Title

Applied equine genetics

Journal

Equine Veterinary Journal, 46(5)

ISSN

0425-1644

Authors

Finno, CJ

Bannasch, DL

Publication Date

2014-09-01

DOI

10.1111/evj.12294

Peer reviewed

Published in final edited form as:
Equine Vet J. 2014 September ; 46(5): 538–544. doi:10.1111/evj.12294.

Applied equine genetics

C. J. FINNO[*] and **D. L. BANNASCH**[†]
Department of Veterinary Population Medicine, College of Veterinary Medicine, University of Minnesota, St. Paul, USA

[†]Department of Population Health and Reproduction, Veterinary Medical Teaching Hospital, University of California, Davis, USA

Summary

Genome sequencing of the domestic horse and subsequent advancements in the field of equine genomics have led to an explosion in the development of tools for mapping traits and diseases and evaluating gene expression. The objective of this review is to discuss the current progress in the field of equine genomics, with specific emphasis on assembly and analysis of the reference sequence and subsequent sequencing of a Quarter Horse mare; the genomic tools currently available to researchers and their implications in genomic investigations in the horse; the genomics of Mendelian and non-Mendelian traits; the genomics of performance traits and considerations regarding genetic testing in the horse. The whole-genome sequencing of a Quarter Horse mare has provided additional variants within the equine genome that extend past single nucleotide polymorphisms to include insertions/deletions and copy number variants. Equine single nucleotide polymorphism arrays have allowed for the investigation of both simple and complex genetic traits while DNA microarrays have provided a tool for examining gene expression across various tissues and with certain disease conditions. Recently, next-generation sequencing has become more affordable and both whole-genome DNA sequencing and transcriptome-wide RNA sequencing are methodologies that are being applied to equine genomic research. Research in the field of equine genomics continues to expand rapidly as the cost of genotyping and sequencing

[*]Correspondence: cjfinno@gmail.com..
Present address: Department of Population Health and Reproduction, University of California-Davis, Room 4206 Vet Med 3A, One Shields Ave, Davis, California 95616, USA.

Authors' declaration of interests
The Grayson Jockey Club Research Foundation commissioned this article as part of a series summarising areas relating to their priorities for research funding. Other workers at the authors' primary institutes hold current and previous research grants funded by the Grayson Jockey Club Research Foundation.

Authorship
C.J. Finno performed the review of the literature, preparation and final approval of the manuscript. D.L. Bannasch contributed to the preparation and final approval of the manuscript.

Additional references are available online (Supplementary Item 5).

Supporting information
Additional Supporting Information may be found in the online version of this article at the publisher's website:
Supplementary Item 1: Equine diseases/traits with known genetic mutation.
Supplementary Item 2: Available equine coat colour genetic tests with published genetic mutations.
Supplementary Item 3: Mendelian diseases/traits in the horse with the causative mutation not yet identified or published.
Supplementary Item 4: Diseases/traits in the horse with unknown modes of inheritance and genetic mutations not yet identified.
Supplementary Item 5: Additional references.

decreases, resulting in a need for quality bioinformatics software and expertise to appropriately handle both the size and complexity of these data.

Keywords

horse; genetics; genomics; DNA test

Introduction

Genome sequencing of the domestic horse and subsequent advancements in the field of equine genomics have led to an explosion in the development of tools for mapping traits and diseases and evaluating gene expression. This review focuses on the assembly and analysis of the equine genome sequence in Twilight, a Thoroughbred mare, and the subsequent sequencing of a Quarter Horse mare; the genomic tools and resources currently available to researchers; the genomics of Mendelian and non-Mendelian traits and diseases; the genomics of performance traits and considerations regarding genetic testing in the horse. A more detailed history of the horse genome project, with additional information on disease traits for which genetic tests are currently available, can be found in a previous publication of the *Equine Veterinary Journal* [1]. Recent textbooks devoted to equine genomics are referenced for further review [2,3]. Emphasis in this review article will be on the development and use of genomic tools, including single nucleotide polymorphism (SNP) arrays, microarrays and next-generations sequencing techniques, along with more in-depth information regarding the genomics of performance traits in the horse.

Equine genome sequence and assembly

The First International Equine Gene Mapping Workshop took place in October 1995 in Lexington, Kentucky, and signified the beginning of an organised equine genomics group. First-generation maps of the equine genome contained markers that were assigned to various equine chromosomes using approaches such as synteny analysis (preserved colocalisation of markers on chromosomes of different species) [4,5], genetic linkage mapping (tendency of markers that are located close together to be inherited together) [6-9] and fluorescent *in situ* hybridisation (the ability to detect specific DNA sequences on chromosomes) [10-14]. Markers consisted of *type I* markers (associated with genes of known function such as expression sequence tags) and *type II* markers (anonymous genomic segments, including microsatellites [repeating sequences of 2–6 bases of DNA]). Radiation hybrid (RH) maps, which use x-ray breakage of chromosomes to determine the distance between markers, were generated for equine chromosomes in order to develop a high-resolution, ordered physical map consisting of uniformly distributed polymorphic markers. The first-generation RH map contained 733 markers [15] and the second-generation RH map contained 4103 markers [16]. These maps provided the initial tools required to assemble the whole genome sequence of the horse.

The first genome sequence of the domestic horse was published in November 2009 as a collaborative effort by the worldwide equine research community [17]. DNA from a single Thoroughbred mare, Twilight, was used to construct the genome sequence. From a panel of

candidate horses, Twilight was selected based on a high level of homozygosity within her major histocompatibility complex, a region of high diversity relevant to immune system function that is normally challenging to assemble. A whole-genome shotgun method was used to sequence Twilight, where large fragments of genomic DNA were randomly sheared and subsequently inserted into libraries for replication and sequencing. DNA libraries are collections of DNA fragments that have been inserted into vectors for sequencing. For Twilight, the libraries were various sizes including 4, 10 and 40 Kb, allowing easier assembly of the sequence data. Due to the shearing process used, different sized DNA fragments were created that were overlapped and joined to form contigs, or consensus regions. Overlapping contigs were then joined together into larger sequences called scaffolds. A high-quality draft assembly was constructed and additional sequences were provided by the inclusion of bacterial artificial chromosome end sequences from a related male Thoroughbred horse [18]. The resulting assembly (EqCab2.0) has 6.8-fold sequence coverage. The genome size of the horse was estimated to be around 2.7 Gb [17].

