Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Veterinary Medicine and Animal Sciences

The role of Myostatin polymorphisms in the Finnhorse and Shetland pony breeds

Laura Bas Conn

Degree Project in Animal Science A2E / Swedish University of Agricultural Sciences Master's Thesis, 30 hp
Department of Animal Breeding and Genetics
Uppsala 2017
528

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Faculty of Veterinary Medicine and Animal Sciences

The role of Myostatin polymorphisms in the Finnhorse and Shetland pony breeds

Laura Bas Conn

Supervisors:

Gabriella Lindgren, SLU, Department of Animal Breeding and Genetics
Brandon Velie, SLU, Department of Animal Breeding and Genetics
Kim Jäderkvist Fegraeus, SLU, Department of Animal Breeding and Genetics

Examiner:
Sofia Mikko, SLU, Department of Animal Breeding and Genetics

Credits: 30 hp
Course title: Degree project in Animal Science
Course code: EX0803
Programme: Degree Project in Animal Science, offered as independent course
Level: Advanced, A2E
Place of publication: Uppsala
Year of publication: 2017
Cover Pictures: Jarno Unkuri for Hevosurheilu (above), Susann Melkersson from Melker stutteri (below).
Name of series: Degree project in Animal Science / Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics
Online publishing: http://epsilon.slu.se
Number of series: 528

Key words: horse, Myostatin, Finnhorse, Shetland pony, harness racing, performance, conformation

Index

1-Abstract

Myostatin, encoded by the *MSTN* gene, is a member of the transforming growth factor β that normally acts to limit skeletal muscle mass by regulating both the number and growth of muscle fibers. Natural mutations that decrease the amounts of myostatin and/or inhibit its function have been identified in several cattle, sheep and dog breeds, where loss-of-function mutations cause increased skeletal muscle mass and produce a phenotype known as "double-muscling". This gene has also been associated with racing performance in Thoroughbreds, where studies have found two Myostatin polymorphisms (PR3737 and PR8604) to be strongly associated with genetic potential and athletic phenotype, affecting both speed and muscularity, although no such associations have been found in harness-racing breeds such as the Swedish-Norwegian Coldblooded trotter. Single nucleotide polymorphisms have also been shown to have an effect on conformation, where these SNPs have been found to have different genotype distributions between horse breeds with different origin, morphology and uses. Such is the case with Icelandic horses, where significantly different genotype distributions for another SNP, PR5826, were observed between horses used for different purposes.

This study investigated the association of these *MSTN* polymorphisms with harness racing performance in the Finnhorse and body conformation in the Shetland pony. Finnhorses used for different disciplines were genotyped for three SNPs (PR5826, PR3737 and PR8604) and were divided into three categories based on their use: harness racing horses (n=223), representing those horses with at least one start in harness racing; riding horses (n=79) used for recreational riding of which owner questionnaire information was available and horses with no performance data (n=112). An association analysis was performed on the raced group, where the genotypes were evaluated for association with life-time racing performance results for: number of starts, victories, placings (1-3) and unplaced, along with proportions for each of these traits, as well as earnings, earnings per start and race times for volt- and autostart. Additionally, the genotypes were evaluated for association with these performance results obtained between 3 and 6 years of age (n=207) and 7 to 10 years of age (n=183). Significant associations were found in the 3 to 6 years of age group as well as the 7 to 10 group. In both cases, TT horses for PR3737 earned more money, were faster for both racing methods and presented a lower number of disqualifications than the CT horses. Concerning PR8604, a similar effect was observed, where the TT horses also earned more money, were faster in auto start racing method and presented a lower number of disqualifications. This study concluded that *MSTN* sequence polymorphisms seem to have an effect on harness racing performance in the Finnhorse. Concerning body conformation on the Shetland pony, a breed which presents a "Heavy" body type and a "Light" body type, no significant associations were found between conformation and *MSTN* genotype. However, further investigation is needed before drawing the conclusion that MSTN has no effect on Shetland pony conformation. This is due to the small sample size and classification method, which could be expanded to include morphological measurements.

2-Literature revision

2.1-Myostatin

Myostatin is a protein encoded by the *MSTN* gene and is a part of the transforming growth factor beta (TGF-β) superfamily. This superfamily encompasses a large group of growth and differentiation factors that play important roles in regulating embryonic development and in maintaining tissue homeostasis in adult animals. Myostatin is specifically expressed in skeletal muscle of several mammalian species, where it acts as a negative regulator of muscle growth. This is achieved by inhibiting myoblast proliferation through the interruption of normal cell cycles, mediated by the p21 family, a group of cyclin-dependent kinase inhibitors which act on all cyclin dependent kinases involved in G_1/S transition (Thomas et al., 2000).

The *MSTN* gene was first described by McPherron et al. in 1997 as a growth/differentiation factor expressed specifically in both developing and adult skeletal muscle. They identified the *MSTN* gene by degenerating oligonucleotides corresponding to conserved regions among the known family members of TGF- β on mouse genomic DNA and could explore its functions by disrupting it through gene targeting in mice. It was observed that the myostatin-null mice were 30% larger than the wild type mice and displayed abnormal body shapes. This was due to an increase in skeletal muscle mass caused by an increase in both thickness and number of muscle fibers. These observations were supported by histological analysis and body weight comparisons where it was observed that the increase in body weight resulted only from increase in muscle mass and not in fat content.

Since it's identification by McPherron et al., there have been numerous studies conducted in order to better understand the function of the *MSTN* gene in other mammalian species.

2.1.1-Myostatin in different species

A) Myostatin in livestock

An exceptional muscle development commonly referred to as 'double-muscling' has been seen in several cattle breeds and has attracted considerable attention from beef producers since it's description in 1807 by G. Culley. 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 mutations of the MSTN gene in heavy-muscled cattle of both breeds. Belgian Blue cattle are homozygous for an 11-bp deletion in the third exon which results in a frame-shift mutation that removes virtually all of the mature, active region of the molecule. On the other hand, the Piedmontese myostatin sequence contains a missense mutation in exon 3, resulting in a substitution of tyrosine for an invariant cysteine in the mature region of the protein. The locus has been termed ''partially recessive'' because there is some effect of a single copy of the allele where the animals present mild abnormalities in muscle phenotype, but generally the truly double muscled phenotype requires that the animal be homozygous (Kambadur et al., 1997).

