The Impact of Myostatin Genetic Polymorphism on Muscle Conformation in the Horse

Honors Thesis
Presented to the College of Agriculture and Life Sciences, Animal Science
of Cornell University
in Partial Fulfillment of the Requirements for the
Research Honors Program

by
Nicole Elizabeth SanGiacomo
May 2013
Dr. Samantha Brooks

ABSTRACT

In this study, we tested the recent hypothesis that a previously discovered polymorphism in the myostatin *(MSTN)* gene could contribute to muscle mass in the Thoroughbred racehorse. The relationship between skeletal and muscle measurements was quantified and compared to the genotypic analysis of myostatin in the horse.

A sample of 101 Thoroughbred racehorses, from a variety of locations, was assessed using a previously published set of 35 body measurements. An additional 8 Thoroughbred horses were examined using a recently developed set of muscle measures in order to better quantify the muscle mass of the animals. In addition, photographs were taken of each individual horse (n=109), in order to assess the muscle conformation. Hair samples were also collected for determination of *MSTN* genotype. Photographs and *MSTN* genotype were then assessed and compared, prior to statistical analysis, in order to determine the existence and correlation between muscle conformation and the *MSTN* genotype. Skewed measures, and failed genotypic analyses were eliminated from statistical analysis, resulting in a total of 86 individuals.

This study demonstrated that there is no direct correlation among conformation, body measurements, muscle measurements and *MSTN* genotype. Horses with a polymorphism in *MSTN* do not exhibit phenotypic differences of the magnitude seen in *MSTN* alleles of cattle and dogs. Subtle changes may be detected in future work with an expanded set of measures, larger population of horses, or by controlling environmental effects due to diet and exercise.

The *MSTN* locus is currently used as a commercial marker of performance in the Thoroughbred although the physiological mechanism of its action is unknown. This research has eliminated the possibility of gross effects on body size and conformation, though subtler changes in muscle mass may exist. We expect future studies to further this exploration in

hopes of determining why the *MSTN* polymorphism is not associated with altered body size or conformation in the equine species.

ACKNOWLEDGEMENTS

The author is extremely grateful for all of the assistance, love and endless support she received throughout the completion of this thesis. She would first like to personally express her deepest appreciation to her supervisor and mentor, Dr. Samantha Brooks, for the patient encouragement, guidance and advice she has provided throughout her time as a Cornell University student. The experience gained as a member of Dr. Brooks' laboratory has provided the author with valuable opportunities. Without her dedication to the efforts of her students and science, this thesis would not have been possible. The author would also like to extend her gratitude to the graduate and undergraduate students of *The Brooks Equine Genetics Laboratory* at Cornell University for their continued assistance, suggestions and friendship.

Second, the author is grateful for the assistance of Elena Guilotto, for sharing not only her knowledge and predictions regarding the *MSTN* polymorphism in the horse, but for sharing the primers and primer sequences used in this study. Similarly, the author would like to thank Dr. Nathan Sutter of The Cornell University College of Veterinary Medicine and his laboratory, for assistance in sample and data collection.

In addition, the author would like to thank the College of Agriculture and Life Sciences Charitable Trust Program at Cornell University, for expressing interest in this topic. Without their funds, this project would not have been addressed, tackled, or confronted.

A very special thanks to her very best friends, who have been a consistent source of encouragement and love throughout her endeavors at Cornell University. Lastly, and most of all, the author wishes to extend the upmost love and appreciation to her beloved family for their understanding, support, and endless love throughout the duration of her studies.

TABLE OF CONTENTS

INTRODUCTION

The physical demands placed on the domesticated horse (*Equus ferus caballus*) require an optimal level of fitness. Despite advances in transportation and agriculture that have largely replaced traditional applications, horses continue to be used for police work, ranching, recreation and competitive sports. Skeletal and muscle conformation in the horse assist in the evaluation of the overall health, prospective performance and the selection of successful breeding characteristics for these animals.

In this study, we examine the impact of a myostatin (*MSTN)* genetic polymorphism on body composition in the horse. This polymorphism was previously identified as a 227 base pair insertion at chromosome 18: 66495327bp, comprising a SINE element labeled ERE-1 (Hill et. al, 2010). We tested the hypothesis that this polymorphism contributes to muscle mass. This was of interested because of its potential to affect performance in the Thoroughbred racehorse. Applications of this knowledge will improve the breeding, conformation, and ultimately the performance of the animal.

Several studies in canine and cattle species have indicated a correlation between myostatin polymorphisms and muscle size and definition (McPherron et. al, 1997 and Mosher et. al, 2007). Muscle composition has proven to be a significant determinant of breeding success, and therefore economic viability, in many livestock species (Casas et. al, 1999). If there is indeed a correlation between the MSTN gene and musculature in the horse, breeders could specifically select for muscle traits using genetic markers and polymorphisms in the gene. As a result, horse-owners could select for purchase or breeding only those animals that are likely to perform well based on specific genetic make-up as opposed to speculative bloodlines, and thus profit competitively and economically.

REVIEW OF THE LITERATURE

Recent scientific evidence suggests that myostatin (a member of the TGF-β superfamily, which is encoded by the *MSTN* gene) regulates skeletal muscle development in a range of mammalian species, including the horse **(**Hill et. al, 2011). The *MSTN* gene is composed of three exons and two introns. Myostatin protein is synthesized as a precursor, and upon proteolytic processing gives an N-terminal latency-associated peptide (myostatin propeptide or LAP-fragment) and a smaller mature peptide at the C-terminus (Dall'Olio et. al, 2010). Myostatin is produced primarily in skeletal muscle cells, circulates in the blood, and acts on muscle tissue by binding to a cell-bound receptor, activin type II. It regulates both the number and growth of muscle fibers, and inhibits muscle differentiation and growth (Hill et. al, 2011).

Scientists have determined that horses lacking the wild-type myostatin alleles, or those treated with substances such as follistatin that block the binding of myostatin to its receptor, possess significantly larger muscles (Cash et. al, 2009). Mutations in myostatin in several mammalian species have been associated with increased muscling and an increase in fast glycolytic type IIB fibers, which could increase power potential (Deveaux et. al, 2001). The *MSTN* gene, therefore, could be significant to the livestock business for the purpose of predicting performance, health, and success in competition.

MSTN gene variants contribute to muscle hypertrophy in a range of mammalian species, including several breeds of cattle. In Belgian Blue, Piedmontese, Marchigiana, and other cattle breeds, the loss of *MSTN* function through mutation within the coding sequence of the *MSTN* gene result in increased skeletal muscle mass, predominantly in the shoulders and thighs. This phenotype is known as "double-muscling" (McPherron et. al, 1997) (Figure 1).

Figure 1. Depiction of double-muscling in cattle according to McPherron et. al, 1997. There is an extreme difference in muscle density, particularly in the shoulders and thighs.

Double muscling originally arose as a natural mutation, and is related to an increase in the number of muscle fibers (Arthur, 1995). The myostatin coding sequence in Belgian Blue cattle has a deletion of nucleotides 937-947 in the third exon. This deletion causes a frame-shift, which results in a truncated protein that terminates 14 codons downstream of the site of the mutation (McPherron et. al, 1997). Cattle breeders specifically select for animals that are homozygous for this mutation, due to their increased musculature, a trait beneficial for beef production. However, the veterinary commissions of Sweden, Denmark and Finland have recently submitted a resolution that proposed a ban on the breeding of cattle for this trait, due to extraordinary biological factors that put these animals at risk. This gene becomes active during the embryonic stage of development in cattle, resulting in calves with excessive muscle development beginning in the womb. Thus, these calves can be 10-38% heavier than average. This often results in extremely difficult calving, and complicated cesarean deliveries (Bassett, 2009). In addition, macroglossia (an over-enlargement or swelling of the tongue), as well as prognathism (where the maxilla is shorter than the mandible) can occur in

calves with this myostatin mutation, interfering with the ability to suckle, and hence to survive (Radostits, 2007). Reduced development of respiratory muscles and the heart, are also typical phenotypic traits of double-muscling in cattle breeds, often resulting in insufficient oxygen intake, and sudden death (Gustin et. al, 1997).

A mutation in the *MSTN* gene, and therefore increased muscle mass, has also been proven to enhance racing performance in heterozygous racing canines. Dogs that contain a single copy of the mutated allele are more muscled than normal and are among the fastest dogs in racing events. Canines with this phenotype carry one copy of a two-base-pair deletion in the third exon of *MSTN,* which leads to a premature stop codon at amino acid 313 (Mosher et. al, 2007). However, dogs with two copies of the same mutation are grossly over-muscled. While cattle breeders specifically select for animals that are homozygous for mutations in *MSTN* because of their increased musculature, a trait beneficial for beef production, canine breeders find this trait undesirable, for they are too bulky for long distance races (Mosher et. al, 2007).

A sequence polymorphism near the equine myostatin gene *(MSTN)* locus has been shown to have a singular genomic influence on optimum race distance in elite Thoroughbred racehorses (Hill et. al, 2010). However, the few studies that have examined the *MSTN* gene thus far have only chosen to examine elite athletes (Hill et. al, 2010 and Tozaki et. al, 2011). As a result of the discovery and function of this gene and its locus, scientific studies have become more prominent, as the basis of equine health, breeding and performance is dependent on muscle composition and build (Hill et. al, 2010). Horses have a very high muscle mass to body weight ratio (55%), as compared to other mammalian species (33%-40%), making the equine genome ideal for selection for muscle strength phenotypes and functional variations that contribute to speed and stamina (Gunn, 1987).

The horse industry is a multi-billion dollar international enterprise that is actively engaged in the breeding, training, and performance of the horse. In the United Kingdom, the horse-racing industry is one of the nation's largest sports enterprises, generating more revenue and attendees than any other sport. In an assessment performed in 2009 by the consulting and analyzing firm Deloitte, the company claimed that in Great Britain alone, the sport grossed an average of 3.39 billion pounds (5.1 billion US dollars), in 2008 (Figure 2). The breeding industry alone was estimated as contributing 207 million pounds, or 310 million US dollars to the United Kingdom's economy in 2008 (Delloite, 2009).

