgcccgg (GATTC) gattctggaga Rcacctgcaa gcagcgggtc
241 cccccaggag acagacgggg gcggaagaag cctgcacagc aggacgtgtc tgggggtgaat
301 YRgggactcc aagactg (GGAAGAAGGCTG) ttctcYgtgcc tagaacacat ctaggagcac
361 tggggggtca gaaaaggcaY gccctggScR tgRaactcctcaRcactt (ACT) gctcaaattg
421 ttacactctg gtacaaaggg aatcagaaRc tcaagaactc ccagggtga tacataRgca
481 atgatcttgY ggcaaaactgt cYcctcctat gcggtttgac cttgtcctcc tgggcgcccc
541 ctgtcagccc tttggcRgtR ggggttgggca tgagScctgc cacgggtggc caagaagcYg
601 tgggaaagggt gccctgcctt ggagtctggg tttgctcatc aatgaaatga gccaacatgg
661 accttgaggg tgtcgaggtg aggttgcYsa ggtgaggtgg ggcataRtag gacctRcacc
721 ggggtgtggg cccaRagcca Ytcacgaggc ccttcccggg ccctgctaga gacacRccca
781 cactacctgg ggtctcggcc tctagatttg gagcaagacc aaccctggtg gcagYggcag
841 ccgctccagg gtcacgcccc tccagtcttc cagagcccgc tctggggacc cccacccccg
901 ggaaagtggc tgccaaacca ttgctcagtc tcctcaggac tggccgatgg gggctgcgct
961 cgctcaggtc ccctcgggtg gcgcgtgtgt tggggatggg cggggtttcc ctccacgtgt
1021 ttgtgagcca cacggactag cgccgggatg cctggcacag gtgccttggg gaggccc

```

The sequences in () are In/Dels.

SNPs:

Two SNPs:	302	599	Estimated frequency
	A	C	47%
	G	C	18%
	G	T	35%

-79-

T at position 301 is low in frequency (~3%)

Three SNPs:	135	392	599	
	T	G	C	13%
	C	A	C	40%
	C	G	T	34%

Frequency estimates do not include individuals heterozygous at position 135

Other Possible SNPs to score:

Position 222

A	36%
G	64%

Position 408 Indel

A	low
G	high

Position 560

A	87%
G	13%

Two SNPs:	55	63	Estimated frequency
	G	C	44%
	G	A	34%
	A	A	22%

Three SNPs:	55	63	135 (In/Del)	Estimated frequency
	G	C	C	44%
	G	A	T	18%
	G	A	C	16%
	A	A	C	22%

	55	63	302	Estimated frequency
	G	C	G	44%
	G	A	G	18%
	G	A	A	16%
	A	A	A	22%

[000171]

BG_mod

```

1  cccagccct ttttccaggt cagcscaggg aaaaaacatg ttctctgtcc ctggttatac
61 tgttttagaaa catcacctcc ctcggcgwwa ctaaaacttg ggggttgcaa tttattcctt
121 gcttctttgt atttcktacc acattgagag agctctaggt tttcatccgc agattcccaa
181 accttcgcag aggagctggt tcacaggacc gtgattcaag tttactctac ttttccatca
241 tttatttggt catatgttta aatgaagaaa gaaaggaatg aagataacctg aatgaaatga
301 gtatttgttt tcttaccagc aggactgaat acaaatgaag agaagaaaaa tacgcacatt
361 taggacttgg gcagaggttt tatccacgct ctccttgttg ttatttccca tattcagaag
421 gcr crggwgw ggattygtct gtatggtcct aaattgaacc acagtgggtca aatccctcca
481 ctttctgctc cttggattct tcgtttgtgt actaagaaaa tggggaggga gtctctaaga
541 gattgctaca gtgggactca actctaaaag ttgtacagac ttgctaagga ggatgaaatt

```

-80-

```

601 agtagcactt tgcactgtga ggatggkacc tagagctccc cagagaaggg ctgaaggtct
661 gaagttggtg ccaggaacgt ctcgaagaca ggtatactgt caacattcaa gcctcaccct
721 gtggaaccac gccctggcct gggccaatct gctcccagaa gcagggaggg caggaggctg
781 ggggggcata aaaggaagag cagagccagc rgccacctac atttgcttct gacacaaccg
841 tgttcactag caactgmaca aacagacaac atggtgcatc tgtctgctga ggagaaggag
901 gccgtcctcg gcctgtgggg caaagtgaat gtggacgaag ttggtggtga ggccctgggc
961 aggttggtat ccagggcttc aggagaggga gygggaggtg ggcaggtggg gacagagcca
1021 cccctgcctt tctgacaggt gctgactccc tcgggcc[ttgcgctcttttcacccctcagg]
1081 ctgctggttg tctacccttg gactcagagg ttcttcgagt cctttgggga cctgtccaat
1141 gccgatgccg tcatgggcaa tcccaagggt aagggccacg gcaagaaggt gctccagtcc
1201 ttcagtgcag gcctgaaaca tctcgacaac ctcaagggca cctttgctaa gctgagYgag
1261 ctgactgtg accagctgca cgtggatcct gagaacttca gggtagagtct gggggaccct
1321 caYgttctcc ttctgtcctt tggatcatggc tgagctcgtg tcatgaagag agggcygaaa
1381 ggcaggatgc cgtttagaat gaaaaggaag cattctggtt acatMgctgg ggactctgca
1441 ggaccactga rtttcttttc cctcttctta ttcacagcca tcatttcttc ttactctctc
1501 ttgttctcck ctgttttggt tkgtttkggt tkgktttggt ttgttttttt ttttttcttc
1561 tgcgatgtct tctctctttt aagttaaact ttttgagtgt ttgtttaaaa aaaaaaaaaac
1621 tttcttcttt taagccactt taaaaaatgt tatctaacag ttgccctta ttctttcttt
1681 ttgaggcaag aaggataaaa tgtttattgc ttcttgcat ggttctaaag cataaaaagg
1741 ataacataaa ttcaagacta aggtagaaag agagaaacat tyctagctgt cattcagggt
1801 gccatgggtg gattcacatc aata[gtaacgtctacacggcagctcact]tt ctgtttatct
1861 tctaggggct cagcttgga tgagactgaa atactcctgg gtctaagctg gtcctctact
1921 aatcaggccc ttctttttta tctctctccc acagctcctg ggcaaYgtga tagtggttgt
1981 tctggctcgc cgccttggcc atgacttcaa cccgatgtg caggctgctt ttcagaaggt
2041 ggtggctggt gtggccaatg ccctggcca caagtaccac taagttcccc ttcctgattt
2101 ccaggaggag ccctttttcc tctgaacca aaaactggat atggaaatat tatgaagagt
2161 tttgagcatc tggcctctgy gtaataaaaa cattatttt cattgcactg gttttttaa
2221 attatttcag tgtctctcac tcagatgggc atatgggaat gaaaggtatt ataaaaaat
2281 rtaataaat gaggggctaa ttgagacttt aagaaagctc tctgtacctt ggaccccatg
2341 aaagragtgg ttgtaaagca cgcagtgtgag tggaataaca ccttatacct cttcctttcc
2401 t

```

The reduced complexity sequence between [] should not be used for the primer design.