Compared with normal cattle, Belgian Blue and Piedmontese animals have an increased proficiency to convert feed into lean muscle and produce a higher percentage of the most desirable cuts of meat (Casas et al., 1997). These animals have less bone, less fat, and 20%

more muscle on average. However, problems associated with the trait, such as reduction in stress tolerance, fertility, and calf viability in Belgian Blue have hindered exploitation of the hypertrophy by classical genetic selection (Arthur 1995).

A similar effect of *MSTN* mutations can be seen in sheep, specifically in Texel sheep, a breed renowned for their exceptional meatiness. The *MSTN* allele of this breed is characterized by a G to A transition in the 3' UTR that creates a target site for microRNAs that are highly expressed in skeletal muscle. This causes translational inhibition of the myostatin gene and hence contributes to the muscular hypertrophy of Texel sheep (Clop et al., 2006).

B) Myostatin in dogs

Double muscling has also been found in dogs due to a mutation in the *MSTN* gene, specifically in the Whippet breed, producing the "bully" whippet as termed by breeders. Mosher et al. (2007) discovered a 2bp deletion in the third exon of the *MSTN* gene which leads to a premature stop codon. All individuals homozygous for the deletion present the double muscle phenotype, while all dogs that sired a "bully" whippet were heterozygous for the deletion.

While the heterozygous individuals were less muscular than the homozygous carriers, they were still more muscular than the wild type. Additionally, an excess of the heterozygote genotype was observed among the fastest racers, which demonstrates that this state carries a performance-enhancing polymorphism that provides a competitive edge. Greyhounds, who share a common ancestral gene pool, were also tested for the mutation, seeing as their genetic proximity and the fact that they're both bred to excel at racing suggested that they might also carry it. However, none of the greyhounds tested presented the mutation.

C) Myostatin in humans

In 2004, Schuelke et al. reported the case of a german child born with muscle hypertrophy linked to a *MSTN* mutation. The child appeared extraordinary muscular at birth and muscular hypertrophy was verified by ultrasonography at 6 days of age, while presenting normal motor and mental development.

Because the double-muscled phenotype has been described in other species, Schuelke et al. sequenced all three exons and flanking intron sequences of the *MSTN* gene of both the child and the mother. Although no mutations were detected in the coding region, a G to A transition at nucleotide gIVS1+5 was present in both alleles of the patient and one allele of his mother. These results strongly indicated that the patient had a loss-of-function mutation in the myostatin gene, thus suggesting that the inactivation of myostatin has similar effects in humans, mice, and cattle.

It is also worth mentioning studies conducted in humans and mice on the role of *MSTN* in bone structure. Genetic studies in human populations have shown that myostatin gene polymorphisms are associated with variation in peak bone mineral density (Zhang et al., 2008) and transgenic overexpression of myostatin propeptide, which inhibits myostatin signalling in vivo and increases bone mineral density in mice (Mitchel et al., 2007). Thus, there is evidence from both human studies and animal models to suggest that myostatin is an important regulator of both muscle mass and bone density. The mechanisms by which myostatin regulates bone formation are not well understood, but it is clear that myostatin has direct effects on the proliferation and differentiation of mesenchymal stem cells (Guo et al., 2008; Artaza et al., 2005) and that myostatin and its receptor are expressed during bone regeneration (Guo et al., 2008; Nagamine et al., 1998). Myostatin therefore appears to be a potent antiosteogenic factor that may function to suppress the proliferation and possibly the survival of osteo- and chondroprogenitors. A study conducted by Elkasrawy and Hamrick in 2010 on myostatin-null mice revealed that these individuals had increased bone mineral density and biomineralization in many aspects of the skeleton. They believed this could well be a direct effect of myostatin deficiency on bone rather than mechanical responses or adaptations to increased muscle mass. In this study it was even observed that mice lacking myostatin presented distinctive features and morphologies of the individual cranial bones and cranial structures (Elkasrawy and Hamrick, 2010)

2.1.2-Myostatin in horses

After the first publication associating polymorphisms in the *MSTN* gene with racing performance in whippets, several studies continued to look for similar associations in other animals used for racing. The English Thoroughbred is the most common used breed for horse racing (The Swedish Horse Racing Authority, 2017) and has thus become the focus of many studies involving the effect of *MSTN* gene. In horses, MSTN is located at the distal end of chromosome 18 (ECA18) and consists 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, while no exonic sequence variants were detected, six Single Nucleotide Polymorphisms (SNP) were detected in intron 1 of *MSTN*, out of which two were included in their association analysis. Considering the relative 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). Among the two distance groups, a highly significant ($P = 3.70 \times 10^{-5}$) association was found with g.66493737C>T, where the C allele was twice as frequent in the short distance than in the long-distance cohort. Additionally, best race distance (BRD) was found to be significantly associated ($P=4.85 \times 10^{-8}$) with the PR3737 SNP. It was observed that as the distance of the races increased, the frequency of the C/C genotype decreased, further supporting the association of the C/C genotype with sprinting power. On the other hand, the T/T genotype was associated with stamina, seeing an increase in the genotype frequency over longer racing distances, while the C/T 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 nuclear gene variants that may contribute to equine performance-related phenotypes. In this cohort-based association test, they evaluated the genotypic variation at 40,977 SNPs between horses suited to short distance and middle-long distance races. The study concluded that PR3737 is the most powerful genome-wide predictor of optimum racing distance in Thoroughbred horses (Hill et al., 2010b).

A later study performed by Tozaki et al. (2010) similar to the work of Hill et al., conducted a genome-wide association study for racing performance in Thoroughbreds that led to the identification of three SNPs on ECA18, namely, g.65809482T>C, g.65868604G>T and g.66539967A>G. The haplotype of these SNPs, together with SNP g.66493737C>T, were found to be associated with lifetime earnings and performance ranks, corroborating the results presented by Hill et al. (2010a). After this first study, Tozaki et al. (2011) performed another study where it was found that one of the analysed SNPs (g.65868604G>T [PR8604], located in the promotor region of ECA18, was also associated with best race distance, performance rank and lifetime earnings. The G allele was found to be suited to short distance races, while the T allele was better suited for long-distance races. The GG genotype had the highest average of lifetime earnings compared to the TG and TT genotypes. For SNP PR3737, the CC genotype had highest average of lifetime earnings compared to CT and TT horses (Tozaki et al., 2011).