Category/Contributor	**United Kingdom Pounds (£)**	**United States Dollars ($)**
Total Economic Impact	£3.39 billion	$ 5,199,921,000
Betting Industry's Gross	£1.05 billion	$ 1, 541,569,500
Race-goers Contribution	£361 million	$ 553,737,900
Media Contribution	£100 million	$ 153,390,000
Owner Gross	£367 million	$ 562,941,300
Owner Prize Money	£92 million	$ 141,118,800
Breeding	£207 million	$ 317,517,300

Figure 2. Economic contributions and thus the impact of Horse-Racing in the United Kingdom in 2008 according to Deloitte, accompanied by United States dollar conversions (Deloitte, 2009).

These surprising figures reflect the enormous economic impact the horse racing industry has on owners, breeders, trainers, riders, gamblers, spectators, venues, the media and the myriad employees associated with the enterprise, all dependent on the ability of the horse to perform at a high level in competition. To remain economically viable, a horse must be selected that will be successful on the track and as a sire to provide an adequate return on investment. The current technique of speculation based on lineage has proven to be just that: speculation, a genetic crap shoot. An ability to replace that speculation with science based on genetic predicting factors would be well received and potentially profitable in accurately selecting

animals with traits predisposing them to athleticism, gait and stature that would yield success in competition and breeding.

Different horse breeds present a variety of morphological phenotypes, and are therefore grouped into several classes and distinctions. Based on size, strength, and build, horse breeds are categorized into draught, light and pony. Skeletal structure, proportions, length, and muscle volume all reflect the goals and uses of the individual horse breeds. Using these proportions, horses are then categorized into brachymorphic, docilchomorphic and intermediate types (Dall'Olio et. al, 2010). From these categories, horse breeders and owners can select specific animals for their intended purpose. Brachymorphic horses are traditionally horses that are tall in stature, heavy boned, and very muscular. They are bred for strength and power, and their conformation is best suited to activities that involve strength and labor, such as pulling a carriage. Long bodies with long, thin muscles characterize Docilchomorphic horses. These horses are selected for sport purposes, such as distance running. Lastly, intermediate types are characterized by a lighter physical structure, but still powerful and compact with heavy muscling. These animals are typically used for trail or leisure riding (Dall'Olio et. al, 2010). As a result of the variance in skeletal and muscle composition within these categories, there is an intense selection for athletic prowess and phenotypes that enable superior equine performance (Hill et. al, 2010).

The majority of *MSTN* research has been performed in the docilchomorphic category of horses. The role of muscle mass development and polymorphism could be extremely beneficial in breeding, and thus racecourse performance, if owners and breeders could identify this polymorphism. Different traits and phenotypes are linked to two types of racecourses within five race distance categories. National Hunt Races involve the obstacle of hurdles or steeplechase fences over a distance of 4.5 miles. Flat races involve no obstacles, but are run over distances ranging form 5-20 furlongs (1 furlong is equivalent to 1/8 of a

mile). Within these two categories, the animals are divided into five distances: sprint/mile, intermediate, long and extended, with each distance becoming progressively longer. Horses, like humans, that are generally good sprinters, are suited to shorter distances and are characteristically more compact and muscular than horses suited to longer distance races, which are characteristically thinner, and more lean. Thus, sprinters are generally shorter, stockier animals with greater muscle mass than animals suited to endurance performance, and tend to mature earlier. A principal characteristic, specifically contributing to the ability of a racehorse to perform well in short distance sprint races, is the extent and maturity of the skeletal musculature (Hill et. al, 2010). Consequently, the analysis of the *MSTN* gene locus could be extremely beneficial for predicting the performance of animals and influencing breeding.

Variation in cardiovascular function contributing to aerobic capacity could perhaps distinguish individuals suited to shorter or longer distance races. Research suggests, however, that although environment, training, and metabolic adaptations may contribute to the race distance for which a horse is best suited; the genetic contribution to optimal performance at certain distances is immensely more significant. In the Australian racehorse, it has been estimated that the affects of environment and training are quite insignificant, with a p-value of 0.94 (Hill et. al, 2010).

Some of the first studies that began to assess the phenotypes associated with the *MSTN* gene were examined in canines. Mosher ***et. al*** examined this phenotypic phenomena in Whippets. He found that animals that were wild-type (+/+) had average muscles and were suited to middle to long distance races and performances. Whippets that were heterozygous for the *MSTN* mutant allele (mh/+) were notably more muscled, and better suited to middle distance races (Figure 3). Lastly, whippets that were homozygous for the mutant *MSTN*

allele possessed extremely dense muscles, and were far more successful at short, sprint races. These animals were far too bulky for longer competitions (Mosher et. al, 2007).

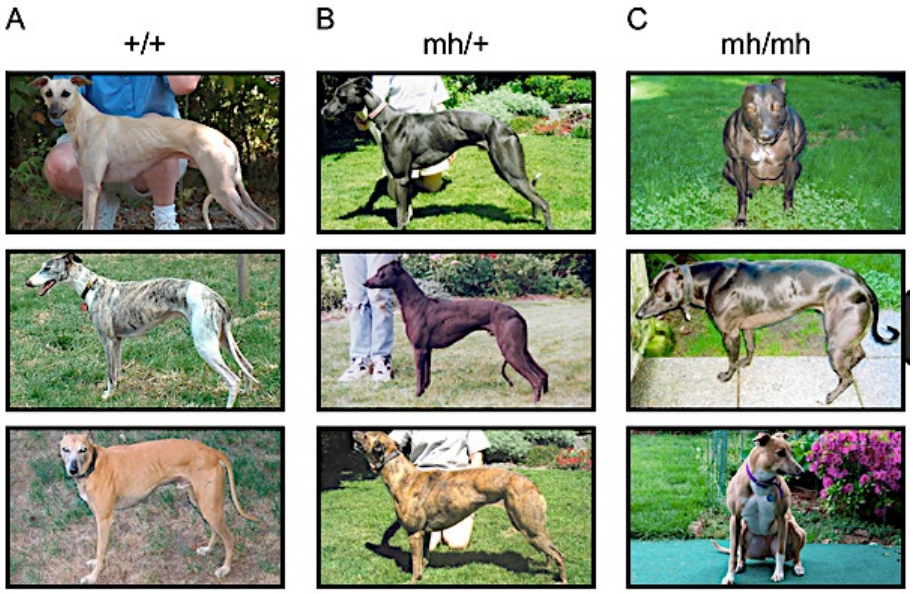


Figure 3. The above images were derived from a Mosher et. al, 2007 experimental protocol. (A) Dogs have two copies of the wild-type allele (+/+). (B) Dogs are heterozygous with one wild-type allele and one mutant (mh/+). (C) Dogs are homozygous for the mutant allele with two copies of the mutation (Mosher et. al, 2007).

Hosoyama and co-workers are credited for isolating and sequencing an *MSTN cDNA* from a Thoroughbred horse, while Caetano mapped this gene to equine chromosome (Caetano et. al, 1999 and Hosoyama et. al, 2002). Hill *et. al*., 2010 plotted the position of the SNPS sorted by chromosome and chromosome position using genome-wide association studies (GWAS), which are commonly used as an analysis of allelic association for genes throughout the genome. Hill indicated the most significant SNP on chromosome 18, therefore validating Caetano's finding (Hill et. al, 2010) (Figure 4).

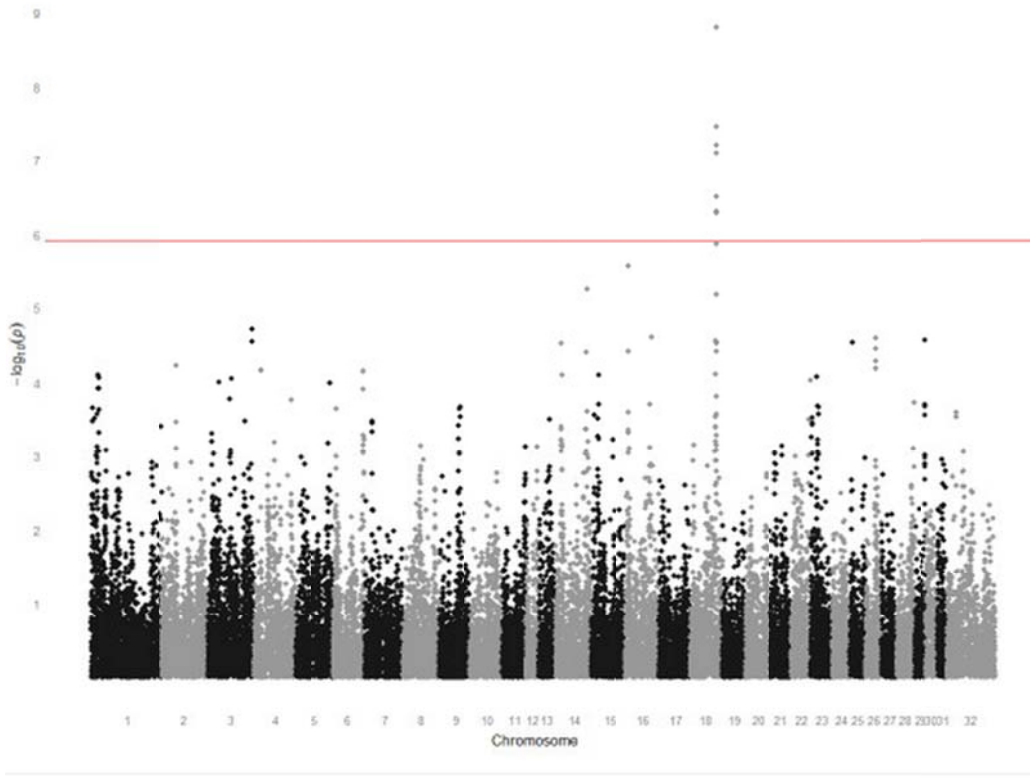


Figure 4. Manhattan plot indicating a quantitative trait GWAS. There is a peak on chromosome 18 in relation to myostatin, thus identifying its location and presence (Hill et. al, 2010.

In order to determine the relationship between the *MSTN* polymorphism and the overall performance of the animal, Hill (2011) then organized this *MSTN* gene into three specific categories: C:C, C:T, and T:T. Hill collected data from horses who possessed each of these genotypes, assessed their physical stature, and their most successful races. A C:C horse was more likely to be a fast, early maturing horse that performed well at a young age. They performed their best at 6.5 furlongs, and excelled in sprint races. These horses were identified as animals with thicker, denser musculature (Equinome Inc., E. Hill co-founder) (Figure 5).