Two SNPs:	Position	1323	1425	Estimated frequency
	C	A		30%
	C	C		46%
	T	C		24%

Three SNPs:	Position	1257	1323	1425	Estimated frequency
	C	C	A		30%
	C	C	C		35%
	C	T	C		24%
	T	C	C		11%

An alternative is to score the following positions:

	1323	1425	1966	Estimated frequency
	C	C	T	11%
	C	C	C	35%
	C	A	C	30%
	T	C	C	24%

[000172]

GLUL

```

1 ggaatttcga tggctctagt acttttcagt ctggaggctc caacagtgac atgtatcttg
61 accctgctgc catgtttcgg gaccctttcc gcaaggaccc caacaagctg gtgttctgtg
121 aggtcttcaa gtacaaccga aagcctgcag gtgggtagag gataggtctg ggtgccccca
181 cccaatcca tgaattggaa aggcggggaa gaggctattc caaaattggg actgagaaga
241 ccagcctgtc ttagaactgt atgtgaagtc tgtattcgct tgggcctcat ctctccatct
301 agaacatcct tgttcaaaga ggttgcggtg cccatcctcg cattattggg ggcataaaac
361 aacattcaat ttctctctct ctctctctct ctcttctttt ttttttttgc tttttagggc
421 tgcacccgtg gcatatggag gttctcaggc taggggtcca atcagagcta cagctgccag
481 ccacagccat agcaacgcag gatccaagcc ccgtctgtga cctacatctg tgacctacac
541 catagctcac agcagtgccg gatccttaac ccattgagca aggccaggta ttgaaccac
601 aacctcatgg ttcttagtcg gattcatttc cgctacacca cggcggaac tccaacatta
661 gattcttcta aatctgcttg ccgttagtcc tctgtggcac agccctcctg cctgcctgta
721 gattttaaga acacagattc taaagttaga atgttgatt tgaattctag ttgtgccact
781 tccttgctRt gaacttgggc aagttacttt tcacctctac tcctcacttc cctgatttta
841 aaatggggaa attaataaaa tccatcttag aaagtgttt Rtggggcgaa ctgagttgat
901 gagtgcggat aattgagaag gctcgtacac cagcactRg ttgccttttc attccacaca
961 gcgcgtttcc tttctctgcc cacgactaa aatgaggaca tccggcttca ccaataacat
1021 acgtttgcct tgagggttgc tggagccccg cccctttgSt gtttgtctag cacaccatt
1081 gctcccagtt ctacttcggt cgtcttttta aatttgccat gctctgccag tgagactttt
1141 tggcaagtca cttagtact ctggacctta ttacttgtct ttgtttggag aggcattatc
1201 atattgtggt aaataacttg agctcggaag ttgagacaag cagagtttcc tggctctgct
1261 gcttttgctg tgaccYgggc aggctacctg ccatctctga gcctcatttt ctcat ttgta
1321 aaaacggaga ctcgtaatcc ccacctcagg gttgttagga agactcaaga ttataggttt
1381 tctttgataa aatatgaagg tcataaaatg gttcctggga catagtggaa tagttatagt
1441 aatgagtcca taaagtaatt ctagcttaga aatagtttca gttgagttca ttgcatttta
1501 ctgggggttt gagggggttt gtttgctttt gcaccagcca atactataga actttgtagt
1561 ggtttacctt ctcaccttga atgtaaggct cttcttctct cagtcaaaat tgtagatttR
1621 gctctaaatg cctgagaaac taggaataga tggaaagcga gaRgcagttc tctagtgaat
1681 agagagttta agtgcttgac atttggcggt tgggaggggg ccagcttgca gataaagtga
1741 aaacgtgctc tattctggat ccactgccgt gtgactttgt tgtaaccag tttagttata
1801 aataaagcct tttgagagtc tccctacaga tacataatta tagcttctct cattgtccRc
1861 ttttacacca gcggggactg aaagcttaca gtgcagcYRa cactcatcat ctcaggtgtg
1921 Yggattcctg gaaaccacct caggcgcatg tgcattcggt acatttgccc atgcgtctct
1981 aataagtgtg ttggctcttt ctttttctta gagaccaact taaggcacac ctgtaaacgg
2041 ataatggaca tgg

```

1 SNP

789

G: 90 %
A: 10 %

939

G: 18.3 %
A: 81.7 %

1069

G: 81.6 %
C: 18.4 %

1276

C: 80 %

-82-

T: 20 %

1620

G: 93.1 %

A: 6.9 %

1663

G: 7.8 %

A: 92.2 %

1859

G: 88.9 %

A: 11.1 %

[000173] IL2R

1861 ctccctgctcc gactttttaa tgaattgcct tccggacaca aagaagtggc gggctgccga
 1921 gctgacgggg ctcatcattc atggctgcac tccttgccat gtgtgatgcc rtcaggcggm
 1981 cggccrgaat ggagtccttc tgggggagac tmatacacia attccatcag aaaatctcag
 2041 ttgagaaata catcRgaggc catggtcRtc tggacaggaa gcccgactga caggctggcr
 2101 ctgggctgac catgaacagr cacaaccaag agcccccca agtgcttctc tttatgacct
 2161 cagtacccca atattggctg gcaggctcgt ggctctaaaa acaccttctt tttgaatgaa
 2221 tacacgaagg gcattttctc ttgtacacat ttgtgaagtt attttcttca tttcacccct
 2281 actcccyagc ctccccgagg atccccgtaa aacaagttac tcctagacct gaagaacaga
 2341 aggaacgaaa aaccacagaa actcagggcc agatgcagcc cccgaaccaa gctaattctc
 2401 caggtaagag gcaaaccatc cccctgcagg tagaggagat gctgccctca ggaggcgtct
 2461 gtcctctttc ctaaggatgc crtgtggaca agaagcgggt ttgcargctg gtcactcagc
 2521 acatcaagcc actctgacag cagccaacct gggttcaggg cttcctctgg gtttgggaac
 2581 cagccaaacr gaattcaact ccattcaatc caactcaact cactkcactt caccaagcac
 2641 atctggggac tacctgcggc caggagatgc aggccctgacc agacagtggg ccccgctctc

2 SNPs: Positions	2055	2069	Estimated Heterozygosity
	A	G	37%
	G	A	39.1%
	G	G	23.9%

[000174] MCP

1 ctccctcmac ctacmawhwt tcctaagggt tgcagaggag ctgccataga gctcaaaaca
 61 yggctctacag acaagcatct tctccatccc tyctcatctt ckacRggcc gYttgacaac
 121 atctctwgng gkgkkgdgg rgcnngncsy vncvmsckg kdwrwscmc yygcgtyyac
 181 scmaagmgct ctgactctgg agttctagtc ctgcggggac cttaggaagt tcacgggtcaa