Additionally, Hill et al. (2011) conducted yet another study focused on associating the PR3737 genotypes with speed indices in Thoroughbred horses. The study demonstrated that genotypes at the *MSTN* g.66493737C>T locus have a significant influence in the determination of individual differences in speed. For all speed parameters that were significantly different among genotype cohorts, horses with the CC genotype outperformed TT horses, with measurements for CT individuals in most cases intermediate to the homozygotes (Hill et al., 2011).

B) Myostatin and morphology

Outside of racing performance, sequence polymorphisms in *MSTN* have been proven to have an effect in conformation as well. Hill et al. (2010a) also investigated whether the different genotypes influenced body mass, finding a significant relationship between the mass to height ratio and PR3737 genotype, where CC horses had greater mass per cm than TT horses.

A study conducted by Velie et al. (2015) on Icelandic horses analysed three SNPs (g.66495826A>G [PR5826], PR3737 and PR8604) that were chosen based on previous studies that associated these genetic variants with morphological type and body composition (Dall'Olio et al., 2010; Tozaki et al., 2011). A significant difference in genotype frequency between horses used for meat production and horses used for riding was present at PR5826, suggesting the likelihood that the different genotypes result in phenotypes that are either less desirable for riding or more desirable for meat production.

Additionally, François et al. conducted an additional study on Icelandic horses in 2016 that demonstrated the possible role of MSTN on both the form and function of horses from non-racing breeds. In the study, SNPs and haplotypes were analysed for association with estimated breeding values (EBV) for conformation traits and gaits. The results showed an association of PR3737 and PR8604 with the EBV for neck, withers and shoulders, along with the association of PR8604 with EBV for total conformation. These associations were all supported by the haplotype analysis. However, while SNP PR5826 showed a significant association with EBVs for leg stance and hooves, haplotype analyses for these traits failed to fully support these associations.

C) Genotype frequencies among horse breeds

Selective breeding for speed in Thoroughbred horses has led to a high frequency (0.51) of the C allele of SNP PR3737. However, analysis of other horse breeds shows an interesting variation in allele frequencies in this particular SNP (Bower et al., 2011).

Hill et. al (2010a) found that in Quarter horses, a breed known for short distance racing and activities requiring short bursts of speed, there was an excess of C alleles (0.90), while in National Hunt racehorses (used for races over obstacles and long distances) the C/C genotype was absent, supporting an association of the T allele with stamina. Also, the genotype frequencies Egyptian Arabian horses, known for endurance exercise, were considerably different to the Thoroughbred population with an excess of T/T (0.90) genotypes Together these findings indicate that the C/C genotype is particularly suited to fast, shorter distance racing and the T/T 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 wild-plains grazing equid populations traversing large distances within expansive ranges. To evaluate the effect of recent selection for different competitions, they compared genotypes from four recently derived breeds with recorded Thoroughbred influences: Standardbred and French trotter as representatives of harness racing horses, Irish Draught Horse representing show jumping and eventing horses and Quarter Horse as sprint racers. It was found that the C-allele occurs at a frequency of 0.51 in the Thoroughbred, is present at frequencies between 0.00 and 0.50 in Eurasian horses (British Isles, Middle East and North Africa and Asia), is absent in Standardbreds and is most common in the Quarter Horse (0.90). The C-allele is therefore not restricted to the Thoroughbred and Thoroughbred-derived populations, is not a new mutation, and seems to occur at variable frequencies depending on the selection pressures on the population (Bower et al., 2012).

A study conducted by Viluma in 2012 where the genotype frequencies and allele segregation of the different *MSTN* polymorphisms were studied in different horse breeds showed that the Shetland Pony, Gotlandsruss and Fjord Horse had the highest frequency of CC homozygotes for PR3737 (not considering Thoroughbreds) out of 25 different breeds analysed. These results were in accordance with those that Bower et al. (2012) presented, where the C allele was found in high frequency in the Shetland Pony. Due to the geographical proximity of this breed and the Galloway breed, which was the preeminent British racing population before the formal foundation Thoroughbred breed, it was assumed that the Shetland Pony might have contributed to the introduction of the C-allele in the Thoroughbred. Viluma also studied allele frequencies in Standardbreds, where it was observed that the T allele was practically fixed for PR8604 and completely fixed for PR3737, presenting no polymorphisms for this last SNP. PR5826 also presented an almost complete fixation of the A-allele.

Stefaniuk et al. (2016) conducted a study to examine the occurrence of *MSTN* variants among four polish breeds: Arabians, Polish Konik, Hucul and and Polish Heavy Draft. The analysis performed in their study showed that SNP PR5826 A>G is present in Hucul and Polish Heavy Draft, but absent in Arabians and Konik Polski. Concerning PR3737, their findings suggested that the C-allele is absent in the studied population of Arabian horses, but considering that it is present in Egyptian bloodlines, they concluded that further investigation was needed. In the other breeds they examined, the C-allele occurred at high frequencies: 0.302 in Konik Polski, 0.281 in Polish Heavy Draft and 0.038 in Hucul horses.

D) Insertion polymorphism of 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 at the myostatin locus may affect the aptitude for specific sport abilities, Santagostino et al. analysed the frequency of the two allelic variants (ERE1 containing allele or ERE1 allele lacking the insertion) in 30 horses competing in show-jumping at various levels, in 90 Italian Trotters bred for harness racing and in 75 horses registered in the Italian Thoroughbred studbook mainly bred for flat racing. Their results showed very 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. Among the flat racing horses, the percentage of ERE1+ alleles was 43. These observations 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 may 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. To test whether the ERE1 insertion influences performance relative to race distance, the elite horses were grouped according to Best Race Distance. It was observed that in short distance races, the majority of winning horses (18 out of 30) were homozygous for the ERE1+ allele and no homozygous individuals for the ERE1- allele were found; in the long distance races, only heterozygotes and ERE1- homozygotes were observed and, in medium distance races all the three genotypes were represented 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.

