Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences
Department of Animal Breeding
and Genetics

The role of polymorphisms of the *MSTN*, *GRIN2B* and *DOCK8* genes in the performance of pace-racing Icelandic horses

Laura Bas Conn

Degree Project in Biology A2E / Swedish University of Agricultural Sciences Master's Thesis, 30 hp
Department of Animal Breeding and Genetics
Examensarbete / SLU, Institutionen för husdjursgenetik, nr 543
Uppsala 2018

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Veterinary Medicine and Animal Sciences

The role of polymorphisms of the MSTN, GRIN2B and DOCK8 genes in the performance of pace-racing Icelandic horses

Laura Bas Conn

Supervisors:
Gabriella Lindgren, SLU, Department of Animal Breeding and Genetics Susanne Eriksson, SLU, Department of Animal Breeding and Genetics Thorvaldur Árnason, International Horse Breeding Consultant (IHBC)

Examiner:
Åsa Viklund, SLU, Department of Animal Breeding and Genetics

Credits: 30 hp
Course title: Degree project in Biology
Course code: EX0802
Course coordinating department: Department of Biomedical Sciences and Veterinary Public Health
Programme: Degree project in Biology, offered as independent course
Level: Advanced, A2E
Place of publication: Uppsala
Year of publication: 2018
Cover Pictures: Laura Bas Conn, Landsmót 2018
Name of series: Examensarbete / SLU, Institutionen för husdjursgenetik, no 543
Online publishing: https://epsilon.slu.se

Keywords: Icelandic horse, SNP, temperament, pace racing, performance, gait

Abstract

This study investigated the association of single-nucleotide polymorphisms (SNPs) of three different genes with pace racing performance in the Icelandic horse. The SNPs chosen were PR3737, PR8604 and PR9482 of the Myostatin gene; SNPs g.41206762 C>T and g.41218272 T>C of the *GRIN2B* gene and SNP g.22496787 T>C of the *DOCK8* gene. Icelandic horses that compete frequently in any of the three pace race disciplines (P1 over 250m, P2 over 100m and P3 over 150m) were genotyped for all SNPs mentioned above. The horses in the sample (n=131) were divided into "Speed groups" based on consistency of racing times and training schemes. Three groups were created: Fast (n=37), Average (n=75) and Slow (n=19).

Differences in genotype distribution among the different Speed groups were evaluated using Fisher's Exact Test. Significant differences were found for all SNPs except PR9482 when comparing only Fast versus Slow horses. Significant differences were also found between all three groups when comparing allele frequencies. Moreover, a drastic difference in allele frequency of the mutant allele for the *DOCK8* SNP was found between horses of the Fast group (frequency of the mutant allele was of 0.70) and 280 Icelandic horses genotyped for other studies (mutant allele frequency ranging between 0.33 (n=95) and 0.43 (n=192)). This implies an important role of this SNP and its mutant allele in pace-racing performance. Additionally, horses were evaluated by owners in terms of three temperament traits: Nervousness, Focus and Motivation. Significant differences were found in the distribution of scores for these temperament traits between horses with the different genotypes of the *GRIN2B* SNPs, suggesting a beneficial effect of the mutant alleles on the horse's temperament.

An association analysis was also performed, where the genotypes were evaluated for association with life-time racing performance results for race times (P1, P2 and/or P3), Breeding Field Test (BFT) assessment scores for pace and estimated breeding values (EBV) for pace at BFTs. Additionally, the genotypes were evaluated for association with performance results obtained when the horses competed at the age of 13 years or under (n=117) and over 13 years of age (n=64). Significant associations were found for PR9482, where the C allele (stamina-related in other breeds) was beneficial to performance in the longer P1 races and the T allele (speed-related in other breeds) was beneficial to performance in P3 for horses of ages 13 and under. Associations were also found for the *GRIN2B* SNP, with the C (mutant) allele having a beneficial effect in the shorter P3 races for both lifetime career and performance in the age group of 13 years and under. As for *DOCK8*, associations were also found in relation to P2, with the mutant (T) allele having a positive effect on the performance. No associations were found with any of the studied polymorphisms with BFT assessment scores for pace or with estimated breeding values.

This study concludes that a SNP at the DOCK8 gene may well be beneficial to performance in pace-racing Icelandic horses due to the associations found with racing times and, more interestingly, the difference found in mutant allele frequencies between an elite group of racers and the rest of the population. Additionally, the study showcases the importance of temperament traits as factors that affect performance in pace-racing, as evidenced by the associations found between the mutant allele of the *GRIN2B* polymorphism and temperament scores, along with racing times. Furthermore, the study concludes that Myostatin SNPs have little effect in pace-racing performance, seeing as only one polymorphism affected performance and only in one of the age groups. Moreover, comparing mutant allele frequencies between the elite racers and a large number of Icelandic horses that were not selected for pace racing showed no variation at all. This further indicates these mutations may not play a major role in performance as they do in other breeds, such as the Thoroughbreds.

Table of contents

3-Introduction

Since the discovery of the DMRT3 mutation, where a single change from cytosine (C) to adenine (A) was found to be permissive for the ability to perform alternate gaits (Andersson et al., 2012), many studies have been conducted in the Icelandic horse (Kristjánsson et al., 2014; Jäderkvist et al., 2017) and other breeds (Jäderkvist et al., 2014a; Jäderkvist et al., 2014b; Jäderkvist et al. 2015, Amano et al., 2018; Novoa-Bravo et al., 2018) in order to understand the role of this mutation and other genetic factors that play a role in gaitedness. Virtually all individuals of the Icelandic horse breed present toelt in addition to walk, trot and canter/gallop. Not only that, but there is also the "flying" pace gait in five-gaited individuals, which are arguably the most valuable within the breed. While the majority of good five-gaited individuals are destined to compete in disciplines such as Five-gait, where pace is exhibited in addition to the other four gaits, the pace-racing disciplines are becoming increasingly popular.

Pace-racing in Icelandic horses is a rather unique sport discipline not only due to the fact that it's raced in an alternate gait, but also to the short distances that it's raced over, with a maximum distance of 250 m to cover. Such a short distance of racing can be comparable to the high explosiveness of human sprinting, where the muscles are demanded to work at high intensity over a short time and draw their power mainly from anaerobic metabolism. This is true also for the Icelandic horses, as a study performed by Stefánsdóttir (2015) demonstrated via simulations of a 100 m pace race.

When discussing horse racing, the Thoroughbred breed also comes to mind. This gallop racing breed has been one of the main subjects of studies for genetics involved in racing performance and speed. Such has been the case with the Myostatin gene, where many SNPs have been proven to affect racing performance in Thoroughbreds, with the common pattern of a reference allele that is beneficial to long-distance racing and therefore stamina and a mutant allele associated to speed (Hill et al., 2010a; Hill et al., 2010b; Hill et al., 2011; Tozaki et al., 2010; Tozaki et al., 2011a; Tozaki et al., 2011b). Studies conducted in other breeds showcase the importance of studying the same gene variants in different breeds, where Myostatin variants important in the Thoroughbred have proven to have no effect in the performance of, for example, Coldblooded Trotters (Petäjistö, 2016). It is therefore interesting to investigate the possible role of these polymorphisms in the performance of such a versatile breed as is the Icelandic horse. While high frequencies of stamina-related alleles have been found in Icelandic horses (Viluma, 2016), no studies have been conducted to associate these performance SNPs in pace-racing Icelandic horses.

Studying genes that affect motor coordination and adaptability may also elucidate some of the mechanisms involved in the process of performing such a demanding gait at high speed. The horses must maintain a focused state of mind in order not to break the gait due to external distractions. Temperament of horses is a trait that is valued by both riders and breeders, even over other performance traits (Gille & Spiller, 2010; Gille et al., 2010; Górecka- Bruzda et al., 2011a; Graf et al., 2013). It is a given that the horse's temperament will affect its performance, both in matters of willingness and focus, but also their response to stress, especially considering the horse has a strong "flight" response to danger.

A recent genome-wide association study performed on Coldblooded trotters (Velie et al., 2018) revealed a series of SNPs that affected harness racing performance in this breed. Out of these, two of them were interesting candidates for a pace-racing study because of their role in a wide spectrum of neurodevelopmental disorders in humans. The first candidate was *GRIN2B*, a gene heavily related to not only neurodevelopmental disorders, but also with learning and memory ability. Two SNPs of this gene (g.41206762 C>T and g.41218272 T>C) proved to be significant to the performance of Coldblooded trotters in the study performed by Velie et al. (2018), but also showed to be borderline significant in a study performed by Jäderkvist et al. (2017) where four and five-gaited Icelandic horses were compared in a genome-wide association study. This highlights the importance of traits related to temperament and mentality in racing performance and therefore, these *GRIN2B* SNPs were selected for this study.

The other candidate chosen was a SNP located in the *DOCK8* gene (g.22496787 T>C), which not only is situated in close proximity to the DMRT3 mutation, but also has been associated with mental retardation and disorders such as ADHD in humans. The same study by Velie et al. (2018) revealed a SNP in this gene as affecting performance in harness-racing horses. Its possible role in gaiting ability through a relationship with DMRT3 and its role in neurodevelopmental disorders make of this gene an interesting candidate to study in pace-racing Icelandic horses.

The aims of this study were thus to investigate the role of known polymorphisms of the Myostatin, *GRIN2B* and *DOCK8* genes in the performance of Icelandic horses competing in pace-racing disciplines. The hypothesis of the study is that mutant alleles of each of these polymorphisms should be associated positively with performance.

2-Literature review

2.1-The Icelandic horse

The Icelandic horse breed originates from Iceland, where it has been purebred for almost 1000 years and therefore all ancestors of the breed are traceable to Iceland. Its closest relatives today are the horse breeds of Scandinavia and of the British Isles (FEIF, 2018). Today, over 250,000 Icelandic horses are registered worldwide, thereof approximately 40% in Iceland (Horses of Iceland, 2018).

These first horses that travelled to Iceland with the initial settlers were already selected using a breeding policy that emerged out of necessity, where only horses that were healthy and strong enough to make the trip were chosen. Since then, the conditions of island life have driven further selection up until the last century, where man has interfered more as knowledge of animal breeding has grown over time. Official assessments of breeding horses have been conducted on this breed for the past fifty years (FEIF, 2018).

2.1.1-Breeding of Icelandic horses

The breeding goal for Icelandic horses is to produce light-bodied, elegant horses that display energetic and attractive movements, focusing on a functional conformation that aids in the performance of all gaits. The breeding goal contains 15 conformation and riding ability traits (Table 1). Assessments are carried out in standardised breeding field tests (BFTs) in Iceland and in membership countries of the International Federation of Icelandic Horse Associations (FEIF). FEIF, 2018.

Table 1. Weighting proportion of single traits evaluated at BFTs and weight of overall score

Conformation		Ridden abilities/gaits	
Head	3.0%	**Toelt**	15.0%
Neck, withers and shoulders	10.0%	**Trot**	7.5%
Back and hindquarters	3.0%	**Pace**	10.0%
Proportions	7.5%	**Canetr/gallop**	4.5%
Legs (quality)	6.0%	**Spirit**	9.0%
Legs (joints)	3.0%	**General impression**	10.0%
Hooves	6.0%	**Walk**	4.0%
Mane and tail	1.5%		
Total	**40%**		**60%**

2.1.2-Training of Icelandic horses

Icelandic horses commence training between the ages of 3 and 5 years. Until then, they often have little contact with humans, although it is becoming more common to carry out pre-training early on to make the future training easier (Horses of Iceland, 2018). A questionnaire developed by Jansson et al. (2014) that aimed to document training strategies for pace horses used by experienced trainers (n=9) showed that the average age of training start for the pace gait was at 5,4 ± 1.0 years.

It can take several years before a sport horse reaches its maximum performance level (Thorén-Hellsten, 2008). According to Thorvaldur Árnason, Agr.Dr. in Animal Breeding and world-

champion rider in pace-racing disciplines, it is most common for pace-racing horses to start competing when they are 8 years old, although some compete as early as when they reach 6 years of age. According to the answers provided in the questionnaire by Jansson et al. (2014), horses reach their peak performance at the age of 14 years, approximately 9 years after beginning of training. At BFTs however, the situation is different, as the goal is to assess the quality of the horse independently of the level of training, riding, and other factors (Albertsdóttir, 2010). Competition horses are often older and more thoroughly trained than horses presented at BFTs. Competition horses are usually more carefully selected individuals destined for specific disciplines as opposed to breeding horses. This is so because competition horses are expected to be potential winners, while breeding assessments aim at evaluating the genetic potential according to the breeding goal of each horse (Albertsdóttir, 2010). The older ages and stronger selection make competition records less suitable as a basis of genetic evaluation compared to BFT records: longer generation intervals and biased estimation of genetic ability is often associated with competition data (Tavernier, 1991; Thorén-Hellsten, 2008). Nevertheless, competitive ability is declared in the breeding goal and many breeders strive to produce capable competition horses because of their value (Albertsdóttir, 2010). Therefore, to study the genetic background of competition performance is of interest.

2.1.3-Competition disciplines in the Icelandic breed

There are a great variety of equestrian competitions for Icelandic horses, which are divided into three types: pace racing, sport competitions and "Gæðinga" competitions (a riding horse quality competition), each including several disciplines. These competitions follow standardised international rules (FIPO) that are reproduced and reviewed annually by FEIF (FEIF, 2018).

The disciplines involve four kinds of gaiting tests: toelt, four-gait, five-gait and pace. Toelt, four-gait and five-gait are ridden on an oval track (200-300m) while pace is ridden on a straight track (100-250m). The pace involves five disciplines where two, the pace tests PP1 and PP2, are sport competitions and three are pace racing disciplines. In the latter, the pace is raced over different lengths: P1 over 250m, P2 with "flying start" over 100m and P3 over 150m. In the pace tests, however, not only the speed of the pace is measured in seconds, as is done in pace racing, but the quality of the pace is also assessed (FEIF, 2018).

Since 2009, competition records from internationally standardised competitions have been registered on the Worldfengur database, which has greatly improved the availability of competition data (Albertsdóttir, 2010).

2.1.4-Metabolic needs for pace versus different competition disciplines and how to train for them

A doctoral project conducted by Stefánsdóttir (2015) which sought to investigate the physiological response to exercise in Icelandic horses uncovered interesting findings in relation to the metabolic needs of pace racing. A simulation of a 100m pace race was performed using 9 horses and 2 riders in two different runs. Heart and respiratory rate were measured and blood samples were taken at different times during the simulations for measurement of haematocrit and plasma-lactate concentration. Lactate is the end product of anaerobic glycolysis (McMiken, 1983) and is produced in muscles at all exercise intensities with increasing levels at higher intensities (Lindholm & Saltin, 1974; Judson et al., 1983; Harris et al., 1991).

Results from Stefánsdóttirs work (2015) showed that flying pace is a demanding gait for the Icelandic horse that requires a high ATP turnover. The faster pace race horses had higher plasma lactate values after pace races, which was supported with results from Standardbred

trotters, where lactate concentrations after maximal exercise were related to performance and were higher in faster horses (Räsänen et al., 1995). The results made clear that the metabolic requirements for the short pace races in Icelandic horses (100, 150 and 250 m) differ from those in most other horse disciplines, which mainly require aerobic metabolism (Marlin & Nankervis, 2002; Gerard et al., 2014) but might be similar to those estimated for Quarterhorse races (Gerard et al., 2014). Therefore, training for anaerobic capacity is key to succeed in pace racing competitions for Icelandic horses (Stefánsdóttir, 2015). Pace trainers of Icelandic horses commonly use interval and uphill training and training in canter/gallop (Jansson et al., 2014), common methods that activate the anaerobic metabolic system (Stefánsdóttir, 2015).