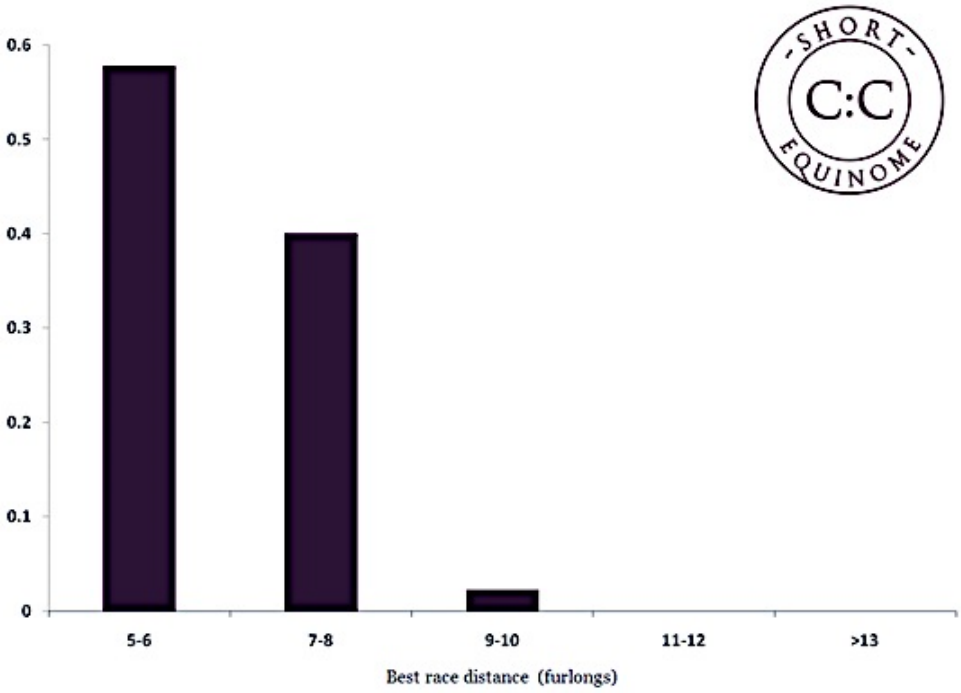


Figure 5. Proportion of Group Racing winning C:C horses that performed best at distances shown. According to this graph, C:C horses were best suited to distances less than 1 mile (Equniome Inc., E. Hill co-founder).

Additionally, Hill addressed the category of C:T horses. These animals performed best at an average distance of approximately 9.1 furlongs, and were depicted as fast, middle-distance types. Horses identified as C:T horses were more likely to have a mixture of speed and stamina (Equinome Inc., E. Hill co-founder) (Figure 6).

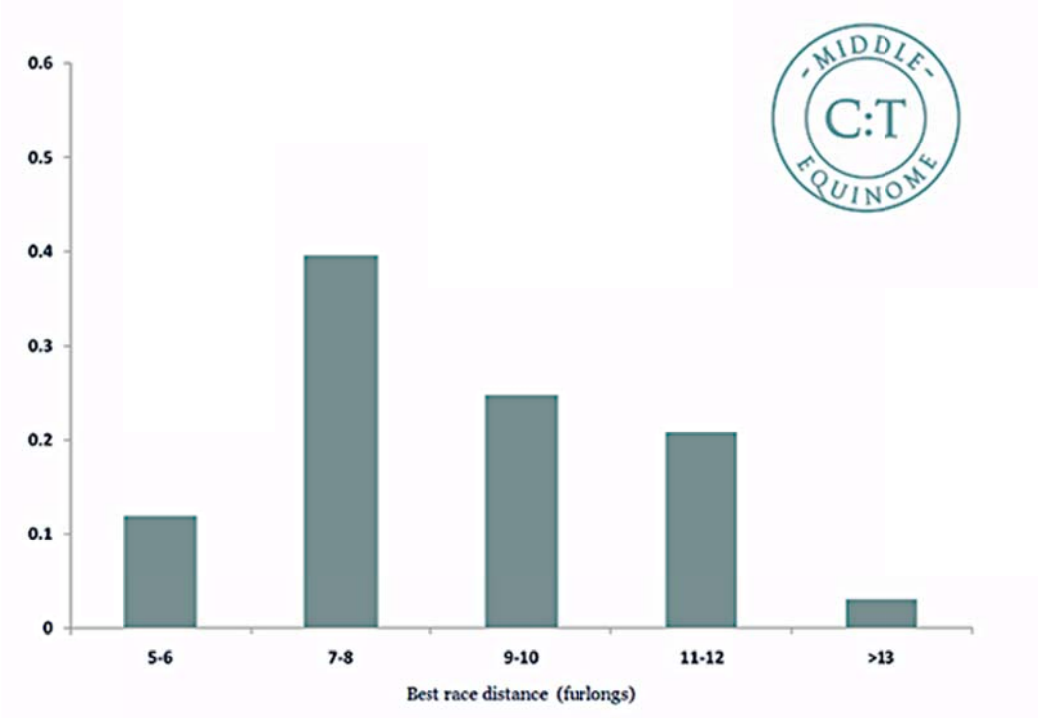


Figure 6. Proportion of Group Racing winning C:T horses that performed best at distances shown. According to this graph, C:T horses were best suited to distances of approximately 7-12 furlongs (Equniome Inc., E. Hill co-founder).

Lastly, T:T horses were described as animals that were best suited to races greater than one mile that require a great deal of stamina. These horses were known for maturing later in life, and were more suited to middle/long distance courses (Equinome Inc., E. Hill co-founder) (Figure 7).

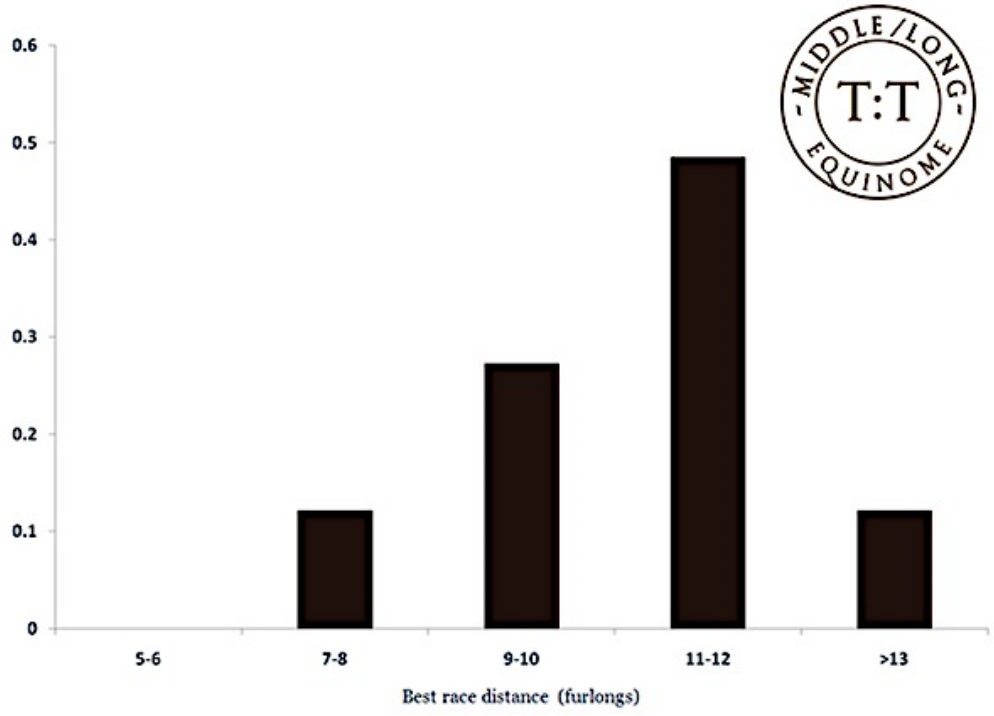


Figure 7. Proportion of Group Racing winning T:T horses that performed best at distances shown. According to this graph, T:T horses were best suited to greater distances that required stamina, such as middle to long distance (Equinome Inc., E. Hill co-founder).

Research has shown that in the Thoroughbred racehorse, the C-allele at the g.66493737C/T SNP has been found at significantly higher frequency in subgroups of the population that are suited to fast, short distance sprint races (Hill, et. al, 2011). Furthermore, horses with 2 copies of the *MSTN* gene variant were suited to fast/long distance races; horses with one copy of the *MSTN* variant were favorably competitive in middle-distance races, while those without a *MSTN* variant had greater stamina (Figures 2-4). Hill *et. al* (2011) and Tozaki *et. al* (2011) suggest that this variation in the *MSTN* genes could influence speed in the Thoroughbred horse.

Thoroughbred horses (n=189) that had won Graded Stakes Races were genotyped for *MSTN* polymorphisms and investigated for correlations with performance by Binns *et al.,* (2010). They observed that Thoroughbred horses that raced over sprint distances of 5-7

furlongs were often characterized by impressive hindquarter musculature, suggesting the association observed in ECA18 SNPs and optimum race distance could be mediated through an effect of the *MSTN* locus on musculature. However the Binns study did not separate effects due to genotypes and specialized training and exercise regimens for sprint races. This same *MSTN* locus on ECA18 was also statistically significant in a study performed by Tozaki *et. al* (2011), and was identified as a candidate region influencing race performance.

Tozaki *et. al* (2011) furthered this study using four SNPs, which were genotyped in 91 Thoroughbred horses in training to evaluate the association between genotype and body composition traits. These traits included body weight, withers height, chest circumference, cannon (3rd metacarpal) circumference, and body weight/withers height. These results were similar to the Hill *et al* (2010) study. Body weight/withers height showed a statistically significant difference at g.658604G>T, g.66493737C>T, and g.66539967A>G ($P<0.01$) only in males during the training period. Horses with a genotype associated with suitability for short-distance racing, C:C at g.66493737C>T, had the highest value (3.17 ± 0.05 kg·cm(-1)) for body weight/withers height, while those with a genotype associated with suitability for long-distance racing, T:T, had the lowest (2.99 ± 0.03 kg·cm(-1)). The trends in the association of body weight/withers height with genotypes for females were similar to those observed in males but did not reach statistical significance. These results suggest that the regulation of *MSTN* gene expression influences racing performance, particularly optimum race distance, in Thoroughbred horses (Tozaki et. al, 2011).