 241 tactcygccc ttgggctcag wactaagak gatctccggg taaagagata gacagtagct
 301 ccatgcctga tttaggaaaa ctgtccgtac agacagttgt aattcatttc ctttcagaga
 361 caaatcctgc tctcttccta gttcctgaag tcattaaaaat caaaagctct cagaaacgtc
 421 ccagcatttg ctaagtccac gctggggggag gatggggcaga gccgtgttca gcarcgtttg
 481 acagcaacac ccacttatctt tcattcagta tccataggca tatatcatgc acctggtata
 541 ggyctctctc tcagcactgg agatacagca aagaaaacgc tattcctgcy ccatggagct
 601 tgnttcagaa aaataractc aaaaccttctc aaatggagct gccgctgggc caagtcaaaa
 661 taagtaaaag aaatccgtga gaaacccttc agttatatta agaagaaata gcttgatg

One SNP: Position 102, Estimated frequency

G	16%
A	84%

One SNP: Position 106, Estimated frequency
 G 47%
 A 53%

2 SNPs: 102 106 Estimated frequency
 G G 15%
 T A 63%
 T G 22%

[000175] mtDNA (15001-16585)

```

15001 ggatacatct cagtagccat agcagtagta taacccaaaa ccaccaacat accccccaaa
15061 taaatcaaaa acgccattaa acctaaaaaa gaccacacaa aattcaatac aataaccacaa
15121 ccaactccac cacttacaat caaccgaagt ccaccataaa taggagaggg cttagaagaa
15181 aaaccaacaa acccaataac aaaaatagta cttaaaataa atgcaatata cattgtcatt
15241 attctcacat ggaatttaac cacgaccaat gacatgaaaa atcatcggtg tacttcaact
15301 acaagaacct taatgaccaa catccgaaaa tcacacccac taataaaaat tatcaacaac
15361 gcattcattg acctcccagc cccctcaaac atctcatcat gatgaaactt cggttccctc
15421 ttaggcatct gcctaattctt gcaaataccta acaggcctgt tcttagcaat acattacaca
15481 tcagacacaa caacagcttt ctcatcagtt acacacattt gtcgagacgt aaattacgga
15541 tgagttattc gctatctaca tgcaaacgga gcatccatat tctttatttg cctattcatc
15601 cacgtaggcc gaggtctata ctacggatcc tatatatcc tagaaacatg aaacattgga
15661 gtagtcctac tatttacgtt tatagcaaca gccttcatag gctacgtcct gccctgagga
15721 caaatatcat tctgaggagc tacggtcatc acaaatctac tatcagctat cccttatatc
15781 ggaacagacc tcgtagaatg aatctgaggg ggcttttccg tcgacaaagc aaccctcaca
15841 cgattcttcg ccttccactt tatcctgcca ttcattatta ccgccctcgc agccgtacat
15901 ctctatttcc tgcacgaaac cggatccaac aaccctaccg gaatctcatc agacatagac
15961 aaaattccat ttcaccata ctacactatt aaagacattc taggagcctt atttataata
16021 ctaatcctac taatccttgt actattctca ccagacctac taggagaccc agacaactac
16081 accccagcaa acccactaaa caccacaccc catattaaac cagaatgata tttcttattc
16141 gcctacgcta ttctacgttc aattcctaata aaactagggtg gagtggtggc cctagtagcc
16201 tccatcctaa tcctaatttt aatgcccata ctacacacat ccaacaacg aggcataata
16261 tttcgaccac taagtcaatg cctattctga atactagtag cagacctcat tacactaaca
16321 tgaattggag gacaaccgtt agaacaccgt ttcattcatc tcggccaact agcctccatc
16381 ttatacttcc taatcattct agtattgata ccaatcacta gcatcatcga aaacaaccta
16441 ttaaaatgaa gagtcttcgt agtatataaa ataccctggt cttgtaaacc agaaaaggag
16501 ggccaccctt cccaagact caaggaagga gactaactcc gccatcagca cccaagctg
16561 aaattctaac taaattattc cctgc

```

[000176] The swine mtDNA sequence is publicly available. For example complete mtDNA sequence of *Sus scrofa* breed Duroc can be obtained from GenBank with accession number AY337045; and breed Landrace (AF486866).

I. SNP detection methodology

[000177] Detection of SNPs in a high throughput mode or in a smaller scale can be performed using standard SNP detection techniques that include specific PCR amplification followed by sequencing, mass spectrometric analysis, and HPLC based

analysis. Some high throughput SNP detection techniques and devices are described in U.S. Pat. Nos. 6,720,143, 6,537,748, 6,337,188, and 6,225,109, which are herein incorporated by reference.

DOCUMENTS

Garber R.A. & Morris J.W. (1983) In: Inclusion Probabilities in Parentage Testing, pp. 277-80.

Hawken et al., (2004), An interactive bovine in silico database (IBISS). Mammalian Genome 15, 819-827

Jamieson A. (1965) Heredity 20, 419-41.

Jamieson A. (1994) Animal Genetics 25, Supplement 1, 37-44.

Weir B.S. (1996) Genetic Data Analysis II. Sinauer Associates Inc., Sunderland, MA, USA.

WE CLAIM:

- [C1]** A method for identification of an animal, the method comprising:
- (a) obtaining a sample from the animal or from a processed product of the animal;
 - (b) performing single nucleotide polymorphism (SNP) analysis comprising analysis of a plurality of markers, wherein a marker comprises at least one SNP;
 - (c) generating SNPTracks of the animal from (b); and
 - (d) comparing the SNPTracks of the animal to a database comprising a plurality of SNPTracks to identify the animal.
- [C2]** The method of claim 1, wherein the animal is a farm animal selected from the group consisting of cow, pig, sheep, and poultry.