Considering that PR3737 is shown to be predictive of athletic performance, with different genotypes suited for different race distances, Santagostino et. al. compared the ERE1 genotypes with that of PR3737. 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), supporting the hypothesis that this is the genotype that drove selection. This suggests a direct influence of the ERE1 insertion on myostatin expression.

2.2 *DMRT3* (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)) is permissive for the ability to perform alternate gaits and has a favorable effect on harness racing performance (Andersson et al., 2012). A high frequency of the *DMRT3* mutation has been found in horses bred for harness racing, where the individuals have the ability to trot or pace at high speed without breaking into a gallop, which is an obvious advantage in these competitions where galloping can lead to disqualification (Andersson et al., 2012).

This mutation in *DMRT3* has shown to have a positive effect in the performance of Standardbreds and Nordic trotters (Andersson et al., 2012; Jäderkvist et al., 2014), where Standardbred horses homozygous for the mutant A allele were proven to be faster, to have a cleaner trot, to earn more money and to win more races. The effect was similar for the Swedish-Norwegian Coldblooded trotters, with the major difference that with increasing age CA horses performed well, or even better in some parameters, than AA horses. This was due to the fact that, where Standardbreds with genotype CA had difficulties keeping a clean, even-beat trot at higher speeds (Andersson et al., 2012), the CA Nordic trotters had a good trotting technique.

A recent study conducted by Jäderkvist et al. (2015) proved that different *DMRT3* genotypes are best adapted for harness racing or riding in Finnhorses. The association of the genotypes with the racing performance yielded different results to those previously described for Swedish-Norwegian Coldblooded trotters. In the Finnhorse, the AA genotype was superior compared to the CA and CC genotypes, presenting a significantly higher proportion of victories and placings, better race times and earned more money compared with C-horses. On the other hand, the horses with CA and CC genotypes seemed to be better adapted to classical riding disciplines, performing better than the AA horses. The AA horses received significantly lower gait scores compared with C-horses for the majority of gaits. Except for rhythm in extended canter, there were no significant differences between CA and CC horses.

2.3-The Finnhorse

The Finnhorse is Finland's national horse breed. Descending from Northern European domestic horses, it has its origins somewhere at the beginning of the 13th, but was not consolidated as a breed until 1907. Originally the horses were selected for haulage and pulling carriages along with agricultural work, requiring animals with good temperament and stamina. In the beginning of the 17th Century, a unified Inn Organization was formed that served as a travelling network, disposing of horses that could do up to four transportations a day (The Finnish Trotting and Breeding Association (Hippos), 2017). All of this suggests that the horses were selected and bred for their stamina from very early on.

Nowadays, the Finnhorse is bred to be either a trotter, a riding horse, a work horse or a pony-sized horse (Figure 1). The horse should be typical of the breeding line into which it is being introduced in the studbook; however, breeding between lines is allowed and the resulting progeny can be registered as either of its parent's type in the stud book. The aim of the studbook is to direct breeding towards well performing, multi-purpose horses abiding by the breed description that are easy to handle, good in movements, tenacious and healthy (Hippos, 2017).

Figure 1. Finnhorses of different types. (a) Trotter type. (b) Riding horse type. (c) Miniature type. (d) Draught horse type. Photo: Johanna Rautio/Wikimedia Commons.

In Scandinavia, harness racing is a more popular sport than gallop racing with Thoroughbreds. In these competitions, the horses will run at a trotting pace while pulling a two-wheeled cart called a sulky. Harness racing with Finnhorses has been held in Finland since the second half of the 19th century. The official annual Finnhorse racing championship, Kuninkuusravit, was first offered 1924 (Hippos, 2017). Nowadays, there are 600 competitions arranged yearly where around 8000 horses of all trotting breeds participate (Hippos, 2017).

2.3-The Shetland Pony

Small ponies have existed in the Shetland Isles for thousands of years, as evidenced by the findings of bones of small ponies that existed during the Bronze Age and it is thought that ponies have been in domestic use there since this time. It is believed that the Shetland Pony has its origin in the Cob type of Tundra Pony and the Mountain Pony type from Southern Europe, which migrated via the ice fields and land masses, with later introduction of a pony brought to the islands by the Celtic people which had evolved from crossing the same Mountain Pony type with the Oriental horse (The Shetland Pony Stud-Book Society, 2017).

Owing to its island existence the pony has evolved with relatively few importations and those that did arrive were by necessity, owing to the difficulties of transportation by sea. Two significant types established themselves within the breed: the heavier boned animal with a longer head and the lighter one, with high tail carriage and small, pretty head. The two types remain established to this day and retain their distinct characteristics. The heavier boned draught animal, with powerful chest and shoulders, is mainly used for leisure driving of carriages and showings, while the lighter, free moving pony for is most commonly used for riding and in both trotting harness-racing competitions and galop races as well (Figure 2). (The Shetland Pony Stud-Book Society, 2017; Sveriges Shetlandssällskap, 2017).

Figure 2. Shetland ponies destined to different uses (a) Showing stallion. Photo: Susann Melkersson, Melkersson Stuteri. (b) Harness racing gelding. Photo: Ranja Melina Bjurman.

The Shetland Pony Stud-Book Society was formed in 1890 with the aim of publishing a Stud-Book which would be the first for a native breed of pony in Britain. Many of the registered ponies today can trace their pedigrees back to the first volumes of the Stud-Book. Sweden's Shetland Society (SSS) was formed in 1967 and since then, a Swedish pedigree has been maintained from the imports procured from the UK and Netherlands. In Sweden, Shetland ponies are commonly used for showings, especially amongst breeders but also particular owners, where the heavy-type animal is usually presented in morphology contests. The lighter type is more commonly found owned by private persons as hobby riding horses for children, but also has a very high popularity as an efficient trotter in harness-racing competitions.

3-Introduction and aims of the study

While analysing genetic traits that can contribute to enhancing athletic performance in Thoroughbreds is undoubtedly important, polymorphisms in the *MSTN* gene have shown to vary among horse breeds, presenting different genotype frequencies for these SNPs depending on each breed's origin, morphology and uses (Dall'Olio et al., 2014; Bower et al., 2011; Stefaniuk et al., 2016). Therefore, it is interesting to analyse the effect these variations may have in different horse breeds in order to provide valuable information towards breeding goals and enhancement of athletic performance relevant to specific competitions.

