POLYMORPHIC GAIT IN THE HORSE:
AN INTERACTION OF GENETICS, MORPHOLOGY, AND BEHAVIOR

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Elizabeth Ann Staiger

January 2015

POLYMORPHIC GAIT IN THE HORSE:
AN INTERACTION OF GENETICS, MORPHOLOGY, AND BEHAVIOR

Elizabeth Ann Staiger, Ph. D.

Cornell University 2015

Selection after domestication has primarily focused on performance, conformation and desirable behaviors in the horse, resulting in breeds that are divergent across these traits. An example are the "gaited" breeds, horses with the ability to perform either a lateral or diagonal four-beat gait without a moment of suspension at intermediate speeds, yet varying in overall size and temperament.

To investigate the contribution of genetics to these divergent traits, we collected DNA samples, 35 body measurements, gait information, horse discipline, and a behavior survey from 801 gaited horses. Utilizing previously genotyped horses, an across-breed genome-wide association study (GWA) identified three novel candidate regions associated with gait type on ECA1, ECA11, and EC4. A GWA in a single gaited breed, the Tennessee Walking Horse (TWH) identified two additional candidate regions on ECA19 and ECA11. Polymorphisms from whole-genome sequences have identified several SNPs within these candidate regions.

We conducted principle component analysis (PCA) on 33 of the body measures for a subset of TWH. A GWA of the first PC, which describes overall size, identified the *LCORL* locus, which has previously been implicated with size in horses, cattle, and humans. No causal variant has been discovered yet due to extensive linkage disequilibrium (LD) in the region, but LD in the TWH is much lower, improving the resolution capabilities for fine-mapping and variant discovery.

To investigate the contribution of genetics to temperament, we utilized factor analysis (FA) on the questionnaire to identify four temperament factors in TWH: neophobia, trainability, hostility, and independence. These four factors account for 64% of the total trait variance. We ran three separate GWAs using the F1-'neophobia', F2-'trainability', and F3-'hostility' scores as the phenotype and identified candidate markers in genes involved with neurodegeneration, steroidogenesis, brain development, and neuronal cell signaling pathways.

The results from this work will hopefully lead to future studies to identify the causal variants of locomotion, size, and behavior traits. This will allow for the development of genetic tests to aid horse owners in their breeding and management decisions and help improve horse welfare as horses are selected for appropriate disciplines.

BIOGRAPHICAL SKETCH

Elizabeth Ann Staiger was born in Portsmouth, Virginia, but spent her formative year in Japan, California, and Maryland. She attended Oklahoma State University and graduated in 2007 with a Bachelor of Science degree in Animal Science and with an Honor's degree. She continued at Oklahoma State University for a Master of Science degree in Animal Breeding and Genetics under Dr. Raluca Mateescu studying genetic markers for increased milk and meat production in sheep. After completing her Masters in 2009, Ann remained in Dr. Mateescu's lab as the lab manager.

Ann began her doctoral studies in Animal Science at Cornell University in 2010. She pursued her research on the genetics of polymorphic gait in the horse under the direction of Dr. Samantha Brooks. While pursuing her degree, Ann has been the recipient of the Everingham Award, Neal A. Jorgenson Travel Award, and winner of the Equine Science Society Genetics Graduate Student competition.

"Intellectual growth should commence at birth and cease only at death"

– Albert Einstein

Dedicated to all of the teachers throughout my life.

ACKNOWLEDGMENTS

The inspiration for this research stemmed from my love and involvement with gaited horses, particularly Tennessee Walking Horses, and a curiosity to understand the genetic underpinnings that makes these horses so unique from other breeds.

This research would not have been possible without the involvement of several people. I owe a great deal of gratitude to Dr. Samantha Brooks for serving as my committee chair and mentor for the past four and a half years, and for allowing me the freedom to work on a project about which I am truly passionate. I know it has been an adventure, and I can't thank you enough. Thanks to Dr. Joe Fetcho for help on neurobiology and all of his helpful advice on designing our behavior trials in the horse. Thanks to Dr. Jon Cheetham for providing important insights on the physiology of the horse, especially in regard to changes over time. And a special thank you to Dr. Heather Huson for her advice on performance genetics and for "adopting" me over this past year.

This work was definitely a collaborative effort from several members of the equine science community. Thank you to Dr. Rebecca Bellone for the Puerto Rican Paso Fino samples, hosting me on some of my data collection trips, and for all of the career and teaching advice/support she provided. Dr. Nate Sutter provided advice on the measurement data and PCA. Dr. Julia Albright was instrumental in the design of the behavior survey and trial, and oversaw all of the behavior trials. I would also like to thank Dr. Ernest Bailey for all of his advice and support, and graciously hosting me on data collection trips.

I could not have survived without the support, friendship, and pet-sitting of my fellow graduate students and friends Kristen Davis, Claire Stephens, Jeremy Allen, and Fernando Migone. A special thanks to my lab mate Dr. Heather Holl for all of the troubleshooting help, discussions, and helping me on data collection trips.

I owe a big thank you to the many undergraduates who have worked in the lab and helped with the project over the years. Thanks to Nicole SanGiacomo for entering records into the

database, Mariya Gugel for genotyping some of the behavior SNPs, Ram Singh for the Marwari samples and genotyping help, and Alexander Thomson, Taylor Baird, Rachel Evanowski, and Janelle Slutsky for help on local farm trips. Special thanks to Chris Posbergh and Lauren Jacobs who were essentially my minions in the lab, but also agreed to go on several out-of-state and overnight trips.

I owe a special thanks to all of the horse breeders, owners, and trainers who have submitted samples over the years, and hosted me on my data collection trips. A big thank you to all of the breed organizations that have promoted my research: TWHBEA, RMHA, KMHSA, USIHC, PFHA, NWHA, FOSH, USMMA, ABCCC, and ABCCMM. Without the help of the associations and the owners, I would never have been able to collect so many horses or have the means to carry out the work. I especially want to thank Joe London, Joyce Moyer, Lynn Kelley, Lori Miller, and Laura Patterson. They have been amazing supporters of the study, and have done whatever they could to get me the material I have needed.

My deepest gratitude for the financial support my research has received from the Cornell Center for Vertebrate Genomics, FAST, the Dorothy Russell Havemeyer Foundation, the ABCCC, and ABCCMM.

And finally, I would like to thank my family for all of their support and love, not to mention their acquiescence when I coerced them into helping on data trips, and for hosting my colleagues and me as we passed through the area on yet another data trip. You've truly played an active role in my research. To my parents, my first teachers, I wouldn't be who and what I am today without your example, encouragement, and sacrifices. I cannot fully express how grateful I am to be your daughter.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION TO THE DISSERTATION

Introduction

Horses have been a part of human culture for over 5000 years, serving several roles throughout. Archeological data indicate that initial interactions were that of predator and prey. Based on the high prevalence of young equine male bones, often found with other wildlife remains discovered at archeological sites (Levine 1990; Levine 1993), it is presumed that ancient people from Dereivka (Ukraine) and Botai (Kazakhstan) often hunted and killed horses for meat and hide (Levine 1999). Eventually, as humans moved toward a more agricultural lifestyle, horses were captured and kept as livestock for milk and meat. Organic residue analysis of fatty acids identified equine milk residues and carcass products in pottery at ancient Eurasian archaeological sites (Outram *et al.* 2009), supporting the change in roles. These horses were likely not bred to produce offspring, but were captured as needed, a practice that eventually led to domestication. With the prevalence of permanent villages, the horse's role changed from hunted wildlife to captive breeding livestock, first for food, then for transportation (Levine 1999).

Horse domestication was vastly different from other livestock domestication events. For most major livestock species, domestication occurred from a limited number of animals and geographic regions, as reflected in their limited genetic basis. In the horse, genetic diversity studies indicate there is extensive variation among domestic horses. Such diversity alludes to multiple domestication events from several genetically diverse populations (Ellegren 2002; Vila *et al.* 2001). Archaeological and geological records indicate domestication occurred between 9400-2000 BC (Ellegren 2002). This time frame, taken in conjunction with mitochondrial DNA sequence data from more than 600 animals, suggests that a minimum of 77 wild mares were required to explain the genetic diversity observed (Jansen *et al.* 2002). The number of males included in the first domesticated populations is still up for debate. The modern day horse population possesses little diversity on the Y-chromosome (Lau *et al.* 2009; Lindgren *et al.* 2004; Wallner *et al.* 2003; Wallner *et*

al. 2004). Yet Y-chromosome sequencing in ancient and wild horse populations has discovered relatively high diversity (Lippold *et al.* 2011). The observed low Y-chromosome diversity may be a consequence of domestication, achieved either through the incorporation of very few wild male horses (Lau *et al.* 2009; Lindgren *et al.* 2004; Wallner *et al.* 2003; Wallner *et al.* 2004), a global selective sweep of the Y chromosome (Wallner *et al.* 2004), or breeding practices developed after domestication that reduced the effective number of males in the domestic species (Cunningham *et al.* 2001; Levine 1999; Lippold *et al.* 2011).

In feral horse populations, there are two types of groupings which form herds: the natal group and the bachelor group (Keiper and Houpt 1984). The natal group is a stable association of mares, their offspring and one or more stallions who defend and maintain the mare group (Linklater 2000). The more dominant stallion often wins the right to copulate more often than the subordinate stallions (Stevens 1990). Bachelor groups consist mostly of males dispersed from natal groups aged two years and up, and less commonly older stallions that have been unsuccessful in defending their bands from other stallions (Cox 1986; Levine 1990). In the feral and wild groups there is the potential for every male to breed and contribute their genetics. However, the breeding practices imposed on domesticated horses often limits the male gene pool by castrating all males deemed inferior based on conformation, pedigree, performance, and/or behavior (i.e. aggressive towards people) before they have bred a mare.

The Horse as a Model for Complex Traits

In contrast to other livestock species, the horse was primarily under selection for the purpose of transportation rather than consumption following domestication. The increased mobility provided by the horse enabled people to travel further, faster and carry more belongings than ever before (Levine 1999). Initially, horses were likely used primarily for pulling heavy loads and in agriculture (McGreevy 2004); after the invention of the wheel, they could pull chariots, and

then as cavalry mounts when equitation improved (van Weeren 2013). After the industrial revolution, horse use predominately changed to sport and recreation. This selective strategy likely enhanced the divergent traits for locomotion and may have resulted in unique qualities like polymorphic gait. Beyond traits for locomotion, selection may also have focused on athletic attributes such as size and strength (Clutton-Brock 1999) and desirable behaviors (Hislop 1992; Houpt and Kusunose 2000; Lloyd *et al.* 2008) to fill specific roles. Indeed, across modern breeds horses display a wide variation in body size and shape (Brooks *et al.* 2010b). Biomechanical studies have shown that conformation plays a considerable role in horse locomotion (Leach and Dagg 1983), but none have yet been able to identify a conformation trait linked to gait type preference.

Abnormal locomotion and temperament in man are often characterized together in complex disorders such as attention deficit hyperactivity disorder (ADHD) and autism (Hildebrand 1989; Kiehn 2006; Kiehn *et al.* 2010; Kuo 2002), implicating behavior and motor control may share neural pathways and genetic factors. A good example of this is the domestication of the silver fox in which individuals were selected for reduced fear of humans resulting in changes in behavior, morphology, and physiology (Trut, Oskina, Kharlamova 2009). Transcriptome sequencing of prefontal brain samples from one tame and one aggressive fox identified functional differences in the prefrontal brain cortex (Kukekova *et al.* 2012).

Studies using mice, cats, and lampreys, uncovered a neural network called the central pattern generator (CPG) responsible for repetitive actions like those used in locomotion (Kuo 2002). However, none of these laboratory animal models have an innate ability to perform a lateral gait. At intermediate speeds beyond the walk, the horse can perform a range of gaits from a two-beat lateral gait ("pace") to a two-beat diagonal gait ("trot"), including a variety of four-beat diagonal and lateral gaits (Harris 1993). Both the trot and pace have a brief moment of suspension where all four feet are off the ground before one pair lands at the same time. A horse with the ability to

perform either a lateral or diagonal four-beat gait without a moment of suspension is commonly called a "gaited" horse. Other than the horse, no other mammalian species is known to discretely segregate for preference in cadence and footfall pattern.

Locomotion is a key factor in determining the quality of equine performance, a key equine economic trait in the horse industry. It has been estimated that the horse industry has a direct economic impact of $39 billion and a total impact of $102 billion annually on the US, with showing and recreation combined accounting for 60% of the total impact (American Horse Council 2005). Of the 9.2 million horses in the US, 29% are involved in showing, 42% in recreation, and less than 19% involved in breeding (American Horse Council 2005). With a small number of horses being used for the production of the major money-makers in the industry, it is extremely important to horse owners and breeders to utilize all available tools in their breeding management decisions. The selection of any horse is based upon conformation, temperament, and gait (Lloyd *et al.* 2008); however, this selection process is more art form than science and lack of precision can lead to poor performing horses, resulting in the loss of both time and money for the breeder. Providing genetic tests that predict gait qualities, skeletal conformation, and behavioral temperament will allow breeders to make better management decisions and improve the marketability of their horses. Use of genetic tests should also improve animal welfare as the horses are selected for, and used in, more appropriate equine disciplines.

Beyond impact on the horse industry, studying horse gait, morphology, and behavior can aid in research of human disease and engineering of robotics circuitry. Animal models have long played an important role in studies of human disease due to similarities in fundamental mammalian biology and physiology as well as experimental convenience. Studies of complex traits in the horse are still relatively rare, but are increasingly feasible as simple traits are flushed out.

As of the writing of this work, there are no other animal models that naturally segregate for lateral locomotion patterns to the same degree as the horse. Inherited polymorphisms in the pattern and timing of equine locomotion offer a unique model for the study of gait development. We predict that genes involved in the CPG pathway are responsible for the different gait patterns observed in the horse, and genes controlling skeletal conformation and temperament play contributing roles in the quality of the gait pattern.

Locomotion

Locomotion is an important trait that separates animals from plants. Locomotion is vital for animals to gather food, protect themselves, and reproduce. Across the different species of animals there are different types of locomotion or gait: symmetrical or asymmetrical, and bipedal or quadrupedal. Symmetrical gaits have an evenly timed footfall, with each foot spending the same amount of time in contact with the ground (Hildebrand 1989). Asymmetrical gaits have couplets in the footfall pattern where two or more feet move together (Hildebrand 1989). Walking is an example of a symmetrical gait, while hopping (e.g. rabbit) is an asymmetrical gait. These gaits can be further divided into lateral or diagonal gaits based on how the front limbs move in relation to the hind limbs. Each gait type is observable across several species, but only in horses are all the gait types inherently observable within a single species (Hildebrand 1989).

The Equine Gait Spectrum

There are four major gait definitions for the horse based upon speed. The fastest gait is the gallop at 20-45 mph or 9-20 m/sec (Barrey 2013); this is a four-beat gait that follows a diagonal footfall pattern. The footfall sequence is right hind, left hind, right front, left front followed by a moment of suspension where all four hooves are off the ground once the left front leaves the ground. This gait can also start with the left hind resulting in a mirror image of the footfall

sequence described above. After the gallop, the next fastest gait is the canter (6.5-20 mph, 2.9-9 m/sec), a three-beat gait that also contains a moment of suspension (Barrey 2013). The pattern of the canter is the same as the gallop, except the diagonal hooves (ex: left hind, right front) leave and hit the ground together at the same time. The slowest gait of the horse is the walk at 2.5-4 mph or 1.2-1.8 m/sec (Barrey 2013). The footfall sequence of the walk is right hind, right front, left hind, left front. This is a lateral four-beat footfall sequence, but it is evenly timed in the lift-off and landing of the hooves, so both lateral and diagonal bi-pedal and tri-pedal supports are seen without a moment of suspension (Barrey 2013; Ziegler 2005). The head nods in action with the placement of the hind hooves. All horses can perform the walk and gallop, and nearly all can perform the canter (Barrey 2013; Ziegler 2005).

Gaits performed at speeds between that of the walk and canter are what are known as the intermediate gaits. These include the trot, pace, amble, running walk, rack, foxtrot, and batida. These gaits can be performed at speeds ranging from 6-35 mph or 2.8-16 m/sec (Barrey 2013; Ziegler 2005). Intermediate gaits are more of a continuum rather than a set distinct classifications due to the spectrum of variation in minute differences of speed, footfall patterns, and stance durations (Hildebrand 1965). Individual performance of the same gaits may also vary between horses (Nicodemus and Clayton 2003). Categorization of intermediate gaits typically relies on the footfall sequence, footfall timing, cadence (ex: lateral four-beat vs diagonal two-beat) and the movement of the head (ex: head shake versus no head shake)(Harris 1993; Ziegler 2005). All of the intermediate gaits, except the trot and pace, have the same fundamental footfall sequence as the ordinary walk (Nicodemus and Clayton 2003).

The lateral-coupled intermediate gaits include the pace, amble, rack, and running walk. The pace marks one end of the intermediate gait spectrum. It is a lateral gait, i.e. the lateral pair of hooves (ex: right hind, right front) traveling together (leaving and hitting the ground at the same

time). There is a moment of suspension where all four hooves are off the ground between the set-down of each lateral pair, resulting in a 1-2 cadence and only lateral bi-pedal support (Barrey 2013). There is no head nod/shake in the pace. The amble, also sometimes referred to as a broken pace, is similar to the pace in that the lateral pairs travel together, but the set down and lift-off is close together in time with the hind hoof hitting the ground slightly before the front hoof in 1-2—3-4 cadence (Ziegler 2005). The amble has tri-pedal and both lateral and diagonal bi-pedal support (Nicodemus and Clayton 2003), and no head shake.

The rack is one of the most common four-beat intermediate gaits, with many different names depending on the region, breed, or speed (Ziegler 2005). At slower speeds (3-12 mph or 1.3-5 m/sec) it is referred to as the saddle rack, single-foot, fino, corto, or slow tölt (Nicodemus and Clayton 2003; Ziegler 2005). At faster speeds (up to 25 mph or 11 m/sec) it is called the rack, largo, fast tölt, or marcha picada (Nicodemus and Clayton 2003; Ziegler 2005). The gait is a four-beat gait with a lateral footfall sequence and lateral pick-up time, but with each hoof hitting at an evenly timed interval (Ziegler 2005). This difference in set-down and pick-up time is due to higher step action in the front legs compared to the other four-beat gaits. At slower speeds, the rack has tri-pedal and bi-pedal support, but as the speed increases this changes to include mono-pedal support (Nicodemus and Clayton 2003). At its fastest speeds, tri-pedal support is no longer observed so that there is never a moment when both hind, or both front hooves, are in contact with the ground at the same time (Nicodemus and Clayton 2003). Often the high speed rack (or tolt) resembles a four-beat pace, so misclassification of these two gaits is common (Robilliard, Pfau, Wilso 2007; Zips *et al.* 2001). There is little to no head shake during performance of the rack.

The running walk is considered a "square" gait in that it is evenly timed in both the set-down and pick-up of the hooves. The footfall sequence and cadence is exactly the same as the walk, but is performed at higher speeds (7-20 mph or 3-9 m/sec) (Ziegler 2005). Like the ordinary walk,

there is a head shake/nod in the running walk, timed so that the head moves down with set-down of the hind hooves (or with the rising of the front knees) (Lane 2011). A unique feature to the gait is the long over-stride, the distance the hind hoof sets down in front of the front hoof track (Tennessee Walking Horse Breeders' and Exhibitors' Association 2011). In the ordinary walk the over-stride is none to 4 inches whereas it can be 6-18 inches in the running walk (Lane 2011).

The diagonal gaits include the trot, broken trot, fox trot, and batida. The trot marks the opposite end of the intermediate gait spectrum as compared to the pace. It is a diagonal gait, with the diagonal pair of hooves (ex: right hind, left front) traveling together (leaving and hitting the ground at the same time). The trot can be performed at speeds ranging from 6-32 mph or 2.8-14.2 m/s (Barrey 2013). There is a moment of suspension where all four hooves are off the ground between the set-down of each diagonal pair, resulting in a 1-2 cadence and only diagonal bi-pedal support (Barrey 2013). There is no head nod/shake in the trot. The broken trot is similar to the trot in that the diagonal pairs travel together, but the set down and lift-off is close together in time with the hind hoof hitting the ground slightly before the front hoof in a 1-2—3-4 cadence (Ziegler 2005). The broken trot has brief tri-pedal and lateral bi-pedal supports and longer diagonal bi-pedal support, and no head shake (Nicodemus and Clayton 2003).

The footfall sequence of the foxtrot is the same as the ordinary walk (right hind, right front, left hind, left front) so the footfall sequence is lateral, but the footfall timing is diagonal (Nicodemus and Clayton 2003). The diagonal hooves (ex: right hind, left front) leave and hit the ground together unevenly in time, with the front hoof hitting the ground before the hind hoof in a 1-2—3-4 cadence (Ziegler 2005). The gait is often described as walking in the front with a head shake and trotting in the hind. The foxtrot is typically performed at speeds ranging from 6-10 mph or 2.7-4.5 m/s (Barrey 2013). The gait has tri-pedal, diagonal bi-pedal, and brief lateral bi-pedal moments of support (Nicodemus and Clayton 2003). The marcha batida has the same footfall sequence and

timing as the fox trot, however the batida resembles the trot more in front leg step height and length than the fox trot and is performed at speeds ranging from 8-11 mph or 3.9-5 m/s (Associacao Brasileira de Criadores do Cavalo Mangalarga Marchador 2006). There is no "walking" in the front end and therefore no head shake in the batida. The batida does utilize tri-pedal, diagonal bi-pedal, and lateral bi-pedal moments of support (Nicodemus and Clayton 2003).

The Innervation of the Central Pattern Generator (CPG)

The nervous system is a network of neurons that coordinates the actions of an animal via transmitted signals to different parts of its body. Neurons can be divided into two major categories: sensory, such as touch, taste, smell, and hearing; and motor: for walking, lifting, turning our head, etc. The Nervous System can be divided into two functioning parts: the Central Nervous System (CNS) and the Peripheral Nervous System (PNS). The CNS is comprised of the brain and spinal cord; the PNS is comprised of the cranial and spinal nerves. In the horse, the brain weighs between 400 and 700 grams representing a ratio of 1:800 when compared to total body weight; by comparison, the ratio in the dog is about 1:100, giving the horse a relatively small brain for its physique (Riegel and Hakola 1999). There are twelve pairs of cranial nerves; five pairs function in motor pathways, three pairs in sensory pathways, and the remaining four pairs are mixed function (Riegel and Hakola 1999). There are usually 42 pairs of spinal nerves further divided into 8 cervical, 18 thoracic, 6 lumbar, 5 sacral, and 5 coccygeal pairs (Riegel and Hakola 1999).

Detailed knowledge is still lacking on the neuroanatomical and neurophysiological basis of locomotion in horses due to the lack of adaptation of non-invasive techniques such as transcranial magnetic stimulation with the recording of EMG or fMRI-techniques (Gramsbergen 2013). Thus the etiology of equine locomotion is based on laboratory animal models and post-mortem examination of equine brain and spinal tissues. From these, it has been determined that locomotion is controlled by four major neural systems: the local spinal cord and brainstem circuits,

descending modulatory pathways, the cerebellum, and the basal ganglia (Purves 2008). All movement produced by skeletal muscles is initiated by lower motor neurons in the spinal cord and brainstem that directly innervate the skeletal muscles (Gramsbergen 2013). The lower motor neurons are controlled by local circuits within the spinal cord and brainstem that coordinate individual muscle groups. Local circuit neurons also receive sensory inputs from the descending projections of upper motor neurons in the higher centers, regulating those local circuits (Purves 2008). Thus the spinal cord and brain stem circuits enable and coordinate the complex sequences required for movement. The upper motor neurons make up the descending modulatory pathways and originate in either the cerebral cortex or the brainstem (Purves 2008). Pathways that arise from the cortex are responsible for voluntary and skilled movements while those from the brainstem are responsible for regulating muscle tone and for orienting the eyes, head, and body with respect to vestibular, somatic, auditory, and visual sensory information (Purves 2008). Circuits in the basal ganglia and cerebellum regulate the upper motor neurons through the suppression of unwanted movements and detecting motor error, respectively, thereby facilitating the initiation and performance of movement with spatial and temporal precision (Purves 2008). A summary of these interactions is presented in Figure 1.1.

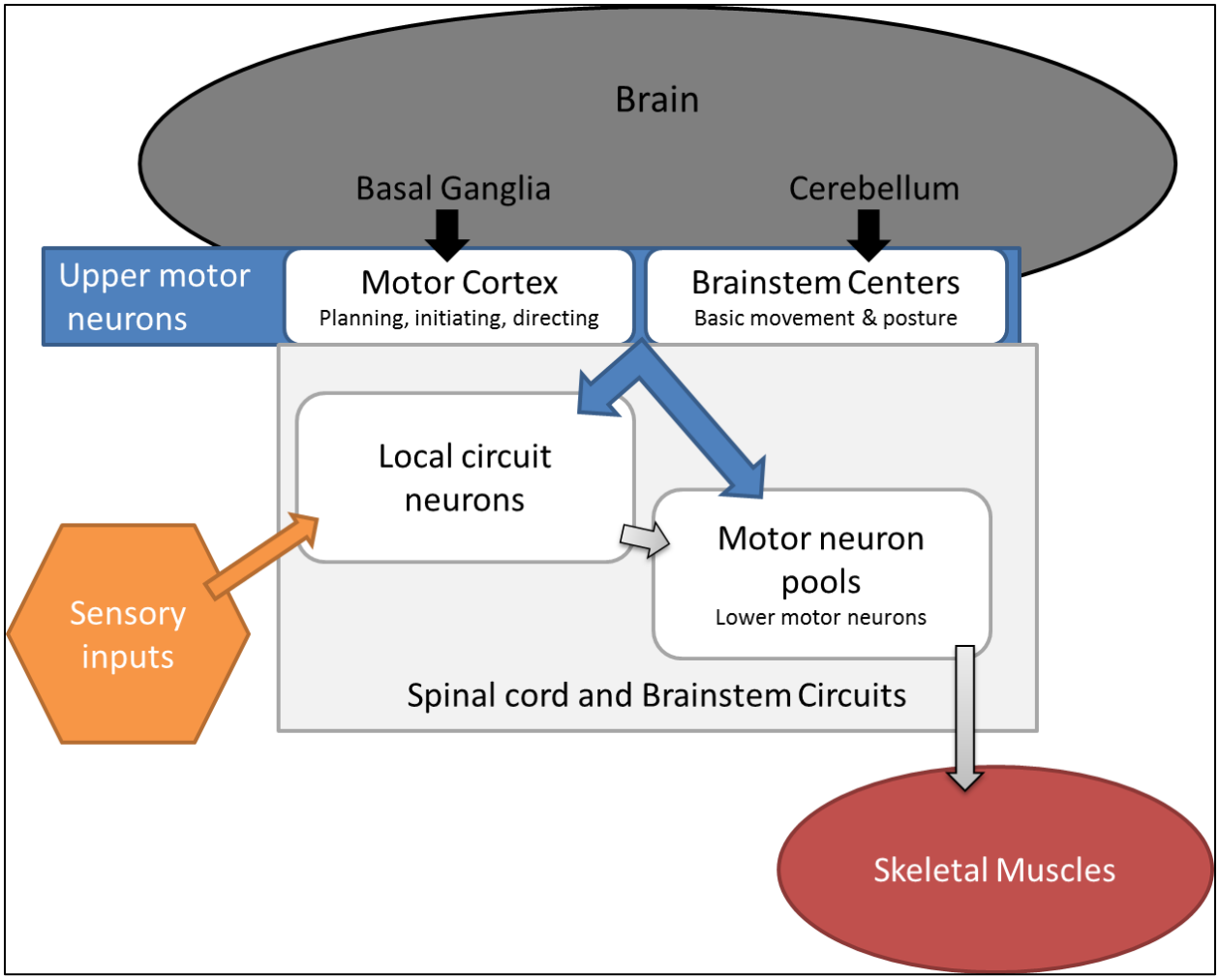


Figure 1.1. Overall organization of the neural circuits involved in locomotion. Adapted from (Purves 2008)

There are two types of lower motor neurons. The α-motor neurons innervate striated muscle fibers to generate the forces necessary for movement and posture while γ-motor neurons innervate specialized muscle fibers to regulate sensory input (Purves 2008). Muscle fiber groups in a single muscle are innervated by a single α motor neuron, called a motor unit (Gramsbergen 2013; Purves 2008). As there are many more muscle fibers than motor neurons, each neuron has axon branches within the muscle to synapse on many different fibers (Purves 2008). Both motor units and α-motor neurons vary in size. Small α-motor neurons innervate type I muscle fibers and are classified as slow-twitch, fatigue-resistant 'S-type units' (Burke 1981). Moderate sized α-motor neurons innervate type IIa muscle fibers and are classified as fast-twitch, fatigue-resistant 'FR-type units' (Burke 1981). The largest α-motor neurons innervated type IIb muscle types and are classified as fast-twitch, fatigable 'FF-type units' (Burke 1981). The smaller motor units innervate fewer fibers while the larger motor units innervate more muscle fibers. All of the motor neurons innervating a single muscle are termed a motor neuron pool and share connectivity with the spinal tracts that runs parallel or perpendicular to the long axis of the spinal cord for one or more spinal cord segments (Gramsbergen 2013; Purves 2008). In cats and humans, there is an orderly relationship between the location of some of these motor neuron pools and the muscles they innervate, providing a spatial map of the body's musculature (Scheibel and Scheibel 1970; Schoenen 1982). Neurons that innervate the postural muscles of the trunk are located medially in the spinal cord, while those that innervate the muscles of the leg lie farther from the midline (Gramsbergen 2013).

Limbed locomotion in mammals involves recurring activation of flexor and extensor muscles within a limb, and coordinated activity between the left and right limbs (Kiehn *et al.* 2010). The actual timing and coordination of rhythmic muscle activity is generated by neuronal circuits in the spinal cord called central pattern generators or CPGs (Kiehn *et al.* 2010). CPGs are found in

the spinal cord, but the spinal neurons receive projections from neurons that originate from the brainstem and midbrain (Yuste *et al.* 2005). Analysis of these CPGs has shown that the core of the network consists of groups of excitatory and inhibitory neurons that serve designated roles in the network operation (Kiehn *et al.* 2010). The key features of the network are rhythm-generation, flexor-extensor alternation, and left-right coordination (Kiehn 2006). The mammalian locomotor CPG is composed of multiple distributed rhythm-generating core networks (Grillner 1981). This dual nature of the CPG was first discovered by Shik and Orlovsky based on the observation that following spinal cord transection the stepping frequency in cats on a treadmill with two belts at differing speeds shifted to a 1:2 relationship (1976). These individual networks are recruited into a functional single CPG when locomotion is initiated. The network controlling hip movement acts as the leading oscillator (Kiehn 2006), with diverse schemes in coupling of the independent fore and hind limb CPGs producing a variety of gaits, especially in the quadruped (Orlovsky, Deliagina, Grillner 1999).

Hindlimb muscles are controlled by the lumbar spinal cord CPGs while the forelimb muscles are controlled by the cervical spinal cord CPGs. Rostral lumbar segments have a greater capacity to generate rhythmic motor outputs than the caudal segments, possibly due to more intraspinal inputs or a larger number of receptors for neuro-modulatory substances such as 5-HT, dopamine or NMDAs (Kiehn 2006). The rostral segment may also contain rhythmogenic CPG interneurons that directly control hip movements (Kiehn 2006).

All locomotor-related neurons are concentrated in a ventral location in the grey matter of the spinal cord, along laminae VII, VIII, and X (Kiehn 2006). Ventral spinal cord interneurons are identified by unique expression of different transcription factors: V0 by Evx1, V1 by En1, V2 by Chx10/GATA2-3, and V3 by Sim1 (Crone *et al.* 2008; Grossmann *et al.* 2010). Ipsilaterally projecting glutamatergic excitatory interneurons are responsible for rhythm generation (Crone *et al.*

2008; Kiehn 2006). There are four classes of excitatory interneurons: 1) neurons in the lower lumbosacral region in the intermediate area of the spinal cord that project to the extensor muscles (Kiehn 2006). 2) Neurons in the intermediate zone of the mid-lumbar (L3-L4) input from group II muscle afferents and projecting to the lower (L7) lumbar cord (Kiehn 2006). 3) EphA4 and ephrin B3, are normally ipsilateral neurons and lie in the ventral region of the lumbar spinal cord (Kiehn 2006). 4) HB9, a transcription factor, is expressed specifically in embryonic motor neurons and in interneurons close to the midline in the upper lumbar spinal cord (Kiehn 2006).

The appropriate alternation between flexor and extensor neurons of the same side requires the action of inhibitory networks. Flexor and extensor motor neurons receive rhythmic glycinergic inhibition alternating with glutamatergic excitation (Kiehn 2006). With glycinergic inhibition blocked, flexors and extensors are activated in synchrony (Kiehn 2006). Renshaw cells (RC), ipsilateral inhibiting interneurons, tune the firing rates of motor neurons and contribute to burst termination (Kiehn 2006). Ia-Ins (similar to RC) are involved with rhythmic inhibition of motor neurons during locomotion and coordinate flexor and extensor interactions (Kiehn 2006).

While the CPGs are capable of generating locomotion, sensory feedback is also utilized to adapt for environmental constraints. Sensory feedback can aid in 1) controlling the timing of different phases in the step cycle via direct access to the rhythm-generating network, 2) shaping the pattern of muscle activities within a step cycle by reflex pathways to motor neurons, 3) contributing to excitatory drive of the motor neurons, and 4) contributing to a long-term adaptation of the locomotion (Hultborn and Nielsen 2007). Exteroceptive and proprioceptive information influences the extension and flexion phases of the step cycle through the segmentally arranged neurocircuitry (Gramsbergen 2013). Proprioceptive feedback from the proximal limb joints, as well as the afferents from the muscles bridging the joints, play important roles in adapting hindlimb movements for changes in speed and turning (Andersson and Grillner 1983; Grillner and Rossignol

1978; Hultborn and Nielsen 2007; McVea 2005; Saltiel and Rossignol 2004a; Saltiel and Rossignol 2004b). Stumbling over unexpected objects in the animal's trajectory leads to adjustments of the step cycle (Rossignol, Lund, Drew 1988) through the cutaneous receptors to initiate knee flexion and create an exaggerated swing phase in the step cycle (Forssberg, Grillner, Rossignol 1977; Forssberg 1979; Rossignol 2006). Other experiments have demonstrated that increasing the load on the extended limb and input from Golgi-tendon organs via group II afferents also leads to adjustments of the step cycle (Duysens and Pearson 1980), and the magnitude of burst activity in the neurons is doubled by sensory feedback (Donelan and Pearson 2004; Gorassini *et al.* 1994; Hiebert and Pearson 1999; Stein, Misiaszek, Pearson 2000).

EPHB3 – A Candidate Gene for Gait Phenotypes

Inhibitory interneurons modulate the homeostasis between excitation and inhibition, precisely gating information and shaping neuronal circuits (Klausberger and Somogyi 2008). Interneurons only account for 20-25% of all cortical neurons and are essential for modulating mature neocortical brain function (Druga 2009). Any imbalance in their proportions can cause instability in neuronal circuits and lead to neurological and mental diseases (Marin 2012). Ephrins and the Eph receptor tyrosine kinases are an important group of brain wiring regulation molecules, often called axonal guidance cues (Bolz and Castellani 1997; Rudolph *et al.* 2014). An interesting feature of the ephrins is their signaling can be induced in both forward and reverse pathways (Davy and Soriano 2005). Ephrins were initially discovered as regulators for the formation of topographic projections (Cheng *et al.* 1995; Drescher *et al.* 1995), but they also play several additional roles in nervous system development (Klein and Kania 2014; Lisabeth, Falivelli, Pasquale 2013). The Eph/ephrin system is involved in the selection of specific migratory routes of interneurons to the cortex (Zimmer *et al.* 2011) and can act as a motogenic signal for migrating cortical interneurons (Steinecke *et al.* 2014).

Ephrin-B3 is a particularly interesting member of the ephrin family. Ephrin-B3 acts as the midline barrier that prevents corticospinal tract projections from re-crossing the midline when they enter the spinal gray matter (Yokoyama *et al.* 2001; Kullander *et al.* 2001a). The absence of both forward and reverse signaling results in the motor cortex on one side of the brain providing bilateral input to the spinal cord. Absence of Ephrin- B3 in knockout mice yields a hopping locomotor pattern apparent in walking, running, swimming, and scratching (Yokoyama *et al.* 2001; Kullander *et al.* 2001a). A similar phenotype is observed in knockout mice for EphA4 receptor (Dottori *et al.* 1998; Kullander *et al.* 2001b), which binds with high affinity to ephrin-B3 (Gale *et al.* 1996a; Gale *et al.* 1996b). However, in the EphA4 receptor knockout mice, but not the ephrin-B3 knockout mice, there are major neuroanatomical defects in the formation of the anterior commissure, a major forebrain axon tract that consists of both an anterior branch, which connects lobes of olfactory bulbs, and a posterior branch, which connects lobes of the temporal cortex (Dottori *et al.* 1998; Kullander *et al.* 2001b). Mutations in equine ephrin-B3 could similarly alter the interneuron pathways in horses, promoting the horses' ability to perform alternate intermediate gaits.

Heritability and Other "Gait" Studies

Although the heritability for gait type has not been reported in all horse breeds, a strong role for genetics is supported by the discrete segregation among breeds for the propensity to perform one gait type over another. Heritability for gait type within the Icelandic horse was reported to be between 0.38-0.58 (Albertsdóttir *et al.* 2008), but there are two gait type subpopulations within the breed undergoing selection.

Early observations of Standardbred racing horses noted that the pace tended to be dominant to the trot and that no matter the gait breeding scheme, some pacing offspring were always observed (Harrison and Baldwin 1968). Pacer by trotter crosses resulted in 100% pacer offspring, whereas pacer by pacer crosses resulted in 99% pacer offspring and 1% trotter offspring, while trotter by

trotter crosses often resulted in pacer offspring (Harrison and Baldwin 1968). However, Cothran *et al.* reported in 1987 that only 20% pacer offspring resulted from trotter by trotter crosses. These seemingly contradictory reports maybe detecting the divergence of pacing and trotting Standardbreds into two separate breeds, first noted in 1983 after examining the loci for Equine Lymphocyte Antigen (Bailey), and confirmed in additional inbreeding studies (Cothran *et al.* 1984; MacCluer *et al.* 1983). In the late 19th and early 20th centuries it was common practice to breed trotters and pacers together, but modern breeding practices have shifted towards breeding trotters to trotters, and pacers to pacers (Cothran *et al.* 1987).

A recent publication describes a mutation in the *DMRT3* (doublesex- and mab-3-related transcription factor 3) gene which the authors claim controls the ability for a horse to perform lateral patterned gaits (Andersson *et al.* 2012). The same group documented that six other gaited breeds are nearly fixed for the homozygous mutant (Andersson *et al.* 2012), although additional genotyping of 141 breeds from throughout the world identified the *DMRT3* mutation at various frequencies in both gaited and non-gaited breeds (Promerová *et al.* 2014). However, breeds in this report were classified as gaited based on a 50% or greater frequency of the mutation and not based on observation of the gait phenotype. Horses with the mutation are expected to produce a truncated version of the DMRT3 protein, yet mRNA expression levels appear the same between mutant and wild-type horses (Andersson *et al.* 2012).

In the *DMRT3* mouse knockout, gait analysis revealed increased stride length and swing times, and incoordination between front and hind limbs (Andersson *et al.* 2012), but no instances of lateral-coupled footfalls were observed by the authors. In the Icelandic horse, the pace has a tendency for significant asymmetry between subsequent flight phases, resulting in a broader distribution of footfall ratios (Robilliard, Pfau, Wilso 2007). For this reason, the fifth gait of the Icelandic is often called the "flying pace" (USIHC), as it can be a broken two-beat gait where one

foot lands before the other and not together as occurs in a true pace. Both four- and five-gaited Icelandic horses exhibit the "tölt", a four-beat lateral gait (Ziegler 2005), yet only 31% of the four-gaited horses possess a homozygous AA genotype at *DMRT3* (Andersson *et al.* 2012). Homozygous AA genotypes are observed in both trotting and pacing Standardbred racehorses (Andersson *et al.* 2012), suggesting the mutation likely controls the transition from an intermediate gait into the canter, or the ability to coordinate the three-beat diagonal-coupled footfall pattern at high speeds. This hypothesis is supported by the observation in Andersson *et al.* (2012) that "Icelandic horse homozygous mutants had inferior scores for gallop..." and examination of the supplementary tables also shows differences ($p=0.07$) between C/- and A/A horses for 'slow gallop'. Most of the American gaited breeds can perform their intermediate gait at speeds up to 28 mph (Dash) or 12.5 m/s and have trouble performing the canter (Ziegler 2005).

The *DMRT3* mutation is likely fixed in the cited US breeds for two reasons. First, these breeds share a common ancestry. Both the Paso Fino and Peruvian Paso derived their ambling gait from the Spanish Jennet, a breed introduced to the Americas in 1493 (Hendricks 1995). All of the American gaited breeds are likely descendants of these Spanish horses; the Missouri Fox Trotter, Tennessee Walking Horse and Mountain Horse breeds all developed from the same family lines in the early 1900s, admixing and developing into breeds based on their geographical regions (Hendricks 1995). Second, the US breeds used are not often evaluated for quality of the three-beat canter in the show ring. Thus, possessing the *DMRT3* allele could lead to more stance time spent in lateral-coupled footfall, which could be an advantage in US gaited horse competitions.

Conformation

Conformation is the overall appearance of the horse, describing how the horse is put together or how all the smaller structures of the horse (i.e. legs, body, neck, and head) fit together. Conformation is largely driven by the skeleton, followed by muscle type, and amount of fat.

Fitness level and age can alter conformation; a young horse ridden everyday will have larger muscle mass and definition and less fat, while an older unexercised horse may have more fat, a large drooping belly, a sway back and no muscle definition.

Conformation has long been a driving force in horse selection and breed identification. The earliest written account evaluating conformation was *De re equestri,* written by the Greek historian and philosopher Xenophon (430-354 BC; cited by (van Weeren 2013)(van Weeren 2013)(van Weeren 2013)(van Weeren 2013)(van Weeren 2013)(van Weeren 2013)(van Weeren 2013)(van Weeren 2013)(van Weeren 2013)(van Weeren 2013)van Weeren 2013). Xenophon provided empirical opinions on the suitability and sturdiness of specific conformation features (van Weeren and Crevier-Denoix 2006). Today, conformation is of particular interest to the racing industry as a predictor for race winnings (Smith, Staniar, Splan 2006; Suontama *et al.* 2013; Weller *et al.* 2006b) and for injury susceptibility (Dolvik and Klemetsdal 1999; Magnusson and Thafvelin 1990; McIlwraith, Anderson, Sanschi 2003; Weller *et al.* 2006a). In any particular discipline there are certain conformation and body dimensions that are reported as desired and advantageous to performance (Back, Schamhardt, Barneveld 1996; Deuel 1995; Dolvik and Klemetsdal 1999). For example, good front limb action is determined mostly by leg and foot stances, slope of the shoulders and pasterns, and length of the leg (Saastamoinen and Barrey 2000). Limb conformation is the major factor in limb soundness and can be a predictor or warning of future unsoundness, especially considering that the front limbs bear 60-65% of the horse's weight (Stashak 1987). Locomotor problems and lameness are often the most common reasons for culling horses and for training problems (Bergsten 1980; Jeffcott *et al.* 1982; Linder and Dingerkus 1993; McGreevy and Thomson 2006; Philipsson *et al.* 1998).

The Skeleton, Muscles and Their Interaction: Biomechanics of Movement

Locomotion assessment requires understanding the interplay of conformation components. The skeleton provides the overall structure and rigidity for muscles, tendons and ligaments to attach and apply forces for movement or stabilization of the joints, with the length of the bones affecting the leverage the muscles/tendons/ligaments can apply during locomotion (Clayton 2004). Bones articulate within the joints, with the bone shapes dictating the amount and types of movement. Tension in the ligaments also limits the amount of movement by stabilizing the joint to prevent dislocations (Clayton 2004). The muscles initiate and control the movements of the limbs, counterbalanced by the action of ligaments and tendons. In the lower limbs, ligaments prevent all the forward flexion, while tendons store energy during stance phases so when the limb moves out of loading/stance phase, the stored energy is released and allows the tendon to recoil. This tendon recoil reduces the amount of energy that must be supplied by active muscle contraction (Clayton 2004). In the body, flexion in the vertebrae and tension in the dorsal ligaments and rectus abdominus and lognissimus dorsi muscles can change the body position of the horse, along with head and neck position. When tension in the ligaments and both muscle groups are increased, the back rises and is said to be in a round position (Ziegler 2005); this also occurs when the horse lowers their head and neck. When the tension is released and the head and neck are held high and extended, the back sags downward into a hollow position (Higgins 2011; Ziegler 2005). With moderate resting tension, and the head and neck carried at a medium height with no flexion, the back is carried in a neutral position (Ziegler 2005).

Front leg motion is determined by the interaction of the shoulder with the humerus bone. The shoulder, or more specifically the scapula, rotates fore and aft, lying against the ribcage. Muscles that move the scapula are attached along the dorsal and ventral ends on the medial and lateral sides of the bone (ref). Several muscles are involved in the movement and stabilization of

the scapula and humerus. In general to extend the limb and lift the knee, the dorsal end of the scapula rotates caudally causing the humerus to move into a more vertical orientation, pulling up the radius and ulna into a horizontal position. As the limb moves from extension to stance to flexion, the dorsal end of the scapula moves cranially causing the humerus to move into a more horizontal position and pushing down the radius and ulna.

A long sloping shoulder is commonly considered an advantage for expressive movement of forelimbs (Ehrengranat 1918; Holmström 2001; Sellet, Albert, Groppel 1981; Van der Veen 1918). Henninges (1933) identified a positive correlation between a sloping shoulder and stride length at the walk, and Back *et al.* (1996) found that a sloping scapula was correlated to a more protracted forelimb. Subjectively, it is very difficult to correctly estimate the real slope of the shoulder due to the discrepancy between the external outline of the shoulder and the real inclination of the scapula (Holmström and Back 2013). While the shoulder governs the entire forelimb, the humerus determines the folding of the elbow, knee and fetlock joints (Bennett 2012). A long humerus results in longer and more sweeping steps, while a short humerus results in short and choppy strides (considered short if the stride only covers half or less of the shoulder length) (Bennett 2012; Clayton 2004; Holmström, Fredricson, Drevemo 1994). A steeper angled humerus results in higher knee action, while a more horizontal humerus results in lower knee action (Bennett 2012; Holmström and Back 2013). The length of the humerus showed the strongest correlation to good gaits in 4-year-old riding horses (Holmström and Philipsson 1993; Holmström and Back 2013) and elite dressage horses have been shown to have a significantly longer humerus than both show jumpers and non-elite or 'normal' horses (Holmström and Back 2013). In Thoroughbred racehorses, the slope of the humerus has shown to be of significant importance for the efficiency in stride (Holmström and Back 2013) where a vertical humerus limits the forward reach of the forelimb and reduces the stance duration of the forelimbs. This can be a critical limitation in an athletic horse; a horse cantering or

galloping can only exhale during the stance phase of the front limbs. Therefore, if the stance phase becomes too short, the horse will be unable to exhale the same volume it inhales resulting in reduced respiratory capacity (Holmström and Back 2013).

Hind leg motion is determined by the length of the femur and tibia, and the angles of the hip, stifle, and hock. A long and forward sloping femur places the hind limb more under the horse, allowing the horse to keep its balance more easily and carry more weight on the hind limbs from a position closer to the center of gravity (Holmström and Back 2013). The angle of the stifle is important for hindlimb strength. A small stifle angle, less than 153º (angle of femur in relation to tibia/fibula), results in significant strain to the quadriceps femoris muscle. The quadriceps femoris muscle extends the stifle and already undergoes significant strain when horses are worked in collected gaits (flexion in the lumbrosacral joint with rounding of the back and neck/poll in bascule). If the muscles cannot 'lock' the stifle in an extended position when maximum load is put on the hind limbs, the horse must transfer weight to the forelimbs (Holmström and Back 2013) which already support ~60% of the horse's weight. A small stifle angle inhibits the ability of the hindlimb to withstand the large load generated at higher speeds (as in racehorses) and store elastic energy at the same time (Holmström and Back 2013). In general, the shorter and straighter the hind limb, the more efficiently it can deliver thrust, generated by the muscles of the horse's rump, downward to the ground (Bennett 2012). The longer a horse's hind limb relative to his croup height, the easier it is for him to bring his hooves forward for a longer step; this is especially noticeable in gaited horses (Bennett 2012).

The slope of the hind pastern is important for overall soundness in the horse; a significantly steeper pastern, ~156º, was found in riding horses with soundness problems (Holmström and Back 2013). This may be a secondary effect due to a smaller hock angle (Holmström and Back 2013), as a straight hock is significantly correlated with a more sloping pastern (Magnusson 1985). A

straighter hind fetlock is correlated to a longer stride and swing duration at the trot (Back, Schamhardt, Barneveld 1996) and a straighter hind fetlock is more efficient in storing elastic energy, which contributes to more power from the hind limb (Holmström and Back 2013).

Performance Longevity

Equine biomechanics has been an active area of study for thousands of years. Aristotle (384-322 BC) was the first to describe the various gaits of the horse, accurately describing the walk in his works *De mou animalium* and *De incessu animalium* (Back and Clayton 2013). Goiffon and Vincent were the first to represent horse gaits in a gait diagram, which is still used today (Back and Clayton 2013; Goiffon and Vincent 1779). One of the earliest studies involved the use of an electrical device connected to bells to measure hoof mechanics by Bayer (1882; Back and Clayton 2013); however, most of these early studies were based on theoretical ideas since the human eye cannot determine the stance durations of the faster gaits.

Groundbreaking experimental data was provided by Eadweard Muybridge and Etienne Marey in the 1880s. In 1872, Muybridge was hired by Leland Stanford to determine if there was a moment of suspension in the trot performed by Stanford's trotter 'Occident'. Initial attempts were unsuccessful due to the lack of a fast shutter speed; however, by 1877 Muybridge was able to carry out the experiment (Back and Clayton 2013). Muybridge setup a series of 24 cameras that took still images when the horse tripped a wire attached to each camera; from these images, suspension and stance phases were observed. Muybridge also invented the 'zoöpraxiscope', a device that consisted of a large glass disc on which successive pictures could be printed then projected rapidly on a screen, giving the impression of a moving picture; Thomas Edison, credited with invention of motion pictures and video as we know it today, derived his basic ideas from Muybridge and his invention (Back and Clayton 2013).

While Muybridge was the first to photograph locomotor patterns, Mayer was the first to accurately discriminate between stance and swing phases. Mayer used rubber pressure-sensitive horse shoes and bracelets that, when compressed, recorded the air pressure differences on a charcoal-blackened rotating cylinder recording the footfall sequence (Back and Clayton 2013). Mayer's calculations of how long each foot remained on the ground were too short, but the sequence order was correct (Leach and Dagg 1983; Back and Clayton 2013).

Studies in the early 1900s focused on identifying relationships between conformation and gait, and understanding the kinematics of the different gaits and muscle function. Bethcke compared various breeds of trotters to identify correlations between morphological data and performance, and was able to identify some anatomical differences among various breeds. Yet, Bethcke concluded that gait performance is primarily determined by factors other than conformation such as training, character, pedigree, and fitness (1930; Back and Clayton 2013). Other studies compared bone/segment lengths and joint angles to analyze performance, and some found skeletal lengths to be related to stride length (Franke 1935), while others were unable to identify a particular conformation feature and drew similar conclusions to Bethcke (Kronacher and Ogrizek 1932; Schmidt 1939; Wagener 1934; Wehner 1944). Studies on the kinematics of the fore and hind limbs were performed prodigiously by Wilhelm Krüger who mounted a camera on a vehicle that moved alongside the horse at the same speed to collect video (1937; 1938). He developed a correction for the artifacts in the data associated with using markers placed on the skin (due to the skin shifting around the joints of the horse),and was quick to blame a lack of this correction for discrepancies between his and others' data (Back and Clayton 2013).

After the Second World War and the increase in mechanized labor, both in agriculture and the military, the horse fell into decline and few studies of its conformation were conducted (van Weeren 2013). Around the 1960s, as the economy grew and prosperity spread through developed

countries, interest was renewed in the horse (van Weeren 2013). The horse was now more readily accessible to the masses rather than the privileged few, resulting in the spread and development of new equine disciplines and competitions. With the renewed interest, scientific investigations into gait analysis and performance prediction flourished.

Today treadmills, accelerometers and computers with motion capture/analysis software are predominant in the classification of gaits (Barrey *et al.* 2001; Barrey *et al.* 2002; Nicodemus, Holt, Swartz 2002; Nicodemus and Clayton 2003; Splan and Hunter 2004). These studies classify gaits based on factors such as footfall pattern and timing, limb support sequences, and speed. So far 18 intermediate gaits have been defined according to breed descriptions of the gaits (Nicodemus and Clayton 2003). Advantageous conformations for sport performance have been identified. Long, forward sloping femurs, a sloping shoulder, a long humerus, large fore pasterns and a flat pelvis were more common among elite dressage and show jumpers compared to average performers (Holmström, Magnusson, Philipsson 1990; Langlois *et al.* 1978).

Height at withers, length of the pelvis and a long neck were found to be advantageous for jumping ability, and straight stifle angles and a wide distance between the wing of the atlas and the mandible advantageous for elite dressage horses (Holmström 2001). Another study found that dressage horses had a longer back than show jumpers, likely related to the required suppleness of dressage horses (Johnston *et al.* 2004). Using skin markers, the angles of the proximal limbs (scapulohumeral angle and coxofemoral angle) were identified as being discriminative for good performance in free jumping horses (Dufosset and Langlois 1984; van Weeren and Crevier-Denoix 2006).

Love *et al.* (2006) found a negative influence of various faulty conformations ('back at the knee', 'turned out') on a number of performance parameters in flat racehorses, but given the strong heritability estimates for conformation (0.16-1.00) it's difficult to determine whether these negative

associations were due to conformation per se, or to other sire-related influences (van Weeren and Crevier-Denoix 2006). In National Hunt racehorses, positive associations were identified between intermandibular width, the flexor angle of the shoulder joint and the coxofemoral angle with performance data (Weller *et al.* 2006a). However, there are no significant differences in the prevalence of mild deviations of optimal conformation in elite horses compared to a population performing at a low level (Holmström, Magnusson, Philipsson 1990).

There is a fine and complex balance to using conformation as a predictor for soundness; one conformational feature may be both beneficial and detrimental depending on how the trait is analyzed. For example, horses with large tarsal angles absorbed less concussion during the impact phase, which may be a factor in the development of degenerative joint disease (Gnagey, Clayton, Lanovaz 2006), but the smaller net joint moment may reduce the risk of plantar ligament desmitis (Baird and Pilsworth 2001; Back and Clayton 2013). Horses with recurrent lameness tend to have a more vertically positioned femur and steeper pasterns, while horses with back problems tend to have a small hock angle (Holmström 2001). Horses with short backs have stronger backs, but there may be more interference between the front and hind limbs (Pritchard 1965). Interestingly, horses with carpus valgus proved to have fewer carpal fractures and carpal effusion, but offset knees contribute to fetlock problems, and long pasterns increase the risk of front limb fractures in Thoroughbreds (Anderson, McIlwraith, Douay 2004).

LCORL/NCAPG – Candidate Genes For Skeletal Size

While heritability estimates for conformation traits, such as height, are relatively high in the horse (Saastamoinen and Barrey 2000), few genetic studies were attempted due to the complex nature of conformation. The introduction of the high density SNP genotyping chips led to the first studies of conformation in humans. Like the horse, humans have a high heritability estimate for overall height of around 0.8 (Visscher, McEvoy, Yang 2010) and meta-analysis have identified

several hundred size loci (Cho *et al.* 2009; Gudbjartsson *et al.* 2008; Kim *et al.* 2010; Lango Allen *et al.* 2010b; Lettre *et al.* 2008; Okada *et al.* 2010; Soranzo *et al.* 2009; Visscher, McEvoy, Yang 2010; Wood *et al.* 2014). Control of human size is mediated by a large number of genes with very small average effects (Perola 2011), with the majority of the identified loci only accounting for a small 15% of the total height variation (Lango Allen *et al.* 2010a). *LCORL* was identified in several of these studies as a candidate size locus (Lango Allen *et al.* 2010a; Perola 2011), and is implicated in birth weight and growth development (Horikoshi *et al.* 2013).

A single gene, *IGF1*, explains ~10-15% of body size variation in the domesticated dog (Hoekstra *et al.* 2010; Sutter *et al.* 2007) and the majority of dog breed-average mass can be explained by as few as six loci (Boyko *et al.* 2010). Several independent studies have identified loci related to height and conformation in cattle (Karim *et al.* 2011; Pausch *et al.* 2011; Pryce *et al.* 2011). These studies have identified approximately 8 loci significantly associated with height, explaining up to 20% of the phenotypic variation (Pryce *et al.* 2011). In several different cattle breeds, a non-coding regulatory mutation in the promoter region of the bovine *PLAG1* gene was found to explain around 1-3.5% of the phenotypic variance (Karim *et al.* 2011).

Within the horse, three separate GWAS studies identified the *LCORL* and *NCAPG* genes as likely candidates for size differences, both across and within breeds (Makvandi-Nejad *et al.* 2012; Metzger *et al.* 2013; Signer-Hasler *et al.* 2012). There is little available literature on the biological function of *LCORL*, but based on a mouse study it is thought to be a transcription factor that functions during spermatogenesis in testes. *LCORL* showed the highest expression level in testes, but could also be detected in the kidney, liver, heart, and brain (Kunieda *et al.* 2003). The *NCAPG* gene overlaps with *LCORL* and encodes a subunit of the condensing complex, which is responsible for the condensation and stabilization of chromosomes during mitosis and meiosis (Jager *et al.* 2000; Murphy and Sarge 2008). In cattle, a mutation in *NCAPG* effects carcass

conformation, growth traits, and feed efficiency (Lindholm-Perry *et al.* 2013; Setoguchi *et al.* 2011). Selective sweeps in sheep identified *LCORL/NCAPG* as candidate regions for production traits (Kijas 2014).

Behavior

Horse behavior is a key factor in their successful domestication. While the horse is a prey animal with a natural flight over fight response, they are social animals with an instinctive understanding of dominance, submission and curiosity, and are relatively non-territorial (Budiansk 1997). They are also generalist plant feeders and able to adapt to a wide variety of climates and circumstances (Budiansk 1997).

Among horse owners and caretakers there is anecdotal evidence for breed typical behavior and temperament. Within breeds, there are anecdotes supporting familial typical behaviors and temperament. Several studies have examined different individual behavior types in the horse, and many have found associations with breed and specific behavior traits (Stashak 1987), likely due to the diversity of disciplines for which a single breed can be utilized and were originally bred (Hausberger *et al.* 1998; McGreevy 2004). For example, Icelandic horses are said to have a good homing instinct, but it is not clear whether this is due to advanced skills or a heightened motivation to return to their home range (McGreevy 2004). In a comparison of 30 temperament traits across eight different horse breeds, anxiousness and excitability showed the greatest variation (Lloyd *et al.* 2008), further supporting the idea that selection for use has created diversity in behavior and temperament. However, differences in temperament across breeds may also be due to underlying physiological differences such as circulating levels of serotonin (Bagshaw, Ralston, Fisher 1994). In other species, breed differences in temperament have been identified (Clutton-Brock 1999; Lloyd *et al.* 2008; Notari and Goodwin 2007; Svartberg 2005; Svartberg 2006; Wolff, Hausberger, Le Scolan 1997) and some of these are attributed to underlying genetic role changes (Notari and Goodwin

2007; Svartberg 2006; Wolff, Hausberger, Le Scolan 1997). In horses, sire influences have been identified in emotionality or nervousness (Houpt and Kusunose 2000; Wolff, Hausberger, Le Scolan 1997) and in the tendency to develop stereotypic behaviors (Wolff, Hausberger, Le Scolan 1997). Crib-biting is a stereotypic oral behavior in horses with a suggested genetic susceptibility, but previously known stereotypic genes are not major risk factors for the trait (Hemmann *et al.* 2014). In cattle, most behavior genetics have focused on feeding and reproductive behavior, rather than temperament, but increased production selection has resulted in increased compulsive disorders such as bar-biting and excessive licking (Adamczyk *et al.* 2013). Additionally, aggressive behavior has been associated with variations in the *MAOA* (monoamine oxidase A) gene in humans, mice and rhesus monkeys (Brunner *et al.* 1993; Cases *et al.* 1995; Craig 2007; Karere *et al.* 2009; Popova *et al.* 2001), but not cattle (Lühken *et al.* 2010). MAOA is often used as an anti-depressant due to its role in the breakdown of serotonin (Geha *et al.* 2002).

Behavior Assessments

There are two types of behavior assessments for horses currently in use. The first is typically comprised of a battery of tests in which an animal is presented with a variety of stimuli and behavior responses assessed by a human observer. Occasionally physiological biomarkers are quantified. These involve analyzing a horse's behavior by scoring the animal's responses to certain stimuli (Hauseberg and Muller 2002; Jezierski, Jaworski, Górecka 1999; McCann *et al.* 1988; Momozawa *et al.* 2003); recording autonomic functions such as heart rate (HR) and respiration (Hada *et al.* 2001; Jezierski, Jaworski, Górecka 1999; McCann *et al.* 1988; Momozawa *et al.* 2005a; Visser *et al.* 2002); determining endocrine function by measuring plasma concentrations of hormones or monoamines (Alexander and Irvine 1998; Anderson *et al.* 1999; Hada *et al.* 2001); or examining behavioral parameters (Le Scolan, Hausberger, Wolff 1997; Visser *et al.* 2001; Wolff, Hausberger, Le Scolan 1997). The second method of studying equine temperament uses a questionnaire survey

completed by caretakers or handlers who are familiar with the target animals (Anderson *et al.* 1999; Le Scolan, Hausberger, Wolff 1997; Momozawa *et al.* 2003). While trials are advantageous in that an unknown animal can be examined objectively, it might be difficult to relate an animal's behavior responses in a specific circumstance back to its temperament and to any behavior problems arising in ordinary care, and may not be a true indication of an animal's character (Momozawa *et al.* 2003; Seaman, Davidson, Waran 2002). Additionally, behavior tests are difficult to perform on a large number of animals while questionnaires are less time intensive and therefore easier to use for large sample sizes. A questionnaire may reflect a respondent's personal impressions and experiences and is thus, subjective; however, the questionnaire, in theory, may be more accurate as it is based on long term observation rather than a "snap shot" in time and could include several innate emotional traits that are stable across time (Lloyd *et al.* 2007; Momozawa *et al.* 2003).

The types of behavior tests that have been reported in the horse include novel object tests, novel area tests, tactile pressure, restraint, handling, and avoidance learning (Forkman *et al.* 2007). In novel object tests, the animal is often confronted with a potentially fear-producing stimulus, such as a sound (pots and pans dropped, puff of air, or popping a balloon), sights (umbrella, a toy, or plastic bags that are either stationary or moving), or footing (a plastic tarp or wooden bridge). Novel area tests are basically conducted as an open-field test or by socially isolating an animal. Behavioral measurements include head, tail and ear carriage, latency to approach novel stimulus, tactile contact with novel stimulus, number of steps taken away from stimulus, gait of movement (walk, trot, canter), frequency of rearing or bucking, total distance traveled, vocalizations (whinnies versus snorts), latency to return to stimulus or area where the stimulus was presented, and number of defecations or urination events. Most behavioral evaluations are used to assess sociability or neophobia/emotionality.

Predictions for Performance

Animals used for a variety of specific tasks often have different characteristics from each other, and it is important to identify the physical abilities, conformation and personality traits that are needed to match the job. Experienced owners are often able to quickly assess the ideal conformation traits; however, less emphasis is often placed on examining personality or temperament (Evans *et al.* 1977; Visser *et al.* 2003b). Few equine studies have investigated how to relate personality with measures of performance (Visser *et al.* 2003b). In terms of poor performance ability, clinical disorders are most often sought (Cator 1991; Hernandez and Hawkins 2001; Martin *et al.* 2000) rather than temperament (Visser *et al.* 2003b). The expectation of good performance is, apart from pedigree, often based on the horse's conformation (Visser *et al.* 2003b).