- [C3]** The method of claim 1, wherein the animal is a mammal.
- [C4]** The method of claim 3, wherein the mammal is a human.
- [C5]** The method of claim 1, wherein the plurality of markers are selected from the group consisting of autosomes, sex chromosomes, and mitochondrial DNA.
- [C6]** The method of claim 1, wherein a plurality of SNPs in a marker are present within a nucleotide region of 0.2 to about 10 kb.
- [C7]** The method of claim 6, wherein the plurality of SNPs in a marker are present within a nucleotide region of 0.2 to about 1 kb.
- [C8]** The method of claim 1, wherein the plurality of markers are selected from the group consisting of swine markers designated ACY-STS7; COX2; EG-STS7; GALT; IKBA; LEPR; P450-STS18; PBE3; PBE42; PBE43; PBE 57; PBE59; PBE 64; PBE 73; PBE 84; PBE132; PBE137; PRKAG-STS3; RYRA-STS6; SCAMP; VAN-STS1; WSCR-STS1; BG; AMG; CTSL; PBE112; and MYF5.
- [C9]** The method of claim 8, wherein the plurality of markers comprise SNPs at ACY-STS7 (245 G/C 421 C/T); COX2 (368 C/T 533 G/A 939 C/T); EG-STS7 (774 G/A 805 G/A 817 G/A); GALT (478 G/A 758 C/G 866 C/G); IKBA (4476 C/T 4679 T/A 4904 G/T); LEPR (426 A/C/G 810 G/C); P450-STS18 (71 C/T 138 G/A 361 G/A); PBE3 (115 G/A 192 A/T 555 T/G); PBE42 (111 G/A 118 T/G 181 C/T); PBE43 (314 C/T 471 C/A 524 C/T); PBE 57 (75 C/T 109 G/A 197 C/T 268 T/G); PBE59 (276 C/T 494 C/T);

PBE 64 (115 G/A 419 C/G 515 C/T); PBE 73 (93 A/C 116 A/G 177 C/T 477 C/T); PBE 84 (14 A/G 97 T/C 428 T/C); PBE132 (102 C/G 127 A/G 193 C/T 371 A/G); PBE137 (121 C/T 278 C/T 409 A/G); PRKAG-STS3 (1845 G/A 1938 G/A 2050 G/A); RYRA-STS6 (402 A/C 408 C/G 567 A/C); SCAMP (184 C/T 389 G/A 516 C/T 582 G/A 939 A/T); VAN-STS1 (889 C/T 950 C/T 1009 G/A 1065 C/T); WSCR-STS1 (411 C/T 599 C/T); BG (1257 C/T 1323 C/T 1425 C/A 1966 C/T); AMG (907 A/C 975 A/G 1467 A/G); LCN (87 C/T 373 G/C 402 C/T); CTSL (252 A/G 272 A/G); PBE112 (61 C/T 87 C/T); and MYF5 (1833 A/G 2204 C/G 2335 A/C).

[C10] The method of claim 1, wherein the plurality of markers are selected from the group consisting of swine mitochondrial markers designated at positions 15543 (C/T), 15558 (A/T), 15615 (C/T), 15616 (C/T), 15675 (C/T), 15714 (C/T), 15840 (C/T), and 16127 (G/A).

[C11] A system for identification of an animal from which a sample is derived, the system comprising:

- (a) obtaining a sample from the animal or from a processed product of the animal;
- (b) performing single nucleotide polymorphism tracking (SNPTrack) analysis of the sample comprising a plurality of markers, wherein the plurality of markers comprise at least one SNP;
- (c) storing one or more SNPTracks of the sample in a computer system;
- (d) comparing the one or more SNPTracks of the sample with known SNPTracks in a database; and
- (e) identifying the animal.

[C12] A computer system for identification of a sample, the computer system comprising:

- (a) a software module comprising instructions operative to provide a searchable database comprising a plurality of SNPTracks of a plurality of animals, the SNPTracks comprising a plurality of markers, wherein the plurality of markers comprise at least one SNP;

- (b) a software module comprising instructions operative to provide an algorithm to determine exclusion probabilities; and
- (c) a software module comprising instructions operative to provide an interface to accept and compare a query SNPTrack with the plurality of SNPTracks in the database to identify the sample.

[C13] A method of developing a database for identification of an animal or a product derived from the animal, the method comprising:

- (a) performing SNPTrack analysis of a plurality of samples obtained from a plurality of animals from one or more sources with a plurality of markers, wherein the plurality of markers comprise one SNP;
- (b) obtaining and storing the SNPTracks of the plurality of samples in a database, wherein the database is searchable;
- (c) performing SNPTrack analysis of the product;
- (d) comparing one or more SNPTracks of the product with the SNPTracks of the plurality of the samples stored in the database; and
- (e) identifying the product.

[C14] The method of claim 13, wherein the database further comprises data selected from the group consisting of production farm, farm of origin, retail outlet, wholesaler, breeding record, animal identification, offspring data, sibling data, and lineage data.

[C15] The method of claim 13, wherein the one or more sources are selected from the group consisting of farm of origin, production farm, processing center, retail outlet, distribution center, and wholesaler.

[C16] The method of claim 13, wherein the SNPTracks of the plurality of the samples stored in the database are associated with an animal identification system.

[C17] The method of claim 13, wherein the SNPTrack analysis is performed on the plurality of samples between birth and slaughter of the plurality of the animals.

[C18] The method of claim 13, wherein the database has limited access to an authorized user.

[C19] A method of obtaining a SNPTrack of a sample, the method comprising:

- (a) selecting a first marker such that the first marker has at least one SNP, wherein more than one SNP are within a 0.2 to about 10 kb region in the genome;
- (b) selecting a second marker such that the second marker has at least one SNP, wherein more than one SNP are within a 0.2 to about 10 kb region in the genome;
- (c) performing SNPTrack analysis of the sample with the first and second markers; and
- (d) obtaining the SNPTrack of the sample.

[C20] The method of claim 18 further comprising performing a SNPTrack analysis with a plurality of markers and obtaining the SNPTrack for the plurality of markers.

[C21] A nucleotide segment comprising a plurality of SNPs, wherein the segment is amplified with primers listed in TABLE 3.

[C22] The nucleotide segment of claim 21, wherein the nucleotide segment is selected from the group consisting of autosomes, sex chromosomes, and mitochondrial DNA.