Based on Stefánsdóttirs results, it is suggested that during training, as with Quarterhorses, it is important to maintain the proportion of type IIB fibers and glycolytic capacity in the muscles as opposed to focusing on endurance conditioning (Nielsen, 2014). These results are supported by studies performed in human athletes, where it is known that type IIA and especially type IIB fibers aid short duration anaerobic exercise and are present in higher proportions in elite strength and power athletes (Wilson et al., 2012). The differences between sprinter-type athletes and long-distance endurance athletes are well studied in humans, where long and middle-distance runners have a proportion of 60 to 70% of slow twitch fibers (type I), while sprinters typically present 80% fast twitch fiber makeup (type II) (Wilson et al., 2012). It is definitely of interest to study the possibility of fiber type interchangeability with training. Although shift between type IIA and type IIB fibers is demonstrated, results are conflicting regarding the capacity of type I and II fibers to interconvert. It is usually assumed that too much aerobic and condition training will cause a shift of fast-twitch muscle fibers into slow twitch muscle fibers and therefore the individual may lose sprinting power due to a lack of sufficient anaerobic muscle fiber types.

Several studies reviewed by Wilson et al. (2012) indicated that exercise-induced muscle fiber shifting only exists between fast twitch fiber types (Karp, 2001; Scott et al., 2001 and Simoneau et al., 1985), whereas other findings reject this possibility (Wilson et al., 2012). However, regardless of the actual controversy on the matter, several findings imply that fast to slow twitch fiber shift and vice versa is possible when exercise variables are carefully controlled (Esbjornsson et al., 1993; Howald et al., 1985; Jansson et al., 1978; Kraemer et al., 1995 and Simoneau et al., 1985). Therefore, while further research is necessary, current findings indicate that altering fiber type percentages might be achievable via exercise (Wilson et al., 2012).

2.1.5-Temperament of Icelandic horses

Pace-racing in Icelandic horses is a discipline that requires the animals to undergo specific training aimed towards improving the specific metabolic needs of pace, where also their temperament and learning ability play a key role apart from their potential to pace at high speed. Overall, temperament traits in horses are considered to be important by both riders and breeders, even over other performance-related traits (Gille & Spiller, 2010; Gille et al., 2010; Górecka- Bruzda et al., 2011a; Graf et al., 2013). Several sport horse breeding associations worldwide mention temperament or behavioural traits as part of the breeding goal (Koenen et al., 2004). However, not all use the information obtained in temperament assessments to estimate breeding values to be used in the selection of breeding animals (von Borstel, 2013). The Icelandic horse is one such breed, where the trait “spirit” is a part of the current selection index (FEIF, 2018) and is included in the genetic evaluation for Icelandic horses. While overall temperament is a commonly evaluated trait in horses, a deeper analysis and more objective evaluations of temperament traits, such as learning ability, is needed (von Borstel et al., 2013).

The specific breeding goal for 'spirit', which aims to evaluate both temperament and willingness of the horse, aims for producing horses that are "fiery, cheerful and brave, but extremely easy to handle". (FEIF, 2018). Temperament has been a part of the official Icelandic horse selection criteria since 1950, but until 2000 evaluated as two separate traits: willingness and character. Since the relationship between the two traits made it difficult to assess these traits separately, along with the fact that it was believed that interplay between them was the trait of true value, they were combined into the trait spirit. A somewhat higher genetic correlation was later estimated between willingness and spirit (0.45), than between character and spirit (0.32) (Árnason, 2004). The estimated heritability used for the genetic evaluation of spirit is 0.37 (Árnason, 2004), and considerable genetic gain for spirit has been reported (Sigurðardóttir, 2012).

2.1.6-Genetic evaluation and genetic correlations in the Icelandic horse

Selection of Icelandic horses destined for breeding has been based on estimated breeding values for over thirty years (Albertsdóttir, 2010). In 1975, the first multiple trait selection indices for individual selection and progeny tests were designed based on estimated genetic parameters on ten BFT traits (Árnason, 1979). In 1979, a BLUP sire model was applied in a single trait genetic evaluation of total score of BFT traits. This evolved in 1983 into a BLUP multivariate animal model (Árnason, 1982, 1984). Since 2005, records have been used from assessments in 11 countries. In 2006, the global genetic evaluations of the Icelandic horse and the genetic connectedness between countries were also evaluated. Results showed that estimated breeding values of individual horses could be compared across countries and within countries with similar accuracy. Additionally, it was concluded that a good genetic connection existed between the countries (Árnason et al., 2006). Today, estimated breeding values for 31 individual traits are calculated from international BFT scores (Árnason et al., 2006).

2.2-Height and sprinting performance

Kristjansson et al (2016) conducted a study with the objective of assessing the phenotypic and genetic relationship between standard conformational measurements and scores for riding ability. Additionally, the study aimed to investigate if more detailed (3-D) morphometric measurements could differentiate high-class and low-class horses based on scores for each gait.

They found that proportions in the top line of the horse, specifically what they named "height at front" (difference between height at withers and height at back and the difference between height at withers and height at croup) compared to hind were found to be most important for riding ability. Furthermore, the estimated heritability and genetic correlation with the total score for riding ability marked these proportions as important indicators for performance. This suggests that an uphill conformation of the horse is advantageous and may discriminate between high and low-class horses. Additionally, length, proportions and angles of bones in fore- and hind limbs also affected scores for the different gaits affected performance in a similar fashion.

However, in this study, height to the withers, back and croup, by themselves, did not have a significant effect on scores for pace. One must take into account that this study used the scores obtained at breeding field tests for pace; in these tests, the best pace is described as "Secure, impressive pace, good 2-beat lateral gait with good suspension and excellent speed" (FEIF, 2018). While the speed is taken into account, this is a subjective evaluation where no measurements are taken. When looking at speed as a performance trait, there have been positive correlations found between height at withers and speed in hand-led trot in horses (Galisteo et al., 1998).

In another study conducted by Albertsdóttir et al. (2008), genetic parameters of competition traits and genetic relationships between competition traits and BFT traits were analysed. The conclusions of this study were that the ability to perform well was evaluated in a similar fashion between BFTs and different competition disciplines in Icelandic horses. In this study, pertaining particularly to pace, Albertsdóttir et al. (2008) found that out of the fifteen breeding field test traits and four competition traits evaluated, the correlation of height to withers and spirit had a moderately strong correlation with PP1 (0.38 and 0.43, respectively), while height to withers had a low correlation with toelt, for example.

In humans, studies provide with controversial results. One study established that greater height confers a higher stride length and therefore, lower stride frequency. These two parameters were deemed to be negatively correlated in maximal velocity running (Vanderka & Kampmiller, 2013). The recommendations offered by the authors were that young athletes were to be evaluated for sprinting talent based on their stride frequency-which suggests that stride length, and therefore height, would be detrimental to speed. However, another study by Watts et al (2012), where the aim was to investigate whether any anthropometric parameters characterized successful world-class sprinters, concluded that reciprocal ponderal index (RPI) is a good predictor of sprinting speed in both men and women. This index is a measure of linearity, with more linear (taller) athletes having greater reciprocal ponderal indices. The authors speculate that high RPI in elite sprinters might be partly explained by the influence of stride length on sprint speed. The authors concluded that body shape should be taken into consideration when selecting potential athletes for sprint events, encouraging more linear (taller) athletes with a high RPI.

To our knowledge, a direct association of speed measured by time taken in finishing a pace race and height at the withers in Icelandic horses has not been done. Therefore, we decided to include it as a trait with the hypothesis that taller horses are faster due to higher length stride.

2.3-*DMRT3* gene (Doublesex and mab-3 related transcription factor 3)

DMRT3 is a gene that has been proven to have a major effect on the pattern of locomotion in horses. A single base-pair mutation in this gene (a change from cytosine (C) to adenine (A)) confers the ability to perform alternate gaits and has a favourable effect on harness racing performance (Andersson et al., 2012).

It has been confirmed that homozygosity for the DMRT3 nonsense mutation relatesalso has a positive effect on scores in BFTs for toelt, demonstrated by better beat quality, speed capacity and suppleness (Kristjánsson et al., 2014). On the other hand, horses with the CA genotype perform better in walk, trot, canter and gallop. This indicates that the AA genotype supports the coordination of ipsilateral legs, with a corresponding negative effect on the movement of diagonal legs (Kristjánsson et al., 2014). Selective breeding for lateral gaits in the Icelandic horse population has apparently altered the DMRT3 genotype frequencies, with a predicted loss of the C allele in relatively few years (Kristjánsson et al., 2014).

A recent study conducted by Amano et al. (2018) investigated the effect of this DMRT3 mutation in Hokkaido Native Horses, a Japanese breed with ability to pace. They examined genetic factors other than DMRT3 by exploring genome-wide Single-Nucleotide Polymorphisms (SNPs) related to gait determination. SNPs on chromosomes 13 and 23 were detected by genome-wide association analysis, although SNPs on chromosome 23 were all located in the vicinity of the DMRT3 mutation. To uncover genetic factors that may work with this mutation in gait determination, animals with AA genotype for DMRT3 were also targeted. Associations with 23 SNPs on six different chromosomes were found (Amano et al., 2018). The possible effects of these SNPs on gait determination need to be further studied with

additional data, as was similarly elucidated by the study Jäderkvist et al. (2017) performed, where a comparison of a small number of four and five-gaited horses yielded no genome-wide significant SNPs. If anything, it is clear that, while the DMRT3 mutation and SNPs in its vicinity seem to have a major role in gait determination, there are bound to be other genes involved in the process.

The probability of DMRT3 genotype is calculated for all horses registered in WorldFengur. The estimates are based on the Mendelian laws, pedigree, phenotypic scores and known genotypes of over 800 tested horses. The method is based on Genotype Elimination algorithm and iterative peeling as described by Kritjansson et al., (2014). The predicted DMRT3 is only presented for estimates with high accuracy.

2.4-*GRIN2B* gene (Glutamate ionotropic receptor NMDA type subunit 2B)

N-methyl-D-aspartate receptors (NMDARs) are a family of ionotropic glutamate receptors that mediate a slow, Ca2+ permeable component of excitatory synaptic transmission in the central nervous system (Hu et al., 2016). NMDARs are a tetrameric assembly of two glycine-binding GluN1 subunits and two glutamate-binding GluN2 subunits. The GluN1 subunit is expressed throughout the central nervous system, whereas the four subtypes of GluN2 (A–D) subunits have differential temporal and spatial expression patterns (Akazawa et al., 1994; Monyer et al., 1994). The GluN2B subunit encoded by *GRIN2B* is highly expressed prenatally, and its expression level starts to decline after birth in most brain regions (Akazawa et al., 1994). The early expression of this NMDAR subunit suggests it may play an important role in brain development, circuit formation, and perhaps cellular migration and differentiation, as well as synaptic plasticity (Cohen et al., 2008). Neonate lethality of *GRIN2B* knock-out mice has been observed (Kutsuwada et al., 1996), while overexpression of *GRIN2B* in the forebrain of mice enhanced hippocampal long-term potentiation and spatial memory performance (Tang et al., 1999).

A large number of mutations have been found in the NMDAR gene family, including the *GRIN2B* gene. The GluN2B subunit it encodes for has been related to many cases of neurodevelopmental disorders, such as language, motor and learning disorders, autism spectrum disorder, attention deficit hyperactivity disorder, developmental delay, epilepsy, and schizophrenia. It has also been shown that overexpression of NMDA receptor 2B (NR2B) in the forebrains of transgenic mice leads to enhanced activation of NMDA receptors, aiding synaptic potentiation in response to stimulation. The mice exhibited superior ability in learning and memory in various behavioural tasks, which suggests that NR2B is critical in gating the age-dependent threshold for plasticity and memory formation (Tang et al., 1999).

A recent genome-wide association study performed in Coldblooded trotters (Velie et al., 2018) yielded two genome-wide significant SNPs associated with career earnings located in the *GRIN2B* gene (g.41206762 C>T and g.41218272 T>C). Interestingly, the same SNPs have also been suggested in a previous study exploring pacing ability in Icelandic horses (Jäderkvist et al., 2017) where both polymorphisms were borderline significant in the analyses. However, minor allele frequencies (MAF) in 4 and 5 gaited horses were practically the same (MAF = 0.21 in four-gaited horses and MAF= 0.20 in five-gaited horses).

2.5-*DOCK8* gene (dedicator of cytokinesis 8)

The *DOCK8* gene provides instructions for making a member of the DOCK family of proteins. The proteins in this family act as guanine nucleotide exchange factors (GEFs). GEFs activate proteins called GTPases, which play an important role in chemical signalling within cells. Signalling stimulated by DOCK family proteins are typically involved in the arrangement of the cytoskeleton, which implies they play a role in cell structure and migration.

The *DOCK8* protein is found most abundantly in cells of the immune system, where it plays a key role in the survival and function of different types of immune system cells, including T cells, NK cells, and B cells. Many *DOCK8* mutations have been identified and strongly associated with hyper-IgE syndrome and other immunodeficiency syndromes (Zhang et al., 2009; Qin et al., 2016; Burbank et al., 2016). While these immune diseases are linked in humans to mutations, it is important to take into account the known effects stress has on the immune system in humans, where chronic stress leads to situations of both immunosuppression or excessive immune activity (Segertsrom and Miller, 2004). Competitions can be a stressful situation and a subsequent immune response can occur to prepare for eventual injury; therefore, horses that cannot manage stress well might eventually suffer from reduced immune activity. Considering that the study performed by Velie et al. (2018) on Coldblooded trotters yielded five SNPs with genome-wide significance affecting performance of harness racing, it is of interest to investigate this gene in other racing breeds for its possible effect in immune system regulation and the relationship of immunity and stress.

While recurrent infections and a predisposition to autoimmune disease is definitely a factor that can affect athletic performance, it is also interesting to note the relationship of *DOCK8* mutations with mental retardation. A study conducted by Griggs et al. (2008) identified disruptions in *DOCK8* in two unrelated patients with mental retardation. The *DOCK8* gene is expressed in several human tissues including adult and fetal brain. A role for *DOCK8* in processes that affect organization of filamentous actin has been suggested (Ruusala et al., 2004). Several genes influencing actin cytoskeleton have been implicated in human cognitive function and thus a possibility exists that the *DOCK8* mutation may contribute to some cases of autosomal dominant mental retardation (Griggs et al., 2008). Actin proteins have a critical role in muscle contraction and relaxation as well, making up 90% of muscle protein along with myosin (Dominiczak and Baynes, 2005). Another study, conducted by Glessner et al. (2017), performed a large scale meta-analysis of CNVs across multiple neurodevelopmental and psychiatric diseases. They uncovered novel significant associations of structural variants in the locus of *DOCK8*/KANK1 shared by five diseases, suggesting common etiology of these clinically distinct neurodevelopmental conditions. This study evidenced for the first time that *DOCK8* duplications were found to be significantly associated with a spectrum of neurospsychiatric disorders, suggesting that a tightly regulated *DOCK8* expression level may be required for normal cellular function.

Out of the five SNPs that the study performed by Velie et al. (2018) identified as significant to performance, only one was chosen for this study (g.22496787 T>C). This was due to the extremely low minor allele frequencies the other SNPs presented in the Icelandic horse, as evidenced by data extracted from the study performed by Jäderkvist et al. (2017).