In humans and dogs, extensive research has been reported in the area of performance ability and personality/temperament. In dogs, a puppy walker's assessment of a young dog's temperament and behavior traits, as obtained through a 40-question survey, was found to be an effective predictor of the dog's suitability as a guide (Momozawa *et al.* 2003; Serpell and Hsu 2001). In humans, extroverts are thought to be more venturesome, inclined to take more risks and they are more in pursuit of a competitive situation (Visser *et al.* 2003b). Athletes are found to be more extroverted when compared to non-exercisers (Visser *et al.* 2003b). Additionally, anxiety has proven to be negatively related to sport performances due to interference with the athlete's concentration on the task (Egloff and Jan Gruhn 1996; Visser *et al.* 2003b).

In horses, overly nervous yearlings tend to have higher overall activity indexes and higher heart rates than normal yearlings, which also influence their frequency of eating and drinking, defecation, and contact with other horses (McCann *et al.* 1988). The level of emotionality (or nervousness) also influences learning ability (McGreevy 2004) and is relatively stable across time (Lansade, Bouissou, Erhard 2008). Early detection of emotionality as an outcome of avoidance

learning tests gives a reliable prediction for the future and is more consistent than responses in reward learning tests (Visser *et al.* 2003a). Therefore, these tests can be applied to help determine individual temperament profiles and discipline suitability at an early age, potentially saving time and money.

The Nervous System and Behavior

One organizational principle of the nervous system is the use of parallel processing in which sensory, motor and cognitive functions can be served by more than one pathway (Kandel, Schwartz, Jessel 2000). The brainstem and thalamic reticular activating system provide arousal and initiate attention, and the caudal part of the forebrain integrates perception. The prefrontal cortex plays a very significant role in the coordination of goal-directed behavior and evaluates the effect of such behavior via reinforcement/punishment outcomes (McGreevy 2004). The frontal lobe is also responsible for the conscious initiation of movement, after which the CPGs take over rhythmic control.

Neurotransmitters play an important role in the nervous system, especially in relation to behavior. As messengers, they communicate information throughout the brain and body via relays between presynaptic and postsynaptic neurons. Abnormal levels of neuropeptides have been shown to be involved in human psychiatric disorders, and differences in the relative amount and location of neurochemicals can influence behavior (McGreevy 2004). There are three major classes of neurotransmitters: biogenic amines, amino acids, and neuropeptides.

The biogenic amines include dopamine, adrenaline, noradrenaline, serotonin, acetylcholine and melatonin. As determined by human medicine studies, dopamine plays an important role in motor and mood control. A strong association has been found between high levels of detachment and the dopaomine D2 receptor (Breier *et al.* 1998), and drugs that bind to this receptor have anti-aggressive effects in human patients (Hector 1998). Two SNPs identified in the equine dopamine

D4 receptor were associated with curiosity and vigilance scores as determined by a questionnaire (Momozawa *et al.* 2005b). The dopaminergic systems are also involved in the brain's reward system, leading to addiction to nicotine, amphetamines, and cocaine, and play an important role in many stereotypies (Dodman 1998), such as cribbing in horses (Rendon, Shuster, Dodman 2001). Adrenaline is the cause of sweating and increased heart rate when horses are stressed or excited (McGreevy 2004). Noradrenaline is more abundant in the brain than adrenaline, and its neurons control cardiovascular and endocrine functions. In humans, noradrenergic activity is correlated with aggressive behavior and specific adrenergic receptor blockers have beneficial effects in violent patients (Elliot 1977). Serotonin (5-HT) or its precursor 5-hydroxytryptophan (5-HTP) may induce sleep and reduce anxiety. The 5-HT system is the mediator of learned and sustained fear responses, and decreased 5-HT levels have been associated with violent psychopathological behavior in humans (Lee and Coccaro 2001). A decrease in serotonergic transmission leads to an inability to adopt passive or waiting attitudes, or to accept situations that necessitate or create strong inhibitory tendencies (McGreevy 2004). Stereotypic behavior patterns in human medicine, often grouped in a classification known as obsessive-compulsive disorders, appear to be partially attributable to decreased 5-HT and abnormal endorphin. This may be particularly relevant to ritualistic and stereotypic behaviors, including wind-sucking and cribbing in horses recognized in veterinary medicine (Overall 1998). Acetylcholine is one of the principle neurotransmitters involved in the propagation of aggressive, predatory behavior and is important in regulating wake and sleep cycles (McGreevy 2004). It has been shown to increase the strength of synaptic connections in the hippocampus, as cholinergic enhancement has been shown to improve memory performance in humans (Furey, Pietrini, Haxby 2000).

Amino acid neurotransmitters are the most abundant in the brain. Glutamate is the major excitatory neurotransmitter. There are five types of glutamate receptors of which N-methyl-D-

aspartate (NMDA) is the best understood because it may play a crucial role in learning and memory as well as in aggressive and defensive behavior (McGreevy 2004). The major inhibitory neurotransmitter is γ-aminobutyric acid (GABA). GABA is the primary neurotransmitter in intrinsic neurons that functions as local mediators for inhibitory feedback loops

Neuropeptides are very different from the other classes of neurotransmitters. They occur at a much lower concentration in the body and more commonly act as modulators of pre- or postsynaptic transmission. They are also synthesized in the cell body and then transported down the axon, unlike the other classes that are typically synthesized at the synapse. Neuropeptides thus take a long time to replenish, cannot be recycled and have a longer duration of action than the other classes (Cooper, Bloom, Rothe 1996).

The Equine Genome

The horse genome consists of 64 chromosomes: 31 pairs of autosomal chromosomes and two sex chromosomes, X and Y. Early equine genetic studies were based on utilizing comparative, cytogenetic fluorescent in situ hybridization (FISH), radiation hybrids, and linkage mapping (Chowdhary *et al.* 2003; Lear *et al.* 2001; Leeb *et al.* 2006; Milenkovic *et al.* 2002; Penedo *et al.* 2005; Raudsepp *et al.* 1999). These methods were successful in identifying several simple genetic traits in the horse, such as hyperkalemic periodic paralysis (HYPP) in Quarter Horses, severe combined immunodeficiency (SCID) in Arabians, overo lethal white syndrome (OWLS), and tobiano coat color (Bernoco and Bailey 1998; Brooks *et al.* 2007; Rudoph *et al.* 1992; Santschi *et al.* 1998; Shin, Perryman, Meek 1997; Yang *et al.* 1998).

In 1990, the human genome project was initiated. This was a major windfall for the equine genomics community. After sequencing of the human genome was completed in 2000, the National Human Genome Research Institute (NHGRI) wanted to sequence 24 additional mammals to identify regions of conservation and diversity shared with the human genome. One of the major

surprising discoveries of the human genome sequence was that it only contained approximately 20,000 genes although previous estimates indicated that there would be 100,000 to 300,000 genes. Additionally, only 2% of the human genome was found to code for genes; the remaining portions were of unknown function.

The horse sequencing project was selected by the NHGRI in 2005, with shot-gun sequencing completed in 2007 by the Broad Institute of MIT and Harvard University in Boston, Massachusetts. Twilight, a Thoroughbred mare that is part of a research herd at Cornell University was selected for complete genome sequencing. The genome assembly consists of 2.43 billion base pairs, with the 6.8x coverage of the initial assembly (EquCab1.0) released in January 2007, and the current assembly (EquCab2.0) released in September 2007 (Wade *et al.* 2009). Little information was available about the genes expressed in the horse, so the horse genome was annotated using transcriptome information from other species, with estimates of 20,322 genes by ENSEMBL and 17,610 genes by NCBI (Coleman *et al.* 2010).

Since the initial genome sequencing project, rapid advancements were made in genetic sequencing technology and data analyses largely due to the human genome project (Shendure and Ji 2008). Next-generation sequencing improved data quality, is faster and cheaper to run than previous Sanger methods, and allows for the detection of SNPs, novel genes or regulatory elements, assessment of transcript expression levels and identification of alternative splice variants and transcription start and stop sites (Bahassi el and Stambrook 2014). Next-generation sequencing of a Quarter Horse mare demonstrated this with improved sequence coverage at 25X and the identification of 3.1 million SNPs, 193,000 INDELS, and 282 CNVs (Doan *et al.* 2012). The downside to next-generation sequencing is that it does not assemble well for species without a reference sequence, and results can be biased based on filtering parameters (Snyder, Du, Gerstein 2010). Sequencing of equine mRNA (RNA-seq) from various horse tissues has also identified

several unannotated equine-specific transcripts in the brain, laminae, and other equine tissues (Coleman *et al.* 2013; Holl *et al.* 2014; McGivney *et al.* 2010; Ouzounis *et al.* 2013). A collaborative effort is currently underway to update the current assembly with improved annotation (EquCab3.0) from transcriptome and genome sequence data (Kalbfleisch *et al.* 2014).

High Throughput SNP Genotyping Assays (The "SNP Chip")

In addition to the whole genome sequencing of Twilight, DNA samples from seven other horses representing breeds from diverse origins were sequenced at random sites to identify 1.5 million SNPs (Wade *et al.* 2009). These breeds included the Akhal-Teke, Andalusian, Arabian, Icelandic, Quarter Horse, Standardbred, and Thoroughbred. The SNPs are identified in the online database EquCab2.0 (horse_snp_release/v2/) from the Broad Institute (Broad Institute of MIT and Harvard) and in dbSNP from NCBI (www.ncbi.nih.gov/snp/). These SNP resources enabled the development of the Illumina EquineSNP50 chip (Brooks *et al.* 2010a; McCue *et al.* 2012) (Illumina, Inc, San Diego, CA). The Illumina SNP chip is based on single-base extension sequencing method. The DNA sample is isothermally amplified, then fragmented and hybridized to the beadchip by annealing to locus-specific 50 bp long probes that are covalently linked to a microbead, and then extended by one base (Steemers *et al.* 2006). The products are fluorescently stained and the emission intensities of each wavelength used to determine the SNP genotype.

The first successful application of the EquineSNP50 chip was Lavender Foal Syndrome (LFS), a recessive disorder characterized by multiple neurological abnormalities and a dilute coat color and often seen in the Arabian horse. Using the SNP chip, Brooks *et al.* (2010a) was able to identify a region associated with LFS with two candidate genes. Sequencing of one of the candidate genes identified a single mutation that changes the amino acid sequence to a shorter protein sequence. Since then, several other genetic studies have successfully used the SNP chip (Andersson *et al.* 2012; Boyko *et al.* 2014; Makvandi-Nejad *et al.* 2012; McCue *et al.* 2012; Meira *et al.*

2014; Petersen *et al.* 2013; Signer-Hasler *et al.* 2012). In 2013, a denser version of the chip was released, the equine SNP70 bead chip, also based off of the EquCab2.0 SNP database (NEOGEN 2013). A collaborative effort is currently underway to develop two high-density SNP chips (2 million and 640K) on the Affymetrix platform from the whole-genome sequences of over 163 horses from 32 distinct breeds (Schaefer *et al.* 2014).

Genome Wide Association Study Analysis Methods

The goal of genome-wide association studies (GWAS) is to understand the variation in complex traits and diseases by relating genotypes of large numbers of markers (usually SNPs) to observed phenotypes (Ball 2013). GWAS scan SNPs across the whole-genome to detect genes or loci that are associated or linked to the phenotypic variation being investigated. However, most of the SNPs interrogated in GWAS are not functionally important. Most polymorphisms resulting in diseases and traits are rare because they are maintained in natural populations by a balance between mutation and genetic drift, and are lost through random sampling effects. Only a third to half of SNPs are "common" in the sense that the more rare allele is present in more than five percent of individuals (this is why the minor allele frequency is set to exclude any SNPs with less than five percent because it requires a large amount of statistical power to make meaningful statements about very rare alleles) (Gibson and Muse 2004).

Linkage Disequilibrium; Case/Control

Association studies work because of a phenomenon known as linkage disequilibrium (LD), a nonrandom association between alleles. LD is often observed when the proportion of one genotype is greater than predicted or expected. The extent of LD is a function of recombination rates; recombination breaks up LD in proportion to genetic and physical distance between sites and creates haplotype blocks (Daly *et al.* 2001; Gabriel *et al.* 2002; Phillips *et al.* 2003), therefore LD is longer in non-recombining regions. There is great variability in the size of the LD blocks due to

recombination hotspots, regions with high recombination and small LD blocks, and cold spots, regions with low recombination and long LD blocks, allowing for haplotype diversity (Wall and Pritchard 2003). Little is known about the molecular mechanism of recombination hotspots, or how rapidly they appear and disappear over evolutionary time, but several studies have found a negative correlation between levels of LD and rates of recombination implying that recombination rates change slowly on time scales of N generations (Badge *et al.* 2000; Chakravarti *et al.* 1984; Jeffreys, Ritchie, Neumann 2000; Jeffreys, Kauppi, Neumann 2001; Kauppi, Sajantila, Jeffreys 2003; May *et al.* 2002; Schneider *et al.* 2002), (Wall and Pritchard 2003). Based on the theory that there is local LD extending over tens of kilobases (Tozaki *et al.* 2007), mapping can be accomplished by only sampling a few SNPs in an LD cluster at a density similar to the level of LD. The SNPs can either contribute to the trait or, more likely, the SNPs are in LD with the causative polymorphism (Wray *et al.* 2013). Once a cluster has been flagged as associated with the phenotype, more detailed sampling can occur to cover all of the variation in the area; this method has been successful in horses, cattle, dogs, and humans (Goldstein *et al.* 2006; Odani *et al.* 2006; Tozaki *et al.* 2007; Wall and Pritchard 2003; Yang *et al.* 2014a).

Despite the many advantages of association studies, there are pitfalls that must be accounted for in association mapping design. These include small effect size, rarity of the variant, potential biases, and level of LD. If a trait has a low genetic contribution, there is no sample size or design that will be sufficient to map the individual loci (Gibson and Muse 2004). This also coincides with the number of genes that influence the trait; the more genes that impact a trait, the smaller the contribution of each individual gene. Multiple genes contributing to a single phenotype will also lower the penetrance or expressivity of the trait, which will make the trait harder to map. Variable LD across the genome can result in genome sections unscored for their measure of LD; LD is affected by the age of the disease allele, the physical distance between the marker and disease allele,

and the effective population size and stability of the population which do not result in equal LD levels across SNPs, an assumption made in running a GWAS (Zondervan and Cardon 2004).

Additional bias can arise from the sampled population. If there are multiple loci in different pedigrees that affect the same trait, they will not be detected in a mapping panel of unrelated individuals. Yet, population stratification can lead to false positive associations that are an artifact of the substructure. Recent contact between previously isolated populations can also result in transient LD and novel population structure that goes undetected; these effects will change over time as heterozygosity returns to normal. Hidden environmental structure can change the results, especially in respect to diseases with multiple pathologies, and lead to erroneous results. Genotype by environment and genotype by genotype interactions still have an unknown magnitude, are difficult to detect statistically, and can reduce the ability to detect the main effects of susceptibility loci (Gibson and Muse 2004). Incomplete genotyping due to gaps in information from both missing individuals and loci, and any genotyping errors can bypass the causal variant and lead to false positives. Yet, despite all these potential pitfalls, association studies offer one of the best options for providing insights and mapping of complex traits. Validation in an independent population is the only evidence that a significant association detected is truly associated (Hayes 2013).

Population/Sample Structure

As mentioned above, any unaccounted for population structure will result in false positives (Pritchard *et al.* 2000). In livestock populations, selection is for specific breeding goals to create breeds, strains, or lines within the population (Hayes 2013). The presence of related individuals within a study sample creates sample structure, a term that encompasses population stratification and hidden relatedness (Kang *et al.* 2010). Population stratification refers to the inclusion of individuals from different populations within the same study sample, while hidden relatedness refers

to the presence of unknown genetic relationships between individuals within the study sample (Voight and Pritchard 2005; Weir, Anderson, Hepler 2006).

There are several methods that can be used to correct for sample structure as it can be difficult in animal studies to include entirely unrelated individuals. One common approach to eliminate the effects of hidden relatedness is to estimate the proportion of genes identical by descent (IBD) between any pair of individuals in the sample and exclude from the analysis those individuals that appear closely related (Burton *et al.* 2007; Voight and Pritchard 2005). Principle component analysis (PCA) of the genomic data is also widely used in human GWAS to account for population structure (Patterson, Price, Reich 2006) and has been in use long before GWAS (Cavalli-Sforza, Menozzi, Piazza 1994; Menozzi, Piazza, Cavalli-Sforza 1978; Novembre and Stephens 2008; Patterson, Price, Reich 2006). PCA is a method used to identify dependencies and simplify the data to minimize their effect. PCA transforms the data into new variables, termed components, that are truly independent of each other. The first component describes the greatest amount of data variation, with the subsequent components describing decreasing amounts of variation. In this way, PCA can be used to identify patterns in the data and highlight effects of differential bias that require additional control (Price *et al.* 2006). Multidimensional scaling (MDS) is an approach related to PCA in that it is a dimensionality reduction technique, but is a reduction of genome-wide identity-by-state pairwise distances (Price *et al.* 2010; Purcell). Some principal components may represent broad differences across individuals within a given data set, effectively capturing a few major axes of population structure, but it is unclear how to interpret the remaining components as surrogates of sample structure (Novembre and Stephens 2008; Novembre *et al.* 2008). Population stratification and hidden relatedness only account for two extreme manifestations of sample structure, so additional measures need to be applied.

A common approach in animal studies, and gaining notice for application in human studies, are mixed models which can model population structure, family structure and cryptic relatedness (Visscher, Hill, Wray 2008; Yu *et al.* 2006). The basic approach is to model phenotypes using a mixture of fixed effects and random effects (Price *et al.* 2010). Fixed effects include the candidate SNP and optional covariates, such as gender or age, whereas random effects are based on a phenotypic covariance matrix, which is modelled as a sum of heritable and non-heritable random variation (Price *et al.* 2010).

One of these models is the efficient Mixed-model association expedited (EMMAX). EMMAX is based on a linear mixed model (also called a mixed linear model) with an empirically estimated relatedness matrix to model the correlation between phenotypes of sample subjects (Kang *et al.* 2010). Similar variance component approaches have been used successfully in animal models (Kang *et al.* 2008; Yu *et al.* 2006; Zhao *et al.* 2007). However, applying even and efficient implementation of a variance component approach, such as EMMA (Kang *et al.* 2008), is computationally demanding for data sets consisting of thousands of individuals, owing to the heavy computational burden in the estimation of variance parameters, leading to run-times of perhaps years in large datasets (Kang *et al.* 2010; Zhou and Stephens 2012). EMMAX takes advantage of the fact that each locus explains only a small fraction of complex traits, so the estimation of the variance parameters only needs to be done once per data set and can be globally applied to each marker. This computational improvement reduces the running time for the analysis of a typical GWAS data set using a variance component model from years to hours, and the use of the empirical relatedness matrix accounts for a wide range of sample structure including hidden relatedness and population stratification (Kang *et al.* 2010).

A comparable method in speed to approximation methods, such as EMMAX, is genome-wide efficient mixed model association (GEMMA); a method for exact calculations that provides

results similar to EMMA, but generated in a more timely manner (Zhou and Stephens 2012). In some settings the approximate methods provide results almost identical to those from the exact method (Kang *et al.* 2010; Zhang *et al.* 2010), but there is no guarantee the results will be the same and the only way to determine the accuracy of the approximation methods is to run the exact test. Large differences were detected in a mouse data set with pervasive relatedness and large effect sizes, but were negligible in a human data set (Zhou and Stephens 2012). Inaccuracies in the approximation can result in a reduction in power as compared to the exact methods (Zhou and Stephens 2012). GEMMA requires complete or imputed genotype data (Guan and Stephens 2008; Howie, Donnelly, Marchini 2009) for all of the SNPs, but does not require variance components to be the same for all the SNPs (Yang *et al.* 2014b). It uses a single eigen-decomposition of the genetic relationship matrix to rotate the data, thereby removing the structure (Zhou and Stephens 2012; Yang *et al.* 2014b).

In the following chapters, we will describe how we applied current genotyping technologies and methodologies to identify genetic loci contributing to polymorphic gait in the horse. We predict that genes involved in the CPG pathway are responsible for the different gait patterns observed in the horse, and genes controlling skeletal conformation and temperament play contributing roles in the quality of the gait pattern. We hypothesize that the identified genetic loci can be used to predict different gait footfall patterns, skeletal size and thickness, and temperament variation in the horse.

References

Adamczyk K, Pokorska J, Makulska J, Earley B, Mazurek M. 2013. Genetic analysis and evaluation of behavioural traits in cattle. Livestock Science 154(1-3):1-12.

Albertsdóttir E, Eriksson S, Näsholm A, Strandberg E, Árnason T. 2008. Genetic correlations between competition traits and traits scored at breeding field-tests in icelandic horses. Livestock Science 114(2-3):181-7.

Alexander SL and Irvine CH. 1998. The effect of social stress on adrenal axis activity in horses: The importance of monitoring corticosteroid-binding globulin capacity. J. Endocrinol. 157:425-32.

National Economic Impact of the U.S. Horse Industry. [Internet]; c2005 [cited 2013 . Available from: www.horsecouncil.org/national-economic-impact-us-horse-industry .

Anderson TM, McIlwraith CW, Douay P. 2004. The role of conformation in musculoskeletal problems in the racing thoroughbred. Equine Veterinary Journal 36:571-5.

Anderson MK, Friend TH, Evans JW, Bushong DM. 1999. Behavioral assessment of horses in therapeutic riding programs. Appl Anim Behav Sci 63(1):11-24.

Andersson LS, Larhammar M, Memic F, Wootz H, Schwochow D, Rubin CJ, Patra K, Arnason T, Wellbring L, Hjalm G, *et al.* 2012. Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. Nature 488(7413):642-6.

Andersson O and Grillner S. 1983. Peripheral control of the cat's step cycle. II. entrainment of the central pattern generators for locomotion by sinusoidal hip movements during "fictive locomotion.". Acta Physiol Scand 118(3):229-39.

[Internet]; c2006 [cited 2012 . Available from: http://www.abccmm.org.br/regulamentos/regulamentos_1.php?regulamento=57 .

Back W and Clayton HM. 2013. Equine locomotion. 2nd ed. London, UK: W.B. Saunders.

Back W, Schamhardt HC, Barneveld A. 1996. The influence of conformation on fore and hind limb kinematics of the trotting dutch warmblood horse. Pferdeheilkunde 12:647-50.

Badge RM, Yardley J, Jeffreys AJ, Armour JAL. 2000. Crossover breakpoint mapping identifies a subtelomeric hotspot for male meiotic recombination. Hum Mol Genet 9(8):1239-44.

Bagshaw CS, Ralston SL, Fisher H. 1994. Behavioral and physiological effect of orally administered tryptophan on horses subjected to acute isolation stress. Appl Anim Behav Sci 40(1):1-12.

Bahassi el M and Stambrook PJ. 2014. Next-generation sequencing technologies: Breaking the sound barrier of human genetics. Mutagenesis 29(5):303-10.

Bailey E. 1983. Population studies on the ELA system in american standardbred and thoroughbred mares. Animal Blood Groups and Biochemical Genetics 14(2):201-11.

Baird DH and Pilsworth RC. 2001. Wedge-shaped conformation of the dorsolateral aspect of the third tarsal bone in the thoroughbred racehorse is associated with development of slab fractures in this site. Equine Veterinary Journal 33:617-20.

Ball RD. 2013. Designing a GWAS: Power, sample size, and data structure. In: Genome-wide association studies and genomic prediction. Gondro C, van der Werf J,Hayes B, editors. London: Humana Press. 37 p.

Barrey E. 2013. Gaits and interlimb coordination. In: Equine locomotion. Back W and Clayton HM, editors. 2nd. ed. New York: Elsevier. 85 p.

Barrey E, Evans SE, Evans DL, Curtis RA, Quinton R, Rose RJ. 2001. Locomotion evaluation for racing in thoroughbreds. Equine Veterinary Journal Supplemental 33:99-103.

Barrey E, Deslinens F, Poirel D, Biau S, Lemaire S, Rivero JL, Langlois B. 2002. Early evaluation of dressage ability in different breeds. Equine Veterinary Journal Supplemental 34:319-24.

Bayer J. 1882. Experimentelles uber hufmechanismus. Oesterr. Monatsschr. Thierheilk 7:72-4.

Bennett D. 2012. Principles of conformation analysis volumes I, II & III. Boulder, Colorado: Equine Network.

Bergsten G. 1980. The durability of the swedish standardbred riding horse judged from a material of insured horses. 32nd annual meeting of the european association of animal production.Zagreb, Yugoslavia: .

Bernoco D and Bailey E. 1998. Frequency of the SCID gene among arabian horses in the USA. Anim Genet 29(1):41-2.

Bethcke H. 1930. Ist es moglich auf grund der mechanischen verhaltnisse die leistungsfahigkeit einse trabers zu bestimmen? Z. Fur Veterinarkunded 42(5):161-70.

Bolz J and Castellani V. 1997. How do wiring molecules specify cortical connections? Cell Tissue Res 290(2):307-14.

Boyko AR, Brooks SA, Behan-Braman A, Castelhano M, Corey E, Oliveira KC, Swinburne JE, Todhunter RJ, Zhang Z, Ainsworth DM, *et al.* 2014. Genomic analysis establishes correlation between growth and laryngeal neuropathy in thorroughbreds. BMC Genomics 15:259.

Boyko AR, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, Lohmueller KE, Zhao K, Brisbin A, Parker HG, vonHoldt BM, *et al.* 2010. A simple genetic architecture underlies morphological variation in dogs. PLoS Biol 8(8):e1000451.

Breier A, Kestler L, Adler C, Elman I, *et al.* 1998. Dopamine D2 receptor density and personal detachment in healthy subjects - ProQuest. The American Journal of Psychiatry 155(10):1440-2.

Horse Genome Project [Internet] [cited 2014 11/14/2014]. Available from: http://www.broadinstitute.org/mammals/horse .

Brooks SA, Gabreski N, Miller D, Brisbin A, Brown HE, Streeter C, Mezey J, Cook D, Antczak DF. 2010a. Whole-genome SNP association in the horse: Identification of a deletion in myosin va responsible for lavender foal syndrome. PLoS Genet 6:e1000909.

Brooks SA, Makvandi-Nejad S, Chu E, Allen JJ, Streeter C, Gu E, McCleery B, Murphy BA, Bellone RR, Sutter NB. 2010b. Morphological variation in the horse: Defining complex traits of body size and shape. Animal Genetics 41(s2):159-65.

Brooks SA, Lear TL, Adelson DL, Bailey E. 2007. A chromosome inversion near the KIT gene and the tobiano spotting pattern in horses. Cytogenet Genome Res 119(3-4):225-30.

Brunner HG, Nelen M, Breakefield XO, Ropers HH, van Oost BA. 1993. Abnormal behavior associated with a point mutation in the structural gene for mono-amine oxidase A. Science 262:578-80.

Budiansk S. 1997. The nature of horses: Exploring equine evolution, intelligence and behavior. New York: The Free Press.

Burke RE. 1981. Motor units: Anatomy, physiology, and functional organization. In: Handbook of physiology. Brooks VB, editor. Bethesda, Maryland: American Physiological Society. 345 p.

Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, *et al.* 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447(7145):661-78.

Cases O, Seif I, Grimsby J, Gaspar P, *et al.* 1995. Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA. . Science 268(5218):1763.

Cator R. 1991. Performance evaluations of racing thoroughbreds. Journal of Equine Veterinary Science 11(3):183-90.

Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton, New Jersey: Princeton University Press.

Chakravarti A, Buetow KH, Antonarakis SE, Waber PG, Boehm CD, Kazazian HH. 1984. Nonuniform recombination within the human beta-globin gene cluster. Am J Hum Genet 36(6):1239-58.

Cheng H, Nakamoto M, Bergemann AD, Flanagan JG. 1995. Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. Cell 82(3):371-81.

Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, Yoon D, Lee MH, Kim DJ, Park M, *et al.* 2009. A large-scale genome-wide association study of asian populations uncovers genetic factors influencing eight quantitative traits. Nat Genet 41(5):527-34.

Chowdhary BP, Raudsepp T, Kata SR, Goh G, Millon LV, Allan V, Piumi F, Guerin G, Swinburne J, Binns M, *et al.* 2003. The first-generation whole-genome radiation hybrid map in the horse identifies conserved segments in human and mouse genomes. Genome Res 13(4):742-51.

Clayton HM. 2004. The dynamic horse: A biomechanical guide to equine movement and performance. First edition ed. Mason, MI: Sport Horse Publications.

Clutton-Brock J. 1999. A natural history of domesticated mammals. Cambridge, UK: Cambridge University Press.

Coleman SJ, Zeng Z, Wang K, Luo S, Khrebtukova I, Mienaltowski MJ, Schroth GP, Liu J, Macleod JN. 2010. Structural annotation of equine protein-coding genes determined by mRNA sequencing. Animal Genetics 41(Supplement 2):121-30.

Coleman SJ, Zeng Z, Hestand MS, Liu J, Macleod JN. 2013. Analysis of unannotated equine transcripts identified by mRNA sequencing. PLoS One 8(7):e70125.

Cooper JR, Bloom FE, Rothe RH. 1996. Neuroactive peptides. the biochmical basis of neuropharmacology. Oxford: Oxford University Press. 410 p.

Cothran EG, MacCluer JW, Weitkamp LR, Pfennig DW, Boyce AJ. 1984. Inbreeding and reproductive performance in standardbred horses. The Journal of Heredity 75:220-4.

Cothran EG, MacCluer JW, Weitkamp LR, Bailey E. 1987. Genetic differentiation associated with gait within american standardbred horses. Anim Genet 18(4):285-96.

Cox JE. 1986. Behavior of the false rig: Causes and treatment. Vet Rec 118:353-6.

Craig IW. 2007. The importance of stress and genetic variation in human aggression. Bioessays 29:227-36.

Crone SA, Quinlan KA, Zagoraiou L, Droho S, Restrepo CE, Lundfald L, Endo T, Setlak J, Jessell TM, Kiehn O, *et al.* 2008. Genetic ablation of V2a ipsilateral interneurons disrupts left-right locomotor coordination in mammalian spinal cord. Neuron 60(1):70-83.

Cunningham EP, Dooley JJ, Splan RK, Bradley DG. 2001. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. Animal Genetics 32:360-4.

Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES. 2001. High-resolution haplotype structure in the human genome. Nat Genet 29(2):229-32.

Speed Racking Association – Racking Horse Breeders Association of America [Internet] [cited 2013 7/29/2013]. Available from: http://www.rackinghorse.org/associations/speed-racking-association/ .

Davy A and Soriano P. 2005. Ephrin signaling in vivo: Look both ways. Developmental Dynamics 232(1):1-10.

Deuel N. R. 1995. Dressage canter kinematics and performances in an olympic three-day event. 46th annual meeting of the european association of animal productionPrague, Czech Republic: .

Doan R, Cohen ND, Sawyer J, Ghaffari N, Johnson CD, Dindot SV. 2012. Whole-genome sequencing and variant analysis of a quarter horse mare. BMC Genomics 13:78.

Dodman NH. 1998. Veterinary models of obsessive-compulsive disorder. In: Obsessive-compulsive disorders: Practical management. Jenicke MA, Baer L,Minichiello WA, editors. St. Louis, MO: Mosby. 318 p.

Dolvik NI and Klemetsdal G. 1999. Conformational traits of norwegian cold-blooded trotters: Heritability and the relationship with performance. Acta Agriculturae Scandinavica, Section A, Animal Science 49:156-62.

Donelan JM and Pearson KG. 2004. Contribution of sensory feedback to ongoing ankle extensor activity during the stance phase of walking. Can J Physiol Pharmacol 82(8-9):589-98.

Dottori M, Hartley L, Galea M, Paxinos G, Polizzotto M, Kilpatrick T, Bartlett PF, Murphy M, Kontgen F, Boyd AW. 1998. EphA4 (Sek1) receptor tyrosine kinase is required for the development of the corticospinal tract. Proceedings of the National Academy of Sciences 95(22):13248-53.

Drescher U, Kremoser C, Handwerker C, Löschinger J, Noda M, Bonhoeffer F. 1995. In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal protein related to ligands for eph receptor tyrosine kinases. Cell 82(3):359-70.

Druga R. 2009. Neocortical inhibitory system. Folia Biologica (Prague) 55:201-17.

Dufosset J- and Langlois B. 1984. Analyse statique du geste a l'obstacle de 122 chevaux de selle francais et interest du jugement du saut en liberte. CEREOPA, Compte-Rendu Journee D'Etude 10:2-26.

Duysens J and Pearson KG. 1980. Inhibition of flexor burst generation by loading ankle extensor muscles in walking cats. Brain Res. 187:321-32.

Egloff B and Jan Gruhn A. 1996. Personality and endurance sports. Personality and Individual Differences 21(2):223-9.

Ehrengranat A. 1918. Om hästens rörelser I deras samband med rikonsten. Lund. .

Ellegren H. 2002. It took many mares to form the domestic horse. Trends in Genetics 18(10):500-1.

Elliot FA. 1977. Propranolol for the control of belligerent behavior following acute brain damage. Annals of Neurology 1:489-91.

Evans JW, Borton A, Fintz HF, VanVleck LD. 1977. The horse. San Francisco, California: Freeman.

Forkman B, Boissy A, Meunier-Salaün M-, Canali E, Jones RB. 2007. A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. Physiol Behav 92(3):340-74.

Forssberg H. 1979. Stumbling corrective reaction: A phase-dependent compensatory reaction during locomotion. J Neurophysiol 42(4):936-53.

Forssberg H, Grillner S, Rossignol S. 1977. Phasic gain control of reflexes from the dorsum of the paw during spinal locomotion. Brain Res 132(1):121-39.

Franke H. 1935. Untersuchungen uber den einglub des korperbaus auf die schrittlange des pferdes. Landwirtschaftliche Hochschule, Berlin.

Furey ML, Pietrini P, Haxby JV. 2000. Cholinergic enhancement and increased selectivity of perceptual processing during working memory. Science 290(5500):2315-9.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, *et al.* 2002. The structure of haplotype blocks in the human genome. Science 296(5576):2225-9.

Gale NW, Flenniken A, Compton DC, Jenkins N, Copeland NG, Gilbert DJ, Davis S, Wilkinson DG, Yancopoulos GD. 1996a. Elk-L3, a novel transmembrane ligand for the eph family of receptor tyrosine kinases, expressed in embryonic floor plate, roof plate and hindbrain segments. Oncogene 13(6):1343-52.

Gale NW, Holland SJ, Valenzuela DM, Flenniken A, Pan L, Ryan TE, Henkemeyer M, Strebhardt K, Hirai H, Wilkinson DG, *et al.* 1996b. Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis. Neuron 17(1):9-19.

Geha RM, Chen K, Wouters J, Ooms F, Shih JC. 2002. Analysis of conserved active site residues in monoamine oxidase A and B and their three-dimensional molecular modeling. J. Biol. Chem. 277:17209-126.

Gibson G and Muse SV. 2004. A primer of genome science. 2nd ed. Sunderland, Massachusetts: Sinauer Associates, Inc.

Gnagey L, Clayton HM, Lanovaz JL. 2006. Effect of standing tarsal angle on joint kinematics and kinetics. Equine Veterinary Journal 38:628-34.

Goiffon GC and Vincent AF. 1779. Memoire artificielle des principes relatifs a la fidelle representation des animaux tant en peinture, qu'en sculpture. .

Goldstein O, Zangerl B, Pearce-Kelling S, Sidjanin DJ, Kijas JW, Felix J, Acland GM, Aguirre GD. 2006. Linkage disequilibrium mapping in domestic dog breeds narrows the progressive rod–cone degeneration interval and identifies ancestral disease-transmitting chromosome. Genomics 88(5):541-50.

Gorassini MA, Prochazka A, Hiebert GW, Gauthier MJ. 1994. Corrective responses to loss of ground support during walking. I. intact cats. J Neurophysiol 71(2):603-10.

Gramsbergen A. 2013. Locomotor neurobiology and development. In: Equine locomotion. Back W and Clayton HM, editors. 2nd ed. London, UK: Saunders Elsevier. 73 p.

Grillner S. 1981. Control of locomotion in bipeds, tetrapods, and fish. In: Handbook of physiology: The nervous system, 2, motor control. Brooks V, editor. Bethesda, MA: Am. Physiol. Soc. 1176 p.

Grillner S and Rossignol S. 1978. On the initiation of the swing phase of locomotion in chronic spinal cats. Brain Res 146(2):269-77.

Grossmann KS, Giraudin A, Britz O, Zhang J, Goulding M. 2010. Genetic dissection of rhythmic motor networks in mice. Prog Brain Res 187:19-37.

Guan Y and Stephens M. 2008. Practical issues in imputation-based association mapping. PLoS Genet 4(12):e1000279.

Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, Sulem P, Thorlacius S, Gylfason A, Steinberg S, *et al.* 2008. Many sequence variants affecting diversity of adult human height. Nat Genet 40(5):609-15.

Hada T, Onaka T, Kusunose R, Yagi K. 2001. Effects of novel environmental stimuli on neuroendocrine activity in thoroughbred horses. Journal of Equine Science 12:33-8.

Harris SE. 1993. Horse gaits, balance, and movement. New York, NY: Howell Book House.

Harrison JC and Baldwin RN. 1968. *Care and training of the trotter and pacer.*. Columbus, Ohio: United States Trotting Association.

Hausberger M, *et al.* 1998. Temperament in the horse: Factors in play and practical implications (french). :159-69.

Hauseberg A and Muller C. 2002. A brief note on some possible factors involved in the reactions of horses to humans. Applied Animal Behaviour Science 76:339-44.

Hayes B. 2013. Overview of statistical methods for genome-wide association studies (GWAS). In: Genome-wide association studies and genomic prediction. Gondro C, van der Werf J,Hayes B, editors. New York: Humana Press. 149 p.

Hector RI. 1998. The use of clozapine in the treatment of aggressive schizophrenia. Canadian Journal of Psychiatry 43(5):466-72.

Hemmann K, Ahonen S, Raekallio M, Vainio O, Lohi H. 2014. Exploration of known stereotypic behaviour-related candidate genes in equine crib-biting. Animal 8(3):347-53.

Hendricks B. 1995. International encyclopedia of horse breeds. First edition ed. Norman, OK: University of Oklahoma Press.

Hennings HE. 1933. Untersuchungen über den einfluss der körpermasse und gliedmassenvinkel auf die leistungsschrittlänge bei pmmerschen warmblutstuten. Berlin.

Hernandez J and Hawkins DL. 2001. Training failure among yearling horses. Am. J. Vet. Res. 62:1418-22.

Hiebert GW and Pearson KG. 1999. Contribution of sensory feedback to the generation of extensor activity during walking in the decerebrate cat. J Neurophysiol 81(2):758-70.

Higgins G. 2011. How your horse moves: A unique visual guide to improving performance. 2nd ed. Newton Abbot, UK: David & Charles.

Hildebrand M. 1989. The quadrupedal gaits of vertebrates. Bioscience 39(11):766-75.

Hildebrand M. 1965. Symmetrical gaits of horses. Science 150:701-8.

Hislop J. 1992. Breeding for racing. London, UK: The Kingswood Press.

Hoekstra HE, Boyko AR, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, Lohmueller KE, Zhao K, Brisbin A, Parker HG, *et al.* 2010. A simple genetic architecture underlies morphological variation in dogs. PLoS Biology 8(8):e1000451.

Holl H., Gao S., Fei Z. and Brooks S. A. 2014. Generation of a de novo transcriptome assembly from equine lamellar tissue. Plant & animal genome XXII; January 11-15, 2014; San Diego, CA: .

Holmström M. 2001. The effects of conformation. In: Equine locomotion. Back W and Clayton HM, editors. 1st ed. London: W.B. Saunders. 281 p.

Holmström M and Back W. 2013. The effects of conformation. In: Equine locomotion. Back W and Clayton HM, editors. 2nd ed. London: Saunders Elsevier. 229 p.

Holmström M and Philipsson J. 1993. Relationship between conformation, performance and health in 4-year old swedish warmblood rding horses. Livestock Production Science 33:293-312.

Holmström M, Fredricson I, Drevemo S. 1994. Biokinematic differences between riding horses judged as good and poort at the trot. Equine Veterinary Journal 17 (Suppl):51-6.

Holmström M, Magnusson LF, Philipsson J. 1990. Variation in conformation of swedish warmblood horses and conformational characteristics of elite sport horses. Equine Veterinary Journal 22:186-93.

Horikoshi M, Yaghootkar H, Mook-Kanamori DO, Sovio U, Taal HR, Hennig BJ, Bradfield JP, St Pourcain B, Evans DM, Charoen P, *et al.* 2013. New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. Nat Genet 45(1):76-82.

Houpt KA and Kusunose R. 2000. Genetics of behaviour. In: The genetics of the horse. Bowling AT and Ruvinsky A, editors. Oxon, UK: CABI. 281 p.

Howie BN, Donnelly P, Marchini J. 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5(6):e1000529.

Hultborn H and Nielsen JB. 2007. Spinal control of locomotion - from cat to man. Acta Physiologica 189(2):111-21.

Jager D, Stockert E, Jager E, Gure AO, Scanlan MJ, Knuth A, Old LJ, Chen YT. 2000. Serological cloning of a melanocyte rab guanosine 5'-triphosphate-binding protein and a chromosome condensation protein from a melanoma complementary DNA library. Cancer Res 60(13):3584-91.

Jansen T, Forster P, Levine M, Oekle H, Hurles M, Renfrew C, Weber J, Olek K. 2002. Mitochondrial DNA and the origins of the domestic horse. Proceedings of the National Academy of Sciences of the United States of America 99(16):10905-10.

Jeffcott LB, Rossdale PD, Freestone J, Frank CJ, Towers-Clark PF. 1982. An assessment of wastage in thoroughbred racing from conception to 4 years of age. Equine Veterinary Journal 14:185-98.

Jeffreys AJ, Ritchie A, Neumann R. 2000. High resolution analysis of haplotype diversity and meiotic crossover in the human TAP2 recombination hotspot. Hum Mol Genet 9(5):725-33.

Jeffreys AJ, Kauppi L, Neumann R. 2001. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat Genet 29(2):217-22.

Jezierski T, Jaworski Z, Górecka A. 1999. Effects of handling on behaviour and heart rate in konik horses: Comparison of stable and forest reared youngstock. Appl Anim Behav Sci 62(1):1-11.

Johnston C, Holm K, Erichsen C, Eksell P, Drevemo S. 2004. Kinematic evaluation of the back in fully functioning riding horses. Equine Veterinary Journal 36:495-8.

Kalbfleisch T., Rebolledo-Mendez J., Orlando L. and Macleod J. N. 2014. Resources and progress toward a fully annotated EquCab3. Plant & animal genome XXII; January 11-15, 2014; San Diego, CA: .

Kandel ER, Schwartz JH, Jessel TM. 2000. Principles of neural science. New York: McGraw-Hill.

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E. 2008. Efficient control of population structure in model organism association mapping. Genetics 178(3):1709-23.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42(4):348-54.

Karere GM, Kinnally EL, Sanchez JN, Famula TR, Lyons LA, Capitanio JP. 2009. What is an "adverse" environment? interactions of rearing experiences and MAOA genotype in rhesus monkeys. Biol Psychiatry 65(9):770-7.

Karim L, Takeda H, Lin L, Druet T, Arias JAC, Baurain D, Cambisano N, Davis SR, Farnir F, Grisart B, *et al.* 2011. Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. Nat Genet 43(5):405-13.

Kauppi L, Sajantila A, Jeffreys AJ. 2003. Recombination hotspots rather than population history dominate linkage disequilibrium in the MHC class II region. Hum Mol Genet 12(1):33-40.

Keiper RR and Houpt KA. 1984. Reproduction in feral horses: An eight-year study. Am. J. Vet. Res. 45:991-5.

Kiehn O. 2006. Locomotoer circuits in the mammalian spinal cord. Annu Rev Neruosci 29:279-306.

Kiehn O, Dougherty KJ, Hagglund M, Borgius L, Talpalar A, Restrepo CE. 2010. Probing spinal circuits controlling walking in mammals. Biochemical and Biophysical Research Communications 396:11-8.

Kijas JW. 2014. Haplotype-based analysis of selective sweeps in sheep. Genome :1-5.

Kim JJ, Lee HI, Park T, Kim K, Lee JE, Cho NH, Shin C, Cho YS, Lee JY, Han BG, *et al.* 2010. Identification of 15 loci influencing height in a korean population. J Hum Genet 55(1):27-31.

Klausberger T and Somogyi P. 2008. Neuronal diversity and temporal dynamics: The unity of hippocampal circuit operations. Science 321(5885):53-7.

Klein R and Kania A. 2014. Ephrin signalling in the developing nervous system. Curr Opin Neurobiol 27:16-24.

Kronacher C and Ogrizek A. 1932. Exterieur und leistungsfahigkeit des pferdes mit besonderen berucksichtigung der gliedmabenwinkelung und schrittlangenverhaltnisse. Z. Zucht Reihe B Tierz. U. Zuchtungsbiologie 32:183-228.

Krüger W. 1938. Ueber den bewegungsablauf an dem oberen teil der hintergliedmabe des pferedes im schritt, trab und galopp. Tierarztl. Rundsch. 44(34):549-57.

Krüger W. 1937. Ueber den bewegungsablauf an dem oberen teil der vordergliedmabe des pferdes im schritt, trab und galopp. Tierarztl. Rundsch 43(49/50):809,816, 825-827.

Kukekova AV, Temnykh SV, Johnson JL, Trut LN, Acland GM. 2012. Genetics of behavior in the silver fox. Mamm Genome 23(1-2):164-77.

Kullander K, Croll SD, Zimmer M, Pan L, McClain J, Hughes V, Zabski S, DeChiara TM, Klein R, Yancopoulos GD, *et al.* 2001a. Ephrin-B3 is the midline barrier that prevents corticospinal tract axons from recrossing, allowing for unilateral motor control. Genes Dev 15(7):877-88.

Kullander K, Mather NK, Diella F, Dottori M, Boyd AW, Klein R. 2001b. Kinase-dependent and kinase-independent functions of EphA4 receptors in major axon tract formation in vivo. Neuron 29(1):73-84.

Kunieda T, Park J, Takeuchi H, Kubo T. 2003. Identification and characterization of Mlr1,2: Two mouse homologues of mblk-1, a transcription factor from the honeybee brain. FEBS Lett 535(1-3):61-5.

Kuo AD. 2002. The relatie roles of feedforward and feedback in the control of rhythmic movements. Motor Control 6:129-45.

Lane G. 2011. Discussion with author in september. .

Langlois B, Froideveaux J, Lamarche L, Legault P, Tassencourt L, Theret M. 1978. Analyse de liaisons centre la morphologie et l'aptitude au galop, au trot et au saut d'obstacle chex le cheval. Ann. Génét. Sél. Anim. 10:443-74.

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, *et al.* 2010a. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467(7317):832-8.

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, *et al.* 2010b. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467(7317):832-8.

Lansade L, Bouissou M, Erhard HW. 2008. Fearfulness in horses: A temperament trait stable across time and situations. Appl Anim Behav Sci 115(3-4):182-200.

Lau AN, Peng L, Goto H, Chemnick L, Ryder OA, Makova KD. 2009. Horse domestication and conservation genetics of przewalski's horse inferred from sex chromosomal and autosomal sequences. Molecular Biology and Evolution 26(1):199-208.

Le Scolan N, Hausberger M, Wolff A. 1997. Stability over situations in temperamental traits of horses as revealed by experimental and scoring approaches. Behav Processes 41(3):257-66.

Leach DH and Dagg AI. 1983. A review of research on equine locomotion and biomechanics. Equine Veterinary Journal 15(2):93-102.

Lear TL, Brandon R, Piumi F, Terry RR, Guérin G, Thomas S, Bailey E. 2001. Mapping 31 horse genes in BACs by FISH: Identification of chromosomal rearrangements and conserved synteny relative to the human gene map. Chromosome Research 9(3):261-2.

Lee R and Coccaro E. 2001. The neuropsychopharmacology of criminality and aggression. Canadian Journal of Psychiatry 46(1):35-44.

Leeb T, Vogl C, Zhu B, de Jong PJ, Binns MM, Chowdhary BP, Scharfe M, Jarek M, Nordsiek G, Schrader F, *et al.* 2006. A human–horse comparative map based on equine BAC end sequences. Genomics 87(6):772-6.

Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, Eyheramendy S, Voight BF, Butler JL, Guiducci C, *et al.* 2008. Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet 40(5):584-91.

Levine MA. 1999. Botai and the origins of horse domestication. Journal of Anthropological Archaeology 18:29-78.

Levine MA. 1993. Social evolution and horse domestication. In: Trade and exchange in prehistoric europe. Scarre C and Healy F, editors. Oxbow, Oxford: . 135 p.

Levine MA. 1990. Dereivka and the problem of horse domestication. Antiquity 64:727-40.

Linder A and Dingerkus A. 1993. Incidence of training failure among thoroughbred horses at coligne, germany. Preventative Veterinary Medicine 16:85-94.

Lindgren G, Backstrom N, Swinburne J, Hellborg L, Einarsson A, Sandberg K, Cothran G, Vila C, Binns M, Ellegren H. 2004. Limited number of patrilines in horse domestication. Nature Genetics 36(4):335-6.

Lindholm-Perry AK, Kuehn LA, Oliver WT, Sexten AK, Miles JR, Rempel LA, Cushman RA, Freetly HC. 2013. Adipose and muscle tissue gene expression of two genes (NCAPG and LCORL) located in a chromosomal region associated with cattle feed intake and gain. PLoS One 8(11):e80882.

Linklater WL. 2000. Adaptive explanation in socio-ecology: Lessons from the equidae. Biological Reviews 75(01):1-20.

Lippold S, Knapp M, Kunetsova T, Leonard JA, Benecke N, Ludwig A, Rasmussen M, Cooper A, Weinstock J, Willerslev E, *et al.* 2011. Discovery of lost diversity of paternal horse lineages using ancient DNA. Nature Communications 2:450.

Lisabeth EM, Falivelli G, Pasquale EB. 2013. Eph receptor signaling and ephrins. Cold Spring Harb Perspect Biol 5(9):10.1101/cshperspect.a009159.

Lloyd AS, Martin JE, Bornett-Gauci HLI, Wilkinson RG. 2008. Horse personality: Variation between breeds. Applied Animal Behaviour Science 112:369-83.

Lloyd AS, Martin JE, Bornett-Gauci HLI, Wilkinson RG. 2007. Evaluation of a novel method of horse personality assessment: Rater-agreement and links to behaviour. Appl Anim Behav Sci 105(1-3):205-22.

Love S, Wyse CA, Stirk A, Stear MJ, Voute L, Calver P, Mellor DJ. 2006. Prevalance, heritability and significance of musculoskeletal conformational traits in thoroughbred yearlings. Equine Veterinary Journal 38:597-603.

Lühken G, Glenske K, Brandt H, Erhardt G. 2010. Genetic variation in monoamine oxidase A and analysis of association with behaviour traits in beef cattle. J Anim Breed Genet. 127:411-8.

MacCluer JW, Boyce AJ, Dyke B, Weitkamp LR, Pfennig DW, Parsons CJ. 1983. Inbreeding and pedigree structure in standardbred horses. Journal of Hereditary 74:394-9.

Magnusson L-. 1985. Studies on the conformation and related traits of standardbred trotters in sweden. II. SLU.

Magnusson L and Thafvelin B. 1990. Studies on the conformation and related traits of standardbred trotters in sweden. J Anim Breed Genet. 107:135-48.

Makvandi-Nejad S, Hoffman GE, Allen JJ, Chu E, Gu E, Chandler AM, Loredo AI, Bellone RR, Mezey JG, Brooks SA, *et al.* 2012. Four loci explain 83% of size variation in the horse. PLos One 7(7):e39929.

Marin O. 2012. Interneuron dysfunction in psychiatric disorders. Nat Rev Neurosci 13(2):107-20.

Martin BB, Reef VB, Parente EJ, Sage AD. 2000. Causes of poor performance of horses during training, racing, or show-jumping: 384 cases (1992–1996). J. Am. Vet. Med. Assoc. 216:554-8.

May CA, Shone AC, Kalaydjieva L, Sajantila A, Jeffreys AJ. 2002. Crossover clustering and rapid decay of linkage disequilibrium in the xp/yp pseudoautosomal gene SHOX. Nat Genet 31(3):272-5.

McCann JS, Heird JC, Bell RW, Lutherer LO. 1988. Normal and more highly reactive horses. I. heart rate, respiration rate and behavioral observations. Appl Anim Behav Sci 19(3-4):201-14.

McCue ME, Bannasch DL, Petersen JL, Gurr J, Bailey E, Binns MM, Distl O, Guerin G, Hasegawa T, Hill EW, *et al.* 2012. A high density SNP array for the domestic horse and extant perissodactyla: Utility for association mapping, genetic diversity, and phylogeny studies. PLoS Genet 8(1):e1002451.

McGivney BA, McGettigan PA, Browne JA, Evans AC, Fonseca RG, Loftus BJ, Lohan A, MacHugh DE, Murphy BA, Katz LM, *et al.* 2010. Characterization of the equine skeletal muscle transcriptome identifies novel functional responses to exercise training. BMC Genomics 11:398.

McGreevy P. 2004. Equine behavior. A guide for veterinarians and equine scientists. 1st ed. London, UK: Saunders Elsevier.

McGreevy PD and Thomson PC. 2006. Differences in motor laterality between breeds of performance horse. Applied Animal Behaviour Science 99:183-90.

McIlwraith CW, Anderson TM, Sanschi EM. 2003. Conformation and musculoskeletal problems in the racehorse. Clinical Techniques in Equine Practices. 2(4):339-47.

McVea DA. 2005. A role for hip position in initiating the swing-to-stance transition in walking cats. J Neurophysiol 94(5):3497-508.

Meira CT, Farah MM, Fortes MRS, Moore SS, Pereira GL, Silva JA, da Mota MDS, Curi RA. 2014. A genome-wide association study for morphometric traits in quarter horse. Journal of Veterinary Science 34(8):1028-31.

Menozzi P, Piazza A, Cavalli-Sforza L. 1978. Synthetic maps of human gene frequencies in europeans. Science 201(4358):786-92.

Metzger J, Schrimpf R, Philipp U, Distl O. 2013. Expression levels of LCORL are associated with body size in horses. PLos One 8(2):e56497.

Milenkovic D, Oustry-Vaiman A, Lear TL, Billault A, Mariat D, Piumi F, Schibler L, Cribiu E, Guerin G. 2002. Cytogenetic localization of 136 genes in the horse: Comparative mapping with the human genome. Mamm Genome 13(9):524-34.

Momozawa Y, Kusunose R, Kikusui T, Takeuchi Y, Mori Y. 2005a. Assessment of equine temperament questionnaire by comparing factor structure between two separate surveys. Appl Anim Behav Sci 92(1-2):77-84.

Momozawa Y, Takeuchi Y, Kusunose R, Kikusui T, Mori Y. 2005b. Association between equine temperament and polymorphisms in dopamine D4 receptor gene. Mammalian Genome 16(7):538-44.

Momozawa Y, Ono T, Sato F, Kikusui T, Takeuchi Y, Mori Y, Kusunose R. 2003. Assessment of equine temperament by a questionnaire survey to caretakers and evaluation of its reliability by simultaneous behavior test. Appl Anim Behav Sci 84(2):127-38.

Murphy LA and Sarge KD. 2008. Phosphorylation of CAP-G is required for its chromosomal DNA localization during mitosis. Biochem Biophys Res Commun 377(3):1007-11.

Equine SNP70 BeadChip Whole Genome SNP profiling. [Internet]; c2013. Available from: http://www.neogen.com/Genomics/pdf/Slicks/EquineGenotypingFlyer.pdf .

Nicodemus MC and Clayton HM. 2003. Temporal variables of four-beat, stepping gaits of gaited horses. Applied Animal Behaviour Science 80:133-42.

Nicodemus MC, Holt KM, Swartz K. 2002. Relationship between velocity and temporal variables of the flat shod running walk. Equine Veterinary Journal Supplemental 34:340-3.

Notari L and Goodwin D. 2007. A survey of behavioural characteristics of pure-bred dogs in italy. Appl Anim Behav Sci 103(1-2):118-30.

Novembre J and Stephens M. 2008. Interpreting principal component analyses of spatial population genetic variation. Nat Genet 40(5):646-9.

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, *et al.* 2008. Genes mirror geography within europe. Nature 456(7218):98-101.

Odani M, Narita A, Watanabe T, Yokouchi K, Sugimoto Y, Fujita T, Oguni T, Matsumoto M, Sasaki Y. 2006. Genome-wide linkage disequilibrium in two japanese beef cattle breeds. Anim Genet 37(2):139-44.

Okada Y, Kamatani Y, Takahashi A, Matsuda K, Hosono N, Ohmiya H, Daigo Y, Yamamoto K, Kubo M, Nakamura Y, *et al.* 2010. A genome-wide association study in 19 633 japanese subjects identified LHX3-QSOX2 and IGF1 as adult height loci. Hum Mol Genet 19(11):2303-12.

Orlovsky G, Deliagina TG, Grillner S. 1999. Neuronal control of locomotion: From mollusc to man. New York: Oxford University Press.

Outram AK, Stear NA, Bendrey R, Olsen S, Kasparov A, Zaibert V, Thorpe N, Evershed RP. 2009. The earliest horse harnessing and milking. Science 323:1332-5.

Ouzounis CA, Coleman SJ, Zeng Z, Hestand MS, Liu J, Macleod JN. 2013. Correction: Analysis of unannotated equine transcripts identified by mRNA sequencing. PLoS ONE 8(9).

Overall KL. 1998. Self-injurious behavior and obsessive-compulsive disorder in domestic animals. In: Psychopharmacology of animal behavior disorders. Dodman NH and Shuster L, editors. Malden, MA: Blackwell Sciences.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. PLoS Genetics 2(12):e190.

Pausch H, Flisikowski K, Jung S, Emmerling R, Edel C, Gotz KU, Fries R. 2011. Genome-wide association study identifies two major loci affecting calving ease and growth-related traits in cattle. Genetics 187(1):289-97.

Penedo MCT, Millon LV, Bernoco D, Bailey E, Binns M, Cholewinski G, Ellis N, Flynn J, Gralak B, Guthrie A, *et al.* 2005. International equine gene mapping workshop report: A comprehensive linkage map constructed with data from new markers and by merging four mapping resources. Cytogenetic and Genome Research 111(1):5-15.

Perola M. 2011. Genome-wide association approaches for identifying loci for human height genes. Best Pract Res Clin Endocrinol Metab 25(1):19-23.

Petersen JL, Mickelson JR, Cothran EG, Andersson LS, Axelsson J, Bailey E, Bannasch DL, Binns MM, Borges AS, Brama P, *et al.* 2013. Genetic diversity in the modern horse illustrated from genome-wide SNP data. PLos One 8(1):e54997.

Philipsson J, Brendow E, Dalin G, Wallin L. 1998. Genetic aspects of disease and lesions in horses. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production 24:408-15.

Phillips MS, Lawrence R, Sachidanandam R, Morris AP, Balding DJ, Donaldson MA, Studebaker JF, Ankener WM, Alfisi SV, Kuo FS, *et al.* 2003. Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. Nat Genet 33(3):382-7.

Popova NK, Skrinskaya YA, Amstislavskaya TG, Vishnivetskaya GB, Seif I, de Meier E. 2001. **Behavioral characteristics of mice with genetic knockout of monoamine oxidase type A**. Neurosci Behav Physiol 31(6):597-602.

Price AL, Zaitlen NA, Reich D, Patterson N. 2010. New approaches to population stratification in genome-wide association studies. Nature Reviews Genetics 11(7):459-63.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38(8):904-9.

Pritchard CC. 1965. Relationship between conformation and lameness in the foot. Auburn Vet. 22:111,126,129.

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. 2000. Association mapping in structured populations. Am J Hum Genet 67(1):170-81.

Promerová M, Andersson LS, Juras R, Penedo MCT, Reissmann M, Tozaki T, Bellone R, Dunner S, Horín P, Imsland F, *et al.* 2014. Worldwide frequency distribution of the 'Gait keeper' mutation in theDMRT3gene. Anim Genet 45(2):274-82.

Pryce JE, Hayes BJ, Bolormaa S, Goddard ME. 2011. Polymorphic regions affecting human height also control stature in cattle. Genetics 187(3):981-4.

PLINK: Whole genome data analysis toolset [Internet] [cited 2014 11/15/2014]. Available from: http://pngu.mgh.harvard.edu/~purcell/plink/contact.shtml#cite .

Purves D. 2008. Neuroscience. 4th ed. Sunderland, Massachusetts: Sinauer Associates, Inc.

Raudsepp T, Kijas J, Godard S, Guérin G, Andersson L. 1999. Comparison of horse chromosome 3 with donkey and human chromosomes by cross-species painting and heterologous FISH mapping. Mammalian Genome 10(3):277-82.

Rendon RA, Shuster L, Dodman NH. 2001. The effect of the NMDA receptor blocker, dextromethorphan, on cribbing in horses. Pharmacology Biochemistry and Behavior 68(1):49-51.

Riegel RJ and Hakola SE. 1999. The nervous system. In: Illustrated atlas of clinical equine anatomy and common disorders of the horse. volume two. 1st ed. Marysille, Ohio: Equistar Publications, Limited. 254 p.

Robilliard JJ, Pfau T, Wilso AM. 2007. Gait characterisation and classification in horses. The Journal of Experimental Biology 210:187-97.

Rossignol S. 2006. Dynamic sensorimotor interactions in locomotion. Physiol Rev 86(1):89-154.

Rossignol S, Lund JP, Drew T. 1988. The role of sensory inputs in regulating patterns of rhythmical movements in higher vertebrates. In: Neural control of rhythmic movements in vertebrates. Cohen AH, Rossignol S,Grillner S, editors. New York: Wiley. 201 p.

Rudolph J, Gerstmann K, Zimmer G, Steinecke A, Döding A, Bolz J. 2014. A dual role of EphB1/ephrin-B3 reverse signaling on migrating striatal and cortical neurons originating in the preoptic area: Should I stay or go away? Frontiers in Cellular Neuroscience 8: 185.

Rudoph J, Spier S, Byrns G, Rojas C, Bernoco D, Hoffman E. 1992. Periodic paralysis in quarter horses, a sodium channel mutation disseminated by selective breeding. Nature Genetics 2:144-7.

Saastamoinen MT and Barrey E. 2000. Genetics of conformation, locomotion and physiological traits. In: The genetics of the horse. Bowling AT and Ruvinsky A, editors. New York, NY: CAB International. 439 p.

Saltiel P and Rossignol S. 2004a. Critical points in the forelimb fictive locomotor cycle and motor coordination: Effects of phasic retractions and protractions of the shoulder in the cat. J Neurophysiol 92(3):1342-56.

Saltiel P and Rossignol S. 2004b. Critical points in the forelimb fictive locomotor cycle and motor coordination: Evidence from the effects of tonic proprioceptive perturbations in the cat. J Neurophysiol 92(3):1329-41.

Santschi EM, Purdy AK, Valberg SJ, Vrotsos PD, Kaese H, Mickelson JR. 1998. Endothelin receptor B polymorphism associated with lethal white foal syndrome in horses. Mammalian Genome 9(4):306-9.

Schaefer R., Schubert M., Orlando L., Mickelson J. R., McCue M. E. and The Equine Genetic Diversity Consortium. 2014. Haplotype discovery and an imputation resource for the domestic horse. Plant & animal genome XXII; January 11-15, 2014; San Diego, CA: .

Schauder W. 1923. Historisch-kritische studie uber die bewegungslehre des pferdes. (1 teil). Berl. Tierarztl. Wschr. 39:123-6.

Scheibel ME and Scheibel AB. 1970. Developmental relationship between spinal motoneuron dendrite bundles and patterned activity in the hind limb of cats. Exp Neurol 29(2):328-35.

Schmidt H. 1939. Beziehungen zwischen schrittlange und bau der gliedmasen des pferdes. Deutsch. Tiearztl. Wschr. 47:689-92.

Schneider JA, Peto TEA, Boone RA, Boyce AJ, Clegg JB. 2002. Direct measurement of the male recombination fraction in the human beta-globin hot spot. Hum Mol Genet 11(3):207-15.