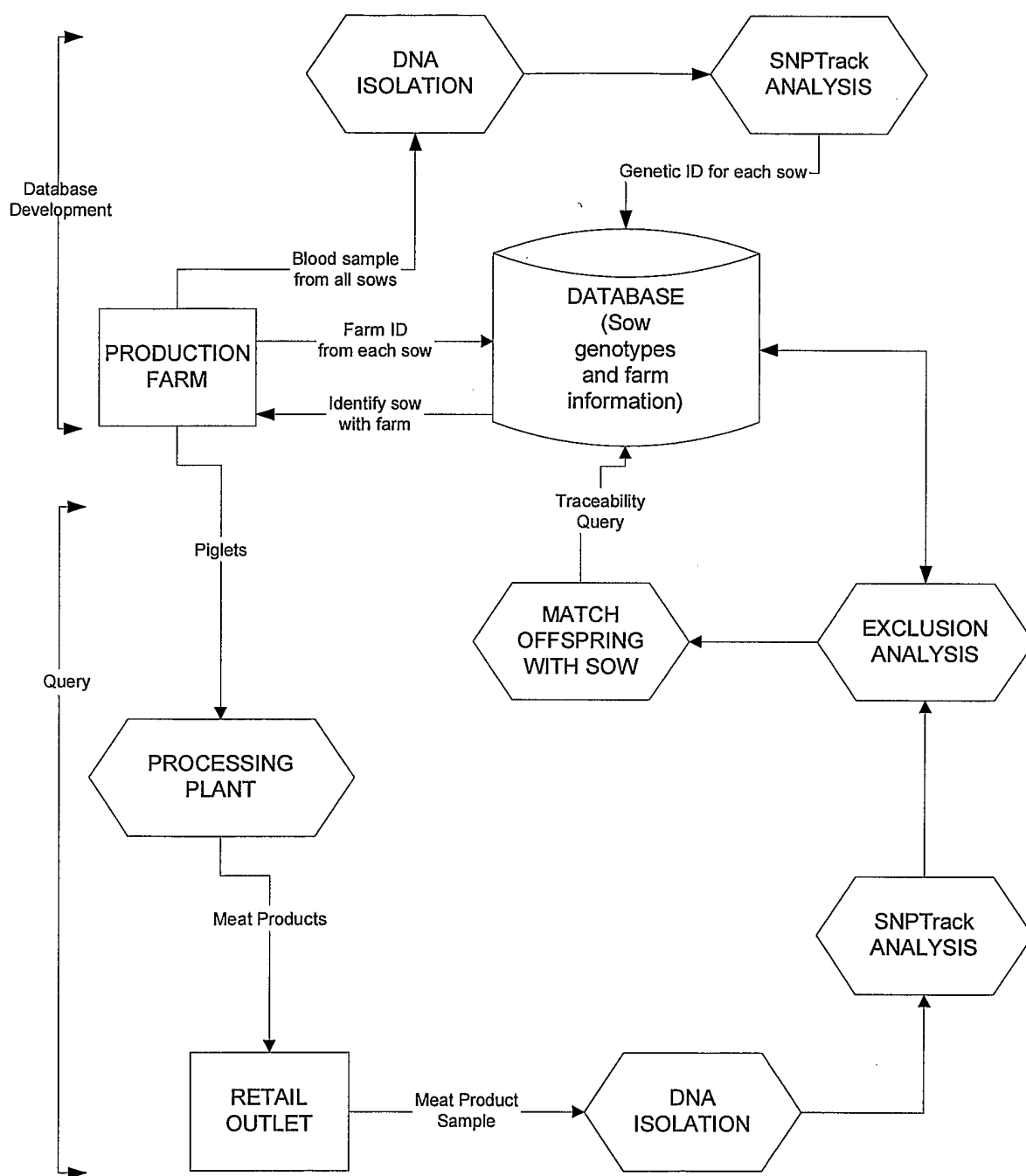
[C23] A high-throughput system for tracking an animal and a meat product through a supply chain, the system comprising:

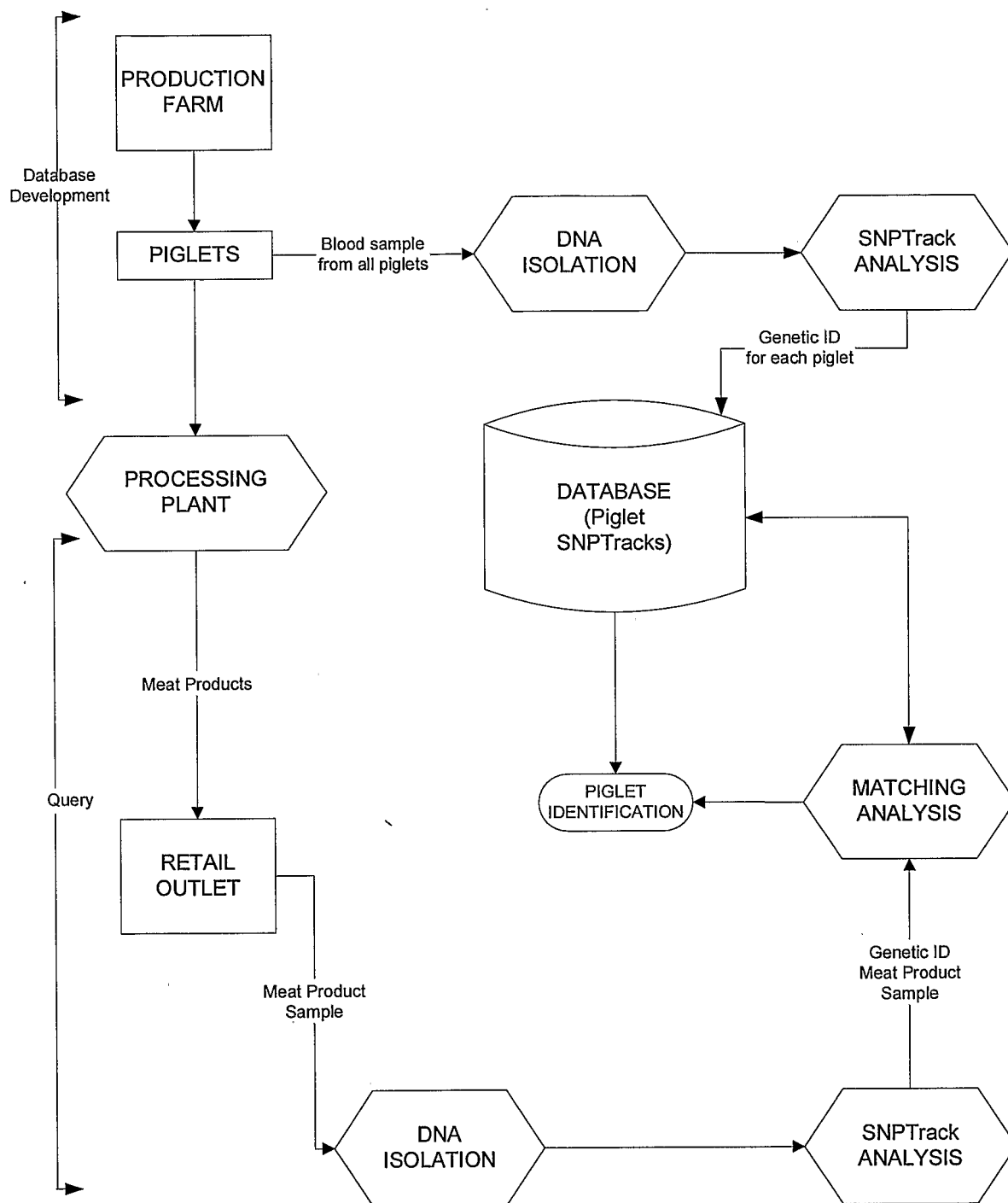
- (a) performing SNPTrack analysis of a plurality of samples obtained from a plurality of animals;
- (b) obtaining and storing a plurality of SNPTracks in a database, wherein the database is searchable;
- (c) obtaining a plurality of samples from meat products;
- (d) performing SNPTrack analysis of the plurality of samples from meat products; and
- (e) tracing the plurality of samples from meat products to a location in the supply chain.