In summary, a number of studies have indicated that the *MSTN* gene could influence skeletal muscle mass and determine racing and overall performance. However, this polymorphism-influenced performance remains unproven. Therefore, the physiology of this polymorphism's effect on conformation of the horse is still widely unknown. Thus, the

identification of a conformational effect for this gene would benefit breeders to selectively produce animals that compete successfully at specific race distances.

MATERIALS AND METHODS

Part I. Animal Sampling and Measurements

The Cornell University Institutional Animal Care and Use Committee has approved all animal procedures (n=109) (experimental protocol #2008-0121). Thoroughbred horses of various ages (n=101) were surveyed for 35 external body-measurements, based on certain skeletal landmarks, in order to assess the muscle and skeletal composition of the animal (As published in Brooks *et. al,* 2005, Appendix A). Their names, dates of birth, purpose, injuries, health concerns, and registration numbers were also recorded. Each of these horses was registered under The Jockey Club USA (TJC) and was involved in various racing and exercise regimens. These animals were collected from a wide range of locations in the US. In addition, hair samples were removed from the tail of each animal in order to be used for further genetic sampling.

In order to accurately assess the association between the *MSTN* gene and the muscle conformation of the animal, an additional eight horses were assessed for 4 additional muscle measures. Because these areas are known to be composed of low adiposity and lean body mass, each of these measurements was assessed with the aim to determine the correlation between the genotype of the animal and their muscle density. The circumference of the neck was designed to target the brachiocephalic region, as well as the cervical part of trapezius muscle. Individuals were instructed to locate the middle of the neck and wrap the measuring tape around the midline of the neck (Figure 8). Additionally, the forearm circumference of the animal was examined to assess the radial carpal extensor muscle. This was performed by wrapping the tape around the top of the forearm, just below the shoulder and point of the elbow. Instructions were to have the tape parallel to the ground at the highest point of the forearm (Figure 8). Next, in order to assess the quality of musculature of the hindquarters of

the animal, we measured the stifle to dock region, which encompassed the lateral vastus muscle, as well as the middle and deep gluteal muscle. In order to assess the composition of muscle density from the stifle to the dock, we made certain that each horse was standing with equal weight distributed on all four limbs. We then palpated the patella, and extended the tape from this point to the dock, just above the tail (Figure 8). The gaskin midpoint was then examined in order to assess the medial gastrocnemius. Once more, in order to accurately assess this measurement, the animal was required to stand with equal weight distributed on all four limbs. The crest of the tibia was then palpated, while wrapping the tape around the gaskin to find the circumference of the region (Figure 8). Additionally, one individual was assigned to perform a visual subjective analysis of the overall variance in muscle density of each horse using a scale of "1" (low muscle density) to "5" (high muscle density).

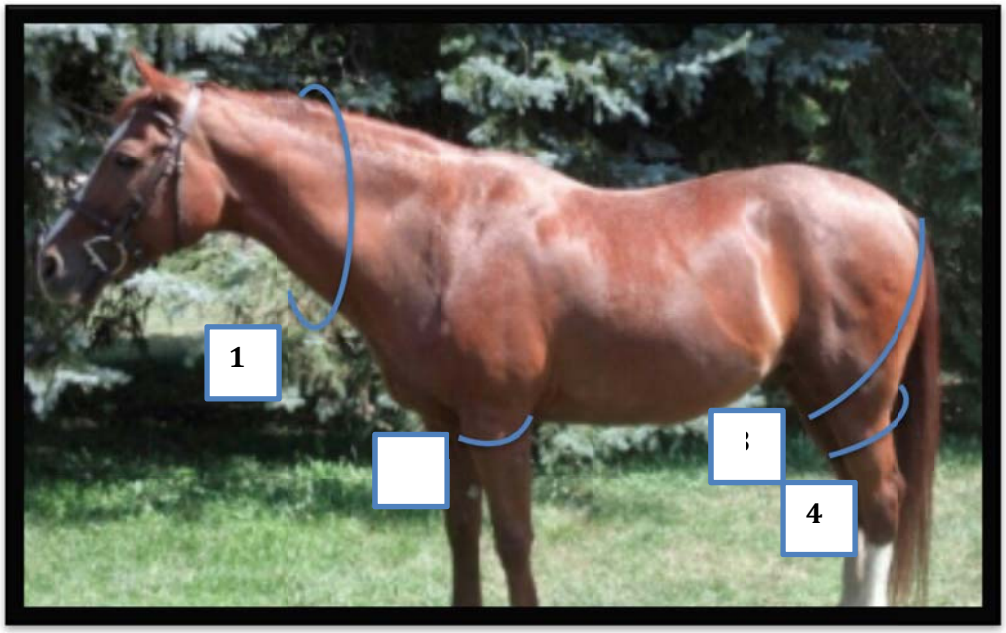


Figure 8. Body measurement of the (1) midline of the neck, (2) forearm circumference (3) stifle to dock (4) gaskin circumference to assess muscle composition in the Thoroughbred horse.

Part II. Polymerase Chain Reaction

Extracts of tail hair samples were used for genetic analysis. The hairs were lysed, in order to free DNA from cells clinging to the hair root (Appendix C). For each individual horse, 5-8 hairs were selected and incubated in 0.6mL tubes with lysis solution at 60° C for 45 minutes, followed by an additional 45 minutes at 95°C to ensure completion. Samples were stored at 4°C for the remainder of the experiment.

Hair DNA samples were used to assess the genotype of each animal using polymerase chain reaction (PCR). A region of the *MSTN* gene was amplified using specific oligonucleotide primer sequences shared pre-publication courtesy of Elena Guilotto: one forward *MSTN* primer (5' CAA ATG AAT CAG CTC ACC CTT 3') and one reverse *MSTN* primer (5' CTG AGA GAC AAC TTG CCA CA 3') (Elena Guilotto, personal communication, October 2012). In addition to these primers, PCR reagents from Roche Diagnostics GmbH including DNA Polymerase Buffer, 10X Concentration with MgCl2, PCR Grade Nucleotide Mix, Fast-Start Taq DNA Polymerase, and nuclease-free water were added to the polymerase chain reaction master-mix and distributed into individual 0.1mL tubes, according to the manufacturers recommendations (Roche Diagnostics GmbH). Lastly, 2µl of hair lysis template from each horse was added to each of the 0.1mL tubes.

The 7-step amplification protocol consisted of a 4 minute initiation step at 95°C, 30 seconds at 95°C, 63°C for 30 seconds, and 72°C at 30 seconds. This was followed by a repetition of 95°C at 30 seconds for 39 cycles, and concluded at 72°C for 7 minutes. The PCR product was then held at 4°C for preservation until examined by electrophoreses.

Part III. Agarose Gel Electrophoreses and Genotyping

PCR products were separated on 3% agarose gels in TBE Buffer (Appendix C). The size of each DNA fragment and thus the genotype of each Thoroughbred horse was

determined. For each gel, 3 grams of powdered agarose were added to 100mL of TBE buffer (Appendix C), and placed in a standard microwave for approximately 1.5 minutes, to ensure in the successful melting of OmniPur Calcibiochem Agarose. 5µl of e-Invitrogen Sybr Safe DNA gel stain was then added to the solution. When sufficiently cooled, the mixtures were placed in UV-transparent plastic electrophoresis trays with combs, and left for 25 minutes to harden.

Once the gel had set, the combs were removed and 1X TBE buffer (Appendix C) was added to both reservoirs in the tray, in order to cover the gel to a depth of approximately 2mm. Using the products from the polymerase chain reaction, 1.5µl of 6X concentration loading dye (New England BioLabs Inc.) was added to each of the PCR products in order to visualize the migration of the product throughout the electrophoresis process. The samples were then carefully loaded, accompanied by a 100bp ladder (New England BioLabs Inc.), in order to compare the bands under UV-fluorescence. The electrodes were then connected to each tray and were run for approximately 45 minutes at 150 volts.

After electrophoresis was complete, the size of each fragment, and thus the genotype of each horse were determined. Horses with one band at the 206 base-pair region were defined as homozygous wild type (+/+), while horses genotyped at 433 base-pairs were labeled as homozygous insertion (ins/ins), and animals that expressed both bands were heterozygous (+/ins). A visual of the regions amplified during this PCR can be seen in Figure 9. Due to failed genotyping, 20 horses were removed from the study.