It is noteworthy to state that the study performed by Amano et al (2018) on Hokkaido Native Horses did not find any of the same location SNPs for *DOCK8* in their study in relation to gaiting ability in the chosen breed-but many of them are still near the *DOCK8* gene itself. One SNPs they located was as close to our selected polymorphism as roughly 100,000 bases upstream (position 22,597,020), with another two still being relatively close (positions 22,967,656 and 22,922,059). Overall, considering that the *DOCK8* SNPs located by Velie et

al. (2018) are in close vicinity to the DMRT3 mutation and that many of the polymorphisms found by Amano et al. (2018) are also in close distance to both genes, *DOCK8* is definitely interesting candidates for gaiting ability.

2.6-*MSTN* (Myostatin)

Myostatin, a protein encoded by the *MSTN* gene, is a part of the transforming growth factor beta (TGF-β) superfamily, which plays an important role in regulating embryonic development and tissue homeostasis maintenance in adults. Myostatin is specifically expressed in skeletal muscle tissue of numerous mammalian species, where it acts as a negative regulator of muscle growth through inhibition of myoblast proliferation. This is achieved through interruption of normal cell cycles mediated by the p21 family, a group of cyclin-dependent kinase inhibitors which act during G_1/S transition (Thomas et al., 2000).

Mutations in the *MSTN* gene have been shown to have different effects in mammalian species. When it was first described in 1997 by McPherron et al, it was discovered that producing myostatin-null mice through gene targeting resulted in individuals that were 30% larger than the wild type mice. Not only that, the mice showed abnormal body shapes, all due to an increase in skeletal muscle mass caused by both hyperplasia and hypertrophy of muscle cells.

A phenotype common to some cattle breeds known as 'double-muscling' has been linked to mutations in this gene. Double-muscled animals are characterized by an increase in muscle mass of about 20% due to general skeletal muscle hyperplasia (increase in number of muscle fibers). The cattle breeds Belgian Blue and Piedmontese show a high frequency of this phenomenom (McPherron & Lee, 1997; Grobet et al., 1997; Kambadur et al., 1997). Sequence analysis of the so-called muscular hypertrophy locus revealed deletion-type mutations of the *MSTN* gene in heavy-muscled cattle of both breeds.

The double muscling phenotype has also been described in dogs, where it is also caused by a *MSTN* mutation. It has been specifically described in the Whippet breed, where some individuals present the so-called "bully" phentoype. Mosher et al. (2007) uncovered a 2bp deletion in the third exon of the *MSTN* gene which causes a premature stop codon. It was found that individuals that were homozygous for the deletion presented the "bully" phenotype, while dogs that sired a "bully" whippet were heterozygous for the mutation. The heterozygous individuals were less muscular than the bully variant; however, they were still more muscular than the wild type. Not only that, but a high frequency of the heterozygote genotype was found among the fastest racers, demonstrating that this mutation enhances performance in this breed.

2.6.1-Myostatin in horses

After the findings Mosher et al. published relating a *MSTN* mutation with racing performance in dogs, other researchers started to look for similar effects of this gene in other species used for racing. The English Thoroughbred is the most common used breed for horse racing (The Swedish Horse Racing Authority, 2018) and has therefore been an interesting breed to study the effect possible effects of the *MSTN* gene in association to performance. In this species, *MSTN* is located at the end of chromosome 18 and is made up of three exons and two intron regions (Grobet et al.,1997).

A) Myostatin and racing performance

A study performed in 2010 by Hill et al. described, for the first time, the identification of a *MSTN* sequence polymorphism (g.66493737C>T [PR3737], located in the first intron of the *MSTN* gene) in these horses that was predictive of genetic potential and athletic phenotype, associated particularly with speed and best racing distance among elite race horses. Hill et al. (2010a) re-sequenced the equine *MSTN* gene and found six Single Nucleotide Polymorphisms (SNP) in intron 1 of *MSTN*, out of which two were included in their association analysis. Considering the contribution of muscle power to sprint and longer distance racing, the elite group within their Thoroughbred sample was divided into those who had won their best race in a short distance (≤1600 m) or long distance (>1600 m). A highly significant ($P = 3.70x10^{-5}$) association was found with PR3737, where the C allele was twice as frequent in the short distance than in the long-distance group. Additionally, best race distance (BRD) was found to be significantly associated ($P=4.85x10^{-8}$) with the PR3737 SNP. It was observed that as the distance of the races increased, the frequency of the CC genotype decreased, further supporting the association of the CC genotype with sprinting power. On the other hand, the TT genotype was associated with stamina, seeing an increase in the genotype frequency over longer racing distances, while the CT genotype was best suited for middle-distance races (Hill et al., 2010a).

Hill et al. (2010b) also conducted a genome-wide SNP-association study in order to investigate the effects of additional gene variants that could have an effect in racing performance. In this study, genotypic variation at 40,977 SNPs was evaluated between horses selected for short distance and middle to long distance racing. It was concluded that PR3737 is the most powerful genome-wide predictor of optimum racing distance in Thoroughbred horses (Hill et al., 2010b).

Following this discovery, Tozaki et al. (2010) also conducted a genome-wide association study in Thoroughbreds that led to the identification of three additional SNPs on chromosome 18, namely, g.65809482T>C, g.65868604G>T and g.66539967A>G. It was found that the haplotype of these polymorphisms, together with genotype for PR3737, were associated with lifetime earnings and rankings, which further supported the results of Hill et al. (2010a). Tozaki et al. (2011) also found in a later study that two of their identified SNPs (g.65868604G>T [PR8604] and g.65809482T>C [PR9482]) were associated with best race distance, performance rank and lifetime earnings. Both there polymorphisms behaved in a similar fashion, where always the alternative allele (G for PR8604 and T for PR9482) was associated to short distance races and the reference allele to long-distance races.

B) Genotype frequencies among horse breeds

Thoroughbred horses have been heavily selected for speed, which has led to a high frequency (0.51) of the C allele of SNP PR3737. However, allele frequencies at this *MSTN* SNP in various other horse breeds present large variation (Bower et al., 2011).

Hill et. al (2010a) found that in Quarter horses, a sprinting breed selected for short distance racing, there was a very high frequency of C alleles (0.90), while in National Hunt racehorses (used for races over obstacles and long distances) there were no CC genotype individuals, which supports the association of the T allele with stamina. Additionally, genotype frequencies in Egyptian Arabian horses, famous for endurance exercise, were quite different to that of the Thoroughbreds, where a high proportion of TT (0.90) genotypes were found. These results further support that the CC genotype is suited to fast, shorter distance racing and the TT genotype confers stamina.

Bower et al. (2012) genotyped the SNP PR3737 in forty donkeys and two zebras in order to determine the ancestral allele in equids. Their results showed that no donkey or zebra had the C variant and therefore ascertained that the T-allele was the wild type, which is consistent with herbivore behaviour of grazing over long distances.

A study performed by Viluma (2012) that evaluated genotype frequencies of different *MSTN* SNPs in various horse breeds showed that the Shetland Pony, Gotlandsruss and Fjord Horse had the highest frequency of CC individuals for PR3737 out of 25 different breeds analysed (excluding Thoroughbreds). Viluma also studied allele frequencies in Icelandic horses, where it was observed that the frequency of the mutant allele for PR3737 was very low (0.09 out of 313 individuals genotyped), with no homozygous CC individuals detected and a frequency of 0.17 CT heterozygous individuals. As for PR8604, the mutant allele G presented a frequency of was of only 0.20, with a slightly higher frequency of heterozygous individuals (0.34). In other studies, where Icelandic horses with either no competition or breeding field test data were genotyped, the allele frequencies for these SNPs remain similar. Out of these 280 horses, the MAF for PR3737 was of 0.08, the MAF for PR8604 of 0.24 and that of PR9482, 0.38.

C) Insertion polymorphism of Equine Repetitive Element 1 (ERE1)

A recent study by Santagostino et al. (2015), where a genome-wide evolutionary and functional analysis of the Equine Repetitive Element 1 (ERE1) was performed, revealed that an insertion in the myostatin promoter affects the expression of this gene. With the hypothesis that the ERE1 insertion could affect the aptitude for specific sport abilities, Santagostino et al. analysed the frequency of the two variants (ERE1 containing allele or ERE1 allele lacking the insertion) in 30 horses competing in show-jumping, in 90 Italian Trotters bred for harness racing and in 75 horses registered in the Italian Thoroughbred studbook bred for flat racing. Their results showed different allele frequencies in the three groups, where the ERE1+ allele was completely absent in the Trotters and, in the Show Jumpers, only one individual was heterozygous for the variant. These results suggested that the ERE1+ allele may have been selected in the Thoroughbreds and in the Quarter Horses together with flat racing aptitude. To test whether the ERE1+ variant could influence racing performance in Thoroughbreds, they selected 117 elite racing horses where this allele was significantly more frequent compared to the general population of Thoroughbreds. The elite horses were grouped according to BRD and it was evidenced that most winning horses (18 out of 30) over short distances were homozygous for the ERE1+ allele. On the other hand, no homozygous individuals for the ERE1- allele were found. For horses competing over long distances, only heterozygotes and ERE1- homozygotes were observed. For horses competing in medium distance races, all the three genotypes were present, although the ERE1+ homozygotes were relatively more frequent in the groups winning up to 1600 m races compared to horses winning 1700–2000 m races.

With the effects of PR3737 in mind, Santagostino et. al. compared the ERE1 genotypes with that of this SNP. Their results showed that in 112 out of 117 horses, the two genotypes were concordant, with the C SNP allele associated with ERE1+ and the T SNP allele associated with the ERE1- promoter. These results show that the two polymorphic loci are tightly linked, as expected by their close proximity in the genome (1605 bp), which supports hypothesis that this is the genotype that drove selection.

It is interesting to note that the study conducted by Viluma (2012) also aimed to genotype this insertion polymorphism, but out of 100 Icelandic horse samples analysed, no insertion alleles were detected.

3-Material and methods

3.1-Population and samples

A sample of 131 Icelandic horses born between 1988 and 2012, originating from different countries (Austria, Denmark, Germany, Iceland, the Netherlands, Norway and Sweden) were used for this study. The sample represented offspring from 98 sires and 128 dams and consisted of 38 stallions, 49 mares and 44 geldings. Sex classification was based on information provided through owner questionnaires for each individual when available. When no questionnaire was available, sex classification was based on recordings in the WorldFengur database up to August 15th, 2018.

The samples consisted of isolated DNA extracted from hair roots, the majority of which were collected via personal contacts and online postings during the year 2018. Some of the samples were provided by the Animal Biobank of the Department of Animal Breeding and Genetics from SLU, Sweden.

Distribution of individuals over country of birth and country of residence can be found in Tables 1 and 2.

Table 2. Distribution of country of birth over the total sample

Country of Birth	Stallion	Mare	Gelding	Total
Austria	1	1	0	2
Denmark	0	1	0	1
Germany	3	6	4	13
Iceland	19	27	27	73
Netherlands	0	1	1	2
Norway	0	1	0	1
Sweden	15	12	12	39

Table 3. Distribution of country of residence over the total sample

Country of Residence	Stallion	Mare	Gelding	Total
Austria	1	4	4	9
Denmark	1	3	0	4
Germany	7	9	4	20
Iceland	3	10	16	29
Netherlands	0	2	1	3
Norway	2	2	2	6
Sweden	24	19	17	60

3.2-SNP genotyping

For this study, three different SNPs for the *MSTN* were chosen based on results from previous studies: g.66493737C>T (PR3737), g.65809482T>C (PR9482) and g.65868604G>T (PR8604). For the *GRIN2B* gene, two SNPs were chosen: g.41206762 C>T (*GRIN2B*.6762) and g.41218272 T>C (*GRIN2B*.8272). For the *DOCK8* gene, a single SNP was chosen, g.22496787 T>C (*DOCK8*.6787). Additionally, all horses of the sample lacking estimated DMRT3 genotype in WorldFengur were genotyped to ensure that they were all homozygous AA (n=21).

Single nucleotide polymorphism genotyping was performed with the StepOnePlus Real-Time PCR System (Life Technologies [Thermo Fisher Scientific], Waltham, MA) using Custom TaqMan Genotyping assays (Applied Biosystems by Life Technologies [Thermo Fisher Scientific]). All samples were genotyped for all six SNPs using 96-well plates and custom mixtures of DNA samples and TaqMan reagents for every plate.

Additionally, Hardy-Weinberg equilibrium was evaluated for all SNPs genotyped using the "SNPassoc" package (González et al., 2015) for the software program for statistical computing R (R Development Core Team, 2017).

3.3-Performance data

Data was obtained from different databases and an owner questionnaire to carry out the statistical analysis based on lifetime performance results up to August 15^{th,} 2018. The main base used was WorldFengur (The Studbook of Origin for the Icelandic horse). Other bases used were the Islandpferde-Reiter- und Züchterverband (German organization for riders and breeders of Icelandic horses); the official website for the Central European Championships 2018 (St. Radegund, Austria) and the website for a Swedish club for pace-racing enthuthiasts called "Passklubben Skeið", which is associated to the Swedish Breeding Association for Icelandic Horses (SIF, Svenska Islandshästförbundet). The data taken from the German database pertained only one horse and constituted all of its available performance data. The rest of bases were used as complementary performance data for a small number of horses (n=23), being their main source of data that from WorldFengur, with the exception of another individual whose only data was provided by the results of the Central European Championships.

The traits extracted from the databases were race times for three different disciplines: P1 (250m pace race), P2 (100m pace race with flying start) and P3 (150 m pace race). Other factors included were sex, country of birth, country of residence and birth year. BLUP values and Breeding Field Test assessment scores obtained for pace were also included when available. The owner questionnaire was designed to obtain temperament information about each horse as well as training routines.

The analysis was divided into three age groups: total career performance, performance up until and including 13 years of age and performance over 13 years of age. This classification was based on the average age at which the horses achieved their best racing times within the samples, which was approximately at 12.6 years of age for all three race disciplines.

Race times

Race times for three different race disciplines were included in the study. Time keeping of all races is measured through an electronic system. When start boxes are used, the time measuring begins when the doors start to open, as the electronic time keeping system should be connected to the opening system of the starting boxes (FEIF, 2018).

Pace race P1, done over 250m, is performed on the pace track and is nowadays started with boxes or automated start machines. Usually, two horses will compete simultaneously in two or more heats. The best time achieved in the event is the one registered. When marks and not time are being registered, they are calculated through the following formula: (32.50 - t) / 1.25, where t = the time of the fastest run.

Pace race P2 with flying start, also known as "Speed Pass", is conducted on a pace track, where there is a stretch of 50 m before the start of the timed stretch of 100 m. Only one horse competes per heat, of which there are two. At the starters' signal, the rider rides to the 50 m marker in any gait he wishes and on crossing the 50 m mark, time keeping starts at a visual signal. From there to the finish line the horse must be in racing pace. Again, the best time achieved is registered. When marks and not time are being registered, they are calculated through the following formula: (12.00 - t) / 0.60, where t = the time of the fastest run.

Pace race P3, done over 150 m, has the same rules as those for P1 and, just as in P2, the first 50 m are run in free gait. Similarly, when marks and not time are being registered, they are calculated through the following formula: 22.00 - t, where t = the time of the fastest run.

Any results recorded in marks were converted to time in seconds using each respective formula in the study.

Breeding field test points

For any horses assessed at a Breeding Field test, the highest score obtained for pace was included in the study. Points range from 5.0 (which means the rider did not show pace for the horse) to 10.0. The pace at breeding field tests is evaluated as a "secure, impressive pace, good 2-beat lateral gait with good suspension and excellent speed" (FEIF, 2018).