[C24] The high-throughput system of claim 23, wherein the plurality of samples from the plurality of animals are obtained prior to slaughtering in a farm or a processing center.

- [C25]** The high-throughput system of claim 23, wherein the supply chain comprises identification of animals and their products from birth to post-slaughter.
- [C26]** The high-throughput system of claim 23, wherein the supply chain comprises from a farm of origin to a consumer.
- [C27]** A software module comprising instructions operative to provide a database comprising SNPtracks identified in TABLE 7.
- [C28]** A SNPTrack analysis kit to identify an animal comprising:
- (a) a plurality of oligonucleotides corresponding to a plurality of markers that contain at least one SNP, wherein more than one SNP are within a 0.2 to about 10 kb region of each marker;
 - (b) reagents to perform SNPTrack analysis; and
 - (c) access to compare SNPTracks with a plurality of SNPTracks in a database to identify the animal.
- [C29]** The kit of claim 25 further comprising an instruction manual to identify the animal.
- [C30]** A method of tracing an infected meat sample to a particular location, the method comprising:
- (a) performing SNPTrack analysis of a plurality of samples obtained from a plurality of animals with a plurality of markers, wherein the plurality of markers comprise at least one SNP;
 - (b) obtaining and storing the SNPTracks of the plurality of samples in a database, wherein the database is searchable;
 - (c) performing SNPTrack analysis of the infected meat sample;
 - (d) comparing one or more SNPTracks of the infected meat sample with the SNPTracks of the plurality of the samples stored in the database; and
 - (e) tracing the infected meat sample to a particular location.
- [C31]** The method of claim 30 wherein the infected meat sample is a beef sample infected with Mad Cow Disease or bovine spongiform encephalopathy (BSE).
- [C32]** The method of claim 30 wherein the infected meat sample is a sheep or goat sample infected with scrapie disease.

- [C33]** The method of claim 30 wherein the infected meat sample is an infected pork sample.
- [C34]** The method of claim 30 wherein the infected meat sample is obtained from a meat product in the market.
- [C35]** The method of claim 30 wherein the particular location is selected from the group consisting of farm of origin, production farm, processing center, retail outlet, distribution center, and wholesaler.