In order to determine the chromosomal locations and orientation of the scaffolds within the equine genome, the genome assembly was compared with the known maps for the horse [8,16,19,20]. The equine gene set, as annotated by the ENSEMBL pipeline, predicts 20,322 protein-coding genes (Ensembl build 52.2b), similar to human (16,617; Ensembl build 73). EqCab2.0 is hosted through public genome browsing sites, including the University of California Santa Cruz, Ensembl and the National Center for Biotechnology Information.

As part of the equine genome project, partial genome sequences were obtained from seven additional horses from seven different breeds (Akhal-Teke, Andalusian, Arabian, Icelandic, Quarter Horse, Standardbred, Thoroughbred) to provide a database of genetic markers [17]. A SNP map of more than one million markers was generated from the approximately 700,000 SNPs discovered in the Twilight genome and the additional 400,000 SNPs discovered from approximately 100,000 whole genome shotgun reads from these seven horses. As a result, in addition to microsatellite markers, the SNP map was also available as a genomic tool to investigate traits and inherited diseases.

In 2011, whole-genome sequencing of an individual American Quarter Horse mare was performed using massively parallel paired-end sequencing [21]. This particular mare was selected based on having no introgression of Thoroughbred lines during the preceding 4 generations. Approximately 97% of the 75-bp paired end reads aligned to the reference genome, resulting in an average of 24.7× sequence coverage of the Quarter Horse mare's genome. Almost 82,000 reads mapped to the reference mitochondrial genome, resulting in an average of 355.6× coverage, and approximately 12.8 million reads were mapped to the unassembled chromosomes. The remaining 12.6 million reads were *de novo* assembled, generating 19.1 Mb of new horse genomic sequence.

One of the most exciting results from the sequencing of the Quarter Horse came from the extensive variant detection analysis performed. Prior to this study, the catalogue of genetic variants in the horse consisted of 1,163,580 SNP polymorphisms, with no annotated insertion/deletion polymorphisms or copy number variants. Upon sequencing of the Quarter Horse, 3.1 million SNPs, 193,000 insertions/deletions, and 282 copy number variants were

detected and subsequently annotated [21]. Pathway analyses of biological pathways containing heterozygous nonsynonymous SNPs were performed and results compared between the Quarter Horse and reference Thoroughbred mare. It was discovered that the Quarter Horse had SNPs enriched in pathways for sensory proprioception, cellular processes and signal transduction. As this particular mare was not selected for sequencing based on homozygosity and is a different breed than the reference sequence, this genome provides an excellent resource for studies of genetic variation.

In addition to sequencing of contemporary breeds, the genome sequence of the ancient horse has recently been investigated, revealing that the *Equus* lineage gave rise to all contemporary horses, zebras and donkeys and that the lineage originated 404.5 million years ago [22]. Additional sequencing of domestic horse breeds and a Przewalski's horse has revealed no evidence of recent admixture between the domestic horse breeds and Przewalski's horse [22], thereby supporting the notion that Przewalski's horses represent the last surviving wild horse population. Readers are directed to publications regarding ancient DNA sequences for further information [23-25].

Efforts are currently underway to improve upon the equine reference sequence through the creation of EquCab3.0 by improving upon the Twilight sequence.

Genomics tools

SNP beadchip

With the discovery of approximately one million SNPs from the sequencing efforts described above [17], sufficient markers were available to construct a whole genome SNP array. Preference was given to SNPs that were discovered in the alternate breeds (Akhal-Teke, Andalusian, Arabian, Icelandic, Quarter Horse, Standardbred, Thoroughbred), resulting in > 67% of the SNPs selected from one of these other breeds relative to Twilight [17]. The first-generation array (Illumina EquineSNP50 Beadchip, San Diego, California, USA; 2008) contained 54,602 SNPs that reliably produced genotypes when assessed on a group of 354 horses representing 14 breeds [26]. Of the ~54 k SNPs, 53,524 were polymorphic (i.e. having at least one heterozygote within the sample set). The EquineSNP50 Beadchip spanned the entire equine genome, with the exception of the Y chromosome, with an average spacing between SNPs of 43.1 kb across the 31 autosomes and few gaps larger than 500 kb.

In the original report describing the sequencing of Twilight, power estimates based on the length of linkage disequilibrium (LD; level of association between markers) in the horse, the number of haplotypes (i.e. combination of adjacent DNA sequences on a chromosome) within haplotype blocks and the polymorphism rate, suggested that more than 100,000 SNPs would be required to map traits within and across breeds [17]. The first generation SNP array was validated on a panel of samples representing 14 domestic horse breeds and 18 evolutionarily related species [26]. Based on the extent of LD in breeds such as the Quarter Horse and Mongolian horse, it has been recommended that more markers are required for effective mapping in ancient breeds and those with a large effective population size [26]. Therefore, the first-generation Equine SNP50 Beadchip represented about one-half of the

estimated marker density required for adequately powered association studies in breeds with an average or high degree of LD.

The Equine SNP50 Beadchip was used to evaluate population structure in 744 individuals from 33 breeds of horses [27]. Variation found among breeds was used to identify genes and genetic variants targeted by selective breeding (i.e. signatures of selection). This study identified variants in the American Paint Horse and American Quarter Horse breeds significantly associated with altered muscle fibre type proportions favourable for sprinting ability, variants in breeds that perform alternative gaits and genomic regions involved in the determination of size [27].

In January 2011, the Equine SNP50 Beadchip was replaced by a second-generation SNP array, the Equine SNP70Beadchip, which contains approximately 74,500 SNP markers with an average of 1.5 SNPs per 50 kb. This platform contains the original 53,500 markers from the Equine SNP50 Beadchip and additional SNPs were chosen to address gaps and improve global coverage across the genome. Additional SNPs were provided from the 7 discovery breeds, Twilight and RNA sequencing (RNA-seq) data [28] (see below). The equine SNP70 Beadchip contains additional SNPs to enhance the coverage of the equine major histocompatibility complex on chromosome 20 as well as SNPs on the X chromosome and 2 SNPs on the Y chromosome.