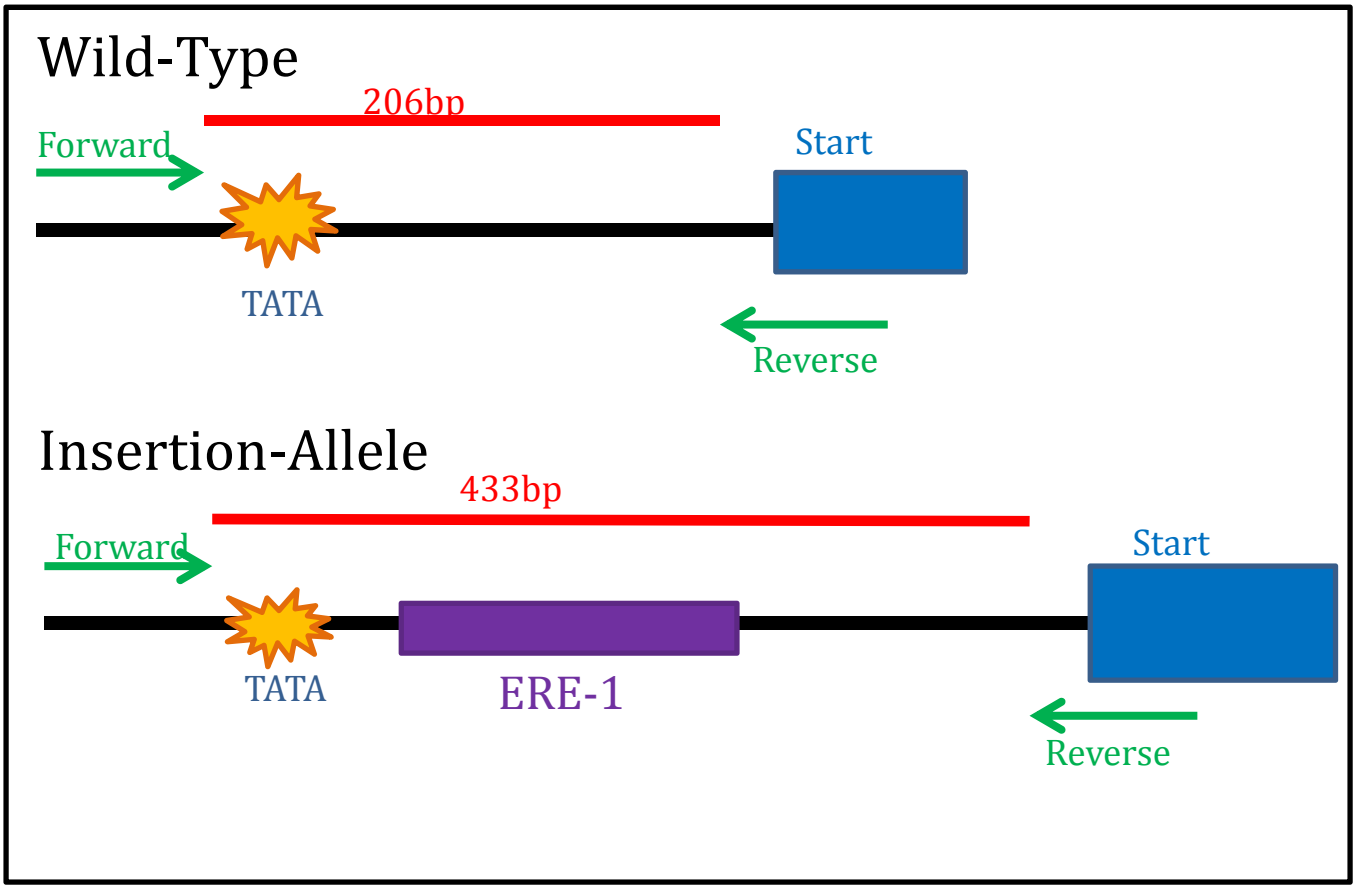


Figure 9. Illustration of the wild-type and insertion-allele expression at the promoter region. The wild-type PCR product is 206 base-pairs, while the ERE-1 insertion PCR product is 433 base-pairs. The ERE-1 insertion allele is located 146 base pairs upstream from the start of transcription, and 17 base pairs downstream from the TATA box.

Part IV. Elimination of Outlying Participants

Prior to analysis, an additional three horses were removed from the study due to significant outlying measurements, likely indicating errors during entry into the database. These errors were more than 4-times the standard deviation for each measure. As stated previously, 20 horses were eliminated from the study due to PCR amplification failures during genotyping. Thus, 23 horses in total were eliminated from this study, leaving a total of 86 horses to be examined by statistical analysis.

Part V. Principle Components Analysis

A Principle Components Analysis (PCA) on selected measures was formulated, similar to the method used in Brooks *et al*., 2005. As a result, four principle components were interpreted to analyze the percent of the variance among the muscle measures. Visual outlying inspection of the distribution of each individual measure was used to identify three outlying, and therefore likely erroneous measures for muzzle circumference, ear-length, and neck-base circumference. Thus, these measures were eliminated from the data due to the hypothesis that these measures were transcribed, measured, or recorded incorrectly, leaving 32 skeletal muscle measurements to be examined.

Part VI. Genotype vs. Measurements

Using the four principle components as traits, they were correlated with genotype through ANOVA. ANOVA tests were also used to examine the relationship between the 32 skeletal muscle measurements and *MSTN* genotypic variation. Similarly, ANOVA tests were used in order to assess the relationship between genotypic variation and the additional four muscle measurements. Multivariate assessments of age and gender as compared to muscle and skeletal muscle measurements were also analyzed using a least squares model.

RESULTS

Part I. *MSTN* Genotyping

Horses were characterized as homozygous wild-type (+/+), homozygous insertion (ins/ins), or heterozygous for the myostatin allele (+/ins) (Appendix B) (Figure 10). Of the 109 horses that were genotyped, 21 failed genotypic analysis and were thus excluded. Figure 11 indicates the frequency of genotypes assessed in the population.

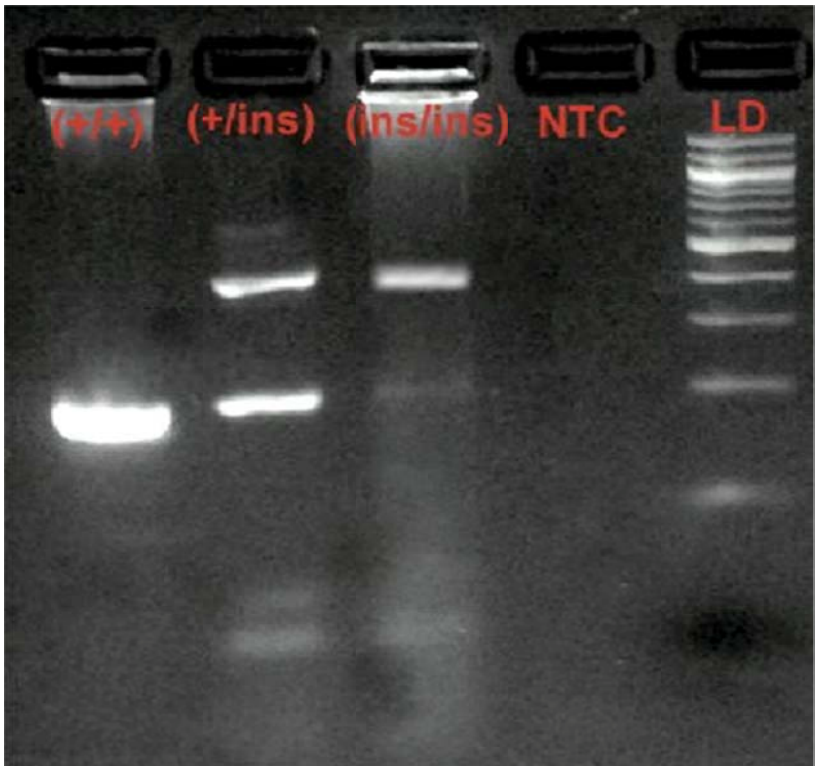


Figure 10. Visual representation of each genotype found in the Thoroughbred population. Samples that produced a single band at 206 base-pairs were homozygous wild-type (+/+), while those samples yielding a single 433 base-pair PCR product were homozygous for the insertion (ins/ins). Animals producing both bands were heterozygous for the *MSTN* insertion (+/ins). A no template control (NTC) 100 base-pair ladder (LD) were also included.

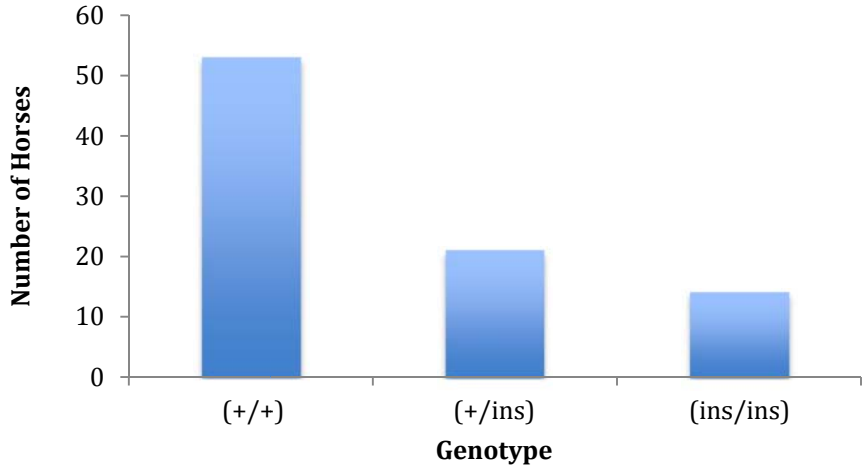


Figure 11. Frequencies of genotypes in 109 Thoroughbred racehorses. Homozygous WT (53), heterozygous (21), homozygous insertion (14). There were 21 horses from the study due to failed genotypic analysis.

Part II. Conformation as Judged in Photographs

Photographs of the horses were examined in order to assess variation in musculature. Using a visual subjective scale from "1" (low muscle density) to "5" (high muscle density), horses were assessed on their muscle conformation. This trait is difficult, however, to separate visually from body condition score (indicating adiposity). The scores were then compared to the *MSTN* genotype for each horse. Horses genotyped as ins/ins did not exhibit gross over-muscling, nor did they appear significantly different than those genotyped as heterozygous, and homozygous wild-type (Figure 12). Yet due to the qualitative nature of this approach, use of body measures was pursued as a more accurate method of conformation assessment.

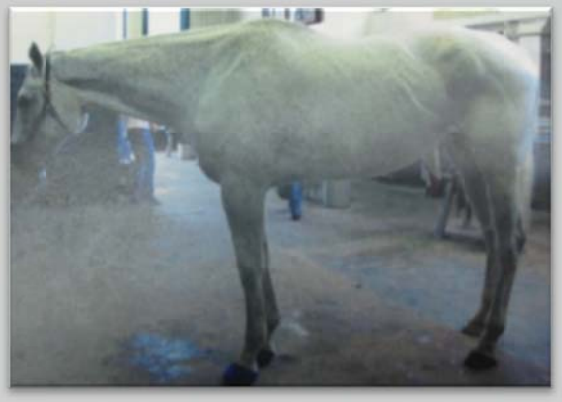
GENOTYPE	PHENOTYPE
HOMOZYGOUS WILD-TYPE (+/+)	
HETEROZYGOUS (+/ins)	
HOMOZYGOUS INSERTION (ins/ins)	

Figure 12. Three representative genotyped horses and their phenotypes demonstrate little to no difference in gross muscle conformation. (A) Homozygous wild-type (B) heterozygous (C) homozygous insertion.