Estimated Breeding Value (EBV)

The BLUP (Best Linear Unbiased Prediction) method with animal model is used for genetic evaluation of Icelandic horses. Using an animal model implies that all available pedigree information is used for determining correct genetic relationships between animals. BLUP methodology uses pedigree data and scores for traits obtained at BFTs in the most efficient way and weighs all available information in accordance with the statistical model used. Therefore, this method is able to adjust for the effects of systematic environmental factors which can be registered in the data such as different countries, years and genders (Árnason, 2014). Somewhat simplified the statistical model for each trait can be written as:

Score assessed at BFT = fixed effects of country and judging year + fixed effects of gender and age class + random effect of breeding value + random residual effect

The following weighting factors are used for computing the total scores (from 2010): Head 3%; Neck; withers and shoulders 10%; Back and croup 3%; Proportions 7.5%; Leg

quality6%; Leg stance 3%; Hooves 6%; Mane and tail 1.5%; Toelt 15%; Trot 7.5%; Pace 10%; Gallop 4.5%; Spirit 9%; Expression 10% and Walk 4%. In this study, the total BLUP score and the score for pace were included as traits in the analysis.

Owner questionnaire: training routines

A questionnaire was designed as a word-document that each owner filled in and sent along with the hair sample of each horse. In this questionnaire, the owners were asked to provide the age (in years) at which each horse had commenced training for pace race. They were also asked to inform about which competitions the horses were regularly trained for in order to classify the horses as pure pace-racers (only train for PP1, P1, P2 and/or P3) or "non-pure" pace racers (trained for Toelt, Five-gait, Breeding show, etc).

Owner questionnaire: temperament traits

The owners were asked to evaluate three temperament traits for each individual on a scale of 1 to 5. The traits evaluated were: Nervousness (where 1 equals easily frightened and nervous horses and 5 equals very calm horses that can handle stressful situations well); Focus (where 1 equals easily distracted and uncooperative horses, while 5 equals very focused and cooperative horses that do not easily break the gait due to external distractions) and finally Motivation (where 1 equals horses that are unmotivated to run fast and are hard to get in motion and 5 equals very motivated to win and energetic horses).

Out of the 131 horses included in the study, a total of eight individuals did not have questionnaire information. Additionally, one individual who only competed in P2 races lacked a score for Motivation, seeing as P2 is a standalone competition.

Height at withers

The height of each horse was also included in the study when available. The majority of recordings came from the owner questionnaire, where the owners were asked to provide the height to withers of each horse. If this information was unknown, the owners were asked to say if the horse was over or under 140 cm. If any horse had attended a field breeding test show, the height recorded there was the one included in the study.

3.4-Summary statistics

Data was structured using custom scripts written in the software program for statistical computing R (R Development Core Team, 2017) to facilitate analysis and production of summary statistics. Summary statistics were calculated for all performance traits previously mentioned stratified by sex, age group, speed group and according to genotype of each SNP, with the exception of those genotypes which were only represented by one individual (n=1 for genotype CC of PR3737).

3.5-Association analysis: group comparisons

Groups were created within the sample to compare Fast, Average and Slow horses. This classification was based on two factors: the consistency of their racing times and what they are regularly trained for. Racing times were considered as fast if the times were below 23 seconds for P1 (250m), below 8 seconds for P2 (100 m) and below 15 seconds for P3 (150 m). Race times were considered slow if the time was above 25 seconds for P1 (250m), above 8.5 seconds for P2 (100 m) and above 16.5 seconds for P3 (150 m).

Group 1 (Fast, n=37) included individuals that are exclusively trained for P1, P2 or P3 and have consistently fast racing times (with the exception of two individuals with exceptionally

fast times that were also trained for Five-gait competitions). Group 2 (Average/Inconsistent, n=75) included horses with inconsistent racing times that also train for other competitions outside of pace racing. They were therefore supposed to have a muscle fiber composition more suited towards aerobic and condition training rather than for anaerobe exercise. Group 3 (Slow, n=19) included horses that are exclusively trained for pace-racing but had consistently slow racing times, with the exception of 4 horses that were trained for Five-gait competitions but had exceptionally slow times.

These groups were compared to evaluate genotype distribution and distribution of temperament trait scores among groups. Additionally, comparisons were carried out between the different *GRIN2B* genotypes and the scores for the different temperament traits in order to find patterns between the scores and the different genotypes. Fisher's exact test was performed in all comparisons to search for significant differences in trait and genotype distributions among the groups.

3.6-Association analysis: models

An association analysis was performed with different models depending on the age group classification. When evaluating the total career of each individual (which included one observation per individual with their best times achieved), simple linear models (linear regression) were used. Sex and Birth year were included as fixed effects in these models. Birth year was made into classes where the different levels were grouped into classes with at least 10 individuals per class.

For the different age groups, where repeated measurements per individual were included, linear mixed models were used. This was achieved through the statistical packages 'lme4' (Bates et al., 2018) and 'lmeTest' (Kuznetsova et al., 2018) for the software program for statistical computing R (R Development Core Team, 2017). These packages allow the building of models that may include the random effect of individual, which was always included in these models. The 'lmeTest' package works within the properties of lme4 and provides p-values for each trait included in the model in the same way that a simple linear model does in the R software.

The temperament traits obtained through the questionnaires were excluded from all regression analyses due to high correlation with race time traits. The age at which the horses started to train for pace was also excluded from the analysis due to the fact that it was not significant in any of the models tested. Concerning the height trait, it was introduced as a covariate in the models for race times (dependent variable) as an interaction with the different myostatin SNPs and removed if not significant.

In order to better understand the roles of the different genotypes analyzed, tables were built to portray the means and medians of all performance traits stratified by genotype and including the respective p-values each genotype had for every trait. These tables can be found in Appendix IV (Table IV.1, Table IV.2, Table IV.3, Table IV.4 and Table IV.5).

4-Results

4.1-Genotype distributions

Genotype distribution, allele frequencies and p-values for Hardy-Weinberg Equilibrium (HWE) can be found in Table 3 for the *MSTN* SNPs, Table 4 for *GRIN2B* SNPs and Table 5 for the *DOCK8* SNP. Five individuals were not successfully genotyped for one? SNP (PR9482). All SNPs were in HWE (p>0.05) except for both *GRIN2B* SNPs (p=0.03).

Table 4. Hardy-Weinberg Equilibrium (HWE), genotype and allele frequencies for the *MSTN* SNPs. Number of individuals for each genotype and frequency of genotypes (within parenthesis). Number of alleles in total sample (131) and frequency (within parenthesis). P-value for HWE

MSTN SNPs	g.65809482T>C			g.66493737C>T			g.65868604G>T		
	CC	CT	TT	TT	TC	CC	TT	TG	GG
Genotype frequency	57 (0.45)	51 (0.41)	18 (0.14)	104 (0.79)	26 (0.20)	1 (0.01)	90 (0.69)	34 (0.26)	7 (0.05)
Allele Frequency	C=165 (0.65)	T= 87 (0.35)	NA's=5	T= 234 (0.89)	C= 28 (0.11)		T=214 (0.82)	G= 48 (0.18)	
HWE p value		p= 0.24			p= 1			p= 0.14	

Table 5. Hardy-Weinberg Equilibrium (HWE), genotype and allele frequencies for the *GRIN2B* SNPs. Number of individuals for each genotype and frequency of genotypes (within parenthesis). Number of alleles in total sample (131) and frequency (within parenthesis). P-value for HWE

GRIN2B SNPs	g.41206762 C>T			g.41218272 T>C		
	TT	TC	CC	CC	CT	TT
Genotype frequency	103 (0.79)	23 (0.18)	5 (0.04)	103 (0.79)	23 (0.18)	5 (0.04)
Allele Frequency	T= 229 (0.87)	C= 33 (0.13)		C= 229 (0.87)	T= 33 (0.13)	
HWE p value		p= 0.03			p= 0.03	

It is important to note that both SNPs for GRIN2B are in complete linkage disequilibrium (LD). LD was calculated using the package 'genetics' (Warnes et al., 2015) for the software program for statistical computing R (R Development Core Team, 2017). Because of this, only g.41206762 C>T was included in the analyses and summary statistics.

Table 6. Hardy-Weinberg Equilibrium (HWE), genotype and allele frequencies for the *DOCK8* SNP. Number of individuals for each genotype and frequency of genotypes (within parenthesis). Number of alleles in total sample (131) appear outside of parenthesis and frequency in parenthesis. P-value for HWE

DOCK8 SNP	g.22496787 T>C		
	CC	CT	TT
Genotype frequency	19 (0.15)	63 (0.48)	49 (0.37)
Allele Frequency	C= 101 (0.39)	T= 161 (0.61)	
HWE p value		p=1	

4.2-Summary statistics

Summary statistics of performance traits for the total sample can be found in Table 6. Summary statistics for total career performance stratified by sex can be found in Appendix I (Table I.1) Summary statistics for performance traits for the different Speed groups can be found in Appendix II (Table II.1). Summary statistics for the two different age groups (horses up to 13 years of age and horses over 13 years) can be found in Appendix III (Tables III.1 and III.2).

Table 7. Summary statistics of performance traits for the total sample (n=131) over their total racing career

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Breeding field test points[1]	6.5	8.5	9.0	9.0	9.0	10.0	0.60
P1 (250m) time record[2]	20.80	22.36	23.13	23.54	24.53	30.20	1.73
P2 (100m) time record[3]	7.20	7.73	8.13	8.23	8.52	10.98	0.68
P3 (150m) time record[4]	13.74	15.08	15.88	16.15	16.77	24.43	1.68
Age of training start for pace[5]	3.0	5.0	6.5	7.0	8.0	16.0	2.67
Height to withers (cm)[5]	128	139	141	141	144	150	3.97
EBV_overall	84.0	102.0	108.0	107.3	113.0	127.0	8.24
EBV_pace	90.0	106.5	112.0	111.2	117.0	128.0	8.29

1: Based on information from72 horses.
2: Based on information from 75 horses.
3: Based on information from 130 horses.
4: Based on information from 66 horses.
5: Based on information from 110 horses.
6: Based on exact height from 106 horses.

4.3-Association analysis: group comparisons

The height distribution in the speed groups (Fast, Average and Slow) was practically the same. Group 1 had 19 horses that were equal or over 140cm and 18 horses under 140 cm; Group 2 had 38 horses over 140 cm and 36 under and Group 3, 10 horses over 140 and 10 under. This equal distribution of heights suggested that for this sample (n=131), Height to withers as a trait relevant to performance was not important.

When performing Fisher's exact test on all three groups (Fast, Average and Slow), no significant differences in genotype distribution of either *MSTN* SNPs were found. However,

when comparing only the Fast and Slow individuals, significant differences were found in genotype distribution for PR3737 and PR8604 ($p<0.01$), but not for PR9482.

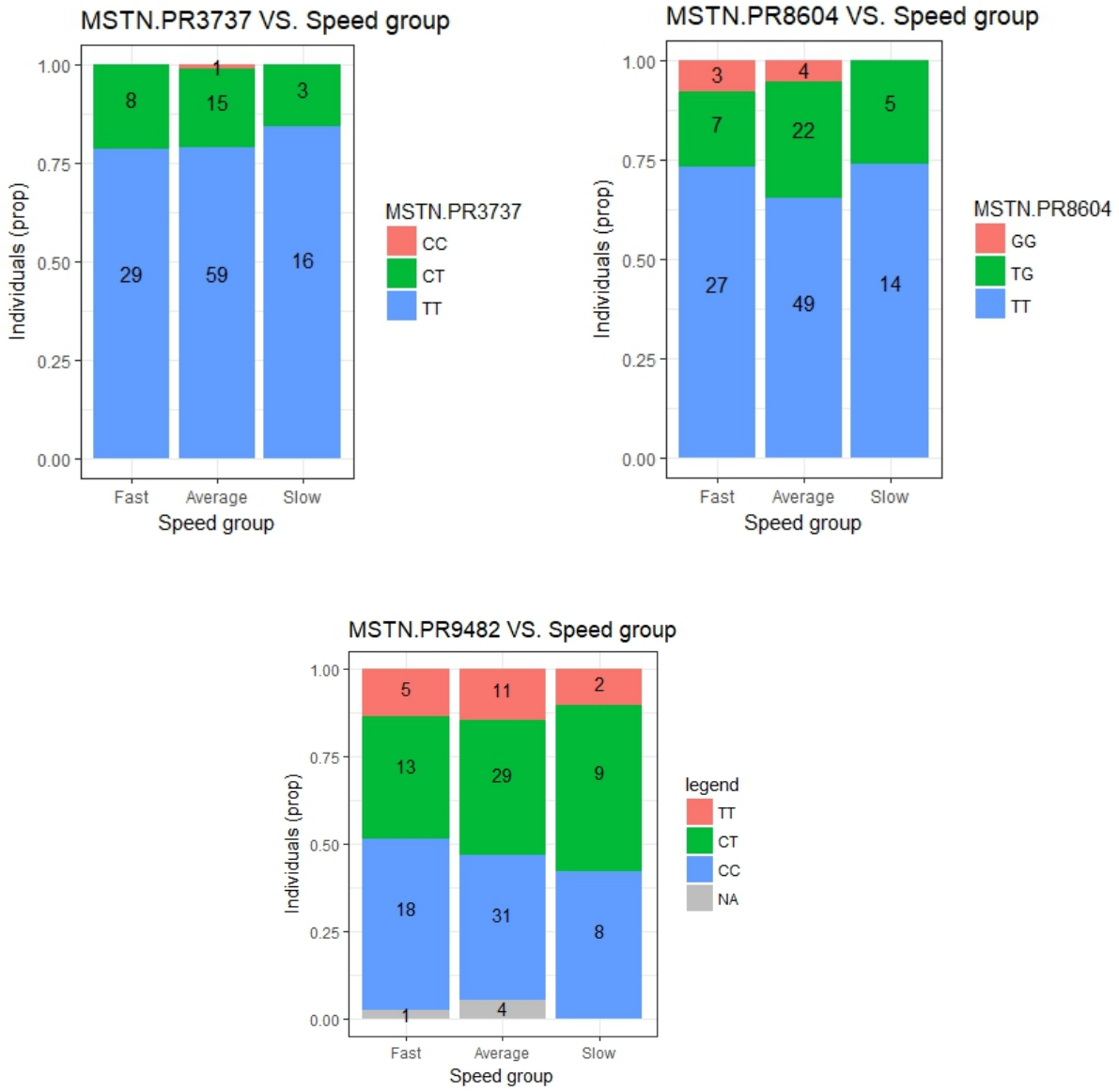

Figure 1. Genotype distribution with number of individuals of the MSTN SNPs among the speed groups. Mutant homozygote individuals appear represented in pink, heterozygotes in green and reference homozygotes in blue.

When performing Fisher's exact test on all three groups (Fast, Average and Slow), no significant differences in genotype distribution of the *GRIN2B* SNP were found. However, comparing Fast versus Slow horses yielded significant differences in genotype distribution ($p<0.01$).