Schoenen J. 1982. Dendritic organization of the human spinal cord: The motoneurons. J Comp Neurol 211(3):226-47.

Seaman SC, Davidson HPB, Waran NK. 2002. How reliable is temperament assessment in the domestic horse (equus caballus)? Appl Anim Behav Sci 78(2-4):175-91.

Sellet LC, Albert WW, Groppel JL. 1981. Forelimb kinematics of the standardbred pacing gait. Proc. Equine Nutr. Physiol. :210-5.

Serpell JA and Hsu Y. 2001. Development and validation of a novel method for evaluating behavior and temperament in guide dogs. Appl Anim Behav Sci 72(4):347-64.

Setoguchi K, Watanabe T, Weikard R, Albrecht E, Kuhn C, Kinoshita A, Sugimoto Y, Takasuga A. 2011. The SNP c.1326T>G in the non-SMC condensin I complex, subunit G (NCAPG) gene encoding a p.Ile442Met variant is associated with an increase in body frame size at puberty in cattle. Anim Genet 42(6):650-5.

Shendure J and Ji H. 2008. Next-generation DNA sequencing. Nat Biotechnol 26(10):1135-45.

Shik ML and Orlovsky G. 1976. Neurophysiology of locomotor automatism. Physiol. Rev. 56:465-501.

Shin EK, Perryman LE, Meek K. 1997. A kinase-negative mutation of DNA-PK(CS) in equine SCID results in defective coding and signal joint formation. J Immunol 158(8):3565-9.

Signer-Hasler H, Flury C, Haase B, Burger D, Simianer H, Leeb T, Reider S. 2012. A genome-wide association study reveals loci influencing height and other conformation traits in horses. PLos One 7:e37282.

Smith AM, Staniar WB, Splan RK. 2006. Associations between yearling body measurements and career racing performance in thoroughbred racehorses. Journal of Equine Veterinary Science 26(5):212-4.

Snyder M, Du J, Gerstein M. 2010. Personal genome sequencing: Current approaches and challenges. Genes and Development 24:423-31.

Soranzo N, Rivadeneira F, Chinappen-Horsley U, Malkina I, Richards JB, Hammond N, Stolk L, Nica A, Inouye M, Hofman A, *et al.* 2009. Meta-analysis of genome-wide scans for human adult stature identifies novel loci and associations with measures of skeletal frame size. PLoS Genet 5(4):e1000445.

Splan RK and Hunter HB. 2004. Temporal variables of the canter of the tennessee walking horse. Equine and Comparative Exercise Physiology 1(1):41-4.

Stashak TS. 1987. The relationship between conformation and lameness. In: Adam's lameness in horses. Stashak TS, editor. Philadelphia, PA: Lea & Febiger. 71 p.

Steemers FJ, Chang W, Lee G, Barker DL, Shen R, Gunderson KL. 2006. Whole-genome genotyping with the single-base extension assay. Nature Methods 3(1):31-3.

Stein RB, Misiaszek JE, Pearson KG. 2000. Functional role of muscle reflexes for force generation in the decerebrate walking cat. J Physiol (Lond) 525(3):781-91.

Steinecke A, Gampe C, Zimmer G, Rudolph J, Bolz J. 2014. EphA/ephrin A reverse signaling promotes the migration of cortical interneurons from the medial ganglionic eminence. Development 141(2):460-71.

Stevens EF. 1990. Instability of harems of feral horses in relation to season and presence of subordinate stallions. Behaviour 112(3/4):149-61.

Suontama M, van der Werf JTN, Juga J, Ojala M. 2013. Genetic correlations for foal and studbook traits with racing traits and implications for selection strategies in the finnhorse and standardbred trotter. J Anim Breed Genet. 130(3):178-89.

Sutter NB, Bustamante CD, Chase K, Gray MM, Zhao K, Zhu L, Padhukasahasram B, Karlins E, Davis S, Jones PG, *et al.* 2007. A single IGF1 allele is a major determinant of small size in dogs. Science 316(5821):112-5.

Svartberg K. 2006. Breed-typical behaviour in dogs—Historical remnants or recent constructs? Appl Anim Behav Sci 96(3-4):293-313.

Svartberg K. 2005. A comparison of behaviour in test and in everyday life: Evidence of three consistent boldness-related personality traits in dogs. Appl Anim Behav Sci 91(1-2):103-28.

The Tennessee Walking Horse Breed: Gaits [Internet]; c2011 [cited 2014 . Available from: www.twhbea.com/breed/gait.php .

Tozaki T, Hirota K, Hasegawa T, Ishida N, Tobe T. 2007. Whole-genome linkage disequilibrium screening for complex traits in horses. Mol Genet Genomics 277(6):663-72.

Trut L, Oskina I, Kharlamova A. 2009. Animal evolution during domestication: The domesticated fox as a model. Bioessays 31(3):349-60.

Five Gaits [Internet] [cited 2014 11/17/2014]. Available from: http://www.icelandics.org/gaits.php .

Van der Veen G. 1918. Het paard in partjes. Tirion Uitgevers, the Netherlands .

van Weeren PR. 2013. History. In: Equine locomotion. Back W and Clayton HM, editors. 2nd ed. New York: Saunders Elsevier. 1 p.

van Weeren PR and Crevier-Denoix N. 2006. Equine conformation: Clues to performance and soundness? Equine Veterinary Journal 38(7):591-6.

Vila C, Leonard JA, Gotherstrom A, Marklund S, Sandberg K, Liden K, Wayne RK, Ellegren H. 2001. Widespread origins of domestic horse lineages. Science 291(5503):474-7.

Visscher PM, McEvoy B, Yang J. 2010. From galton to GWAS: Quantitative genetics of human height. Genet Res (Camb) 92(5-6):371-9.

Visscher PM, Hill WG, Wray NR. 2008. Heritability in the genomics era--concepts and misconceptions. Nat Rev Genet 9(4):255-66.

Visser E, Vanreenen C, Van der Werf J, Schilder M, Knaap J, Barneveld A, Blokhuis H. 2002. Heart rate and heart rate variability during a novel object test and a handling test in young horses. Physiol Behav 76(2):289-96.

Visser EK, van Reenen CG, Schilder MBH, Barneveld A, Blokhuis HJ. 2003a. Learning performances in young horses using two different learning tests. Appl Anim Behav Sci 80(4):311-26.

Visser EK, Van Reenen CG, Engel B, Schilder MBH, Barneveld A, Blokhuis HJ. 2003b. The association between performance in show-jumping and personality traits earlier in life. Appl Anim Behav Sci 82(4):279-95.

Visser EK, van Reenen CG, Hopster H, Schilder MBH, Knaap JH, Barneveld A, Blokhuis HJ. 2001. Quantifying aspects of young horses' temperament: Consistency of behavioural variables. Appl Anim Behav Sci 74(4):241-58.

Voight B and Pritchard J. 2005. Confounding from cryptic relatedness in case-control association studies. PLoS Genetics preprint(2005):e32.

Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, Lear TL, Adelson DL, Bailey E, Bellone RR, *et al.* 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science 326(5954):865-7.

Wagener H. 1934. Untersuchungen an spitzenpferden des spring- und schulstalles der kavalerie-schule hannover. Arbeiten Der Deutsch. Gesellsch. Fur Zuchtungsbiol. 65:1-117.

Wall JD and Pritchard JK. 2003. Haplotype blocks and linkage disequilibrium in the human genome. Nature Reviews Genetics 4(8):587-97.

Wallner B, Brem G, Muller M, Achmann R. 2003. Fixed nucleotide differences on the Y chromosome indicate clear divergence between equus przewalskii and equus cabalus. Animal Genetics 34:453-6.

Wallner B, Piumi F, Brem G, Muller M, Achmann R. 2004. Isolation of Y chromosome-specific microsatellites n the horse and corss-species amplification of the genus equus. Journal of Hereditary 95:158-64.

Wehner R. 1944. Lassen sich beziehungen der knochenachsen und gliedmabenwinkel und schrittlange beim rheinisch-deutschen kaltblut naschweisen? Z. Tierz. Zucht. Biol. 56:321-53.

Weir BS, Anderson AD, Hepler AB. 2006. Genetic relatedness analysis: Modern data and new challenges. Nature Reviews Genetics 7(10):771-80.

Weller R, Pfau T, May SA, Wilson AM. 2006a. Variation in conformation in a cohort of national hunt racehorses. Equine Veterinary Journal 38:616-21.

Weller R, Pfau T, Babbage D, Brittin E, May SA, Wilson AM. 2006b. Reliability fo conformational measurements in the horse using a three-dimensional motion analysis system. Equine Veterinary Journal 38(7):610-5.

Wolff A, Hausberger M, Le Scolan N. 1997. Experimental tests to assess emotionality in horses. Behav Processes 40(3):209-21.

Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z, *et al.* 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46(11):1173-86.

Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. 2013. Pitfalls of predicting complex traits from SNPs. Nat Rev Genet 14(7):507-15.

Yang G, Croaker D, Zhang AL, Manglick P, Cartmill T, Cass D. 1998. A dinucleotide mutation in the endothelin-B receptor gene is associated with lethal white foal syndrome (LWFS); a horse variant of hirschsprung disease. Hum Mol Genet 7(6):1047-52.

Yang J, Zhu W, Chen J, Zhang Q, Wu S. 2014a. Genome-wide two-marker linkage disequilibrium mapping of quantitative trait loci. BMC Genet 15:20.

Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. 2014b. Advantages and pitfalls in the application of mixed-model association methods. Nat Genet 46(2):100-6.

Yokoyama N, Romero MI, Cowan CA, Galvan P, Helmbacher F, Charnay P, Parada LF, Henkemeyer M. 2001. Forward signaling mediated by ephrin-B3 prevents contralateral corticospinal axons from recrossing the spinal cord midline. Neuron 29:85-97.

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, *et al.* 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38(2):203-8.

Yuste R, MacLean JN, Smith J, Lansner A. 2005. Opinion: The cortex as a central pattern generator. Nature Reviews Neuroscience 6(6):477-83.

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, *et al.* 2010. Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42(4):355-60.

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, *et al.* 2007. An arabidopsis example of association mapping in structured samples. PLoS Genetics 3(1):e4.

Zhou X and Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. Nat Genet 44(7):821-4.

Ziegler L. 2005. Easy-gaited horses. First edition ed. North Adams, MA: Storey Publishing.

Zimmer G, Rudolph J, Landmann J, Gerstmann K, Steinecke A, Gampe C, Bolz J. 2011. Bidirectional ephrinB3/EphA4 signaling mediates the segregation of medial ganglionic eminence- and preoptic area-derived interneurons in the deep and superficial migratory stream. J Neurosci 31(50):18364-80.

Zips S, Peham C, Scheidl M, Licka T, Girtler D. 2001. Motion pattern of the telt of icelandic horses at different speeds. Equine Veterinary Journal Supplemental 33:109-11.

Zondervan KT and Cardon LR. 2004. The complex interplay among factors that influence allelic association. Nature Reviews Genetics 5(2):89-100.

CHAPTER 2

GENOME-WIDE ASSOCIATION STUDIES OF POLYMORPHIC GAIT

Introduction

Horses are polymorphic in the pattern and timing of locomotion. At intermediate speeds, beyond the flat walk, the horse can perform a range of diagonal and lateral, two-beat or four-beat, footfall patterns (Harris 1993). The diagonal gaits include the two-beat trot and four-beat foxtrot; the lateral gaits include the two-beat pace, and four-beat rack and running walk. Both the trot and pace have a moment of suspension where all four feet are off the ground before one pair lands at the same time. A horse with the ability to perform any of the four-beat gaits without a moment of suspension is commonly called a "gaited" horse.

The horse is also a unique model for the study of gait development, as no other mammalian species is known to discretely segregate for congenital differences in stride cadence and footfall pattern. Laboratory induced models of variation in gait exist in mice, cats, and lampreys via transverse sectioning of the spinal cord (Grillner and Zangger 1979), induced mutations or knockouts in mice (Coonan *et al.* 2001; Crone *et al.* 2008; Dottori *et al.* 1998; Hinckley *et al.* 2005; Kullander *et al.* 2001a; Kullander *et al.* 2001b; Wilson *et al.* 2005; Yokoyama *et al.* 2001), and pharmacologically induced blockage of synaptic pathways in lampreys (Grillner 2003). These laboratory models enabled discovery of localized neural networks known as central pattern generators (CPGs) (Kuo 2002). CPGs contribute to a number of repetitive reflexive actions, including gait. When locomotion is initiated, activity in the CPG is turned on and maintained by inputs from descending locomotor commands originating from neurons in the brainstem and midbrain (Kiehn 2006). Yet existing laboratory animal models do not possess an innate ability to perform a lateral gait. Naturally occurring models provide a novel resource for study of CPG structure and function.

Although the heritability for gait type has not been extensively reported in the horse, a strong role for genetics is supported by the discrete segregation among breeds for the propensity to

perform one gait type over another. Of the approximately 515 international breeds of horse, 78 possess polymorphic gait type (Hendricks 1995; Ziegler 2005). Additionally there are 15 breeds known to not be polymorphic, yet rare individuals with polymorphism in gait type are reported (Hendricks 1995; Ziegler 2005). The Icelandic horse breed is the most extensively studied in regards to heritability for gait type, which has been reported to be between 0.38-0.58, but there are at least two gait subpopulations within the breed undergoing selection (Albertsdóttir *et al.* 2008). These subpopulations were utilized to identify a premature stop mutation within the *DMRT3* gene (Andersson *et al.* 2012), and subsequent analysis has shown a high frequency rate of the mutation within several gaited breeds and racing Standardbreds (Promerová *et al.* 2014). While the mutation does impact locomotion, the high frequency of the mutation within the racing trotting Standardbreds suggests more of an impact on speed regulation. The Tennessee Walking Horse (TWH) is a US breed renowned for their ability to perform an even-timed four-beat gait (the "running-walk") at intermediate speeds. The TWH is nearly fixed for the *DMRT3* mutation (Promerová *et al.* 2014), yet within the TWH breed there is variation in gait type. TWHs are capable of performing the whole range of intermediate gaits from two-beat lateral to four-beat to two-beat diagonal (Tennessee Walking Horse Breeders' and Exhibitors' Association 2011). At liberty, these horses show an innate preference to either trot, pace, or running walk from birth (Figure 2.1). Through training and shoeing, this innate preference can occasionally be enhanced to perform a more even-timed running walk.

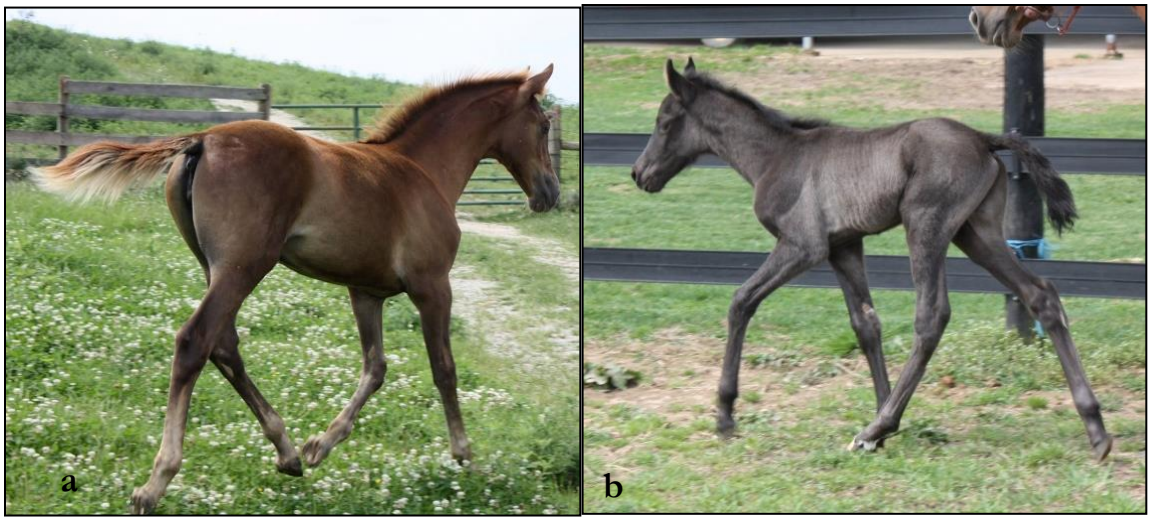


Figure 2.1. Two foals that exemplify variation in gait types in TWH. The diagonal trot is shown in pane "a" (foal is two months old) and the lateral amble in pane "b" (foal is two days old).

There are two objectives for this study. The first is to identify loci and mutations, other than *DMRT3,* that are common in gaited breeds and rare in trotting breeds (Promerová *et al.* 2014). The second objective of this study is to identify candidate loci and SNPs unique to gait type variation within the TWH breed. Both of these objectives will be attempted through genome-wide association studies (GWAS) and filtering of polymorphisms in available whole-genome sequences

Materials & Methods

Animals and Phenotypes – Across Breeds

From previous studies, 98 horses, two to six horses each from 30 diverse breeds (Table 2.1) were genotyped and were available for use in this study. For three of these breeds (American Saddlebred, Icelandic Horse, and Puerto Rican Paso Fino), the ability to perform a lateral four-beat gait is a key characteristic of the breed. Yet, gait traits are not fixed in these breeds and some individuals within these breeds are still able to trot. Within the American Saddlebred, there are subpopulations for gait type where some individuals cannot perform a lateral four-beat gait. Therefore, we investigated breed association competition records for these individuals and their relatives (ASHA). These records revealed that three individuals were exhibited strictly in three-gaited classes (only able to trot) and three individuals appeared at least once in the five-gaited division (able to trot and rack). Such information is not available for the Icelandic and Puerto Rican Paso Fino horses. Therefore, we assumed all individuals from the Icelandic Horse and Puerto Rican Paso Fino horse to be gaited based on the lack of known subpopulations for trotting-only ability.

Table 2.1. Breed distribution of animals used in the across breed GWA. *Breeds known to have some individuals polymorphic for gait. Genome project origin are samples contributed by the Equine Genetics Consortium; size project origin are samples published by Makvandi-Nejad *et al.* (2012).

Breed	Abbrev.	Total Number	Number Gaited	Project Origin
American Minature Horse	AAM	3	0	Size
Andalusian	AND	3	0	Genome
Ardennais	ARD	3	0	Size
Arabian	ARA	3	0	Genome
Belgian	BEL	6	0	Genome, Size
Brabant	BRA	3	0	Size
Caspian	CAS	3	0	Size
Clydesdale	CLY	3	0	Size
Dartmoor Pony	DMP	3	0	Size
Falabella	FAL	3	0	Size
Franches Montagnes	FM	3	0	Genome
Friesian	FRI	3	0	Size
French Trotter	FT	3	0	Genome
Hanoverian	HAN	3	0	Genome
Hokkaido*	HOK	3	0	Genome
Icelandic Horse*	ICE	3	3	Genome
Mongolian*	MON	3	0	Genome
Morgan*	MOR	2	0	Size
Norwegian Fjord	NORF	3	0	Genome
Percheron	PER	3	0	Size
Puerto Rican Paso Fino*	PRF	3	3	Size
Quarter Horse	QH	3	0	Genome
Saddlebred*	SB	6	3	Genome
Shetland Pony	SHP	3	0	Size
Shire	SHR	3	0	Size
Standardbred*	STBD	3	0	Genome
Suffolk Punch	SUF	3	0	Size
Swiss Warmblood	SZWB	3	0	Genome
Thoroughbred	TB	6	0	Genome, Size
Welsh Mountain Pony (Section A)	WMP	3	0	Size

Animals and Phenotypes – TWH Specific

We collected 139 TWH samples from private farms and public horse shows across the US for use in this study. The group consisted of 82 mares and 57 geldings and stallions ranging from 1-34 years in age at time of collection. We video recorded all horses either under saddle, at liberty in a round pen, or in-hand at the walk, and at all intermediate gaits using a Canon FS30 camcorder (Canon Inc.). The camcorder was setup so that it was perpendicular to the horse's plane of travel, approximately at the horse's wither height, and far enough away that the horse fit within the camera's view finder. We recorded each horse for at least 10 seconds at the walk and all intermediate gaits. We also collected level of training under saddle, horse discipline (breeding, trail, show), type of shoes (barefoot, aluminum, steel, pad), and pedigree information. Pedigree information, as contained in the registry or reported by the owner, was recorded for 137 horses; the remaining two horses were registered, but the registration papers including the pedigree had been lost. Horses were selected so that no horse was related to any other within a single generation for each gait type to reduce population stratification. Half-siblings (n=17 pairs) or parent-offspring pairs (n=8 pairs) were only used if one member of each pair had been classified as lateral-only gaited and the other as multi-gaited. Pedigraph v2.4 (Garbe and Da 2008) and the five-generation pedigrees were used to evaluate the inbreeding within the 136 horses. In the two cases where pedigree information was not available, we calculated genome-sharing (pi hat) values and inbreeding (F) for the autosomes in PLINK (Purcell *et al.* 2007) after pruning for minor allele frequency of 0.05 and genotyping rate of 0.05.

A single experienced observer (EAS) classified horses as either lateral-only or multi-gaited (able to perform both lateral and diagonal gaits, including the trot) after review of video recordings in slow-motion. The individual classified horses as lateral-only gaited if the lateral pair of legs traveled together in a four-beat pattern (with the hind hoof landing before the front hoof). If the

footfall landing followed a two-beat or four-beat diagonal pattern, and the lateral gait pattern previously described was also observed, the individual classified the horse as multi-gaited. Horses were also classified as multi-gaited if the owner reported viewing the horse trotting at liberty, but not in the video analysis (n=9). 82 horses were classified as lateral-only and 57 as multi-gaited, with the lateral-only horses considered controls and the multi-gaited horses as cases in PLINK.

DNA Extraction – Across Breed and TWH Specific

We isolated genomic DNA from blood samples using the Gentra® Puregene® Blood Kit, following the manufacturer's protocol for whole blood (Qiagen Inc., Valencia, CA). We performed extraction of DNA from hair using the Gentra® Puregene® DNA Isolation Kit, following the manufacturer's protocol with modifications to optimize for hair root bulbs (Qiagen Inc., Valencia, CA) (Cook, Gallagher, Bailey 2010).

Genotyping and Quality Control – Across Breeds

Genotyping yielded 59,351 loci using the Illumina Equine SNP50 beadchip (Illumina, Inc., San Diego, CA) at GeneSeek Inc. (Lincoln, NE) or the Genotyping Shared Resource at the Mayo Clinic (Rochester, MN). We excluded SNPs from analysis if the genotyping rate was less than 95% (n=13195) or if the minor allele frequency was less than 10% (n=17398) across all individuals. No individuals were removed from the analysis due to poor genotyping (missing >10%). After quality control filtering, we analyzed 98 individuals and 39,208 SNPs. We used a Bonferroni significance cutoff of 1.41×10^{-6}, conservatively estimating 35,380 independent comparisons (3,828 markers that are in complete LD, $r^2>0.99$, and subtracted from our 39,208 total markers in the study). Genome-wide IBD estimates were calculated using an IBS similarity matrix to evaluate population structure; one individual was removed from a pair with an IBS distance greater than 0.90.

Statistical Analysis – Across breeds

PLINK V1.07 (Purcell *et al.* 2007) simultaneously tested the resulting 39,208 genotypes for significant association with naturally gaited breeds within the sample population. We applied a Mixed Model linear analysis (EMMAX) to correct for population stratification and relatedness (Kang *et al.* 2010)(Huson, 2014) using the Golden Helix SVS software. The linkage disequilibrium (LD) structure between markers was examined using Haploview v4.2 (Barrett *et al.* 2005). We visualized statistical results using JMP Pro 11 (SAS Institute, Inc., Cary, NC).

Genotyping and Quality Control – TWH Specific

We genotyped 139 TWH with complete gait phenotypes at 65,157 loci using the Equine SNP70K beadchip (Illumina Inc., San Diego, CA) at GeneSeek Inc. (Lincoln, NE), and added two additional SNPs genotyped in-house as described below. Genome-wide IBD estimates were calculated using an IBS similarity matrix to evaluate population structure after filtering for minor allele frequency and genotyping rate as described above. Twelve pairs (but only 22 individuals) had IBS similarities greater than 0.80; only one pair was greater than 0.90 and the remaining 10 pairs were less than 0.85. We removed five individuals from these pairs; one individual per pair if the pair was of the same gait type or if an individual appeared in two or more pairwise comparisons. SNPs with less than 90% genotyping rate (n=10,035) and with a minor allele frequency <0.05 (n=15,417) were excluded. Five individuals were excluded for a genotyping rate <86%. The final dataset consisted of 129 individuals (52 cases, 77 controls) genotyped at 42,226 markers with a mean call rate of 96.56%. We used a Bonferroni significance cutoff of 1.32×10^{-6}, conservatively estimating 37,770 independent comparisons (4,456 markers that are in complete LD, $r^2>0.99$, subtracted from our 42,226 total markers in the study).

Genome-wide LD was estimated using the r^2 statistic in PLINK (Purcell *et al.* 2007) under the following filters: maf <0.05 and deviation from HWE p<0.0001. Ten individuals previously genotyped on the EquineSNP50 chip were chosen from the Arabian, Eqyptian Arabian,

Thoroughbred and Saddlebred breeds and compared to 10 unrelated TWHs selected from this study. Values were binned in groups of 5000 and average r^2 and inter SNP distance graphed using Excel 2007 (Microsoft Corp., New York, NY).

DMRT3 Amplification and Genotyping

A recent publication describes a mutation in the *DMRT3* gene which the authors claim controls the ability for a horse to perform lateral patterned gaits (Andersson *et al.* 2012). We examined the effect of this mutation in our study population. We performed PCR amplification in a 20μL volume containing 2μL of DNA (diluted to a concentration of 25ng/μL), 2μL of 10X PCR reaction buffer with 20mM $MgCl_2$, 0.2μL of FastStart Taq DNA Polymerase (Roche Diagnostics), 2μL of 2mM dNTP's, 2μL each of forward and reverse 5μM primers, and 9.8μL PCR-grade water. We obtained the primer sequences from Andersson *et al.* (2012) to produce a PCR product of 681bp. We performed the PCR on an Eppendorf Mastercycler Ep Gradient (Eppendorf Corp.) under the following conditions: 95°C for 4 min, followed by 40 cycles of 95°C for 30 sec, 58°C for 30 sec, 72°C for 45 sec, and a final extension of 72°C for 7 min and cooling to 4°C.

For convenience we devised a novel RFLP test for this polymorphism. The restriction digest used 10μL of *DMRT3* PCR product, 0.4 μL Dde I (1.0U per reaction, New England Biolabs Inc. (NEB)), 1 μL 10X NEB CutSmart buffer and 8.6 μL MilliQ water to bring the reaction volume to 20 μL. Following incubation at 37°C overnight, the resulting products were visualized by electrophoresis following standard conditions on a 3% agarose gel (Omnipur Agarose, EMD Chemicals Inc). The wild type (*C* allele) produces 31bp, 73bp and 577bp fragments following digestion, while the novel *A* allele produced 31bp, 73bp, 145bp and 432bp fragments.

EPHB3 Amplification and Genotyping

The Ephrin and EPH receptors have been shown to play several important roles in the development of the CNS, including regulating axon-path growth, cell proliferation, migration and

synaptic plasticity. In a mouse strain exhibiting a hopping gait (similar to rabbits where the left and right hind limbs move synchronously), a single mutation in two genes, *ephrin-B3* or its receptor *EPHA4*, were identified as sufficient to produce the phenotype (Kullander *et al.* 2001a). We examined a non-synonymous mutation identified from whole-genome sequencing of a TWH in our study population for differences in gait type (Al Abri *et al.*). We performed PCR amplification in a 20μL volume containing 2μL of DNA (diluted to a concentration of 25ng/μL), 2μL of 10X PCR reaction buffer with 20mM $MgCl_2$, 0.2μL of FastStart Taq DNA Polymerase (Roche Diagnostics), 2μL of 2mM dNTP's, 2μL each of forward and reverse 5μM primers, and 9.8μL PCR-grade water. We used Primer3 (Rozen and Skaletsky 2000) to design primer sequences around the second exon to produce a PCR product of 308bp and carried out the PCR on an Eppendorf Mastercycler Ep Gradient (Eppendorf Corp.) under the following conditions: 95°C for 4 min, followed by 40 cycles of 95°C for 30 sec, 63°C for 30 sec, 72°C for 30 sec, and a final extension of 72°C for 7 min and cooling to 4°C.

For convenience, we developed a novel RFLP test for this polymorphism. The restriction digest used 10μL of *EPHB3* PCR product, 0.4 μL Sau3AI (1.0U per reaction, New England Biolabs Inc. (NEB)), 1 μL of 10X NEB CutSmart buffer and 8.6 μL MilliQ water to bring the reaction volume to 20 μL, which was incubated at 37°C for 2 hours. The resulting products were visualized by electrophoresis following standard conditions on a 2.5% agarose gel (Omnipur Agarose, EMD Chemicals Inc). The *G* allele is uncut (308bp) following digestion, while the *A* allele produces 109bp and 199bp fragments.

Statistical Analysis – TWH Specific

PLINK V1.07 (Purcell *et al.* 2007), and SVS software from Golden-Helix (reference) simultaneously tested the sample population's resulting 42,226 genotypes for significant association with gait type. We applied case/control allelic association and adaptive permutation to the dataset,

with lateral-only individuals run as controls and multi-gaited as cases due to the higher prevalence of lateral-only horses in our sampling population. Multidimensional scaling was performed using PLINK V1.07 on all SNPs that passed quality control filters using the "--mds-plot 4" option. To assess association to haplotypes, candidate regions of 2 Mb centered on the most significant SNP from ECA11, ECA12, ECA19, and ECA25 were selected and filtered to include only markers with a 99% genotyping rate and MAF >0.01. We tested for a haplotype association to gait type by a general logistic model, permutation, and chi-square test using a sliding window of between 2 and 20 SNPs in PLINK. The LD structure between markers was examined using Haploview v4.2 (Barrett *et al.* 2005). Epistasis between candidate loci and all other SNPs were tested in PLINK (--epistasis). Statistical results were visualized using the JMP Pro v11 software package (SAS Institute Inc., Cary, NC).

Polymorphism Detection from Whole-Genome Sequencing

We collected 801 gaited horse samples from across North and South America, India and China at private farms and public horse shows for this study by personal visits and mail. From our 801 samples, we selected three gaited horses from three diverse breeds for whole genome sequencing on-going in another study (Al Abri *et al.*). The first, a Tennessee Walking Horse gelding also included in the TWH association study, was selected for his preference to perform the more lateral end of four-beat gait spectrum (rack and running walk) and an inability to trot (owner has kept the horse since birth). The second was a Chakouyi (CH) stallion from China able to perform a lateral four-beat gait, including a gait similar to the 'flying pace' observed in Icelandic horses (DNA provided by Dr. Chuzhao Lei's lab at Northwest A&F University, Yangling, Shaanxi, China). The third was a Mangalarga Marchador (MM) stallion from Brazil able to perform a diagonal four-beat gait (Marcha Batida). All three horses were video recorded under saddle for at least 10 seconds at all their intermediate gaits. One experienced individual analyzed the videos in

slow-motion to determine if lateral or diagonal pairs of legs traveled together and if there are any moments of suspension in the intermediate gaits; the analyzer observed no moments of suspension for all three horses (EAS). All three horses were genotyped for the *DMRT3* mutation. In addition to gait type, we also selected the Mangalarga Marchador for his homozygous wild-type *DMRT3* mutation genotype (*CC*); the other two horses were both homozygous for the mutation (*AA* genotype).

As described in Al Abri *et al.,* we submitted 10 μg of DNA from each horse for 150 bp paired-end sequencing on the Illumina HiSeq 2500 operating in the 'rapid run mode' at the Cornell University Core Laboratories Center (Ithaca, NY). The Cornell University Core Laboratories Center (Ithaca, NY) performed the sequencing library construction according to their standard operating procedure. Quality control, alignment and annotation were completed as described Al abri *et al.* (manuscript in progress).

Results

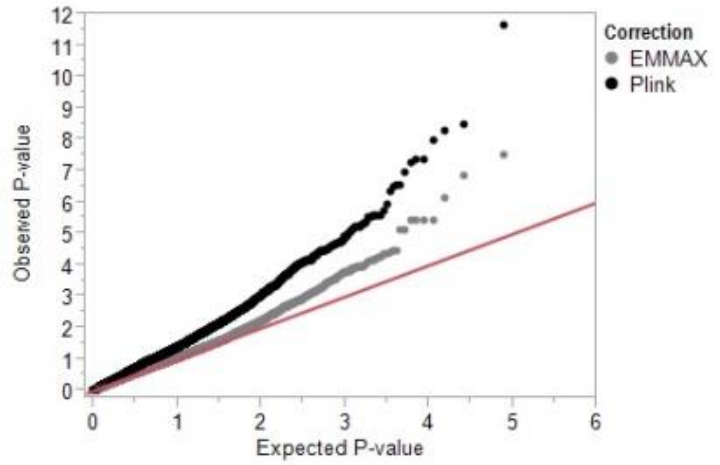
The goal of this study was to identify loci and mutations, other than *DMRT3*, that are contributing to polymorphic gait in the horse. We utilized two different genome-wide association sample populations, the first consisting of 30 diverse breeds that are known to be or not to be polymorphic in gait type, and the second consisting of one breed known to be polymorphic in gait type with several variations.

Three Genetic Loci are Significantly Associated with Variation in Gait Type Across Horse Breeds

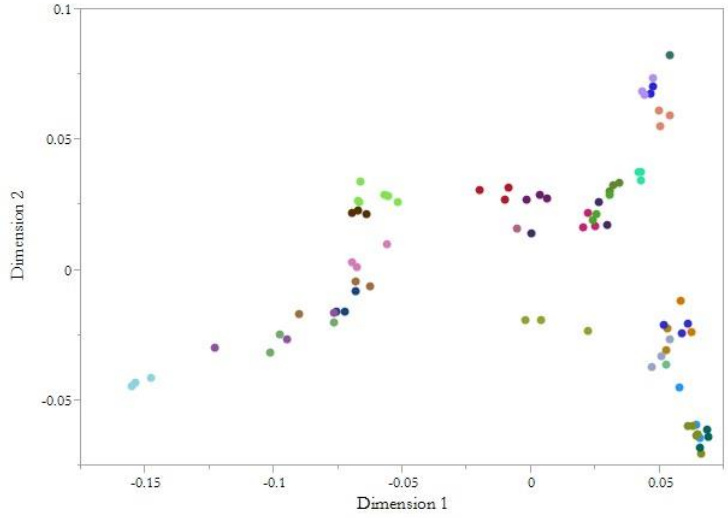
Genotypes were available for 98 horses from a previous project run on the Equine 50K beadchip (Illumina, Inc), enabling a GWA study to identify loci controlling for alternate gaits across breeds. Allelic association in Plink identified 13 candidate markers on ECA1, ECA7, ECA11, ECA16, ECA17, ECA18, ECA19, ECA21, ECA23, and ECA24 with p-values ranging from 2.40e-12 to 1.20e-6 (Table 2.2). However, the genomic inflation factor was 1.69326 before correction, indicating significant population stratification due to breed. This was detectable via the quantile-quantile plot (Figure 2.2a) and breed clusters could be differentiated in plotting of genotype principle components/metric multidimensional scaling of pair-wise genetic distances (Figure 2.2c & d). Dimension 1 separates the draft and pony breeds from the light breeds. Dimension 2 separates the pony breeds from the draft breeds. Dimension 3 separates the Saddlebreds from the other breeds, and dimension 4 separates the largest breeds (Shire and Clydesdale) from the other breeds.

Figure 2.2. Genome wide association across breeds identifies three candidate loci. a) Quantile-quantile plot for the GWA scans. The p-values for the uncorrected allelic association are plotted in black and have a genomic inflation factor of X. The p-values for the EMMAX association are plotted in grey and have a genomic inflation factor of 1.0758. Horse breed population structure were inferred using MDS dimensions of SNP genotypes. The first four dimensions are plotted, (b) dimension 1 and dimension 2 and in (c) dimension 3 and dimension 4. D) Manhattan plot for the GWA scan from the Emmax association. The red horizontal line indicates genome-wide significance with $\alpha<0.05$ and Bonferroni correction for multiple hypothesis testing.

A


Correction
EMMAX
Plink
Observed P-value
Expected P-value

B


Dimension 2
Dimension 1

C

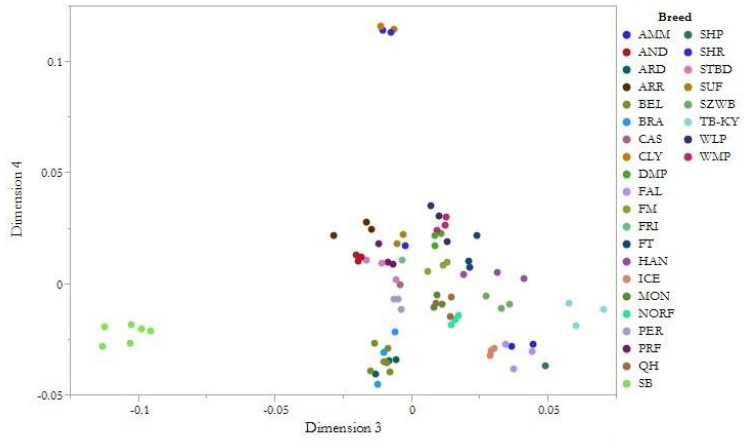
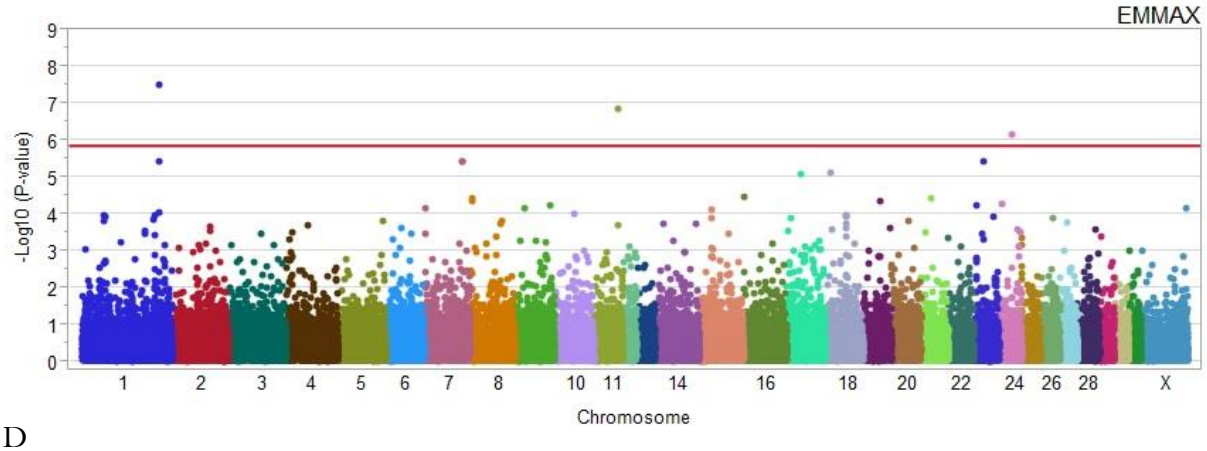

Breed
AMM
AND
ARD
ARR
BEL
BRA
CAS
CLY
DMP
FAL
FM
FRI
FT
HAN
ICE
MON
NORF
PER
PRF
QH
SB
SHP
SHR
STBD
SUF
SZWB
TB-KY
WLP
WMP
Dimension 4
Dimension 3

Figure 2.2. (Continued)



D

Table 2.2. Across-breed GWA associated SNPs.

SNP	Chr	BP	Major Allele	Minor Allele	Raw P-value	EMMAX P-value	Markers in/near Genes
BIEC2-159954	11	53413674	G	A	2.40E-12	1.435E-07*	*PIRT, SHISA6*
BIEC2-67309	1	155759738	A	G	3.43E-09	3.224E-08*	*OLR14, OLR778, OLR749*
BIEC2-620109	23	22967656	G	A	5.38E-09	3.859E-06	*DMRT1, DMRT3*
BIEC2-641566	24	27856416	A	G	1.12E-08	7.265E-07*	*CSK, DDX28*
BIEC2-1008603	7	82561206	A	G	4.53E-08	3.684E-06	*FAR1*
BIEC2-1008622	7	82582932	G	A	4.53E-08	3.684E-06	*SPON1*
BIEC2-375024	17	30210377	A	G	5.61E-08	8.194E-06	1.05Mb region with no annotation
BIEC2-159977	11	53420235	C	A	1.09E-07	2.104E-04	*PIRT*
BIEC2-557401	21	23968210	G	A	2.93E-07	4.031E-05	*GHR*
BIEC2-401575	18	12161633	C	A	2.93E-07	7.981E-06	*EN1*
BIEC2-69568	1	158412743	C	A	3.21E-07	3.959E-06	*OLR291*
BIEC2-436491	19	34516426	G	A	4.76E-07	4.430E-05	*SLC12A8*
BIEC2-326760	16	368598	A	C	1.20E-06	3.595E-05	*MGLL*

The EMMAX model applied to the same dataset reduced genome-wide inflation and was able to correct for some of the population substructure, as observed in the QQ-plot (Figure 2.2a), decreasing genomic inflation to a factor of 1.07581. The EMMAX model verified nine of the candidate markers on ECA1, ECA7, ECA11, ECA16, ECA17, ECA18, ECA23, and ECA24 (Figure 2.2d) with p-values ranging from 3.22e-8 to 8.19e-6 (Table 2.2), but only the three markers on ECA1, ECA11 and ECA24 surpassed our Bonferroni genome-wide significance cutoff of 1.32e-6. The ECA1 marker lies within a large 25 Kb intronic region of several olfactory receptor genes identified in the rat (OLR14, OLR778, and OLR749). The ECA11 marker falls within a 365 Kb segment devoid of annotated genes; the *PIRT* gene is approximately 125 Kb downstream of the significant marker, and the *SHISA6* gene is approximately 248 Kb upstream from the marker. The ECA24 marker did not contain any annotated genes for an interval of 26,838,220 to 28,120,719 bp surrounding the significant marker. There is a plant specific invertase/pectin methylesterase inhibitor gene located 6.7 Kb upstream of the marker, with the vertebrate genes *CSK* located 264 Kb upstream of the marker and *DDX28* located 1.02 Mb downstream of the marker. The ECA23 marker falls within the intron of DMRT1/DMRT3, 31,999 bp downstream from the known mutation (Andersson *et al.* 2012).

Permutation Identifies a Suggestive Candidate Locus for Gait Type in the TWH

Genotypes generated for 139 horses on the Equine 70K beadchip (Illumina, Inc) enabled a GWA study to identify loci contributing to gait type within TWH. Allelic association identified several candidate markers on ECA19, ECA12, ECA25, ECA11, and ECA8 with raw p-values ranging from 1.71e-5 to 1.46e-4 (Table 2.3); however, none surpassed Bonferroni significance. Examination of the quantile-quantile plot indicates very little association across the genome (Figure 2.3a), likely due to insufficient power in such a small sample size. After one million permutations,

the top candidate locus from the allelic association on ECA19 (Figure 2.4b) had a p-value of 1.50e-5 (Table 2.3).

The two top candidate markers on ECA19 are in complete LD and span just 51 bp within an approximately 2 Kb intron of *FBX040*. Additional loci on ECA19 are approximately 26 Mb downstream from this region, spanning 23 Kb in a 967 Kb region that did not contain any annotated genes (Figure 2.4). The associated locus on ECA12 falls within a 220 Kb region lacking any annotated features. The nearest genes include *NUP160* (93 Kb downstream), *FNBP4* (148 Kb downstream), and *PTPRJ* (129 Kb upstream). The locus on ECA25 falls within a 207 Kb also without known features. The nearest genes include *RAD23B* (62 Kb upstream), *OOEP* (132 Kb upstream), and *ZNF462* (148 Kb downstream). The locus on ECA11 falls within a 315 bp intron of *GGT6*. There was no evidence for epistatic interaction between the top allelic association loci from ECA19, ECA25, ECA11, and ECA12 as none of the tests were significant at P=1.00e-4.

Haplotype analysis revealed associations of gait type with a three SNP block spanning ECA19: 37,674,757-37,719,196 and a seven SNP block spanning ECA11: 47,893,127-48,015,457 (uncorrected P=3.79e-5 and P=3.92e-5, respectively, with Bonferroni cutoffs of 6.25e-4 and 5.95e-4, respectively). The ECA19 SNP block genomic region includes the *FBXO40, ARGFX,* and *POLQ* genes (Figure 2.5a). The ECA11 SNP block genomic region includes the *SPNS3, SPNS2, MYBBP1A, GGT6,* and *SMTNL2* genes (Figure 2.5c). There was no evidence for an epistatic interaction between the ECA19 and ECA11 loci as none of the tests were significant at P=1.00e-4.

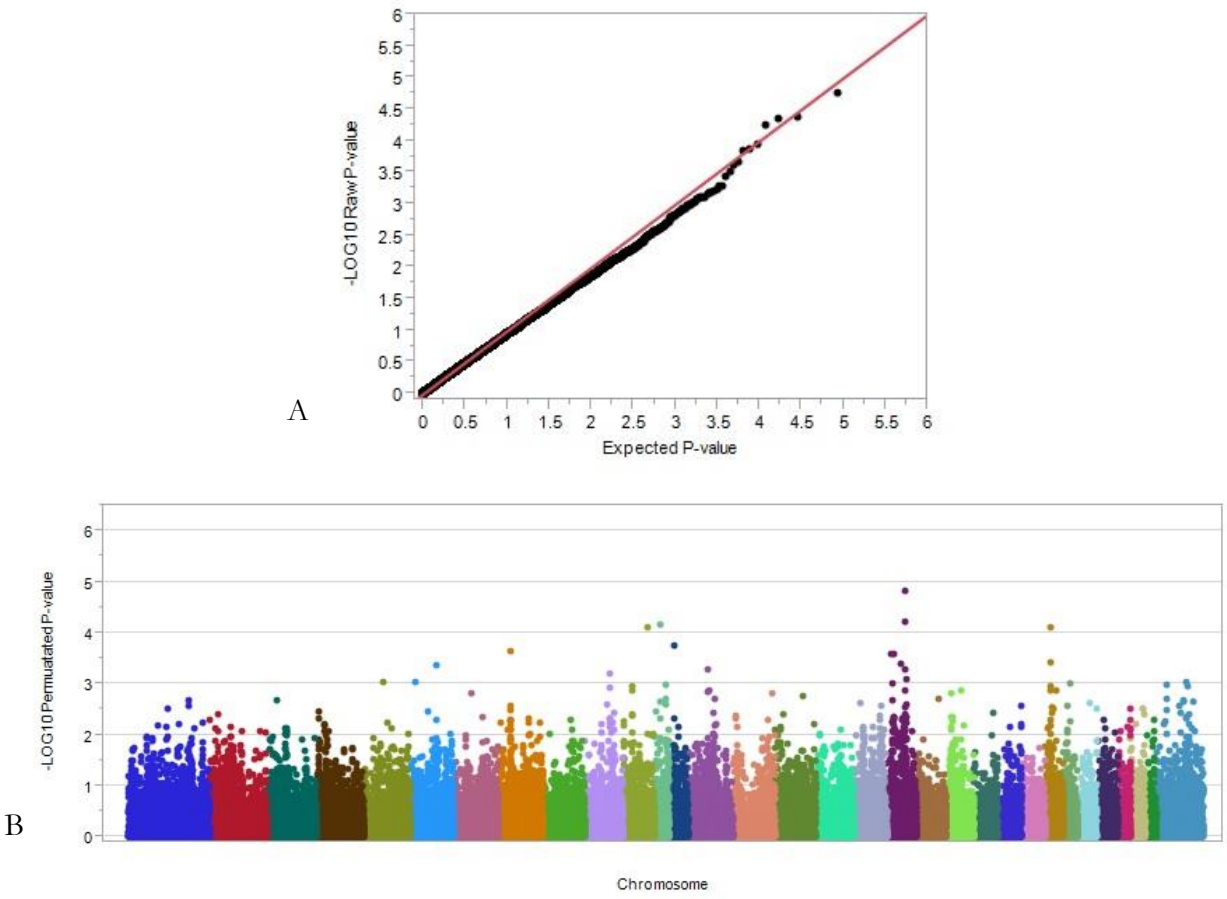


Figure 2.3. Genome wide association analysis identified peaks on ECA19 and ECA11 for gait types. A) Quantile-quantile plot of the allelic association model illustrates no association due to lack of deviation of the expected red line. B) Manhattan plot of the permutated –Log10 p-values.

Table 2.3. TWH GWA SNPs associated with gait type.

SNP	Chr	BP	Major Allele	Minor Allele	Raw P-value	Permutated P-values	# Permutations	Gene in Locus
19	BIEC2_437782	37674807	G	A	1.71E-05	1.50E-05	1000000	*FBXO40*
19	BIEC2_437781	37674757	C	T	4.12E-05	6.06E-05	594405	*FBXO40*
12	BIEC2_175614	12173695	T	C	4.60E-05	7.08E-05	508491	*NUP160, PTPRJ*
25	BIEC2_659946	12939295	C	A	5.58E-05	7.83E-05	460000	*RAD23B*
11	BIEC2_156863	47974247	G	A	0.000115	7.89E-05	456543	*GGT6*
19	BIEC2_456655	30903484	A	C	0.000134	0.000417	88822	*HES4,HES1*
8	BIEC2_1037747	24686529	G	A	0.000146	0.000225	160195	*AACS, TMEM132B*
25	BIEC2_658647	10385019	A	C	0.000223	0.000383	93906	*OLFR275, OLR845*
13	BIEC2_210428	11381404	T	C	0.000239	0.000174	207376	*CLDN3, CLDN4*
19	BIEC2_428949	11617833	A	G	0.000311	0.00026	138583	*TNIK/MINK1*
10	BIEC2_124370	53570894	T	G	0.00037	0.000636	56601	*ZP4*
19	BIEC2_429943	13181184	A	G	0.00052	0.001	36000	*PRKRIPL, SPATA16*
19	BIEC2_452486	16010941	A	G	0.001159	0.000264	136178	967Kb region with no annotation
19	BIEC2_430358	16034924	T	G	0.001159	0.000264	136178	967Kb region with no annotation

Figure 2.4. The 26Mb region surrounding the candidate markers on ECA19 spans several candidate genes (image from the UCSC genome bowser, http://genome.ucsc.edu (Kent *et al.* 2002)). A) The top candidate markers and B) the additional markers. C) Gene ontology of both regions identifies several important biological functions based on human annotation (Panther v9.0 (Mi, Muruganujan, Thomas 2013)).

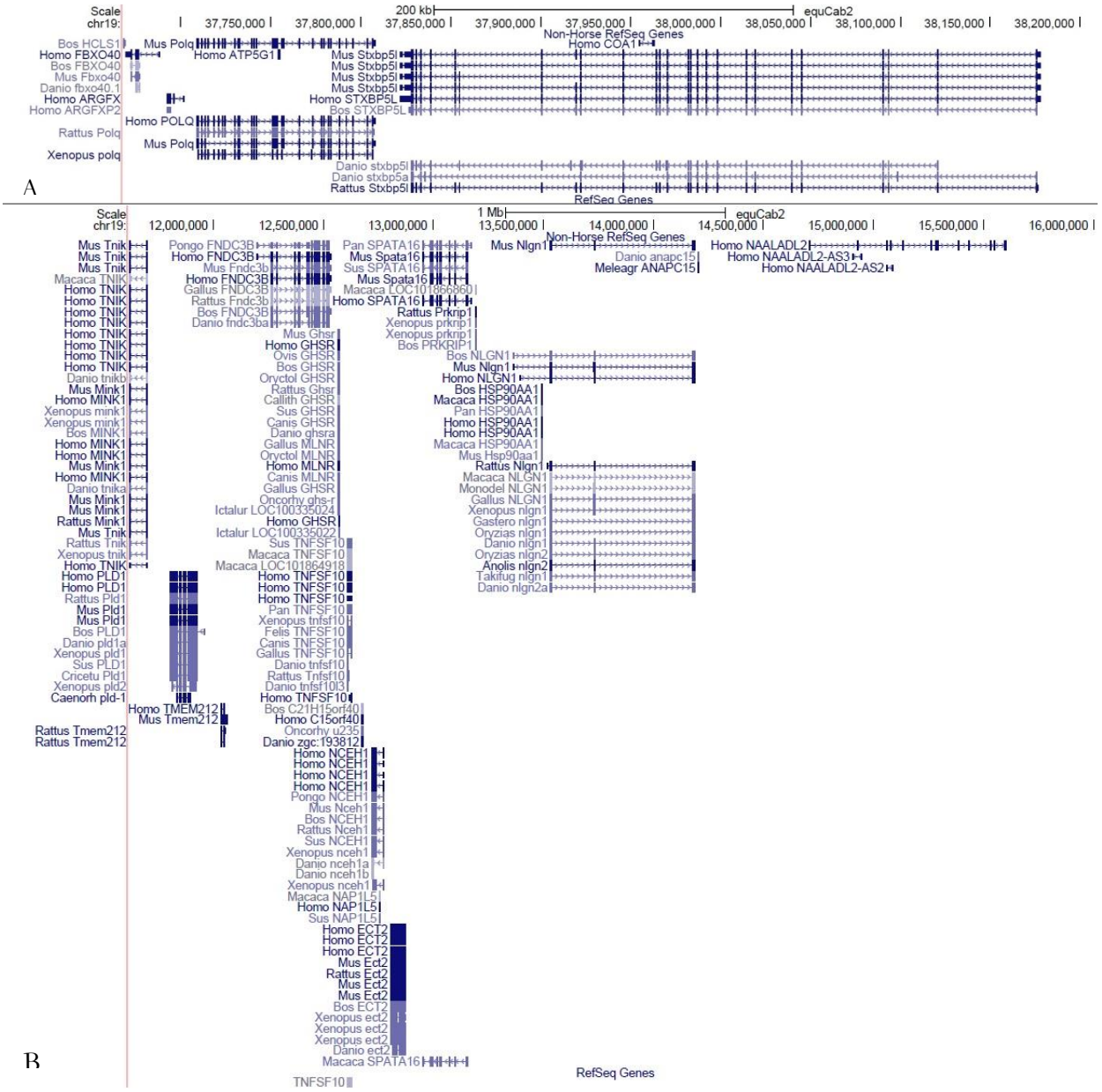
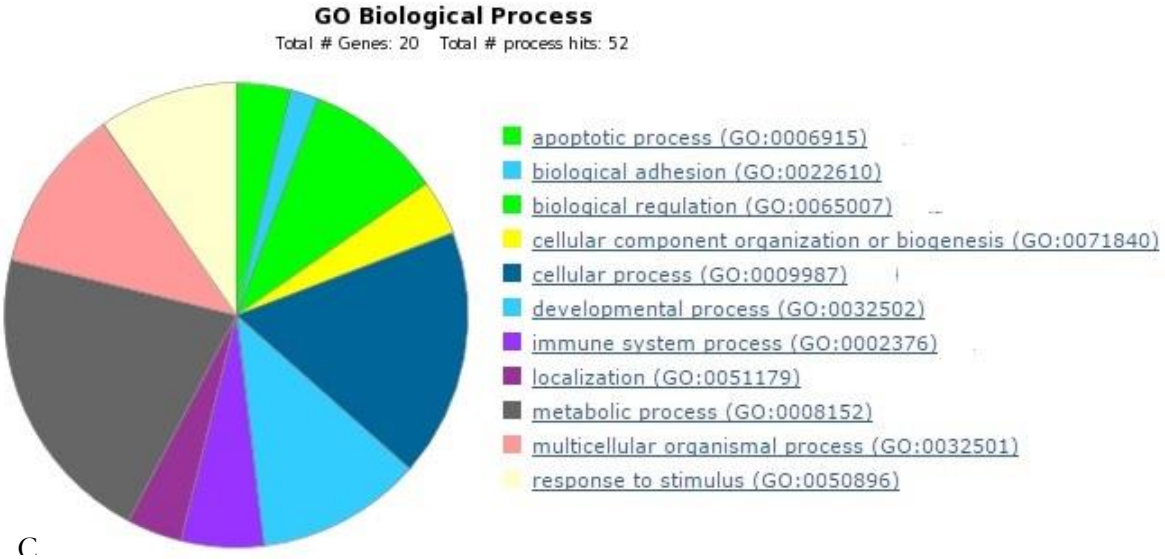

A
Scale
chr19:
200 kb
equCab2
37,750,000
37,800,000
37,850,000
37,900,000
37,950,000
38,000,000
38,050,000
38,100,000
38,150,000
38,200,000
Non-Horse RefSeq Genes
Bos HCLS1
Homo FBXO40
Bos FBXO40
Mus Fbxo40
Danio fbxo40.1
Homo ARGFX
Homo ARGFXP2
Rattus Polq
Xenopus polq
Mus Polq
Homo ATP5G1
Homo COA1
Mus Stxbp5l
Homo STXBP5L
Bos STXBP5L
Homo POLQ
Danio stxbp5l
Danio stxbp5a
Rattus Stxbp5l
RefSeq Genes
B
Scale
chr19:
1 Mb
equCab2
12,000,000
12,500,000
13,000,000
13,500,000
14,000,000
14,500,000
15,000,000
15,500,000
16,000,000
Non-Horse RefSeq Genes
Mus Tnik
Macaca TNIK
Homo TNIK
Danio tnikb
Xenopus tnik
Rattus Tnik
Mus Mink1
Homo MINK1
Xenopus mink1
Bos MINK1
Danio tnika
Rattus Mink1
Homo PLD1
Rattus Pld1
Mus Pld1
Bos PLD1
Danio pld1a
Xenopus pld1
Sus PLD1
Cricetu Pld1
Xenopus pld2
Caenorh pld-1
Homo TMEM212
Mus Tmem212
Rattus Tmem212
Pongo FNDC3B
Homo FNDC3B
Mus Fndc3b
Gallus FNDC3B
Rattus Fndc3b
Bos FNDC3B
Danio fndc3ba
Mus Ghsr
Homo GHSR
Ovis GHSR
Bos GHSR
Oryctol GHSR
Rattus Ghsr
Callith GHSR
Sus GHSR
Canis GHSR
Danio ghsra
Gallus MLNR
Oryctol MLNR
Homo MLNR
Canis MLNR
Gallus GHSR
Oncorhy ghs-r
Ictalur LOC100335024
Homo GHSR
Ictalur LOC100335022
Sus TNFSF10
Macaca TNFSF10
Macaca LOC101864918
Homo TNFSF10
Pan TNFSF10
Xenopus tnfsf10
Felis TNFSF10
Canis TNFSF10
Gallus TNFSF10
Danio tnfsf10
Rattus Tnfsf10
Danio tnfsf10l3
Bos C21H15orf40
Homo C15orf40
Oncorhy u235
Danio zgc:193812
Homo NCEH1
Pongo NCEH1
Mus Nceh1
Bos NCEH1
Rattus Nceh1
Sus NCEH1
Xenopus nceh1
Danio nceh1a
Danio nceh1b
Xenopus nceh1
Macaca NAP1L5
Homo NAP1L5
Sus NAP1L5
Homo ECT2
Mus Ect2
Rattus Ect2
Bos ECT2
Xenopus ect2
Danio ect2
Macaca SPATA16
Pan SPATA16
Mus Spata16
Sus SPATA16
Macaca LOC101866860
Homo SPATA16
Rattus Prkrip1
Xenopus prkrip1
Bos PRKRIP1
Mus Nlgn1
Homo NLGN1
Bos NLGN1
Macaca NLGN1
Monodel NLGN1
Gallus NLGN1
Xenopus nlgn1
Gastero nlgn1
Oryzias nlgn1
Danio nlgn1
Oryzias nlgn2
Anolis nlgn2
Takifug nlgn1
Danio nlgn2a
Danio anapc15
Meleagr ANAPC15
Homo NAALADL2
Homo NAALADL2-AS3
Homo NAALADL2-AS2
Homo HSP90AA1
Macaca HSP90AA1
Pan HSP90AA1
Mus Hsp90aa1
Rattus Nlgn1
RefSeq Genes
TNFSF10

Figure 2.3 (Continued)



C.

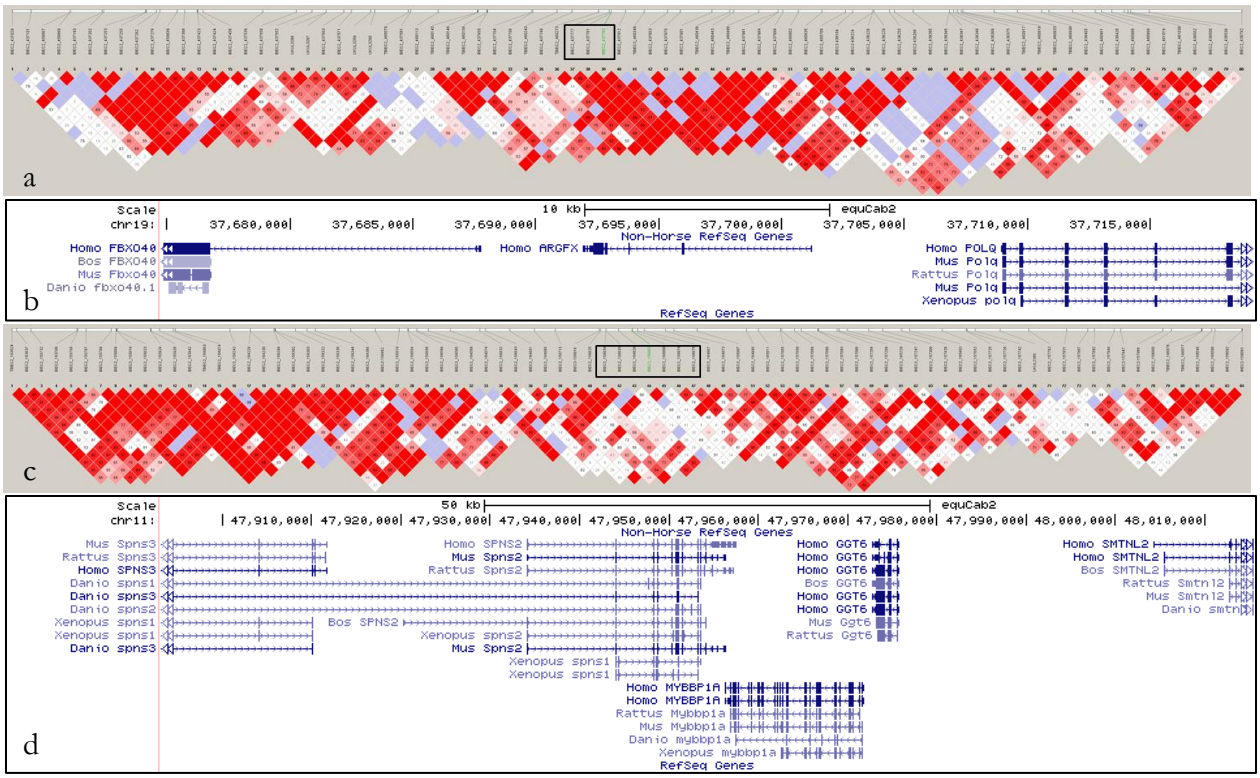
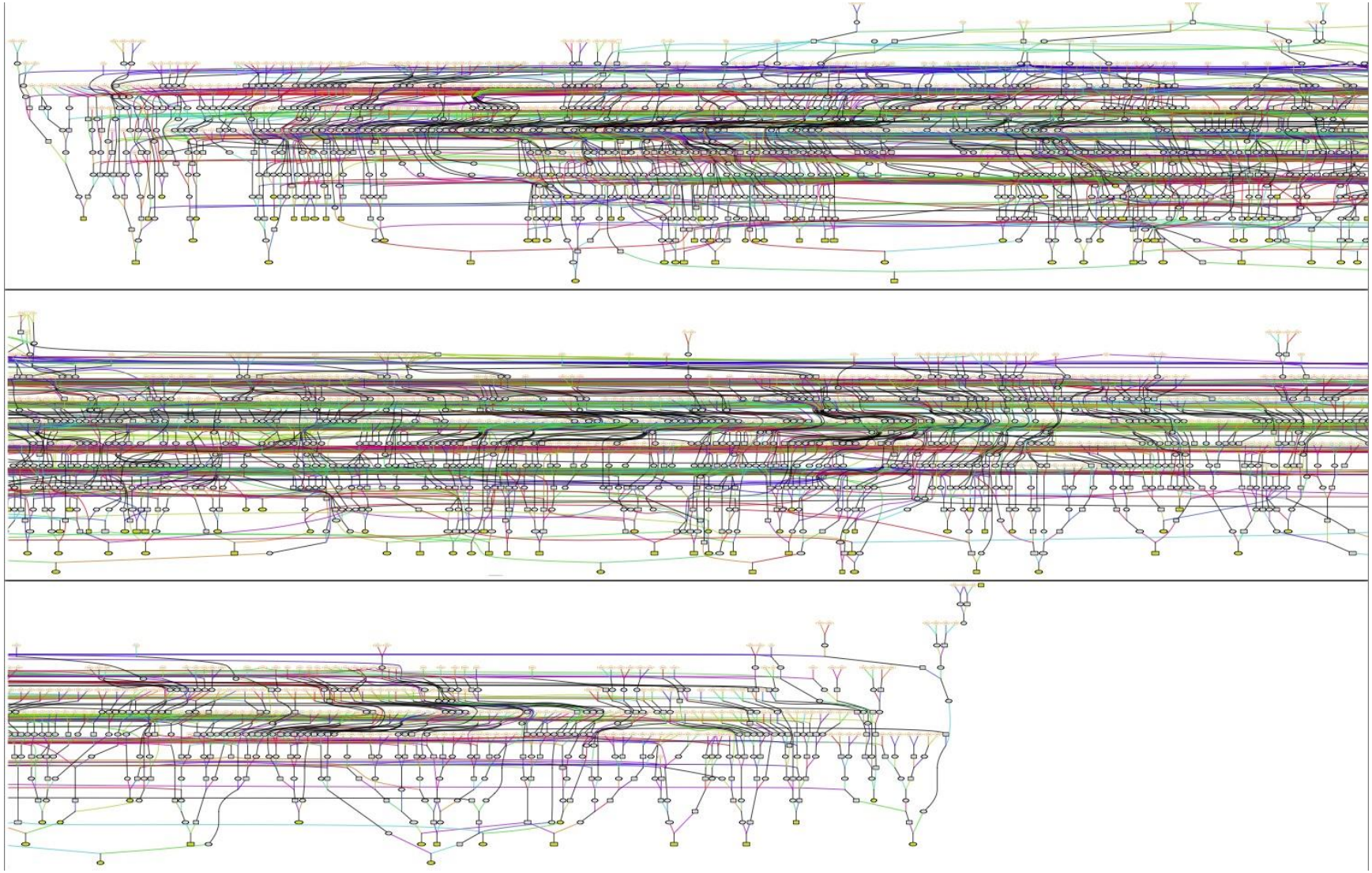


Figure 2.5. Haplotype association identified a a) three SNP haplotype block outlined in black that b) spans several genes on ECA19 based on gene annotation in UCSC. Haplotype association on ECA11 identified a c) seven SNP block not in LD, outlined in black, that d) spans several genes based on gene annotation in UCSC.