In Nordic countries there are three distinct breeds of harness racing horses: the globally popular Standardbred trotter, the Swedish-Norwegian Coldblooded trotter and the Finnhorse. Swedish-Norwegian Coldblooded trotters and Finnhorses originate from northern European heavy horse breeds (Hendricks, 2007) and while the Swedish-Norwegian Coldblooded trotter is today bred mainly for its trotting abilities, the breeding goal of the Finnhorse is orientated towards obtaining a multi-purpose horse that can be used for working, riding or trotting, although the trotter type is still the most popular (Hippos, 2017). Previous studies have shown that there is very little variation in allele frequency of the same polymorphisms of *MSTN* in Standardbreds. Concerning the Swedish-Norwegian Coldblooded trotter trotter, a study was conducted to find associations between three known variations of the *MSTN* gene (PR5826, PR3737 and PR8604) and racing performance, but no association could be found (Petäjistö, 2016).

The primary aim of this study is to investigate whether three *MSTN* variants (PR5826, PR3737 and PR8604) are associated with racing performance in the Finnhorse. Despite the fact that these polymorphisms don't play a role in the Swedish-Norwegian Coldblooded trotter, the Finnhorse can still be considered a different breed, considering it is the only native horse breed in Finland and has its own stud-book separate from that of the Swedish-Norwegian Coldblooded trotter (Hippos, 2017). It has also been bred with different warmblood horses in order to produce a multi-purpose horse not just orientated towards trotting. The fact that the optimum genotypes of DMRT3 for harness racing differ between these two breeds is also an indication that the Swedish-Norwegian Coldblooded trotter and Finnhorse present differences despite having similar origins and thus present a promising subject of study for the possible role of *MSTN* polymorphisms.

The secondary aim of the study is to analyse whether genotypes of the three mentioned variants of *MSTN* are associated with body conformation in the Shetland Pony, a breed that presents two distinct morphological types: the heavy-type pony, used mainly as a showing animal and for leisure driving of carriages and the light-type pony, used mainly for riding and harness racing (The Shetland Stud-Book Society, 2017). In a similar way to the Icelandic breed, where the horses may present body conformations that are suitable to different purposes but no such classification is done within the breeding system, the Shetland pony presents two clearly differentiated types that are not registered as such in the Stud-Book. Because of this, it is interesting to investigate which have an effect on these morphological differences. *MSTN* polymorphisms present different genotype frequencies in different breeds that have drastically different morphologies, such as draft horses versus lighter riding horses and ponies (Dall'Olio et. Al, 2010; Viluma, 2012). Genotype frequencies may even vary within a same breed, such as the study Velie et. al (2015) highlighted. Because of this it is an interesting gene to study on the Shetland pony in an attempt to explain the morphological differences within the breed. This is not only due to the association of PR3737 allele C with muscularity, but it is also interesting to take into account the effect on bone mass and structure described both in myostatin-null mice (Mitchel et al., 2007; Elkasrawy and Hamrick, 2010) and humans (Zhang et al., 2008).

4-Material and methods

4.1-Finnhorse study

4.1.1-Population and samples

A sample of 395 Finnhorses born between 1981 and 2012, all originating from Finland, were used for this study. The sample represented offspring from 178 sires and 355 dams and consisted of 81 stallions, 224 mares and 90 geldings (note that sex classification is based on sex recorded from the latest information in the system up to December 31st, 2016). The samples came from isolated DNA from hair roots which were collected via personal contacts and online postings during the year 2015. Distribution of birth year over the sample can be found in Table 1.

The total sample was divided into three groups to carry out the statistical analysis: raced horses, consisting of those horses that have at least one start in harness racing; riding horses, which are used for recreational riding purposes and of which there was available information on performance from subjective questionnaires and, finally, horses with no available data in either field, such as horses used for working and other miscellaneous purposes. Additionally, a "non-raced" group was created with the horses used for riding and those with no data.

Table 1. Distribution of birth year over the total sample.

Year	Stallion	Mare	Gelding	Total
1981	0	1	0	1
1985	0	0	1	1
1986	1	0	0	1
1987	0	0	1	1
1988	0	1	0	1
1991	0	2	0	2
1992	1	2	1	4
1993	0	2	1	3
1994	2	5	0	7
1995	1	2	0	3
1996	1	6	3	10
1997	1	6	1	8
1998	1	3	1	5
1999	2	8	2	12
2000	3	12	4	20
2001	3	13	7	23
2002	3	12	7	22
2003	10	8	7	25
2004	7	21	6	34
2005	7	19	6	32
2006	12	23	9	44
2007	6	26	11	43
2008	11	17	7	35
2009	1	16	7	24
2010	3	9	4	16
2011	3	2	4	9
2012	2	8	0	10

4.1.2-SNP genotyping

For this study, three different SNPs from the *MSTN* gene were chosen based on results from previous studies: PR8604, PR3737 and PR5826 (Figure 3). 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 three SNPs using 96-well plates and custom mixtures of DNA samples and TaqMan reagents for every plate.

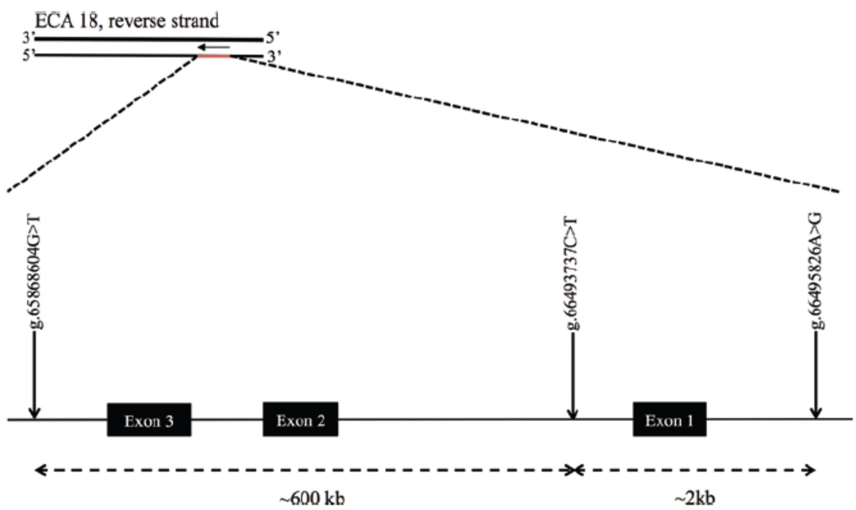

Figure 3. Relative locations of the analyzed SNP markers in horse *MSTN* (Velie et al., 2015).