Association studies, using the equine SNP chips, were used to identify a chromosomal region containing a strong candidate gene for lavender foal syndrome and subsequent sequencing discovered the genetic mutation responsible for the disease [29]. In addition to lavender foal disease, the SNP50 Beadchip was used to identify associations with SNP markers and lead to the subsequent identification of genetic mutations for foal immunodeficiency syndrome [30] and a mutation that is permissive for gaitedness in the horse [31]. Association studies using the equine SNP chips have also identified quantitative trait loci for further investigation in osteochondritis dissecans in Thoroughbreds [32], risk loci for recurrent laryngeal neuropathy [33], loci for body size [34], and candidate regions for guttural pouch tympany [35], equine uveitis [36] and insect bite hypersensitivity [37]. As the estimated marker density of 100,000 SNPs has still not been achieved, efforts are currently underway to develop a third generation SNP Beadchip, with a targeted 700,000 SNPs. Estimated availability of this array is scheduled for 2014.

DNA microarrays

The study of tissue-specific gene expression in the horse under particular conditions and considering certain disease processes is an ever-expanding area of research. The first tools developed to study gene expression, through evaluation of the mRNA transcriptome (all of the RNAs transcribed from the genome that code for proteins), included expressed sequence tags [38], serial analysis of gene expression [39] and microarrays [40]. Until recently, microarrays were used as the primary experimental method for analysing gene expression in the horse at the transcriptome level. Microarray technology involves isolation of RNA (target) and subsequent hybridisation to specific, known DNA-sequences on the microarray (probe). Hybridisation patterns are then compared to enable the identification of mRNAs that differ in abundance in ≥2 target samples [41].

Initially, human and mouse-specific arrays were used to profile gene expression in the horse. Upon completion of the sequencing of the equine genome, several groups initiated efforts to improve equine-specific microarrays, using the gene prediction models from Ensembl and National Center for Biotechnology Information. Equine-specific microarrays have been used to evaluate gene expression in laminitis [42] and articular cartilage repair [43]. A recent study using microarray technology on placental tissues identified a >900-fold upregulation of mRNA encoding the cytokine interleukin (IL)-22 in chorionic girdle, which is the first time IL-22 has been reported in any cells other than immune cells [44]. As is required for any expression study using microarray technology, these results were confirmed using quantitative RT-PCR. Currently, Agilent provides a horse gene expression microarray with 43,803 probes that can be customised to meet specific research needs (Agilent eArray custom microarray; available at (https://earray.chem.agilent.com/earray/).

Next-generation sequencing

Most recently, RNA-seq methods have been used to refine gene structure models and evaluate gene expression patterns [28,45]. Using RNA-seq generates quantitative and qualitative data concurrently, while also providing insight into alternative transcripts from the same gene. These RNA-seq techniques are currently being applied in the investigation of many equine diseases. Also, RNA-seq methods were used to investigate the role of genomic imprinting in the horse. Genomic imprinting is an epigenetic phenomenon by which certain genes can be expressed in a parent-of-origin-specific manner. In 2012, Wang *et al*. used RNA-seq methodologies to exclude the role of X-imprinted activation, which has been demonstrated to occur in extra-embryonic studies of other species, in X-inactivation within the horse and mule placenta [46]. The role of imprinting was further evaluated, using RNA-seq, to determine that there is a paternal bias to expressed genes in the horse placenta, thereby highlighting the importance of the placenta as the tissue for genomic imprinting [47].

Next-generation sequencing techniques are also being employed at a DNA level (DNA-seq). These more efficient technologies allow sequences representing the entire genome of a horse to now be obtained at an affordable cost and offer unprecedented insight into the number of variants (SNPs, insertions, deletions, rearrangements) within the genome. Many studies that initially employed genome-wide association techniques (see above) are currently using next-generation sequencing to investigate further a region of association.

As the next-generation sequencing continues to become more and more affordable, whole genome sequences in the horse are being generated worldwide. Analysis and storage of these large amounts of sequence data can become problematic. At this time, there are many software programs available to perform quality control, alignment to the reference sequence, assembly of unaligned sequences, and variant detection in next-generation sequence data [48]. For RNA-seq, there are various algorithms for quantifying and comparing gene expression between conditions [49]. An extensive knowledge of bioinformatics has become essential in processing the sequences obtained through next-generation sequencing.

Appropriate strategies to investigate genetic traits

With the increased availability and affordability of genetic tools in horses, ample opportunities exist to apply these molecular tools to a wide array of equine diseases. To investigate genetic diseases in horses, both the sample population and the tools chosen warrant consideration. For disease susceptibility traits, accurate phenotyping is essential. Often, selection of affected cases is more straightforward than selection of an appropriate control group. However, the phenotype to be studied should be decided *a priori* and its severity graded, as this becomes useful when selecting cases for sequencing. Whenever possible, a control population should be selected to maximise the chance that the control horses would never manifest the phenotype and with consideration placed on the age of disease onset, environmental risk factors and degree of relatedness between the control and the affected populations. The utility of genetic tools is affected by the sample size available. With moderate sized populations with easily discernible phenotypes and an autosomal recessive mode of inheritance, an association between phenotype and a chromosomal locus (loci) may be identified with the currently available techniques through a genome-wide association study using the current Equine SNP BeadArrays. However, small sample sizes, a phenotype influenced by multiple risk factors, diagnostic accuracy or short LD in certain breeds (i.e. Quarter Horse) can fail to detect significant associations due to low statistical power. Both DNA-seq and RNA-seq can be used concurrently to explore gene expression and correlate with whole genome sequence variants. When designing an RNA-seq experiment, it is necessary to factor in the fundamental aspects of sound experimental design; replication, randomisation and blocking. Biological replicates are more important than technical replicates in RNA-seq study design [50]. It is also imperative that the RNA obtained in all cases is from a standardised site in the tissue of interest. Simultaneously performing RNA-seq and whole genome DNA-seq on matching samples enhances the power to detect biologically relevant variants in smaller sample sizes and should be considered for investigating complex genetic traits.