TWH Population Structure Does Not Influence Gait Type Preference

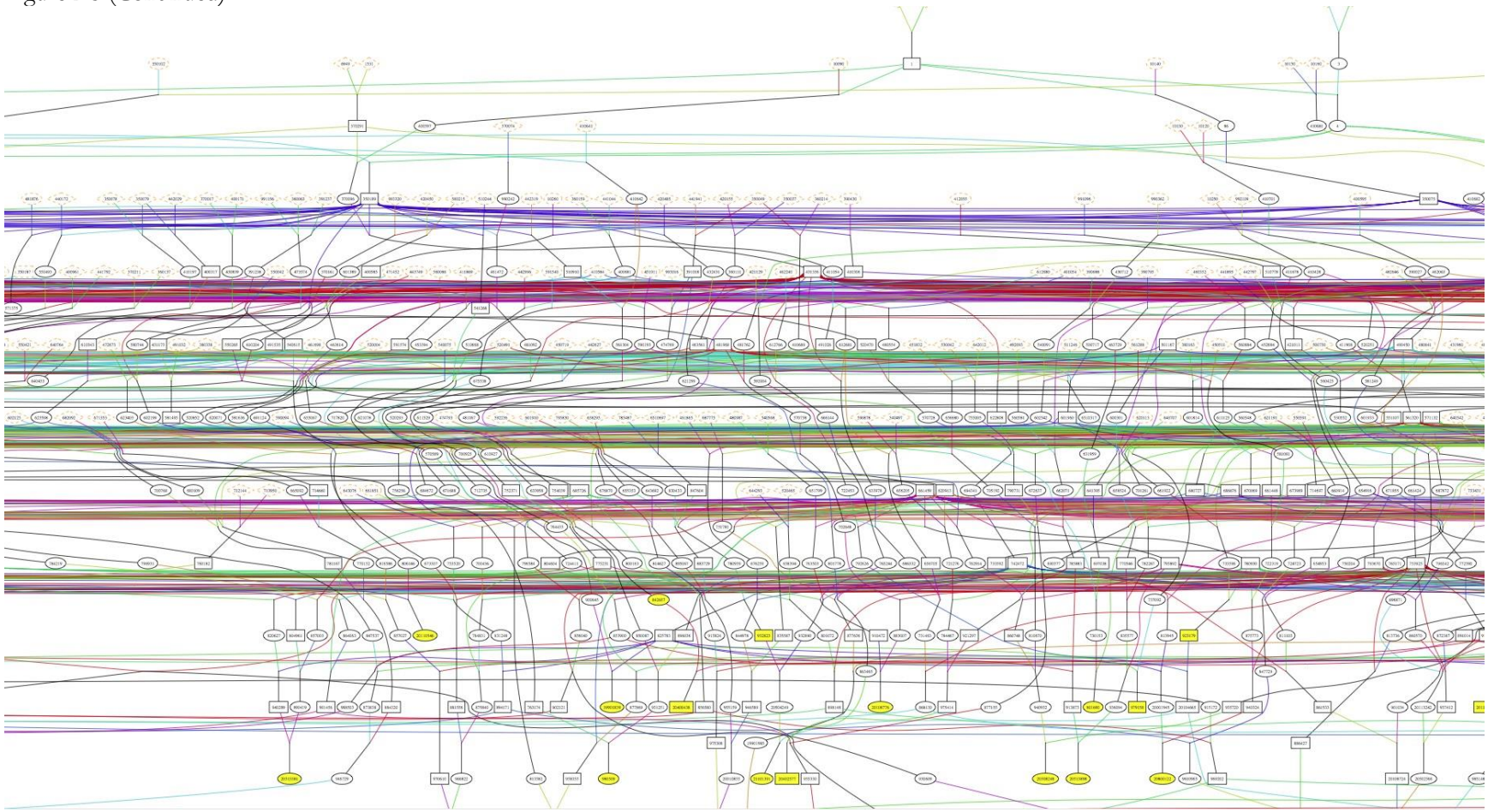
The TWH is recognized as a breed with moderate to high levels of inbreeding due to common line-breeding practices early in breed development (Fletcher 1946), which can also be observed in the five-generation pedigree of horses utilized in this study (Figure 2.6a). The available genome-wide genotypes and pedigrees allowed us to further examine what, if any substructure within the breed may be influencing gait type preference within individuals, as observed in the Icelandic horse (Albertsdóttir *et al.* 2008). Genomic inflation was limited to a factor of 1, indicating adequate control of population stratification. Further examination of the MDS analysis did not identify significant population stratification due to gait type in the TWH (Figure 2.7a and b). However, two subpopulations are detectable in the comparison of dimensions 1 and 3, but these are not correlated with gait-type (Figure 2.7c). The two gait types are evenly distributed throughout all clusters, indicating that either the subpopulations are not altering the alleles contributing to polymorphic gait type in the breed or there are several different alleles that can give rise to the same phenotype. The identified subpopulations are likely due to popular sire effects as observed in the pedigree tree (Figure 2.6b) and could explain the slight deviation observed higher in the quantile-quantile plot (Figure 2.3a). Pedigree inferred inbreeding coefficients were moderate in both lateral-only and multi-gaited horses, 7.9% (range: 23.5%-0.32%) and 8% (range: 23.4%-0.30%) respectively and were not different between the two groups (Students T-test, $p=0.92$). The two individuals without pedigrees had low average pi hat values of 0.002744 and 0.003671 compared to the sample mean of 0.037314 (max pi hat value=0.792, min=0).

Figure 2.6. Pedigree tree of 139 TWH included in the genome wide association for gait type. A) The full pedigree tree displays the level of inbreeding across individuals (F=7.9%), while B) a zoomed view of the pedigree tree highlights the practice of popular sire use within the sample population. Boxes indicate males, ovals indicate females, and the small yellow blocks highlight the horses included in the GWA.



A

Figure 2.5 (Continued)



B

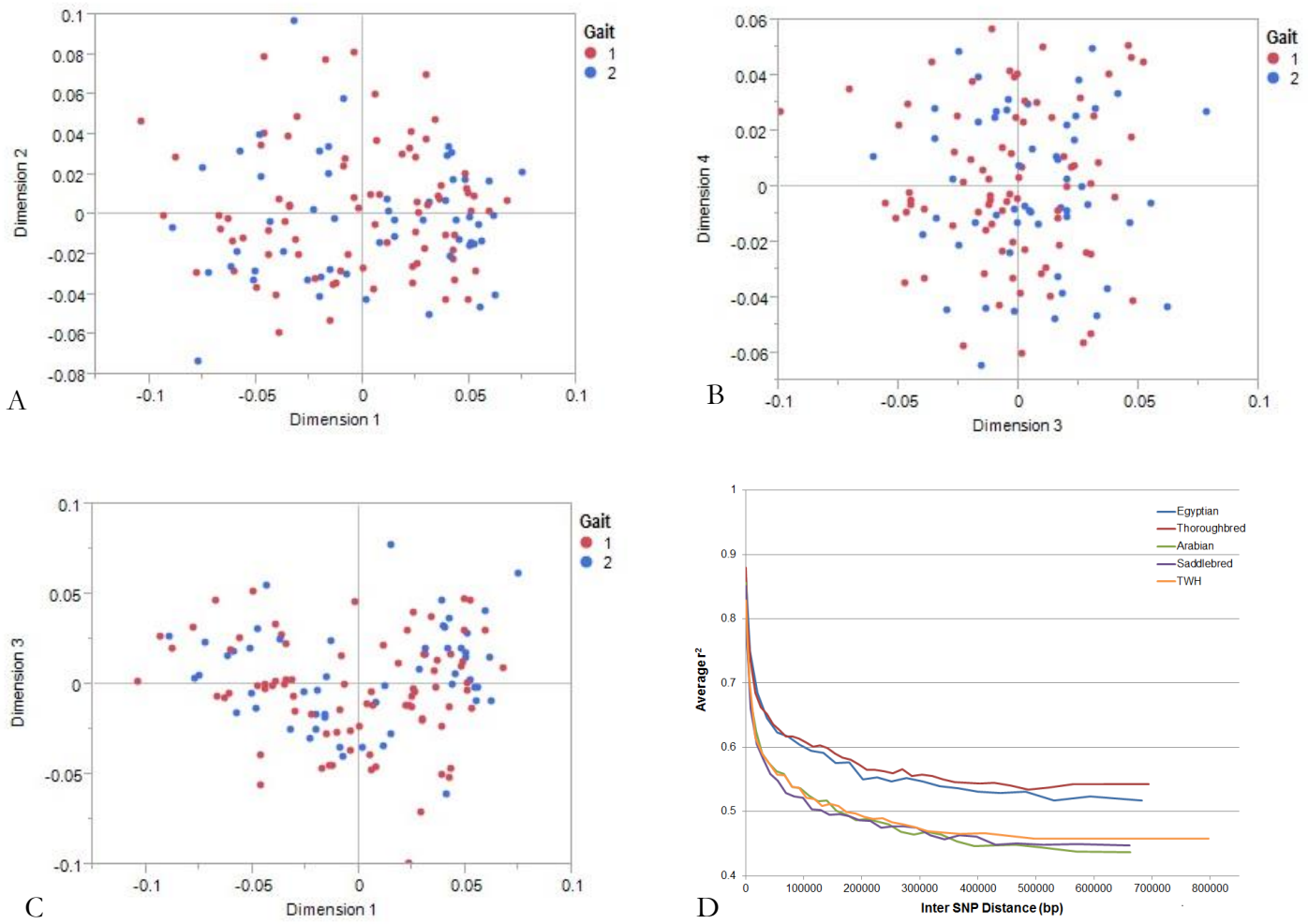


Figure 2.7. Population structure analysis of the TWH. Genome wide association MDS A) dimensions 1 versus 2, B) dimension 3 versus 4, and C) dimensions 1 versus 3. Red dots represent lateral-only gaited horses, blue dots multi-gaited horses. D) Average genome-wide LD in four different breeds, with two subpopulations from one breed. The TWH LD is calculated from genotypes generated on the Illumina SNP70 beadchip; all other breeds on the Illumina SNP50 beadchip.

Ten unrelated individuals from this study, as well as an additional 10 unrelated individuals from the Eqyptian Arabian, Thoroughbred, Arabian (non-Eqyptian) and Saddlebred breeds were used to calculate average genome-wide LD (Figure 2.5d). LD in the TWH is shorter than that of the Eqyptian Arabian and Thoroughbred, but is most similar to the Arabian (non-Egyptian) and the Saddlebred.

Polymorphisms Discovered by Whole-Genome Sequencing Support the ECA19 Candidate Locus

Comparison of the three gaited horse whole-genome sequences to three non-gaited individuals from Arabian, Percheron, and American Miniature (sequenced for another project) identified several SNPs segregating with gait within the candidate regions identified in the both the across-breed and TWH-specific GWAS (Table 2.4). The *DTX3L* on ECA19 has 12 non-synonymous SNPs shared between the gaited horses, but not the non-gaited. On ECA11, we identified seven non-synonymous SNPs, five in different olfactory genes and two in other genes, as well as a start-gained in *ABCA9* (ATP-binding cassette, sub-family A, member 9). No functional SNPs were identified on ECA1 that were homozygous alternate in the gaited horses and homozygous reference in the non-gaited horses.

Table 2.4. Whole genome sequencing SNPs from regions identified in genome-wide associations.

Chr	BP	Major Allele	Minor Allele	Predicted Effect	Gene	CH Geno	MM Geno	TWH Geno	Mini Geno	Perch Geno	Arab Geno
11	11789294	T	C	START_GAINED	*ABCA9*	1/1	0/1	0/1	0/0	0/0	0/0
11	20297978	T	C	NON_SYNONYMOUS	*WNK4*	1/1	0/1	0/1	0/0	0/0	0/0
11	45418023	A	G	NON_SYNONYMOUS	*SERPINF1*	1/1	0/1	1/1	0/0	0/0	0/0
11	46743050	T	C	NON_SYNONYMOUS	Olfactory genes	0/1	1/1	0/1	0/0	0/0	0/0
11	46743368	G	A	NON_SYNONYMOUS	Olfactory genes	0/1	0/1	0/1	0/0	0/0	0/0
11	46845680	G	A	NON_SYNONYMOUS	Olfactory genes	0/1	0/1	0/1	0/0	0/0	0/0
11	46876073	G	A	NON_SYNONYMOUS	Olfactory genes	0/1	0/1	0/1	0/0	0/0	0/0
11	46912956	G	A	NON_SYNONYMOUS	Olfactory genes	0/1	0/1	0/1	0/0	0/0	0/0
19	8017732	T	G	NON_SYNONYMOUS	*ZBBX*	1/1	0/1	0/1	0/0	0/0	0/0
19	8058766	T	A	NON_SYNONYMOUS	*ZBBX*	0/1	1/1	1/1	0/0	0/0	0/0
19	19501811	G	A	NON_SYNONYMOUS	*CCDC39*	0/1	0/1	0/1	0/0	0/0	0/0
19	27953788	G	A	NON_SYNONYMOUS	Predicted	1/1	0/1	1/1	0/0	0/0	0/0
19	36830875	C	T	SYNONYMOUS	*PARP15*	0/1	0/1	0/1	0/0	0/0	0/0
19	36885894	C	A	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36885929	C	T	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36885969	C	T	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36886693	C	A	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36886965	T	C	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36887036	G	A	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36887102	T	G	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36887247	G	A	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36887292	G	T	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36887316	G	A	NON_SYNONYMOUS	*DTX3L*	0/1	0/1	0/1	0/0	0/0	0/0
19	36887767	C	T	NON_SYNONYMOUS	*DTX3L*	1/1	0/1	0/1	0/0	0/0	0/0
19	36887807	A	G	NON_SYNONYMOUS	*DTX3L*	1/1	0/1	0/1	0/0	0/0	0/0
19	43942086	A	G	NON_SYNONYMOUS	*KIAA1407*	0/1	0/1	0/1	0/0	0/0	0/0
25	2521541	G	A	NON_SYNONYMOUS	*GRHPR*	1/1	1/1	0/1	0/0	0/0	0/0
25	3149552	C	T	EXON (MODIFIER|pseudogene)	Predicted	0/1	1/1	0/1	0/0	0/0	0/0
25	4205931	G	T	NON_SYNONYMOUS	*SPATA31/OLF*	0/1	-/-	1/1	0/0	0/0	0/0
25	35513566	T	C	NON_SYNONYMOUS	ABO Blood group	0/1	0/1	-/-	0/0	-/-	0/0
25	36804320	G	C	NON_SYNONYMOUS	*FCN2*	1/1	1/1	1/1	0/0	0/0	0/0

Beyond the GWAs candidate regions, we identified 11 functional SNPs uniquely present in the three gaited individuals (Table 2.5). Two out of the 11 are stop-gained on ECA7 and ECA25; the remaining SNPs are frameshifts on ECA7, ECA10, ECCA16, ECA17, ECA18, and ECA20. By chance, we would expect to find approximately 9.6 million out of 10 million SNPs that are unique in three out of three horses.

Candidate Loci, Association and Frequency Estimates

All horses in this study were genotyped for the *DMRT3* mutation (Andersson *et al.* 2012). 138 horses had the AA genotype and 1 horse had the CA genotype indicating that *DMRT3* is not a predictor for gait type within TWH (P=0.2227). All 139 TWH GWA horses were genotyped for the newly discovered *EPHB3* SNP. 50 horses had the AA genotype, 72 horses had the GA genotype, and 17 horses had the GG genotype, which was not significant for gait type (P=0.5887).

Table 2.5. Novel whole-genome SNPs potentially associated with gait compared across six breeds. CH, MM, and TWH are gaited individuals, while Min, Perch, and Arab are not gaited individuals.

Chr	BP	Major Allele	Minor Allele	Predicted Effect	CH Geno	MM Geno	TWH Geno	Mini Geno	Perch Geno	Arab Geno
7	4632414	GC	G	FRAME_SHIFT	-/-	-/-	0/1	-/-	0/0	-/-
7	21820317	C	A	STOP_GAINED	0/1	0/1	0/1	0/0	0/0	0/0
10	10899078	TC	T	FRAME_SHIFT	1/1	1/1	1/1	-/-	-/-	-/-
16	2925878	CAA	C	FRAME_SHIFT	0/1	1/1	0/1	0/0	0/0	0/0
17	32765644	CA	C	FRAME_SHIFT	0/1	0/1	1/1	0/0	0/0	0/0
17	46325120	A	AT	FRAME_SHIFT	1/1	1/1	1/1	-/-	-/-	-/-
18	49218698	G	GAA	FRAME_SHIFT	0/1	1/1	0/1	0/0	-/-	0/0
18	49218702	AAATC	A	FRAME_SHIFT	0/1	1/1	0/1	0/0	-/-	0/0
18	49218707	GC	G	FRAME_SHIFT	0/1	1/1	0/1	0/0	-/-	0/0
20	28711656	TG	T	FRAME_SHIFT	0/1	0/1	0/1	0/0	0/0	0/0
25	28052634	C	T	STOP_GAINED	0/1	1/1	0/1	0/0	0/0	0/0

Discussion

We identified several genomic regions that are likely associated with gait type across several different breeds. We also demonstrate the importance of precise phenotyping in a complex trait, such as with gait type. Within a small sample set from a single breed, we were able to detect a suggestive association with gait type after correction with permutation that warrants further investigation. We were unable to detect any Bonferroni significant regions associated with gait type, likely due to the continuous nature of gait type. Additional research in a larger sample size, utilizing a denser SNP array, and employing a more quantitative means of measuring the phenotype would improve the power of association and likely confirm our identified regions.

A cross-breed GWA approach can lead to spurious associations due to population stratification, and SNPs identified can represent genetic drift or other loci for traits also selected within the same breeds. However, a low genome-wide inflation factor of 1.07 indicates quality filters were adequate to account for population stratification effects due to breed differences and that our significant associations were due to differences in gait type.

TWH Specific GWAS

The intermediate gaits of the Tennessee Walking horse are a continuous spectrum that blend together and fluctuate based on speed, terrain, as well as influence and skill level of the rider. In our study we attempted to phenotype based on owner reported gaits and a short video of the intermediate gaits, and then grouped horses based on presence/absence of diagonal pair footfall landings. All horses within the TWH study were able to utilize lateral pair footfall landings, indicating that alleles conferring a lingering ability to utilize a diagonal footfall pattern are at a low frequency in this breed. Nine of the horses may have been misclassified as able to trot due to the inexperience of some owners; we sampled horses from a broad range of owner experience level. For some owners, this was their first gaited horse, and for others they had been breeding, training,

and/or showing gaited horses for over 60 years. We also sampled from owners who were not involved with the daily care of their horses (i.e. boarded their horses or kept the horse with a professional trainer). Due to this difference in owner experience, some of the horses may have been able to perform a broader spectrum of the intermediate gaits but the owner didn't know how to ask for it or had never observed the horse out in pasture.

Variation in horse age (1-34 years), and therefore training and experience level, also introduces error in our phenotyping scheme. Older horses are more likely to have more training and are often better conditioned to perform one end of the gait spectrum. Horses trained to stay in one gait type over another are likely to have adapted muscle conformations and muscle memory for that specific gait as skeletal muscle properties are dynamic and reflect the history of use (Riley and Van Dyke 2012; Rivero 2007; Snow and Valberg 1994) and differences in muscle conformation and mass can influence how the horses perform their gaits (Ziegler 2005). Horses with extensive training are typically used for showing, which in turn means they are also more likely to wear associated equipment, like heavier shoes, to enhance the lateral end of the gait spectrum and discourage the use of the trot by increasing the swing phase of the limb (Clayton 2004). Shoeing with weights and wedges (as observed in lite shod, plantation and performance show horses) also influences tendon tension by reducing the tension, prolonging the breakover of the stride and changing the interaction between tendons and muscles (Biewener 1998; Clayton 2004; Lawson *et al.* 2007; Riemersma *et al.* 1996), thereby changing the mechanics of the gait. Older horses with longer training histories could have greater muscle mass and therefore be more conditioned to perform one gait type over another. However, it is important to note that age did not differ significantly between the two gait types in our study.

Based on the previously mentioned confounding environmental aspects to gait, a more quantitative and objective phenotyping method could improve association. We grouped horses

based on the presence/absence of diagonal pair footfall landings, which likely was not a specific enough phenotyping method for comprehensive mapping a complex trait such as gait. Additionally, genotyping a larger set of horses and/or markers would enable detection of multiple loci with small effects contributing to our phenotype. Improving phenotype ascertainment could overcome difficulties encountered in this study. The use of accelerometers has proven useful in gait analysis (Pfau, Witte, Wilson 2005; Pfau, Witte, Wilson 2006; Robilliard, Pfau, Wilso 2007; Witte, Knill, Wilson 2004) and in distinguishing lateral from diagonal four-beat gaits (Östlund 2011). The accelerometers can be used to determine stride kinematics, speed, and potentially head motion (head motion is associated with the running walk and fox trot). Equipment consistently measuring the speed of the gait would allow for a more uniform comparison between horses moving at the same speed. In addition, the maximum speed before a horse breaks gait into the canter and kinematics of the canter could be measured and tested for association to the *DMRT3* mutation.

Footing has a huge impact on locomotion (Barrey, Landjerit, Wolter 1991; Clayton 2004; Kai *et al.* 1999). Documenting footing conditions, or providing a uniform surface on which to observe horses, would improve accuracy of phenotyping. Intermediate gaits are a continuum, and horses can change gaits over their lifetime and environmental conditions, longer videos and accelerometer collection can catch a larger window of the gaits the horse can perform and improve the accuracy of phenotyping.

Due to the continuous nature of the intermediate gaits, pleiotropic genes are likely candidates. There are several examples of pleiotropic genes in humans, mice, and horses. For example, mutations in the same gene, *ARX*, cause two highly diverse diseases: ambiguous genitalia and lissencephaly, a disease of the brain (Gustafson *et al.* 2014), illustrating the disparate roles a single gene can play. In horses, several genes that code for popular coat colors are also associated with a variety of diseases that affect vision (Andersson *et al.* 2008; Bellone *et al.* 2013), tumor growth

(Rosengren Pielberg *et al.* 2008), intestinal tract development (Metallinos, Bowling, Rine 1998; Santschi *et al.* 1998; Yang *et al.* 1998), and deafness (Baldwin *et al.* 1992; Hauswirth *et al.* 2012). Several of the genes we identified were related to immunology; while there are a disproportionally large number of immune related genes, there is the potential that we could have either inadvertently run an association on disease status or we have identified pleiotropic genes that play important roles in gait type and complex disease susceptibility. By collecting more detailed information on disease status and health management, future studies can genotype horses that have not been diagnosed with these complex diseases. Gaited breeds possess a predisposition to some diseases and may be a valuable resource for their study. For instance, equine metabolic syndrome (EMS) is reported to be very common among TWH, Paso Finos and Saddlebreds (Frank 2011) and several of our horses were diagnosed with EMS and other metabolic diseases, such as cushings, laminitis, and founder.

Despite insufficient power to definitively identify association with gait type within the TWH, our studies did highlight some suggestive candidate regions containing genes that may prove to play a role in modulation of this trait. Several of these regions include genes involved in developmental process, responses to stimuli, and biological regulation (Figure 2.3c). It may prove that gaited horses have greater capacity for motor neuron signaling, longer axon lengths, or different neuron growth patterns.

Candidate Genes

For example, we identified two genes within our across-breed ECA11 locus, *PIRT* and *SHISA6*, which may play important roles in sensory integration and head development. The *PIRT* (phosphoinositide-interacting protein) gene is a novel gene encoding a membrane protein specifically expressed in sensory neurons (Kim *et al.* 2008; Patel *et al.* 2011) known to play a vital part in sensing pain through modulation of the transient receptor potential vanilloid 1 (TRPV1) channel. *PIRT* has been implicated in playing a role in itch sensation (Patel *et al.* 2011) and heat pain (Tang *et*

al. 2013). *PIRT* -/- mice exhibit decreased behavioral responses to cold and cool temperatures (Tang *et al.* 2013), whereas horses seem to be more active in cool temperatures (Janczarek *et al.* 2014), which could be highlighting differences in sensory integration. The *SHISA6* (protein shisa-6 homolog) gene has been identified as a candidate locus for myopia (nearsightedness) in human populations (Oishi *et al.* 2013). The SHISA protein is a transcription factor that is important in early embryonic head development (Yamamoto *et al.* 2005) that physically interacts with immature forms of the Wnt receptor Frizzled and the FGF receptor within the endoplasmic reticulum to inhibit their post-translational maturation and trafficking to the cell surface (Hedge and Mason 2008). In chapter 3 we report head morphology differences in gaited breeds, and the *SHISA6* gene may play a functionally important role in the early head development of gaited breeds.

Our across-breed candidate locus on ECA24 was flanked by two genes: *CSK* and *DDX28,* which play important roles in cell signaling. The *CSK* gene is a member of the Src family of protein tyrosine kinases (SFK) that have been implicated in controlling an array of signaling networks regulating metabolism, viability, proliferation, differentiation and migration within many different cell lineages. The DEAD box proteins, characterized by the conserved motif Asp-Glu-Ala-Asp (DEAD), are putative RNA helicases (Valgardsdottir and Prydz 2003). They are implicated in a number of cellular processes involving alteration of RNA secondary structure, such as translation initiation, nuclear and mitochondrial splicing, and ribosome and spliceosome assembly (Valgardsdottir and Prydz 2003). Based on their distribution patterns, some members of the DEAD box protein family are believed to be involved in embryogenesis, spermatogenesis, and cellular growth and division (Valgardsdottir and Prydz 2003). *DDX28* is a nucleocytoplasmic shuttling protein that is also localized to the mitochondria (Valgardsdottir and Prydz 2003), and could play a role in signaling pathways.

Our TWH ECA19 candidate locus encompasses two different, but functionally interesting genes. The *FBXO40* gene is a member of the F-box protein family that are characterized as components of SCF (Skp1-Cullin1-F-box protein) E3 ubiquitin ligase complexes, in which they bind substrates for ubiquitin-mediated proteolysis (Cardozo and Pagano 2004; Deshaies 1999; Lin and Diehl 2004; Ye *et al.* 2007). The gene is expressed in skeletal muscle and functions as a regulator involved in postnatal myogenesis (Ye *et al.* 2007) and regulating insulin-like growth factor 1 (IGF1) (Shi et al, 2011). Muscles initiate and control movement through innervation by motor neurons, which originate in the spinal cord but receive projections from neurons originating from the brainstem and midbrain (Yuste *et al.* 2005). Differences in the muscles of gaited horses could be altering the signaling pathways of the neurons involved in the CPG pathways.

The *ARGFX* (arginine-fifty homeobox) gene is a member of the homeobox genes that encode DNA-binding proteins, many of which are thought to be involved in early embryonic development. Interestingly, an open reading frame for *ARGFX* has only been identified in the human genome, but not in mice, horse, cow, or guinea pig and cannot be distinguished as either a true functional gene or a nonfunctional pseudogene (Li and Holland 2010). The gene likely originated by from a gene duplication of *OTX1, OTX2* or *CRX* during early mammalian evolution (Li and Holland 2010). *OTX1* is a member of the homeodomain-containing transcription factors, and in the mouse, is required for proper brain and sensory organ development (Booth and Holland 2007). *OTX2* is also a member of the homeodomain-containing transcription factors and plays a role in brain, craniofacial, and sensory organ development (Booth and Holland 2007). *CRX* is a photoreceptor-specific transcription factor that plays a role in the differentiation of photoreceptor cells (Booth and Holland 2007; Li and Holland 2010).

The *SMTNL2* gene falls within our TWH ECA11 candidate locus and has been identified as a new substrate for mitogen-activated protein kinases (MAPK) which execute several critical cell

processes, including cell division, differentiation, and stress response (Gordon *et al.* 2013). *SMTNL2* is expressed in several mammalian tissues, but shows the highest expression in skeletal muscle and appears to be involved in myogenic differentiation (Gordon *et al.* 2013). As we described in Chapter 1, the rigidity and flexibility of the body frame can influence which gait is performed; a genetic difference in muscle composition can predispose a horse towards a body frame that is easier to maintain and therefore a preferred gait type.

Sequence Variation in Candidate Regions

We were unable to identify functionally important polymorphisms within our GWAS candidate genes that only occurred in our three gaited individuals. However, we did identify three genes bordering the candidate loci that contained functionally important polymorphisms. The 12 non-synonymous SNPs in *DTX3L* fell just outside our TWH ECA19 locus and is a member of deltx (DTX) family of ubiquitin ligases with known roles in Notch signaling (Holleman and Marchese 2014; Matsuno *et al.* 1995; Matsuno *et al.* 2002; Matsuno *et al.* 1998; Takeyama *et al.* 2003; Yamada *et al.* 2011). *DTX3L* has been implicated in regulating CXCR4 trafficking to lysosomes (Holleman and Marchese 2014). G protein-couple receptor (GPCR) sorting into the degradative pathway is important for limiting the duration and magnitude of signaling for hormones and neurotransmitters, and *CXCR4* plays an important role in the regulatory control of GPCRs (Holleman and Marchese 2014). This gene may play a role in how neurotransmitters are released in gaited horses, allowing them to switch between intermediate gaits. The presence of the 12 polymorphims in the region, spanning 1916 bp, is highly unusual as we would only expect to see this distribution of genotypes 1/277 times by chance, with only 18 observations over 10 million SNPs. To add further to the significance of this finding, all 12 of the SNPs are non-synonymous which are less common than intergenic and intron SNPs (Al Abri *et al.* ; Doan *et al.* 2012; Wade *et al.* 2009).

The stop-gained SNP on ECA7 identified from comparing the six horse genomes falls within a predicted ENSEMBLE exon for the horse and an intron for humans, mice, and zebra fish of the *ANKK1* gene. This gene is involved in signal transduction pathways and is in LD with the *DRD2* gene in humans. Both genes are part of the dopaminergic reward system and have been implicated with drug addiction and neuropsychiatric disorders.

Determination of polymorphisms in next generation sequencing datasets is not perfect, and many of these newly discovered changes will need to be validated, and their association to gait type measured in a larger sample. But the wealth of information that can be pulled from next-generation sequencing outweighs the potential downfalls, and can provide a way to detect rare variants and structural variants which is not possible with genome-wide association studies.

DMRT3 Is Fixed In a Breed of Horse Segregating for Different Intermediate Gaits

Following correction for population stratification *DMRT3* was the sixth ranked locus in our across-breed scan. Prior works have suggested that a stop mutation in DMRT3 is permissive for gaitedness and that there are other modifying loci determining the various gait types found in gaited breeds (Andersson *et al.* 2012; Petersen *et al.* 2013); however, there are gaited breeds such as Mangalarga Marchador where all of the individuals are gaited, but segregate for the DMRT3 mutation (~50% presence) (Patterson, Staiger, Brooks 2014; Promerová *et al.* 2014). Our sample TWH population is nearly fixed for the nonsense mutation in *DMRT3* despite the differences in gait type observed.

In the *DMRT3* mouse knockout, longer strides were observed as speed increased, but the mouse was unable to gallop (Andersson *et al.* 2012). Anecdotally, this is a common feature observed in the TWH. While speed was not measured in our gait analysis, the TWH can travel 10 to 20 miles per hour at an intermediate gait (Tennessee Walking Horse Breeders' and Exhibitors' Association 2011), rather than using the canter which has a typical speed of 10 to 17 miles per hour

(Clayton 1993; Clayton 1994). Harness racing horses reach speeds around 25-30 miles per hour (Barrey 2013) at both the trot and pace, and yet these horses were also observed to have the *DMRT3* mutation at high frequency (Andersson *et al.* 2012). Even though previous studies have shown an association between the mutation and the ability to perform alternate gaits, the frequency of *DMRT3* mutation was detected in a wide distribution of breeds, gaited and non-gaited (Promerová *et al.* 2014), indicating DMRT3 is more likely associated with speed. We hypothesize that the *DMRT3* mutation is not permissive for "gaitedness", but plays an inhibitory role in switching from intermediate gaits into the canter. Since many of the gaited breeds are able to perform their alternate four-beat gaits at higher speeds equivalent to a canter, it is likely that in selecting for lateral gaits, there was simultaneous selection against use of the canter. Additionally, the presence of this mutation in other, non-gaited, breeds indicates there must be another region controlling the ability to perform a four-beat gait.

TWH Population Structure

LD length was shortest in the Saddlebred, TWH, and Arabian with r^2 values dropping below 0.5 within the first 10-20 kb. LD was the longest in the Thoroughbred and Eqyptian Arabians, which did not drop below 0.5, and reflects the breeds' high inbreeding, low diversity and closure of the studbooks to outside genetic influence for several hundred years. The Saddlebred breed registry was closed in 1917 and contributed to the TWH breed registry, which itself was closed in 1947. The longer extent of LD in the TWH compared to the Saddlebred and Arabian may reflect a higher level of inbreeding, or a smaller founding population in the TWH than the latter two breeds. Higher individual inbreeding levels are observed in the TWH compared to the Arabian and Saddlebred (f mean=0.148, 0.060, and 0.103, respectively) by McCue, *et al.* (2012). This can be explained by frequent use of line-breeding in the horse, defined as a form of inbreeding where individuals are bred together to maintain a substantial degree of relationship to a highly regarded

ancestor (Bourdon 2000), which has occurred in the TWH since its foundation (Womack 1994). Indeed every TWH horse in existence today in the TWH can trace back to one stallion born in 1940, Midnight Sun. The individuals in this study with the highest inbreeding values traced their ancestry back to Midnight Sun multiple times in the pedigree.

Typically the modern TWH is classified in to three major stallion lines descending from Midnight Sun. Two of the lines are grandsons from the same son of Midnight Sun, Pride of Midnight, who was bred to daughters of his half-brothers (Spirit of Midnight and Sun's Delight D), The third is a great-grandson who is blended with another popular sire during Midnight Sun's lifespan, Merry Go Boy. These lines are likely what have been detected in the comparison of the dimensions 1 and 3 in the MDS plots, based on the examination of the pedigree tree for the study horses.

Conclusions

Due to the complex nature of locomotion, with its continuous phenotype and multiple environmental factors, genetic mapping of gait genes is a difficult task. At a rudimentary level of phenotyping we demonstrate with the current genotyping technology that there are detectable genetic differences both within and across breed. However, with more precise phenotyping and the use of a denser SNP array and whole genome sequencing, we will be able to map the location of the gait genes and the causal variants with improved precision.

With the identification of the gait genes, several advancements can be made in science and to the horse industry. Genetic tests can be developed that will aid breeders in their selection and breeding plans, improving the marketability of their horses. These tests will also decrease the amount of time and money spent on training, which will improve horse welfare; training schemes can be personalized to each individual horse's genetic potential and ineffective techniques can be avoided. Beyond the impact to horse owners, studying horse gait can aid in human disease

research and engineering of robotics circuitry. Animal models have long played an important role in human disease research due to similarities in basic biology, physiology and experimental convenience. By understanding the different gait types in the horse, new insights can be provided on motor coordination as no other mammalian species is known to discretely segregate for lateral stride cadence and footfall pattern.

References

Al Abri M, Kalla SE, Sutter NB, Brooks SA. Genomic polymorphism in six diverse horse breeds. Manuscript in Progress. .

Albertsdóttir E, Eriksson S, Näsholm A, Strandberg E, Árnason T. 2008. Genetic correlations between competition traits and traits scored at breeding field-tests in icelandic horses. Livestock Science 114(2-3):181-7.

Andersson LS, Juras R, Ramsey DT, Eason-Butler J, Ewart S, Cothran G, Lindgren G. 2008. Equine multiple congenital ocular anomalies maps to a 4.9 megabase interval on horse chromosome 6. BMC Genet 9:88.

Andersson LS, Larhammar M, Memic F, Wootz H, Schwochow D, Rubin CJ, Patra K, Arnason T, Wellbring L, Hjalm G, *et al.* 2012. Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. Nature 488(7413):642-6.

American Saddlebred Horse Association Saddlebred Record [Internet] [cited 2011 7/23/2011]. Available from: http://www.asha.net/Prize-Saddlebred-Record .

Baldwin CT, Hoth CF, Amos JA, da-Silva EO, Milunsky A. 1992. An exonic mutation in the *HUP2* paired domain gene causes waardenburg's syndrome. Nature 355:637-8.

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-5.

Barrey E. 2013. Gaits and interlimb coordination. In: Equine locomotion. Back W and Clayton HM, editors. 2nd. ed. New York: Elsevier. 85 p.

Barrey E, Landjerit B, Wolter R. 1991. Shock and vibration during the hoof impact on different track surfaces. In: Equine exercise physiology 3. Persson SGB, Lindholm A,Jeffcott LB, editors. Davis, California: ICEEP Publications. 97 p.

Bellone RR, Holl H, Setaluri V, Devi S, Maddodi N, Archer S, Sandmeyer L, Ludwig A, Foerster D, Pruvost M, *et al.* 2013. Evidence for a retroviral insertion in TRPM1 as the cause of congenital stationary night blindness and leopard complex spotting in the horse. PLos One 8(10):e78280.

Biewener AA. 1998. Muscle-tendon stresses and elastic energy storage during locomotion in the horse. Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology 120(1):73-87.

Booth HAF and Holland PWH. 2007. Annotation, nomenclature and evolution of four novel homeobox genes expressed in the human germ line. Gene 387(1-2):7-14.

Bourdon RM. 2000. Mating strategies based on pedigree relationship: Inbreeding and outbreeding. In: Understanding animal breeding. 2nd ed. Upper Saddle River, NJ: Prentice Hall. 333 p.

Cardozo T and Pagano M. 2004. The SCF ubiquitin ligase: Insights into a molecular machine. Nature Reviews Molecular Cell Biology 5(9):739-51.

Clayton HM. 2004. The dynamic horse: A biomechanical guide to equine movement and performance. First edition ed. Mason, MI: Sport Horse Publications.

Clayton HM. 1994. Comparison of the collected, working, medium and extended canters. Equine Veterinary Journal Supplemental 17:16-9.

Clayton HM. 1993. **The extended canter: A comparison of some kinematic variables in horses trained for dressage and for racing.** . Acta Anatomica 146(2-3):183-7.

Cook D, Gallagher PC, Bailey E. 2010. Genetics of swayback in american saddlebred horses. Animal Genetics 41:64-71.

Coonan JR, Greferath U, Messenger J, Hartley L, Murphy M, Boyd AW, Dottori M, Galea MP, Bartlett PF. 2001. Development and reorganization of corticospinal projections in EphA4 deficient mice. J Comp Neurol 436(2):248-62.

Crone SA, Quinlan KA, Zagoraiou L, Droho S, Restrepo CE, Lundfald L, Endo T, Setlak J, Jessell TM, Kiehn O, *et al.* 2008. Genetic ablation of V2a ipsilateral interneurons disrupts left-right locomotor coordination in mammalian spinal cord. Neuron 60(1):70-83.

Deshaies RJ. 1999. SCF and cullin/ring H2-based ubiquitin ligases. Annu Rev Cell Dev Biol 15:435-67.

Doan R, Cohen ND, Sawyer J, Ghaffari N, Johnson CD, Dindot SV. 2012. Whole-genome sequencing and variant analysis of a quarter horse mare. BMC Genomics 13:78.

Dottori M, Hartley L, Galea M, Paxinos G, Polizzotto M, Kilpatrick T, Bartlett PF, Murphy M, Kontgen F, Boyd AW. 1998. EphA4 (Sek1) receptor tyrosine kinase is required for the development of the corticospinal tract. Proceedings of the National Academy of Sciences 95(22):13248-53.

Fletcher JL. 1946. A study of the first fifty years of tennessee walking horse breeding. The Journal of Heredity 37(12):369-73.

Frank N. 2011. Equine metabolic syndrome. Veterinary Clinics of North America: Equine Practice 27(1):73-92.

Garbe JR and Da Y. 2008. Pedigraph: A software tool for the graphing and analysis of large complex pedigree. User Manual Version 2.4. Department of Animal Science, University of Minnesota .

Gordon EA, Whisenant TC, Zeller M, Kaake RM, Gordon WM, Krotee P, Patel V, Huang L, Baldi P, Bardwell L. 2013. Combining docking site and phosphosite predictions to find new substrates: Identification of smoothelin-like-2 (SMTNL2) as a c-jun N-terminal kinase (JNK) substrate. Cellular Signalling 25(12):2518-29.

Grillner S. 2003. The motor infrastructure: From ion channels to neuronal networks. Nat Rev Neurosci 4(7):573-86.

Grillner S and Zangger P. 1979. On the central generation of locomotion in the low spinal cat. Exp Brain Res 34(2):241-61.

Gustafson M, Edstrom M, Gawel D, Nestor CE, Wang H, Zhang H, Barrenas F, Tojo J, Kockum I, Olsson T, *et al.* 2014. Integrate genomic and prospective clinical studies show the importance of modular pleiotropy for disease susceptibility, diagnosis and treatment. Genome Med. 6(2):17.

Harris SE. 1993. Horse gaits, balance, and movement. New York, NY: Howell Book House.

Hauswirth R, Haase B, Blatter M, Brooks SA, Burger D, Drogemuller C, Gerber V, Henke D, Janda J, Jude R, *et al.* 2012. Mutations in MITF and PAX3 cause "splashed white" and other white spotting phenotypes in horses. PLoS Genet 8(4):e1002653.

Hedge TA and Mason I. 2008. Expression of Shisa2, a modulator of both wnt and fgf signaling, in the chick embryo. Int J Dev Biol 52(1):81-5.

Hendricks B. 1995. International encyclopedia of horse breeds. First edition ed. Norman, OK: University of Oklahoma Press.

Hinckley CA, Hartley R, Wu L, Todd A, Ziskind-Conhaim L. 2005. Locomotor-like rhythms in a genetically distinct cluster of interneurons in the mammalian spinal cord. J Neurophysiol 93(3):1439-49.

Holleman J and Marchese A. 2014. The ubiquitin ligase deltex-3l regulates endosomal sorting of the G protein-coupled receptor CXCR4. Mol Biol Cell 25(12):1892-904.

Janczarek I, Wilk I, Zalewska E, Bocian K. 2014. Correlations between the behavior of recreational horses, the physiological parameters and summer atmospheric condidtions. Animal Science Journal :Epub ahead of print.

Kai M, Takahashi T, Aoki O, Oki H. 1999. Influence of rough track surfaces on components of vertical forces in cantering thoroughbred horses. Equine Veterinary Journal Supplemental July(30):214-7.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42(4):348-54.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. Genome Res. 12(6):996-1006.

Kiehn O. 2006. Locomotoer circuits in the mammalian spinal cord. Annu Rev Neruosci 29:279-306.

Kim AY, Tang Z, Liu Q, Patel KN, Maag D, Geng Y, Dong X. 2008. Pirt, a phosphoinositide-binding protein, functions as a regulatory subunit of TRPV1. Cell 133(3):475-85.

Kullander K, Croll SD, Zimmer M, Pan L, McClain J, Hughes V, Zabski S, DeChiara TM, Klein R, Yancopoulos GD, *et al.* 2001a. Ephrin-B3 is the midline barrier that prevents corticospinal tract axons from recrossing, allowing for unilateral motor control. Genes Dev 15(7):877-88.

Kullander K, Mather NK, Diella F, Dottori M, Boyd AW, Klein R. 2001b. Kinase-dependent and kinase-independent functions of EphA4 receptors in major axon tract formation in vivo. Neuron 29(1):73-84.

Kuo AD. 2002. The relatie roles of feedforward and feedback in the control of rhythmic movements. Motor Control 6:129-45.

Lawson SEM, Chateau H, Pourcelot P, Denoix J, Crevier-Denoix N. 2007. Effect of toe and heel elevation on calculated tendon strains in thehorse and the influence of the proximal interphalangeal joint. Journal of Anatomy 210(5):583-91.

Li G and Holland PW. 2010. The origin and evolution of ARGFX homeobox loci in mammalian radiation. BMC Evolutionary Biology 10(1):182.

Lin DI and Diehl JA. 2004. Mechanism of cell-cycle control: Ligating the ligase. Trends Biochem Sci 29(9):453-5.

Makvandi-Nejad S, Hoffman GE, Allen JJ, Chu E, Gu E, Chandler AM, Loredo AI, Bellone RR, Mezey JG, Brooks SA, *et al.* 2012. Four loci explain 83% of size variation in the horse. PLos One 7(7):e39929.

Matsuno K, Diederich RJ, Go MJ, Blaumueller CM, Artavanis-Tsakonas S. 1995. Deltex acts as a positive regulator of notch signaling through interactions with the notch ankyrin repeats. Development 121:2633-44.

Matsuno K, Ito M, Hori K, Miyashita F, Suzuki S, Kishi N, Artavanis-Tsakonas S, Okano H. 2002. Involvement of a proline-rich motif and RING-H2 finger of deltex in the regulation of notch signaling. Development 129(4):1049-59.

Matsuno K, Eastman D, Mitsiades T, Quinn AM, Carcanciu ML, Ordentlich P, Kadesch T, Artavanis-Tsakonas S. 1998. Human deltex is a conserved regulator of notch signalling. Nat Genet 19(1):74-8.

McCue ME, Bannasch DL, Petersen JL, Gurr J, Bailey E, Binns MM, Distl O, Guerin G, Hasegawa T, Hill EW, *et al.* 2012. A high density SNP array for the domestic horse and extant perissodactyla: Utility for association mapping, genetic diversity, and phylogeny studies. PLoS Genet 8(1):e1002451.

Metallinos DL, Bowling AT, Rine JT. 1998. A missense mutation in the endothelin-B receptor gene is associated with lethal white foal syndrome: An equine version of hirschsprung disease. Mammalian Genome 9:426-31.

Mi H, Muruganujan A, Thomas PD. 2013. PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res 41(Database issue):D377-86.

Oishi M, Yamashiro K, Miyake M, Akagi-Kurashige Y, Kumagai K, Nakata I, Nakanishi H, Yoshikawa M, Oishi A, Gotoh N, *et al.* 2013. Association between ZIC2, RASGRF1, and SHISA6 genes and high myopia in japanese subjects. Invest Ophthalmol Vis Sci 54(12):7492-7.

Östlund V. 2011. Limb phasing icelandic horses. Swedish University of Agricultural Sciences.

Patel KN, Liu Q, Meeker S, Undem BJ, Dong X. 2011. Pirt, a TRPV1 modulator, is required for histamine-dependent and -independent itch. PLoS One 6(5):e20559.

Patterson L, Staiger EA, Brooks SA. 2014. DMRT3 is associated with gait type in mangalarga marchador horses, but does not control gait ability. Animal Genetics In Review.

Petersen JL, Mickelson JR, Cothran EG, Andersson LS, Axelsson J, Bailey E, Bannasch DL, Binns MM, Borges AS, Brama P, *et al.* 2013. Genetic diversity in the modern horse illustrated from genome-wide SNP data. PLos One 8(1):e54997.

Pfau T, Witte TH, Wilson AM. 2006. Centre of mass movement and mechanical energy fluctuation during gallop locomotion in the thoroughbred racehorse. J Exp Biol 209(19):3742-57.

Pfau T, Witte TH, Wilson AM. 2005. A method for deriving displacement data during cyclical movement using an inertial sensor. J Exp Biol 208(13):2503-14.

Promerová M, Andersson LS, Juras R, Penedo MCT, Reissmann M, Tozaki T, Bellone R, Dunner S, Horín P, Imsland F, *et al.* 2014. Worldwide frequency distribution of the 'Gait keeper' mutation in theDMRT3gene. Anim Genet 45(2):274-82.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, *et al.* 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 81:559-75.

Riemersma DJ, van der Bogert AJ, Jansen MO, Schamhardt HC. 1996. Influence of shoeing on ground reaction forces and tendon strain in the forelimbs of ponies. Equine Veterinary Journal 28(2):126-32.

Riley DA and Van Dyke JM. 2012. The effects of active and passive stretching on muscle length. Physical Medicine and Rehabilitation Clinics of North America 23(1):51-7.

Rivero J-L. 2007. A scientific background for skeletal muscle conditioning in equine practice. Journal of Veterinary Medicine Series A 54(6):321-32.

Robilliard JJ, Pfau T, Wilso AM. 2007. Gait characterisation and classification in horses. The Journal of Experimental Biology 210:187-97.

Rosengren Pielberg G, Golovko A, Sundstrom E, Curik I, Lennartsson J, Seltenhammer MH, Druml T, Binns M, Fitzsimmons C, Lindgren G, *et al.* 2008. A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. Nat Genet 40(8):1004-9.

Rozen S and Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics methods and protocols: Methods in molecular biology. Krawetz S and Misener S, editors. Totowa, NJ: Humana Press. 365 p.

Santschi EM, Purdy AK, Valberg SJ, Vrotsos PD, Kaese H, Mickelson JR. 1998. Endothelin receptor B polymorphism associated with lethal white foal syndrome in horses. Mammalian Genome 9(4):306-9.

Snow DH and Valberg SJ. 1994. Muscle anatomy, physiology and adaptations to exercise and training. In: The athletic horse: Principles and practice of equine sports medicine. Hodgson DR and Rose RJ, editors. Philadelphia, PA: WB Saudners Company. 145 p.

Takeyama K, Aguiar RC, Gu L, He C, Freeman GJ, Kutok JL, Aster JC, Shipp MA. 2003. The BAL-binding protein BBAP and related deltex family members exhibit ubiquitin-protein isopeptide ligase activity. J Biol Chem 278(24):21930-7.

Tang Z, Kim A, Masuch T, Park K, Weng H, Wetzel C, Dong X. 2013. Pirt functions as an endogenous regulator of TRPM8. Nat Commun 4:2179.

The Tennessee Walking Horse Breed: Gaits [Internet]; c2011 [cited 2014 . Available from: www.twhbea.com/breed/gait.php .

Valgardsdottir R and Prydz H. 2003. Transport signals and transcription-dependent nuclear localization of the putative DEAD-box helicase MDDX28. J Biol Chem 278(23):21146-54.

Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, Lear TL, Adelson DL, Bailey E, Bellone RR, *et al.* 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science 326(5954):865-7.

Wilson JM, Hartley R, Maxwell DJ, Todd AJ, Lieberam I, Kaltschmidt JA, Yoshida Y, Jessell TM, Brownstone RM. 2005. Conditional rhythmicity of ventral spinal interneurons defined by expression of the Hb9 homeodomain protein. J Neurosci 25(24):5710-9.

Witte TH, Knill K, Wilson AM. 2004. Determination of peak vertical ground reaction force from duty factor in the horse (equus caballus). J Exp Biol 207(Pt 21):3639-48.

Womack B. 1994. The echo of hoofbeats: A history of the tennessee walking horse. Third ed. Shelbyville, Tennessee: DABORA, INC.

Yamada K, Fuwa TJ, Ayukawa T, Tanaka T, Nakamura A, Wilkin MB, Baron M, Matsuno K. 2011. Roles of drosophila deltex in notch receptor endocytic trafficking and activation. Genes to Cells 16(3):261-72.

Yamamoto A, Nagano T, Takehara S, Hibi M, Aizawa S. 2005. Shisa promotes head formation through the inhibition of receptor protein maturation for the caudalizing factors, wnt and FGF. Cell 120(2):223-35.

Yang G, Croaker D, Zhang AL, Manglick P, Cartmill T, Cass D. 1998. A dinucleotide mutation in the endothelin-B receptor gene is associated with lethal white foal syndrome (LWFS); a horse variant of hirschsprung disease. Hum Mol Genet 7(6):1047-52.

Ye J, Zhang Y, Xu J, Zhang Q, Zhu D. 2007. FBXO40, a gene encoding a novel muscle-specific F-box protein, is upregulated in denervation-related muscle atrophy. Gene 404(1-2):53-60.

Yokoyama N, Romero MI, Cowan CA, Galvan P, Helmbacher F, Charnay P, Parada LF, Henkemeyer M. 2001. Forward signaling mediated by ephrin-B3 prevents contralateral corticospinal axons from recrossing the spinal cord midline. Neuron 29:85-97.

Yuste R, MacLean JN, Smith J, Lansner A. 2005. Opinion: The cortex as a central pattern generator. Nature Reviews Neuroscience 6(6):477-83.

Ziegler L. 2005. Easy-gaited horses. First edition ed. North Adams, MA: Storey Publishing.

CHAPTER 3

MORPHOLOGICAL VARIATION IN GAITED BREEDS OF HORSES

Introduction

Conformation has long been a driving force in horse selection and breed identification, particularly as a predictor for performance and injury susceptibility. In any equine discipline there are certain conformations and body dimensions that are reported as desired and advantageous to performance. For example, good front limb action is determined mostly by leg and foot stances, slope of the shoulders and pasterns, and length of the leg (Harris 1993; Nicodemus, Holt, Swartz 2002; Nicodemus and Clayton 2003). Limb conformation is the major factor in limb soundness and can be a predictor for future lameness (Splan and Hunter 2004; Stashak 1987). Locomotor problems and lameness are the most common reason for training problems and culling horses (Bergsten 1980; Jeffcott *et al.* 1982; Linder and Dingerkus 1993; McGreevy and Thomson 2006; Philipsson *et al.* 1998). The development of lameness and other locomotor problems can be due to genetic disorders affecting skeletal development, conformation, and repair.

Gaited horses, those able to perform a two-beat lateral or four-beat gait at speeds equivalent to a trot, have different body proportions than non-gaited horses (Ziegler 2005). Differences in body posture or "frame" relate to the amount of flexion and tension in the ligaments and muscles of the horse's body and the shape of the back (see Chapter 1); both a hollow and round frame have high tension, but there is more flexion in the round frame with the back in a convex shape while the back is concave in the hollow frame. In the neutral frame, there is some tension, but no flexion in the back in either direction. Horses that utilize a pace or stepping pace work in a hollow frame and possess slightly longer backs and loins, shorter necks, and steeper, shorter hips than those that work in more neutral frames (Ziegler 2005). Horses that are expected to work in the racking gaits often have a longer loin and shorter hip than those that work in more neutral frame gaits (the running walk or paso llano) and also work in a hollow frame, but less so than horses that pace (Ziegler 2005). The neutral frame gaits favor a moderate to long neck, long back, and relatively long hip (Ziegler

2005). Horses that work in neutral to round frames, as in the fox trot (diagonal four-beat gait), may have shorter backs and loins than those that prefer the running walk or paso llano, but they will also have long necks and relatively long hips (Ziegler 2005). Multi-gaited horses worked in a fully round frame will start to trot (Ziegler 2005).

The Tennessee Walking Horse (TWH) breed ranges in size from 14.3 to 17 hands (Figure 3.1) and is renowned for the ability to perform an even-timed four-beat gait at intermediate speeds. Yet, within the breed there remain variations in gait type (Tennessee Walking Horse Breeders' and Exhibitors' Association 2011) which may be a reflection of the biomechanical influence of diverse body conformations, as well as genetics.

The TWH is used for a number of activities, including exhibition in the show ring and for pleasure and trail riding. Different body conformations are desired for these different activities and likely gives rise to the variation observed in TWH body size. By identifying the different conformation types and determining which conformation types are more prevalent in each discipline, selection can be more objective and effective. Reducing the production of foals with sub-optimal conformation, and the injuries that result from poor conformations will thereby improve horse welfare. Improved accuracy of selection is especially important for the TWH, as many prominent sires are often from one particular discipline, yet are expected to produce offspring for multiple disciplines.

The aim of this study is to assess the impact of morphological traits and patterns on gait type and performance in horses. These morphological traits will then be used to identify genes associated with skeletal variation in the TWH.

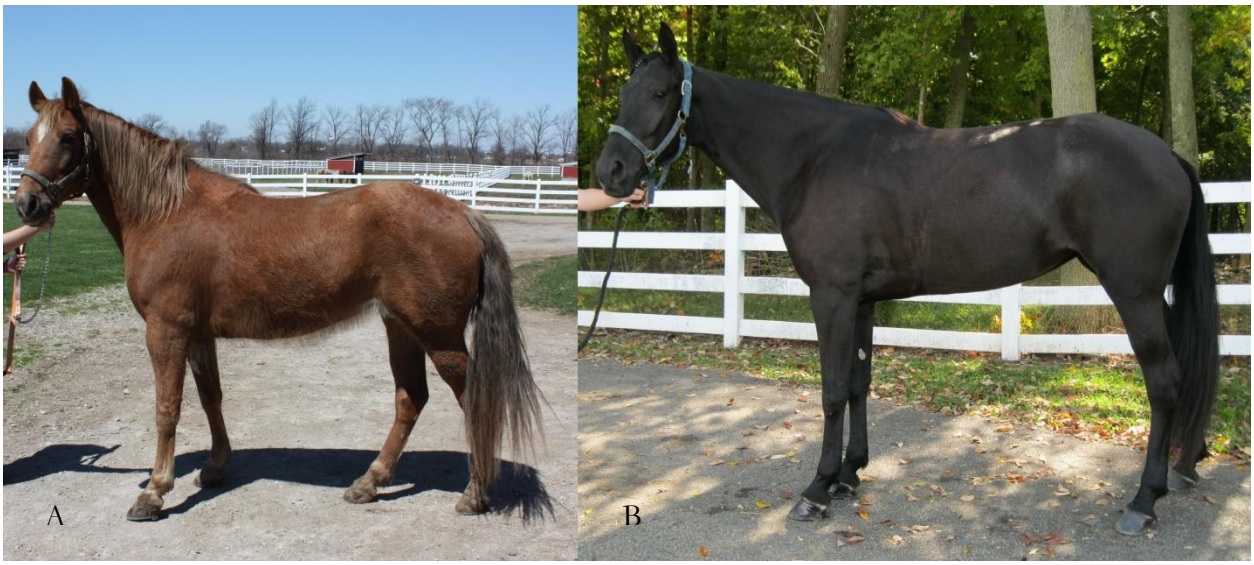


Figure 3.1. Two TWH that exemplify skeletal size variation; horse in pane "a" is 55.25 inches tall at the withers, while the horse in pane "b" is 66 inches tall.

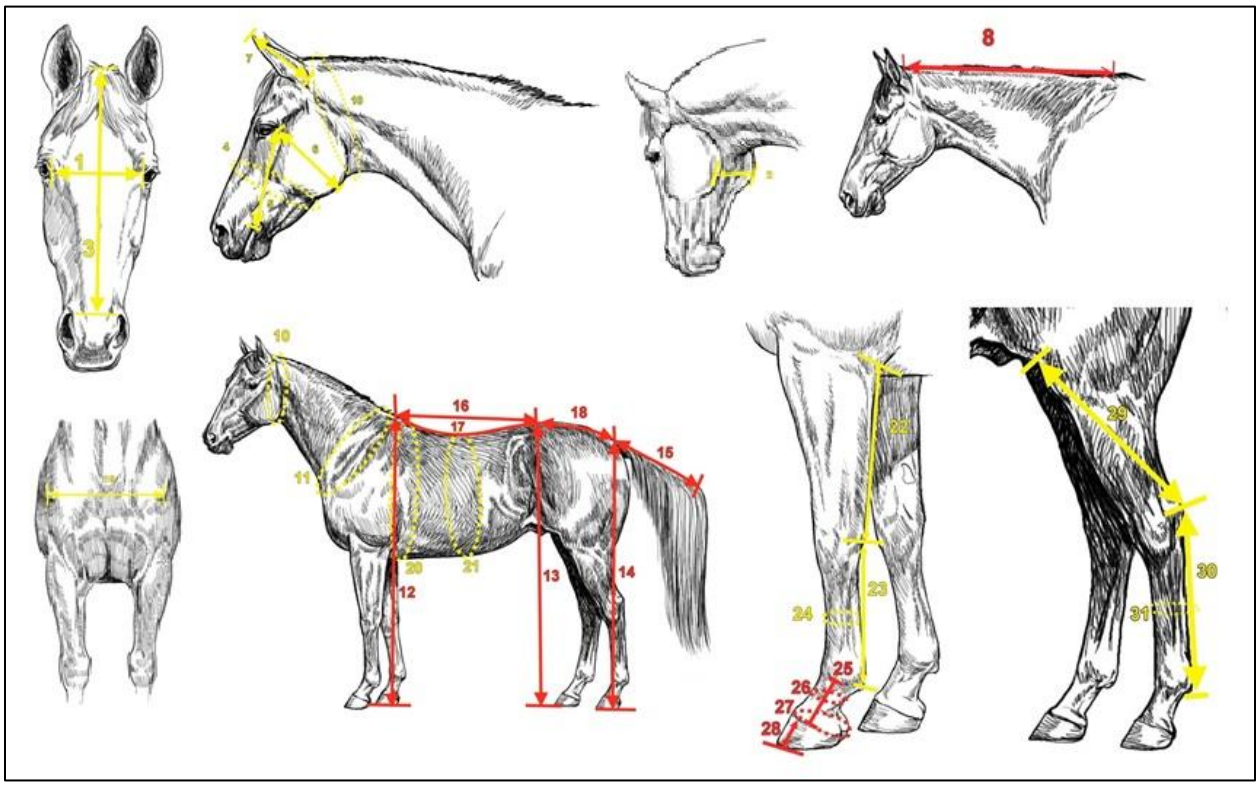


Figure 3.2. Example of body measurements collected from each horse. Solid lines are lengths, while dashed lines are circumferences. (Illustrations by Susan Harris.)