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).

4.1.3-Harness racing horses

4.1.3.1-Samples

As previously mentioned, all horses that have at least one start in harness racing were extracted from the total sample as raced horses. The raced sample consisted of 223 horses born between 1981 and 2011 representing offspring from 86 sires and 204 dams. Out of the sample, 49 were stallions, 119 were mares and 55 were geldings. Distribution of birth year over the sample of raced horses can be found in Table 2.

Table 2. Distribution of birth year over the sample of raced Finnhorses.

Year	Stallion	Mare	Gelding	Total
1981	0	1	0	1
1985	0	0	1	1
1993	0	0	1	1
1994	0	2	0	2
1995	1	0	0	1
1996	1	3	0	4
1997	1	2	1	4
1998	1	1	0	2
1999	1	7	2	10
2000	2	8	1	11
2001	3	8	4	15
2002	1	5	4	10
2003	6	4	6	16
2004	3	14	2	19
2005	6	15	5	26
2006	9	17	8	34
2007	4	16	10	30
2008	7	7	3	17
2009	1	7	4	12
2010	2	2	2	6
2011	0	0	1	1

4.1.3.2-Performance data

Data was obtained from Hippos to carry out the statistical analysis based on lifetime performance results up to December 31, 2016. The traits included in the analysis were number of starts, wins and wins frequency, placings and placings frequency, unplaced and unplaced frequency, earnings, earnings per start (EPS), record race times for volt- and auto start and finally number of disqualifications and disqualification frequency. The analysis was divided into four age groups: total career performance, performance at 3 years of age, performance between 3 to 6 years of age and performance between 7 to 10 years of age.

Number of starts

The number of starts was defined as the number of official races a horse had competed in as registered by Hippos (2017).

Finish position

The number of victories (Wins) was calculated as the number of times a horse had finished a race in first position. The number of placings was calculated as the number of times a horse finished a race in first, second or third position and the number of unplaced was calculated as the times the horse failed to finish in either position. Frequencies for all three traits were also calculated.

Earnings

Cumulative earnings were calculated as the total amount of prize money in Euros that the horses had earned during from their lifetime and up to December 2016. Earnings per start (EPS) were also calculated, defined as the amount of prize money each horse earned per race start.

Race times

Race times for two different start methods (volt and autostart) were included in the study. Volt-start is a starting method in which the horses start in pens that dispose 20 metres of volt space and trot in a circular pattern to then hit the starting line as a group. Auto-start is a starting method where a car is used to set the starting line. The car is placed 350 metres before the start line and gradually increases the speed until it hits the start line (Finnish Trotting and Breeding Association, 2017). Best race times per competition year are recorded for each horse, presenting their lifetime best time achieved in the summary of their total career. The race times were calculated as average time per kilometer and are presented in seconds.

Disqualifications

The disqualifications were calculated as the number of times a horse had been disqualified from a race per race start. In harness racing, horses may be disqualified if they transition to gaits other than trotting, namely pace or gallop, under a series of conditions (such as the number of times they break the trot-2 per race- or the distance that they cover while on a different gait, applying certain conditions depending on the point of the race that they are currently in-Finnish Trotting and Breeding Association, 2017).

Best linear unbiased predictor (BLUP)-Estimated Breeding Value (EBV)

BLUP values for each horse were provided by Hippos. The breeding values for trotting horses are estimated by animal model using BLUP –methodology, which allows for essential environmental factors affecting single race results to be taken into account in the estimation. The indices consist of three parts: pedigree, horse's race results and the results of offspring. The total index is composed of single indices that are calculated using the following coefficients: 40 % time difference, 40 % earnings, 10 % age at first start and 10 % best annual time (Hippos, 2017). For this study, the latest BLUP values were used, which represent their breeding value up to 2016 (since it is calculated yearly).

4.1.3.3-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. For the sample of raced Finnhorses, summary statistics were calculated for all performance traits previously mentioned stratified by sex, age group, genotype and age group according to genotype of each SNP, with the exception of those genotypes which were only represented by one individual (PR5826 GG, PR3737 CC, PR8604 GG).

4.1.3.4-SNP genotyping
All individuals in the sample of riding Finnhorses were genotyped as described above in point 4.1.2. Some individuals were not successfully genotyped for all three SNPs (n=5) but were nonetheless included in the study due to the small sample number available.

4.1.3.5-Statistycal analysis
Data was structured for analysis using custom scripts written in the statistical software R (R Development Core Team, 2017). All traits were tested for normality using the Shapiro-Wilks test. All traits that were non-normally distributed were log transformed using log (racing time-68.2) for both time record traits (Árnason 1994). For the remainder of non-normally distributed traits, log_{10} transformations were used.

4.1.3.6-Models
Linear models were built to determine association of the *MSTN* polymorphisms with racing performance. SNP PR5826 was excluded due to low polymorphism frequency of the SNP, with only 7 individuals representing the AG genotype and 1 for the GG genotype.

All models included the fixed effects of sex, age, *MSTN* genotypes for all three SNPs and *DMRT3* genotype, while the rest of performance traits were included as covariates depending on logical assumptions made for each of the dependent variables analyzed. Models were built for the total career of all horses, horses between the ages of 3 and 6 years and horses between the ages of 7 and 10 years. Analysis of only 3-year old horses was not conducted due to a small sample size (n=28). Only those traits with significant p-values were included in the final models. Fixed effects that were not significant in some models but were at least significant in half of the traits analyzed were included in all models.

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. Note that CC genotype of PR3737 and GG genotype of PR8604 were excluded from the analysis due to low representation in the sample (n=1 for both cases).