- [C36]** A method of assuring food safety of a meat product, the method comprising:
- (a) performing SNPTrack analysis of a plurality of samples obtained from a plurality of animals prior to slaughtering with a plurality of markers, wherein the plurality of markers comprise at least one SNP;
 - (b) obtaining and storing a plurality of SNPTracks of the plurality of samples in a database, wherein the database is searchable;
 - (c) tracing the meat product to its source by performing SNPTrack analysis of the meat product and comparing a SNPTrack of the meat product with the SNPTracks of the plurality of samples stored in the database; and
 - (d) determining if the meat product from the source is safe.

**FIG. 1**

**FIG. 2**

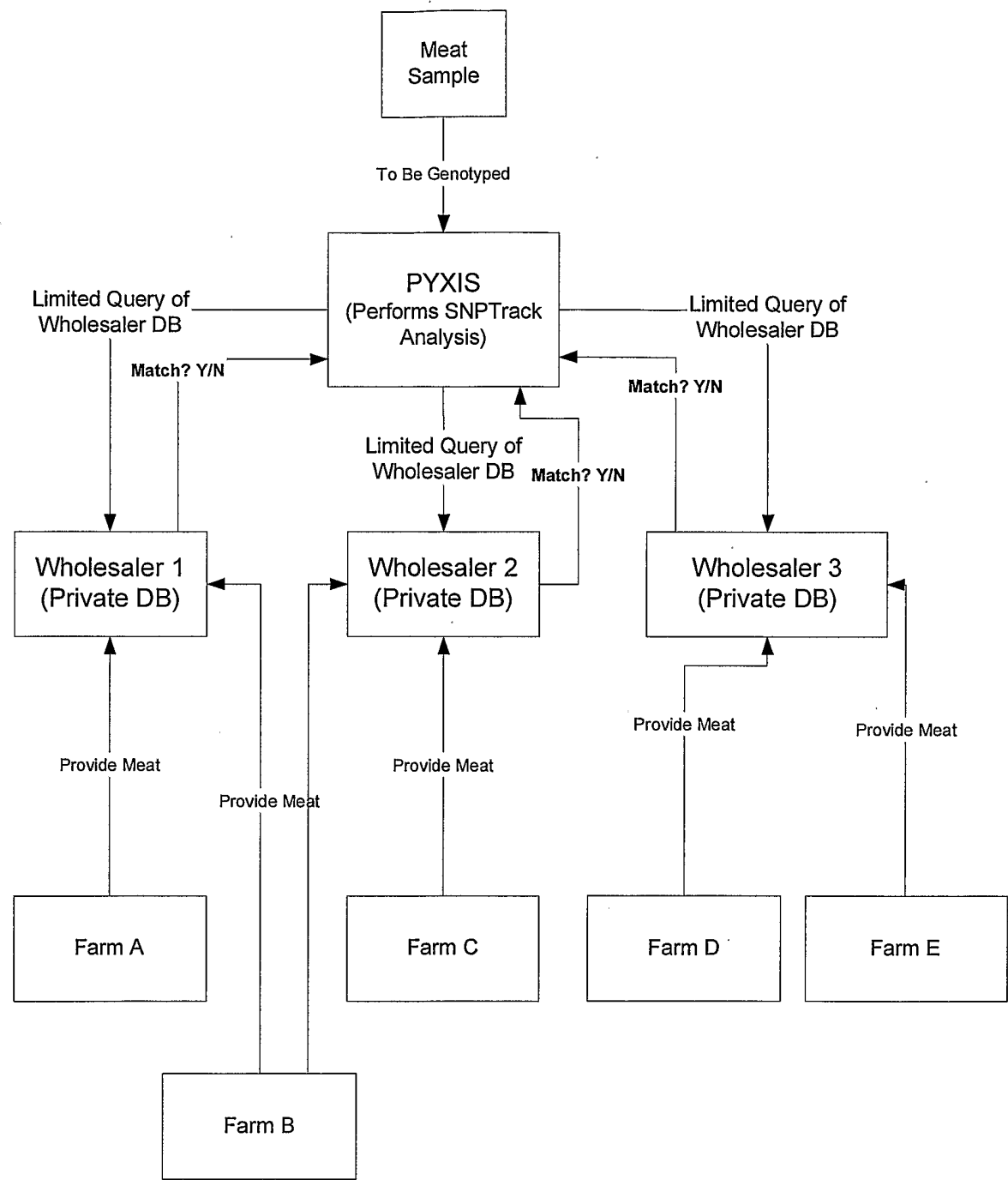


FIG. 3A

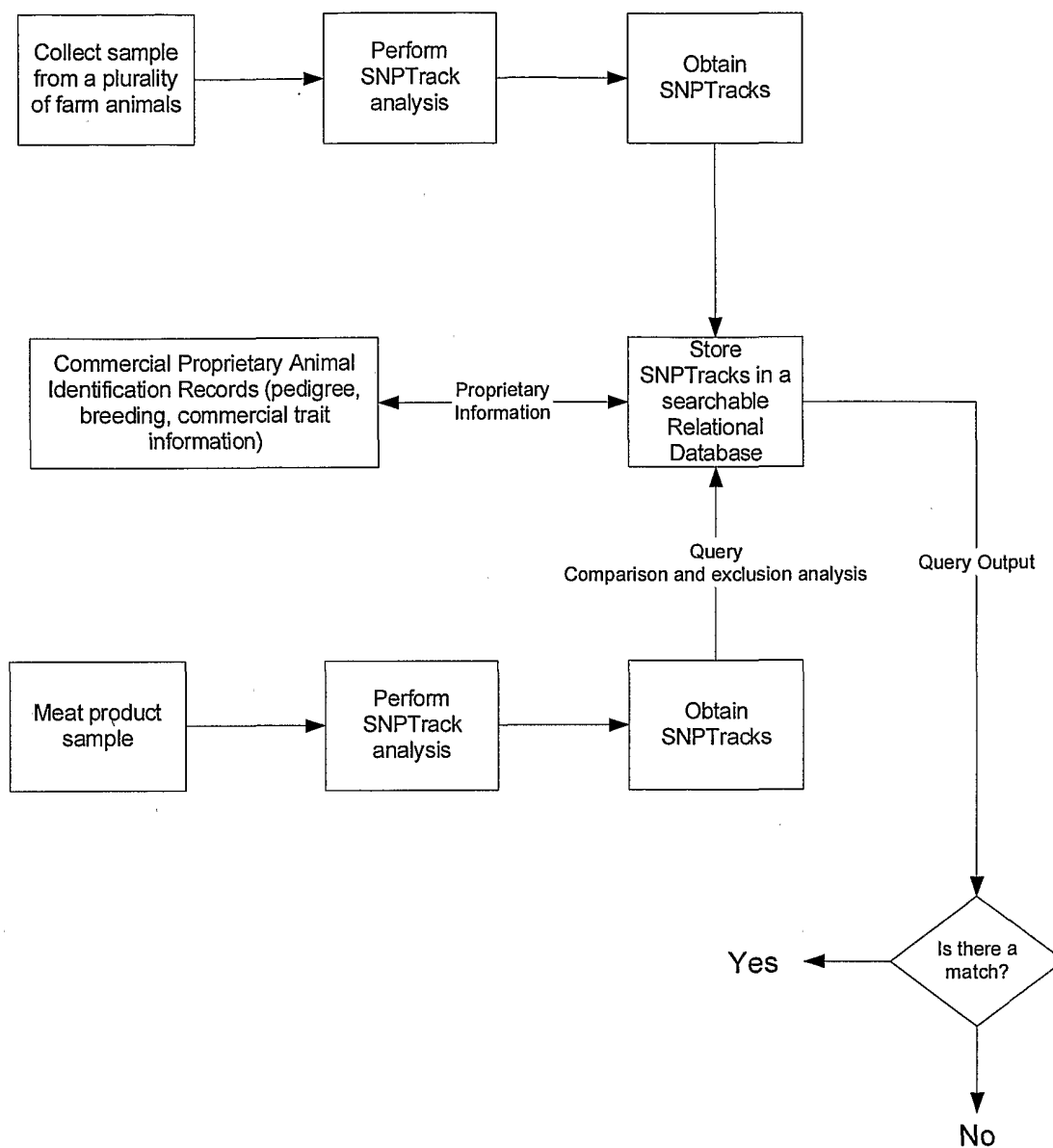


FIG. 3B

SNPTrack determination (2 SNP example)

Position	1	2			
	XXTXXXXX	GXXXXXXXXX	♂	Population	Original Allele (TG)
	XXAXXXXX	CXXXXXXXXX			
	XXTXXXXX	GXXXXXXXXX	♀		
	XXAXXXXX	CXXXXXXXXX			
	XXGXXXXX	GXXXXXXXXX	♂	Founder Animal 1	Mutant Allele (GG)
	XXCXXXXX	CXXXXXXXXX			
	XXTXXXXX	GXXXXXXXXX	♀		Original Allele (TG)
	XXAXXXXX	CXXXXXXXXX			
	XXGXXXXX	GXXXXXXXXX	♂		Two Alleles (TG/GG)
	XXCXXXXX	CXXXXXXXXX			
	XXTXXXXX	GXXXXXXXXX	♀		
	XXAXXXXX	CXXXXXXXXX			
	XXGXXXXX	AXXXXXXXXX	♂	Founder Animal 2	2nd Mutant Allele (GA)
	XXCXXXXX	TXXXXXXXXX			
	XXTXXXXX	GXXXXXXXXX	♀		
	XXAXXXXX	CXXXXXXXXX			
	♂♀				
	TG/TG	homozygous			Three Alleles (TG/GG/GA)
	TG/GG	heterozygous			
	TG/GA	heterozygous			
	GG/GG	homozygous			
	GG/GA	heterozygous			
	GA/GA	homozygous			Six Genotypes

	Genotype Data		Genotype Data		Predicted Alleles
TG/TG homozygous	T	G	T	G	TG/TG
TG/GG heterozygous	T/G	G	T/G	G	TG/GG
TG/GA heterozygous	T/G	G/A	T/G	G/A	TG/GA
GG/GG homozygous	G	G	G	G/A	GG/GA
GG/GA heterozygous	G	G/A	G	A	GA/GA
GA/GA homozygous	G	A			

FIG. 4

SNPTrack determination (3 SNP Example)

Position	1	2	3			
	XXTXXXXX	GXXXXXX	AXXX	♂	Population	Original Allele (TGA)
	XXAXXXXX	CXXXXXX	TXXX	♀		
	XXTXXXXX	GXXXXXX	AXXX	♂	Founder Animal 1	Mutant Allele (GGA)
	XXCXXXXX	CXXXXXX	TXXX	♀		Original Allele (TGA)
	XXGXXXXX	GXXXXXX	AXXX	♂		Two Alleles (TGA/GGA)
	XXCXXXXX	CXXXXXX	TXXX	♀		
	XXGXXXXX	AXXXXXX	AXXX	♂	Founder Animal 2	2nd Mutant Allele (GAA)
	XXCXXXXX	TXXXXXX	TXXX	♀		Original Allele (TGA)
	♂♀					
	TGA/TGA	homozygous			Three Alleles (TGA/GGA/GAA)	