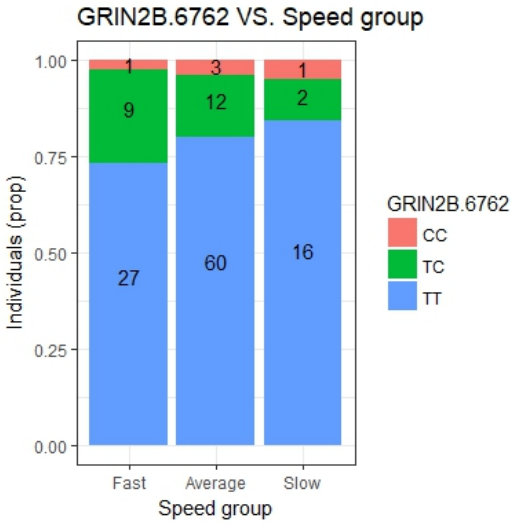

Figure 2. Genotype distribution with number of individuals of the GRIN2B SNP among the speed groups. Mutant homozygote individuals appear represented in pink, heterozygotes in green and reference homozygotes in blue.

When performing Fisher's exact test on all three groups (Fast, Average and Slow), no significant differences in genotype distribution were found. However, comparing Fast versus Slow horses did give significant differences in distribution ($p<0.01$).

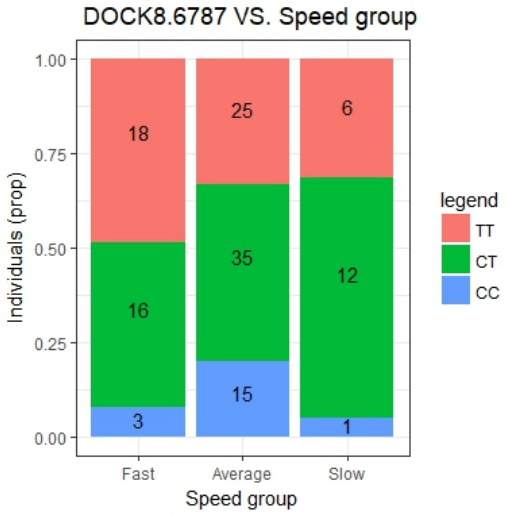

Figure 3. Genotype distribution with number of individuals of the DOCK8 SNP among the speed groups. Mutant homozygote individuals appear represented in pink, heterozygotes in green and reference homozygotes in blue.

Table 8. Comparison of allele frequencies and number of alleles (in parenthesis) among the different Speed groups. Included are p-values for Fisher's exact test performed on the number of reference versus mutant alleles for each SNP between all three Speed groups

Group 1 Allele frequencies					
SNP Name	PR3737	PR8604	PR9482	GRIN2B.6762	DOCK8.6787
Reference Allele	T= 0.89 (66)	T= 0.82 (61)	C= 0.68 (49)	T= 0.85 (63)	C= 0.30 (22)
Mutant Allele	C= 0.11 (8)	G= 0.18 (13)	T=0.32 (23)	C= 0.15 (11)	T= 0.70 (52)
Group 2 Allele frequencies					
SNP Name	PR3737	PR8604	PR9482	GRIN2B.6762	DOCK8.6787
Reference Allele	T= 0.89 (133)	T= 0.80 (120)	C= 0.64 (91)	T= 0.88 (132)	C= 0.43 (65)
Mutant Allele	C= 0.11 (17)	G= 0.20 (30)	T=0.36 (51)	C= 0.12 (18)	T= 0.57 (85)
Group 3 Allele frequencies					
SNP Name	PR3737	PR8604	PR9482	GRIN2B.6762	DOCK8.6787
Reference Allele	T= 0.92 (35)	T= 0.86 (33)	C= 0.66 (25)	T= 0.89 (34)	C= 0.37 (14)
Mutant Allele	C= 0.08 (3)	G= 0.14 (6)	T=0.34 (13)	C= 0.10 (4)	T= 0.63 (24)
Fisher's exact test	p= <0.01	p= <0.01	p= <0.01	p= <0.01	p= <0.01

Significant differences in allele frequency distribution are found both when comparing all three groups (Fast, Average and Slow) and only Fast versus Slow, with $p<0.01$ in both cases.

Lastly, significant differences in temperament trait score distribution were found when comparing all three genotypes of the *GRIN2B* SNP ($p<0.01$) as well as when comparing distribution between only TT and TC individuals ($p>0.01$).

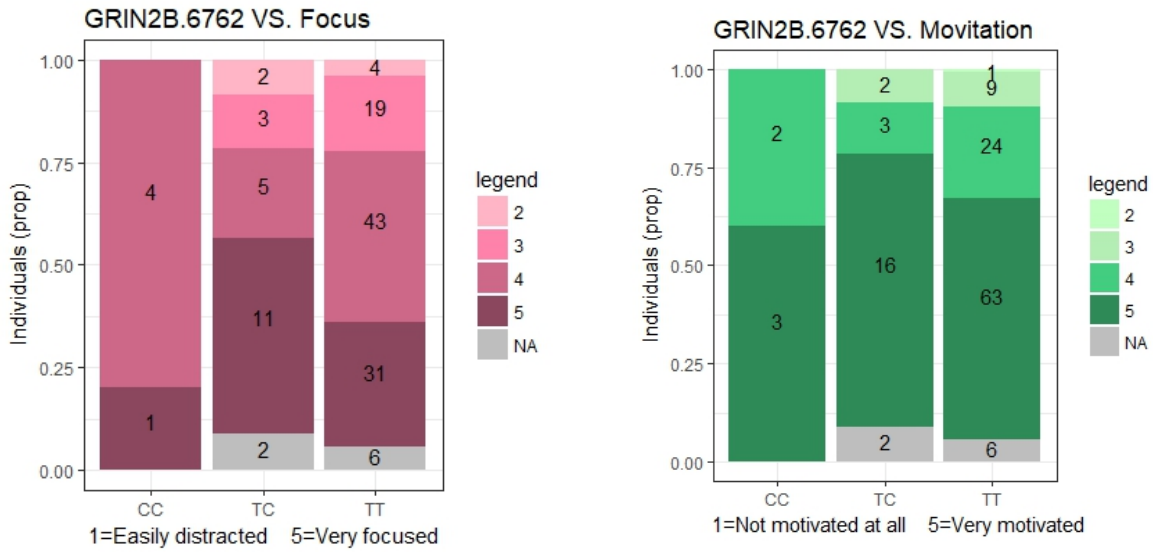

Figure 4. Distribution of scores for temperament traits Focus and Motivation among the different speed groups.

4.4-Association analysis: models

Summary of performance traits and p-values for all age groups (total career, <=13 years and >13 years) can be found in Appendix IV. Specifically, results are found in Table IV.1a to Table IV.1c for PR3737; Table IV.2a to Table IV.2c for PR8604; Table IV.3a to Table IV.3c for PR9482; Table IV.4a to Table IV.4c for *GRIN2B*.6762 and Table IV.5a to Table IV.5c for *DOCK8*.6787.

5-Discussion

The most interesting results of this study are arguably those associated with the *DOCK8* SNP. Starting with the allele frequencies, it was observed that the sample of 131 horses used in the study presented a remarkably low number of homozygote reference-allele individuals (14.5%) as opposed to the high heterozygote (48%) and homozygote mutant-allele (37%) individuals. Not only that, but comparing frequencies of the mutant allele of the Fast group (0.70) and horses from other studies yielded surprising results. This particular SNP presented an interesting range of allele frequencies between four and five gaited horses, as evidenced by genotypes extracted from the study conducted by Jäderkvist et al. (2017). In their study, the four-gaited horses had a frequency of the mutant allele of 0.34, while in the five-gaiters it was of 0.41. Out of 95 Icelandic horses sampled in Iceland and Sweden with performance data from breeding shows, the mutant allele frequency was 0.33, while in data from 185 horses with no competition or BFT data it was of 0.43. It was thus a dramatic difference in the mutant allele frequencies between horses of our Fast group and what may represent the rest of the population. This indicates that the mutant allele of this *DOCK8* SNP is beneficial to pace-racing, since it is found in such high frequency in the elite pace-racing horses. A weak association was also found between *DOCK8* genotype and performance when conducting the analysis of racing time records using linear models. Associations were found between this SNP and P2 (100 m) racing for Total career and in the age group horses under and including 13 years of age (see Table IV.5), where homozygote mutant TT horses performed better than CT, who also performed better than CC horses. However, this was not the case for the older age group, where CT horses seemed to perform better or similarly to TT and always better than CC individuals. Nonetheless, the overall p-value for the effect of DOCK8 genotype was not below 0.05 in the younger age group and no associations were found for the older cohort. The association with only one of the three race disciplines and in only one group opens questions as to why this SNP only affects performance over 100 m and warrants further research.

Concerning the comparisons between the speed groups, results are somewhat conflicting. While it is safe to trust that the horses in the Fast group are correctly classified, the rest of horses may not be as fast for a myriad of reasons that have nothing to do with genetics. Horses in the Average group may have been trained for aerobic metabolism and condition, which would inevitably decrease their chances at excelling in a sport like pace-racing due to lack of sufficient fast-twitch, anaerobic muscle fibers, proven to be vital in this discipline (Stefándsóttir, 2015). The horses classified as Slow, while strictly categorised, could still be unsuccessful pace-racers at least partly due to incorrect training or due to having less skilled riders, for example.

Comparisons between the Fast and Slow groups also yielded interesting results, where significant differences are found for the *DOCK8* SNP. These differences become more obvious when comparing allele frequency and distribution among the groups (Table 7), where significant differences were found between all three Speed groups. If we look only at Fast versus Slow horses, we can see that the frequency of mutant alleles for *DOCK8* is slightly lower in the Slow group versus the Fast-again reinforcing that the T allele may be influencing pace-racing performance positively. This pattern is also observed when comparing genotype distribution, where a slight tendency for more T (mutant) alleles in the Fast group versus the Slow group was observed.

Concerning the *GRIN2B* SNPs, which were in complete LD, frequencies for the minor allele were quite low (0.13), with only 18% heterozygote individuals in the study sample of 131 horses. Comparing mutant allele frequencies of the Fast group (0.15) with 280 Icelandic horses genotyped for other studies yielded no differences whatsoever (also 0.15). When comparing genotype distribution between Speed groups, a significant difference was found in both genotype and allele frequency distribution between the Fast and Slow groups, with the Fast cohort showing a frequency of 15% mutant allele and the Slow cohort, 10%.

Significant differences were also found in the distribution of temperament scores in relation to the different *GRIN2B* SNP genotypes, even when excluding the 5 CC individuals. Said CC individuals always counted with the highest scores for all three traits. Similarly, TC individuals also had a slight tendency to include more highly-scored individuals for all traits than the TT individuals.

As for the analysis of racing times using linear models for this *GRIN2B* SNP, significant associations were found for their Total career and for horses under and including 13 years of age, for P2 and P3 (Table IV.4). However, this changed depending on whether the 5 CC individuals were included or not. Except for P3 models in Total career data, where the association remained significant in favor of the mutant allele, all significant associations were lost. Considering that the association remains significant for the Total career dataset even when removing the 5 CC individuals, one could conclude that the C allele is beneficial to performance in P3 150m race. The question remains why the C allele would be beneficial for P3 but not for P1. As always, a larger sample number could clear up the veracity of these associations.

When studying the allele frequencies of the *MSTN* SNPs in the Icelandic horse, selection towards the so-called "stamina" alleles can be observed for all three SNPs included in the study. If we begin by examining the distribution for PR3737, the same pattern of high frequency of T allele can be observed in the sample of 131 horses used in this study (T= 0.89). These results do not differ from frequencies seen in Icelandic horses genotyped for other studies. Comparing the frequencies of the mutant allele for this SNP in the group of Fast horses (n=37) and in horses from other studies (n= 408) yields no remarkable differences (MAF = 0.11 and MAF= 0.09 respectively). For PR8604, a relatively low frequency of the mutant allele was found (0.26) for both the total sample and only the Fast group individuals. These results again concur with other studies, where MAF of 408 Icelandic horses was of approximately 0.20. PR9482

counted with the highest frequency of heterozygote individuals among the three *MSTN* SNPs (41 %). MAF of the sample was of 0.35 (0.32 for the Fast group horses), which once more did not differ greatly from 280 Icelandic horses genotyped for other studies, where the MAF was of 0.37. Comparisons between the speed groups did not show a clear, different distribution among the Fast and Slow groups as with the other SNPs. Nonetheless, significant differences in genotype distribution were found when comparing Fast and Slow horses for both PR3737 and PR8604. Additionally, allele frequency comparisons between the groups yielded significant differences for all three SNPs, although no clear patterns could be observed the between Fast and Slow groups.

Concerning results from the analysis of racing time records using linear models, no significant associations were found between genotypes of PR3737 and performance (Table IV.1). Moreover, no clear patterns of better performance of one genotype or the other can be observed when looking at means and medians of the different traits. Concerning PR8604, associations were only found in the group of horses raced up until 13 years of age with P1 (Table IV.2), where TG horses performed better than GG horses. However, the analyses excluding the 6 GG individuals gave no significant results. A larger number of GG individuals would be required to determine whether this genotype actually influences performance in 250 m pace race. According to the literature, the G allele should rather enhance performance at short distances.

For PR9482, no associations were found in the Total career performance, but there were some in both age groups (Table IV.3). For horses 13 years and under, significant but unclear effects were seen for both P1 and P3, where CC and CT horses performed best for P1, but were worse in P3 when comparing means, but not when comparing medians. In any case, if the C allele associated to stamina enhances performance for P1, while the speed allele T is better in shorter distances (150m), it is surprising that no associations were seen with the T allele and P2 (100 m) times. Similarly, in the age group of horses over 13 years, the CC individuals performed worse than the CT individuals for P3, which was consistent with results from the younger age group. Nevertheless, these results are still unclear and somewhat inconsistent and no strong conclusions can be drawn from them.

6-Conclusions

This study analyzed the effect of various SNPs of the *MSTN*, *GRIN2B* and *DOCK8* genes on the racing performance of Icelandic horses competing in pace-races. A significant difference in genotype distribution was found between horses categorized as Fast and those as Slow for all studied SNPs except *MSTN* PR9482 when using Fisher's exact test ($p<0.01$). Furthermore, significant differences in allele frequency distribution were found among horses categorized as Fast and Icelandic horses used for other purposes at the *DOCK8* SNP, with a drastic increase in mutant allele frequency in the Fast group. These comparisons suggest that there is a difference in genotype distribution of these SNPs between fast, elite athlete horses competing in pace racing and other individuals of the breed and warrants further investigation. Furthermore, the results in this study suggest a possible relationship with the mutant C allele of a GRIN2B SNP and temperament traits in association with racing performance.

However, results from the analysis of pace-racing time records indicate that MSTN, a crucial gene in gallop racing breeds, plays little to no role in pace-racing. This opens the door to further

research in search of other genetic factors that affect racing performance in this breed, which may well be associated also with gaiting ability.

A genome-wide association study on this sample may shed light on genes that play a role in gaiting ability at high speed and in neurological aspects related to temperament as well. This study also portrays the importance of studying the effects a single gene may have in different breeds, seeing as allele frequencies alone may vary greatly between horses bred for different goals and destined to compete in radically different sport disciplines.

7- Acknowledgements

I would like to thank many people for this project. First and foremost, to Gabriella Lindgren, my main supervisor, for giving me the opportunity to continue working in her research group and carry out this project. Special thanks as well to Susanne Eriksson and Thorvaldur Árnason, also supervisors, for all their teachings, patience and guidance during the project. To Åsa Viklund for being my examiner. Thank you to all the horse owners and breeders who shared their interest in the project and contributed with samples-without you, this study would not have been possible. Special thanks go to Anna Munters for constantly helping me with her contacts and for interviewing me. To Elfa Ágústdóttir, whose invaluable work sampling in Iceland granted me with a fourth of the samples included in the study. A great thanks also to the Horses of Iceland project, who not only showed interest in the project through an interview, but also invited me to assist to Landsmót Hestamanna 2018 and give a seminar of my research on site. I am incredibly grateful for all the interest and support everyone has shown in this study.