Part III. Principle Components Analysis and Correlation with Genotype

Before statistical analysis, we examined the use of previously published body measurements for detection of variation in musculature. Based on factor loading, we interpret PC1 as body size, PC2 as overall body thickness, and PC3 and 4 as conformational traits likely related to variation in joint angularity. These four principal components were derived from these body measures to interpret the percent of variation and influence of different measures, in a method similar to Brooks *et. al*, 2005 (Figure 13). None of the four PC traits were associated with *MSTN* genotype as demonstrated by Figure 13.

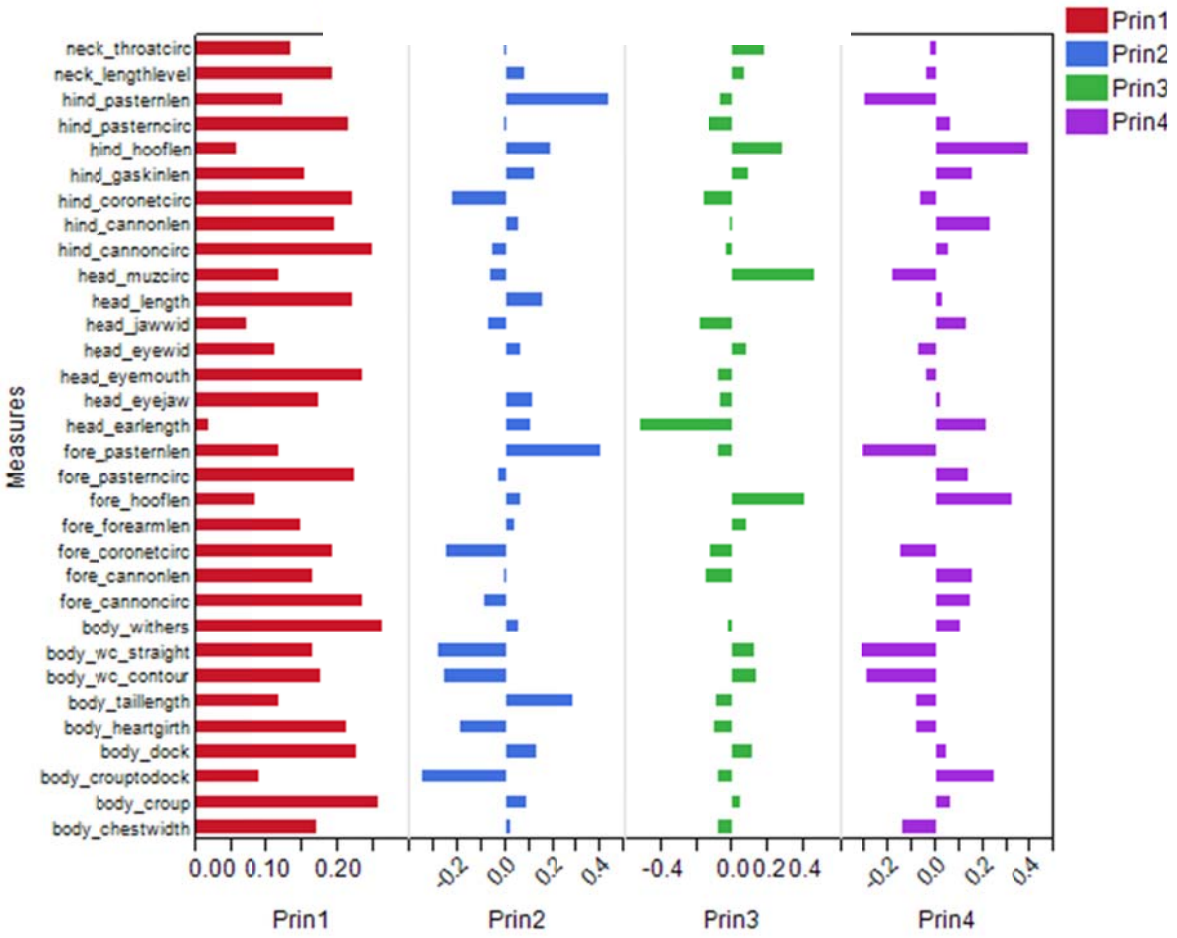


Figure 13. 32 muscle measurements as compared to the four principal components. Tail length, neck-length down and barrel girth were eliminated from principal component analysis as they were subject to frequent errors or impact due to environment (i.e. docking).

Part IV. Statistical Analysis for Correlation with Genotype

On statistical testing of each of the 35 previously published body measurements on the original 101 Thoroughbred horses, 23 horses were excluded due to outlying measures, and/or failed *MSTN* genotypes. Each measure was divided by the withers height of each horse in order to normalize for variation in overall body size. There was no significant difference between any of the 35 body measurements and their genotype in 78 horses. Figure 15 demonstrates the analysis of one of the 35 measures.

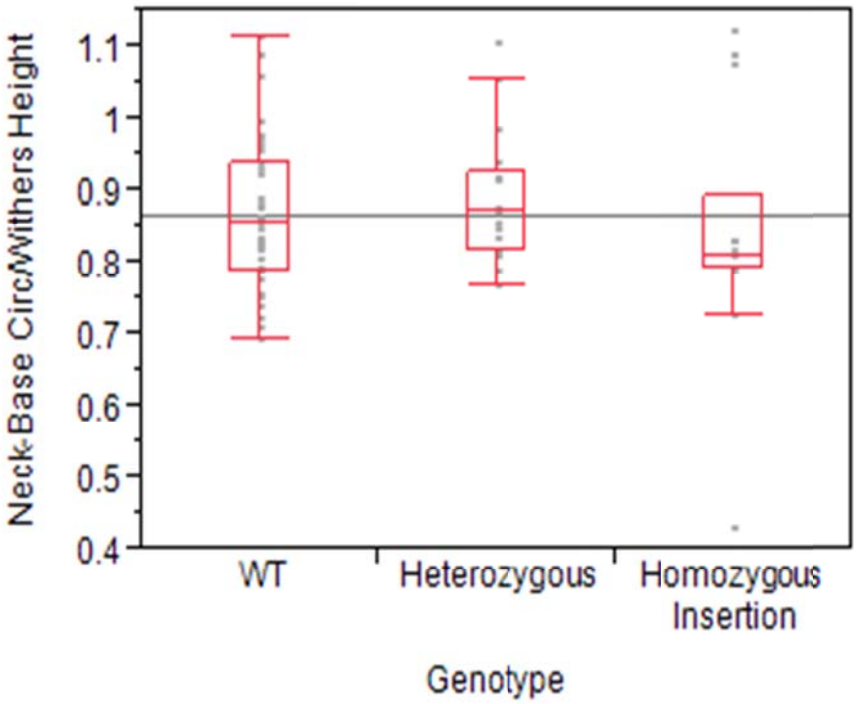


Figure 14. Oneway analysis of variance of neck-base circumference/withers height vs. genotype in original 78 Thoroughbred racehorses (p= 0.421).

We then analyzed the variance in the 35 body measures against the genotypes of the eight additional horses used for newly developed muscle measurements. These eight Thoroughbred racehorses were analyzed observing the four additional muscle measurements we added to the protocol. No significant difference resulted from the analyses of the four muscle measures as compared to genotype (Figure 17).

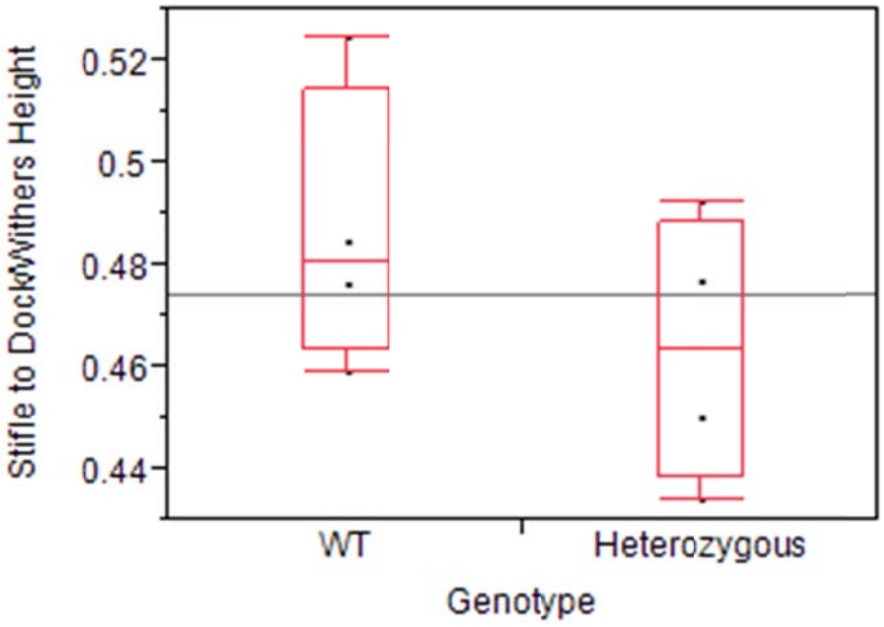


Figure 15. Oneway analysis of variance of stifle to dock/withers height vs. genotype in additional eight Thoroughbred racehorses (p= 0.4633).

Given the findings of Tozaki *et. al*, 2011, inclusion of sex as a covariate in statistical test could be important. Therefore, genotype and sex of the Thoroughbred racehorses were placed into a least-squares regression model, to fit the data to a model including these important variables. In the least-squares model of all 86 horses, the original 78 Thoroughbreds and 8 additional Thoroughbreds, neck-base circumference was significant (p=0.0248) for genotype, but not for sex (p=0.767). However, this test did not pass the Bonferroni correction for multiple testing given the many independent variables examined in this study (Figure 18).

Table 1. Results of effect tests for a least squares regression model of croup to dock length/withers height on 86 Thoroughbred horses.

Source	**Nparm** [A*]	**DF** [B*]	**Sum of Squares**	**Prob > F**
Genotype	2	2	0.00023787	0.8831
Sex	1	1	0.00090258	0.3345

[A*] Nparm gives the number of parameters estimated for the model
[B*] DF examines the error degrees of freedom

An additional least-squares regression model was derived on the 86 Thoroughbred horses. Probability describes the likelihood that the variable has a significant effect on the model. Genotype and sex had no significant effect on muscle measurements (p=0.8831 and p=0.3345).