Mendelian diseases and traits

Initial genetic mutations in horses were discovered through the use of comparative genomics. For a certain disease, specific ‘candidate genes’ were investigated based on equivalent diseases in man. In the horse, the genetic mutations for many diseases that have genetic tests currently available, including hyperkalaemic periodic paralysis [51] and severe combined immunodeficiency [52], were uncovered by evaluating candidate genes that had been associated with similar diseases in man. With the sequencing and annotation of whole genome maps, other diseases were discovered through whole genome linkage mapping (hereditary equine regional dermal asthenia [53]), genome-wide association studies with microsatellites (*type I* polysaccharide storage myopathy [54]) and genome-wide association studies using SNP array technology (lavender foal syndrome [29]). At the time of publication, 35 Mendelian diseases and traits have their key genetic mutations identified in the horse, including the mutations encoding for coat colour loci (Supplementary Items 1 and 2). There are an additional 13 diseases or traits that appear to be inherited in a Mendelian fashion but an underlying genetic mutation has not yet been identified or published (Supplementary Item 3). An updated list of equine diseases can be found at the Online

Mendelian Inheritance for Animals webpage (http://omia.angis.org.au/home). With the current technologies available through SNP-association mapping and next-generation sequencing, we should expect to further our understanding of Mendelian traits and diseases.

Diseases with suspected heritable basis

Currently, there are many diseases in the horse with a suspected heritable basis for which a genetic test is not currently available and the mode of inheritance is unclear (Supplementary Item 4). Diseases and traits in this table include both those with strong evidence for a genetic basis based on current research and those that have strong comparative correlates in other species. Many of these more complex diseases and traits have strong environmental influences and may be polygenic. Researchers studying these diseases have the opportunity to utilise new technologies, including the equine SNP70 Beadchip, DNA-seq and RNA-seq, to advance our understanding of genetic variants and gene expression.

Performance traits

Selection for performance traits has been most extensively studied in Warmblood horses [55-57] and Thoroughbred racehorses [58-60]. In Warmbloods, based on the genetic correlations between conformation, performance and radiographic health of the limbs [61], research has been targeted at selecting breeding horses based on these multiple traits [55]. Heritabilities for showjumping was estimated at 0.39–0.61 in Hanoverian horses [62] and 0.12–0.28 for Swedish Warmblood horses [63]. Recently, a genome-wide association study was performed for quantitative trait loci for showjumping abilities in Hanoverians [56]. This study identified 6 QTL regions that contained genes previously identified as performance-related genes in man, including *PAPSS2* (3′-phosphoadenosine 5′-phosphosulfate synthase 2), *MYL2* (myosin, light chain 2, regulatory, cardiac, slow), *TRHR* (thyrotropin-releasing hormone receptor) and *GABPA* (GA binding protein transcription factor, α subunit 60 kDa) [56].

In Thoroughbred racehorses, groups of genes associated with the control of substrate utilisation, insulin signalling and muscle strength seem to be of the greatest importance to performance [58-60,64]. Sequence variants associated with performance traits have been reported in the genes *MSTN* (myostatin) [59,60,64,65], *CKM* (creatine kinase), *COX4I2* (cytochrome c oxidase, subunit 4 isoform 2) [66] and *PDK4* (pyruvate dehydrogenase kinase isoenzyme 4, mitochondrial gene) [67]. Of these genes, the most extensively studied locus has been the myostatin gene (*MSTN*, *GDF-8*). Myostatin is a secreted growth differentiation factor that inhibits muscle differentiation and growth during myogenesis. Sequence and structural variation has been discovered in the proximal upstream, downstream and intergenic sequences of the *MSTN* gene that are associated with optimum racing distance in Thoroughbreds [59,60,65]. There are no variants identified within coding sequence of this gene; all of these associated variants are outside of exons. The variants include: 2 SNPs in intron 1, a 227 bp insertion located 145 bp upstream of the transcriptional start site and four 3′ untranslated region SNPs [60]. A BLAST search identified the 227 bp insertion as a horse-specific repetitive DNA sequence element (SINE) known as ERE-1 [60]. Of these variants, one of the SNPs in intron 1 (g.66493737C>T; $P = 5.24 \times 10^{-13}$) and the SINE

insertion ($P = 5.54 \times 10^{-10}$) were highly associated with the quantitative trait of best race distance in 165 samples [60]. A genetic test was made available, termed the *Equinome Speed Gene Test*, to predict the type of distance best suited to a particular horse based on the genotype at this intronic locus. Individuals homozygous for the 'C' allele (i.e. CC) appear to compete best in faster, shorter distance races, heterozygous horses (CT) are best at middle-distance races and horses homozygous for the 'T' allele (i.e. TT) are best suited in longer-distance races [64]. Short-distance races were considered to have a mean distance of 6.5 ± 1.5 furlongs, medium-distance 9.1 ± 2.3 furlongs and long-distance 11.0 ± 2.1 furlongs [64].

Many horse breeds demonstrate alternate gaits, including pace, regular rhythm ambling, lateral ambling and diagonal ambling. The Icelandic horse has a characteristic gait, termed the tölt, which is a regular ambling gait. A genome-wide association study was performed in 70 Icelandic horses who segregated by gait. Thirty of these horses were classified as 4-gaited (walk, tölt, trot and gallop) and 40 were classified as 5-gaited (walk, tölt, trot, gallop and pace) [31]. A significant association between the ability to pace and an SNP on chromosome 23 was discovered. Subsequent re-sequencing of the region revealed a homozygous haplotype block in the 5-gaited horses, which contained a family of doublsex and mab-3 related transcription factors, DMRT1-3. Upon whole-genome re-sequencing, a single base pair change at codon 301 in *DMRT3* was discovered that led to a premature stop at codon 301. All 5-gaited Icelandic horses are homozygous for this nonsense mutation while nongaited horses are homozygous wild-type. A high frequency of the *DMRT3* mutation was found in horses bred for harness racing [31]. These researchers created *Dmrt3*-null mice that demonstrated that *Dmrt3* is expressed in the spinal cord and is critical for normal development of coordinated locomotor network controlling limb movements [31]. A recent study demonstrated worldwide distribution of the *DMRT3* mutation, occurring in 68 out of the 141 breeds genotyped, most abundant in breeds classified as gaited [68].

Genetic testing

DNA tests can be divided into two categories: mutation tests and linked-marker or haplotype tests. *Mutation* tests are based on assaying an actual mutation that causes disease, whereas the *linked-marker* or *haplotype* test is based on an assay of the genomic region that is known to cause disease, but which is not necessarily the actual mutation. Usually, haplotype tests are offered instead of a mutation test where the functional mutation has not yet been identified.