4.1.4-Riding horses

4.1.4.1-Samples
From the total sample available, those horses categorized as riding horses were extracted. The sample consisted of 72 individuals born between 1991 and 2011 representing offspring from 52 sires and 58 dams. The sample consisted of 11 stallions, 40 mares and 21 geldings. As previously mentioned, these samples were recollected during the year 2015 from particular horse owners in Finland who used the animals for hobby riding, show jumping and dressage. Distribution of birth year over the sample can be found in Table 3.

Out of this group, 11 horses in total had been previous harness racers with available performance data and are thus overlapped in both groups. It was suspected from their performance data that these horses were not efficient racers and were thus sold to use for other purposes. The overlapped individuals were kept in both groups in order to provide the maximum information available in consideration of the small sample size.

Table 3. Distribution of birth year over the sample of riding Finnhorses.

Year	Stallion	Mare	Gelding	Total
1991	0	2	0	2
1992	1	0	1	2
1995	0	1	0	1
1996	0	2	1	3
1997	0	2	0	2
1998	0	1	1	2
1999	0	1	0	1
2000	1	2	2	5
2001	0	4	1	5
2002	2	3	2	7
2003	2	1	0	3
2004	0	3	4	7
2005	0	2	2	4
2006	0	4	0	4
2007	2	4	1	7
2008	1	4	3	8
2009	0	3	0	3
2010	1	1	1	3
2011	1	0	2	3

4.1.4.2-SNP genotyping
All individuals in the sample of riding Finnhorses were genotyped as described above in point 4.1.2. Four individuals were not successfully genotyped for all SNPs but were nonetheless included in all analysis.

4.1.4.3-Performance data
A questionnaire was designed to provide information on each horse's ability to perform different gaits on a scale from 1 (poor) to 6 (perfect), which were filled out by the owners. The questionnaire provided information on the rhythm, balance and transitions of 5 gaits (walk, extended trot, collected trot, extended canter and collected canter), whether they could perform "tölt" and/or pace gaits and whether the horses had competed in show jumping and/or dressage.

4.1.4.4-Summary statistics
Summary statistics were calculated using the subjective estimation of their ability to perform different gaits (with values from 1 to 6) for the whole sample and stratified by genotype of the different SNPs.

4.1.4.5-Statistical analysis
No association analysis was performed for the sample of riding horses due to the small sample number available (n=73) and the visibly insignificant variation in their performance as observed in the questionnaires.

4.1.5-Horses with no performance data

4.1.5.1-Samples

Out of the total sample, horses with no available performance data were extracted, consisting of 111 individuals born between 1986 and 2012. They represented offspring from 83 sires and 98 dams and consisted of 21 stallions, 74 mares and 16 geldings. Distribution of birth year over the sample can be found in Table 4.

Table 4. Distribution of birth year over the sample of horses with no performance data.

Year	Stallion	Mare	Gelding	Total
1986	1	0	0	1
1987	0	0	1	1
1988	0	1	0	1
1992	0	2	0	2
1993	0	2	0	2
1994	2	3	0	5
1995	0	1	0	1
1996	0	2	2	4
1997	0	2	0	2
1998	0	1	0	1
1999	1	1	0	2
2000	0	3	1	4
2001	0	3	2	5
2002	0	5	1	6
2003	2	4	1	7
2004	4	5	0	9
2005	1	2	0	3
2006	3	3	1	7
2007	0	6	0	6
2008	3	6	1	10
2009	0	6	3	9
2010	0	6	1	7
2011	2	2	2	6
2012	2	8	0	10

4.1.5.2-SNP genotyping

All individuals in the sample were genotyped as described above in point 4.1.2. Five individuals were not successfully genotyped for all SNPs but were nonetheless included in all analysis.

4.1.5.3-Summary statistics

Genotype frequencies were calculated for this sample. Additionally, summary statistics were produced to present the different uses of the horses in this group and their use stratified by genotype.

4.1.6-Non-raced horses

Horses used for riding and horses with no performance data were grouped into a "non-raced" category (n= 183) in order to carry out comparisons with the raced sample of horses (excluding those horses that were overlapped in both groups). An analysis was conducted to determine differences in genotype frequencies, EBVs and inbreeding coefficients between the group of raced horses and non-raced horses.

4.2-Shetland pony study

4.2.1-Population and samples

A sample of 158 Shetland ponies born between 1993 and 2017 originating from Sweden (n=149) and Denmark (n=9) were used for this study. The sample represented offspring from 99 sires and 138 dams and consisted of 43 stallions, 101 mares and 14 geldings (note that sex classification is based on sex recorded from the latest information in the Sveriges Shetlandssällskap database for the Swedish ponies and Landbrug Fødevarer for the Danish ponies, 2017). The samples consisted of DNA extracted from hair roots that were collected via personal contacts and online postings from private owners and breeders. Distribution of birth year over the sample can be found in Table 17.

Table 17. Distribution of birth year over the sample of Shetland ponies.

Year	Stallion	Mare	Gelding	Total
1993	0	2	0	2
1994	0	2	1	3
1995	2	3	0	5
1996	0	2	1	3
1997	2	3	0	5
1998	0	1	0	1
1999	0	4	1	5
2000	0	3	0	3
2001	0	2	1	3
2002	3	6	0	9
2003	1	6	1	8
2004	1	0	1	2
2005	0	4	2	6
2006	0	2	2	4
2007	1	2	0	3
2008	3	8	2	13
2009	0	9	0	9
2010	2	7	1	10
2011	4	5	0	9
2012	4	3	1	8
2013	2	4	0	6
2014	3	9	0	12
2015	8	9	0	17
2016	6	4	0	10
2017	1	1	0	2

4.2.2-DNA extraction and SNP genotyping

Genomic DNA was extracted from hair samples, using the hair roots in a standard hair-preparation procedure where 7 μL proteinase K (20 mg/mL; Merck KgaA, Darmstadt, Germany) and 100 μL 5 % Chelex Resin (Bio-Rad Laboratories, Hercules, CA) were added to each sample. The mix was incubated for 1 hour in 56°C and the proteinase K was inactivated for 10 min in 95°C.

All individuals were genotyped as described in point 4.1.2. All individuals were successfully genotypes for all three SNPs. Hardy-Weinberg equilibrium was evaluated for all SNPs genotyped for the Shetland pony sample using the “SNPassoc” package (González et al., 2015) for the software program for statistical computing R (R Development Core Team, 2017).