Materials & Methods

Sampled Horses – Skeletal Variation Across Breeds

A total of 2019 horses, ranging in age from 1 to 36 years, from 96 breeds were measured for a total of 35 body measurements following a previously developed protocol (Brooks, *et al.*, 2010). 1496 horses were previously published in Brooks *et al.* (2010), 523 were gathered to more closely investigate these variables in the gaited breeds. Measurements gathered for that study included, in part, 35 body measurements that span the head, neck, body, and limbs collected from horses, vital statistics (breed, age, sex, etc.) and owner reported gait information. The 35 body measurements use easily palpable landmarks and a measuring tape to capture skeletal length and circumference as quantitative variables (Figure 3.2). Horses were classified as "gaited" if they were reported by their owner as able to perform an intermediate gait other than trot, or belong to a breed known to be "gaited" (n=740 total).

In total, we examined 714 geldings, 1076 mares, 215 stallions, and 5 of unknown sex from both the previous and current study. Breeds included 79 registered breeds, 8 mixed breeds, 2 mule breeds, 4 donkey breeds, a "tarpan", *E. asinus* and 45 horses of unknown breed. 64 breeds fell under the "not gaited" classification, 23 breeds as "gaited", and 9 breeds were shared across the two categories.

Data Quality Control and Statistics – Comparison Across Breeds

Head to ground neck length, tail length and max barrel girth measurements were excluded due to inconsistent compliance among horses attempting to reach the ground, owners-error in measuring tail hair length or presence of a docked tail, and confounding impact of body condition score or pregnancy in mares, respectively. Horses missing a value for any of the remaining measurements (n=20) were excluded to reduce measurement overrepresentation weighting bias for the principle component analysis. Horses aged two years or less were excluded from analysis

(n=112) due to skeletal immaturity. 96 horses of unknown breed were excluded from analysis as their gait could not be inferred from their breed standards. Three horses were excluded for unknown sex. In breeds with at least three samples, extreme outlier measurements were identified by using quantile plots and calculating the median and interquartile range for each measurement. 20 horses were excluded that had at least one measurement value greater than ten times the interquartile range from the median value for that measurement in their breed. 65 horses were excluded with at least one measurement that was between five and ten times the interquartile range from the median for their breed; these outliers were presumed to be errors in either measuring or recording of the measurements. 20 horses were excluded that had two or more outlier measurements and that had at least one measurement that was between 4.5 and five times the interquartile range from the median for their breed; these outliers were also presumed to be errors in either measuring or recording of the measurements. 1683 horses met inclusion criteria; these included 598 gaited, 1085 non-gaited, 606 geldings, 910 mares, and 167 stallions.

A second dataset was generated from the filtered first dataset to compare balanced numbers of male and female gaited versus non-gaited horses. This set negated differences due to sexual dimorphism. 524 mares, 345 geldings that were non-gaited and 242 mares, 117 geldings, and 23 stallions that are gaited were randomly excluded from analysis to balance the number of gaited and non-gaited horses in the dataset. 432 horses met inclusion criteria; these included 216 gaited, 216 non-gaited, 144 geldings, 144 mares, and 144 stallions.

Principle Component Analysis (PCA) was conducted on the 32 body measurements for each dataset. A correlation matrix was used due to the wide variation in the scale of the measurements (median pastern length of 4.25 inches compared to median heart girth of 73.5 inches). Principle Component (PC) scores were retained for interpretation and analysis by examining when the scree plot curve plateaus, if the eigenvalue is greater than 1.0, and if percent variance explained is greater

than $1/n$ where n is the number of body measurements we collected ($1/n = 3.125\%$). Analysis of variance (ANOVA) and pairwise correlations were conducted to identify differences in the PC scores and measurements (normalized by wither height) due to gait type, sex, gender or age. PCA, ANOVA and correlations were conducted and visualized using JMP Pro 10.0.2 (SAS Institute, Inc., Cary, NC).

Sampled Horses – Skeletal Variation Within the TWH Breed

We pulled a subset of 282 TWH, ages 1-34, from the dataset mentioned above that met inclusion criteria. Vital statistics and basic history (date or year of birth, registry, barn name, gait information, brief notes on any injury or disease, a photo and pedigree) was also available for each horse. Pedigrees were inspected to avoid including horses related within one generation. Video of the horse performing an intermediate gait and the owner reported horse discipline were recorded on all of the included horses. Video recording and analysis followed the same protocol as described in Chapter 2. For 38 horses, the video was only of walking or had not been submitted by the owner, resulting in 244 horses with complete gait phenotypes.

Statistical Tests - Skeletal Variation Within the TWH Breed

Principle Component Analysis (PCA) was conducted on 32 body measurements for the 282 horses. The correlation matrix was used due to the wide variation in the scale of the measurements. Principle Component (PC) scores were retained for interpretation and analysis by examining the scree plot and identifying the component at which the curve plateaus, if the eigenvalue for a given component was greater than 1.0, or if percent variance explained was greater than $1/n$ where n is the number of body measurements we collected ($1/n = 3.125\%$). Based on these guidelines, we retained the first eight PCs. We used ANOVA and pairwise correlations to identify differences in the measurements (normalized by wither height) due to gait type, sex, gender, and age, and then applied an analysis of covariance (ANCOVA) to control for the effects of gender.

PCA, ANOVA, ANCOVA, and correlations were conducted and visualized using JMP Pro 11 (SAS Institute, Inc., Cary, NC).

Genotyping and Genome Wide Association Studies

Of the 282 TWH that passed filtering above, we selected 109 TWH with complete size PC and video-analyzed gait phenotypes for study so that no horse was related to any other within a single generation within each gait type to reduce population stratification in the subsequent GWAS. We isolated genomic DNA from blood samples using the Gentra® Puregene® Blood Kit, following the manufacturer's protocol for whole blood (Qiagen Inc., Valencia, CA). Using the same kit, we followed the manufacturer's protocol with modifications optimized for hair root bulbs to extract DNA from hair (Cook, Gallagher, Bailey 2010). GeneSeek Inc. (Lincoln, NE) genotyped the 109 TWH samples with complete size phenotypes at 65,000 loci using the Equine SNP70K beadchip (Illumina Inc., San Diego, CA).

To supplement coverage in an important candidate region with poor coverage, we designed restriction fragment length polymorphism (RFLP) tests for four additional SNPs from chromosome 9. This region was previously identified as one of four candidate loci contributing to skeletal size variation in horses (Makvandi-Nejad *et al.* 2012). The SNPs were identified from whole-genome sequencing of four individuals (TWH, Percheron, American Miniature, and Arabian) for another project (Al Abri *et al.*). Three of the SNPs were identified in the *ZFAT* gene, one SNP in exon 13 and two SNPs in intron 9. We identified the fourth SNP in exon 9 of the *KHDRBS3* gene. For all four SNPs, we performed PCR amplification in a 20μL volume containing 2μL of DNA (diluted to a concentration of 25ng/μL), 2μL of 10X PCR reaction buffer with 20mM $MgCl_2$, 0.2μL of FastStart Taq DNA Polymerase (Roche Diagnostics), 2μL of 2mM dNTP's, 2μL each of forward and reverse 5μM primers, and 9.8μL PCR-grade water. We used Primer3 (Rozen and Skaletsky 2000) to design primer sequences around each SNP to produce a single PCR product (Table 3.1) and carried out the

PCR on an Eppendorf Mastercycler Ep Gradient (Eppendorf Corp.) under the following conditions: 95°C for 4 min, followed by 40 cycles of 95°C for 30 sec, 63°C for 30 sec, 72°C for 30 sec, and a final extension of 72°C for 7 min and cooling to 4°C. For the *KHDRBS3* SNP, we carried out the PCR with an annealing temperature of 58°C instead of 63°C and with a 45 sec cycling extension time instead of 30 sec. The restriction digest used 10μL of the PCR product, 0.4 μL restriction enzyme (1.0U per reaction, New England Biolabs Inc. (NEB)), 1 μL of 10X NEB CutSmart buffer and 8.6 μL MilliQ water to bring the reaction volume to 20 μL, which was incubated according to each enzyme's specifications (Table 3.1). The resulting products were visualized by electrophoresis following standard conditions on a 2.5% agarose gel (Omnipur Agarose, EMD Chemicals Inc). A summary of the enzymes used, incubation temperatures and times, and fragment sizes of each allele for each SNP are provided in Table 3.1.

Genotyping Quality Control

We excluded SNPs with a genotyping rate <95% across all individuals (n=14068) and MAF< 0.05 (n=14494). After we filtered for genotyping rate <90% across all remaining SNPs, three individuals were removed. The total genotyping rate in the remaining 106 individuals was 99.1% across 39,262 SNPs. We used a Bonferroni significance cutoff of 1.38×10^{-6}, conservatively estimating 36,171 independent comparisons (3,091 markers that are in complete LD, $r^2>0.99$, and subtracted from our 39,262 total markers in the study). To evaluate population structure, we used an identity by state (IBS) similarity matrix to calculate genome-wide IBD estimates; we removed one individual from a pair with an IBS greater than 0.90.

We ran three genome-wide tests for association with the quantitative trait scores for PC 1, 2, and 3 unique to the TWH. In the software program Golden Helix SVS (reference), we applied basic quantitative association, linear associations with covariates of sex, gender, and/or a discipline category, and a Mixed Model linear analysis (EMMAX) (Kang, 2010) with and without the covariates

of sex, gender, and/or a discipline category, under additive, dominant, and recessive models. We utilized quantile-quantile plots and genomic inflation factors to determine the model of best fit for each PC. We also examined the linkage disequilibrium (LD) structure between markers on candidate regions using Haploview v4.2 (Barrett *et al.* 2005). For PC1, we applied the EMMAX dominant model with sex added as a covariate. For both PC2 and PC3 we found linear models ineffective and EMMAX models over effective in correcting for population structure. Therefore, we applied adaptive permutation in PLINK (Purcell *et al.* 2007) for PC2 and 3 after standard test failed to surpass our Bonferroni cutoff (EMMAX models).

Table 3.1. RFLP genotyping parameters.

SNP Name	Gene & Location	PCR product size	PCR Annealing temp.	PCR Extension time	Restriction Enzyme	Incubation temp. & time	Major Allele: digest sizes (bp)	Minor Allele: digest sizes (bp)
ZFAT e13	*ZFAT* exon 13	476 bp	63°C	30 sec	EciI	37°C 1 hour	A: 476	G: 345, 141
ZFAT i9a	*ZFAT* intron 9, 5' end	457 bp	63°C	30 sec	TaqI	65°C Overnight	C: 324, 144	T: 457
ZFAT i9b	*ZFAT* intron 9, 3' end	499 bp	63°C	30 sec	Tsp45I	65°C 1 hour	C: 398, 101	T: 398, 85, 16
KHDe9	*KHDRBS3* exon 9	590 bp	58°C	45 sec	HhaI	37°C Overnight	A: 590	G: 460, 130

Results

Skeletal Variation Across Breeds

From 1683 horses of diverse breeds, we collected and analyzed 32 body measurements using PCA to reduce the measurements down to components that explain the most size variance. Only PC1, PC2, and PC3 were retained based on eigenvalue, percentage explained, and scree plot (Figure 3.3a). PC1 explains 65.3% of the variance, PC2 explains 6.6% of the variance, and PC3 explains 3.2% of the variance (Figure 3.3b). All measurements loaded in a positive direction for PC1, indicating a positive correlation between all 32 measures, and therefore PC1 likely explains overall body size (Figure 3.3c). For PC2, length measurements loaded in a positive direction while circumference measurements loaded in a negative direction, with the exception of both hoof lengths, croup to dock length, and ear length which loaded negatively with circumference measurements. These deviations could be due to an undetectable relationship with the circumferences of their respective body parts (croup to dock length) or environmental impacts (hoof and ear length). For example, hoof length can be altered through routine (or not routine) trimming of the horny growth encapsulating the hoof. Based on the overall pattern of factor loading, PC2 likely represents body "thickness". PC3 explains an even smaller proportion of the variance than PC1 or PC2 (3.2% versus 65.3% and 6.6%, respectively), and represents a more understated pattern of body shape variation. The top contributors to PC3 include fore and hind pastern lengths and jaw width measures loading in a positive direction, with lengths in general loading in a negative direction and circumferences in a positive direction. PC3 likely represents skull thickness and lower limb length based on the pattern of factor loadings. A horse with a high positive PC3 score has a broad eye and jaw width, long thick pasterns, a short back and forearm, and short hooves; a horse with a negative PC3 score will demonstrate the opposite pattern.

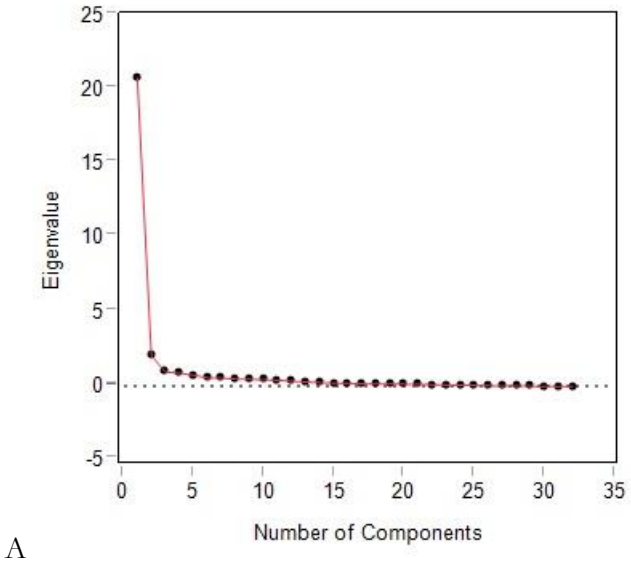
In addition to the body measurements, we collected gender (mare, stallion or gelding), age, and gait information for each horse. PC1 and PC2 scores were significantly higher in geldings and mares, respectively, indicating they are overall taller and thinner than stallions, while PC3 scores were not significantly different for sex or gender (Table 3.2). PC1 and PC3 scores were higher in horses that could only trot, while PC2 scores were higher in gaited horses (Table 3.2), indicating gaited horses are shorter, but thinner than horses that only trot. We found a negative correlation in PC2 and a positive correlation in PC3 with age, indicating that as horses age they generally become thicker. This is not unlike patterns associated with age in other mammals (Etherton 2009).

We also collected factor scores for several qualitative traits that are anecdotally correlated with body conformation, including head profile shape, incisor alignment (or bite), degree of long hair growth(feathering) on lower limbs, and a subjective measure of bone thickness. For example, draft breeds often have "Roman" or convex head profiles, copious amounts of feathering, and thick bones. In contrast, the Arabian, a light horse breed, often has a "Dish" or concave head profile, little to no feathering, and thin bones. We identified positive correlations in head profile, feathering and bone thickness scores with PC1 scores, while these were negatively correlated with PC2 scores. This matches with our breed stereotypes, *i.e.* draft breeds are often the tallest breeds, have more feathering, thicker bones, and a roman head profile. For PC3, we identified a negative correlation with head profile scores, but positive correlations with feathering and bone thickness scores. We did not detect any significant correlations of incisor alignment with any of the PC scores. Correlation coefficients and Pearson's p-values for all factors are presented in Table 3.2.

All three components were identified as significant with gait type, along with gender in the first two components. Therefore, we examined the individual measurements for association with gait type, sex and gender classification and detected sexual dimorphism across breeds. For the 32 body measurements normalized by wither height, 23 measurements were significantly different

between males and females (Table 3.3). These included shorter ear length, shorter back length, smaller heart girth, thicker cannons, and longer front hooves and shorter croup and dock heights in males. Males were taller at non-normalized wither height. If geldings were separated from stallions into their own gender category, 22 measurements were significantly different (Table 3.3). For wither height, geldings were taller than mares and mares taller than stallions. Stallions had thicker throat latches and muzzles, shorter dock heights, shorter backs, smaller heart girths, longer hooves, a longer gaskin and longer cannons.

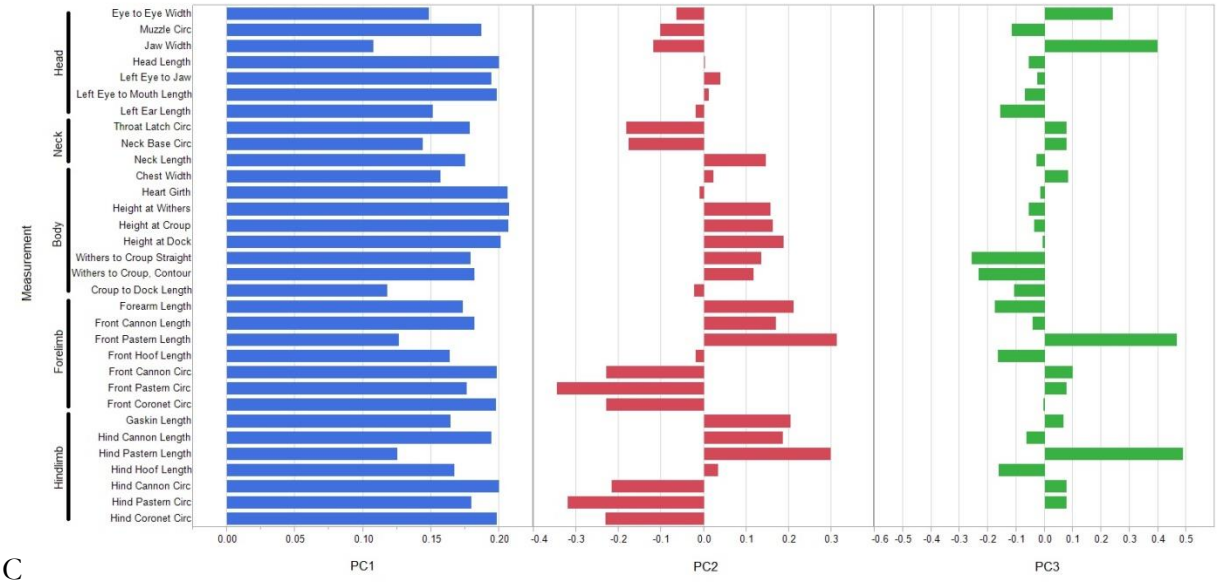
Due to the presence of sexual dimorphism and unequal sampling across genders in our sample set, we examined correlations between conformation and gait ability within each gender class. Thirteen body measurements were significantly different in gaited horses for all three gender classes, plus 12 additional measurements that were gender specific (Table 3.3). These included gaited geldings with shorter wither heights, and gaited mares with shorter heads, longer eye to mouth lengths, thinner throatlatches, shorter dock heights, and longer hind hooves. Across all genders, gaited horses had narrower eye and jaw widths, thinner cannons and pasterns, and longer forearms and cannons.



A

Number	Percent	Cum Percent
1	66.780	66.780
2	6.646	73.425
3	3.306	76.731
4	2.503	79.235
5	2.326	81.560
6	1.923	83.484
7	1.853	85.337
8	1.616	86.953
9	1.501	88.454
10	1.384	89.838
11	1.247	91.085
12	1.010	92.095
13	0.900	92.995
14	0.890	93.886
15	0.798	94.683
16	0.644	95.327
17	0.595	95.922
18	0.556	96.478
19	0.527	97.005
20	0.483	97.487
21	0.466	97.953
22	0.397	98.350
23	0.371	98.721
24	0.276	98.997
25	0.247	99.244
26	0.220	99.464
27	0.185	99.649
28	0.106	99.754
29	0.100	99.854
30	0.069	99.923
31	0.047	99.970
32	0.030	100.000

B



C

Figure 3.3. Principle component analysis of the 32 across-breed body measurements. The a)scree plot and b)percent explained were used to determine the number of components retained for further analysis. The bars represent the individual components percentage contribution, while the line represents the cumulative percentage. C) factor loading of body measurements onto principle components.

Table 3.2. Associations between Across breed PC scores and other traits of interest

Trait	PC1	PC2	PC3
Sex†	P<0.0001*	P=0.0354*	P=0.1590
Gender†	P<0.0001*	P=0.0052*	P=0.3077
Gait ability†	P=0.0009*	P<0.0001*	P<0.0001*
Age‡	0.0304	-0.0742	0.0651
	P=0.2124	P=0.0023*	P=0.0075*
Dish‡	0.3374	-0.1218	-0.0724
	P<0.0001*	P<0.0001*	P=0.0030*
Feathering‡	0.1388	-0.6213	0.1168
	P<0.0001*	P<0.0001*	P<0.0001*
Bone‡	0.3998	-0.5325	0.1647
	P<0.0001*	P<0.0001*	P<0.0001*
Bite‡	0.0053	-0.0395	0.0004
	P=0.8282	P=0.1058	P=0.9862

† designates traits analyzed by ANOVA. ‡ designates traits analyzed by pairwise correlation with Pearson's p-values. * p-value significant at $\alpha<0.05$.

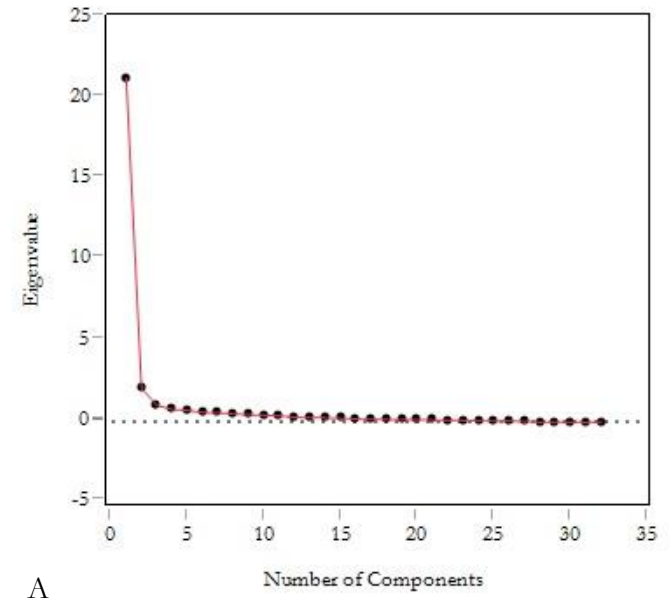
Table 3.3. ANOVA p-values for the full across breed dataset of body measurements normalized by wither height versus traits of interest.

Measure	Sex	Gender	Gaited?	Gait Mares	Gait Geldings	Gait Stallions
Wither Height#	0.0005**	<0.0001**	0.0418*	0.6752	0.0011**	0.0813
Eye Width	0.2607	0.3399	<0.0001**	<0.0001**	<0.0001**	<0.0001**
Jaw Width	0.0382*	0.0991	<0.0001**	<0.0001**	<0.0001**	<0.0001**
Head Length	0.0441*	0.0431*	<0.0001**	<0.0001**	0.0136*	0.7531
Muzzle Circumference	0.2785	0.0011*	0.1989	0.1106	0.3917	0.0038*
Eye to Mouth Length	0.0052*	0.0198*	0.0002**	0.0002**	0.0077*	0.1902
Eye to Jaw Length	0.9489	0.1581	0.3163	0.1036	0.9831	0.1856
Ear Length	<0.0001**	<0.0001**	0.3486	0.1235	0.8975	0.8704
Neck Length	0.0144*	0.0154*	0.3112	0.8560	0.6875	0.0516
Throatlatch Circumference	0.0016*	<0.0001**	0.0340*	0.0006**	0.6273	0.0529
Neck Base Circumference	0.0035*	<0.0001**	<0.0001**	<0.0001**	0.0001**	0.0097*
Croup Height	0.0006**	0.0011**	0.0343*	0.0034*	0.7211	0.5646
Dock Height	<0.0001**	<0.0001**	<0.0001**	<0.0001**	0.0071*	0.6536
Withers to Croup, Straight	<0.0001**	<0.0001**	0.0328*	0.0039*	0.0570	0.1265
Withers to Croup, Contoured	<0.0001**	<0.0001**	0.6717	0.5654	0.4358	0.0176*
Croup to Dock Length	0.0282*	0.0472*	0.0048*	0.0307*	0.5442	0.0437*
Chest Width	0.8817	0.7068	0.1600	0.2209	0.2194	0.7991
Heart Girth	<0.0001**	<0.0001**	<0.0001**	<0.0001**	0.0044*	0.0003**
Forearm Length	0.9955	0.4194	<0.0001**	<0.0001**	<0.0001**	0.0074*
Front Cannon Length	0.0004**	0.0020*	<0.0001**	<0.0001**	0.0148*	<0.0001**
Front Cannon Circumference	<0.0001**	<0.0001**	<0.0001**	<0.0001**	<0.0001**	<0.0001**
Front Pastern Length	0.5181	0.5878	0.0278*	0.1473	0.3072	0.1769
Front Pastern Circumference	0.0244*	0.0412*	<0.0001**	<0.0001**	0.0002**	<0.0001**
Front Coronet Circumference	0.0426*	0.0707	<0.0001**	<0.0001**	0.0175*	<0.0001**
Front Hoof Length	<0.0001**	<0.0001**	<0.0001**	<0.0001**	0.0053*	0.1378
Gaskin Length	0.3403	0.0008**	0.0501	0.2277	0.4483	0.8341
Hind Cannon Length	<0.0001**	<0.0001**	<0.0001**	<0.0001**	0.0231*	0.0006**
Hind Cannon Circumference	0.0213*	0.0550	<0.0001**	<0.0001**	<0.0001**	<0.0001**
Hind Pastern Length	0.1337	0.3015	0.8389	0.8228	0.5294	0.3546
Hind Pastern Circumference	0.1653	0.1019	<0.0001**	<0.0001**	<0.0001**	<0.0001**
Hind Coronet Circumference	0.0081*	0.0268*	<0.0001**	<0.0001**	0.0037*	<0.0001**
Hind Hoof Length	0.0434*	0.0243*	0.0008**	<0.0001**	0.9047	0.7570

#Not normalized. *P-value significant at $\alpha<0.05$. **P-value significant at Bonferroni cutoff $\alpha<0.0015$.

Gaited Breeds Have Longer Individual Limb Lengths than Non-Gaited Breeds

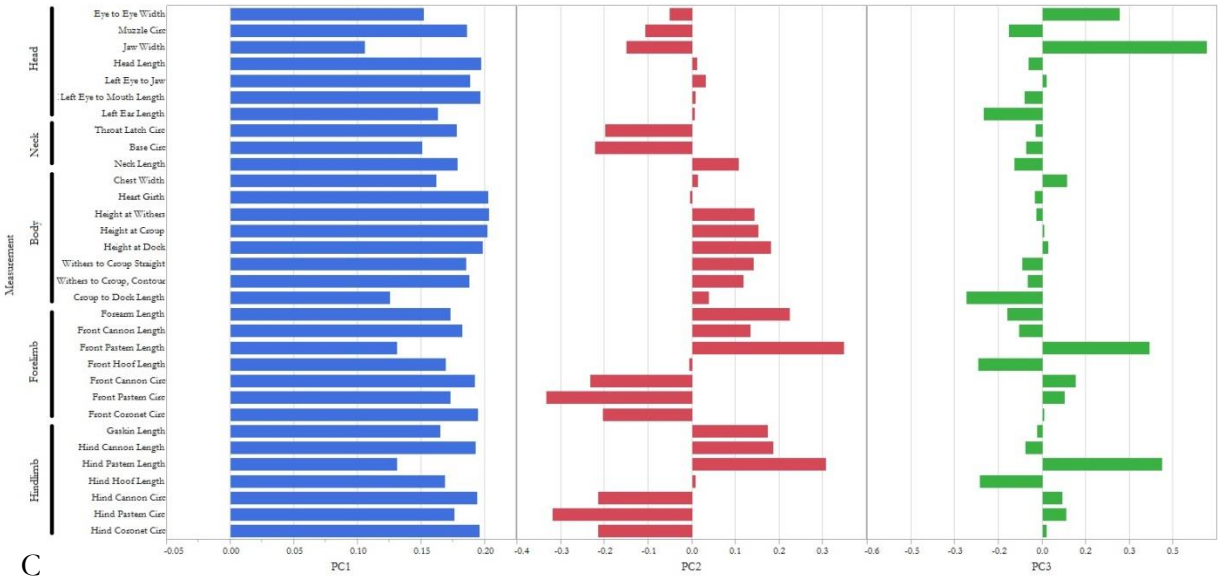
To better compare the influence of morphological traits on gait type, a pruned dataset was generated which was balanced for an equal number of mares, geldings and stallions gaited and non-gaited horses. In this balanced dataset PC1-3 were retained based on the eigenvalue, percent variance explained and scree plot (Figure 3.4a). PC1 accounted for 66.8% of variance, PC2 for 6.6%, and PC3 for 3.3% (Figure 3.4b) with each PC representing similar conformation trends in "size", "thickness" and "skull/pastern" traits as for the full dataset (Figure 3.4c). Only PC2 and PC3 scores were significantly different in gaited horses versus non-gaited; non-gaited horses had more negative PC2 (thickness) scores, but more positive PC3 (skull/pastern) scores. Only PC2 scores were significantly different for sex, while PC1 and PC2 were significantly different for gender (Table 3.3).



A

Number	Percent	Cum Percent
1	66.780	66.780
2	6.646	73.425
3	3.306	76.731
4	2.503	79.235
5	2.326	81.560
6	1.923	83.484
7	1.853	85.337
8	1.616	86.953
9	1.501	88.454
10	1.384	89.838
11	1.247	91.085
12	1.010	92.095
13	0.900	92.995
14	0.890	93.886
15	0.798	94.683
16	0.644	95.327
17	0.595	95.922
18	0.556	96.478
19	0.527	97.005
20	0.483	97.487
21	0.466	97.953
22	0.397	98.350
23	0.371	98.721
24	0.276	98.997
25	0.247	99.244
26	0.220	99.464
27	0.185	99.649
28	0.106	99.754
29	0.100	99.854
30	0.069	99.923
31	0.047	99.970
32	0.030	100.000

B



C

Figure 3.4. Principle component analysis of the 32 across breed balanced for gender and gait body measurements. The a)scree plot and b)percent of variation explained were used to determine the number of components retained for further analysis. C)factor loading of body measurements onto the first three principle components.

Table 3.4. Associations between balanced across breed PC scores and other traits of interest.

Trait	PC1	PC2	PC3
Sex†	P=0.4791	P=0.0035*	P=0.7364
Gender†	P=0.0029*	P=0.0006*	P=0.5972
Gait ability†	P=0.2553	P<0.0001*	P<0.0001*
Age‡	-0.0290	-0.1539	0.1137
	P=0.5482	P=0.0013*	P=0.0181*
Dish‡	0.4077	-0.0904	-0.1123
	P<0.0001*	P=0.0605	P=0.0196*
Feathering‡	0.1610	-0.6060	0.1052
	P=0.0009*	P<0.0001*	P=0.0311*
Bone thickness‡	0.4431	-0.5453	0.0962
	P<0.0001*	P<0.0001*	P=0.0462*
Bite‡	0.0729	0.0018	0.0354
	P=0.1310	P=0.9699	P=0.4635

† designates traits analyzed by ANOVA. ‡ designates traits analyzed by pairwise correlation with Pearson's p-values. * p-value significant at $\alpha<0.05$.

Table 3.5. ANOVA p-values for the balanced across breed body measurements normalized by wither height vs traits of interest.

Measure	Sex	Gender	Gaited?	Gait Mares	Gait Geldings	Gait Stallions
Wither Height#	0.3759	0.0003**	0.9769	0.8137	0.0263*	0.1203
Eye Width	0.8852	0.4880	<0.0001**	0.2391	0.0110*	<0.0001**
Jaw Width	0.0433*	0.0640	<0.0001**	<0.0001**	0.0274*	<0.0001**
Head Length	0.2373	0.4497	0.0791	0.0186*	0.3164	0.8577
Muzzle Circumference	0.4402	0.0124*	0.0713	0.4996	0.8342	0.0060*
Eye to Mouth Length	0.0867	0.1089	0.5489	0.7256	0.9529	0.1270
Eye to Jaw Length	0.3553	0.5334	0.0945	0.5491	0.0824	0.1199
Ear Length	0.0248*	0.0706	0.4556	0.2224	0.6043	0.7179
Neck Length	0.3434	0.1515	0.2553	0.9569	0.6712	0.0149*
Throatlatch Circumference	0.0008**	<0.0001**	0.0050*	0.0113*	0.5619	0.0457*
Neck Base Circumference	0.0003**	<0.0001**	<0.0001**	0.0625	0.0044*	0.0259*
Croup Height	0.1418	0.0766	0.0722	0.0233*	0.3132	0.8614
Dock Height	0.0069*	0.0072*	0.0568	0.0345*	0.2823	0.7269
Withers to Croup, Straight	0.0005**	<0.0001**	0.4121	0.5262	0.4212	0.2368
Withers to Croup, Contoured	<0.0001**	<0.0001**	0.0540	0.9777	0.2531	0.0333*
Croup to Dock Length	0.0197*	0.0612	0.0243*	0.6342	0.1650	0.0542
Chest Width	0.8611	0.3707	0.7173	0.6404	0.4315	0.5426
Heart Girth	0.0001**	<0.0001**	<0.0001**	0.0007**	0.0314*	0.0009**
Forearm Length	0.3846	0.5192	<0.0001**	0.0022*	<0.0001**	0.0038*
Front Cannon Length	0.6354	0.4981	<0.0001**	0.0027*	0.0016*	<0.0001**
Front Cannon Circumference	0.0132*	0.0061*	<0.0001**	<0.0001**	<0.0001**	<0.0001**
Front Pastern Length	0.8247	0.7513	0.1868	0.5506	0.0472*	0.3296
Front Pastern Circumference	0.0883	0.0270*	<0.0001**	0.0004**	0.0279*	<0.0001**
Front Coronet Circumference	0.3425	0.5336	<0.0001**	0.0025*	0.0197*	<0.0001**
Front Hoof Length	0.0804	0.0360*	0.0130*	0.2408	0.0782*	0.1711
Gaskin Length	0.9602	0.0040*	0.0740	0.1065	0.2015	0.8453
Hind Cannon Length	0.5556	0.5397	<0.0001**	0.0812	0.0115*	0.0009**
Hind Cannon Circumference	0.0910	0.1869	<0.0001**	<0.0001**	<0.0001**	<0.0001**
Hind Pastern Length	0.3042	0.3537	0.8323	0.3720	0.9580	0.7091
Hind Pastern Circumference	0.2141	0.0266*	<0.0001**	<0.0001**	0.0020*	<0.0001**
Hind Coronet Circumference	0.1220	0.3027	<0.0001**	0.0010**	0.0109*	<0.0001**
Hind Hoof Length	0.8057	0.2975	0.6264	0.3236	0.7846	0.8957

#Not normalized. *P-values significant at $\alpha<0.05$. **P-values significant at Bonferroni cutoff $\alpha<0.0015$.

To examine differences in the individual measures specific to gait type, the 32 body measures were normalized by wither height and compared between the gaited (216 horses) and non-gaited horses (216 horses) using ANOVA. 16 measurements were found to be statistically different in gaited breeds (p-values in Table 3.5). These include smaller eye and jaw width, smaller neck base, smaller heart girth, longer forearm and cannon lengths, and smaller limb circumferences in gaited horses.

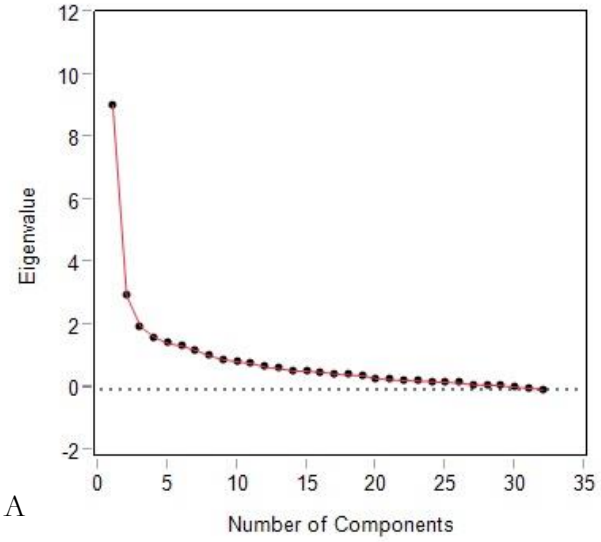
To better discern which sexually dimorphic measurements are influenced by alternate gait ability, we examined each gender class separately. Across all gender classifications, ten measurements were significantly different for alternate gait ability (Table 3.5). These included narrower jaw widths, longer forearms, longer cannons, and thinner cannons, pasterns, and coronets in gaited horses.

TWH Skeletal Variation is Driven not by Gait Type, but by Training Discipline

In order to better examine the influence of morphology on gait type within a single breed, we pulled a subset comprising only registered TWH from the full across-breed dataset. Based on the eigenvalues, percent explained and scree plot (Figure 3.5a) we retained PC1-8 for further analysis. PC1 explains 28.55% of the variance, PC2 explains 9.53%, PC3 6.4%, PC4 5.4%, and the remaining four PCs explaining between approximately 5-3.5%, totaling 66.6% of the total variance for all 8 PCs (Figure 3.5b). For PC1 all measurements loaded in a positive direction, indicating a positive correlation and quantifying overall body "size" within the TWH (Figure 3.5c). Circumferences and lengths loaded in opposite directions, with the exceptions of eye to mouth length, both withers to croup lengths, and croup to dock lengths, which loaded with circumferences and eye to eye width which loaded with lengths on PC2 (Figure 3.5c). Based on the overall pattern of factor loading, PC2 quantifies body variation due to bone "thickness". The major contributors to PC3 are the withers to croup measurements which load in a negative direction, indicating PC3

likely explains "relative back length" (Figure 3.5c). The remaining PCs represent more understated patterns of body shape variation (Figure 3.5c&d).

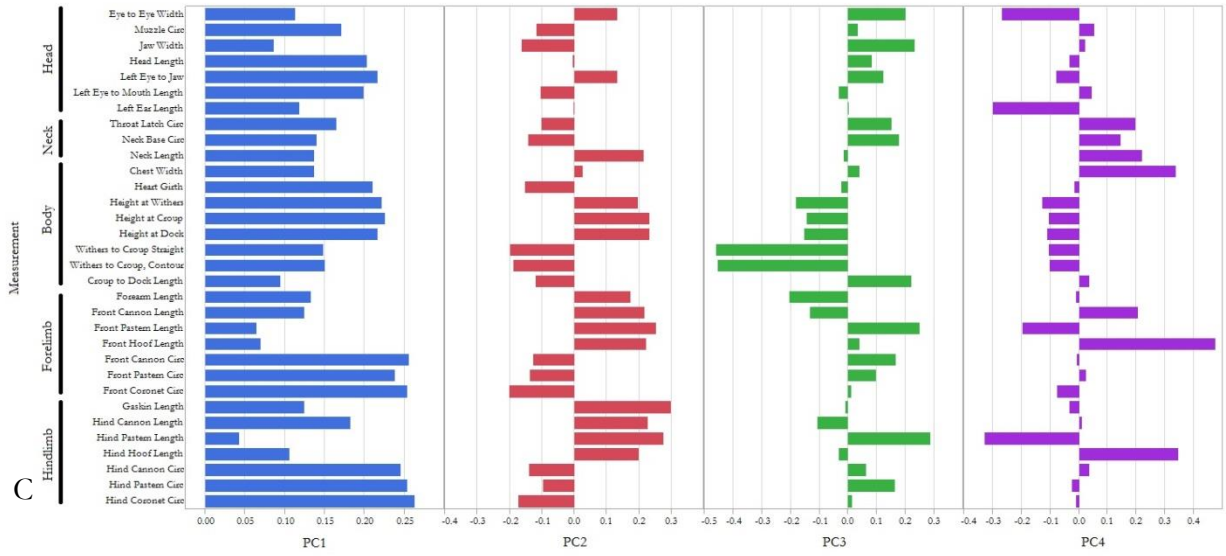
Figure 3.5. Principle component analysis of the 32 TWH body measurements. The a)scree plot and b)percent explained were used to determine the number of components retained for further analysis. c)Factor loading of body measurements onto principle components.



A

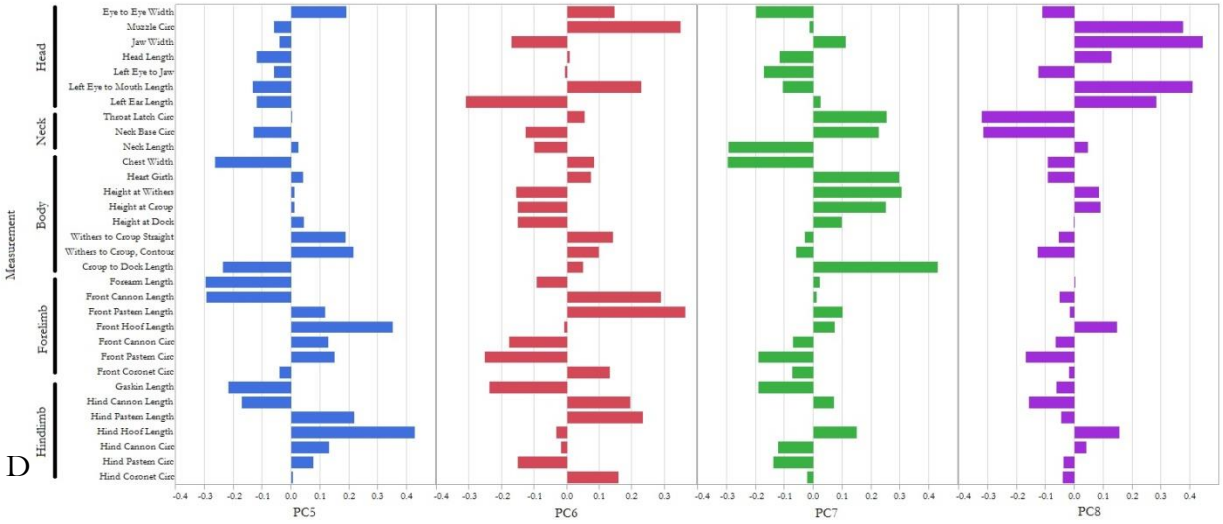
Number	Percent	Cum Percent
1	28.548	28.548
2	9.524	38.072
3	6.433	44.506
4	5.358	49.864
5	4.728	54.592
6	4.461	59.053
7	4.055	63.108
8	3.487	66.595
9	3.104	69.699
10	2.917	72.616
11	2.737	75.353
12	2.368	77.721
13	2.218	79.939
14	2.011	81.950
15	1.970	83.919
16	1.871	85.790
17	1.698	87.488
18	1.583	89.071
19	1.518	90.589
20	1.234	91.824
21	1.113	92.937
22	1.049	93.986
23	0.989	94.976
24	0.923	95.899
25	0.862	96.761
26	0.781	97.542
27	0.617	98.159
28	0.593	98.752
29	0.517	99.269
30	0.361	99.629
31	0.266	99.895
32	0.105	100.000

B



C

Figure 3.5 (Continued)



D

When we compared the TWH component scores to gait type, after including gender in the statistical model, PC1 scores were significantly different for gait type (Table 3.6), with lateral horses having higher PC1 scores, and therefore larger bodies, than multi-gaited horses. We found that age was negatively correlated with PC2, 4, 7 and 8 scores (Table 3.6). We examined the correlations between the additional factor scores of head profile, level of feathering on lower limbs, subjective bone thickness and incisor alignment, but identified fewer correlations than in the across breed analysis. Head profile was only positively correlated with PC1 scores. Feathering was positively correlated with PC5 scores and negatively correlated with PC6 scores. Bone thickness was positively correlated with PC1 and PC5 scores, and negatively correlated with PC4 scores. Within the TWH there is a significant correlation for incisior alignment with skeletal conformation: PC8 scores, which seem to describe head shape patterns (Figure 3.4d) were negatively correlated with scores for incisor alignment (Table 3.6).

Table 3.6. Associations between TWH PC scores and other traits of interest.

Trait	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Sex†	P<0.0001**	P=0.0735	P=0.0836	P<0.0001**	P=0.6077	P=0.4330	P=0.5747	P=0.0113*
Gender†	P<0.0001**	P=0.0472*	P=0.0016**	P<0.0001**	P=0.0828	P=0.5864	P=0.8156	P=0.0246*
Gait + Gender#	P=0.0018**	P=0.3144	P=0.3514	P=0.1449	P=0.7054	P=0.0520	P=0.4802	P=0.9204
Trail + Gender#	P=0.2078	P=0.0043**	P<0.0001**	P=0.1893	P=0.0131*	P=0.1234	P=0.3571	P=0.1787
Show + Gender#	P=0.8345	P<0.0001**	P=0.1939	P=0.0027**	P=0.3025	P=0.1493	P=0.6880	P=0.3942
Breeding + Sex#	P=0.0091**	P=0.0702	P<0.0001**	P<0.0001**	P=0.5461	P=0.3345	P=0.3133	P=0.0520
Age‡	0.1158	-0.2148	-0.0702	-0.039	0.0989	0.0013	-0.2394	-0.1784
	P=0.0525	P=0.0003**	P=0.2408	P=0.5154	P=0.098	P=0.9828	P<0.0001**	P=0.0027**
Dish‡	0.156	-0.0037	0.0151	-0.0722	-0.0539	0.0154	-0.0185	-0.0933
	P=0.0091*	P=0.9511	P=0.8021	P=0.2292	P=0.3697	P=0.7979	P=0.7586	P=0.1199
Feathers‡	0.0787	-0.0242	0.0974	0.0508	0.1286	-0.1332	-0.0079	0.0748
	P=0.2162	P=0.7043	P=0.1252	P=0.4250	P=0.0426*	P=0.0357*	P=0.9016	P=0.2395
Bone‡	0.4722	-0.089	0.0092	-0.1375	0.1391	-0.0599	0.0597	0.0448
	P<0.0001**	P=0.1383	P=0.8785	P=0.0216*	P=0.0201*	P=0.3190	P=0.3201	P=0.4564
Bite‡	-0.0437	0.0786	0.0562	0.1048	-0.087	-0.1247	-0.0505	-0.1217
	P=0.4684	P=0.1913	P=0.3506	P=0.0812	P=0.1479	P=0.0377*	P=0.4015	P=0.0426*

† designates traits analyzed by ANOVA. # designates traits analyzed by ANCOVA. ‡ designates traits analyzed by pairwise correlation with Pearson's p-values. * designates p-value significant at $\alpha<0.05$. ** designates p-value significant at $\alpha<0.0062$

Various skeletal conformations are better suited to distinct uses, and are often under artificial selection based on training discipline. In addition to body measurements, we collected discipline information from 244 TWH horses classified into the following groups: trail (includes animals used for pleasure), show (used exclusively for competition), and breeding (animal is only used for reproduction purposes). Trail use was associated with negative PC2 and 5 scores and positive PC3 scores (Table 3.6), describing a thick horse with a narrow chest, short back and rump, and short forelimbs. Show use was associated with positive PC2 and PC4 scores (Table 3.6), describing a thin horse with narrow set eyes, a wide chest, short pasterns and ears, and long hooves.

To confirm the observed differences within a single breed, we normalized the 32 body measures by wither height and compared the measures between the gait types and disciplines. After correcting for gender, six measurements were significantly different for gait type (Table 3.7). However, for three of these measurements there is a significant interaction between gait and the covariate gender, indicating we cannot separate the effects of gender from the effects of gait type. Therefore, only wither height, front pastern length and front pastern circumference are significantly different for gait type, in which lateral gaited horses have taller withers, shorter and wider front pasterns than multi-gaited horses. Lateral gaited geldings were 1.22 to 1.96 inches taller at the withers than other gaited horses.

Eight individual measurements were significantly different for trail use, 18 for show use, and 10 for breeding (Table 3.7). After correcting for gender, three measurements were significantly different for trail and nine for show; five measurements were significantly different for breeding use after correcting for sex. These include narrower chest width, longer ears, and shorter gaskins in trail horses, narrower muzzles, smaller heart girths, narrower front coronets, and longer hooves in show horses, and wider heart girths, longer front cannons, and longer gaskins in animals used for breeding.

Table 3.7. ANOVA p-values for TWH body measurements normalized by wither height vs traits of interest.

Measure	Gender	Gait + Gender	Trail	Trail + Gender	Show	Show + Gender	Breeding	Breeding + Sex
Wither Height#	<0.0001**	0.0005**	0.1754	0.5524	0.0134*	0.0598	0.0074*	0.0006**[i]
Eye Width	0.6494	0.9388	0.7198	0.2636	0.1233	0.5041	0.4162	0.1958
Jaw Width	0.2211	0.0111*[i]	0.0779	0.0037*[i]	0.0111*	0.0136*	0.6926	0.0883
Head Length	0.6125	0.4976	0.6804	0.3343	0.6588	0.5237	0.2705	0.6086
Muzzle Circumference	0.4122	0.9221	0.0296*	0.1507	0.0015**	0.0270*	0.9425	0.4739
Eye to Mouth Length	0.2607	0.6089	0.7847	0.3595	0.0089*	0.1039	0.0670	0.1034
Eye to Jaw Length	0.8681	0.0423*[i]	0.8061	0.3813	0.7110	0.9641	0.9657	0.2356
Ear Length	0.0007**	0.9897	0.2257	0.0011**	0.0500*	0.5070	0.4731	0.2019
Neck Length	0.1875	0.8321	0.7931	0.3798	0.0229*	0.1150	0.9852	0.0132*[i]
Throatlatch Circumference	<0.0001**	0.2083	0.5227	0.1222	0.3976	0.2865	0.0314*	0.0002**[i]
Neck Base Circumference	0.0357*	0.0761	0.6051	0.2154	0.6080	0.2311	0.2117	0.0292*[i]
Croup Height	0.7381	0.5836	0.6332	0.9286	0.8580	0.6998	0.1937	0.3844
Dock Height	0.4300	0.3739	0.0837	0.5609	0.9569	0.6425	0.0025*	0.0167*
Withers to Croup, Straight	0.0001**	0.1973	0.2848	<0.0001**[i]	0.0027*	<0.0001**	0.4034	0.0006**[i]
Withers to Croup, Contoured	0.0016*	0.1495	0.1381	<0.0001**[i]	0.0174*	0.0021*	0.2069	0.0029*[i]
Croup to Dock Length	0.3193	0.2220	0.0206*	0.1212	0.6825	0.5020	0.0167*	0.0573
Chest Width	0.0108*	0.5107	0.0042*	0.0029*	0.1697	0.8276	0.0040*	0.0002**
Heart Girth	<0.0001**	0.6494	0.1501	0.5152	<0.0001**	<0.0001**	0.0131*	0.0791
Forearm Length	0.5963	0.8943	0.0189*	0.3222	0.5773	0.7180	0.0125*	0.0898
Front Cannon Length	0.0107*	0.2909	0.0704	0.0640	0.0373*	0.0092*	0.0078*	0.0081*
Front Cannon Circumference	0.1087	0.6947	0.0079*	0.0679	0.0019*	0.0091*	0.2612	0.1565
Front Pastern Length	0.9242	0.0326*	0.8792	0.7199	0.1227	0.3712	0.6752	0.9502

#Not normalized. *P-values significant at $\alpha<0.05$. **P-value significant at Bonferroni cutoff $\alpha<0.0015$. [i] designates significance driven by interaction between gender and gait.

Table 3.7. (Continued).

Measure	Gender	Gait + Gender	Trail	Trail + Gender	Show	Show + Gender	Breeding	Breeding + Sex
Front Pastern Circumference	0.0800	0.0266*	0.1082	0.4967	0.0737	0.0891	0.2038	0.1912
Front Coronet Circumference	0.0480*	0.4900	0.1975	0.2054	0.0003**	0.0024*	0.0552	0.0632
Front Hoof Length	0.0007**	0.1998	0.3756	<0.0001**[i]	<0.0001**	<0.0001**	0.6905	<0.0001**[i]
Gaskin Length	0.0087*	0.6413	0.0089*	0.0010**	0.0471*	0.2258	0.0795	0.0028*
Hind Cannon Length	0.0492*	0.1081	0.0381*	0.0924	0.6415	0.1126	0.0139*	0.0145*
Hind Cannon Circumference	0.7917	0.9405	0.2167	0.7824	0.0491	0.1236	0.1646	0.3651
Hind Pastern Length	0.2020	0.0419*[i]	0.3786	0.0121*[i]	0.1819	0.5850	0.6557	0.1310
Hind Pastern Circumference	0.3338	0.6802	0.0094*	0.0607	0.0381*	0.0411*[i]	0.9842	0.6683
Hind Coronet Circumference	0.0793	0.5154	0.3786	0.3870	0.0122*	0.0719	0.0102*	0.0225*
Hind Hoof Length	0.3485	0.1704	0.4060	0.0032*[i]	0.0250*	0.1958	0.9710	0.1984

*P-values significant at $\alpha<0.05$. **P-value significant at Bonferroni cutoff $\alpha<0.0015$. [i] designates significance driven by interaction between gender and gait.

TWH PC1 Maps to LCORL/NCAPG

Genotypes were available for 105 horses enabling a GWA study to identify loci controlling body size in TWH. Linear quantitative association run under a dominant model with sex added as a covariate using Golden Helix SVS software on PC1 identified candidate loci on ECA3, ECA18, and ECA20, with p-values ranging from 2.35e-9 to 5.37e-5 (Table 3.8). Genomic inflation was limited to a factor of 1.06, indicating limited stratification due to breed substructure. This uncorrected substructure can also be detected in the QQ-plot (Figure3.6a&b) where there is an obvious deviation from the expected, identity line. The EMMAX model applied to the same dataset verified candidate loci on ECA3 and ECA18 (Figure3.6c) with p-values ranging from 3.86e-9 to 6.00e-5 (Table 3.8) and was able to correct for some of the population substructure, as observed in the QQ-plot (Figure3.6d) and with the genomic inflation reduced to a factor of 1.01. Only loci on ECA3 surpassed Bonferroni correction in both analyses. The three markers on ECA3 span a 0.38 Mb region that encompasses the *LCORL/NCAPG* genes. The best associated SNP is located approximately 100 Kb upstream from the *LCORL/NCAPG* genes (Figure 3.7a&b). The haplotype structure reveals several SNPs that are in LD around the genes (Figure 3.7c). The additional four SNPs included in the dataset from ECA9 in the *ZFAT* and *KHDRBS3* genes were not significantly associated with PC1 in the TWH.

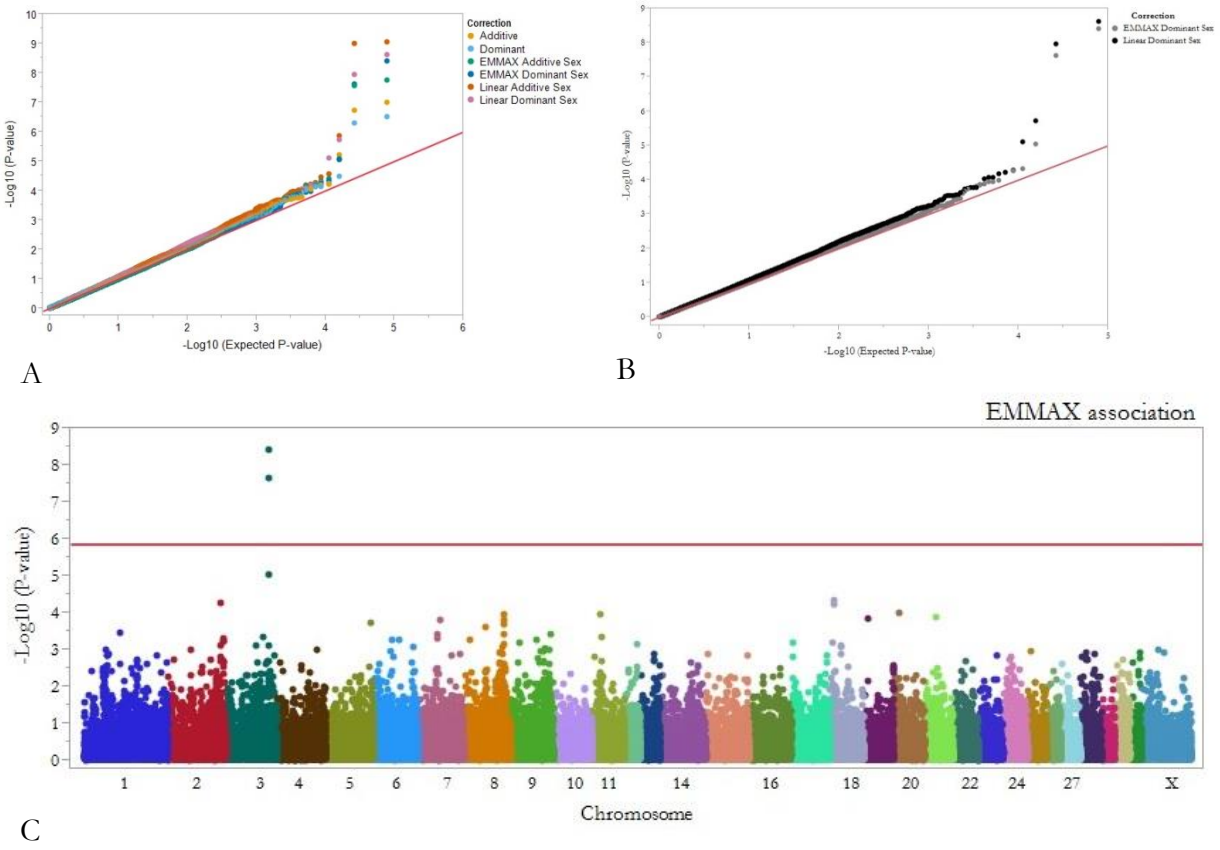


Figure 3.6. Quantitative association of PC1. Quantile-quantile plot of a) all associations and b) linear and EMMAX dominant models. C) Manhattan plot of the EMMAX dominant model. The red line signifies Bonferroni significance.

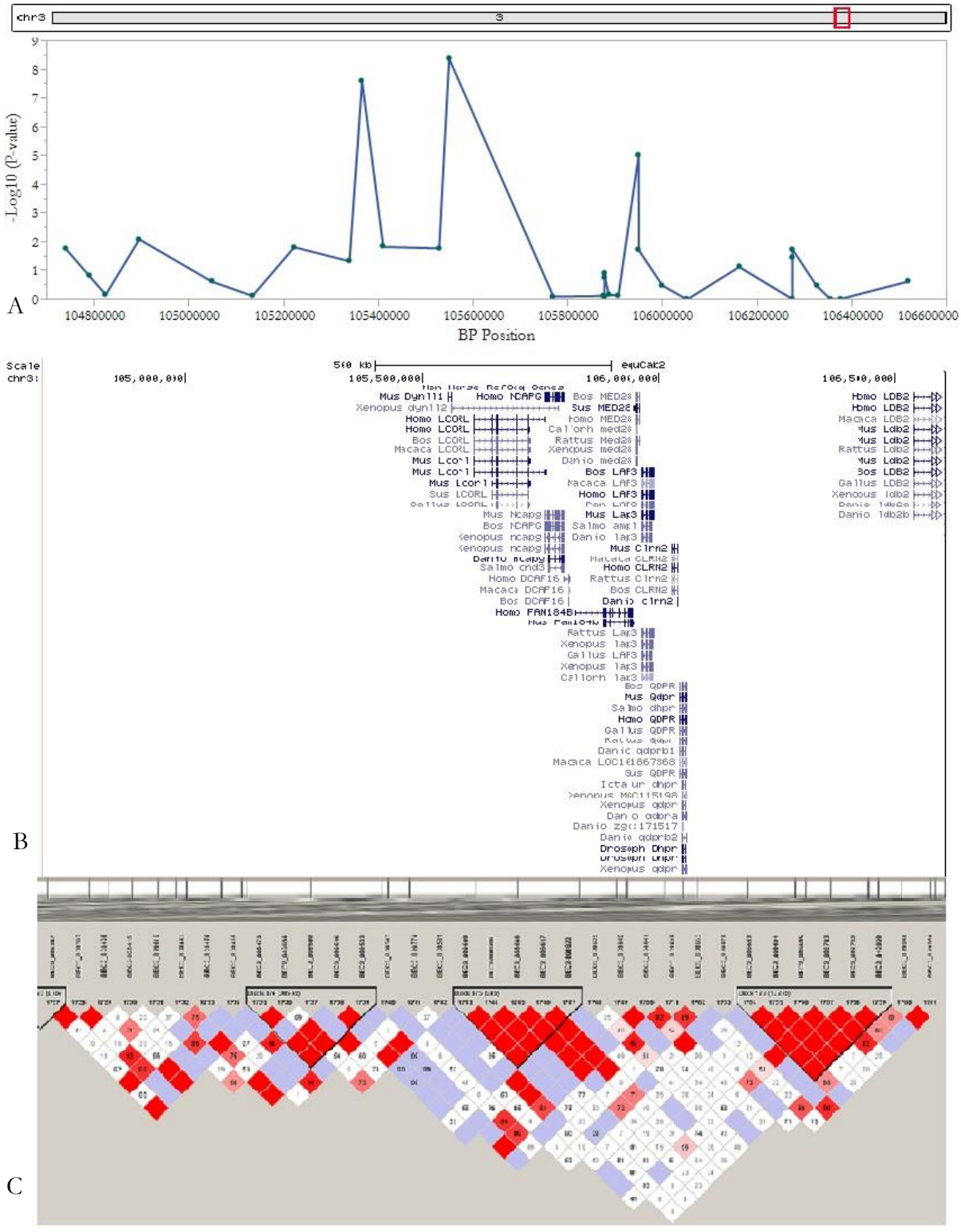


Figure 3.7. PC1 ECA3 candidate region a) p-values, b) other species RefSeq genes adapted from UCSC Genome Browser reveals c) associated haplotypes in *LCORL/NCAPG* region.

Table 3.8. SNPs associated with TWH PC1 from linear and EMMAX dominant associations.

SNP	Chr	Position (bp)	Minor Allele	Major Allele	Linear P-value	EMMAX P-value	Marker in Gene
BIEC2_808543	3	105547002	C	T	2.35E-09*	3.86E-09*	*LCORL, NCAPG*
BIEC2_808500	3	105363241	T	G	1.10E-08*	2.33E-08*	*LCORL, NCAPG*
BIEC2_808640	3	105947243	C	T	1.87E-06	9.01E-06	*LCORL, NCAPG*
BIEC2_544015	20	10132609	T	G	7.96E-06	1.06E-04	396 Kb region with no annotations
BIEC2_397289	18	8489067	T	G	5.37E-05	4.70E-05	2.3 Mb region with no annotations
BIEC2_506191	2	111224478	G	A	1.36E-04	5.62E-05	850 Kb region with no annotations
BIEC2_397290	18	8489231	T	C	6.58E-05	6.00E-05	2.3 Mb region with no annotations

*P-value surpassed Bonferroni threshold of 1.38×10^{-6}.

TWH PC2 Maps to 10 Potential Candidate Loci

Genotypes were available for 105 TWH to identify loci contributing to PC2- "thickness" and PC3 body shape patterns. For both associations we were unable to detect any Bonferroni significant loci. Suggestive loci for PC2 are on ECA7, ECA23, and ECA30 (Figure 3.8a and Table 3.9). Initial genomic inflation was a factor of 1.13976 which reduced to 1.06178 after EMMAX, as demonstrated by early deviations from the expected line in the QQ-plot (Figure3.8b) suggesting population structure may have inflated p-values across the dataset. For PC3, initial genomic inflation was a factor of 1.33366 which reduced to 0.99471 after EMMAX correction, indicating stratification due to breed substructure, as observed in the qq-plots (Figure 3.9a). The EMMAX models were able to correct for the population stratification, but overcorrected due to likely insufficient power in the GWAs (Figure 3.9a). After one million permutations, 12 candidate loci were identified on ECA2, ECA8, ECA9, ECA19, and ECAX (Figure 3.9b) with p-values ranging from 3.00e-6 to 3.30e-5 (Table 3.10). Power may have been limited in these studies by the small sample size. PC2 only accounted for 9.53% of the total size variation and PC3 6.4%, and 105 horses is likely an insufficient sample size for these more subtle and complex trait.