Mutations that cause disease appear in many different forms. A SNP can cause a disease either by changing an amino acid ('missense' mutation), truncating the amino acid chain ('nonsense' mutation), or altering expression or proper splicing. For example, a missense mutation has been shown to cause *type I* polysaccharide storage myopathy [54] (Supplementary Item 1). Insertions or deletions of a single base pair can cause mutations in the coding sequence by altering the translational frame, which ultimately causes either protein truncation or an elongated abnormal protein.

The basis for DNA testing is PCR. Primers can be designed specifically to amplify the DNA fragment containing either the disease-causing allele or the normal allele. Direct sequencing

of a section of DNA can also be used to determine the animal's genotype. Alternatively, the PCR product can be digested with a restriction enzyme that cleaves the DNA at a particular sequence of bases. To test for the mutation, a restriction enzyme is chosen that shows a different cleavage pattern between the mutant and the normal forms of the DNA. Many different methods are available to assay changes in DNA that lead to disease. Each company that offers a test may choose a different type of assay for the same mutation.

There are limits to all genetic testing. In mutation tests, the specific mutation being assayed is the only factor being evaluated. An animal may have an alternative mutation in the same gene or a mutation in a different gene that causes the same phenotype (phenocopy). It is therefore correct to state that an animal has been 'DNA tested negative' for this specific mutation rather than 'DNA tested clear' of the disease. Linked-marker tests have these same and additional sources of error. In the case of linked-marker tests, recombination events between the marker(s) and the true disease mutation can lead to false-positive and false-negative results. The use of multiple markers that flank the gene of interest (haplotype test) can increase the probability that a recombination event will be identified; and if one is identified, the laboratory will know that the test is not valid for that individual.

It is important to recognise that no authority, association or committee examines quality control of DNA tests that are available in animals. Most tests are published in the scientific literature, not as diagnostic tests, but as articles describing the discovery of the mutation. Much of the research done to identify the mutations involved in the tests is performed at universities and funded by granting agencies that have both financial and intellectual interest in patenting the tests. Companies then license the rights to offer the tests. Veterinarians should contact the laboratories to inquire about available genetic tests for horses and determine if the laboratory maintains a license to run a particular test.

Conclusions

The past two decades have resulted in an explosion of research in the field of equine genomics. With the creation of the original marker maps in the horse, subsequent sequencing and annotation of the complete equine genome and the availability of genomic tools to investigate specific traits and diseases, the study of equine genomics has rapidly accelerated. Efforts are currently underway to improve upon the equine reference sequence through the creation of EquCab3.0 and develop variant databases to expand our knowledge of common variants in the equine genome. Undoubtedly, the next decade will continue to see an increase in the amount of available DNA tests for horses, in addition to an enhanced understanding of specific traits and diseases at the molecular level.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

The authors thank Jessica Petersen for reviewing this manuscript.

Sources of funding

Dr Finno is supported by an NIH K01 award (1K01OD015134-01A1) and a post-doctorate fellowship through Morris Animal Foundation (D12EQ-401).

References

1. Brosnahan MM, Brooks SA, Antczak DF. Equine clinical genomics: a clinician's primer. Equine Vet. J. 2010; 42:658–670. [PubMed: 20840582]
2. Chowdhary, BP., editor. Equine Genomics. Wiley-Blackwell; Ames: 2013.
3. Bailey, E.; Brooks, SA., editors. Horse Genetics 2nd. CABI; Oxford: 2013.
4. Caetano AR, Shiue YL, Lyons LA, O'Brien SJ, Laughlin TF, Bowling AT, Murray JD. A comparative gene map of the horse (*Equus caballus*). Genome Res. 1999; 9:1239–1249. [PubMed: 10613847]
5. Shiue YL, Bickel LA, Caetano AR, Millon LV, Clark RS, Eggleston ML, Michelmore R, Bailey E, Guerin G, Godard S, Mickelson JR, Valberg SJ, Murray JD, Bowling AT. A synteny map of the horse genome comprised of 240 microsatellite and RAPD markers. Anim. Genet. 1999; 30:1–9. [PubMed: 10050277]
6. Lindgren G, Sandberg K, Persson H, Marklund S, Breen M, Sandgren B, Carlsten J, Ellegren H. A primary male autosomal linkage map of the horse genome. Genome Res. 1998; 8:951–966. [PubMed: 9750194]
7. Guerin G, Bailey E, Bernoco D, Anderson I, Antczak DF, Bell K, Binns MM, Bowling AT, Brandon R, Cholewinski G, Cothran EG, Ellegren H, Forster M, Godard S, Horin P, Ketchum M, Lindgren G, McPartlan H, Meriaux JC, Mickelson JR, Millon LV, Murray J, Neau A, Roed K, Sandberg K, Shiue Y-L, Skow LC, Stott M, Swinburne J, Valberg SJ, Van Haeringen H, Van Haeringen WA, Ziegle J. Report of the International Equine Gene Mapping Workshop: male linkage map. Anim. Genet. 1999; 30:341–354. [PubMed: 10582279]
8. Guerin G, Bailey E, Bernoco D, Anderson I, Antczak DF, Bell K, Biros I, Bjornstad G, Bowling AT, Brandon R, Caetano AR, Cholewinski G, Colling D, Eggleston M, Ellis N, Flynn J, Gralak B, Hasegawa T, Ketchum M, Lindgren G, Lyons LA, Millon LV, Mariat D, Murray J, Neau A, Roed K, Sandberg K, Skow LC, Tammen I, Tozaki T, Van Dyk E, Weiss B, Young A, Ziegle J. The second generation of the International Equine Gene Mapping Workshop half-sibling linkage map. Anim. Genet. 2003; 34:161–168. [PubMed: 12755815]
9. Swinburne J, Gerstenberg C, Breen M, Aldridge V, Lockhart L, Marti E, Antczak D, Eggleston-Stott M, Bailey E, Mickelson J, Roed K, Lindgren G, von Haeringen W, Guerin G, Bjarnason J, Allen T, Binns M. First comprehensive low-density horse linkage map based on two 3-generation, full-sibling, cross-bred horse reference families. Genomics. 2000; 66:123–134. [PubMed: 10860657]
10. Raudsepp T, Kijas J, Godard S, Guerin G, Andersson L, Chowdhary BP. Comparison of horse chromosome 3 with donkey and human chromosomes by cross-species painting and heterologous FISH mapping. Mamm. Genome. 1999; 10:277–282. [PubMed: 10051324]
11. Godard S, Vaiman A, Schibler L, Mariat D, Vaiman D, Cribiu EP, Guerin G. Cytogenetic localization of 44 new coding sequences in the horse. Mamm. Genome. 2000; 11:1093–1097. [PubMed: 11130977]
12. Lear TL, Brandon R, Piumi F, Terry RR, Guerin G, Thomas S, Bailey E. Mapping of 31 horse genes in BACs by FISH. Chromosome Res. 2001; 9:261–262. [PubMed: 11330401]
13. Mariat D, Oustry-Vaiman A, Cribiu EP, Raudsepp T, Chowdhary BP, Guerin G. Isolation, characterization and FISH assignments of horse BAC clones containing type I and II markers. Cytogenet. Cell Genet. 2001; 92:144–148. [PubMed: 11306814]
14. Milenkovic D, Oustry-Vaiman A, Lear TL, Billault A, Mariat D, Piumi F, Schibler L, Cribiu E, Guerin G. Cytogenetic localization of 136 genes in the horse: comparative mapping with the human genome. Mamm. Genome. 2002; 13:524–534. [PubMed: 12370783]
15. Chowdhary BP, Raudsepp T, Kata SR, Goh G, Millon LV, Allan V, Piumi F, Guerin G, Swinburne J, Binns M, Lear TL, Mickelson J, Murray J, Antczak DF, Womack JE, Skow LC. The first-generation whole-genome radiation hybrid map in the horse identifies conserved segments in human and mouse genomes. Genome Res. 2003; 13:742–751. [PubMed: 12671008]