8-Study Contributions

I, the author, performed the preparation and genotyping of DNA, structured the data for analysis with custom scripts written in the statistical software R and performed all statistical analysis. Preliminary results of this study have been presented in a seminar given at Landsmót Hestamanna 2018 (Iceland) and through two interviews by Horses of Iceland and Islandshästpodden. Gabriella Lindgren, Susanne Eriksson and Thorvaldur Árnason, scientists at the department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Sweden, helped the author conceive and design the experiment for this study. The Animal Genetics Laboratory, Swedish University of Agricultural Sciences, Uppsala, Sweden contributed with materials and analysis tools as well as a small number of samples available in the Animal Biobank. Gabriella Lindgren, Susanne Eriksson and Thorvaldur Árnason supervised the project and contributed with input on my thesis and guidance throughout the whole project.

9-References

Akazawa C., Shigemoto R., Bessho Y., Nakanishi S. and Mizuno N. (1994). Differential expressionof five Nmethyl- D-aspartate receptor subunit mRNAs in the cerebellum of developing and adult rats. J Comp Neurol. 1994; 347:150–160. [PubMed: 7798379].

Albertsdóttir, E. (2010). Integrated genetic evaluation of breeding field test traits, competition traits and test status in Icelandic horses. Doctoral thesis. Agricultural University of Iceland, Faculty of Land and Animal Resources.

Amano T., Onogi A., Yamada F., Kawai M., Shirai K. and Ueda J. (2018). Genome-wide association mapping and examination of posible maternal effect for the pace trait of horses. Animal genetics 2018 Aug 14. doi: 10.1111/age.12711.

Andersson LS., Larhammar M., Memic F., Wootz H., Schwochow D., Rubin CJ., Patra K, Árnason T., Wellbring L., Hjälm G., Imsland F., Petersen JL., McCue ME., Mickelson JR., Cothran G., Ahituv N., Roepstorff L., Mikko S., Vallstedt A., Lindgren G., Andersson L. and Kullander K. (2012) Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice, *Nature*, vol. 488, pp. 642-646.

Árnason T. (1979). Studies on traits in the Icelandic toelter horses. J. Agr. Res. Icel., 11, 81–102.

Árnason T. (1982). Prediction of breeding values for multiple traits in small non-random mating (horses) populations. Acta Agric. Scand., 32, 171–176.

Árnason T. (1984). Genetic studies on conformation and performance of Icelandic Toelter horses: IV Best linear unbiased prediction of ten correlated traits by use of an animal model. Acta Agric. Scand. 34, 450–462.

Árnason T. (2004). Worldfengur project 2004 – genetic parameters used in BLUP analysis. Available at: http://www. ihbc.se/app/servlets/arfgengien.html

Árnason T., Sigurdsson A. (2004). International genetic evaluations of the Icelandic horse. 55th Annual Meeting of the EAAP, Bled, Slovenia, September 5th-9th, 2004. Commission on Animal Genetics (G) and Commission on Horse Production (H), Session V. Paper HG5.3.

Árnason, T. (2014). International Genetic Evaluations with the BLUP method 2014. Available at: http://www.worldfengur.com/temp/International_BLUP.pdf#zoom=100

Bates D., Maechler M., Bolker B. and Walker S. (2018). Package 'lme4': Linear Mixed-Effects Models using 'Eigen' and S4. Available at: https://cran.r-project.org/web/packages/lme4/lme4.pdf

Bower MA., McGivney BA., Campana MG., Gu J., Andersson LS., Barrett E., Davis CR., Mikko S., Stock F., Voronkova V., Bradley DG., Fahey AG., Lindgren G., MavHugh DE., Sulimova G. and Hill E. (2012) The genetic origin and history of speed in the Thoroughbred racehorse. *Nature Communications*, 3:643, DOI: 10.1038/ncomms1644.

Burbank AJ., Shah SN., Montgomery M., Peden D., Tarrant TK. and Weimer ET. (2016). Clinically focused exome sequencing identifies an homozygous mutation that confers *DOCK8* deficiency. Pediatr Allergy Immunol. 2016;27:96-98.

Central European Championship 2018 at St. Radegund, Austria. Available at: https://www.mem2018.at/ergebnisse/

Cohen S. and Greenberg ME. (2008). Communication between the synapse and the nucleus in neuronal development, plasticity, and disease. Annu Rev Cell Dev Biol. 2008; 24:183–209. [PubMed: 18616423].

Dominiczak MH. and Baynes J. (2005). Bioquimica Medica: con acceso a Student Consult (Spanish Edition). Elsevier Espana. ISBN 84-8174-866-8.

Esbjornsson M., Hellsten-Westing Y., Balsom, PD, Sjodin B., and Jansson E. (1993). Muscle fibre type changes with sprint training: Effect of training pattern. Acta Physiol Scand 149: 245–246.

FEIF (2018). FEIF Breeding rules and regulations. Available at: https://www.feif.org/files/documents/FEIF_Rules_Regulations2018_breeding(1).pdf

FEIF (2018). FEIF Sport rules and regulations. Available at: https://www.feif.org/files/documents/FEIF_Rules_Regulations2018_sport.pdf

Gerard MP., de Graaf-Roelfsema E., Hodgson, DR. & van der Kolk JH. (2014). Energetic considerations of exercise. In: D.R. Hodgson, McKeever K.H. & McGowan C.M. (eds.) The athletic horse, principles and practice of equine sports medicine. 2. ed. USA: Elsevier Saunders, pp. 19-33.

Gille C, Kayser M. & Spiller A. (2010). Target group segmentation in the horse buyers' market against the background of equestrian experience. Journal of Equine Science, 21, 67–71.

Glessner JT., Li J.,Wang D., March M., Lima L., Desai A., Hadley D., Kao C., Gur RE., Cohen N., Sleiman PMA., Li Q., Hakonarson H. and the Janssen-CHOP Neuropsychiatric Genomics Working Group. (2017). Copy number variation meta-analysis reveals a novel duplication at 9p24 associated with multiple neurodevelopmental disorders. Genome Medicine (2017) 9:106. DOI 10.1186/s13073-017-0494-1.

González JR., Armengol L., Solé X., Guíno E., Mercader, JM., Estivill X. and Moreno V. (2007) SNPassoc: an R package to perform whole genome association studies. Available: https://cran.r-project.org/web/packages/SNPassoc/SNPassoc.pdf

Górecka-Bruzda A., Chruszczewski MH., Jaworski Z., Golonka M., Jezierski T., Długosz B. and Pieszka M. (2011a). Looking for an ideal horse: Rider preferences. Anthrozoös, 24, 379–392.

Graf P., von Borstel UK. and Gauly M. (2013). Importance of personality traits in horses to breeders and riders. Journal of Veterinary Behavior: Clinical Applications and Research, 8, 316–325.

Griggs BL., Ladd S., Saul RA., DuPont BR. and Srivastava AK. (2008). Dedicator of cytokinesis 8 is disrupted in two patients with mental retardation and developmental disabilities. Genomics. 2008;91:195-202.

Grobet, L., Martin, L. J. R., Poncelet, D., Pirottin, D., Brouwers, B., Riquet, J., Schoeberlein, A., Dunner, S., Ménissier, F., Massabanda, J., Fires, R., Hanset, R. and Georges, M. (1997) A deletion in the bovine Myostatin gene causes the double-muscled phenotype in cattle, *Nature genetics*, vol. 17, pp. 71-74.

Harris RC., Martin DJ, Snow DH. and Harkness A. (1991). Muscle ATP loss and lactate accumulation at different work intensities in the exercising thoroughbred horse. European Journal of Applied Physiology, 62, pp. 235-244.

Hill EW., Fonseca RG., McGivney BA., Gu J., MacHugh DE. and Katz LM. (2011). *MSTN* genotype (g.66493737C/T) association with speed indices in Thoroughbred racehorses. *J Appl Physiol*, doi: 10.1152/japplphysiol.00793.2011.

Hill EW., Gu J., Eivers SS., Fonseca RG., McGivney BA., Govindarajan P., Orr N., Katz LM. And MacHugh D. (2010a) A Sequence Polymorphism in *MSTN* Predicts Sprinting Ability and Racing Stamina in Thoroughbred Horses. PLOS one, vol. 5(1).

Hill EW., McGivney BA., Gu J., Whiston R. and MacHugh, DE. (2010b) A genome wide SNP association study confirms a sequence variant (g.66493737C>T) in the equine myostatin (*MSTN*) gene as the most powerful predictor of optimum racing distance for Thoroughbred racehorses, *BMC Genomics*, vol. 11, pp. 552.

Horses of Iceland (2018). "Marketing plan developed by stakeholders in the Icelandic horse community to increase the awareness and strengthen the image of the Icelandic horse in international markets". Available at: https://www.horsesoficeland.is

Howald H., Hoppeler H., Claassen H., Mathieu O., and Straub R. (1985). Influences of endurance training on the ultrastructural composition of the different muscle fiber types in humans. Pflugers Arch 403: 369–376.

Hu C., Chen W., Myers SJ., Yuan H. and Traynelis SF. (2016). Human *GRIN2B* variants in neurodevelopmental disorders. J Pharmacol Sci. 2016;132:115-121.

Islandpferde-Reiter- und Züchterverband (2018). Available at: https://www.ipzv.de/home.html

Jansson A, Hedenström U, Ragnarsson S. (2014). A survey of training regimes used by top trainers of the Icelandic pace racing horse. Comparative Exercise Physiology, 2014; 10 (4): 253-257.

Jansson E., Sjodin B. and Tesch P. (1978). Changes in muscle fibre type distribution in man after physical training. A sign of fibre type transformation? Acta Physiol Scand 104: 235–237.

Jansson A., Hedenström U. and Ragnarsson S. (2014). A survey of training regimes used by top trainers of the Icelandic pace racing horse. Comparative Exercise Physiology, 2014; 10 (4): 253-257.

Judson GJ., Frauenfelder HC. and Mooney GJ. (1983). Biochemical changes in thoroughbred racehorses following submaximal and maximal exercise. In: Snow DH., Persson SGB. & Rose RJ. eds. Equine exercise physiology 1. Cambridge UK: Granta Editions. pp. 408-415.

Jäderkvist K., Andersson, LS., Johansson AM., Árnason T., Mikko S., Eriksson S., Andersson L. and Lindgren G. (2014a) The DMRT3 'Gait keeper' mutation effects performance of Nordic and Standardbred trotters*, J. Anim. Sci*, vol. 92, pp. 4279-4286.

Jäderkvist K., Kangas N., Andersson LS. and Lindgren G. (2014b). Gaitedness is associated with the DMRT3 'Gait keeper' mutation in Morgan and American Curly horses. Animal Genetics, Vol 45, Issue 6, p 908-909. https://doi.org/10.1111/age.12228.

Jäderkvist Fegraeus K., Johansson L., Mäenpää M., Mykkänen A., Andersson LS., Velie BD., Andersson L., Árnason T and Lindgren G. (2015). Different DMRT3 Genotypes Are Best Adapted for Harness Racing and Riding in Finnhorses. Journal of Heredity, Volume 106, Issue 6, 1 January 2015, Pages 734–740, https://doi.org/10.1093/jhered/esv062.

Jäderkvist Fegraeus K, Hirschberg I, Arnason T, Andersson L, Velie BD, Andersson LS, Lindgren G (2017). To pace or not to pace: a pilot study of four- and five-gaited Icelandic horses homozygous for the DMRT3 'Gait Keeper' mutation. Anim Gent. 2017;48:694 697.

Kambadur R., Sharma M., Smith, TPL. and Bass JJ. (1997) Mutations in Myostatin (GDF8) in Double-Muscled Belgian Blue and Piedmontese Cattle. *Cold Spring Harbor Laboratory Press*, vol. 7, pp. 910-915.

Karp JR. (2001). Muscle fiber types and training. Strength Cond J 23: 21–26.

Kim JI., Yoo JH, Kim D., Jeong B. and Kim BN. (2017). The effects of *GRIN2B* and DRD4 gene variants on local functional connectivity in attention-deficit/hyperactivity disorder. Brain Imaging Behav. 2017; doi: 10.1007/s11682-017-9690-2.

Koenen EPC, Aldridge LI and Philipsson J. (2004). An overview of breeding objectives for warmblood sport horses. Livestock Production Science, 88, 77–84.

Kraemer WJ., Patton JF., Gordon SE., Harman EA., Deschenes MR., Reynolds K., Newton RU., Triplett NT., and Dziados JE. (1995). Compatibility of high-intensity strength and endurance training on hormonal and skeletal-muscle adaptations. J Appl Physiol 78: 976–989.

Kristjansson T, Bjornsdottir S, Sigurdsson A, Andersson L.S., Lindgren G, Helyar S.J, Klonowski A.M & Arnason T (2014). The effect of the 'Gait keeper' mutation in the DMRT3 gene on gaiting ability in Icelandic horses. J. Anim. Breed. Genet. ISSN 0931-2668.

Kristjansson T, Bjornsdottir S, Albertsdóttir E, Sigurdsson A, Pourcelot P, Crevier-Denoix N, Árnason T (2016). Association of conformation and riding ability in Icelandic horses. Journal of Livestock Science. DOI: 10.1016/j.livsci.2016.05.010.

Kutsuwada T., Sakimura K., Manabe T., Takayama C., Katakura N. and Kushiya E. (1996). Impairment of suckling response, trigeminal neuronal pattern formation, and hippocampal LTD in NMDA receptor epsilon 2 subunit mutant mice. Neuron. 1996; 16:333–344. [PubMed: 8789948].

Kuznetsova A., Brockhoff PB. and Bojesen Christensen RH. (2018). Package ´lmerTest': Tests in Linear Mixed Effects Models. Available at: https://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf

Landsmót Hestmanna (2018). The National Icelandic Horse Competition. Available at: http://www.landsmot.is/en/about-landsmot

Lindholm A. and Saltin B. (1974). The physiological and biochemical response of Standardbred horses to exercise of varying speed and duration. Acta Veterinaria Scandinavica, 15, pp. 310-324.

Marlin D. & Nankervis K. (2002). Equine Exercise Physiology. UK: Wiley-Blackwell.

McMiken DF. (1983). An energetic basis of equine performance. Equine Veterinary Journal, 15(2), pp. 123-133.

McPherron AC. and Lee SJ. (1997) Double muscling in cattle due to mutations in the myostatin gene. *Proc. NAtl. Acad. Sci. USA*, vol. 94, pp. 12457-12461.

McPherron AC., Lawler AM. and Lee SJ. (1997) Regulation of skeletal muscle mass in mice by a new TGF-β superfamily member, *Letters to nature*, vol. 387, pp. 83-90.

Mosher DS., Quignon P., Bustamante CD., Sutter NB., Mellersh CS., Parker HG. And Ostrander EA. (2007) A Mutation in the Myostatin Gene Increases Muscle Mass and Enhances Racing Performance in Heterozygote dogs. *PLOS Genetics*, vol. 3(5): e79. Doi: 10.1371/journal.pgen.0030079.

Monyer H., Sprengel R., Schoepfer R., Herb A., Higuchi M. Lomeli H., Burnashev N., Sakmann B. and Seeburg PH. (1992). Heteromeric NMDA receptors: molecular and functional distinction of subtypes. Science 256, 1217±1221.