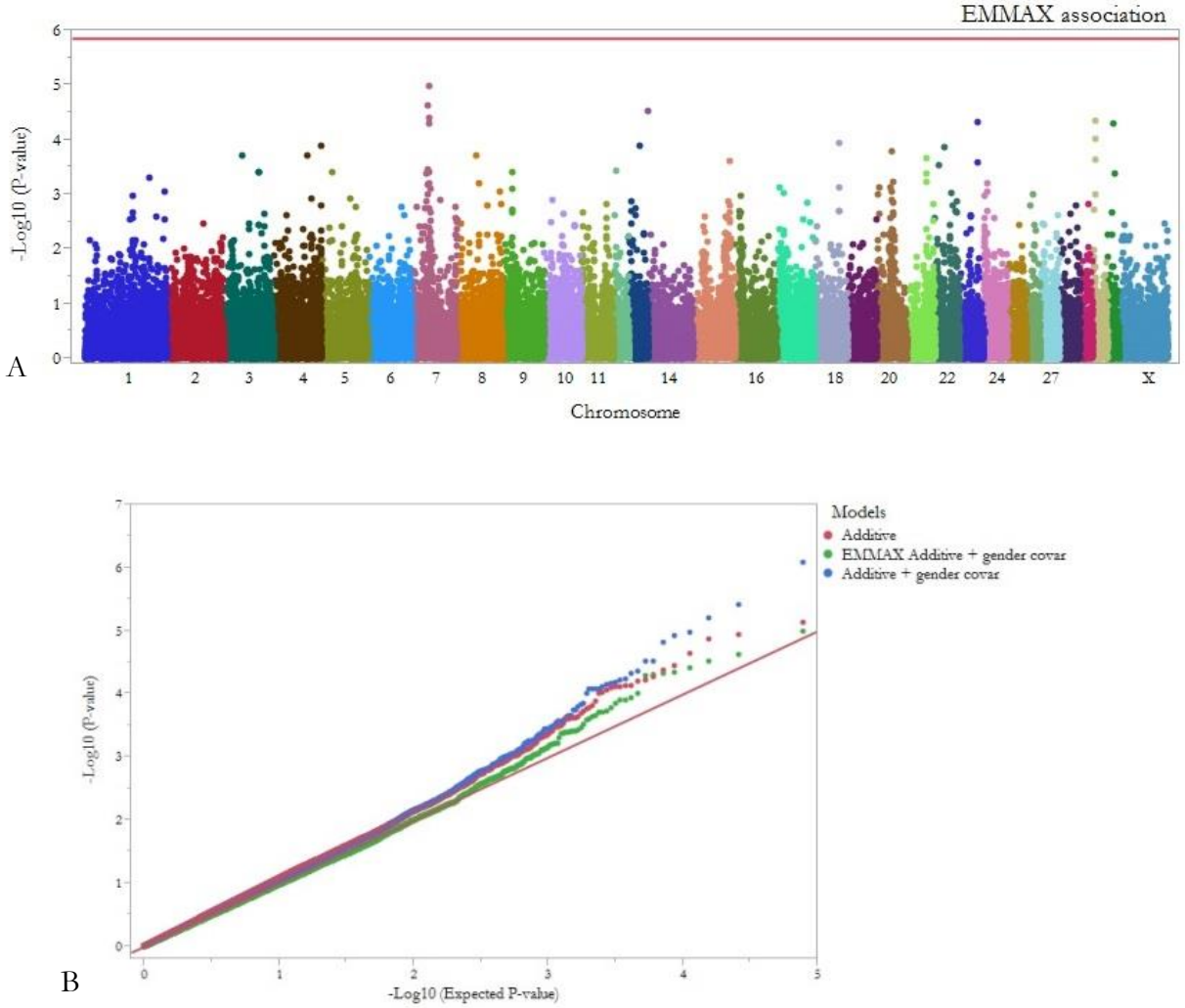


Figure 3.8. Quantitative association of TWH PC2-"thickness". A) EMMAX additive model with gender covariate manhattan plot highlights several suggestive loci. Red line designates Bonferroni significance. B) Quantile-quantile plot of additive model associations displaying deviations from expected line and convex curvature.

Table 3.9. SNPs associated with PC2 - "thickness" from EMMAX model.

SNP	Chr	Position	Minor Allele	Major Allele	EMMAX P-Value	Marker in Gene
BIEC2_994454	7	34979113	T	C	1.03E-05	48 Kb upstream of *RPUSD4, FAM118B*
BIEC2_992373	7	33305275	A	G	2.42E-05	*NRGN*
BIEC2_248285	14	697905	A	G	2.98E-05	868Kb region no annotation
BIEC2_994466	7	35006118	C	A	3.90E-05	22Kb upstream: *RPUSD4, FAM118B*
BIEC2_814617	30	4038552	C	A	4.48E-05	*FMN2*
BIEC2_656537	23	41038852	T	C	4.69E-05	*CDKN2B-AS1*
BIEC2_879217	31	11273216	C	T	5.00E-05	216Kb downstream: *ARID1B*
BIEC2_1048923	7	35134279	A	G	5.08E-05	*DCPS*
BIEC2_578956	22	7176739	G	A	3.00E-04	*PCSK4, BFSP1*
TBIEC2_1102523	8	37506696	T	C	1.94E-04	*PSMG2*
BIEC2_529555	20	31537319	A	G	1.67E-04	*VARS*
TBIEC2_915904	4	79641162	T	C	1.98E-04	*HYAL4*
BIEC2_624649	23	41041826	T	G	1.98E-04	*CDKN2B-AS1*

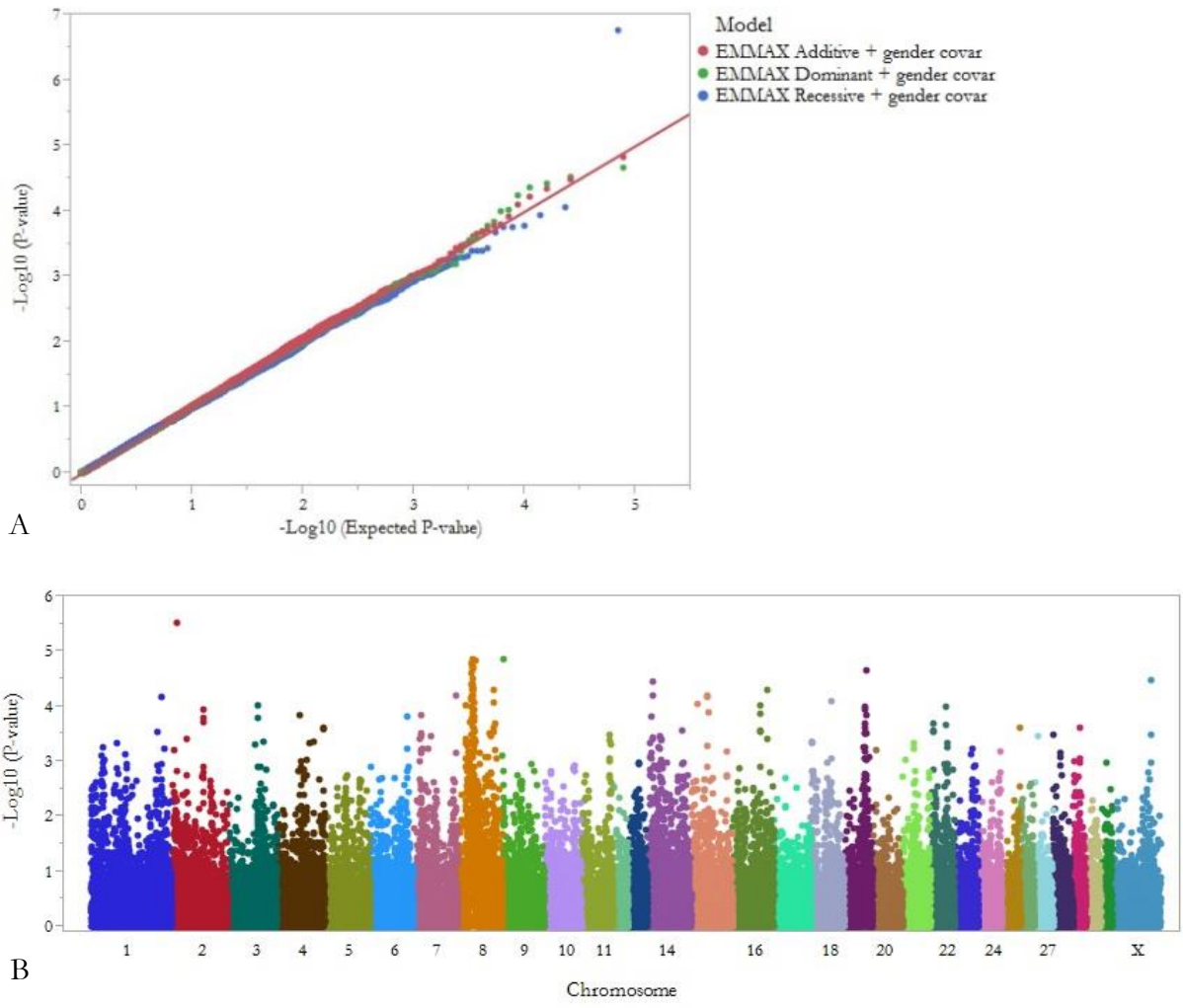


Figure 3.9. Quantitative association of TWH PC3 body shape highlights several suggestive loci. A) Quantile-quantile plot of EMMAX specific models. B) Manhattan plot of permutated p-values.

Table 3.10. Permutated p-values of candidate SNPs for TWH PC3 body shape.

SNP	Chr	Position	Minor Allele	Major Allele	Permutated P-value	Permutations	Marker in Gene
BIEC2_457100	2	12896028	A	G	3.00E-06	1000000	14Kb upstream: *PRDX1*
BIEC2_1041042	8	29419499	C	T	1.40E-05	1000000	*GALNT9*
BIEC2_1068497	9	2203592	A	G	1.40E-05	1000000	49Kb upstream: *CNGB3;* 60Kb downstream *CNDB1*
BIEC2_1043683	8	35025281	C	T	1.50E-05	1000000	intergenic
BIEC2_1039670	8	27419544	C	T	1.70E-05	1000000	*TMEM132D*
BIEC2_1041126	8	29716343	A	C	1.90E-05	1000000	*GOLGA3*
BIEC2_1041132	8	29729387	T	C	2.00E-05	1000000	*GOLGA3*
TBIEC2_1098553	8	30756405	T	C	2.10E-05	1000000	21Kb upstream *EPB41L3*
BIEC2_443350	19	47016375	T	G	2.20E-05	1000000	1.3Mb region with no gene annotation
BIEC2_1040563	8	28452448	A	G	2.50E-05	1000000	93Kb upstream *GPR133*
BIEC2_1041531	8	30729188	G	A	3.10E-05	1000000	*EPB41L5*
BIEC2_1148144	32	1.02E+08	G	T	3.30E-05	1000000	179Kb downstream: *TTC14*

Discussion and Conclusions

Body conformation is a critically important trait in nearly all horse breeds. It has been well documented that conformation influences movement; hip and should angles influence the range of motion of limbs, while limb segment lengths influence shock absorption and strength of the limbs. For example, longer pasterns increased the odds of a fracture in the front limbs of racing Thoroughbreds (Anderson, McIlwraith, Douay 2004). Yearlings taller at the withers and hip were more likely to win or come in second in graded stakes races on turf tracks (Smith, Staniar, Splan 2006). Our initial overall goal for this study was to identify morphological features unique to alternate intermediate gaits. Gait specific conformations exist across breeds, but within a single breed conformation is predominantly specific to training discipline.

The inclusion of approximately 500 more horses slightly reduced the percent variance explained by each principle component, but allowed for an additional PC and pattern to be retained and examined compared to earlier work (Brooks *et al.* 2010). Gaited breeds are typically 0.75 inches (1.905 cm) smaller than average for non-gaited, with median whither heights ranging from 53 inches (Icelandic Horse) to 64.5 inches (Saddlebred), but most are around 59 inches tall (Figure 3.10). Within the TWH dataset, the decreased percent variance explained is a result of the limited variation within the breed as compared to the total variation seen across multiple diverse breeds.

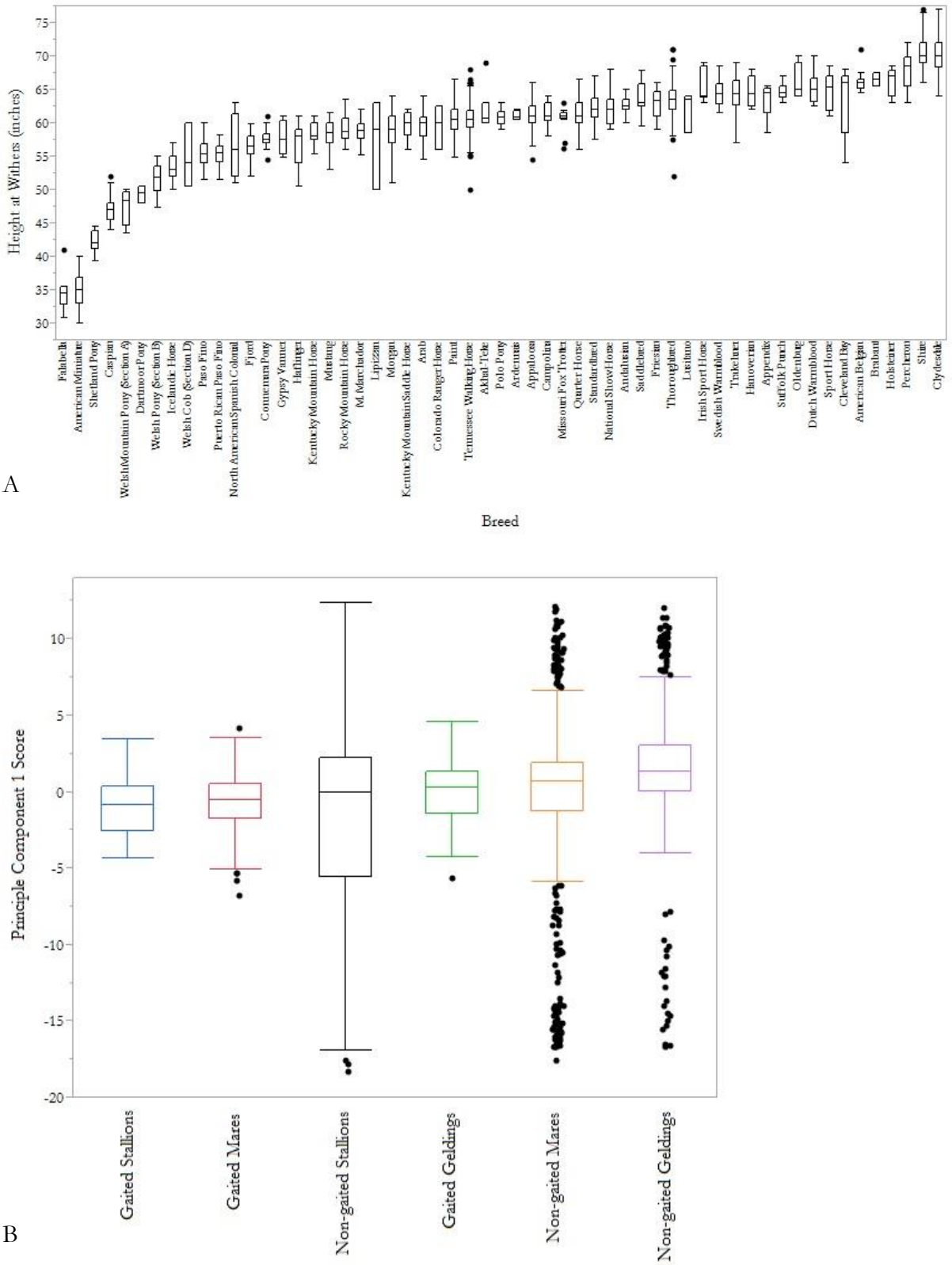


Figure 3.10. Median (a) wither heights across breeds versus median (b) PC1 scores across breeds.

PCA identifies relationships by generating linear combinations of variables showing common trends of variations (Gauch 1982; Peres-Neto, Jackson, Somers 2005); the first component extracted accounts for the largest amount of total variation observed in the sample set (O'Rourke and Hatcher 2013), so with larger sample sets there is more variation across the variables to be pulled out. The overall pattern of PC2 is the same, so the reduction in horse number was not detrimental to the comparison of gaited versus non-gaited conformation types. Within breed, the pattern for PC2 slightly changes due to the decrease in variation observed within a breed; breeds are typically selected to match a specific standard suitable for the breed's purpose and therefore should be fairly uniform in conformation. Within the TWH dataset, the pattern is fairly similar for PC2 with lengths and circumferences loading opposite of each other. However, gaskin length was the largest contributor to PC2 in TWH whereas it was front pastern circumference in the balanced dataset. PC2 was significantly associated with use category (show, trail, breeding, etc.) indicating that PC2 has likely captured and explains/predicts the different typical body shapes of TWH used in different jobs. For example, a horse with a positive PC2 score will have longer leaner legs, a longer thinner neck, a shorter back, a wider forehead, but a narrower muzzle and jaw width resulting in a more "refined" looking horse (Figure 3.11). Refined horses may be preferred in the TWH show ring, while thicker horses may be more durable for trail riding (London 2012; Womack 1994; Ziegler 2005).

The body shape pattern described by PC3 is likely breed specific. In TWH, the major contributors are back length followed by pastern lengths while in the balanced-for-gait-type dataset, pastern lengths are the major contributors.

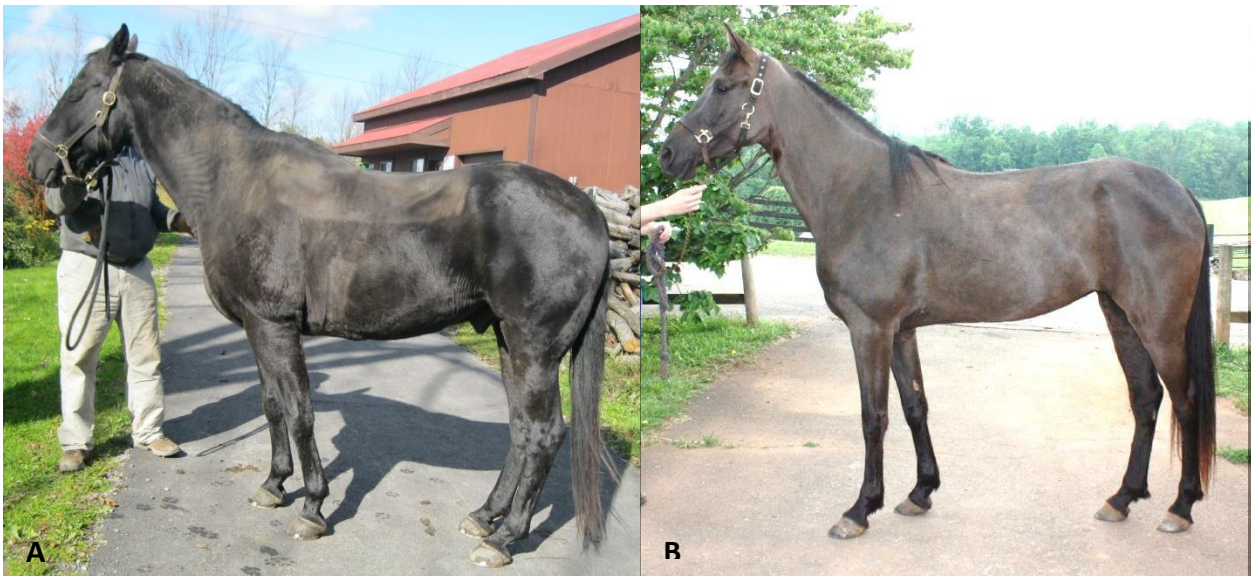


Figure 3.11. Examples of extreme PC2 scores at the same wither height (61.5 inches) in the TWH. Negative score horse in panel A used for trail and positive score horse in panel B used for show.

Sex and gender play important roles in determining body size and shape. Across breeds, sex (male vs female) and gender (stallion vs gelding vs mare) were significantly different for PC1 and PC2 with males typically having more positive PC scores than females. This difference is largely driven by geldings (606 geldings vs 910 mares, vs 167 stallions). However, with the underrepresentation of stallions in the PCA, our ability to detect stallion variation is diminished and favors mares and geldings. In examining the individual measurements (both across breeds and in TWH), geldings are identified as the tallest at the withers followed by mares then stallions. This pattern has also been observed in several other breeds, including Arabians (Sadek, Al-Aboud, Ashmawy 2006), Thoroughbred (Biedermann and Schmucker 1989), and more. There is a common misconception held by horse owners and breeders that stallions are taller than mares. This may be due to a human stereotype of a larger, more muscular, male physique. For example, for official registration of a Mangalarga Marchador (ABCCMM) stallion he must be between 57.8 and 62 inches tall (ideal height around 59.8 inches), yet mares must only achieve a height of between 55.1 to 61 inches tall (ideal around 57.8 inches); this could be a problem for an owner with a tall mare, though she is an ideal choice as a dam of future tall stallions (Associacao Brasileira de Criadores do Cavalo Mangalarga Marchador 2006).

Only PC2 and 3 from the balanced-for-gait-type dataset were significant for gait type. Gaited horses were not significantly different from non-gaited horses at wither, croup and dock heights, but their limb segment lengths were longer indicating that limb proportion and angles play a role in distinguishing gaited horses from non-gaited horses. Longer hind limb proportion in gaited horses has been observed by professional gaited horse trainers (Lane 2011; London 2012; Ziegler 2005), but this is the first scientific work to systematically compare conformation traits between gaited and non-gaited breeds. Furthermore, we have identified different proportions in the front limbs not previously observed.

In comparing gait type within a single breed, the TWH, the differences in conformation are much more subtle, and likely relate to the quality of the animation of the limbs. The multi-gaited horses (those able to perform lateral and diagonal gaits) were shorter yet have longer front pasterns. The pasterns act as both shock absorbers and propulsion generator, with longer sloping pasterns aiding in cushioning the impact (Clayton 2004) and lengthening the arc of forward motion (Bennett 2012).

Conformation may better predict suitable training discipline than ability to perform one gait type over another. Several studies highlight conformation characteristics of elite performance horses (Albertsdóttir *et al.* 2008; Anderson, McIlwraith, Douay 2004; Aranason 1984; Dolvik and Klemetsdal 1999; Holmström and Philipsson 1993; Holmström, Fredricson, Drevemo 1994; Magnusson and Thafvelin 1990; McIlwraith, Anderson, Sanschi 2003; Smith, Staniar, Splan 2006), but few have examined conformation across horses of the same breed but used for different disciplines. The majority of the horses collected for this study were used for trail riding (n=87), followed closely by use in show (n=59). Some animals were considered dual purpose, as trail and show animals (n=55), and others strictly for breeding (mostly broodmares never trained to be ridden, n=35). A few young horses, over the age of two years, were just starting their under-saddle training (n=8).

TWH is promoted as a dual-purpose (show and trail horse). Early determination of animals with conformation types ideal for trail, just show or both could be beneficial for breeders with untrained broodmares or young stock. In the comparison of conformation measures between horses that were trail or show, there were more measurements associated with being a show horse than with a trail horse, indicating there is likely greater selection pressure on show horse conformation than for other disciplines. Key features important for show horses were overall lengths and refinement; show horses were shorter in the back, narrower in limb circumferences,

smaller heart girths and tended to have relatively smaller heads. The inverse of these measurements were true for horses not involved in show, which would be important if the horse was used for trail work.

Horses not involved in showing had wider jaws, muzzles and heart girths, which would be important for increased respiration rate and blood flow (Weeren and Crevier-Denoix 2006). The thicker limbs are important for carrying weight, reducing limb stress over long distances and varied terrain. Show horses are usually ridden on level ground, typically sand or dirt arenas, for short periods of 15 to 20 minutes at a time in the show ring or 30 to 60 minutes in training sessions (shows for American gaited breeds such as the TWH; this time is longer for some European and South American breeds), so endurance is usually less of a concern. Hoof length was longer in show horses, but this is not surprising as hoof length, angle, and shoe weight are often manipulated to change the flight and timing of the footfall in show horses (Ziegler 2005).

Interrogation of the Genome for Loci Contributing to Size Variation in TWH

Quantitative association analysis for PC1 revealed three candidate markers on ECA3 that are located upstream from the *LCORL/NCAPG* genes. These genes have already been identified in other association studies for height in humans, cattle, and horses (Boyko *et al.* 2014; Karim *et al.* 2011; Lango Allen *et al.* 2010; Lindholm-Perry *et al.* 2013; Makvandi-Nejad *et al.* 2012; Metzger *et al.* 2013; Pausch *et al.* 2011; Pryce *et al.* 2011; Signer-Hasler *et al.* 2012; Soranzo *et al.* 2009; Tetens *et al.* 2013; Visscher, McEvoy, Yang 2010; Yang *et al.* 2010). In humans and horses no causative mutation in *LCORL* or *NCAPG* has been identified (Boyko *et al.* 2014; Signer-Hasler *et al.* 2012). In cattle, a non-synonymous variant in the *NCAPG* gene is proposed as a potential causative variant for various growth-related traits (Setoguchi *et al.* 2009; Setoguchi *et al.* 2011), but there is no functional proof for the causality of this variant (Signer-Hasler *et al.* 2012). In horses, decreased expression levels of *LCORL* gene in cDNA samples from hair are associated with larger, heavier

breeds of horses (Metzger *et al.* 2013) and the same locus on ECA3 has been associated with wither height in Hanoverians (Distl *et al.* 2011), Franches-Montagnes (Signer-Hasler *et al.* 2012), German Warmbloods (Tetens *et al.* 2013) and Thoroughbreds (Boyko *et al.* 2014; Makvandi-Nejad *et al.* 2012). Identification of a causative mutation is complicated by due to long linkage disequilibrium (LD) between markers in the region for some breeds (Boyko *et al.* 2014). Further complications arise from a highly homologous 5 Kb retrogene copy of the 3' end of the *LCORL* gene on another chromosome (Baird, Raudsepp, Brooks 2012; Distl *et al.* 2011)(unpublished data). Yet in the TWH we have observed smaller blocks of LD between the markers in the *LCORL/NCAPG* region (figure 3.7c). Therefore, the TWH would be an ideal breed to use for fine-mapping of the region to identify the causal variant.

Interestingly, initial studies into the function of the *LCORL/NCAPG* gene were not for size or height variation. It was originally identified as a gene, *Mblk-1,* preferentially expressed in the mushroom bodies of the honey bee (Takeuchi *et al.* 2001); the mushroom bodies are important regions for learning, memory, and sensory integration in the insect brain (Davis 1993; Heisenberg 1998). Isolated mouse homologues, *Mlr1* and *Mlr2,* were identified as transcription factors with *Mlr1* predominantly expressed in the spermatocytes of the testis and suggesting a role in spermatogenesis (Kunieda *et al.* 2003).

Quantitative associations for PC2 and PC3 did not identify any loci that surpassed a Bonferroni cutoff. Each association also showed evidence of population stratification, that only increased as fixed covariates such as sex, geographic region, and popular sires were included (data not shown). The population stratification could be due to the high level of inbreeding among the sampled horses. However, while removal of highly inbred individuals (one individual with greater than 18% homozygous markers) did reduce the genomic inflation factor to 1.10866, there was only a

slight detectable change in the quantile-quantile plot. Likely there is not enough power to detect these more subtle conformation pattern PCs.

To improve on our work, we can incorporate several different methodologies in the future. It would be interesting to measure the scapula, humerus, pelvis, and femur to better quantify skeletal size and examine the influence of limb angles on gait ability. The addition of these measures would also allow for a better comparison of leg length across the gaits and could be determined from biomechanical analysis software, such as Quintic software (Centaur Biomechanics, Warwick, UK) or Ariel Performance Analysis System (Ariel Dynamics Inc., Trabuco Canyon, CA). It would also be beneficial to integrate more specific gait type groupings (i.e. running walk vs rack vs foxtrot) to truly examine how body size and shape influence diverse gaits. These determinations would require a more objective measure of gait type, perhaps through the use of accelerometers.

Whether by chance, or design, gaited horses have developed a unique set of conformational traits that may influence gait performance. For example, a longer humerus would extend the stride length, increasing the reach in the front limbs, and would be advantageous in performing a running walk over a rack or pace. Additional data, particularly digital analysis of horses in motion, may further help shed light on how these individual components influence gait type and quality.

With the identification of genes responsible for conformation, especially genes responsible for gait-specific conformation several advancements can be made in basic science and to the horse industry. Genetic tests for these loci will aid breeders in their selection of bloodstock and breeding plans, improving the marketability of their horses. Beyond the impact to horse owners, understanding equine conformation can aid in disease research for both humans and equines. Animal models have long played an important role in human disease research due to similarities in basic biology, physiology and experimental convenience. By understanding the different genetic interactions influencing conformation and gait, new insights can be provided on orthopedic disease

susceptibility and athletic performance. This will allow for the identification of genetic predispositions toward discipline-specific injury in both humans and horses.

References

Al Abri M, Kalla SE, Sutter NB, Brooks SA. Genomic polymorphism in six diverse horse breeds. Manuscript in Progress. .

Albertsdóttir E, Eriksson S, Näsholm A, Strandberg E, Árnason T. 2008. Genetic correlations between competition traits and traits scored at breeding field-tests in icelandic horses. Livestock Science 114(2-3):181-7.

Anderson TM, McIlwraith CW, Douay P. 2004. The role of conformation in musculoskeletal problems in the racing thoroughbred. Equine Veterinary Journal 36:571-5.

Aranason T. 1984. Genetic studies on conformation and performance of icelandic toelter horses. Acta Agric Scand. 34:409-27.

[Internet]; c2006 [cited 2012 . Available from: http://www.abccmm.org.br/regulamentos/regulamentos_1.php?regulamento=57 .

Baird TL, Raudsepp T, Brooks SA. 2012. Copy number variation of a novel retrogene for *LCORL* in the horse. .

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-5.

Bennett D. 2012. Principles of conformation analysis volumes I, II & III. Boulder, Colorado: Equine Network.

Bergsten G. 1980. The durability of the swedish standardbred riding horse judged from a material of insured horses. 32nd annual meeting of the european association of animal production.Zagreb, Yugoslavia: .

Biedermann G and Schmucker F. 1989. Body measurements of thoroughbreds and their relationship with racing performance.(in german with english summary). Zuchtungskunde 61:181-9.

Boyko AR, Brooks SA, Behan-Braman A, Castelhano M, Corey E, Oliveira KC, Swinburne JE, Todhunter RJ, Zhang Z, Ainsworth DM, *et al.* 2014. Genomic analysis establishes correlation between growth and laryngeal neuropathy in thorroughbreds. BMC Genomics 15:259.

Brooks SA, Makvandi-Nejad S, Chu E, Allen JJ, Streeter C, Gu E, McCleery B, Murphy BA, Bellone RR, Sutter NB. 2010. Morphological variation in the horse: Defining complex traits of body size and shape. Animal Genetics 41(s2):159-65.

Clayton HM. 2004. The dynamic horse: A biomechanical guide to equine movement and performance. First edition ed. Mason, MI: Sport Horse Publications.

Cook D, Gallagher PC, Bailey E. 2010. Genetics of swayback in american saddlebred horses. Animal Genetics 41:64-71.

Davis RL. 1993. Mushroom bodies and drosophilia learning. Neuron 11:1-14.

Distl O., Schröder W., Dierks C. and Klostermann A. 2011. Genome-wide association studies for performance and conformation traits in hanoverian warmblood horses.
. 9th dorothy russel havemeyer foundation, international equine genome mapping workshopOak Ridge Conference Center, Chaska, Minnesota: . 12 p.

Dolvik NI and Klemetsdal G. 1999. Conformational traits of norwegian cold-blooded trotters: Heritability and the relationship with performance. Acta Agriculturae Scandinavica, Section A, Animal Science 49:156-62.

Etherton TD. 2009. ASAS centennial paper: Animal growth and development research: Historical perspectives1. Journal of Animal Science 87(9):3060-4.

Gauch HG. 1982. Noise reduction by eigenvector ordinations. Ecology 63(6):1643-9.

Harris SE. 1993. Horse gaits, balance, and movement. New York, NY: Howell Book House.

Heisenberg M. 1998. What do the mushroom bodies do for the insect brain? an introduction. Learning and Memory 5:1-10.

Holmström M and Philipsson J. 1993. Relationship between conformation, performance and health in 4-year old swedish warmblood rding horses. Livestock Production Science 33:293-312.

Holmström M, Fredricson I, Drevemo S. 1994. Biokinematic differences between riding horses judged as good and poort at the trot. Equine Veterinary Journal 17 (Suppl):51-6.

Jeffcott LB, Rossdale PD, Freestone J, Frank CJ, Towers-Clark PF. 1982. An assessment of wastage in thoroughbred racing from conception to 4 years of age. Equine Veterinary Journal 14:185-98.

Karim L, Takeda H, Lin L, Druet T, Arias JAC, Baurain D, Cambisano N, Davis SR, Farnir F, Grisart B, *et al.* 2011. Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. Nat Genet 43(5):405-13.

Kunieda T, Park J, Takeuchi H, Kubo T. 2003. Identification and characterization of Mlr1,2: Two mouse homologues of mblk-1, a transcription factor from the honeybee brain. FEBS Lett 535(1-3):61-5.

Lane G. 2011. Discussion with author in september. .

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, *et al.* 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467(7317):832-8.

Linder A and Dingerkus A. 1993. Incidence of training failure among thoroughbred horses at coligne, germany. Preventative Veterinary Medicine 16:85-94.

Lindholm-Perry AK, Kuehn LA, Oliver WT, Sexten AK, Miles JR, Rempel LA, Cushman RA, Freetly HC. 2013. Adipose and muscle tissue gene expression of two genes (NCAPG and LCORL) located in a chromosomal region associated with cattle feed intake and gain. PLoS One 8(11):e80882.

London J. 2012. Discussion with author in october. .

Magnusson L and Thafvelin B. 1990. Studies on the conformation and related traits of standardbred trotters in sweden. J Anim Breed Genet. 107:135-48.

Makvandi-Nejad S, Hoffman GE, Allen JJ, Chu E, Gu E, Chandler AM, Loredo AI, Bellone RR, Mezey JG, Brooks SA, *et al.* 2012. Four loci explain 83% of size variation in the horse. PLos One 7(7):e39929.

McGreevy PD and Thomson PC. 2006. Differences in motor laterality between breeds of performance horse. Applied Animal Behaviour Science 99:183-90.

McIlwraith CW, Anderson TM, Sanschi EM. 2003. Conformation and musculoskeletal problems in the racehorse. Clinical Techniques in Equine Practices. 2(4):339-47.

Metzger J, Schrimpf R, Philipp U, Distl O. 2013. Expression levels of LCORL are associated with body size in horses. PLos One 8(2):e56497.

Nicodemus MC and Clayton HM. 2003. Temporal variables of four-beat, stepping gaits of gaited horses. Applied Animal Behaviour Science 80:133-42.

Nicodemus MC, Holt KM, Swartz K. 2002. Relationship between velocity and temporal variables of the flat shod running walk. Equine Veterinary Journal Supplemental 34:340-3.

O'Rourke N and Hatcher L. 2013. Introduction: The basics of principle component analysis. In: A step-by-step approach to using SAS for factor analysis and structural equation modeling. 2nd ed. Cary, North Carolina: SAS Institute, Inc. 1 p.

Pausch H, Flisikowski K, Jung S, Emmerling R, Edel C, Gotz KU, Fries R. 2011. Genome-wide association study identifies two major loci affecting calving ease and growth-related traits in cattle. Genetics 187(1):289-97.

Peres-Neto PR, Jackson DA, Somers KM. 2005. How many principal components? stopping rules for determining the number of non-trivial axes revisited. Comput Stat Data Anal 49(4):974-97.

Philipsson J, Brendow E, Dalin G, Wallin L. 1998. Genetic aspects of disease and lesions in horses. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production 24:408-15.

Pryce JE, Hayes BJ, Bolormaa S, Goddard ME. 2011. Polymorphic regions affecting human height also control stature in cattle. Genetics 187(3):981-4.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, *et al.* 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 81:559-75.

Rozen S and Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics methods and protocols: Methods in molecular biology. Krawetz S and Misener S, editors. Totowa, NJ: Humana Press. 365 p.

Sadek MH, Al-Aboud AZ, Ashmawy AA. 2006. Factor analysis of body measurements in arabian horses. J Anim Breed Genet 123(6):369-77.

Setoguchi K, Furuta M, Hirano T, Nagao T, Watanabe T, Sugimoto Y, Takasuga A. 2009. Cross-breed comparisons identified a critical 591-kb region for bovine carcass weight QTL (CW-2) on chromosome 6 and the ile-442-met substitution in NCAPG as a positional candidate. BMC Genet 10:43.

Setoguchi K, Watanabe T, Weikard R, Albrecht E, Kuhn C, Kinoshita A, Sugimoto Y, Takasuga A. 2011. The SNP c.1326T>G in the non-SMC condensin I complex, subunit G (NCAPG) gene encoding a p.Ile442Met variant is associated with an increase in body frame size at puberty in cattle. Anim Genet 42(6):650-5.

Signer-Hasler H, Flury C, Haase B, Burger D, Simianer H, Leeb T, Reider S. 2012. A genome-wide association study reveals loci influencing height and other conformation traits in horses. PLos One 7:e37282.

Smith AM, Staniar WB, Splan RK. 2006. Associations between yearling body measurements and career racing performance in thoroughbred racehorses. Journal of Equine Veterinary Science 26(5):212-4.

Soranzo N, Rivadeneira F, Chinappen-Horsley U, Malkina I, Richards JB, Hammond N, Stolk L, Nica A, Inouye M, Hofman A, *et al.* 2009. Meta-analysis of genome-wide scans for human adult stature identifies novel loci and associations with measures of skeletal frame size. PLoS Genet 5(4):e1000445.

Splan RK and Hunter HB. 2004. Temporal variables of the canter of the tennessee walking horse. Equine and Comparative Exercise Physiology 1(1):41-4.

Stashak TS. 1987. The relationship between conformation and lameness. In: Adam's lameness in horses. Stashak TS, editor. Philadelphia, PA: Lea & Febiger. 71 p.

Takeuchi H, Kage E, Sawata M, Kamikouchi A, Ohashi K, Ohara M, Fujiyuki T, Kunieda T, Sekimizu K, Natori S, *et al.* 2001. Identification of a novel gene, mblk-1, that encodes a putative transcription factor expressed preferentially in the large-type kenyon cells of the honeybee brain. Insect Mol Biol 10(5):487-94.

The Tennessee Walking Horse Breed: Gaits [Internet]; c2011 [cited 2014 . Available from: www.twhbea.com/breed/gait.php .

Tetens J, Widmann P, Kuhn C, Thaller G. 2013. A genome-wide association study indicates LCORL/NCAPG as a candidate locus for withers height in german warmblood horses. Animal Genetics 44(4):467-71.

Visscher PM, McEvoy B, Yang J. 2010. From galton to GWAS: Quantitative genetics of human height. Genet Res (Camb) 92(5-6):371-9.

Weeren PR and Crevier-Denoix N. 2006. Equine conformation: Clues to performance and soundness? Equine Vet J 38(7):591-6.

Womack B. 1994. The echo of hoofbeats: A history of the tennessee walking horse. Third ed. Shelbyville, Tennessee: DABORA, INC.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, *et al.* 2010. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42(7):565-9.

Ziegler L. 2005. Easy-gaited horses. First edition ed. North Adams, MA: Storey Publishing.

CHAPTER 4

HERITABLE TEMPERAMENT VARIATION IN GAITED BREEDS OF HORSES

Introduction

Temperament is a key criterion in the selection of horses to ensure optimal performance and safety for both leisure and competitive riding. Throughout domestication, horses have been selected for desirable traits resulting in the development of several different breeds. Gaited breeds are renowned for their typically easy-going temperaments (Ziegler 2005), but there is variation across and within the breeds. For example, the Tennessee Walking Horse (TWH) is described as a calm, docile breed and is often used as a trail, show, and pleasure horse. However, among horse owners and caretakers, there are anecdotes supporting familial trends in behavior and personality.

Prior work in our lab comparing gaited to non-gaited horses identified a potential candidate region associated with behavior and startle response (Staiger *et al.* 2011). To further investigate this, we adapted a 20 item questionnaire from Momozawa *et al.* (2003) to be included with data collection for another ongoing project. The efficacy of questionnaire-based behavioral assessment has already been demonstrated in a variety of species, including horses (Anderson *et al.* 1999; French 1993; Momozawa *et al.* 2003; Momozawa *et al.* 2005; Morris, Gale, Duffy 2002; Pervin and John 1997). Surprise and startle tests are often used in behavioral assessments to measure a phenomenon known as neophobia. Neophobia is an emotional state induced by the perception of actual danger, and can be defined as a predisposition to react in a similar manner to various fear-provoking events (Lasande, 2008). Surprise and startle tests utilize a novel object that elicits fear due to its novelty or movement which in turn triggers a startle response due to the innate startle reflex (Valls-Sole 2012). The startle reflex is a fast response to a sudden, intense stimulus, such as a loud sound, and consists of contraction of skeletal musculature (Quednow *et al.* 2006). The startle reflex shows several forms of behavioral plasticity, such as prepulse inhibition (PPI) and habituation (Quednow *et al.* 2006). PPI refers to the reduction of the magnitude of the acoustic startle response (ASR) when a distinctive, non-startling stimulus is presented 30-500 ms before the startling

stimulus. PPI is used as an operational measure for sensorimotor gating that reflects the ability of an organism to properly inhibit sensory information (Graham 1975; Hoffman and Ison 1980; Quednow *et al.* 2006). Habituation is a theoretical construct that refers to the reduction in the magnitude of the ASR after repeated presentation of the startling stimulus that is not due to muscle fatigue or blunting of sensory receptor responsiveness (Groves and Thompson 1970; Quednow *et al.* 2006; Siddle and Kroese 1985).

Several investigations have reported changes of habituation and/or PPI of ASR in neuropsychiatric disorders such as schizophrenia (Braff, Grillion, Geyer 1992; Geyer and Braff 1982; Parwani *et al.* 2000), schizotypal personality disorder (Cadenhead, Geyer, Braff 1993), obsessive-compulsive disorder (Swerdlow, 1993), and Huntington's disease (Swerdlow *et al.* 1995). Changes in PPI and habituation may provide trait markers for psychiatric disorders with altered neurotransmitter regulation (Cadenhead *et al.* 1999; Quednow *et al.* 2006). Horses suffer from several different stereotypic behaviors, including wind-sucking and cribbing (Overall 1998). To date, no PPI testing has been performed in horses. However, habituation is a commonly applied technique used to attenuate startle response in riding horses.

Studies on the genetics of temperament and behavior have already been conducted on mice, cattle, and humans. In mice, studies frequently utilize inbred lines to detect quantitative trait loci (QTL) (Flint, 2003; Willis-Owen & Flint, 2006 – Gutierrez refs), candidate genes (Yalcin, 2004), and improve the overall understanding of behavior and fear (Koch, 1997). Cattle experimental designs typically involve identifying QTL for response in behavior trials (Glenske *et al.* 2011; Gutierrez-Gil *et al.* 2008) or estimated breeding values (Haskell, Simm, Turner 2014) and have been successful in detecting candidate loci. Human studies have been less successful, utilizing different temperament scales across large sample sizes (de Moor *et al.* 2012; Munafò *et al.* 2009; Service *et al.* 2012; Terracciano *et al.* 2010; Verweij *et al.* 2010). One study was able to detect loci at genome-wide

significance levels (de Moor *et al.* 2012), but the results could not be replicated (Service *et al.* 2012). Another study did not reach genome-wide significance, but the top signals were found in genes that are believed to influence behavioral traits and mental disorders (Terracciano *et al.* 2010).

In the present study, we have utilized a behavior questionnaire to identify temperament components shared across breeds and with breed unique components. We have also mapped several breed-specific temperament components determined from the questionnaire. Within a small subset, we have conducted a preliminary behavior trial that identified auditory tests that could prove useful in testing PPI in horses.

Materials & Methods

Animals Used – Across Breed Temperament Survey

We used a total of 561 horses for this study from 20 gaited breeds, collected at private farms and horse shows from North and South America, from September 2011 through August 2014. These breeds include four different cross-breed groups as well as mules, and donkeys. The group consisted of 293 mares, 99 stallions, and 169 geldings, ranging in age from 1 to 32 years. Amount of training, horse's use, type of shoes, gait type, coat color and pattern, and pedigree information were also collected, as described in Chapter 2. 311 horses were classified as lateral-only and 244 as multi-gaited. Horses were additionally classified on whether they were able to trot or not; 435 horses were classified as non-trotters and 120 were classified as trotters. 6 horses were unclassified in both categories due to lameness at the time of collection.

Temperament Questionnaire

A survey for horse caretakers was adapted from Momozawa *et al.* (2005) and consisted of 20 questions (Table 4.1) with responses based on a scale of 1 to 9. The survey included information on the number of observations per week and the relationship of the rater to the horse (owner,

trainer, barn manager, etc.). The survey was completed by a caretaker who was well acquainted with the respective horses (observed the horse daily for at least a year). The rater's experience level was subjectively interpreted based on the relationship reported. Trainers and barn managers were rated as the most experienced since their livelihood is dependent upon horses. Owners were rated as the least experienced since some of these included owners who board their horses and are therefore not necessarily responsible for the horse's daily care, such as feeding, or have the experience with the number of horses commensurate to industry professional. Owners who reported themselves as also breeders and trainers of the horses were rated as moderately experienced.

Table 4.1. Survey Questions.

Item	Question/Statement	Score 1	Score 9
Nervousness	Tends to become nervous or agitated by objects, noises, etc.	Nervous	Calm
Concentration	Tends to be focused and unaffected by the environment.	Excellent	Poor
Self-reliance	Tends to be at ease if left alone away from the herd.	At ease	Restless
Trainability	Tends to be trained easily and promptly	Excellent	Poor
Excitability	Tends to get excited easily or responds too quickly to new stimuli.	Excitable	Not excitable
Friendly-People	Tends to rarely become aggressive or fearful of people.	Friendly	Unfriendly
Curiosity	Tends to be interested in novel objects and approaches them.	Frequently	Rarely
Memory	Tends to memorize/remember what it learned or was trained.	Excellent	Poor
Panic	Tends to get excited/fearful to an abnormal extent.	Frequently	Rarely
Cooperation	Tends to be cooperative/willing to work with a caretaker when handled.	Always	Never
Inconsistent Emotionally	Tends to be unpredictable from day to day.	Inconsistent	Consistent
Stubbornness	Tends to be obstinate once it resists a command.	Stubborn	Obedient
Docility	Tends to be docile in general.	Docile	Active
Vigilance	Tends to be vigilant/watchful about surroundings.	Always	Never
Perseverance	Tends to be patient with various stimuli	Patient	Impatient
Friendly-Horses	Tends to interact with horses in a friendly manner	Friendly	Unfriendly
Competitiveness	Tends to be dominant in antagonistic encounters with other horses.	Dominant	Subordinate
Skittishness	Tends to get surprised easily.	Skittish	Not skittish
Timidity	Tends to be timid/lacks courage in novel environments.	Timid	Audacious
Trailer Loading	Tends to load easily into a trailer.	Always	Rarely

Quality Control and Statistical analysis

21 horses were excluded due to incomplete surveys (missing or ambiguous response to one or more questions). 85 horses with mixed/unknown breed origin or related within one generation to another horse in the data set were excluded. Principle component analysis (PCA) was performed on the scores for the 20 question items for the remaining 455 horses. The covariance matrix was applied as all questions used the same scale (whole numbers of 1 to 9). Principle Component (PC) scores were retained for interpretation and analysis if the eigenvalue is greater than 1.0, if percent variance explained is greater than $1/n$ where n is the number of body measurements we collected ($1/n = 5.00\%$), and by examining the scree plot for the point where the eigenvalue slope plateaus. Additionally, a factor method, prior communality, and Varimax rotation for orthogonal transformation, was used to correct for skewing in the responses (human nature not to use the scores 1-9 in a normally distributed manner). In the remaining 455 horses, three horses had missing gait information and 15 horses had missing age information; these horses were kept in the dataset for PCA and factor analysis, but were excluded in additional analysis. Analysis of variance (ANOVA) was used to evaluate differences of the extracted factor mean scores within sex (male or female), gender (mare, gelding, or stallion), breed, discipline (trail, show, or breeding), training level (trained to ride, not trained to ride) and gait type (able to trot or not able to trot, lateral-only or multi-gaited). Correlation coefficients or Spearman's rank correlation coefficients were calculated to determine whether or not any relationship existed between the results of the survey with age, the number of observations, and rater experience level. To compare extracted factors across an individual breed, we conducted breed specific PCA and FA on the 20 question items for a subset of TWH that passed quality control filters (n=211). All statistics were calculated and visualized using JMP Pro 11 (SAS Institute, Inc., Cary, NC).

Startle Trial - Animals

Experiments were carried out from March 19 through March 22, 2012 at two privately-owned farms in middle Tennessee. Farm 1 only stables young horses for showing in conformation classes. Farm 2 is a breeding operation that houses breeding stallions, broodmares and foals, and young riding horses. Sixteen TWH horses were tested in total, 9 horses at farm 1, and 7 horses at farm 2. The group consisted of 10 stallions and 6 mares, ranging in age from 10 to 22 months (average age 15 months). All of the horses at farm 1 also have a completed behavior questionnaire, while only one horse at farm 2 has a completed questionnaire. We conducted testing at farm one over two consecutive days.

At farm 1, the trial was conducted in a covered pen (3.7m x 7.3m) with metal gate panels, a shared wall with two stalls (lower half wood, upper half diamond mesh), and a dirt floor. Blue painters tape makers were placed on the wooden wall spaced at 2 ft intervals. The camera was placed on a tripod outside the pen along the 7.3m panel side. A 0.6m x 1.2m wooden board was propped inside the pen, against the 3.7m panel opposite of the opening 3.7m panel to obscure the tester from the horses' plane of vision. All stimuli were presented next to the board.

At farm 2, the trail was conducted in an indoor stall (3.7m x 7.3m) with three full wooden walls (ground to ceiling), an open metal bar stall door, a half stone wall with metal bars to the ceiling, and a dirt floor covered with sawdust. Blue painters tape makers were placed on the wooden wall spaced at 2-foot intervals. The camera was placed on a tripod outside of the stall by the half stone wall. All stimuli were presented by the stall door.

Test stimuli included a pressure algometer (Wagner Force Ten FDX) for a tactile test, a red automatic-open 42" umbrella (Wal-mart brand) for a visual test, and 3" wide Velcro strips with one strip adhered to a 12" long x 4" wide foam board for auditory tests.

Before the test, each subject at Farm 1 was placed in cross-ties and equipped with a HR monitor. We continuously recorded the heart rate (HR) before and during the trials using a Polar S625X HR monitor (Polar Electro Ltd.). The horse was kept in cross-ties quietly for 3 min to obtain a baseline heart rate level. The first test consisted of a tactile pressure test until the horse reacted (i.e. skin twitch or other movement). We measured tactile pressure on the scapula just above the shoulder joint with a Force Ten FDX force gage/algometer (Wager Instruments) while the horse remained in the cross-ties. The initial reactivity was observed and scored on a three grade scale (1=low; 2=moderate; 3=high). Horses at Farm 2 were not equipped with a HR monitor or tested with the algometer due to lack of training and violent defensive reactions that threatened handler and experimenter safety.

The test horse was kept in cross-ties quietly for 1-2 min before it was led by a familiar caretaker into the pen for the remaining tests. The handler removed the lead line, left the pen and closed the gate/door. The subject was allowed 10 min to acclimate and explore the pen. After the initial acclimation period, the tests commenced when the horse faced the stimulus presentation area.

The first series of tests involved an auditory challenge. The auditory stimulus tested was the sound of separating Velcro strips. After the tester separated the Velcro, the reaction level and distance of movement were measured.

After the auditory challenge, the horse was presented with a novel visual stimulus. Five minutes after the auditory stimulus, the tester placed a closed umbrella inside the pen. The intentional touching of the closed umbrella and latency to the intentional touching of the umbrella were measured. The test ended either when the horse touched the umbrella or after five minutes had passed since presenting the umbrella.

The last test was a startle test. When the horse faced the stimulus presenting area, the tester opened the umbrella outside the pen (Farm 1) or just inside the stall door (Farm 2). We measured the distance of movement, intentional touching of the umbrella, and latency to the intentional touching of the umbrella. The test ended either when the horse approached and touched the umbrella or after five minutes had elapsed since opening the umbrella.

Investigation of a candidate gene for startle response

We genotyped all 16 horses for a five-SNP haplotype in the *NXPH2* gene, identified as a potential target for gait type and behavior differences in horses (Staiger *et al.* 2011). We extracted DNA from hair using the Gentra® Puregene® DNA Isolation Kit, following the manufacturer's protocol with modifications to optimize for hair root bulbs (Qiagen Inc., Valencia, CA) (Cook, Gallagher, Bailey 2010). We used Primer3 (Rozen and Skaletsky 2000) to design primer sequences around the 3'-untranslated region to produce a PCR product of 1664bp and performed PCR amplification in a 20μL volume containing 2μL of DNA (diluted to a concentration of 25ng/μL), 2μL of 10X PCR reaction buffer with 20mM $MgCl_2$, 0.2μL of FastStart Taq DNA Polymerase (Roche Diagnostics), 2μL of 2mM dNTP's, 2μL each of forward and reverse 5μM primers, and 9.8μL PCR-grade water. We carried out the PCR on an Eppendorf Mastercycler Ep Gradient (Eppendorf Corp.) under the following conditions: 95°C for 4 min, followed by 40 cycles of 95°C for 30 sec, 62°C for 30 sec, 72°C for 2 min, and a final extension of 72°C for 7 min and cooling to 4°C. PCR products were sequenced at the Cornell University Core Laboratories Center (Ithaca, NY) on an ABI 3730xl capillary electrophoresis unit with the forward primer. Sequence reads were aligned and analyzed using CodonCode Aligner software (CodonCode Corporation, Dedham, MA).

Behavior Trial Statistical analysis

Correlation coefficients or Spearman's rank correlations coefficients were calculated to determine relationships between results of the behavior trial and with the results of the behavior

questionnaire. Analysis of variance (ANOVA) or contingency analysis likelihood ratio tests were conducted to identify differences in the behavior trial results due to sex (male or female), base coat color, farm and day of test (for farm one). Correlations and ANOVA were calculated and visualized using JMP Pro 11 (SAS Institute, Inc., Cary, NC).

Genotyping and Quality Control – Behavior Traits

Genomic DNA was isolated from blood samples using the Gentra® Puregene® Blood Kit, following the manufacturer's protocol for whole blood (Qiagen Inc., Valencia, CA). Extraction of DNA from hair was performed as described above.

116 TWH with complete phenotypes as described above were genotyped at 65,000 loci using the Equine SNP70K beadchip (Illumina Inc, San Diego, CA) at GeneSeek Inc. (Lincoln, NE). We excluded SNPs with a genotyping rate <95% (n=14,709) and MAF<0.05 (n=14,820) across all individuals. Three individuals with a genotyping rate <90% across the remaining 38,541 SNPs were excluded. Remaining individuals had a 99.2% genotyping rate. We used a Bonferroni significance cutoff of 1.42×10^{-6}, conservatively estimating 35,173 independent comparisons (subtracted 3,368 markers that are in complete linkage disequilibrium (LD), $r^2>0.99$). Genome-wide IBD estimates were calculated using an IBS similarity matrix to evaluate population structure; all individuals had IBS distances less than 0.74 and thus were retained.

Genome-wide Association Studies

We ran three genome-wide tests for associations with quantitative factor traits 1, 2, and 3 unique to TWH. We applied basic quantitative association, linear associations with covariates of age, rater experience, and gender, and a in A Mixed Model linear analysis (EMMAX) (Kang *et al.* 2010) with and without the covariates of age, rater experience, and gender under additive, dominant, and recessive models using Golden Helix SVS software (Golden Helix, Inc., Bozeman, MT). We utilized quantile-quantile plots and genomic inflation factors to determine the model of best fir for

each factor. We examined the LD structure between markers using Haploview v4.2 (Barrett *et al.* 2005). For factor 1, we applied adaptive permutation to a linear model with age and rater experience level added as covariates under a dominant mode of inheritance in PLINK (Purcell *et al.* 2007). In the data transfer from SVS to PLINK, PLINK identified 1780 SNPs from the X chromosome as heterozygous haploid genotypes and set them to missing, resulting in the exclusion of the X chromosome in the permutation testing for Factor 1. For factor 2, we used a recessive model test with no covariates. We applied a dominant linear model with gender added as a covariate for Factor 3.

Results

The overall goal of this study was to determine if different gait types could be predicted by differences in temperament and behavior responses. To test this, we utilized a temperament questionnaire from several gaited breeds, and applied a behavior trial within a single breed to ensure accuracy of the subjective questionnaire. To identify common genetic inputs/pathways, we also performed a genome-wide scan of three temperament factors.

Four Factors Explain 59.3% of Temperament Variation across Gaited Breeds

We conducted principle component analysis (PCA) and factor analysis (FA) to explore the degree of temperament variation in gaited breeds of horses. PC1, PC2, PC3, and PC4, cumulatively accounting for 59.3% of the total variance (Figure 4.1a), were retained for FA with orthogonal transformation due to the skewed distributions of the question responses (data not shown). FA of the gaited breeds identified two separate but highly similar groups (tractable and anxious). Factor 1 accounted for 22.9% of the total variance (Figure 4.1b) and loads strongly on the positive side for excitability, skittishness, nervousness, and panic, while perseverance, concentration and docility load negatively (Figure 4.1c). Therefore, horses with higher Factor 1

scores are less excitable, skittish, nervous, and panicky, and more patient, docile and attentive, while horses with negative Factor 1 scores are the opposite. Based on this, we labeled Factor 1 as 'Neophobia'. Factor 2 (14.6%) separates emotionally inconsistent and skittish horses from curious and attentive horses (Figure 4.1c), which we labeled as 'Trainability'. Positive Factor 2 scores are horses that are more emotionally inconsistent and skittish and less curious and attentive, while horses with a negative score are the opposite. Factor 3 (12.4%) loaded strongly and positively on self-reliance and concentration (Figure 4.1c), so we labeled it as 'Independence' as horses with a positive score are less self-reliant and attentive, while horses with a negative score are more self-reliant. Factor 4 (9.3%) loads strongly on competition and friendly to horses, in opposite directions (Figure 4.1c) and seems to describe social interaction, so we labeled it as 'Hostility'. Horses with a positive Factor 4 score are less competitive and friendlier to horses than horses with a negative score.

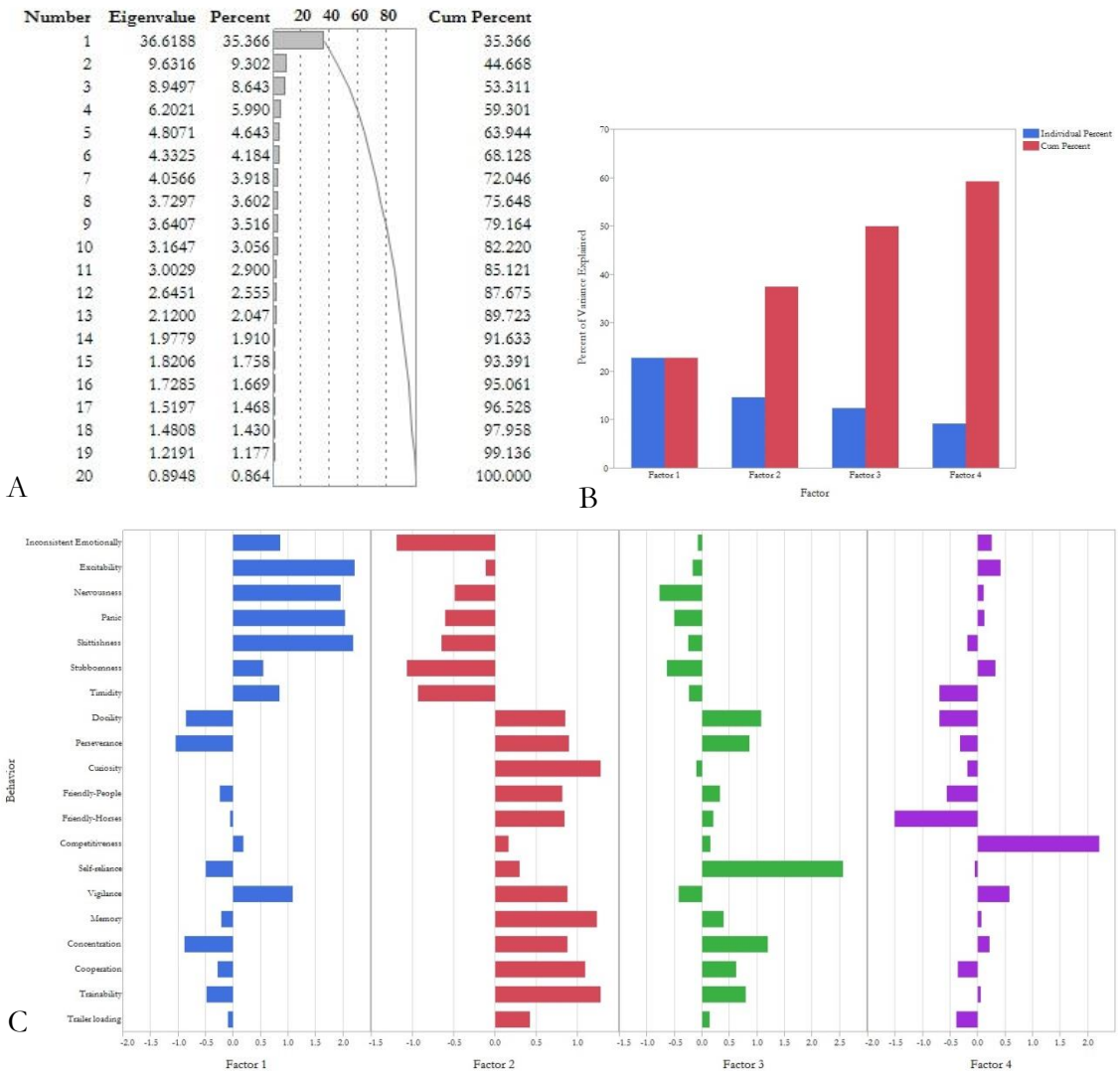


Figure 4.1. Temperament traits determined from across-breed behavior surveys. A) PCA identified four components that cumulatively account for ~60% of the behavior variance. Bars indicate individual component percentage, with the line indicating cumulative percentage. B) FA percent variance of four traits account for ~60% of the variance. Blue bars are individual factor percentages, red bars are cumulative percentages. C) Plot of factor loading for each individual behavior trait across the four identified factors.

Temperament Factors Across Breeds are More Predictive of Discipline Than Gait Ability

Factor 1 (neophobia) accounted for the greatest proportion of variance in behavior across breeds following varimax rotation. However, Factor 1 scores appear to be heavily influenced by environment as do the remaining factors 2-4. There is a positive correlation between age and Factor 1 scores, indicating horses become less neophobic as they age, yet are not different due to sex or gender (Table 4.2). Additionally, horses trained to ride under saddle and used for trail riding had higher mean factor 1 scores than those untrained or not used for trail (Table 4.2). There was no difference in mean factor 1-3 scores across different gait type categories. Rater experience was negatively correlated with Factor 1 and 3 scores, indicating horses with negative Factor 1 scores were rated by more experienced observers (Table 4.2). The only significant difference in Factor 2 scores were that show horse had lower mean scores (Table 4.2), indicating that show horses are more "trainable" than non-show horses. Stallions had lower Factor 3 mean scores than mares and geldings, and show horses have lower mean scores than non-show horses (Table 4.2), suggesting stallions and show horses are more "independent". However, the differences in Factor 2 and 3 scores may be an artifact of management and training. Geldings have higher Factor 4 scores than mares and stallions (Table 4.2) implying they are less "hostile" to other horses. Interestingly, lateral-only gaited horses had higher mean Factor 4 scores than multi-gaited horses (Table 4.2) also indicating they are less "hostile" than multi-gaited horses. Across Factors 1-3, mean scores were different based on breed. However, there was only one breed (TWH) with over 100 samples, the rest ranged from 71 to five individuals.

Table 4.2. Across-breed ANOVA p-values for comparison of mean factor scores across traits of interest.