	TGA/GGA	heterozygous				
	TGA/GAA	heterozygous				
	GGA/GGA	homozygous			Six Genotypes	
	GGA/GAA	heterozygous				
	GAA/GAA	homozygous				
	XXGXXXXX	AXXXXXX	CXXX	♂	Founder Animal 3	3 rd Mutant Allele (GAC)
	XXCXXXXX	TXXXXXX	GXXX	♀		
	XXTXXXXX	GXXXXXX	AXXX	♂		
	XXAXXXXX	CXXXXXX	TXXX	♀		

FIG. 5A

SNPTrack determination (3 SNP example continued)

♂♀		
TGA/TGA	homozygous	
TGA/GGA	heterozygous	
TGA/GAA	heterozygous	
TGA/GAC	heterozygous	
GGA/GGA	homozygous	
GGA/GAA	heterozygous	
GGA/GAC	heterozygous	
GAA/GAA	homozygous	
GAA/GAC	heterozygous	
GAC/GAC	homozygous	
		Four Alleles (TGA/GGA/GAA/GAC)
		Ten Genotypes

TGA →	GGA	
TGA		
GGA →	GAA	
TGA		
GGA		
GAA →	GAC	
TGA		
GGA		
		#1 #2 #3 T/G G/A A/C
		Four Alleles (TGA/GGA/GAA/GAC)
		TAA/TAC/TGC/GGC Do not exist

- 1. TGA
- 2. GGA
- 3. GAA
- 4. GAC

		All possible Genotypes			Genotype Data
A)	1/1	TGA/TGA	T	G	A
B)	2/2	GGA/GGA	G	G	A
C)	3/3	GAA/GAA	G	A	A
D)	4/4	GAC/GAC	G	A	C
E)	1/2	TGA/GGA	T/G	G	A
F)	1/3	TGA/GAA	T/G	G/A	A
G)	1/4	TGA/GAC	T/G	G/A	A/C
H)	2/3	GGA/GAA	G	G/A	A
I)	2/4	GGA/GAC	G	G/A	A/C
J)	3/4	GAA/GAC	G	A	A/C

* Genotype data predicts the exact allele combination!

FIG. 5B

SNPTrack determination (3 SNP example continued)

1. TGA
2. GGA
3. GAA
4. GAC

Offspring Genotype (A)

T G A

All Parent Genotypes

Genotype Data

A)	1/1	TGA/TGA	T	G	A	**	**
B)	2/2	GGA/GGA	G	G	A		
C)	3/3	GAA/GAA	G	A	A		
D)	4/4	GAC/GAC	G	A	C		
E)	1/2	TGA/GGA	T/G	G	A	**	**
F)	1/3	TGA/GAA	T/G	G/A	A	**	**
G)	1/4	TGA/GAC	T/G	G/A	A/C	**	**
H)	2/3	GGA/GAA	G	G/A	A		
I)	2/4	GGA/GAC	G	G/A	A/C		
J)	3/4	GAA/GAC	G	A	A/C		

1. TGA
2. GGA
3. GAA
4. GAC

Offspring Genotype (B)

G G A

All Parent Genotypes

Genotype Data

A)	1/1	TGA/TGA	T	G	A	**	**
B)	2/2	GGA/GGA	G	G	A		
C)	3/3	GAA/GAA	G	A	A		
D)	4/4	GAC/GAC	G	A	C		
E)	1/2	TGA/GGA	T/G	G	A	**	**
F)	1/3	TGA/GAA	T/G	G/A	A	**	
G)	1/4	TGA/GAC	T/G	G/A	A/C	**	
H)	2/3	GGA/GAA	G	G/A	A	**	**
I)	2/4	GGA/GAC	G	G/A	A/C	**	**
J)	3/4	GAA/GAC	G	A	A/C		

1. CAG
2. GTG
3. GAG
4. GAC

10 possible Genotypes

Parent Genotype

SNP/SNP Match

SNPTrack Match

A)	1/1	CAG/CAG	A)	4	AEFG	AEFG	4
B)	2/2	GTG/GTG	B)	4	BEHI	BEHI	4
C)	3/3	GAG/GAG	C)	7	CEFGHIJ	CFHJ	4
D)	4/4	GAC/GAC	D)	4	GDIJ	GDIJ	4
E)	1/2	CAG/GTG	E)	9	ABCEFGHIJ	ABEFGHI	7
F)	1/3	CAG/GAG	F)	8	ACEFGHIJ	ACEFGHIJ	7
G)	1/4	CAG/GAC	G)	9	ACDEFGHIJ	ADEFGIJ	7
H)	2/3	GTG/GAG	H)	8	BCEFGHIJ	BCEFGHIJ	7
I)	2/4	GTG/GAC	I)	9	BCDEFGHIJ	BDEFGHIJ	7
J)	3/4	GAG/GAC	J)	8	CDEFGHIJ	CDEFGHIJ	7

FIG. 5C