16. Raudsepp T, Gustafson-Seabury A, Durkin K, Wagner ML, Goh G, Seabury CM, Brinkmeyer-Langford C, Lee EJ, Agarwala R, Stallknecht-Rice E, Schaffer AA, Skow LC, Tozaki T, Yasue H, Penedo MC, Lyons LA, Khazanehdari KA, Binns MM, MacLeod JN, Distl O, Guerin G, Leeb T, Mickelson JR, Chowdhary BP. A 4,103 marker integrated physical and comparative map of the horse genome. Cytogenet. Genome Res. 2008; 122:28–36. [PubMed: 18931483]

17. Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, Lear TL, Adelson DL, Bailey E, Bellone RR, Blocker H, Distl O, Edgar RC, Garber M, Leeb T, Mauceli E, MacLeod JN, Penedo MC, Raison JM, Sharpe T, Vogel J, Andersson L, Antczak DF, Biagi T, Binns MM, Chowdhary BP, Coleman SJ, Della Valle G, Fryc S, Guerin G, Hasegawa T, Hill EW, Jurka J, Kiialainen A, Lindgren G, Liu J, Magnani E, Mickelson JR, Murray J, Nergadze SG, Onofrio R, Pedroni S, Piras MF, Raudsepp T, Rocchi M, Roed KH, Ryder OA, Searle S, Skow L, Swinburne JE, Syvanen AC, Tozaki T, Valberg SJ, Vaudin M, White JR, Zody MC, Broad Institute Genome Sequencing Platform; Broad Institute Whole Genome Assembly Team. Lander ES, Lindblad-Toh K. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science. 2009; 326:865–867. [PubMed: 19892987]

18. Leeb T, Vogl C, Zhu B, de Jong PJ, Binns MM, Chowdhary BP, Scharfe M, Jarek M, Nordsiek G, Schrader F, Blocker H. A human-horse comparative map based on equine BAC end sequences. Genomics. 2006; 87:772–776. [PubMed: 16603334]

19. Penedo MC, Millon LV, Bernoco D, Bailey E, Binns M, Cholewinski G, Ellis N, Flynn J, Gralak B, Guthrie A, Hasegawa T, Lindgren G, Lyons LA, Roed KH, Swinburne JE, Tozaki T. International Equine Gene Mapping Workshop Report: a comprehensive linkage map constructed with data from new markers and by merging four mapping resources. Cytogenet. Genome Res. 2005; 111:5–15. [PubMed: 16093715]

20. Swinburne JE, Boursnell M, Hill G, Pettitt L, Allen T, Chowdhary B, Hasegawa T, Kurosawa M, Leeb T, Mashima S, Mickelson JR, Raudsepp T, Tozaki T, Binns M. Single linkage group per chromosome genetic linkage map for the horse, based on two three-generation, full-sibling, crossbred horse reference families. Genomics. 2006; 87:1–29. [PubMed: 16314071]

21. Doan R, Cohen ND, Sawyer J, Ghaffari N, Johnson CD, Dindot SV. Whole-genome sequencing and genetic variant analysis of a Quarter Horse mare. BMC Genomics. 2012; 13:78. [PubMed: 22340285]

22. Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, Schubert M, Cappellini E, Petersen B, Moltke I, Johnson PL, Fumagalli M, Vilstrup JT, Raghavan M, Korneliussen T, Malaspinas AS, Vogt J, Szklarczyk D, Kelstrup CD, Vinther J, Dolocan A, Stenderup J, Velazquez AM, Cahill J, Rasmussen M, Wang X, Min J, Zazula GD, Seguin-Orlando A, Mortensen C, Magnussen K, Thompson JF, Weinstock J, Gregersen K, Roed KH, Eisenmann V, Rubin CJ, Miller DC, Antczak DF, Bertelsen MF, Brunak S, Al-Rasheid KA, Ryder O, Andersson L, Mundy J, Krogh A, Gilbert MT, Kjaer K, Sicheritz-Ponten T, Jensen LJ, Olsen JV, Hofreiter M, Nielsen R, Shapiro B, Wang J, Willerslev E. Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. Nature. 2013; 499:74–78. [PubMed: 23803765]

23. Orlando L, Male D, Alberdi MT, Prado JL, Prieto A, Cooper A, Hanni C. Ancient DNA clarifies the evolutionary history of American Late Pleistocene equids. J. Mol. Evol. 2008; 66:533–538. [PubMed: 18398561]

24. Orlando L, Metcalf JL, Alberdi MT, Telles-Antunes M, Bonjean D, Otte M, Martin F, Eisenmann V, Mashkour M, Morello F, Prado JL, Salas-Gismondi R, Shockey BJ, Wrinn PJ, Vasil'ev SK, Ovodov ND, Cherry MI, Hopwood B, Male D, Austin JJ, Hanni C, Cooper A. Revising the recent evolutionary history of equids using ancient DNA. Proc. Natl Acad. Sci. U.S.A. 2009; 106:21754–21759. [PubMed: 20007379]