Nielsen BD. (2014). Training the racing Quarterhorse. In: Hodgson DR., McKeever KH. & McGowan CM. (eds). The athletic horse, principles and practice of equine sports medicine. 2 ed. USA: Elsevier Saunders. pp. 354-360.

Novoa-Bravo M., Jäderkvist Fegraeus K., Rhodin M., Strand E., García LF. and Lindgren G. (2018). Selection on the Colombian paso horse's gaits has produced kinematic differences partly explained by the DMRT3 gene. PLoS ONE 13(8): e0202584. https://doi.org/10.1371/journal.pone.0202584.

Passklubben Skeið (2018). Available at: http://www.skeid.se/

Petäjistö K. (2016). The role of Myostatin in Coldblooded Trotter harness racing performance. Unpublished master's thesis. Faculty of Veterinary Medicine and Animal Science, Swedish University of Agricultural Sciences, Uppsala, Sweden. Available at: https://stud.epsilon.slu.se/9762/

Qin T., An Y., Liu C., Wu J., Dai R., Liu D., Li X., Jiang L., Wu D., Tang X., Song W., Wang T. and Zhao X. Novel *DOCK8* gene mutations lead to absence of protein expression in patients with hyper-IgE syndrome. Immunol Res. 2016;64:260-271.

R Development Core Team (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ruusala A, Aspenström P (2004). Isolation and characterization of *DOCK8*, a member of the DOCK-180 related regulators of cell morphology. FEBS Letters 2004;572:159–166. [PubMed: 15304341].

Räsänen, L.A., Lampinen, K.J. & Pösö, A.R. (1995). Responses of blood and plasma lactate and plasma purine concentrations to maximal exercise and their relation to performance in Standardbred trotters. American Journal of Veterinary Research, 56(12), pp. 1651-1656.

Santagostino M., Khoriauli L., Gamba R., Bonuglia M., Klipstein O., Piras MF., Vella F., Russo A., Badiale C., Mazzagatti A., Raimondi E., Nergadze SG. and Giulotto, E (2015). Genome-wide evolutionary and functional analysis of the Equine Repetitive Element 1: an insertion in the myostatin promoter affects gene expression. *BMC Genetics*, 16:126. DOI 10.1186/s12863-015-0281-1.

Scott W., Stevens J. and Binder-Macleod SA. (2001). Human skeletal muscle fiber type classifications. Phys Ther 81: 1810–1816.

Segerstrom SC. and Miller GE (2004). Psychological Stress and the Human Immune System: A MetaAnalytic Study of 30 Years of Inquiry. Psychol Bull. 2004 July ; 130(4): 601–630.

Sigurðardóttir H., Albertsdóttir E. & Eriksson S. (2017): Analysis of new temperament traits to better understand the trait spirit assessed in breeding field tests for Icelandic horses, Acta Agriculturae Scandinavica, Section A — Animal Science, DOI: 10.1080/09064702.2017.1383507.

Simoneau JA., Lortie G., Boulay MR., Thibault MC., Theriault G. and Bouchard C. (1985). Skeletal-muscle histochemical and biochemical characteristics in sedentary male and female subjects. Can J Physiol Pharm 63: 30–35.

Stefánsdóttir GJ. (2015). Physiological response to exercise in the Icelandic horse. Doctoral Thesis. Swedish University of Agricultural Sciences, Uppsala 2015. Acta Universitatis agriculturae Sueciae 2015:84. Available at: https://pub.epsilon.slu.se/12538/

Tang YP., Shimizu E., Dube GR., Rampon C., Kerchner GA., Zhuo M., Liu G. and Tsien JZ. Genetic enhancement of learning and memory in mice. Nature. 1999;401:63-69.

Thomas M., Langley B., Berry C., Sharma M., Kirk S., Bass J. and Kambadur R. (2000) Myostatin, a Negative Regulator of Muscle Growth, Functions by Inhibiting Myoblast Proliferation, The Journal of Biological Chemistry, vol. 257, pp. 40235-40243.

Thorén Hellsten E., 2008. International sport horse data for genetic evaluation. Doctoral thesis No. 2008:32. Faculty of Veterinary Medicine and Animal Science. Swedish university of Agricultural Sciences.

Tozaki T., Miyake T., Kakoi H., Gawahara H., Sugita S., Hasegawa T., Ishida N., Hirota K.

and Nakano Y. (2010). A genome-wide association study for racing performances in Thoroughbreds clarifies a candidate region near the *MSTN* gene. *Animal Genetics*, vol. 41 (Suppl. 2), pp 28-35.

Tozaki T. Hill EW., Hirota K., Kakoi, H., Gamahara H., Miyake T., Sugita S., Hasegawa T., Ishida N., Nakano Y. and Kurosawa, M. (2011a) A cohort study of racing performance in Japanese Thoroughbred racehorses using genome information on ECA18. *Animal Genetics*, vol. 43, pp. 42-52.

Tozaki T., Sato F., Hill E. M., Miyake T., Endo Y., Kakoi H., Gawahara H., Hirota K., Nakano Y., Nambo Y. and Kurosawa M. (2011b) Sequence Variants at the Myostatin Gene Locus Influence the Body Composition of Thoroughbred Horses, *J. Vet. Med. Sci*, vol. 73(12), pp. 1617-1624.

Vanderka M. and Kampmiller T. (2013). Kinematics of Sprinting in children and youths. The IAAF technical quarterly. 28:1/2; 35-45, 2013.

Velie BD., Jäderkvist Fegraeus K., Solé M., Rosengren MK., Røed KH., Ihler CF., Strand E. and Lindgren G. (2018). A genome-wide association study for harness racing success in the Norwegian-Swedish coldblooded trotter reveals genes for learning and energy metabolism. BMC Genetics (2018) 19:80. https://doi.org/10.1186/s12863-018-0670-3.

Viluma A. (2012). Polymorphism in *Myostatin* gene and athletic performance in Nordic horse breeds. *Unpublished master's thesis*. Faculty of Veterinary Medicine and Animal Science, Swedish University of Agricultural Sciences, Uppsala, Sweden. Available at: https://stud.epsilon.slu.se/5083/13/viluma_a_121205.pdf

Von Borstel UK. (2013). Assessing and influencing personality for improvement of animal welfare: a review of equine studies. CAB Reviews Perspectives in Agriculture Veterinary Science Nutrition and Natural Resources. DOI: 10.1079/PAVSNNR20138006

Warnes G., Gorjanc G., Leisch F. and Man M. (2015). Package 'genetics': Population genetics. Available at: https://cran.r-project.org/web/packages/genetics/genetics.pdf

Watts AS., Coleman I. and Nevill A. (2012) The changing shape characteristics associated with success in world-class sprinters, Journal of Sports Sciences, 30:11, 1085-1095, DOI: 10.1080/02640414.2011.588957

Wilson JM., Loenneke JP., Jo E., Wilson GJ., Zourdos MC. And Kim JS. (2012). he effects of endurance, strength, and power training on muscle fiber type shifting. Journal of Strength and Condition research, 2012 Jun;26(6):1724-9. doi: 10.1519/JSC.0b013e318234eb6f.

Zhang Q., Davis JC., Lamborn IT., Freeman AF., Jing H., Favreau AJ., Matthews HF., Davis J., Turner ML., Uzel G., Holland SM. and Su HC. (2009). Combined immunodeficiency associated with *DOCK8* mutations. N Engl J Med. 2009;361:2046-2055.

APPENDIX I

Table I.1-Summary statistics of performance traits for the total sample (n=131) stratified by sex (n= 38 stallions, n= 49 mares and n=44 geldings).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Breeding field test points[1]							
Stallions	8.0	9.0	9.0	9.0	9.0	10.0	0.43
Mares	6.5	8.0	8.5	8.5	9.0	9.5	0.66
Geldings	7.5	8.0	8.5	8.5	9.0	9.0	0.53
P1 (250m) time record[2]							
Stallions	21.93	22.80	23.46	23.85	25.07	26.14	1.37
Mares	21.90	22.93	23.75	23.99	24.57	28.40	1.58
Geldings	20.80	21.57	22.36	22.85	23.66	30.20	1.96
P2 (100m) time record[3]							
Stallions	7.35	7.80	8.29	8.35	8.80	10.60	0.70
Mares	7.29	7.93	8.11	8.27	8.45	10.98	0.70
Geldings	7.20	7.65	8.00	8.10	8.39	9.81	0.64
P3 (150m) time record[4]							
Stallions	14.93	16.57	16.74	16.96	17.48	19.12	1.23
Mares	13.86	15.11	15.59	15.86	16.64	18.44	1.09
Geldings	13.74	14.59	15.48	16.05	16.52	24.43	2.25
Age of training start for pace[5]							
Stallions	3.0	5.0	6.0	7.0	8.0	12.0	3.58
Mares	5.0	6.0	6.0	7.0	8.0	12.0	1.79
Geldings	5.0	6.0	8.0	8.0	8.0	16.0	2.63
Height to withers (cm)[6]							
Stallions	136	142	143	143	145	150	2.99
Mares	130	138	140	140	142	150	3.93
Geldings	128	140	141	141	145	148	4.00
EBV overall							
Stallions	94.0	108.2	113.5	112.4	118.8	127.0	7.71
Mares	84.0	102.0	106.0	105.5	110.0	116.0	7.09
Geldings	85.0	100.5	106.0	105.0	110.0	120.0	8.11
EBV pace							
Stallions	94.0	109.2	116.5	114.3	119.0	125.0	7.21
Mares	91.0	106.0	111.0	111.0	118.0	128.0	8.39
Geldings	90.0	102.0	110.0	108.8	114.0	123.0	8.38

1: Based on information from 31 stallions, 32 mares and 9 geldings.
2: Based on information from 23 stallions, 25 mares and 27 geldings.
3: Based on information from 38 stallions, 48 mares and 44 geldings.
4: Based on information from 13 stallions, 28 mares and 25 geldings.
5: Based on information from 32 stallions, 44 mares and 34 geldings.
6: Based on exact height from 31 stallions, 45 mares and 30 geldings.

.

APPENDIX II

Table II.1-Summary statistics of performance traits stratified by Speed group classification (Group 1= fast (n=37); Group 2= average/inconsistent (n=75) and Group 3= slow (n=19)).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Breeding field test points[1]							
Group1	8.0	9.0	9.0	9.0	9.0	10.00	0.45
Group2	6.5	8.0	9.0	9.0	9.0	9.5	0.62
Group3	7.5	8.0	8.5	8.0	8.5	8.5	0.39
P1 (250m) time record[2]							
Group1	20.80	21.62	22.08	22.15	22.76	24.01	0.76
Group2	21.90	23.12	23.73	24.20	24.94	30.20	1.66
Group3	23.85	24.31	25.07	25.18	25.80	27.32	1.29
P2 (100m) time record[3]							
Group1	7.20	7.42	7.57	7.59	7.74	8.19	0.23
Group2	7.57	8.03	8.31	8.50	8.83	10.98	0.68
Group3	8.02	8.20	8.40	8.51	8.70	9.41	0.43
P3 (150m) time record[4]							
Group1	13.74	14.34	14.77	15.10	15.35	18.34	1.22
Group2	14.42	15.39	16.57	16.69	17.11	24.43	1.88
Group3	14.79	15.48	16.31	16.05	16.63	16.89	0.79
Age of training start for pace[5]							
Group1	4.0	5.0	6.0	7.0	8.0	16.0	2.45
Group2	3.0	6.0	7.0	7.0	8.0	21.0	2.63
Group3	5.0	5.5	6.0	8.0	11.0	15.0	3.24
Height to withers (cm)[6]							
Group1	130	140	142	142	145	148	3.62
Group2	128	138	141	141	143	150	4.06
Group3	132	138	140	140	143	147	4.28
EBV overall							
Group1	85.0	103.0	108.0	107.7	113.0	120.0	8.13
Group2	88.0	103.0	108.0	108.2	114.0	127.0	8.38
Group3	84.0	101.0	104.0	103.3	107.5	116.0	6.92
EBV pace							
Group1	90.0	111.0	116.0	114.5	120.0	128.0	8.03
Group2	91.0	106.0	111.0	110.5	117.0	125.0	8.35
Group3	98.0	104.0	108.0	107.6	110.5	122.0	6.48

1: Based on information from 18 horses from Group 1, 47 horses from Group 2 and 7 horses from Group 3.
2: Based on information from 29 horses from Group 1, 36 horses from Group 2 and 10 horses from Group 3.
3: Based on information from 37 horses from Group 1, 74 horses from Group 2 and 19 horses from Group 3.
4: Based on information from 16 horses from Group 1, 37 horses from Group 2 and 13 horses from Group 3.
5: Based on information from 29 horses from Group 1, 66 horses from Group 2 and 15 horses from Group 3.
6: Based on exact height from 29 horses from Group 1, 62 horses from Group 2 and 11 horses from Group 3.

APPENDIX III

Table III.1. Summary statistics of performance traits for horses that competed at the ages of 13 years and under (n=117).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
P1 (250m) time record[1]							
Stallions	21.93	23.35	24.08	24.31	25.20	27.83	1.33
Mares	21.90	23.59	24.07	24.63	25.59	28.47	1.57
Geldings	21.15	22.63	23.48	23.69	24.40	29.79	1.43
Total	21.15	23.05	23.85	24.04	24.87	29.79	1.49
P2 (100m) time record[2]							
Stallions	7.35	8.04	8.47	8.68	9.02	13.50	0.95
Mares	7.32	8.05	8.40	8.68	9.08	13.35	0.93
Geldings	7.20	7.82	8.15	8.27	8.55	12.61	0.68
Total	7.20	7.95	8.32	8.53	8.84	13.50	0.87
P3 (150m) time record[3]							
Stallions	14.93	16.14	16.98	17.05	17.91	19.14	1.34
Mares	14.35	15.07	15.64	16.04	16.68	20.23	1.28
Geldings	13.74	15.03	15.61	15.80	16.41	24.43	1.32
Total	13.74	15.08	15.68	16.02	16.63	24.43	1.34

1: Based on information from 62 horses.
2: Based on information from 115 horses.
3: Based on information from 54 horses.

Table III.2. Summary statistics of performance traits for horses that competed at ages over 13 years (n=64).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
P1 (250m) time record[1]							
Stallions	22.46	23.20	23.86	24.24	24.43	30.28	1.54
Mares	22.37	24.16	25.09	25.10	25.88	28.89	1.53
Geldings	20.80	22.32	22.99	23.29	24.11	30.20	1.42
Total	20.80	22.83	23.74	23.96	24.87	30.28	1.66
P2 (100m) time record[2]							
Stallions	7.47	8.02	8.45	8.64	9.06	11.23	0.81
Mares	7.29	8.18	8.59	8.67	8.99	10.72	0.66
Geldings	7.39	7.87	8.22	8.34	8.71	11.49	0.66
Total	7.29	7.96	8.40	8.50	8.88	11.49	0.71
P3 (150m) time record[3]							
Stallions	16.57	16.79	17.16	17.50	18.18	18.98	0.76
Mares	13.86	15.07	15.72	16.01	16.85	19.02	1.29
Geldings	13.84	14.73	15.34	15.44	15.96	17.80	0.95
Total	13.84	14.95	15.69	15.88	16.80	19.02	1.24

1: Based on information from 43 horses.
2: Based on information from 60 horses.
3: Based on information from 29 horses.

APPENDIX IV

Table IV.1a. Summary of performance traits and p-values according to *MSTN* PR3737 genotype for total career of horses (n=131).