Trait of Interest	Factor 1	Factor 2	Factor 3	Factor 4
Age Corr	0.1949	-0.0270	0.0068	-0.1162
Age P-value	<0.0001**	0.5660	0.8848	0.0134*
Experience Corr	-0.1986	0.0721	-0.1703	0.0026
Experience p-value	<0.0001**	0.1254	0.0003*	0.9561
Observations Corr	-0.0013	0.0095	0.0367	-0.0537
Observation p-value	0.9788	0.8405	0.4366	0.2547
Breed	0.0215*	0.0008**	0.0001**	0.1249
Sex	0.4269	0.6277	0.2419	0.2431
Gender	0.3099	0.8224	0.0030**	0.0113**
Gait type	0.8701	0.1059	0.3431	0.0095**
Trot ability	0.0576	0.1191	0.9551	0.7806
Training level/started to ride	0.0319*	0.9638	0.1332	0.8688
Trail use	0.0002**	0.5565	0.2118	0.0488*
Show use	0.0272*	0.0007**	<0.0001**	0.6425
Breeding use	0.3163	0.4628	0.4701	0.0154*

*P-value significantly different across mean scores at $\alpha<0.05$. **P-values significantly different across mean scores at $\alpha<0.0125$.

Four Factors Explain 64% of Temperament Variation, but Do Not Correlate with Gait Ability, in TWH

We conducted principle component analysis (PCA) and factor analysis (FA) to explore the degree of temperament variation within the TWH. PCs1-4 cumulatively account for 64% of the total variance, were retained for factor analysis with orthogonal transformation (Figure 4.2a).

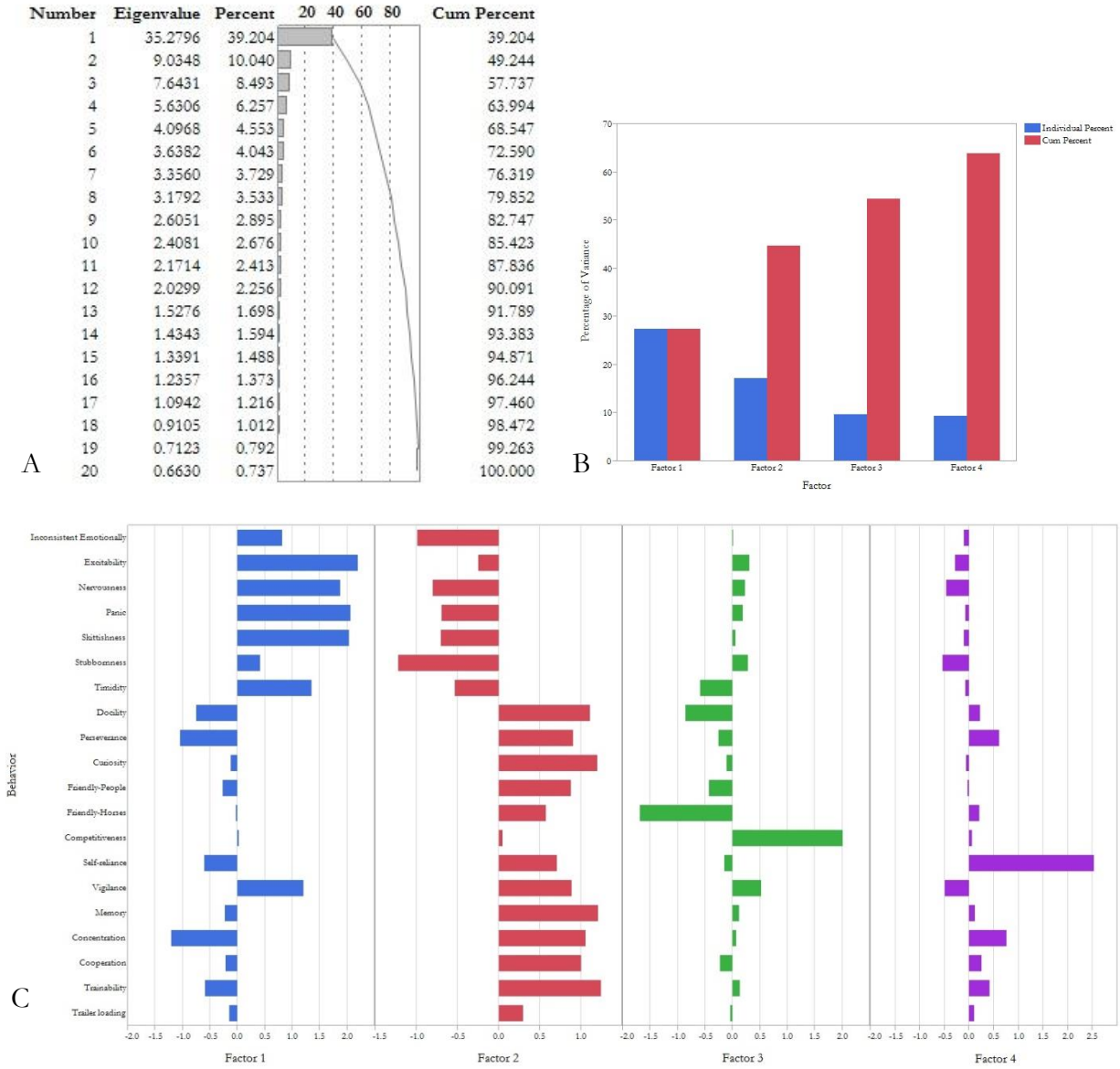
Factor 1 accounted for 27.6% of the total variance and loads strongly on the positive side for excitability, skittishness, nervousness, and panic, while perseverance, concentration and docility load negatively (Figure 4.2c). Therefore, horses with higher Factor 1 scores are less excitable, skittish, nervous, and panicky, and more patient, docile and attentive, while horses with negative Factor 1 scores are the opposite. Based on this, we labeled Factor 1 as 'Neophobia'. Factor 1 was positively correlated with age (Table 4.3), indicating horses are rated as less neophobic as they age.

Factor 2 (17.2%) separates emotionally inconsistent and skittish horses from curious and attentive horses (Figure 4.2c), which we labeled as 'Trainability'. Positive Factor 2 scores are horses that are more emotionally inconsistent and skittish and less curious and attentive, while horses with a negative score are the opposite. Factor 2 scores were not significantly different for any of our additional traits of interest (Table 4.3).

Factor 3 (9.8%) loaded strongly on competition and friendly to horses, in opposite directions (Figure 4.2c) and seems to describe social interaction, so we labeled it as 'Hostility'. Horses with a positive Factor 3 score are less competitive and friendlier to horses than horses with a negative score. Stallions have the lowest mean Factor 3 scores, while geldings have the highest (Table 4.3), indicating stallions are less 'hostile' than mares and geldings.

Factor 4 (9.4%) loads strongly and positively on self-reliance and concentration (Figure 4.2), so we labeled it as 'Independence' as horses with a positive score are less self-reliant and attentive, while horses with a negative score are more self-reliant. Factor 4 was not significantly different for

any of our traits of interest except show use; show horses have lower Factor 4 scores (Table 4.3), indicating show horses are more 'independent'.



Number	Eigenvalue	Percent	Cum Percent
1	35.2796	39.204	39.204
2	9.0348	10.040	49.244
3	7.6431	8.493	57.737
4	5.6306	6.257	63.994
5	4.0968	4.553	68.547
6	3.6382	4.043	72.590
7	3.3560	3.729	76.319
8	3.1792	3.533	79.852
9	2.6051	2.895	82.747
10	2.4081	2.676	85.423
11	2.1714	2.413	87.836
12	2.0299	2.256	90.091
13	1.5276	1.698	91.789
14	1.4343	1.594	93.383
15	1.3391	1.488	94.871
16	1.2357	1.373	96.244
17	1.0942	1.216	97.460
18	0.9105	1.012	98.472
19	0.7123	0.792	99.263
20	0.6630	0.737	100.000

Figure 4.2. Temperament traits determined from TWH behavior surveys. A) PCA identified four components that cumulatively account for ~60% of the behavior variance. Bars indicate individual component percentage, with the line indicating cumulative percentage. B) FA percent variance of four traits account for ~60% of the variance. Blue bars are individual factor percentages, red bars are cumulative percentages. C) Plot of factor loading for each individual behavior trait across the four identified factors.

Table 4.3. TWH ANOVA p-values for comparison of factors across traits of interest.

Trait of Interest	Factor 1	Factor 2	Factor 3	Factor 4
Age Corr	0.2221	-0.1277	-0.1292	0.0607
Age P-value	0.0012**	0.0654	0.0623	0.3826
Experience Corr	-0.1243	0.1018	0.045	-0.0429
Experience p-value	0.0683	0.1360	0.5108	0.5307
Observations Corr	-0.0307	0.0672	-0.0723	-0.0485
Observation p-value	0.6534	0.3257	0.2902	0.4786
Sex	0.2365	0.8746	0.5918	0.2485
Gender	0.2105	0.6053	0.0152*	0.0503
Gait type	0.3404	0.9720	0.2026	0.0711
Trot ability	0.3450	0.7951	0.3895	0.1593
Training level/broke to ride	0.1609	0.9875	0.7655	0.6084
Trail use	0.2170	0.3842	0.1631	0.1543
Show use	0.8222	0.1137	0.7197	0.0176*
Breeding use	0.0941	0.1290	0.1365	0.3423

*P-value significantly different across mean scores at $\alpha<0.05$. **P-values significantly different at $\alpha<0.0125$.

Startle Trial

We administered a preliminary four-test startle trial to a subset of 16 TWH yearlings from two different farms with several goals in mind. The first was to validate the accuracy of our subjective temperament questionnaire in assessing individual behavior. The second goal was to assess the feasibility of testing for PPI in the horse, and if so, could we predict gait type from variable PPI responses.

A Larger Sample Size Is Needed To Determine Congruence between the Startle Trial and Questionnaire

Across all of the horses with a completed survey (n=10), only Factor 1 (Neophobia) and Factor 4 (Independence) showed significant correlations with the results from the behavior trail (Table 4.4). Factor 1 scores had a positive correlation with the latency to touch the novel object. Factor 4 scores had a negative correlation with the latency to touch the startle object. This indicates that less neophobic horses were slower to touch a novel object, while less independent horses were faster to touch the open umbrella. We would expect to see a less neophobic horse be faster to touch a novel object as they would be less afraid, therefore while we do have significant correlations, they do not exactly correspond with the questionnaire results and are likely a reflection of the small sample size.

Startle Trial Responses Highlight Important Factors for Future Studies

Across the 16 horses, measuring the HR proved difficult and inconsistent to record in young horses and therefore the method was abandoned. It seems likely that securely anchored recording or telemetry units will have to be developed, in conjunction with more consistent acclimatization in client-owned animals.

After analyzing the startle trial results, several significant differences were identified. There were significant differences between the two farms for the time it took to touch the startle object, whether the horses actually touched the object, the response to the auditory stimulus, and age (Table

4.5), indicating that environment likely plays a strong role in some behavior responses. Horses at Farm One were younger (average 13.5 months) than horses at Farm Two (average 17.3 months). Horses at Farm Two reacted less to the auditory stimulus, touched the startle object more and were faster to approach the startle object than horses at Farm One. However, there was no difference between the two farms in touching the novel stimulus or the distance traveled from the startle test, suggesting environment plays less of a role in flight response. Within Farm One, there was a difference between horses for approaching the novel object between the first and second test day (Table 4.5), likely due to the setup of the first farm resulting in unintentional exposure and acclimation to the startle response for horses tested on the second day. Horses that did touch the novel object were faster to approach the stimulus, were closer to the startle stimulus, and touched the startle object more than horses that did not touch the novel object. Horses that touched the startle object were faster at approaching both the novel and startle object.

Table 4.4. Comparison of startle trial results versus other cofactors.

Trial Measure	Gait	Sex	Farm	Day Tested	Day of Test for Farm1	Touch Stationary Umbrella?	Touch Open Umbrella?	A1 geno	A2 geno	A3 geno	A4 geno	A5 geno
Age (months)	0.8563	0.2120	0.0012**	0.0061**	0.5165	0.7167	0.2238	0.5175	0.6233	0.4929	0.9460	0.6560
Sex	0.8895	NA	0.1523	0.3189	0.5708	0.3296	0.5153	0.5866	0.7008	0.2019	0.1682	0.1682
Time to Touch Stationary Umbrella (seconds)	0.3150	0.0682	0.2474	0.1302	0.1192	0.0018**	0.0231*	0.2984	0.5600	0.2258	0.1207	0.2175
Touch Stationary Umbrella?	0.5128	0.3296	0.1967	0.3818	0.6353	N/A	0.0488*	0.6218	0.1663	0.2701	0.4126	0.1858
Distance Away from Umbrella When Opened	0.6263	0.2510	0.3161	0.5334	0.6845	0.0010**	0.1486	0.8290	0.7350	0.5601	0.6911	0.4216
Distance Traveled When Umbrella Opened	0.0099**	0.5346	0.9757	0.9494	0.7397	0.8044	0.3738	0.8262	0.8755	0.9055	0.8446	0.7025
Time to Touch Open Umbrella	0.3920	0.4599	0.0481*	0.0236*	0.0369*	0.1761	0.0002**	0.5266	0.9205	0.0076**	0.1446	0.0300*
Touch Open Umbrella?	0.8385	0.5153	0.0361*	0.0403*	0.1336	0.0488*	N/A	0.1289	0.3745	0.4346	0.1289	0.1431
Tactile Pressure Level	0.7201	0.0896	N/A	0.8618	0.8618	0.1239	0.5399	0.8713	0.3148	0.8523	0.8713	0.8523
Initial reactivity level	0.8357	0.8116	0.3659	0.2088	0.1705	0.6472	0.3659	0.9266	0.0080**	0.8713	0.9430	0.9430
Auditory Response	0.2655	0.1946	0.0104*	0.0120*	0.2576	0.0628	0.1329	0.7395	0.9204	0.7217	0.6207	0.4591
Factor 1	0.7612	0.7653	0.7109	0.3030	0.1464	0.8753	0.7563	0.3299	0.6281	0.3978	0.3299	0.2769
Factor 2	0.1662	0.2515	0.2306	0.5098	0.9214	0.6187	0.6035	0.2148	0.1033	0.2718	0.2148	0.2869
Factor 3	0.3133	0.6466	0.4148	0.6963	0.7547	0.4596	0.9712	0.5216	0.3800	0.3976	0.5216	0.9335
Factor 4	0.5906	0.3570	0.6086	0.3726	0.2039	0.2351	0.3000	0.7066	0.2964	0.6621	0.7066	0.5565

*Pearson p-values significantly different across mean scores at $\alpha<0.05$. **P-values significantly different at $\alpha<0.01$. †Only one individual from farm two/day three.

Table 4.5. Correlations between the startle trial results and trial results with the TWH-specific temperament factors.

Variable	Age (months) ϱ	P-value	I.R.L ϱ	P-value	Velcro response ϱ	P-value	Time Stat Umbrella (sec) ϱ	P-value	Distance away from Umbrella Open ϱ	P-value	Distance Traveled Upon Opening ϱ	P-value	Time Open Umbrella (sec) ϱ	P-value	Algometer (Pressure) ϱ	P-value
I.R.L	-0.1491	0.6811	NA	NA	0	1	0.0257	0.9439	0.1156	0.7506	-0.1486	0.6821	0.5005	0.1406	-0.2673	0.4869
Auditory response	-0.2946	0.4415	0	1	NA	NA	0	1	0.6489	0.0587	-0.0265	0.9461	-0.0536	0.8911	0.378	0.3559
Time Stat Umbrella (sec)	-0.3216	0.3648	0.0257	0.9439	0	1	NA	NA	0.0873	0.8106	-0.2805	0.4325	0.2884	0.4191	0.5667	0.1116
Distance away from Umbrella Open	0.0517	0.8872	0.1156	0.7506	0.6489	0.0587	0.0873	0.8106	NA	NA	-0.2258	0.5306	-0.101	0.7814	0.5805	0.1012
Distance Traveled Upon Opening	0	1	-0.1486	0.6821	-0.0265	0.9461	-0.2805	0.4325	-0.2258	0.5306	NA	NA	0.2654	0.4586	-0.6276	0.0704
Time Open Umbrella (sec)	-0.2035	0.5728	0.5005	0.1406	-0.0536	0.8911	0.2884	0.4191	-0.101	0.7814	0.2654	0.4586	NA	NA	-0.1017	0.7946
Algometer (Pressure)	0	1	-0.2673	0.4869	0.378	0.3559	0.5667	0.1116	0.5805	0.1012	-0.6276	0.0704	-0.1017	0.7946	NA	NA

*Spearman correlations and p-values significantly correlated at $\alpha<0.05$. **Spearman correlations and p-value significantly correlated at $\alpha<0.0125$.

Table 4.5 (Continued)

Variable	Age (months) ϱ	P-value	I.R.L ϱ	P-value	Velcro response ϱ	P-value	Time Stat Umbrella (sec) ϱ	P-value	Distance away from Umbrella Open ϱ	P-value	Distance Traveled Upon Opening ϱ	P-value	Time Open Umbrella (sec) ϱ	P-value	Algometer (Pressure) ϱ	P-value
Factor 1	-0.3819	0.2761	0.3917	0.263	0.1054	0.7872	0.7091	0.0217*	0.081	0.8239	-0.5	0.1411	0.2761	0.44	0.35	0.3558
Factor 2	-0.1407	0.6982	-0.4302	0.2146	-0.0527	0.8929	0.1394	0.7009	0.1995	0.5806	-0.2134	0.5538	-0.1534	0.6723	0.3	0.4328
Factor 3	0.5226	0.1212	0.2633	0.4624	-0.0527	0.8929	-0.2121	0.5563	-0.2182	0.5449	0.4634	0.1774	0.3865	0.2699	-0.2167	0.5755
Factor 4	0.3618	0.3043	-0.2312	0.5205	-0.1581	0.6845	-0.0909	0.8028	0.1621	0.6547	-0.0671	0.8539	-0.7915	0.0064*	-0.1333	0.7324

*Spearman correlations and p-values significantly correlated at $\alpha<0.05$. **Spearman correlations and p-value significantly correlated at $\alpha<0.0125$.

Relationship of Behavior with SNP Genotypes

All 16 horses were genotyped for the five 3'-UTR SNPs in the *NXPH2* gene and all three genotypes were detected for each SNP. All genotypes were detected at each farm, except the AA genotype for the A2 SNP at farm two (Figure 4.3). There was no difference in the allele frequency for all SNPs between the two farms (p-values ranged from 0.6879 to 1).

The A2 SNP was significantly associated with initial reactivity level from the behavior trial and factor 2 scores from the questionnaire, with the GA genotype having higher scores for both traits (Table 4.8). The A3 SNP was significant for amount of time to touch the open umbrella and competitive score (Table 4.4), with the CT genotype being faster and the CC genotype less competitive. The A5 genotype was also significant for the amount of time to touch the open umbrella (Table 4.4), with the TA genotype taking the least amount of time.

Startle Response and Gait Ability

Horses unable to trot traveled an average of 9ft farther from the moving novel object than multi-gaited horses, indicating a potential difference in startle reactivity of horses with different gait capabilities. Measures of PPI might be more appropriate for quantifying innate differences in startle response, but as demonstrated by the logistical difficulties we encountered in preliminary experiments, these are not easily adapted for the horse.

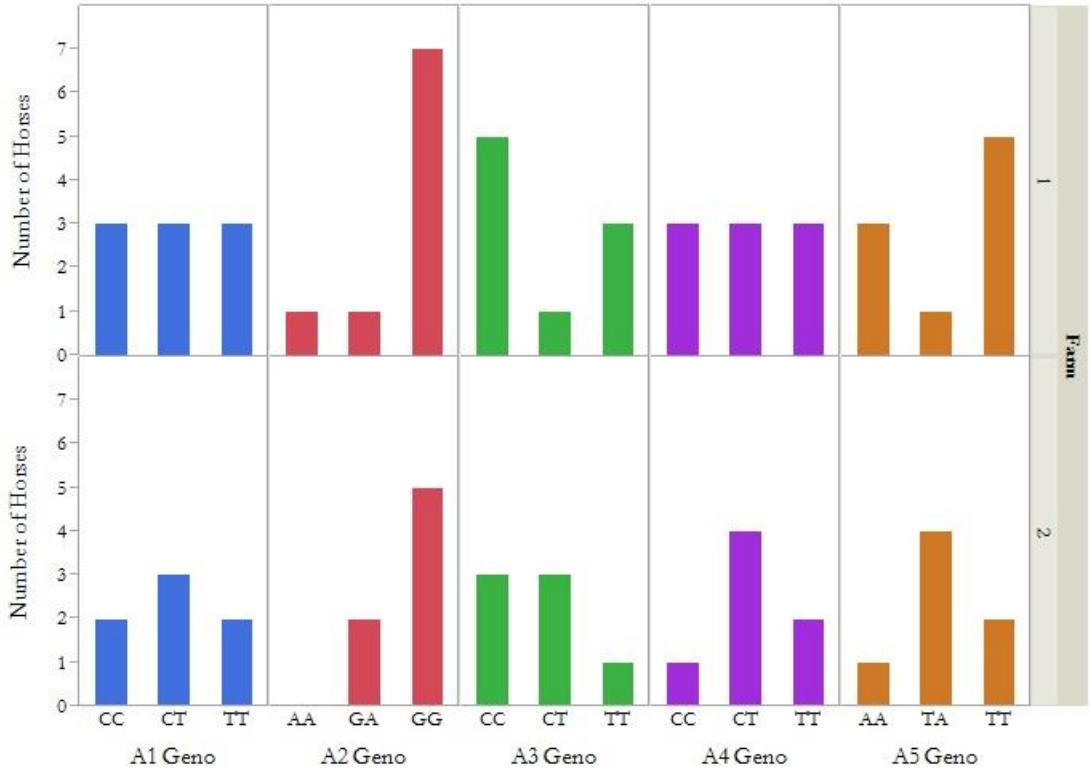
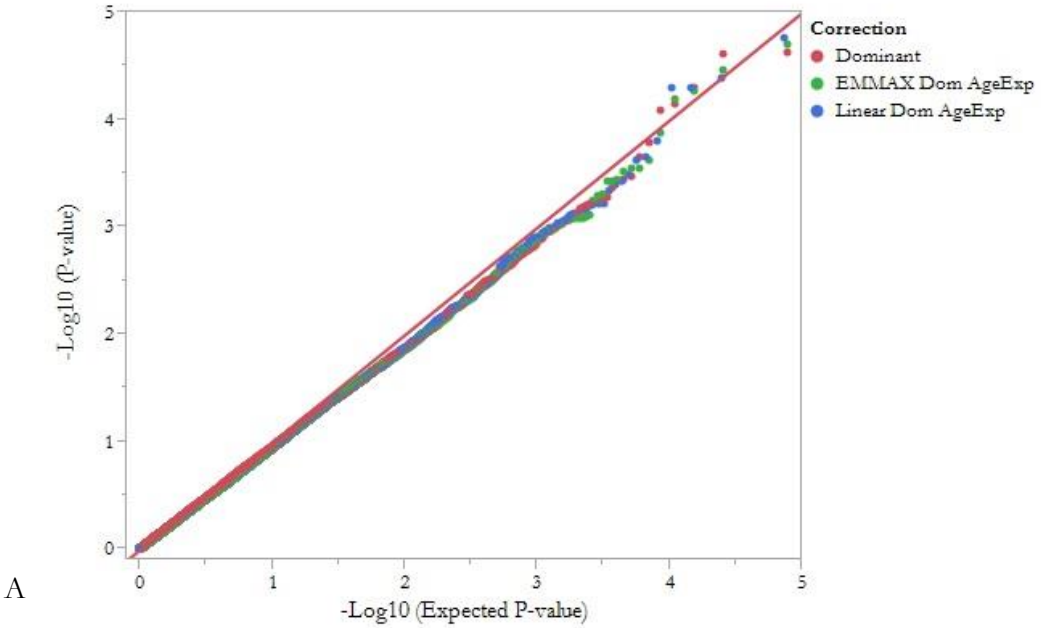


Figure 4.3. Genotype distribution of each SNP across the two test farms highlights the lack of the AA genotype for the A2 *NXPH2* SNP at Farm 2.

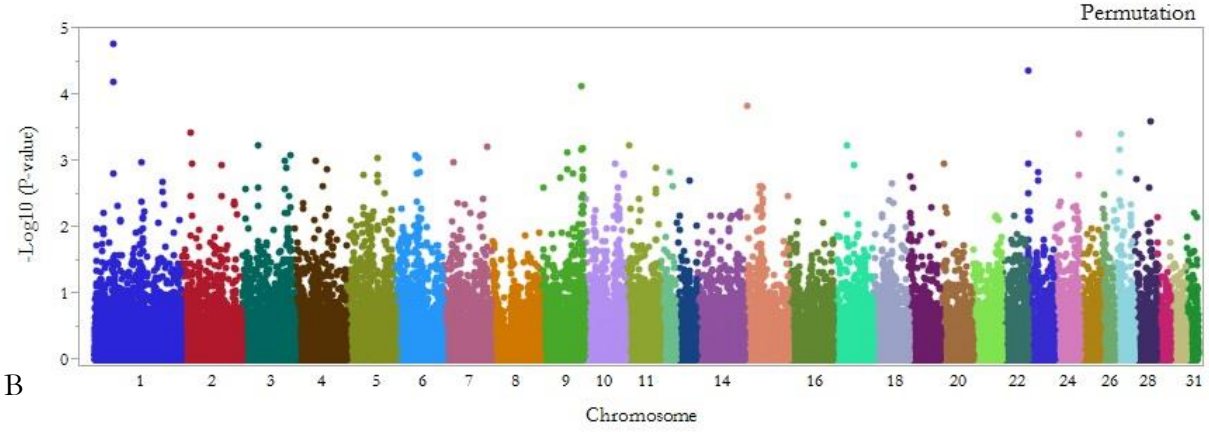
Candidate Loci on ECA1 and ECA23 Contribute to Variation in Factor 1 'Neophobia'

To identify loci controlling factor 1-neophobic variation in TWH, we utilized genotypes for 65160 loci in 113 horses to perform a genome wide association study. Linear quantitative trait association run under a dominant model identified candidate loci on ECA1, ECA9, and ECA23 with p-values ranging from 1.74e-5 to 5.02e-5 (Table 4.6). The significant markers on ECA1 fall within a 19.6 Kb block spanning nine exons of the *ALDH18A1* gene, respectively. The best marker on ECA23 falls within a 2 Kb intron of the *HSD17B3* gene. However, none of the identified markers surpassed our Bonferroni cutoff. After one million permutations, the top candidate marker on ECA1 had a p-value of 1.70e-5 (Figure 4.4b). The second candidate marker on ECA23 reached 835,000 permutations with a p-value of 4.31e-5. PLINK genomic inflation was limited to a factor of 1, indicating no population stratification. Examination of the QQ-plot (Figure 4.4a) confirmed the control for stratification and the limited association signal, likely due to insufficient statistical power. A linear model applied using EMMAX to the same dataset verified the candidate loci on ECA1, ECA9, and ECA23 with p-values ranging from 1.98e-5 to 6.53e-5 (Table 4.6).

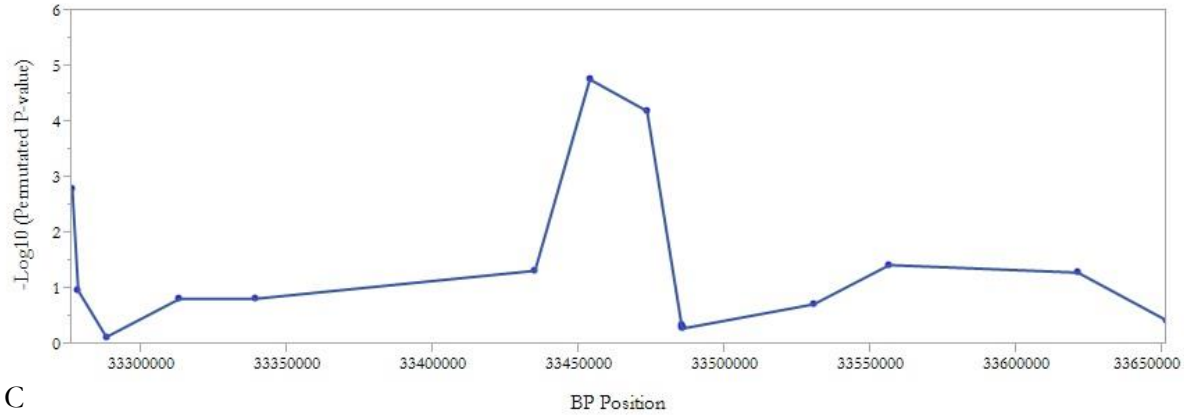
Figure 4.4. Genome wide association analysis identifies a suggestive peak on ECA1 for Factor 1 – 'Neophobia'. A) Quantile-quantile plots of different association models indicates no p-value inflation due to substructure after EMMAX correction. B) Manhattan plot of the permutated –Log10 p-values. C) the 596Kb region surrounding the candidate locus on ECA1 spans several candidate genes (D) from the UCSC genome bowser and shows E) several blocks of LD in the region. The two candidate loci on ECA1 are in complete LD.


Correction
Dominant
EMMAX Dom AgeExp
Linear Dom AgeExp
-Log10 (P-value)
-Log10 (Expected P-value)

A

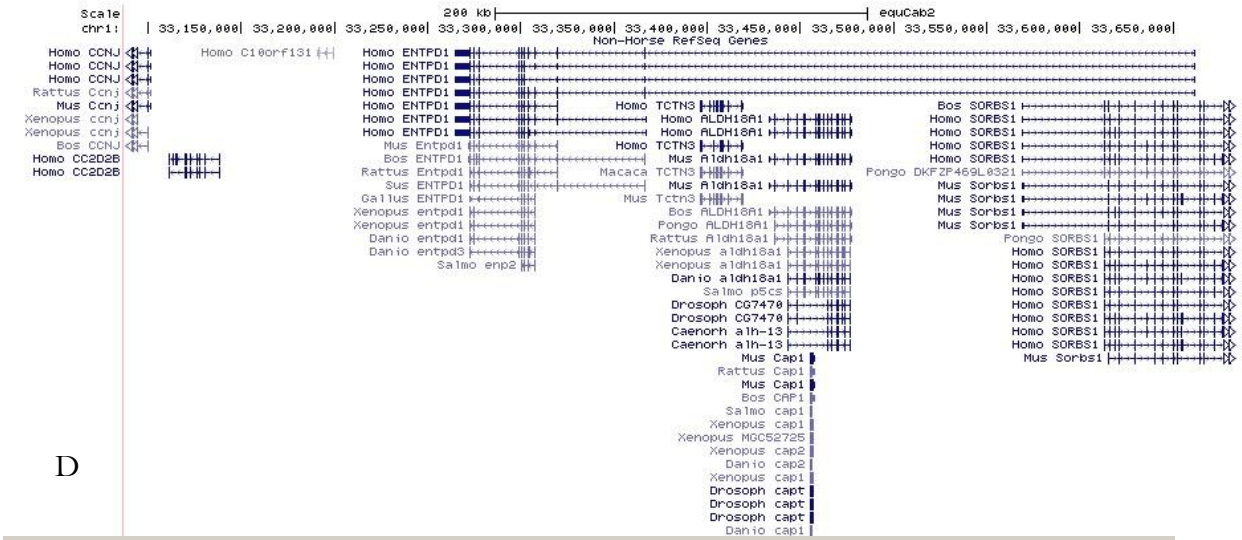

Permutation
-Log10 (P-value)
Chromosome

B

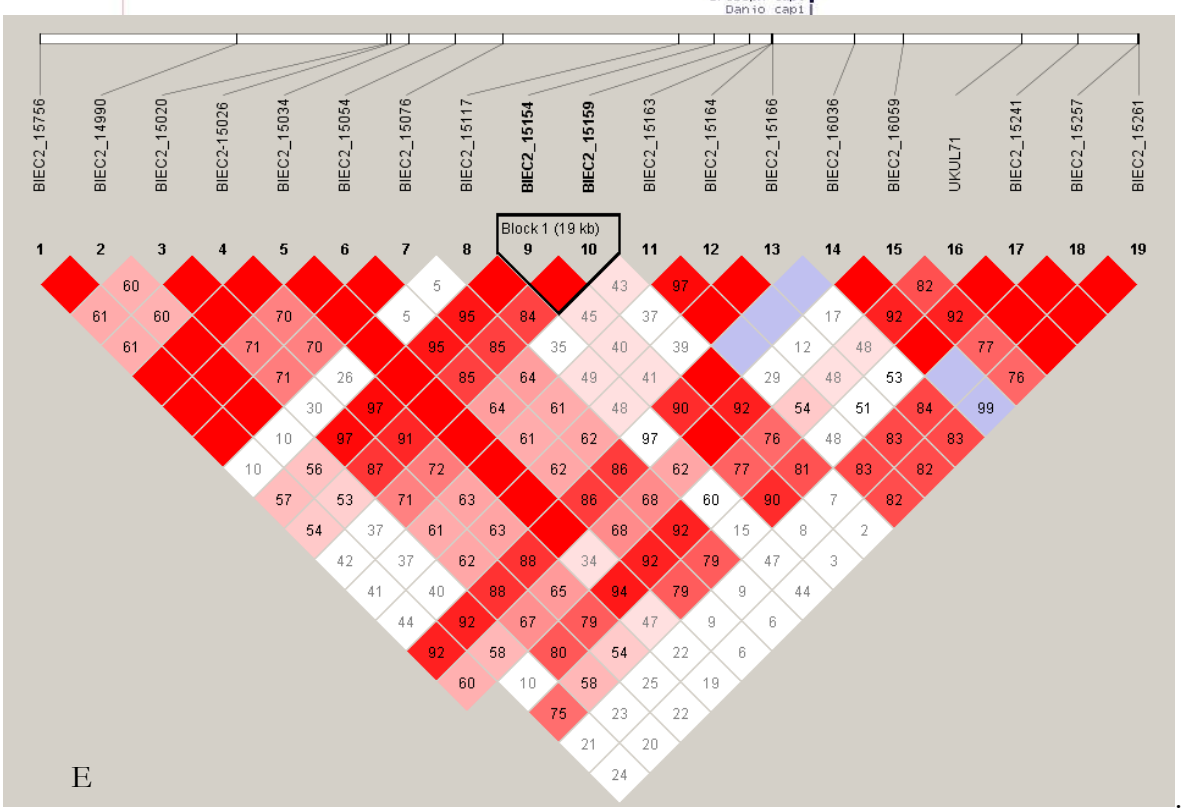

-Log10 (Permutated P-value)
BP Position

C

Figure 4.4 (Continued)



D



E

Table 4.6. P-values for association with factor 1-'Neophobia'.

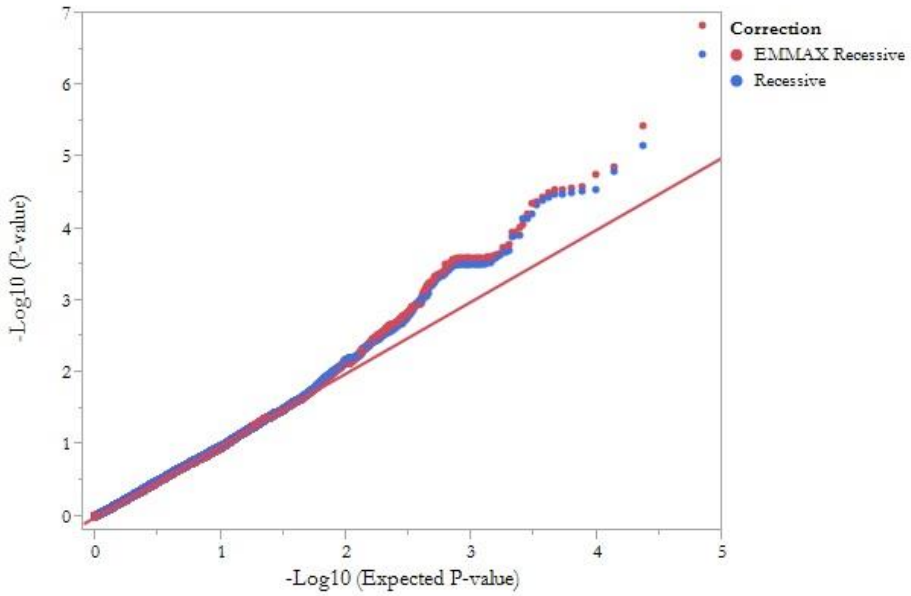
SNP	Chr	BP	Major Allele	Minor Allele	Raw P-value	Permutated P-values	# Permutations	EMMAX P-value	Marker in Gene
BIEC2_15154	1	33453429	A	C	1.74E-05	1.70E-05	1000000	1.98E-05	*ALDH18A1, ENTPD1*
BIEC2_604717	23	1197047	T	C	4.15E-05	4.31E-05	835000	3.44E-05	*HSD17B3*
BIEC2_15159	1	33473046	T	C	5.11E-05	6.32E-05	570000	6.53E-05	*ALDH18A1, ENTPD1*
BIEC2_1106102	9	77622379	T	C	5.02E-05	7.40E-05	486513	5.50E-05	1.8Mb region no gene annotation
BIEC2_283421	15	6057803	G	A	1.58E-04	1.49E-04	241274	1.33E-04	1.9Mb region no gene annotation
BIEC2_738608	28	32996239	C	T	2.21E-04	2.54E-04	142000	2.36E-04	*MCM5*

Loci on Chr 13 Contribute to Variation in Factor 2 'Trainability'

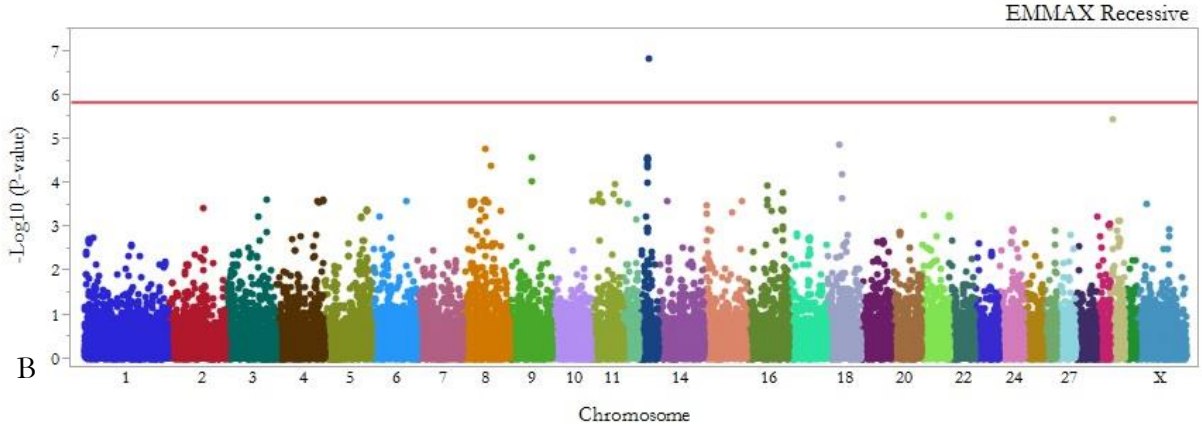
Quantitative trait association run under a recessive model identified candidate markers on ECA8, ECA9, ECA13, ECA18, and ECA30 with p-values ranging from 3.69e-7 to 3.58e-5 (Table 4.7). The top candidate marker on ECA13 falls within a 28 Kb intron of the *PRKCB* gene (Figure 4.5d). The other markers on ECA13 span a 1 Mb gene-dense region, but fall within *PRTT2, TUFM,* and *GSG1L* (Figure 4.5f) based on the LD structure of the region (Figure 4.5g). The candidate marker on ECA30 falls within a 87 Kb intron of the *SMYD3* gene. However, only the top marker on ECA13 surpassed our Bonferroni cutoff. Genomic inflation was limited to a factor of 1, indicating adequate control for stratification. Examination of the QQ-plot (Figure 4.5a) identified limited inflation and association with factor 2. The linear model applied in EMMAX verified the candidate loci on ECA8, ECA13, ECA18, and ECA30 (Figure 4.5b) with p-values ranging from 1.48e-7 to 4.53e-5 (Table 4.7).

Figure 4.5. Genome wide association analysis shows a significant peak on ECA13 for Factor 2 – 'Trainability'. A) Quantile-quantile plots of different association models show no p-value inflation due to substructure after EMMAX correction. B) Manhattan plot of the –Log10 p-values of the recessive EMMAX model. C) The 742 Kb region surrounding the candidate locus on ECA13 spans several candidate genes (D) from the UCSC genome bowser and shows E) small blocks of LD in the region.

A


Correction
EMMAX Recessive
Recessive
-Log10 (P-value)
-Log10 (Expected P-value)

B


EMMAX Recessive
-Log10 (P-value)
1 2 3 4 5 6 7 8 9 10 11 14 16 18 20 22 24 27 X
Chromosome

C

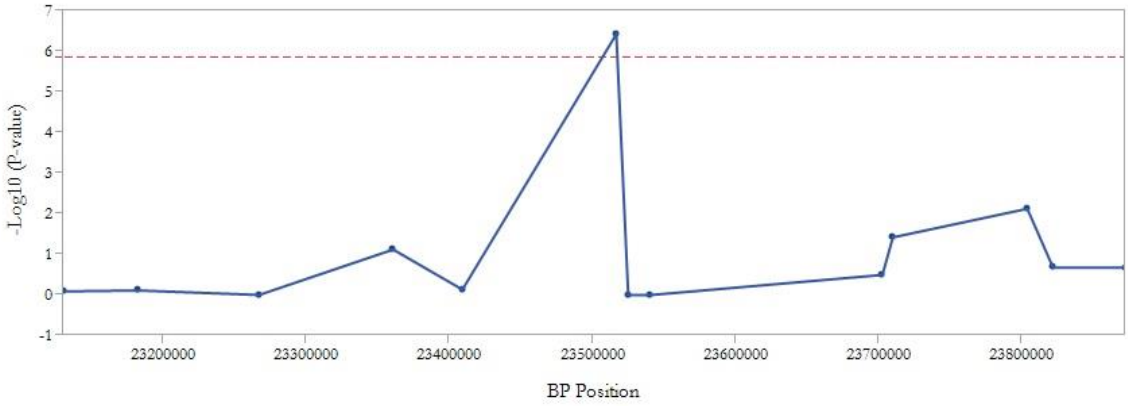
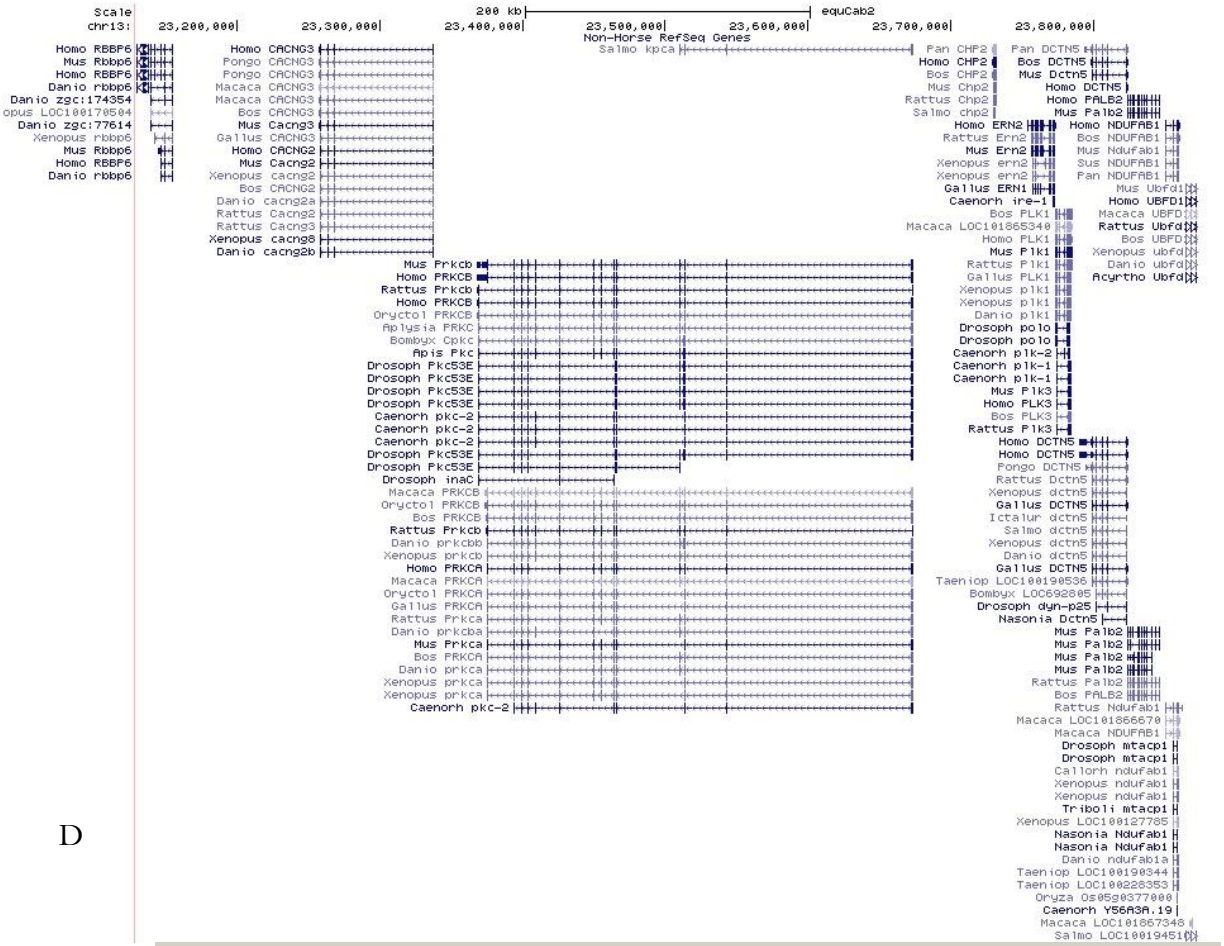

-Log10 (P-value)
23200000 23300000 23400000 23500000 23600000 23700000 23800000
BP Position

Figure 4.5 (Continued)

D



E

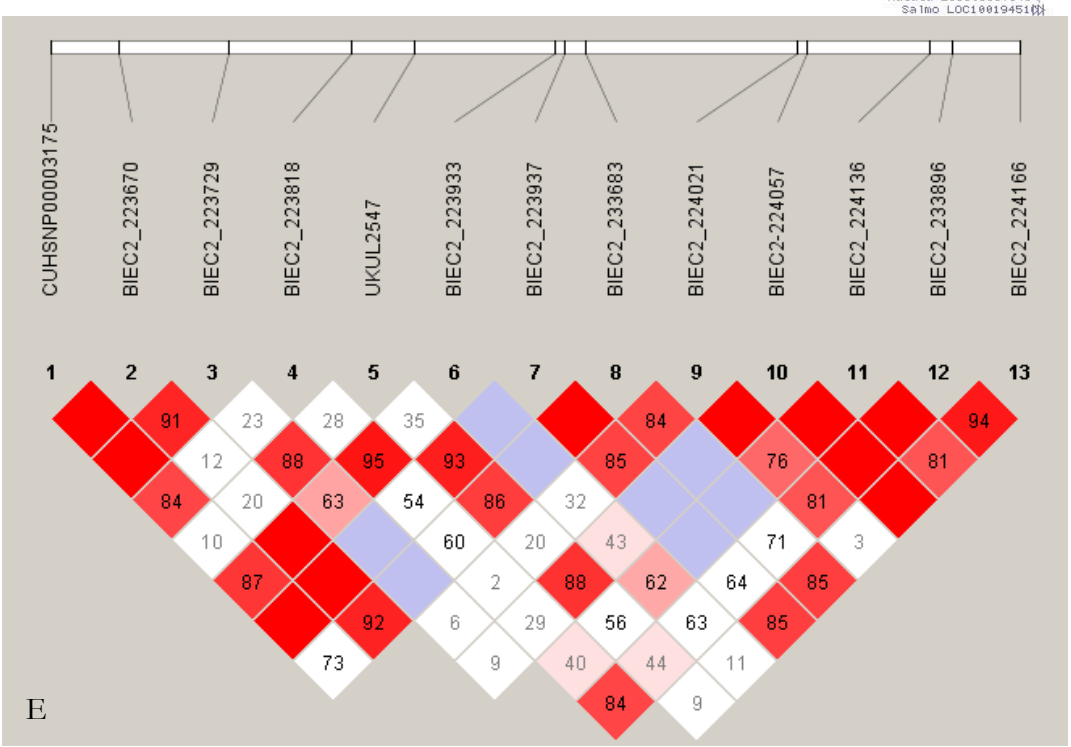
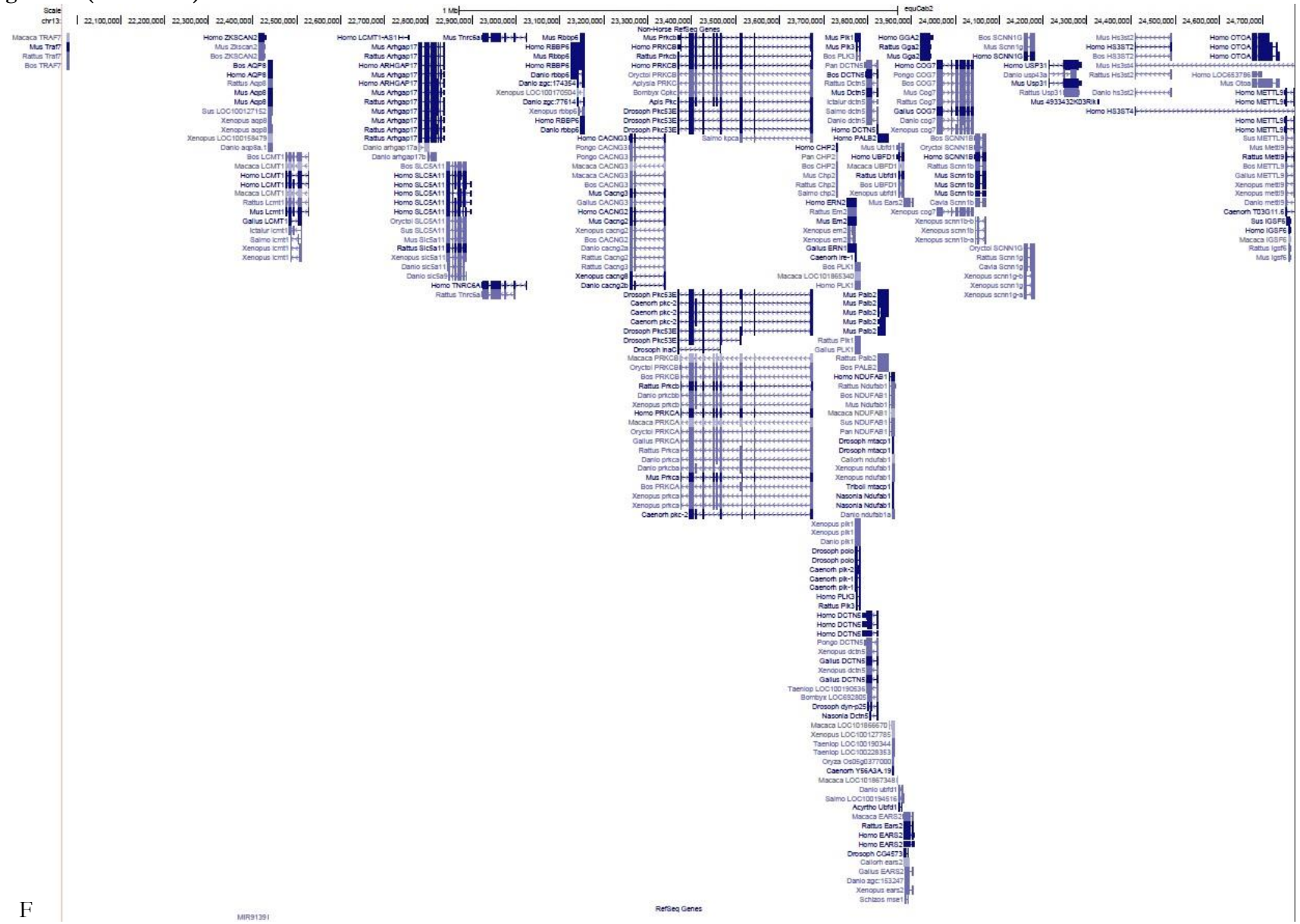


Figure 4.5 (Continued)



F

Figure 4.5 (Continued)

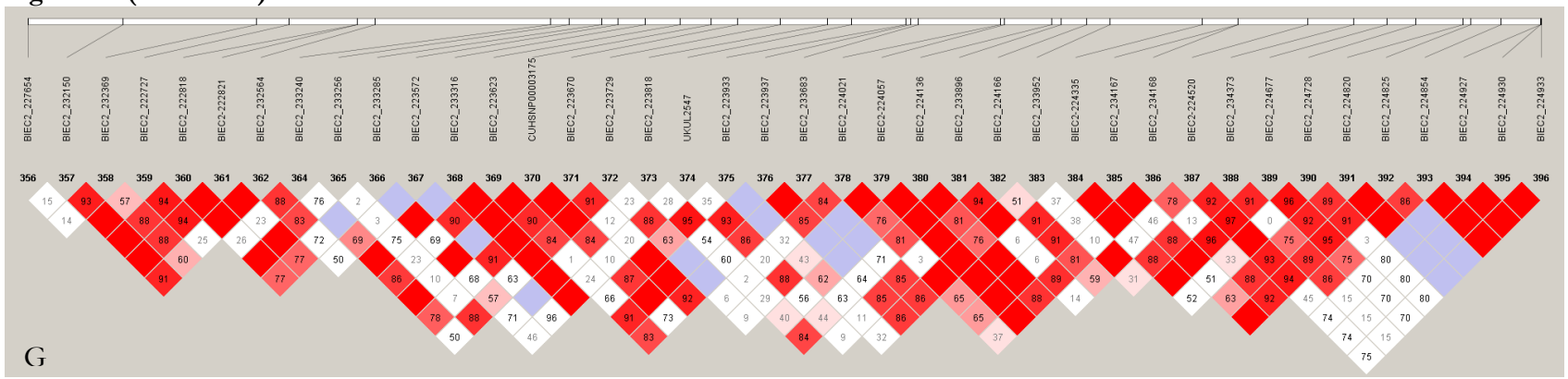


Table 4.7. P-values for association with factor 2-'Trainability'.

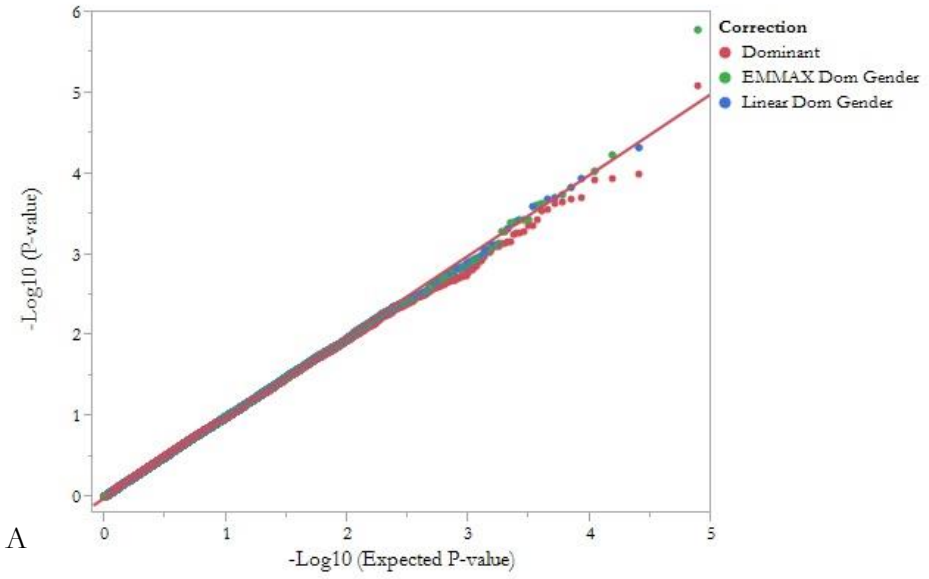
SNP	Chr	BP	Major Allele	Minor Allele	Raw P-value	EMMAX P-value	Marker in Gene
BIEC2_223933	13	23516316	T	G	3.69E-07*	1.48E-07*	*PRKCB, PRKCA*
BIEC2_856922	30	5124578	C	T	6.88E-06	3.69E-06	*SMYD3*
BIEC2_409342	18	25280342	C	T	1.56E-05	1.35E-05	*LRP1B*
BIEC2_215531	13	20345422	G	A	2.88E-05	3.71E-05	*GSG1L*
BIEC2_215014	13	19421156	C	A	3.07E-05	2.58E-05	*PRRT2*
BIEC2_1047477	8	44164828	A	G	3.22E-05	1.74E-05	*DLGAP1*
BIEC2-215206	13	19890320	A	G	3.27E-05	2.80E-05	*TUFM*
BIEC2_1091441	9	45279882	A	C	3.38E-05	2.70E-05	*VPS13B*
BIEC2_214950	13	19300483	G	A	3.58E-05	2.92E-05	*TBS1D10B*
BIEC2_215310	13	20096505	A	G	4.03E-05	3.12E-05	*XPO6*
BIEC2_215313	13	20096676	G	T	4.73E-05	4.53E-05	*XPO6*

*SNP surpassed Bonferroni cutoff of 1.42×10^{-6}.

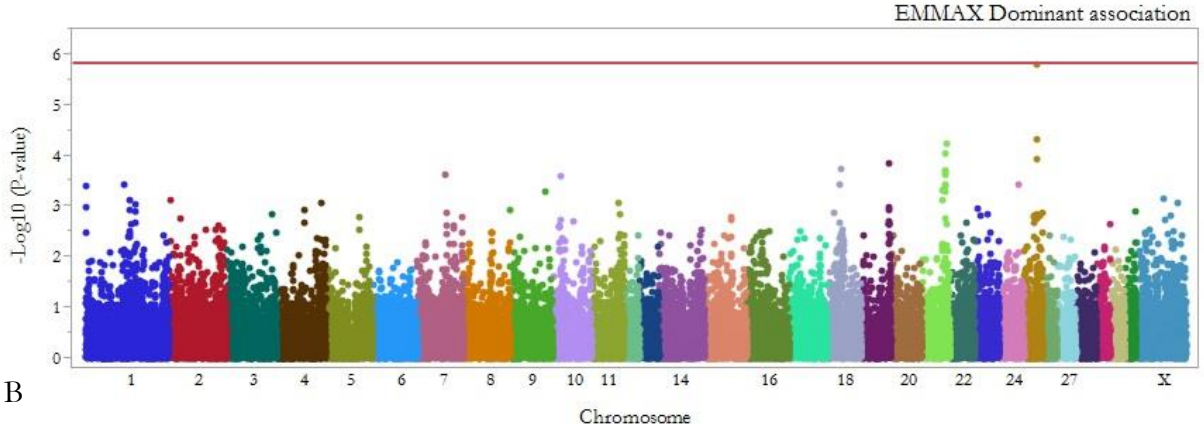
Loci on Chr 21 and Chr 25 May Contribute To Variation in Factor 3 'Hostility'

Quantitative trait association for Factor 3, "hostility", identified candidate markers on ECA21 and ECA25 with p-values ranging from 1.62e-6 to 9.16e-5 (Table 4.8). The top candidate locus on ECA25 falls within the 5' intron of the salmon *STOM* gene; the second locus falls within a 5Kb intron of *STOM* (Figure 4.6d). The loci on ECA21 span a 2.69Mb region that includes the *ROPN1L*, *MARCH6*, *CMBL*, *CCT5, TAS2R1, SEMA5A, MTRR, FASTKD3,* and *ADCY7* genes. However, only the top locus on ECA25 surpassed our Bonferroni cutoff. PLINK genomic inflation was limited to a factor of 1.0, indicating adequate control for stratification. Examination of the QQ-plot (Figure 4.6a) confirmed the control for stratification and the low association, likely due to not enough power. The EMMAX model applied to the same dataset verified the candidate loci on ECA21 and ECA25 with p-values ranging from 1.62e-6 to 1.17e-5 (Figure 4.6b). After EMMAX correction, the genomic inflation was reduced to a factor of 0.99918. However, comparison of the EMMAX association results with linear association tests revealed no evidence of P-value inflation due to additional population structure in our cohort.

Figure 4.6. Genome wide association analysis shows a significant peak on ECA25 for Factor 3 – 'Hostility'. A) Quantile-quantile plots of different association models illustrates no p-value inflation due to substructure after EMMAX correction. B) Manhattan plot of the –Log10 p-values of the EMMAX dominant model with gender added as a covariate indicates no difference in p-values. C) The 515Kb region surrounding the candidate locus on ECA25 spans several candidate genes (D) from the UCSC genome bowser and reveals a E) small block of LD in the region.