25. Orlando L, Ginolhac A, Raghavan M, Vilstrup J, Rasmussen M, Magnussen K, Steinmann KE, Kapranov P, Thompson JF, Zazula G, Froese D, Moltke I, Shapiro B, Hofreiter M, Al-Rasheid KA, Gilbert MT, Willerslev E. True single-molecule DNA sequencing of a pleistocene horse bone. Genome Res. 2011; 21:1705–1719. [PubMed: 21803858]

26. McCue ME, Bannasch DL, Petersen JL, Gurr J, Bailey E, Binns MM, Distl O, Guerin G, Hasegawa T, Hill EW, Leeb T, Lindgren G, Penedo MC, Roed KH, Ryder OA, Swinburne JE, Tozaki T, Valberg SJ, Vaudin M, Lindblad-Toh K, Wade CM, Mickelson JR. A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. PLoS Genet. 2012; 8:e1002451. [PubMed: 22253606]

27. Petersen JL, Mickelson JR, Rendahl AK, Valberg SJ, Andersson LS, Axelsson J, Bailey E, Bannasch D, Binns MM, Borges AS, Brama P, da Camara Machado A, Capomaccio S, Cappelli K, Cothran EG, Distl O, Fox-Clipsham L, Graves KT, Guerin G, Haase B, Hasegawa T, Hemmann K, Hill EW, Leeb T, Lindgren G, Lohi H, Lopes MS, McGivney BA, Mikko S, Orr N, Penedo MC, Piercy RJ, Raekallio M, Rieder S, Roed KH, Swinburne J, Tozaki T, Vaudin M, Wade CM, McCue ME. Genome-wide analysis reveals selection for important traits in domestic horse breeds. PLoS Genet. 2013; 9:e1003211. [PubMed: 23349635]
28. Coleman SJ, Zeng Z, Wang K, Luo S, Khrebtukova I, Mienaltowski MJ, Schroth GP, Liu J, MacLeod JN. Structural annotation of equine protein-coding genes determined by mRNA sequencing. Anim. Genet. 2010; 41(Suppl. 2):121–130. [PubMed: 21070285]
29. Brooks SA, Gabreski N, Miller D, Brisbin A, Brown HE, Streeter C, Mezey J, Cook D, Antczak DF. Whole-genome SNP association in the horse: identification of a deletion in myosin Va responsible for Lavender Foal Syndrome. PLoS Genet. 2010; 6:e1000909. [PubMed: 20419149]
30. Fox-Clipsham LY, Carter SD, Goodhead I, Hall N, Knottenbelt DC, May PD, Ollier WE, Swinburne JE. Identification of a mutation associated with fatal foal immunodeficiency syndrome in the Fell and Dales pony. PLoS Genet. 2011; 7:e1002133. [PubMed: 21750681]
31. Andersson LS, Larhammar M, Memic F, Wootz H, Schwochow D, Rubin CJ, Patra K, Arnason T, Wellbring L, Hjalm G, Imsland F, Petersen JL, McCue ME, Mickelson JR, Cothran G, Ahituv N, Roepstorff L, Mikko S, Vallstedt A, Lindgren G, Andersson L, Kullander K. Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. Nature. 2012; 488:642–646. [PubMed: 22932389]
32. Corbin LJ, Blott SC, Swinburne JE, Sibbons C, Fox-Clipsham LY, Helwegen M, Parkin TD, Newton JR, Bramlage LR, McIlwraith CW, Bishop SC, Woolliams JA, Vaudin M. A genome-wide association study of osteochondritis dissecans in the Thoroughbred. Mamm. Genome. 2012; 23:294–303. [PubMed: 22052004]
33. Dupuis MC, Zhang Z, Druet T, Denoix JM, Charlier C, Lekeux P, Georges M. Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. Mamm. Genome. 2011; 22:613–620. [PubMed: 21698472]
34. Makvandi-Nejad S, Hoffman GE, Allen JJ, Chu E, Gu E, Chandler AM, Loredo AI, Bellone RR, Mezey JG, Brooks SA, Sutter NB. Four loci explain 83% of size variation in the horse. PLoS One. 2012; 7:e39929. [PubMed: 22808074]
35. Metzger J, Ohnesorge B, Distl O. Genome-wide linkage and association analysis identifies major gene loci for guttural pouch tympany in Arabian and German warmblood horses. PLoS One. 2012; 7:e41640. [PubMed: 22848553]
36. Kulbrock M, Lehner S, Metzger J, Ohnesorge B, Distl O. A genome-wide association study identifies risk loci to equine recurrent uveitis in German warmblood horses. PLoS One. 2013; 8:e71619. [PubMed: 23977091]
37. Schurink A, Wolc A, Ducro BJ, Frankena K, Garrick DJ, Dekkers JC, van Arendonk JA. Genome-wide association study of insect bite hypersensitivity in two horse populations in the Netherlands. Genet. Sel. Evol. 2012; 44:31. [PubMed: 23110538]
38. Nagaraj SH, Gasser RB, Ranganathan S. A hitchhiker's guide to expressed sequence tag (EST) analysis. Brief. Bioinform. 2007; 8:6–21. [PubMed: 16772268]
39. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science. 1995; 270:484–487. [PubMed: 7570003]
40. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995; 270:467–470. [PubMed: 7569999]
41. Murphy D. Gene expression studies using microarrays: principles, problems, and prospects. Adv. Physiol. Educ. 2002; 26:256–270. [PubMed: 12443997]
42. Noschka E, Vandenplas ML, Hurley DJ, Moore JN. Temporal aspects of laminar gene expression during the developmental stages of equine laminitis. Vet. Immunol. Immunopathol. 2009; 129:242–253. [PubMed: 19128842]
43. Mienaltowski MJ, Huang L, Frisbie DD, McIlwraith CW, Stromberg AJ, Bathke AC, Macleod JN. Transcriptional profiling differences for articular cartilage and repair tissue in equine joint surface lesions. BMC Med. Genomics. 2009; 2:60. [PubMed: 19751507]