DISCUSSION

This study demonstrates that there is no significant correlation between *MSTN* genotype and conformation as judged in a photograph, body measurements or muscle measurements in the Thoroughbred horse. As the results suggest (Figures10-15 and Table 1), the Thoroughbred horse does not exhibit phenotypic differences of the magnitude seen in cattle (Figure 1) and canines (Figure 3). Thus, this research has eliminated the possibility of gross muscular effects on body size and conformation, although subtler changes may still exist.

When the original 101 Thoroughbred horses were enrolled in this study, the 35 previously published body measurements seemed appropriate to utilize in order to quickly examine the skeletal muscle composition of the animal as compared to *MSTN* genotype. However, these analyses failed to identify any significant effect. To address this, an additional four muscle measurements at locations on the body with little adiposity were applied to an additional eight Thoroughbreds in this study. These horses also demonstrated no significant association between *MSTN* genotypes and muscle density and composition. However, sample size of eight horses may be insufficient to detect an effect. A significant association between genotype and muscle measures might have been possible if the number of participants had been increased.

Limitations to this study could have contributed to the lack of significant difference between the measurements and *MSTN* genotype. Sample sizes were limited by material readily available in existing DNA banks. Some horses were eliminated from the study due to erroneous body measurements and failed DNA samples. In order to eliminate the possibility of lurking variables, it would have been advantageous to control diet and exercise.

Regardless of these limitations of sample number, the subjective and quantitative phenotypic analysis (Figure 12) both indicate remarkable conformational similarity in all three *MSTN* genotypes. Certainly in the equine species, a phenotype comparable to "double-muscling" is not expressed. In future studies it may be valuable to determine why this phenotypic expression is diminished or non-existent.

According to our results, the *MSTN* insertion is expected to eliminate expression of *MSTN* by preventing transcription of the normal mRNA. Lack of *MSTN* releases the inhibition of muscle growth, thus yielding the "double-muscling" phenotype expressed in other species. Yet, horses with the *MSTN* insertion that were observed in our study lacked the overt phenotype observed in other species with alleles causing loss of normal *MSTN* transcript (McPherron et. al, 1997 and Mosher et. al, 2007). Hypotheses as to why this might be the case include a *MSTN* gene duplication at this or another location in the genome, an alteration in the equine *MSTN* receptor activin type II, or the use of a novel alternative *MSTN* promoter in the horse.

Gene duplication, a process that may occur as an error in recombination, is a mechanism through which new genetic material is generated during molecular evolution. In the case of *MSTN*, this duplication could be present in a location that cannot be amplified using the specific forward and reverse primers used in this protocol. Thus a second functional *MSTN* gene could protect ERE-1 insertion bearing horses from the phenotypic effects produced by loss of *MSTN* transcription in this copy. In addition, using a computational prediction of promoter motifs from The Berkeley Drosophila Genome Project, we found a TATAA box just 17 base pairs upstream from the ERE-1 insertion which in turn lies 146 base pairs upstream from the start of translation (Berkeley Drosophila Genome Project, 2008) (Hill et. al, 2010). Therefore, due to the proximity of the ERE-1 insertion to

the TATAA box, the ERE-1 insertion is predicted to reduce or eliminate expression of *MSTN*.

While these are certainly plausible explanations for why there is little phenotypic expression in the equine species for the *MSTN* ERE-1 insertion polymorphism, future work should examine these possible mechanisms potentially through an expanded set of muscle-analysis measures, the examination of myostatin protein measurements through biopsy, a larger population of horses, and by controlling environmental effects such as diet and exercise to their studies.

LITERATURE CITED

Arthur , P.F. "Double Muscling in Cattle: A Review." *Australian Journal of Agricultural Research* 46 (1995): 1493-1515.

Bassett, Anna. "Welfare and Belgian Blue cattle." *Animal Welfare Approved.* Web. 20 May 2013. <http://www.animalwelfareapproved.org>.

Binns, M.M, D.A Boehler, and D. H. Lambert. "Identification of the myostatin locus (MSTN) as having a major effect on optimum racing distance in the Thoroughbred horse in the USA." *Animal Genetics* 41 (2010): 154-158.

Brooks, S.A, and E. Bailey. "Exon skipping in the KIT gene causes a Sabino spotting pattern in horses." *Journal of Veterinary Medicine* 11 (2005): 883-902.

Casas, E, J. W. Keele, L. Smith, L.V. Cundiff, and R. T. Stone. "Quantitative analysis of birth, weaning and yearling calving weights and calving difficulty in Piedmontese crossbreeds segregating in inactive myostatin allele." *Journal of Animal Science* 77 (1999): 1686-1692.

Cash, J.N, C.A Rejon, A.C McPherron, D. J Barnard, and T.B Thompson. "The structure of myostatin: follistatin 288: insights into receptor utilization and heparin binding." *EMBO* 17.1662-2676 (2009): 1662-1676.

Dall'Olio, S, L Fontanesi, L Costa, N Tassinari, M Minieri, and A Falaschini. "Analysis of horse myostatin gene and identification of single nucleotide polymorphism in breeds of different morphological types." *Journal of Biomedical Technology* 10 (2010): 1155-1165.

Deveaux, V, I Cassar-Malek, and B Picard. "Comparison of contractile characteristics of muscle from Holstein and double-muscled Belgian Blue fetuses." *Comparative Biochemistry and Physiology* 131 (2001): 21-29.

Gunn, H.M.. "Muscle, bone and fat proportions and muscle distribution of thoroughbreds and quarter horses." *Equine Exercise Physiology* 1 (1987): 253-264.

Gustin, P, B Detry, A Robert, M.L Cao, F Lessire, C Gambier, V Katz, M Ansay, A Frans, and T Clerbaux. "Influence of age and breed on the binding of oxygen to red blood cells of bovine calves." *Journal of applied Physiology* 3 (1997): 784-790.

Hill, E. W, R.G Fonseca, B.A McGivney, J Gu, D.E Machugh, and L.M Katz. "MSTN genotype (g.66493737C/T) association with speed indices in Thoroughbred racehorses." *Journal of Applied Physiology* 1 (2011): 86-90.

Hill, E.W, B.A McGivney, J Gu, R Whiston, and D.E MacHugh. "Equinome - Test for the Best." *Test for the Best*. Web. 20 May 2013. <http://www.equinome.com/>.

Hill, E.W, J Gu, and S.S Eivers. "A sequence polymorphism in MSTN predicts sprinting ability and racing stamina in thoroughbred horses." *PLoS ONE* 5 (2010): 1371.

Hosoyama, T, S Kawada, and R Oshiumi. "Molecular cloning of equine (thoroughbred) myostatin cDNA and detection of myostatin precursor proteins in the serum." *Journal of Reproduction and Development* 4 (2002): 335-342.

McGivney, B.A, S.S Eivers, D.E MacHugh, J.N MacLeod, G.M O'Gorman, S.D Park, L.M Katz, and E.W Hill. "Transcriptional adaptations following exercise in thoroughbred horse skeletal muscle highlights molecular mechanisms that lead to muscle hypertrophy." *BMC Genomics* 10 (2009): 1471-2164.

McGivney, B.A, P.A McGettigan, L.M Katz, J.A Browne, A.C Browne, A.C Evans, R.G Fonseca, B.J Loftus , A Lohan, D.E MacHugh, and B.A Murphy. "Characterization of the equine skeletal muscle transcriptome identifies novel functional responses to exercise training." *BMC Genomics* 10 (2010): 1471-1164.

McPherron, A.C, and S.J Lee. "Double muscling in cattle due to mutations in the myostatin gene." *Proc. Natl. Academy of Science* 23 (1997): 12457-12461.

Mosher, D.S, P Ouignon, C.D Bustamante, N.B Sutter, and C.S Mellersh. "A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs." *PLoS Genetics* 3 (2007): 79.

Radostits, O. M., and S. H Done. *Veterinary medicine: a textbook of the diseases of cattle, sheep, pigs, goats, and horses*. 10th ed. New York: Elsevier Saunders, 2007.

Sambrook, J, and D.W Russell. *Molecular cloning: a laboratory manual*. 3. ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001.

Tozaki, T , F Sato, M Kurosawa, E.W Hill, T Miyake, Y Endo, H Kakoi, H Gawahara, K Hirota, Y Nakano, and Y Nambo. "Sequence variants at the myostatin gene locus influence the body composition of Thoroughbred horses." *Journal of Veterinary Medical Science* 12 (2011): 1617-1624.

APPENDIX A

Horse Measurement Instructions

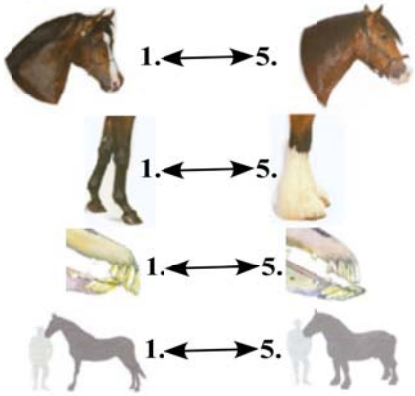
For the Sutter and Brooks Genetics Laboratories

Cornell University
College of Veterinary Medicine

Before you begin, make sure your horse is standing on even ground. Measurements should be recorded in inches. Please do not attempt to measure an unwilling horse. We appreciate your contribution to our equine genetics research.

Body Condition:
Choose a value from 1-9, where 1 is extremely underweight, 9 is overweight and 5 is ideal.

Skull Shape: Rate the skull shape where 1 is deeply dish-faced, 3 is a straight profile, and 5 is strongly Roman nosed.



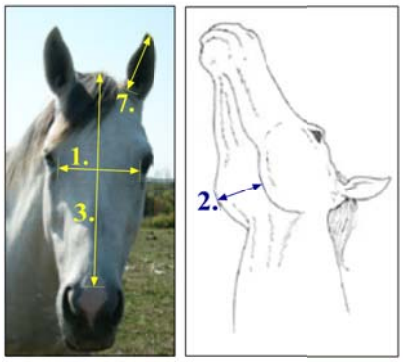
Feathering: Rate the amount of feathering where 1 is absolutely no feathering and 5 is copious feathering.

Bite Conformation: Rate the bite where 1 is an extreme overbite and 5 is an extreme underbite.

Bone Thickness: Rate bone thickness or heaviness without consideration of the horse's size, where 1 is a light, thin skeleton and 5 is a heavy, thick skeleton.