Correction
Dominant
EMMAX Dom Gender
Linear Dom Gender
-Log10 (P-value)
-Log10 (Expected P-value)
0
1
2
3
4
5
6

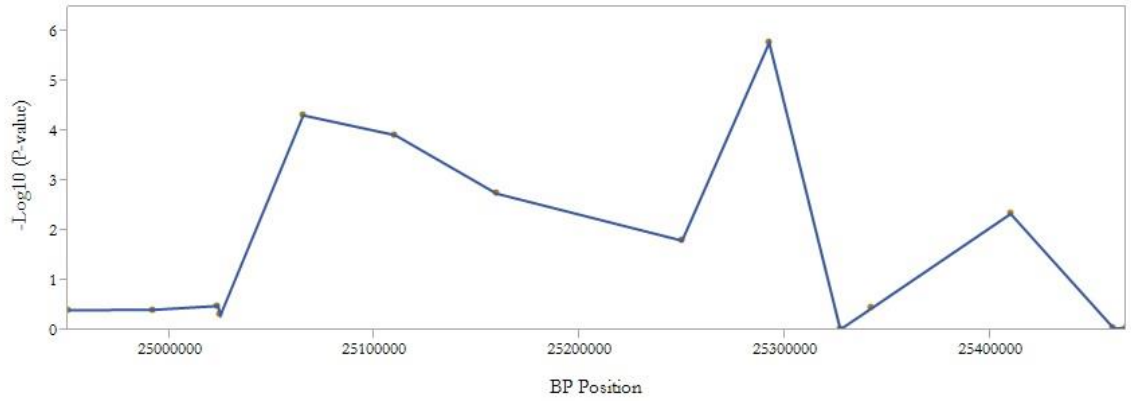
A


EMMAX Dominant association
-Log10 (P-value)
Chromosome
1
2
3
4
5
6
7
8
9
10
11
14
16
18
20
22
24
27
X

B

Figure 4.6 (Continued)

C



D

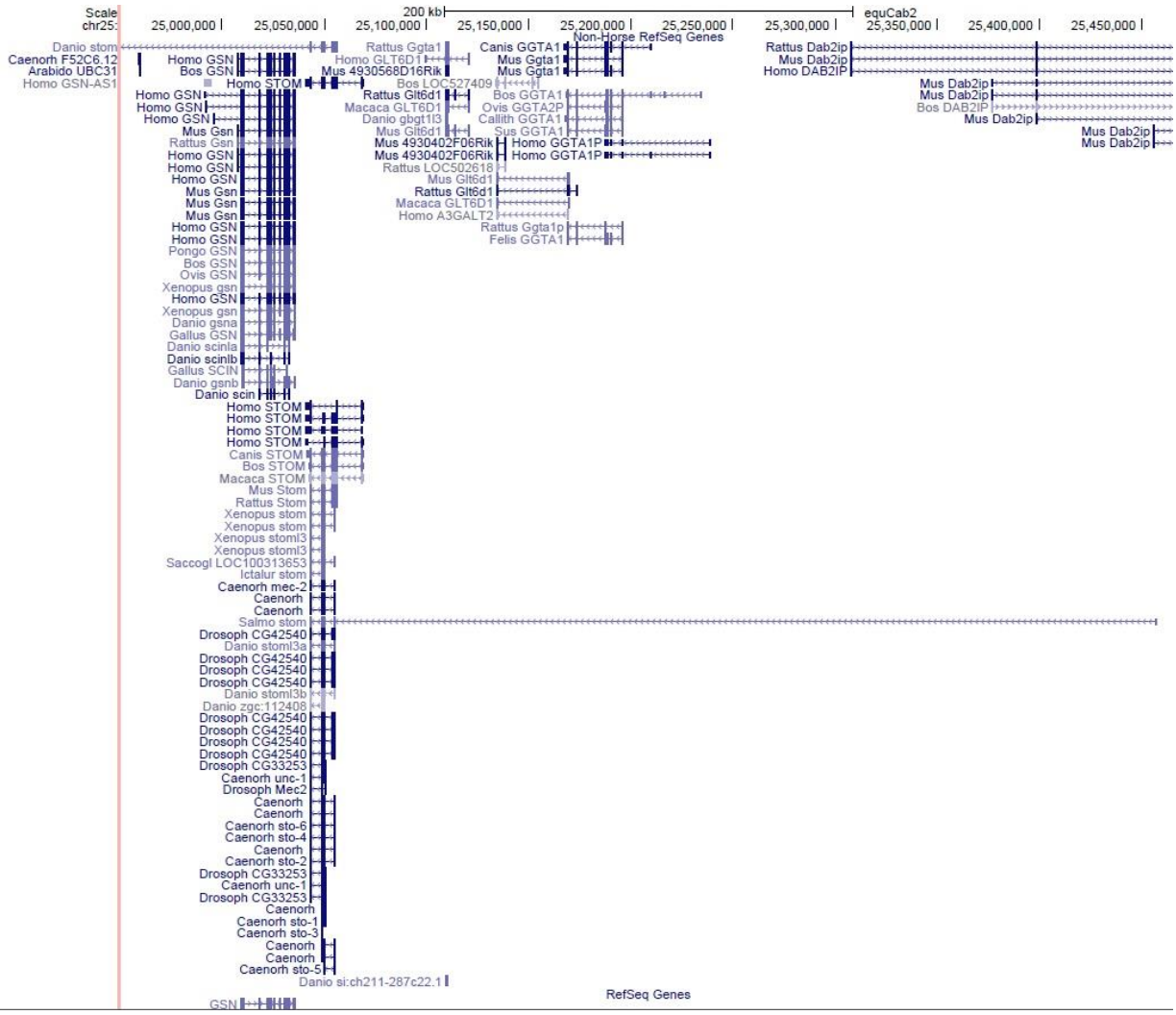
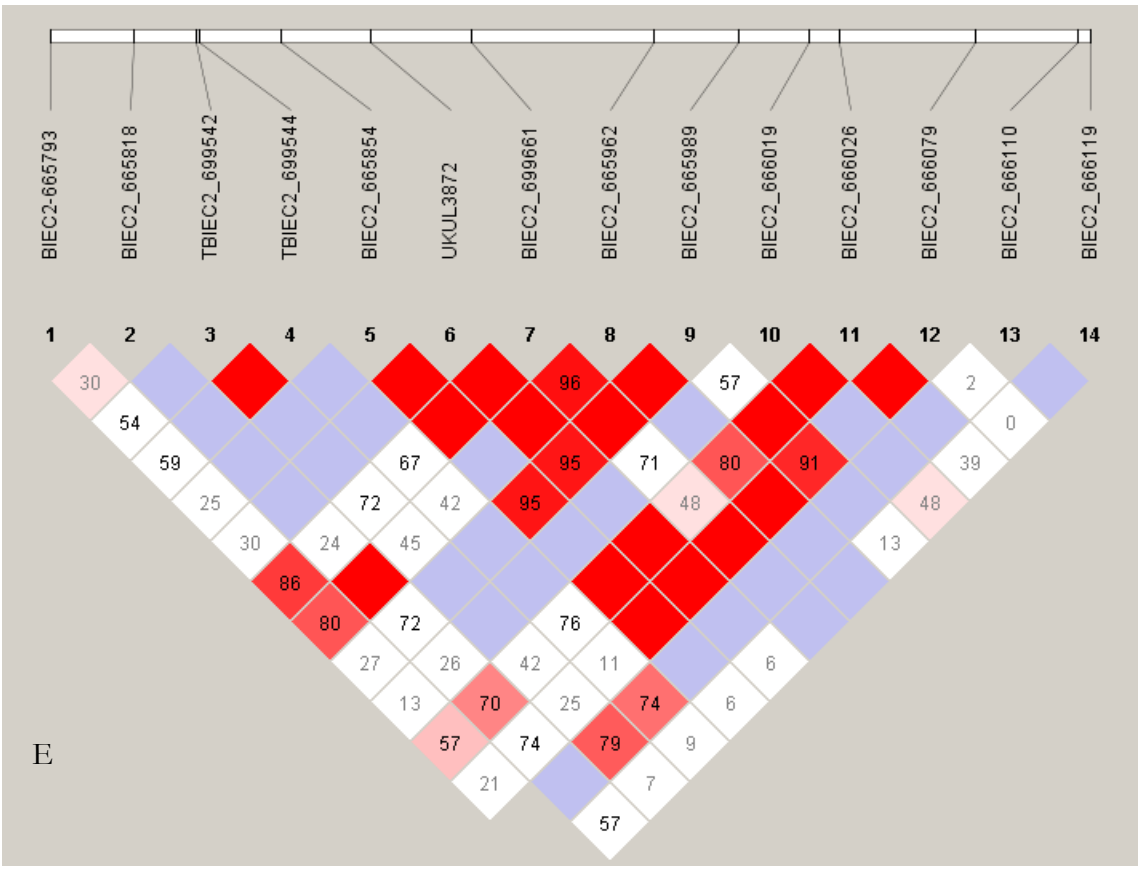


Figure 4.6 (Continued)



E

Table 4.8. P-values for association with factor 3-'Hostility'.

SNP	Chr	BP	Major Allele	Minor Allele	Raw P-value	EMMAX P-value	Marker in Gene
BIEC2_665989	25	25292122	T	G	1.62E-06	1.62E-06	*STOM*
BIEC2_665854	25	25064974	T	C	4.69E-05	4.69E-05	*STOM*
BIEC2_601248	21	51892904	C	T	5.79E-05	5.79E-05	*ADCY2*
BIEC2_598944	21	49202410	C	T	9.16E-05	9.16E-05	*ROPN1L*
UKUL3872	25	25109769	C	T	1.17E-04	1.17E-04	*STOM, GLT6D1*

Discussion and Conclusions

Behavior and temperament assessment is an extremely important aspect for selection and discipline suitability in all horse breeds and both are considered a key issue in horse health and performance (Buckley, Dunn, More 2004). Improvement in temperament assessment can lead to improved selection for appropriate jobs and therefore lead to improved management and welfare for the horse, and improved safety for their handlers.

Our results demonstrate that there are differences in behavior/temperament across gaited breeds of horses, and that there are key behavioral trait differences among sexes, gait types and training disciplines across and within breed. In this study, we identified four temperament components shared across breeds: 'Neophobia', 'Trainability', 'Independence', and 'Hostility'. Within the TWH breed, we identified the same four temperament components and have successfully identified candidate regions for 'neophobia', 'trainability', and 'hostility' factors. We have also shown a wide variation in response to our preliminary startle trial and with modification, it may be feasible to expand this method in to a measure of PPI in the horse. Candidate genes identified in this study (Chapter 2) along with the neurophysiology of central pattern generators, suggest that PPI metrics may be altered in gaited horses.

Questionnaires Highlight Discipline Specific Temperament Traits

Collection of data at horse shows and private farms requires a behavior assessment that is relatively quick to complete yet still able to accurately capture temperament variation. The 20 item questionnaire developed by Momozawa (2005) had already proven a successful tool for study of equine temperament and is well suited to provide preliminary insights to temperament variation in gaited breeds of horses. Given the limited amount of time it is reasonable to ask owners to spend completing a survey, we felt the 20 question format would be faster to complete than the longer questionnaire developed by Lloyd *et al.* (2007). Principle component analysis and factor analysis of

behavior traits from questionnaires from diverse breeds revealed fairly consistent trends in anxiety, nervousness, trainability, and sociability traits. Traits related to anxiety and nervousness consistently account for the greatest amount of variation across and within breeds for both methods (Lloyd *et al.* 2008; Lloyd *et al.* 2007; Momozawa *et al.* 2005). Within our own data, our factor 1 labeled as 'neophobia' (incorporates nervous, panic, and excitability traits) and factor 2 labeled as 'trainability', accounted for the greatest and second greatest variation both across breeds and within the TWH.

Traits related to 'independence, and 'hostility' account for the third and fourth largest component of variation across and within breeds in our dataset. In our across breed dataset, 'independence' ranks as the third largest component and 'hostility' the fourth. However, this order is swapped when we examined a single breed, the TWH. The TWH accounts for nearly half of the horses in the sample set; there are almost three times as many TWH than Rocky Mountain horses, the breed with the second highest contribution. Yet, despite the overwhelming influence of the TWH and low numbers of the other gaited breeds, we were able to detect breed related temperament differences. For example, the Icelandic horse is generally regarded as a stoic breed that is willing to cooperate and work (Hendricks 1995; USIHC ; Ziegler 2005). In our results, Icelandic horses had the highest scores for Factor 1 (neophobia) and one of the lowest Factor 3 (independence) scores, indicating they are the least excitable, skittish, nervous and panicky of the breeds surveyed and are more at ease being alone, matching their breed temperament description. However, temperament differences may also reflect an artifact of the sampling methods used in our study. For example, the majority of the horses from the Mangalarga Marchador, Campolina, and Rocky Mountain horse breeds were sampled at horse shows, and so only represents a strict sect of one of the behavior types that are present within the breed. Whereas the Morgan samples primarily came from one private farm, again only representing a limited amount of the variation that

may be occurring in the breed, and the TWH were sampled at both shows and private farms providing a more diverse temperament background. Additionally, some of the raters may have answered differently if completing the survey in front of a client or colleague, for example, when a caretaker is completing the survey in front of the owner, etc. With additional samples from more diverse disciplines across each breed we may be able to detect individual breeds' unique temperament traits.

Across all samples, horses were subjectively categorized into discipline groups based on owner reported discipline use. From these, we identified behavior traits typical for each training discipline (show, trail, and breeding). Trail horses were scored as less excitable, less panicky, and less skittish than their non-trail counterparts. These findings seem appropriate since trail horses are typically ridden in novel areas with variable amounts of stimulation and would be important for both rider and horse safety that the horse be stoic in strange situations. Interestingly, trail horses were scored as less friendly to other horses both within and across breeds. Trail horses may be less friendly to other horses due to frequent introduction to new horses in unfamiliar surroundings and the lack of an established hierarchy with the new horse. This result may also reflect a bias on the part of the rater who may use the horse's current interaction with herd mates as the basis for all new interactions rather than considering all occurrences of new meetings.

Show horses are typically preconceived to be flighty and neurotic, but our results demonstrate quite the opposite personality for show horses (as perceived by their owners). Show horses were rated to be more patient, better able to concentrate and remember what they were taught, more docile, less timid, and more self-reliant. In considering the show-ring environment with lots of people and horses around, it makes sense that show horses have these traits to make them safer to be around and get through the chaos at a horse show. The one exception is that

show horses were rated as more excitable, which would be a beneficial counterpoint to the other traits in the show ring so that the horse would appear more animated and expressive in the ring.

Breeding horses were rated as being less friendly to people and to other horses. However, compared across the sexes, both mares and stallions were rated as less friendly to other horses than geldings, likely due the decreased level of steroid hormones in geldings (castrated males). Geldings were included in this analysis as there was a single horse used for breeding before he was castrated.

Cumulative time in training influenced how horses were scored in the survey. Horses that had been trained to wear a saddle and allow a rider on their back were rated as less excitable, friendlier to people, less likely to panic, more cooperative, less stubborn and have better memory. This is all a reflection of the desensitization that has likely occurred during training for the horse to willingly accept a rider on their back. Within the TWH breed, untrained horses were rated as less vigilant; this could be due to the inclusion of a large number of young horses who are part of herd with older and more protective members.

Across breeds, horses unable to trot were rated as more excitable than horses able to trot. This is opposite of what we expected to find, as light trotting breeds such as the Arabian and Thoroughbred are stereotypically considered fractious so we assumed horses able to trot would be rated as more excitable. However, this may be unique to gaited breeds as we did not sample from non-gaited breeds such as the Arabian and Thoroughbred. In our preliminary behavior trial, the lateral-gaited horses also traveled the farthest distance after being startled by an opening umbrella, potentially implicating a difference in startle reaction between horses of different gait types. This makes sense since startle reaction and locomotion share similar pathways in the spinal interneurons and motor neurons to influence skeletal muscle contraction (Koch 1999).

Prior behavior studies have identified that breed (Hauseberg and Muller 2002; Lloyd *et al.* 2008; Mader and Price 1980; Momozawa *et al.* 2003), age (Mader and Price 1980; Visser *et al.* 2002),

social environment (Søndergaard and Ladewig 2004) or the stage of training (Visser *et al.* 2002) could influence temperament. To address these, in future studies the collection of more detailed training and owner experience information would allow horses from similar backgrounds to be grouped and compared for a more meaningful analysis. This could include information on the level of training (i.e. basic, intermediate or advanced), the type of training equipment used (draw reins, shackles, bell boots, etc.), and the type of training used (desensitization vs sensitization, reward vs punishment, etc.). To better evaluate the owner's experience level, a questionnaire can be applied where the owner can self-rate themselves on their experience level (beginner, intermediate, advanced) based on their horse-interaction experiences. The questionnaire would cover topics such as the number of years they have owned a horse, if they board or provide self-care, are they involved in breeding horses, their riding experience, and any training certification or clinics attended, etc. This is important information to collect because an owner may have owned horses for several years, but actually never be involved in the daily care and handling of the horse, leading to misinformed responses about the horse's temperament. Additional information on the background of the horse will also be useful, such as how long the horse has been owned, the age of the horse when acquired, and the age when the horse was started under saddle, etc. to help account for the some of the variability in the owners' responses as horses owned or known longer should have a better representation of the horse's temperament than less familiar horses. To be statistically meaningful, we would need to collect a large number of horses and consider applying a mixed model to account for all of these environmental effects.

Beyond the additional training and background information, including a few more questions on additional behaviors such as aggression, hierarchy level, protectiveness, opportunistic/sneaky behavior and other behaviors, such as those from Lloyd *et al.* (2007), would allow us to better define our temperament factors. For example, our 'hostility' factor would either be strengthened or

changed to a different trait by the incorporation of an aggression score. Since one of the original reasons we developed the questionnaire was to test the claims and anecdotes that gaited breeds are calmer and more docile than non-gaited breeds, our next step will be to include non-gaited breeds. As we incorporate more breeds, we have to include larger representative numbers that are relatively balanced across the breeds to better evaluate the breed-specific traits.

Genome-Wide Association Mapping of TWH Temperament Traits Is Possible

Using the genotypes from 116 TWH with complete behavior surveys, we were able to identify regions associated with the temperament factors described above. These horses were genotyped for another project, and therefore most of these horses had moderate factor scores; individuals with extreme factor scores did not meet inclusion thresholds for the other project and thus were not genotyped. If we were to repeat this study, but utilize a population of individuals with extreme phenotypes, it is possible different regions will be identified as there will be a greater difference in the phenotype in the new study compared to our current study. When we applied models of inheritance to our factors we were able to identify significant associations in genes related to cell signaling and behavior differences in the mouse model, with some stratification that could not be accounted for with EMMAX correction and may be due to environmental differences.

The top candidates for Factor 1 include the *ALDH18A1* (aldehyde dehydrogenase 18 family, member A1) gene and the *HSD17B3* (hydroxysteroid 17-beta dehydrogenase 3) gene. The *ALDH18A1* gene is a member of the aldehyde dehydrogenase family and encodes a bifunctional ATP- and NADPH-dependent mitochondrial enzyme with both gamma-glutamyl kinase and gamma-glutamyl phosphate reductase activities (Baumgartner *et al.* 2000). The encoded protein catalyzes the reduction of glutamate to detla1-pyrroline-5-carboxylate, a critical step in the *de novo* biosynthesis of proline, ornithine and arginine (Baumgartner *et al.* 2005). Mutations in this gene lead to hyperammonemia, hypoornithinemia, hypocitrullinemia, hypoargininemia and

hypoprolinemia and may be associated with neurodegeneration, cataracts and connective tissue diseases (Baumgartner *et al.* 2000; Patel *et al.* 2011). Factor 1 (neophobia) scores were correlated with age, indicating horses become less neophobic with time. This is likely attributable to cumulative training and life experiences, or this could be due to increased visual or neurological degeneration as the horse ages.

The *HSD17B3* gene produces an isoform of 17 beta-hydroxysteroid dehydrogenase that is expressed predominantly in the testes and catalyzes the conversion of androstenedione to testosterone. It preferentially uses NADP as a cofactor, and deficiencies can result in male pseudohermaphroditism with gynecomastia in humans (Legeza *et al.* 2013). Aggressive behavior in horses has been associated with elevated testosterone levels (Beaver and Amoss 1982), and there may be an undetected correlation between the amount of circulating testosterone and neophobia rating. Future work could investigate a correlation between circulating testosterone and this polymorphism in a random sampling of horses to validate the association with neophobia scores.

The candidate gene for Factor 2 (trainability) is the *PRKCB* (protein kinase C beta type) genes. The *PRKCB* gene encodes for protein kinase C (PKC), a family of serine- and threonine-specific protein kinases that can be activated by calcium and second messenger diacylglycerol (Lee *et al.* 2014). The PKC family members phosphorylate a wide variety of protein targets and are known to be involved in diverse cellular signaling pathways (Farren *et al.* 2014; Lee *et al.* 2014). The specific protein kinase encoded by this gene has been reported to be involved in many different cellular functions, such as B cell activation, apoptosis induction, endothelial cell proliferation, and intestinal sugar absorption (Lee *et al.* 2014). Studies in mice also suggest that this kinase may regulate neuronal functions and correlate fear-induced conflict behavior after stress (de Diego-Otero *et al.* 2009). Training can be stressful and induce fear responses in horses, yet horses have different

reactions to the same training regime; this difference may be due to a genetic predisposition in the horse's perception of training.

Another interesting and potential candidate gene for Factor 2 (trainability) includes the *DLGAP1* (discs large-associated protein 1) gene. The postsynaptic synapse-associated protein 90 (SAP90)/postsynaptic density 95 (PSD95)-associated proteins (SAPAPs) constitute a part of the NMDA receptor-associated postsynaptic density proteins, and are involved in the stabilization of synaptic junction and regulation of neurotransmitter receptors (Li *et al.* 2013; Takeuchi *et al.* 1997). The *DLGAP1* gene encodes the SAPAP1 protein which localizes at the post-synaptic density and interacts with PSD95 protein to maintain normal brain function and development (Kawashima *et al.* 1997; Kim *et al.* 1997). Kajimoto (2003) demonstrated that SAPAP1 protein levels increased in the nucleus accumbens of patients with schizophrenia; the nucleus accumbens is a region in the basal forebrain with a significant role in the cognitive processing of motivation, pleasure, reward, and reinforcement learning. In humans, the *DLGAP1* gene is located in a region linked to schizophrenia (Mukherjee *et al.* 2006; Segurado *et al.* 2003) and has also been implicated as a candidate gene for obsessive-compulsive disorder (Grados *et al.* 2014; Stewart *et al.* 2013). With links already to behavior-related disorders, *DLGAP1* is a promising candidate gene for behavior differences in horses.

The top candidate gene for Factor 3 (hostility) is ADCY2 (adenylate cyclase type 2) gene. The *ADCY2* gene encodes a member of the family of adenylate cyclases, which are membrane-associated enzymes that catalyze the formation of the secondary messenger cyclic adenosine monophosphate (cAMP) (Purves 2008). These play an important role in post-synaptic signaling pathways (Dell'Acqua *et al.* 2006), implicating differences in social hostility may be due to changes in the strength of neurotransmitter and synaptic transmission.

Startle Trials

A concern with questionnaires is that they are a subjective measure of behavior and can be confounded due to rater's experience level and temperament. However, we were able to identify some concordance between the questionnaires and an experimental startle response trial. Reactivity level, as determined by an experienced behaviorist, was correlated with identified behavior factors within our limited number of horses that were evaluated via both methods. However, only one farm has been sampled by both techniques so far. Prior studies between surveys and behavior trials found relationships between the resting heart rate and questionnaire results (Momozawa *et al.* 2003; Visser *et al.* 2003)(Visser, 2003), but due to technical difficulties with heart rate equipment we were unable to provide that additional evidence.

Environmental differences between the two test farms may also have confounded our results. Yet, testing at individual farms is a very cost effective sampling method. Horses are large animals, and transporting them can be very costly (~$1-2 per mile). Young horses are not typically acclimated to transport in a trailer, and therefore transport to a new facility can be extremely stressful and potentially introduce additional errors. Owners are also not likely to accept the risk of potential injury with the transport of young horses, especially if the animal is considered valuable. Therefore, it would be extremely difficult to find enough horses for a statistically significant study if we transported privately-owned horses to a single location. The ideal setup would be either purchasing young horses or working with a single large operation. However, these too have their pitfalls. Purchasing enough horses would be the most costly with the upfront purchase price, transport and maintenance costs, plus the dilemma of dispersal after the completion of the study. Single large operations with young horses are relatively rare in the horse industry and are typically very busy with the daily maintenance of the facility, so scheduling to not interfere with daily activities can be an issue; smaller facilities tend to be a little more flexible with their time. And due the large operations relative rarity, travel to these facilities can also be expensive. Therefore, we

sought out smaller farms with several young horses that were within a two-three hour driving radius to maximize participation numbers at a reduced cost for our preliminary trial.

This preliminary startle trial shows the potential feasibility of testing PPI in a larger subset of horses, with modifications. The PPI model utilizes a distinctive non-startling tone followed by a strong startling tone; in normal PPI, the weak tone reduces the magnitude of the reaction to the startle tone (Quednow *et al.* 2006). Our initial attempts to measure startle reaction with auditory tones, using similar protocols as in mice, resulted in no detectable startle reaction. A more novel auditory stimulus, the separating Velcro was sufficient to elicit a response. Further investigation into whether the sound of separating Velcro could be used as the startling tone to test for PPI needs to be done. In this study we were physically separating the Velcro for auditory startle, resulting in sound variability and could potentially be responsible for our variable startle responses. To improve the accuracy of the test, we need to determine if a recording of the sound elicits a startle response. With a recording, we could test if there is variable response to the speed the Velcro is separated and select the best frequency and decibel level for a startle reaction. This test would need to be conducted on young horses less than a year in age due to potential desensitization to Velcro separation in older horses accustom to the use of blankets and fly-masks which often feature Velcro fixtures. Once we have optimized our startle tone, we could continue testing to identify a unique auditory tone that does not elicit a startle response and then setup our PPI test. Differences in PPI in these young horses could allude to changes underlying their sensory reactivity and information processing, which may manifest in different gait type abilities and behavior traits. With a larger sample size, we would be able to draw more concise conclusions about the feasibility of utilizing PPI responses as a method for identifying neural circuits correlated with gait ability.

During testing, we also noted other traits that could be included in future behavior studies. These include latency to graze, total distance traveled, total time at standstill, number of defecations,

and number of vocalizations. With these additional traits, we should be able to improve upon the congruence between the startle trial and temperament questionnaire as latency to graze, defecation, and vocalization can be in response to fear (McGreevy 2004). To improve the measurability of all the behavior traits, we would need to setup additional cameras to cover the entire stall area. For our startle trial, we only utilized one camera which left an approximately six-foot square corner of the test stall unrecorded and prevented us from measuring the traits mentioned above. Another measure to consider recording during the startle trial is the eye blink reflex. The blink test is the best known test of brainstem excitability differences that can indicate physiological differences in the central nervous system (Valls-Sole 2012) and is a measure for startle response (Koch 1999). We could setup a slow motion camera near the startle testing area to record the eye blink and use the latency to blink as an additional dimension to improve our measure of startle response.

We genotyped all 16 horses from the startle trial for SNPs within the *NXPH2* gene we hypothesized would play a role in startle behavior. Prior work in our lab identified the region as potentially associated with behavior and startle response based on mouse models (Beglopoulos *et al.* 2005). We identified significant associations for two of the SNPs in the 3' UTR for the latency to touch the open umbrella, and one of the SNPs with the individual questionnaire item ratings. These results should be considered preliminary, due to the low number of horses genotyped (16 for the trial result, 10 for the questionnaire responses), but these results are promising and warrant genotyping horses used in future behavior trials.

Conclusions

The results of this study provide evidence for a genetic basis of temperament in the horse. We have demonstrated that within gaited breeds, there appear to be breed typical temperaments, and that there are key behavioral trait differences among sexes, gait types and equine disciplines across and within a breed. Our adapted owner questionnaire method also proved successful in generating

a phenotype for genetic association with temperament. Future adaptations of this work will aid in unraveling the mystery surrounding the nature versus nurture debate of behavior and temperament, and hopefully lead to a genetic test that may allow for more informed decisions and improved selection of horses in specific disciplines and roles. This will improve handler safety and lead to increased horse welfare.

References

Anderson MK, Friend TH, Evans JW, Bushong DM. 1999. Behavioral assessment of horses in therapeutic riding programs. Appl Anim Behav Sci 63(1):11-24.

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-5.

Baumgartner MR, Rabier D, Nassogne MC, Dufier JL, Padovani JP, Kamoun P, Valle D, Saudubray JM. 2005. Delta1-pyrroline-5-carboxylate synthase deficiency: Neurodegeneration, cataracts and connective tissue manifestations combined with hyperammonaemia and reduced ornithine, citrulline, arginine and proline. Eur J Pediatr 164(1):31-6.

Baumgartner MR, Hu CA, Almashanu S, Steel G, Obie C, Aral B, Rabier D, Kamoun P, Saudubray J-, Valle D. 2000. Hyperammonemia with reduced ornithine, citrulline, arginine and proline: A new inborn error caused by a mutation in the gene encoding Delta1-pyrroline-5-carboxylate synthase. Hum Mol Genet 9(19):2853-8.

Beaver BV and Amoss MS. 1982. Aggressive behavior associated with naturally elevated serum testosterone in mares. Applied Animal Ethology 8(5):425-8.

Beglopoulos V, Montag-Sallaz M, Rohlmann A, Piechotta K, Ahmad M, Montag D, Missler M. 2005. Neurexophilin 3 is highly localized in cortical and cerebellar regions and is functionally important for sensorimotor gating and motor coordination. Mol Cell Biol 25(16):7278-88.

Braff DL, Grillion C, Geyer MA. 1992. Gating and habituation of the startle reflex in schizophrenic patients. Arch Gen Psychiatry 49(3):206.

Buckley P, Dunn T, More SJ. 2004. Owners' perceptions of the health and performance of pony club horses in australia. Prev Vet Med 63(1-2):121-33.

Cadenhead KS, Geyer MA, Braff DL. 1993. Impaired startle prepulse inhibition and habituation in pati*ents with schizotypal personality disorder*
. The American Journal of Psychiatry 150(12):1862-7.

Cadenhead KS, Carasso BS, Swerdlow NR, Geyer MA, Braff DL. 1999. Prepulse inhibition and habituation of the startle response are stable neurobiological measures in a normal male population. Biol Psychiatry 45(3):360-4.

Cook D, Gallagher PC, Bailey E. 2010. Genetics of swayback in american saddlebred horses. Animal Genetics 41:64-71.

de Diego-Otero Y, Romero-Zerbo Y, Bekay Re, Decara J, Sanchez L, Fonseca FR, Arco-Herrera Id. 2009. A-tocopherol protects against oxidative stress in the fragile X knockout mouse: An experimental therapeutic approach for the Fmr1 deficiency. Neuropsychopharmacology 34(4):1011-26.

de Moor MH, Costa PT, Terracciano A, Krueger RF, de Geus EJ, Toshiko T, Penninx BW, Esko T, Madden PA, Derringer J, *et al.* 2012. Meta-analysis of genome-wide association studies for personality. Mol Psychiatry 17(3):337-49.

Dell'Acqua ML, Smith KE, Gorski JA, Horne EA, Gibson ES, Gomez LL. 2006. Regulation of neuronal PKA signaling through AKAP targeting dynamics. Eur J Cell Biol 85(7):627-33.

Farren MR, Carlson LM, Netherby CS, Lindner I, Li P-, Gabrilovich DI, Abrams SI, Lee KP. 2014. Tumor-induced STAT3 signaling in myeloid cells impairs dendritic cell generation by decreasing PKC II abundance. Science Signaling 7(313):ra16.

French JM. 1993. Assessment of donkey temperament and the influence of home environment. Appl Anim Behav Sci 36(2-3):249-57.

Geyer MA and Braff DL. 1982. Habituation of the blink reflex in normals and schizophrenic patients. Psychophysiology 19:1-6.

Glenske K, Prinzenberg EM, Brandt H, Gauly M, Erhardt G. 2011. A chromosome-wide QTL study on BTA29 affecting temperament traits in german angus beef cattle and mapping of DRD4. Animal 5(2):195-7.

Grados M, Sung HM, Kim S, Srivastava S. 2014. Genetic findings in obsessive-compulsive disorder connect to brain-derived neurotropic factor and mammalian target of rapamycin pathways: Implications for drug development. Drug Development Research 75(6):372-83.

Graham FK. 1975. The more or less startling effects of weak prestimulation. Psychophysiology 12:238-48.

Groves PM and Thompson RF. 1970. Habituation: A dual-process theory. Psychological Review 77:419-50.

Gutierrez-Gil B, Ball N, Burton D, Haskell M, Williams JL, Wiener P. 2008. Identification of quantitative trait loci affecting cattle temperament. J Hered 99(6):629-38.

Haskell MJ, Simm G, Turner SP. 2014. Genetic selection for temperament traits in dairy and beef cattle. Front Genet 5:368.

Hauseberg A and Muller C. 2002. A brief note on some possible factors involved in the reactions of horses to humans. Applied Animal Behaviour Science 76:339-44.

Hendricks B. 1995. International encyclopedia of horse breeds. First edition ed. Norman, OK: University of Oklahoma Press.

Hoffman HS and Ison JR. 1980. Reflex modifications in the domain of startle. I. some empirical findings and their implications for how the nervous system processes sensory input. Psychological Review 87:175-89.

Kajimoto Y, Shirakawa O, Lin X-, Hashimoto T, Kitamura N, Murakami N, Takumi T, Maeda K. 2003. Synpase-associated protein 90/postsynaptic density-95-associated protein (SAPAP) is expressed differentially in phencyclidine-treated rats and is increased in the nucleu accumbens of patients with schizophrenia. Neuropsychopharmacology 28:1831-9.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42(4):348-54.

Kawashima N, Takamiya K, Sun J, Kitabatake A, Sobue K. 1997. Differential expression of isoforms of PSD-95 binding protein (GKAP/SAPAP1) during rat brain development. FEBS Letters 418:301-4.

Kim E, Naisbitt S, Hsueh YP, Rao A, Rothschild A, Craig AM, Sheng M. 1997. GKAP, a novel synaptic protein that interacts with the guanylate kinase-like domain of the PSD-95/SAP90 family of channel clustering molecules. The Journal of Cell Biology 136(3):669-78.

Koch M. 1999. The neurobiology of startle. Prog Neurobiol 59(2):107-28.

Lee BK, Yoon JS, Lee MG, Jung YS. 2014. Protein kinase C-beta mediates neuronal activation of na(+)/H(+) exchanger-1 during glutamate excitotoxicity. Cell Signal 26(4):697-704.

Legeza B, Balazs Z, Nashev LG, Odermatt A. 2013. The microsomal enzyme 17beta-hydroxysteroid dehydrogenase 3 faces the cytoplasm and uses NADPH generated by glucose-6-phosphate dehydrogenase. Endocrinology 154(1):205-13.

Li JM, Lu CL, Cheng MC, Luu SU, Hsu SH, Chen CH. 2013. Genetic analysis of the DLGAP1 gene as a candidate gene for schizophrenia. Psychiatry Research 205(1-2):13-7.

Lloyd AS, Martin JE, Bornett-Gauci HLI, Wilkinson RG. 2008. Horse personality: Variation between breeds. Applied Animal Behaviour Science 112:369-83.

Lloyd AS, Martin JE, Bornett-Gauci HLI, Wilkinson RG. 2007. Evaluation of a novel method of horse personality assessment: Rater-agreement and links to behaviour. Appl Anim Behav Sci 105(1-3):205-22.

Mader DR and Price EO. 1980. Discrimination-learning in horses: Effects of breed age and social-dominance. Journal of Animal Science 50:962-5.

McGreevy P. 2004. Equine behavior. A guide for veterinarians and equine scientists. 1st ed. London, UK: Saunders Elsevier.

Momozawa Y, Kusunose R, Kikusui T, Takeuchi Y, Mori Y. 2005. Assessment of equine temperament questionnaire by comparing factor structure between two separate surveys. Appl Anim Behav Sci 92(1-2):77-84.

Momozawa Y, Ono T, Sato F, Kikusui T, Takeuchi Y, Mori Y, Kusunose R. 2003. Assessment of equine temperament by a questionnaire survey to caretakers and evaluation of its reliability by simultaneous behavior test. Appl Anim Behav Sci 84(2):127-38.

Morris PH, Gale A, Duffy K. 2002. Can judges agree on the personality of horses? Personality and Individual Differences 33(1):67-81.

Mukherjee O, Meera P, Ghosh S, Kubendran S, Kiran K, Manjunath KR, Subhash MN, Benegal V, Brahmachari SK, Majumder PP, *et al.* 2006. Evidence of linkage and asociation on 18p11.2 for psychosis. American Journal of Medical Genetics Part B, Neuropsychiatric Genetics 141B:868-73.

Munafò MR, Freimer NB, Ng W, Ophoff R, Veijola J, Miettunen J, Järvelin M, Taanila A, Flint J. 2009. 5-HTTLPR genotype and anxiety-related personality traits: A meta-analysis and new data. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 150B(2):271-81.

Overall KL. 1998. Self-injurious behavior and obsessive-compulsive disorder in domestic animals. In: Psychopharmacology of animal behavior disorders. Dodman NH and Shuster L, editors. Malden, MA: Blackwell Sciences.

Parwani A, Duncan EJ, Bartlett E, Madonick SH, Efferen TR, Rajan R, Sanfilipo M, Chappell PB, Chakravorty S, Gonzenbach S, *et al.* 2000. Impaired prepulse inhibition of acoustic startle in schizophrenia. Biol Psychiatry 47(7):662-9.

Patel A, Rees SD, Kelly MA, Bain SC, Barnett AH, Thalitaya D, Prasher VP. 2011. Association of variants within APOE, SORL1, RUNX1, BACE1 and ALDH18A1 with dementia in alzheimer's disease in subjects with down syndrome. Neurosci Lett 487(2):144-8.

Pervin LA and John OP. 1997. Personality theory and research. New York: John Wiley and Sons.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, *et al.* 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 81:559-75.

Purves D. 2008. Neuroscience. 4th ed. Sunderland, Massachusetts: Sinauer Associates, Inc.

Quednow BB, Kühn K, Beckmann K, Westheide J, Maier W, Wagner M. 2006. Attenuation of the prepulse inhibition of the acoustic startle response within and between sessions. Biol Psychol 71(3):256-63.

Rozen S and Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics methods and protocols: Methods in molecular biology. Krawetz S and Misener S, editors. Totowa, NJ: Humana Press. 365 p.

Segurado R, Detera-Wadleight SD, Levinson DF, Lewis CM, Gill M, Nurnberger Jr. JI, Craddock N, DePaulo JR, Baron M, Gershon ES, *et al.* 2003. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. American Journal of Human Genetics 73:49-62.

Service SK, Verweij KJH, Lahti J, Congdon E, Ekelund J, Hintsanen M, Räikkönen K, Lehtimäki T, Kähönen M, Widen E, *et al.* 2012. A genome-wide meta-analysis of association studies of cloninger's temperament scales. Translational Psychiatry 2(5):e116.

Siddle DA and Kroese BS. 1985. Orienting, habituation, and short-term memory. Psychophysiology 22:535-44.

Søndergaard E and Ladewig J. 2004. Group housing exerts a positive effect on the behaviour of young horses during training. Appl Anim Behav Sci 87(1-2):105-18.

Staiger EA, Bellone RR, Sutter NB, Brooks SA. 2011. Genome-wide association of polymorphic gait in the horse. Journal of Animal Science 89(E-Suppl. 1):321.

Stewart SE, Yu D, Scharf JM, Neale BM, Fagerness JA, Mathews CA, Arnold PDea. 2013. Genome-wide association study of obsessive-compulsive disorder. Molecular Psychiatry 18(7):788-98.

Swerdlow NR, Paulsen J, Braff DL, Butters N, Geyer MA, Swenson MR. 1995. Impaired prepulse inhibition of acoustc and tactile startle response in patients with huntingtn's disease. Journal of Neurology, Neurosurgery, and Psychiatry 58:192-200.

Takeuchi M, Hata Y, Hirao K, Toyoda A, Irie M, Taki Y. 1997. SAPAPs: A famil of PSD-95/SAP90-associated proteins localized at postsynaptic density. Journal of Biological Chemistry 272:11943-51.

Terracciano A, Sanna S, Uda M, Deiana B, Usala G, Busonero F, Maschio A, Scally M, Patriciu N, Chen WM, *et al.* 2010. Genome-wide association scan for five major dimensions of personality. Mol Psychiatry 15(6):647-56.

Breeding Standards [Internet] [cited 2014 . Available from: http://www.icelandics.org/standards.php .

Valls-Sole J. 2012. Assessment of excitability in brainstem circuits mediating the blink reflex and the startle reaction. Clin Neurophysiol 123(1):13-20.

Verweij KJ, Zietsch BP, Medland SE, Gordon SD, Benyamin B, Nyholt DR, McEvoy BP, Sullivan PF, Heath AC, Madden PA, *et al.* 2010. A genome-wide association study of cloninger's temperament scales: Implications for the evolutionary genetics of personality. Biol Psychol 85(2):306-17.

Visser E, Vanreenen C, Van der Werf J, Schilder M, Knaap J, Barneveld A, Blokhuis H. 2002. Heart rate and heart rate variability during a novel object test and a handling test in young horses. Physiol Behav 76(2):289-96.

Visser EK, van Reenen CG, Rundgren M, Zetterovist M, Morgan K, Blokhuis HJ. 2003. Responses of horses in behavioural tests correlate with temperament assessed by riders. Equine Veterinary Journal 35:176-83.

Ziegler L. 2005. Easy-gaited horses. First edition ed. North Adams, MA: Storey Publishing.

CHAPTER 5

SUMMARY AND CONCLUSIONS

"Wherever man has left his footprint in the long ascent from barbarism to civilization we will find the hoofprint of the horse beside it."
~John Moore

Few species have had the same impact on humans as the horse subsequent to domestication. While initial contact with the horse was for the purpose of consumption, the recognition of the horse as a beast of burden fueled the spread of human cultures and moved the horse into a position of prestige. Like other domesticated species, the horse underwent artificial selection by man in order to enhance traits viewed as desirable for the different roles the horse was asked to fill. Selection would have focused on physical attributes such as size and strength (Clutton-Brock 1999), desirable behaviors (Hislop 1992; Houpt and Kusunose 2000; Lloyd *et al.* 2008), and locomotion.

Locomotion is one of the most valuable traits under selection in the horse, and in comparison to other domesticated livestock, is one of the more unique traits in the horse. Horses can perform a range of footfall patterns from a two-beat lateral gait ("pace") to a two-beat diagonal gait ("trot"), including a variety of four-beat diagonal and lateral gaits (Harris 1993). This gait plasticity is an inherent trait apparent from birth and is not found in any other mammalian species to the same extent (Hildebrand 1989). However, polymorphic gait is not fixed among the equids, or within *Equus caballus*. Not all individual horses can perform the full range of intermediate gaits. Therefore, any horse with the innate ability to perform any intermediate gait other than the trot (two-beat diagonal) is commonly called a "gaited" horse, regardless of the official registration of the horse. Locomotor pattern is controlled by neural networks in the spinal cord called central pattern generators (CPGs). Studies using laboratory animal models such as mice, cats, and lampreys (Kiehn 2006; Kuo 2002) have characterized the walk, trotting, swimming, and "galloping" gaits, but not lateral gaits like those performed by the gaited horse. So what factors are contributing to the locomotor plasticity of the horse?

Biomechanical studies have shown that conformation plays a considerable role in horse locomotion (Leach and Dagg 1983), but none have yet been able to identify a musculoskeletal trait linked to gait type preference. Abnormal locomotion and temperament in man are often characterized together in complex disorders such as attention deficit hyperactivity disorder (ADHD) and autism (Hildebrand 1965; Kiehn 2006; Kiehn *et al.* 2010; Kuo 2002), implying behavior and locomotion control may share genetic factors. Recently, a functional mutation in *DMRT3* (an isoform of the doublesex and mab-3 related transcription factor) on ECA23 was identified as permissive for pacing in the Icelandic horse (Andersson *et al.* 2012) and appears at varying frequencies in gaited and non-gaited breeds (Promerová *et al.* 2014). While most of the gaited breeds were fixed for the mutation, some gaited breeds, such as the Mangalarga Marchador, were not (Patterson, Staiger, Brooks 2014; Promerová *et al.* 2014). To further complicate the matter, trotting Standardbreds have high frequencies of the mutation (Andersson *et al.* 2012), suggesting a functional role in speed regulation is better suited for *DMRT3*. With this dissertation, we start to shed some light on the genetics of polymorphic gait in the horse, and the influences that conformation and behavior may hold on polymorphic gait.

Multiple Loci, in Addition to DMRT3, Contribute to Polymorphic Gait in the Horse.

In Chapter 2, we describe two different approaches to identify genetic loci contributing to gait type. First, we ran a genome wide association study for gait type across multiple breeds, including three breeds known to discretely segregate for gait type. For this study, we wanted to see if we would still identify the *DMRT3* locus, and if there were other loci contributing to gait type. The results from this work did identify the *DMRT3* locus as contributing to gait type, but also four additional loci on ECA1, ECA7, ECA11, and ECA24 that are statistically more significant than the *DMRT3* locus. Notably, the marker identified on ECA11 lies in the gene *PIRT* which encodes a protein specifically expressed in sensory neurons (Kim *et al.* 2008; Patel *et al.* 2011). *PIRT* has been

implicated in playing a role in itch sensation (Patel *et al.* 2011) and heat pain (Tang *et al.* 2013), with *PIRT* -/- mice exhibiting decreased behavioral responses to cold and cool temperatures (Tang *et al.* 2013), indicating this gene may play a role in the efficiency of signal transduction from sensory neurons to aid in movement control.

While several of the intermediate speed four-beat gaits are shared across the gaited breeds, not all breeds are able to perform the same gaits and there are breed-specific deviations (Nicodemus and Clayton 2003). For example, the Tennessee Walking Horse (TWH) is renowned for the ability to perform the 'running walk', an even-timed four-beat intermediate gait characterized by both a headshake and over-stride. Yet the breed is still able to perform the whole spectrum of intermediate gaits (Tennessee Walking Horse Breeders' and Exhibitors' Association 2011) despite being nearly fixed for the *DMRT3* mutation (Andersson *et al.* 2012). The second study in Chapter 2 was aimed to investigate the genetic contribution to gait specifically within the TWH. For this study, we ran a genome wide association study utilizing 97 horses phenotyped by slow-motion video analysis into two gait type categories. The results from this work identified a suggestive candidate locus on ECA19 after permutation. None of the detected loci surpassed a Bonferroni significance, likely due to the small sample size, imprecision in phenotyping by a presence/absence scheme and confounding effects of environment. However, several of our markers fell within or near genes involved in myogenesis (*FBXO40* and *SMTNL2*) and brain development (*ARGFX*), both critical for locomotion. Muscles initiate and control movement through innervation by motor neurons, which originate from the spinal cord but also receive projections from neurons in the brainstem and midbrain (Yuste *et al.* 2005).

To verify the results of the GWAs and to identify additional novel causal variants, we filtered through polymorphisms of whole genome sequences of six horses from phylogenetically diverse breeds available from another project in our lab. For three of the horses, a single observer (EAS)

phenotyped the horses as gaited from available video footage of each horse traveling at intermediate speed. From this data, we were able to identify 12 non-synonymous SNPs within the *DTX3L* gene bordering the candidate region on ECA19; all 12 SNPs were only present in the three gaited individuals. While the large number of SNPs in a single gene may be a reflection of a poor gene prediction model, the occurrence in only the gaited individuals is intriguing. Indeed occurrence of 12 such non-synonymous and polymorphic SNPs within a span of just 1918 bp is a rare event (we only observe 18 other non-synonymous SNPs in the entire 10 million SNP dataset). *DTX3L* is a member of a family of ubiquitin ligases with known roles in Notch signaling (Holleman and Marchese 2014; Matsuno *et al.* 1995; Matsuno *et al.* 2002; Matsuno *et al.* 1998; Takeyama *et al.* 2003; Yamada *et al.* 2011); Notch signaling is an important signaling pathway and is involved in neurogenesis and neuronal differentiation (Aster 2014; Purves 2008). This gene may play an important role in neurotransmitter release and neuronal differentiation in gaited horses, potentially allowing for polymorphic gait. Accumulation of so many non-synonymous mutations may also indicate that a loss of function mutation was acquired by this *DTX3L* allele a very long time ago (Halliburton 2004).

Additionally, we identified a novel stop-gained SNP on ECA7 in an ENSEMBL predicted gene model for *ANKK1* (ankyrin repeat and kinase domain containing 1). The *ANKK1* gene is involved in signal transduction pathways and is part of the dopaminergic reward system (Neville, Johnstone, Walton 2004), indicating a potential tie between gait and behavior. While the determination of polymorphisms from next generation sequencing datasets is not perfect and requires validation in a larger sample sets, these findings are exciting and warrant further investigation.

Conformation Is Different in Gaited Horses, but Contributes Only to the Quality of the Gait

While there is work left to be done to conclusively define causative loci for gait, we were able to identify conformation traits unique to gaited horses and map skeletal size variation in the Tennessee Walking Horse (TWH). We used body measures from several hundred gaited and non-gaited horses, as described in Chapter 3, to analyze conformation and identify body shape patterns. After accounting for sexual dimorphism, gaited horses have significantly longer, but narrower heads, longer forearms and front pasterns, longer cannons, thinner coronets and hind pasterns, and longer hind hooves than non-gaited horses. Gaited horse trainers have long reported that gaited horses have proportionately longer hind limbs (Lane 2011; London 2012; Ziegler 2005), but none have ever noted the longer forearm of the gaited breeds. Within a single breed, we were unable to identify measurements that were associated with a broad grouping of gait type, but we were able to identify morphological differences between horses used for different training disciplines. Taken in conjunction with the across-breed data, this indicates that morphology likely contributes to the quality of the gait, but is not necessarily permissive for gait type.

As a side project to investigating gait, we interrogated the genome for loci that might contribute to the wide variation in size that occurs in the TWH, as they range in skeletal size from 56 to 68 inches at the withers. Utilizing principle component analysis, we were also able to identify two body shape patterns that describe overall size and thickness within the TWH and matches the patterns previously observed across diverse breeds. Additionally, we identified a third body shape pattern in TWH that explains relative back size length in relation to other body lengths; this body shape pattern was unique to the TWH. In 105 TWH, we were able to map the PC1-size scores to the *LCORL/NCAPG* region on ECA3. Previous studies have identified *LCORL* as a candidate locus for overall size and wither height in horses (Boyko *et al.* 2014; Distl *et al.* 2011; Makvandi-Nejad *et al.* 2012; Metzger *et al.* 2013; Signer-Hasler *et al.* 2012). However a causal variant has not

been discovered due to the long extent of linkage disequilibrium (LD) in the region in other breeds (Boyko *et al.* 2014) and to a highly homologous retrogene copy of *LCORL* on another chromosome (Baird, Raudsepp, Brooks, unpublished data ; Distl *et al.* 2011). Unlike other breeds such as the Thoroughbred, we found the length of LD in the region to be much shorter in the TWH. Smaller LD blocks in this region point suggest the TWH as an ideal candidate breed for fine-mapping of the region and identification of the causal variant in *LCORL* responsible for size variation.

In Gaited Breeds, Temperament Variation Is Related More to Training Discipline than Gait Type

Behavior or temperament is a key criterion in the selection of an animal like the horse that is often in close contact with humans. In Chapter 4, we adapted a questionnaire from Momozawa *et al.* (2005) and surveyed over 500 horses from 20 different gaited breeds. We utilized PCA to identify the unknown behavior structure in our data, and then used factor analysis (FA) to correct for skewed responses, but modeled on the PCA structure assumptions. FA and PCA are both variable reduction techniques, but FA assumes that the covariation in the observed variables is due to the presence of underlying causal structure (O'Rourke and Hatcher 2013). From the FA we were able to identify four temperament patterns both across breeds and within a single breed (TWH). Factor analysis reduced produced traits of 'neophobia' (FC1), 'trainability' (FC2), 'independence' (FC3), and 'hostility' (FC4).

Probably one of the more interesting results of chapter 4 was the higher average excitability score in multi-gaited horses versus those were unable to trot. Our assumption at the start of the study was that horses able to trot would be rated as more excitable, especially if we consider typically non-gaited trotting breeds, such as Arabians and Thoroughbreds, who are stereotypically considered volatile. Expansion of the study to include horses from more diverse breeds, including Arabians and Thoroughbreds, would be interesting to refine this discovery. However, the survey measures

the owner's perception of their horses temperament, so there is the potential for Thoroughbred owners to become desensitized toward their horse's behavior. For this future work it would be best to target equine professionals with vast breed experience. This was the only significant result with gait type, but comparison with disciplines highlighted more temperament differences. These identified traits were appropriate considering the situations each discipline is placed in. For example, trail horses were identified with more stoic trait scores, an appropriate finding considering trail horses are often exposed to fearful situations, as would be perceived by a prey animal like the horse. Show horses were rated with increased attention, patience, and self-reliance, yet also with increased excitability. The first few traits are important for safety on the show grounds, while the latter trait would be a beneficial counterpoint in the show ring, where it is more desirable to appear lively, animated, and expressive.

As the TWH was the most widely surveyed of the gaited breeds, it provided a good sample set for identification of within-breed temperament variation, and to subsequently map this variation as a temperament trait. The within breed analysis of the TWH revealed similar temperament factor patterns of 'neophobia' (FC1), 'trainability' (FC2), but a swapping in the last two factors: 'hostility' as FC3 and 'independence' as FC4. This suggests that either of the first two factors are stable across breeds, and that the remaining factors are breed dependent. Alternatively, this change in trait ranking may be due to the relative over representation of the TWH breed (n=211 out of 455 total) in the full, across-breed dataset. We did consider balancing for breed, but with only two representatives from some breeds the dataset would be very small and under represent the breeds' characteristics. The expansion of this work to include more diverse individuals would clarify this point.

Following GWAs using the top three factor traits and 113 phenotyped horses, we were able to identify some suggestive candidate loci contributing to the first three temperament factors in

TWH. While a small sample size likely hindered statistical power in this study, the identified candidate loci offer provocative functional hypothesis for this these traits. Factor 1, "neophobia", appeared to be mostly driven by age and environmental influences, and returned the most ambiguous results likely reflecting the many genetic or gene by environment components at play. Yet, factor 1 did identify candidate markers in genes with roles in visual and neurodegeneration, and in steroidogenesis. Factor 2, "trainability", mapping provided the best statistical significance; the top candidate marker fell within *PRKCB,* may regulate neuronal function and correlates with fear-induced conflict behavior after stress in mice (de Diego-Otero *et al.* 2009). In horses, this could correlate to horse perception of training and ability to overcome their flight instincts when they are afraid. The most significant marker for Factor 3-hostility falls within a gene that mediates ligand-gated ion channels important in sensory transduction processes (Purves 2008) and post-synaptic signaling pathways (Dell'Acqua *et al.* 2006). Differences in social hostility may be due to changes in the strength of the neurotransmitter and synaptic transmission, potentially changing how horses perceive each other and influencing their interactions.

Experimental Measures of Startle Response Suggest an Interaction between Behavior and Gait Type

A behavior questionnaire is a subjective measure of temperament variation, therefore we also applied an objective experimental startle trial utilizing a small subset of young horses, as described in Chapter 4. We applied four different types of tests: a tactile pressure test, an auditory test, a novel stimulus test, and a novel startle test. Across all tests, we were able to capture variable responses, with differences in gait type associated with some of the results of the startle test. This preliminary startle trial shows the potential feasibility of modifying protocols for measuring aspects of startle reflexes in the horse, including pre-pulse inhibition (PPI). The murine PPI experimental model utilizes a distinctive non-startling tone followed by a strong startling tone. A normal PPI response the weak tone reduces the magnitude of the reaction to the startle tone (Quednow *et al.*

2006). Abnormal PPI can be increases or decreases in the startle response to the startle tone (Quednow *et al.* 2006). Abnormal PPIs are associated with neuropsychiatric disorders such as schizophrenia and obsessive compulsive disorder (Braff, Grillion, Geyer 1992; Geyer and Braff 1982; Parwani *et al.* 2000; Swerdlow *et al.* 1995). In young horses, differences in PPI could allude to changes underlying their sensory reactivity and information processing, which may manifest in different gait type abilities and behavior traits. With a larger sample size, we will be able to draw more concise conclusions about the feasibility of utilizing PPI responses as a method for identifying underlying neural circuits correlated with these traits.

Overall Significance and Impact

Due to the complex nature of behavioral traits like locomotion and temperament, with their continuous phenotype and multiple environmental factors, genetic mapping of causal genes is a difficult task. At a rudimentary level of phenotyping, we demonstrate with the current genotyping technology that there are detectable genetic differences both within and across breeds segregating for these traits, and we present some convincing candidate genes that warrant further investigation. With a larger sample size, a more objective and quantitative phenotype (for gait type), and additional markers provided by a denser SNP array, used in conjunction with data available from whole genome sequencing, it will be possible to map the causal variants for these traits with improved precision. This will likely occur within the next few years, as improvements are made in the annotation of the equine reference genome and with the imminent release of high density equine SNP chips and broader use of genotyping by sequencing technologies (GBS).

These causal variants will provide genetic tests that will predict gait type, conformation, and behavior traits in horses. It is extremely important to horse owners and breeders to utilize all available tools in their breeding management decisions, as only a small number of horses are being used for the production of the highest priced animals (totaling ~$33.2 billion in 2005) in the horse

industry (American Horse Council 2005). The selection of any horse is based upon conformation, temperament, and gait (Lloyd *et al.* 2008); however, this selection process is currently practiced more art form than science. Lack of precision in selection can lead to poor performing horses, resulting in the loss of both time and money for the breeder. Providing genetic tests that predict different gait types, conformation, and temperament will allow breeders to make better management decisions and improve the marketability of their horses. The genetic tests should also improve animal welfare as the horses are selected for and used in more appropriate equine disciplines.

Beyond the impact to horse owners, studying horse gait, morphology, and temperament can aid in human disease research and engineering of robotics circuitry. Animal models have long played an important role in human disease research due to similarities in basic biology, physiology and experimental convenience. The horse, as a large, long-lived, diurnal mammal, shares many biological and physiological processes with humans. Horses have complex social interactions and a greater level of complexity in their central nervous system than laboratory model systems; plus, they display social stress, compulsive behaviors, and pathological aggression (McGreevy 2004), all akin to human disorders. Inherent variation in temperament and reactivity traits in the horse could serve as models for human disorders and diseases, such as schizophrenia, attention deficit hyperactivity disorder, and obsessive compulsive disorders. Identification of the causal variants influencing gait type, morphology, and temperament in the horse will highlight interactions between these different systems, and provide more information on the pathology of the underlying disorders, perhaps leading to novel treatments for these complex conditions in both man and horse.

References

National Economic Impact of the U.S. Horse Industry. [Internet]; c2005 [cited 2013 . Available from: www.horsecouncil.org/national-economic-impact-us-horse-industry .

Andersson LS, Larhammar M, Memic F, Wootz H, Schwochow D, Rubin CJ, Patra K, Arnason T, Wellbring L, Hjalm G, *et al.* 2012. Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. Nature 488(7413):642-6.

Aster JC. 2014. In brief: Notch signalling in health and disease. J Pathol 232(1):1-3.

Baird TL, Raudsepp T, Brooks SA. 2012. Copy number variation of a novel retrogene for *LCORL* in the horse. .

Boyko AR, Brooks SA, Behan-Braman A, Castelhano M, Corey E, Oliveira KC, Swinburne JE, Todhunter RJ, Zhang Z, Ainsworth DM, *et al.* 2014. Genomic analysis establishes correlation between growth and laryngeal neuropathy in thorroughbreds. BMC Genomics 15:259.

Braff DL, Grillion C, Geyer MA. 1992. Gating and habituation of the startle reflex in schizophrenic patients. Arch Gen Psychiatry 49(3):206.

Clutton-Brock J. 1999. A natural history of domesticated mammals. Cambridge, UK: Cambridge University Press.

de Diego-Otero Y, Romero-Zerbo Y, Bekay Re, Decara J, Sanchez L, Fonseca FR, Arco-Herrera Id. 2009. A-tocopherol protects against oxidative stress in the fragile X knockout mouse: An experimental therapeutic approach for the Fmr1 deficiency. Neuropsychopharmacology 34(4):1011-26.

Dell'Acqua ML, Smith KE, Gorski JA, Horne EA, Gibson ES, Gomez LL. 2006. Regulation of neuronal PKA signaling through AKAP targeting dynamics. Eur J Cell Biol 85(7):627-33.

Distl O., Schröder W., Dierks C. and Klostermann A. 2011. Genome-wide association studies for performance and conformation traits in hanoverian warmblood horses.
. 9th dorothy russel havemeyer foundation, international equine genome mapping workshopOak Ridge Conference Center, Chaska, Minnesota: . 12 p.

Geyer MA and Braff DL. 1982. Habituation of the blink reflex in normals and schizophrenic patients. Psychophysiology 19:1-6.

Halliburton R. 2004. Introduction to population genetics
. Upper Saddle River, NJ: Pearson/Prentice Hall.

Harris SE. 1993. Horse gaits, balance, and movement. New York, NY: Howell Book House.

Hildebrand M. 1989. The quadrupedal gaits of vertebrates. Bioscience 39(11):766-75.

Hildebrand M. 1965. Symmetrical gaits of horses. Science 150:701-8.

Hislop J. 1992. Breeding for racing. London, UK: The Kingswood Press.

Holleman J and Marchese A. 2014. The ubiquitin ligase deltex-3l regulates endosomal sorting of the G protein-coupled receptor CXCR4. Mol Biol Cell 25(12):1892-904.

Houpt KA and Kusunose R. 2000. Genetics of behaviour. In: The genetics of the horse. Bowling AT and Ruvinsky A, editors. Oxon, UK: CABI. 281 p.

Kiehn O. 2006. Locomotoer circuits in the mammalian spinal cord. Annu Rev Neruosci 29:279-306.

Kiehn O, Dougherty KJ, Hagglund M, Borgius L, Talpalar A, Restrepo CE. 2010. Probing spinal circuits controlling walking in mammals. Biochemical and Biophysical Research Communications 396:11-8.

Kim AY, Tang Z, Liu Q, Patel KN, Maag D, Geng Y, Dong X. 2008. Pirt, a phosphoinositide-binding protein, functions as a regulatory subunit of TRPV1. Cell 133(3):475-85.

Kuo AD. 2002. The relatie roles of feedforward and feedback in the control of rhythmic movements. Motor Control 6:129-45.

Lane G. 2011. Discussion with author in september. .

Leach DH and Dagg AI. 1983. A review of research on equine locomotion and biomechanics. Equine Veterinary Journal 15(2):93-102.

Lloyd AS, Martin JE, Bornett-Gauci HLI, Wilkinson RG. 2008. Horse personality: Variation between breeds. Applied Animal Behaviour Science 112:369-83.

London J. 2012. Discussion with author in october. .

Makvandi-Nejad S, Hoffman GE, Allen JJ, Chu E, Gu E, Chandler AM, Loredo AI, Bellone RR, Mezey JG, Brooks SA, *et al.* 2012. Four loci explain 83% of size variation in the horse. PLos One 7(7):e39929.

Matsuno K, Diederich RJ, Go MJ, Blaumueller CM, Artavanis-Tsakonas S. 1995. Deltex acts as a positive regulator of notch signaling through interactions with the notch ankyrin repeats. Development 121:2633-44.

Matsuno K, Ito M, Hori K, Miyashita F, Suzuki S, Kishi N, Artavanis-Tsakonas S, Okano H. 2002. Involvement of a proline-rich motif and RING-H2 finger of deltex in the regulation of notch signaling. Development 129(4):1049-59.

Matsuno K, Eastman D, Mitsiades T, Quinn AM, Carcanciu ML, Ordentlich P, Kadesch T, Artavanis-Tsakonas S. 1998. Human deltex is a conserved regulator of notch signalling. Nat Genet 19(1):74-8.

McGreevy P. 2004. Equine behavior. A guide for veterinarians and equine scientists. 1st ed. London, UK: Saunders Elsevier.

Metzger J, Schrimpf R, Philipp U, Distl O. 2013. Expression levels of LCORL are associated with body size in horses. PLos One 8(2):e56497.

Momozawa Y, Kusunose R, Kikusui T, Takeuchi Y, Mori Y. 2005. Assessment of equine temperament questionnaire by comparing factor structure between two separate surveys. Appl Anim Behav Sci 92(1-2):77-84.

Neville MJ, Johnstone EC, Walton RT. 2004. Identification and characterization of ANKK1: A novel kinase gene closely linked to DRD2 on chromosome band 11q23.1. Hum Mutat 23(6):540-5.

Nicodemus MC and Clayton HM. 2003. Temporal variables of four-beat, stepping gaits of gaited horses. Applied Animal Behaviour Science 80:133-42.

O'Rourke N and Hatcher L. 2013. Introduction: The basics of principle component analysis. In: A step-by-step approach to using SAS for factor analysis and structural equation modeling. 2nd ed. Cary, North Carolina: SAS Institute, Inc. 1 p.

Parwani A, Duncan EJ, Bartlett E, Madonick SH, Efferen TR, Rajan R, Sanfilipo M, Chappell PB, Chakravorty S, Gonzenbach S, *et al.* 2000. Impaired prepulse inhibition of acoustic startle in schizophrenia. Biol Psychiatry 47(7):662-9.

Patel KN, Liu Q, Meeker S, Undem BJ, Dong X. 2011. Pirt, a TRPV1 modulator, is required for histamine-dependent and -independent itch. PLoS One 6(5):e20559.

Patterson L, Staiger EA, Brooks SA. 2014. DMRT3 is associated with gait type in mangalarga marchador horses, but does not control gait ability. Animal Genetics In Review.

Promerová M, Andersson LS, Juras R, Penedo MCT, Reissmann M, Tozaki T, Bellone R, Dunner S, Horín P, Imsland F, *et al.* 2014. Worldwide frequency distribution of the 'Gait keeper' mutation in theDMRT3gene. Anim Genet 45(2):274-82.

Purves D. 2008. Neuroscience. 4th ed. Sunderland, Massachusetts: Sinauer Associates, Inc.

Quednow BB, Kühn K, Beckmann K, Westheide J, Maier W, Wagner M. 2006. Attenuation of the prepulse inhibition of the acoustic startle response within and between sessions. Biol Psychol 71(3):256-63.

Signer-Hasler H, Flury C, Haase B, Burger D, Simianer H, Leeb T, Reider S. 2012. A genome-wide association study reveals loci influencing height and other conformation traits in horses. PLos One 7:e37282.

Swerdlow NR, Paulsen J, Braff DL, Butters N, Geyer MA, Swenson MR. 1995. Impaired prepulse inhibition of acoustc and tactile startle response in patients with huntingtn's disease. Journal of Neurology, Neurosurgery, and Psychiatry 58:192-200.

Takeyama K, Aguiar RC, Gu L, He C, Freeman GJ, Kutok JL, Aster JC, Shipp MA. 2003. The BAL-binding protein BBAP and related deltex family members exhibit ubiquitin-protein isopeptide ligase activity. J Biol Chem 278(24):21930-7.

Tang Z, Kim A, Masuch T, Park K, Weng H, Wetzel C, Dong X. 2013. Pirt functions as an endogenous regulator of TRPM8. Nat Commun 4:2179.

The Tennessee Walking Horse Breed: Gaits [Internet]; c2011 [cited 2014 . Available from: www.twhbea.com/breed/gait.php .

Yamada K, Fuwa TJ, Ayukawa T, Tanaka T, Nakamura A, Wilkin MB, Baron M, Matsuno K. 2011. Roles of drosophila deltex in notch receptor endocytic trafficking and activation. Genes to Cells 16(3):261-72.

Yuste R, MacLean JN, Smith J, Lansner A. 2005. Opinion: The cortex as a central pattern generator. Nature Reviews Neuroscience 6(6):477-83.

Ziegler L. 2005. Easy-gaited horses. First edition ed. North Adams, MA: Storey Publishing.