44. Brosnahan MM, Miller DC, Adams M, Antczak DF. IL-22 is expressed by the invasive trophoblast of the equine (Equus caballus) chorionic girdle. J. Immunol. 2012; 188:4181–4187. [PubMed: 22490443]
45. Coleman SJ, Zeng Z, Liu J, MacLeod J. Analysis of equine protein-coding gene stucture and expression by RNA-sequencing. BMC Bioinformatics. 2010; 11:O8.
46. Wang X, Miller DC, Clark AG, Antczak DF. Random X inactivation in the mule and horse placenta. Genome Res. 2012; 22:1855–1863. [PubMed: 22645258]
47. Wang X, Miller DC, Harman R, Antczak DF, Clark AG. Paternally expressed genes predominate in the placenta. Proc. Natl Acad. Sci. U.S.A. 2013; 110:10705–10710. [PubMed: 23754418]
48. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. Cell. 2013; 155:27–38. [PubMed: 24074859]
49. Seyednasrollah F, Laiho A, Elo LL. Comparison of software packages for detecting differential expression in RNA-seq studies. Brief. Bioinform. 2013 Epub ahead of print; doi: 10.1093/bib/bbt086.
50. Auer PL, Doerge RW. Statistical design and analysis of RNA sequencing data. Genetics. 2010; 185:405–416. [PubMed: 20439781]
51. Rudolph JA, Spier SJ, Byrns G, Rojas CV, Bernoco D, Hoffman EP. Periodic paralysis in quarter horses: a sodium channel mutation disseminated by selective breeding. Nat. Genet. 1992; 2:144–147. [PubMed: 1338908]
52. Shin EK, Perryman LE, Meek K. A kinase-negative mutation of DNA-PK(CS) in equine SCID results in defective coding and signal joint formation. J. Immunol. 1997; 158:3565–3569. [PubMed: 9103416]
53. Tryon RC, White SD, Bannasch DL. Homozygosity mapping approach identifies a missense mutation in equine cyclophilin B (PPIB) associated with HERDA in the American Quarter Horse. Genomics. 2007; 90:93–102. [PubMed: 17498917]
54. McCue ME, Valberg SJ, Miller MB, Wade C, DiMauro S, Akman HO, Mickelson JR. Glycogen synthase (GYS1) mutation causes a novel skeletal muscle glycogenosis. Genomics. 2008; 91:458–466. [PubMed: 18358695]
55. Stock KF, Distl O. Multiple-trait selection for radiographic health of the limbs, conformation and performance in Warmblood riding horses. Animal. 2008; 2:1724–1732. [PubMed: 22444077]
56. Schroder W, Klostermann A, Stock KF, Distl O. A genome-wide association study for quantitative trait loci of show-jumping in Hanoverian warmblood horses. Anim. Genet. 2012; 43:392–400. [PubMed: 22497689]
57. Viklund A, Nasholm A, Strandberg E, Philipsson J. Effects of long-time series of data on genetic evaluations for performance of Swedish Warmblood riding horses. Animal. 2010; 4:1823–1831. [PubMed: 22445143]
58. Gu J, Orr N, Park SD, Katz LM, Sulimova G, MacHugh DE, Hill EW. A genome scan for positive selection in thoroughbred horses. PLoS One. 2009; 4:e5767. [PubMed: 19503617]
59. Tozaki T, Miyake T, Kakoi H, Gawahara H, Sugita S, Hasegawa T, Ishida N, Hirota K, Nakano Y. A genome-wide association study for racing performances in Thoroughbreds clarifies a candidate region near the MSTN gene. Anim. Genet. 2010; 41(Suppl. 2):28–35. [PubMed: 21070273]
60. Hill EW, McGivney BA, Gu J, Whiston R, Machugh DE. A genome-wide SNP-association study confirms a sequence variant (g.66493737C>T) in the equine myostatin (MSTN) gene as the most powerful predictor of optimum racing distance for Thoroughbred racehorses. BMC Genomics. 2010; 11:552. [PubMed: 20932346]
61. Stock KF, Distl O. Genetic correlations between conformation traits and radiographic findings in the limbs of German Warmblood riding horses. Genet. Sel. Evol. 2006; 38:657–671. [PubMed: 17129565]
62. Stock KF, Distl O. Genetic correlations between performance traits and radiographic findings in the limbs of German Warmblood riding horses. J. Anim. Sci. 2007; 85:31–41. [PubMed: 17179537]
63. Viklund A, Braam A, Nasholm A, Strandberg E, Philipsson J. Genetic variation in competition traits at different ages and time periods and correlations with traits at field tests of 4-year-old Swedish Warmblood horses. Animal. 2010; 4:682–691. [PubMed: 22444120]

64. Hill EW, Gu J, Eivers SS, Fonseca RG, McGivney BA, Govindarajan P, Orr N, Katz LM, MacHugh DE. A sequence polymorphism in MSTN predicts sprinting ability and racing stamina in thoroughbred horses. PLoS One. 2010; 5:e8645. [PubMed: 20098749]

65. Binns MM, Boehler DA, Lambert DH. Identification of the myostatin locus (MSTN) as having a major effect on optimum racing distance in the Thoroughbred horse in the USA. Anim. Genet. 2010; 41(Suppl. 2):154–158. [PubMed: 21070290]

66. Gu J, MacHugh DE, McGivney BA, Park SD, Katz LM, Hill EW. Association of sequence variants in CKM (creatine kinase, muscle) and COX4I2 (cytochrome c oxidase, subunit 4, isoform 2) genes with racing performance in Thoroughbred horses. Equine Vet. J. 2010; 42(Suppl. s38): 569–575.

67. Hill EW, Gu J, McGivney BA, MacHugh DE. Targets of selection in the Thoroughbred genome contain exercise-relevant gene SNPs associated with elite racecourse performance. Anim. Genet. 2010; 41(Suppl. 2):56–63. [PubMed: 21070277]

68. Promerova M, Andersson LS, Juras R, Penedo MC, Reissmann M, Tozaki T, Bellone R, Dunner S, Horin P, Imsland F, Imsland P, Mikko S, Modry D, Roed KH, Schwochow D, Vega-Pla JL, Mehrabani-Yeganeh H, Yousefi-Mashouf N, Cothran EG, Lindgren G, Andersson L. Worldwide frequency distribution of the 'Gait keeper' mutation in the DMRT3 gene. Anim. Genet. 2014; 45:274–282. [PubMed: 24444049]