HEAD:

1. Eye to Eye Width: Standing in front of the horse, measure across the forehead between the inside corners of the eyes.



2. Jaw Width: Across the underside of the head, measure straight across from the deepest points of the cheekbones.

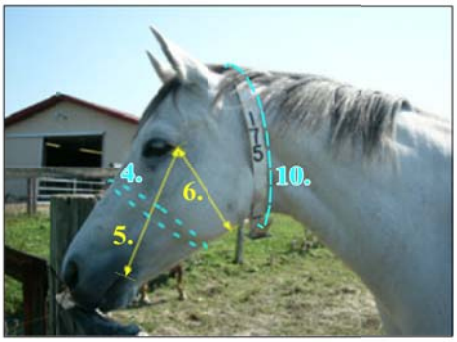
3. Head Length: Starting between the top corners of the two nostrils, measure straight to the front of the poll.

4. Muzzle Circumference: Take the circumference of the muzzle, settling the tape directly in front of the cheekbones.

5. Left Eye to Mouth Length: Measure from the corner of the mouth to the back corner of the eye.

6. Left Eye to Jaw Length: Measure from the back corner of the eye to the deepest point of the cheek curve.

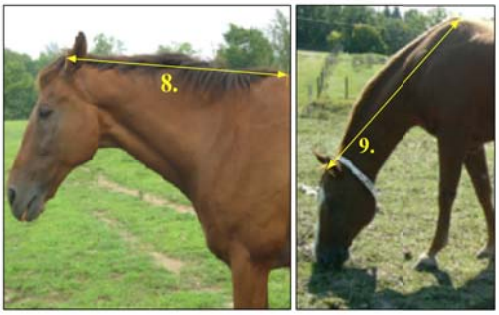
7. Left Ear Length: On the side of the ear closest to the poll, measure from the base to the tip of the ear.



NECK

8. Length, *Head Level with Withers*: Measure from the back base of the ears to the withers, with your horse's head level with the withers.

9. Length, *Head Down to the Ground*: Measure as in **8.**, but with your horse's head stretched as close to the ground as possible.



10. Circumference, *at Throat Latch*: Settle the tape where the throat latch of the bridle goes. Pull it snugly but not uncomfortably for your horse.

11. Circumference, *at Base*: Settle the tape just in front of the withers. Let it rest on the chest and curve comfortably around the base of the neck.

BODY:

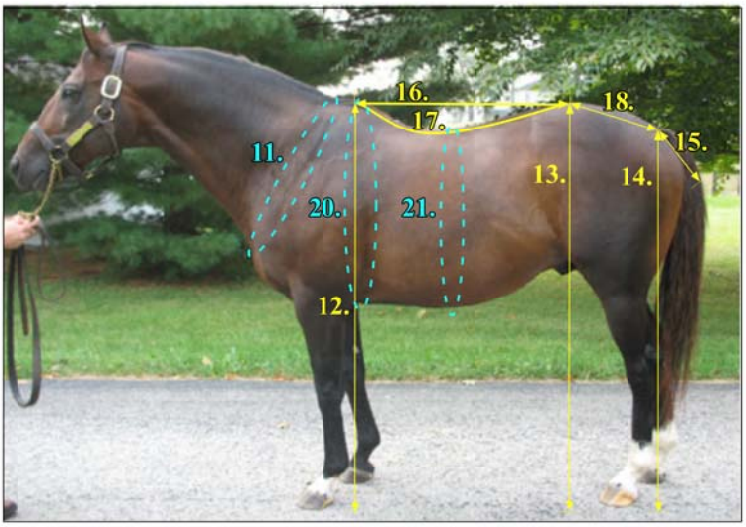
12. Height, *at Withers*: Measure from the ground to the highest point of the withers.

13. Height, *at Croup*: Measure from the ground straight up to the highest point of the rump.

14. Height, *at Dock*:Measure from the ground straight up to the base of the tail.

15. Tail Length: Start at the end of the bony portion of the tail and measure to the dock.

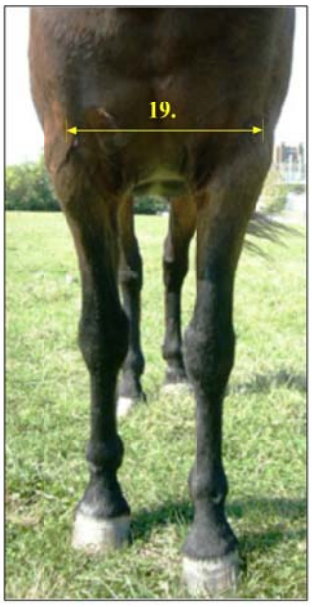
16. Withers to Croup, *Straight*: Measure across the back from the withers to the point of the croup. The tape won't touch the back except at the ends.



17. Withers to Croup, *Contour*: Measure the back from the withers to the croup, allowing the tape to relax and touch the entire length of the spine.

18. Length, Croup to Dock: Measure from the point of the croup to the base of the tail.

19. Chest Width: Feel for the humeral bones that project forward out of the chest and measure the distance between the outside edges.



20. Barrel Girth, *at Heart*: Settle the measuring tape where the girth of the saddle fits, directly behind the forelegs.

21. Barrel Girth, *Maximum*: Measure around the barrel at greatest circumference.

FORE LIMB:

22. Left Forearm Length: Measure from the point of the elbow to the back of the kneecap.

23. Left Fore Cannon Length: Measure from the back of the knee cap to the ergot.

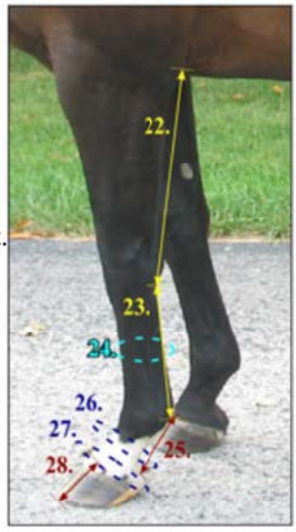
24. Left Fore Cannon Midpoint Circumference: Measure around the cannon bone halfway between the knee and the ergot.

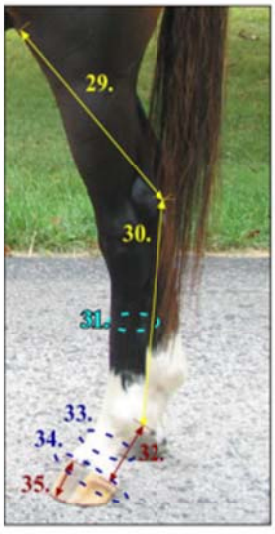
25. Left Fore Pastern Length: Measure from the ergot to the top of the coronet.

26. Left Fore Pastern Circumference: Measure around the pastern.

27. Left Fore Coronet Circumference: Measure around the coronet.

28. Left Hoof Length: Measure from the coronet to the bottom of the hoof.





HIND LIMB:

29. Left Gaskin Length: Measure from the point of the hip (feel under the flank muscles) to the point of the hock.

30. Left Hind Cannon Length: Measure from the point of the hock to the top of the fetlock.

For Measurements 31-35 of the hind limb, please follow the directions for 24-28. Thank you!

APPENDIX B

Horse HSID	Genotype
23	Failed
307	Failed
1083	Failed
1177	Failed
1611	Failed
24	Failed
255	Failed
298	Failed
316	Failed
319	Failed
447	Failed
702	Failed
736	Failed
867	Failed
1366	Failed
1383	Failed
1384	Failed
1387	Failed
1395	Failed
1396	Failed
1606	Failed
231	(+/ins)
300	(+/ins)
375	(+/ins)
448	(+/ins)
751	(+/ins)
1066	(+/ins)
1131	(+/ins)
1247	(+/ins)
1386	(+/ins)
1448	(+/ins)
1449	(+/ins)
1452	(+/ins)
1454	(+/ins)
1523	(+/ins)
1605	(+/ins)
1619	(+/ins)
1677	(+/ins)
1928	(+/ins)
1931	(+/ins)

Horse HSID	Genotype
1932	(+/ins)
1933	(+/ins)
636	(ins/ins)
662	(ins/ins)
696	(ins/ins)
1249	(ins/ins)
1385	(ins/ins)
1431	(ins/ins)
1450	(ins/ins)
1455	(ins/ins)
1456	(ins/ins)
1462	(ins/ins)
1526	(ins/ins)
1528	(ins/ins)
1614	(ins/ins)
1880	(ins/ins)
124	(+/+)
224	(+/+)
226	(+/+)
228	(+/+)
239	(+/+)
248	(+/+)
295	(+/+)
301	(+/+)
302	(+/+)
304	(+/+)
305	(+/+)
309	(+/+)
310	(+/+)
311	(+/+)
379	(+/+)
489	(+/+)
689	(+/+)
703	(+/+)
711	(+/+)
714	(+/+)
716	(+/+)
725	(+/+)
731	(+/+)
740	(+/+)

Horse HSID	Genotype
747	(+/+)
787	(+/+)
878	(+/+)
914	(+/+)
988	(+/+)
1010	(+/+)
1071	(+/+)
1112	(+/+)
1129	(+/+)
1133	(+/+)
1137	(+/+)
1207	(+/+)
1382	(+/+)
1397	(+/+)
1398	(+/+)
1400	(+/+)
1415	(+/+)
1419	(+/+)
1451	(+/+)
1453	(+/+)
1461	(+/+)
1663	(+/+)
1914	(+/+)
1927	(+/+)
1929	(+/+)
1930	(+/+)
1934	(+/+)
3729	(+/+)
GS140	(+/+)

APPENDIX C
Protocols Utilized in this Study

I. Electrophoresis and Gel-Loading Buffers (Sambrook et. al, 2001)

Buffer	**Working Solution (0.5X)**	**Stock Solution/Liter (5X)**
TBE	45 mM Tris-borate 1mM EDTA	54g of Tris base 27.5 g of boric acid 20ml of 0.5 M EDTA (pH 8.0)

II. Hair Lysis Solutions for DNA Extraction (Brooks and Bailey, 2005)

Buffer	**Working Solution**
Hair Lysis Buffer (HLB)	1mL 10X PCR Buffer without MgCl2 1mL MgCl2 50uL Tween 20 Bring up to 10mL with autoclaved MilliQ H2O