TOTAL CAREER								
PR3737	**TT (n=104)**		**TC (n=26)**		**CC (n=1)**	**P-value**	**P-value**	**P-value**
Performance trait	**Median**	**Mean**	**Median**	**Mean**	**Value**	**TT/TC**	**TC/CC**	**TT/CC**
Breeding field test points[1]	9.0	9.0	9.0	9.0	8.0	0.94	0.41	0.40
P1 (250m) time record[2]	23.36	23.58	23.02	23.39	22.99	0.45	0.82	0.65
P2 (100m) time record[3]	8.15	8.20	8.11	8.37	7.88	0.49	0.66	0.77
P3 (150m) time record[4]	16.15	16.11	15.65	16.36	15.53	0.66	0.77	0.87
Age of training start for pace[5]	7.00	7.00	6.00	7.00	8.00	0.45	0.55	0.66
EBV_overall	108.0	107.4	105.0	107.3	100.0	0.28	0.85	0.66
EBV_pace	112.0	111.1	113.0	111.9	108.0	0.98	0.74	0.73

1: Based on information from72 horses.
2: Based on information from 75 horses.
3: Based on information from 130 horses.
4: Based on information from 66 horses.
5: Based on information from 110 horses.

Table IV.1b. Summary of performance traits and p-values according to *MSTN* PR3737 genotype for horses under and including 13 years of age (n=117).

PR3737	**TT (n=93)**		**TC (n=23)**		**CC (n=1)**		**P-value**			
Performance trait	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**TT/TC**	**TC/CC**	**TT/CC**	**Overall***
P1 (250m) time record[1]	23.89	24.08	23.75	23.96	23.64	23.64	0.66	0.58	0.48	0.71
P2 (100m) time record[2]	8.34	8.53	8.24	8.51	8.26	8.24	0.33	0.33	0.44	0.46
P3 (150m) time record[3]	15.70	16.04	15.60	15.94	15.98	15.87	0.89	0.70	0.65	0.89

1: Based on information from 62 horses.
2: Based on information from 115 horses.
3: Based on information from 54 horses.

Table IV.1c. Summary of performance traits and p-values according to *MSTN* PR3737 genotype for horses over 13 years of age (n=64).

PR3737	**TT (n=53)**		**TC (n=10)**		**CC (n=1)**		**P-value**			
Performance trait	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**TT/TC**	**TC/CC**	**TT/CC**	**Overall***
P1 (250m) time record[1]	23.61	23.85	24.01	24.67	24.69	24.77	0.38	0.66	0.96	0.68
P2 (100m) time record[2]	8.36	8.48	8.54	8.66	8.49	8.50	0.77	0.67	0.73	0.90
P3 (150m) time record[3]	15.67	15.84	15.70	16.14	17.03	17.03	0.73	0.52	0.59	0.81

1: Based on information from 43 horses.
2: Based on information from 60 horses.
3: Based on information from 29 horses.

All analyses were conducted both including and excluding the one CC individual to ensure the trustworthiness of the results. *Overall p-values were calculated through anova analysis using mixed-effects models.

Table IV.2a. Summary of performance traits and p-values according to *MSTN* PR8604 genotype for total career of horses (n=131).

TOTAL CAREER									
PR8604	TT (n=90)		TG (n=34)		GG (n=7)		P-value	P-value	P-value
Performance trait	Median	Mean	Median	Mean	Median	Mean	TT/TG	TG/GG	TT/GG
Breeding field test points[1]	9	9	9	9	9	9	0.14	0.18	0.54
P1 (250m) time record[2]	23.13	23.53	23.46	23.58	22.85	23.40	0.88	0.78	0.81
P2 (100m) time record[3]	8.08	8.19	8.14	8.38	8.29	8.07	0.38	0.52	0.83
P3 (150m) time record[4]	16.40	16.25	15.31	16.06	15.09	15.37	0.72	0.21	0.11
Age of training start for pace[5]	6.00	7.00	7.00	7.00	7.50	7.00	0.51	0.90	0.68
EBV_overall	107.5	106.9	108.5	108.7	108.0	107.0	0.13	0.83	0.31
EBV_pace	113.0	111.6	111.0	109.9	113.0	112.3	0.68	0.31	0.38

1: Based on information from72 horses.
2: Based on information from 75 horses.
3: Based on information from 130 horses.
4: Based on information from 66 horses.
5: Based on information from 110 horses.

Table IV.2b. Summary of performance traits and p-values according to *MSTN* PR8604 genotype for horses under and including 13 years of age (n=117).

<=13 YEARS OLD										
PR8604	TT (n=82)		TG (n=29)		GG (n=6)		P-value			
Performance trait	Median	Mean	Median	Mean	Median	Mean	TT/TG	TG/GG	TT/GG	Overall*
P1 (250m) time record[1]	23.88	24.05	23.69	23.89	25.43	25.60	0.08	**0.03**	0.09	**0.04**
P2 (100m) time record[2]	8.26	8.47	8.46	8.77	8.40	8.44	0.92	0.63	0.64	0.89
P3 (150m) time record[3]	16.38	16.39	15.48	15.78	15.35	15.44	0.67	0.17	0.13	0.31

1: Based on information from 62 horses.
2: Based on information from 115 horses.
3: Based on information from 54 horses.

Table IV.2c. Summary of performance traits and p-values according to *MSTN* PR8604 genotype for horses over 13 years of age. (n=64).

>13 YEARS OLD										
PR8604	TT (n=44)		TG (n=16)		GG (n=4)		P-value			
Performance trait	Median	Mean	Median	Mean	Median	Mean	TT/TG	TG/GG	TT/GG	Overall*
P1 (250m) time record[1]	24.00	24.11	23.21	23.45	23.09	23.42	0.31	0.95	0.47	0.47
P2 (100m) time record[2]	8.47	8.55	8.39	8.53	7.92	8.06	0.83	0.55	0.58	0.83
P3 (150m) time record[3]	16.07	16.18	15.76	15.86	15.07	15.20	0.77	0.20	0.10	0.24

Based on information from 43 horses.
2: Based on information from 60 horses.
3: Based on information from 29 horses.

*Overall p-values were calculated through anova analysis using mixed-effects models.

Table IV.3a. Summary of performance traits and p-values according to *MSTN* PR9482 genotype for total career of horses (n=126).

TOTAL CAREER									
PR9482	CC (n=57)		CT (n=51)		TT (n=18)		P-value		
Performance trait	Median	Mean	Median	Mean	Median	Mean	CC/CT	CT/TT	CC/TT
Breeding field test points[1]	9.0	9.0	9	9	9	9	0.23	0.52	0.89
P1 (250m) time record[2]	23.11	23.52	23.36	23.31	23.61	24.32	0.67	0.07	0.12
P2 (100m) time record[3]	8.18	8.20	8.07	8.27	8.33	8.24	0.77	0.91	0.74
P3 (150m) time record[4]	16.46	16.34	15.32	15.71	15.13	16.60	0.10	0.42	0.74
Age of training start for pace[5]	7.0	7.5	6.00	7.42	6.00	6.50	0.88	0.22	0.26
EBV_overall	106.0	106.6	109.0	108.3	108.0	106.3	0.29	0.24	0.68
EBV_pace	112.0	111.4	112.0	111.3	111.0	109.2	0.82	0.40	0.49

1: Based on information from 52 horses.
2: Based on information from 56 horses.
3: Based on information from 110 horses.
4: Based on information from 48 horses.
5: Based on information from 90 horses.

Table IV.3b. Summary of performance traits and p-values according to *MSTN* PR9482 genotype for horses under and including 13 years of age (n=113).

<=13 YEARS OLD										
PR9482	CC (n=52)		CT (n=45)		TT (n=16)		P-value			
Performance trait	Median	Mean	Median	Mean	Median	Mean	CC/CT	CT/TT	CC/TT	Overall*
P1 (250m) time record[1]	23.75	23.97	23.68	23.80	25.03	25.08	0.18	**0.01**	0.11	**0.03**
P2 (100m) time record[2]	8.30	8.53	8.32	8.53	8.40	8.54	0.79	0.99	0.85	0.96
P3 (150m) time record[3]	16.40	16.41	15.51	15.82	15.44	15.85	0.48	**<0.01**	**<0.01**	**<0.01**

1: Based on information from 58 horses.
2: Based on information from 111 horses.
3: Based on information from 51 horses.

Table IV.3c. Summary of performance traits and p-values according to *MSTN* PR9482 genotype for horses over 13 years of age. (n=61)

>13 YEARS OLD										
PR9482	CC (n=30)		CT (n=24)		TT (n=7)		P-value			
Performance trait	Median	Mean	Median	Mean	Median	Mean	CC/CT	CT/TT	CC/TT	Overall*
P1 (250m) time record[1]	23.71	23.98	23.41	23.79	23.29	23.89	0.65	0.4	0.82	0.57
P2 (100m) time record[2]	8.51	8.53	8.47	8.57	7.98	8.14	0.61	0.25	0.39	0.52
P3 (150m) time record[3]	16.56	16.57	15.67	15.75	15.08	15.23	**0.04**	0.10	0.05	**<0.01**

1: Based on information from 41 horses.
2: Based on information from 57 horses.
3: Based on information from 28 horses.

*Overall p-values were calculated through anova analysis using mixed-effects models.

Table IV.4a. Summary of performance traits and p-values according to *GRIN2B*.6762 genotype for total career of horses (n=131).

TOTAL CAREER									
GRIN2B g.41206762 T>C	TT (n=103)		TC (n=23)		CC (n=5)		P-value	P-value	P-value
Performance trait	Median	Mean	Median	Mean	Median	Mean	TT/TC	TC/CC	TT/CC
Breeding field test points[1]	9.0	9.0	8.5	9.0	9.0	9.0	0.65	0.32	0.40
P1 (250m) time record[2]	23.20	23.47	23.32	23.87	21.93	21.93	0.51	0.26	0.33
P2 (100m) time record[3]	8.13	8.20	8.03	8.13	8.91	9.20	0.37	**<0.01**	**<0.01**
P3 (150m) time record[4]	16.29	16.27	15.25	15.35	18.18	18.18	**0.02**	**0.02**	0.10
Age of training start for pace[5]	6.00	7.00	7.00	7.50	8.00	7.00	0.66	0.80	0.60
EBV_overall	108.0	107.3	109.0	107.2	111.0	108.0	0.99	0.43	0.39
EBV_pace	113.0	111.5	110.0	109.3	111.0	112.8	0.45	0.70	0.43

1: Based on information from72 horses.
2: Based on information from 75 horses.
3: Based on information from 130 horses.
4: Based on information from 66 horses.
5: Based on information from 110 horses.

Table IV.4b. Summary of performance traits and p-values according to *GRIN2B*.6762 genotype for horses under and including 13 years of age (n=117).

<=13 YEARS OLD										
GRIN2B g.41206762 T>C	TT (n=91)		TC (n=21)		CC (n=5)		P-value			
Performance trait	Median	Mean	Median	Mean	Median	Mean	TT/TC	TC/CC	TT/CC	Overall*
P1 (250m) time record[1]	23.80	24.02	24.12	24.23	22.71	22.91	0.61	0.26	0.33	0.51
P2 (100m) time record[2]	8.26	8.50	8.49	8.65	8.22	8.55	0.57	**0.02**	**0.02**	**0.04**
P3 (150m) time record[3]	15.73	16.12	15.31	15.59	18.42	18.43	0.07	0.03	0.11	**0.04**

1: Based on information from 62 horses.
2: Based on information from 115 horses.
3: Based on information from 54 horses.

Table IV.4c. Summary of performance traits and p-values according to *GRIN2B*.6762 genotype for horses over 13 years of age. (n=64)

>13 YEARS OLD										
GRIN2B g.41206762 T>C	TT (n=54)		TC (n=9)		CC (n=1)		P-value			
Performance trait	Median	Mean	Median	Mean	Median	Mean	TT/TC	TC/CC	TT/CC	Overall
P1 (250m) time record[1]	23.77	23.98	23.59	23.86	-	-	0.61	0.26	0.33	0.51
P2 (100m) time record[2]	8.41	8.50	8.36	8.51	-	-	0.91	-	-	0.91
P3 (150m) time record[3]	15.73	16.07	15.04	15.19	17.73	17.73	0.07	0.03	0.11	0.19

1: Based on information from 43 horses.
2: Based on information from 60 horses.
3: Based on information from 29 horses.

*Overall p-values were calculated through anova analysis using mixed-effects models.

Table IV.5a. Summary of performance traits and p-values according to *DOCK8*.6787 genotype for total career of horses (n=131).

DOCK8 g.22496787 T>C	CC (n=19)		CT (n=63)		TT (n=49)		P-value		
Performance trait	Median	Mean	Median	Mean	Median	Mean	CC/CT	CT/TT	CC/TT
Breeding field test points[1]	9.0	9.0	9.0	9.0	9.0	9.0	0.94	0.75	0.54
P1 (250m) time record[2]	24.12	24.20	23.27	23.69	22.89	23.14	0.48	0.60	0.31
P2 (100m) time record[3]	8.42	8.51	8.18	8.21	8.00	8.14	0.05	0.70	**0.04**
P3 (150m) time record[4]	16.61	16.46	15.75	16.03	15.48	16.11	0.92	0.98	0.91
Age of training start for pace[5]	6.00	7.00	7.00	7.50	6.00	7.00	0.33	0.47	0.69
EBV_overall	103.0	105.6	108.0	107.4	109.0	107.9	0.27	0.77	0.21
EBV_pace	108.0	109.9	114.0	111.8	113.0	111.0	0.18	0.73	0.31

1: Based on information from72 horses.
2: Based on information from 75 horses.
3: Based on information from 130 horses.
4: Based on information from 66 horses.
5: Based on information from 110 horses.

Table IV.5b. Summary of performance traits and p-values according to *DOCK8*.6787 genotype for horses under and including 13 years of age (n=117).

<=13 YEARS OLD										
DOCK8 g.22496787 T>C	CC (n=18)		TC (n=55)		TT (n=44)		P-value			
Performance trait	Median	Mean	Median	Mean	Median	Mean	TT/TC	TC/CC	TT/CC	Overall
P1 (250m) time record[1]	25.19	25.15	23.90	24.20	23.58	23.76	0.07	0.28	0.07	0.13
P2 (100m) time record[2]	8.66	8.81	8.33	8.59	8.20	8.37	0.05	**0.04**	0.05	0.11
P3 (150m) time record[3]	16.24	16.33	15.68	16.01	15.62	15.90	0.81	0.69	0.81	0.75

1: Based on information from 62 horses.
2: Based on information from 115 horses.
3: Based on information from 54 horses.

Table IV.5c. Summary of performance traits and p-values according to *DOCK8*.6787 geotype for horses over 13 years of age (n=64).

>13 YEARS OLD										
DOCK8 g.22496787 T>C	CC (n=12)		CT (n=30)		TT (n=22)		P-value			
Performance trait	Median	Mean	Median	Mean	Median	Mean	TT/CT	CC/CT	TT/CC	Overall
P1 (250m) time record[1]	24.51	24.93	23.28	23.69	23.44	23.79	0.65	0.3	0.65	0.54
P2 (100m) time record[2]	8.56	8.61	8.24	8.40	8.46	8.56	0.68	0.41	0.68	0.69
P3 (150m) time record[3]	15.52	15.93	15.73	15.88	15.73	15.85	0.82	0.81	0.82	0.97

1: Based on information from 43 horses.
2: Based on information from 60 horses.
3: Based on information from 29 horses.

*Overall p-values were calculated through anova analysis using mixed-effects models