4.2.3-Pony classification

A questionnaire was filled out for each pony by their respective owners where it was asked to provide information on the use they gave each individual, namely: Showing, Riding, Trotting or Other (such as hobby horses, companion animals, reproduction purposes, etc.). Based on this information, the horses were classified as Heavy Type if they were used for Showing and as Light Type if used for either riding or trotting. All ponies owned by private owners used as Other (company, reproduction, etc.) were classified based on the owner’s perception of their body type along with previous or future uses of the ponies (for example, a young pony that will start competing in harness racing when it reaches maturity).

4.2.4-Summary statistics

Genotype distributions were calculated for the Shetland pony sample stratified by sex and by pony type (Heavy or Light), excluding ponies classified as Other (n=3).

4.2.5-Association analysis

Data was structured for analysis using custom scripts written in the statistical software R (R Development Core Team, 2017). Generalized linear models were built to determine association of *MSTN* polymorphisms with the two different body types. This was achieved by logistic regression where the pony type was defined as a binomial variable where “1” represented the Heavy type and “0” the Light type. All models included the fixed effect of sex and age. Ponies classified as Other (n=3) were excluded from this analysis.

5-Results

5.1-Finnhorse study

5.1.1-Total sample

Distribution of birth year over the sample is presented above in Table 1. Genotype distribution over the sample can be found in Table 5. Some individuals were not successfully genotyped for all SNPs (n=15). HWE results with genotype and allele frequencies can be seen in Table 6. All SNPs deviated from HWE (p>0.05).

Table 5. Genotype distribution for the total sample.

PR5826 (n=386)	AA	AG	GG
Stallion	0.99 (76)	0.01 (1)	0.00 (0)
Mare	0.94 (205)	0.06 (13)	0.004 (1)
Gelding	0.99 (89)	0.01 (1)	0.00 (0)
Total	0.96 (370)	0.04 (15)	0.002 (1)

PR3737 (n=392)	TT	CT	CC
Stallion	0.80 (65)	0.20 (16)	0.00 (0)
Mare	0.75 (167)	0.23 (51)	0.02 (4)
Gelding	0.78 (69)	0.20 (18)	0.02 (2)
Total	0.77 (301)	0.22 (85)	0.01 (6)

PR8604 (n=392)	TT	TG	GG
Stallion	0.78 (63)	0.22 (18)	0.00 (0)
Mare	0.78 (173)	0.22 (49)	0.00 (0)
Gelding	0.78 (69)	0.20 (18)	0.02 (2)
Total	0.78 (305)	0.22 (85)	0.01 (2)

Table 6. HWE equilibrium for all genotyped SNPs.

SNP	PR5826			PR3737			PR8604		
	AA	AG	GG	TT	CT	CC	TT	TG	GG
Genotype frequency	370 (95.86%)	15 (3.89%)	1 (0.26%)	301 (76.79%)	85 (21.68%)	6 (1.53%)	305 (77.81%)	85 (21.68%)	2 (0.51%)
Allele Frequency	A= 755 (97.80%)	G= 17 (2.20%)	NAs= 18	T=687 (87.63%)	C= 97 (12.37%)	NAs= 6	T= 695 (88.65%)	G= 89 (11.35%)	NAs= 6
HWE p value	p = 0.17			p = 1			p = 0.20		

5.1.2-Harness racing horses

5.1.2.1-Distributions and summary statistics

Distribution of birth year over the sample of raced horses is presented above in Table 2. Summary statistics of performance traits for the sample of raced horses (n=223) can be found in Table 7. Summary statistics for the sample of raced horses stratified by age groups and *MSTN* genotypes can be found in Appendix I.

Table 7. Summary statistics of performance traits for the sample of raced horses.

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Number of starts	1	22	43	59	79	329	53
Victories	0	1	4	6	9	56	8
Victories (frequency)	0.00	0.03	0.08	0.10	0.14	0.63	0.11
Placings (1-3)	0	4	11	18	26	123	18
Placings (Frequency)	0.00	0.17	0.27	0.28	0.37	1.00	0.18
Unplaced	0	15	29	41	56	269	39
Unplaced (frequency)	0.00	0.63	0.73	0.72	0.83	1.00	0.18
Disqualifications	0	3	7	9	12	42	8
Disqualifications (Frequency)	0.00	0.07	0.14	0.18	0.25	1.00	0.16
Earnings (Euro)	0	1778	8020	32163	37710	516230	66040
Earnings per start	0	73	172	431	455	7309	815
Racing time auto (seconds)	78.6	83.8	86.8	87.6	90.2	105.3	4.9
Racing time volt (seconds)	80.3	85.6	88.8	90.2	92.5	126.6	6.7
EBV	68.0	107.0	114.0	112.9	120.0	148.0	11.8

5.1.2.2-Genotyping

Genotype distribution for the sample of raced horses can be found in Table 8.

Table 8. Genotype distribution for the sample of raced horses.

PR5826 (n=221)	AA	AG	GG
Stallion	1 (48)	0.00 (0)	0.00 (0)
Mare	0.93 (110)	0.06 (7)	0.01 (1)
Gelding	1 (55)	0.00 (0)	0.00 (0)
Total	0.96 (213)	0.03 (7)	0.005 (1)

PR3737 (n=222)	TT	CT	CC
Stallion	0.83 (40)	0.18 (9)	0.00 (0)
Mare	0.77 (91)	0.22 (26)	0.01 (1)
Gelding	0.82 (45)	0.18 (10)	0.00 (0)
Total	0.79 (176)	0.20 (45)	0.005 (1)

PR8604 (n=221)	TT	TG	GG
Stallion	0.81 (39)	0.21 (10)	0.00 (0)
Mare	0.79 (93)	0.21 (25)	0.00 (0)
Gelding	0.78 (42)	0.20 (11)	0.02 (1)
Total	0.79 (174)	0.21 (46)	0.005 (1)

5.1.2.3-Association analysis

Fixed effects and covariates used to build the explanatory models can be found in Appendix 2. Summary of performance traits and p-values according to the different genotypes and grouped by age can be found in Tables 9 and 10**.**

Table 9. Summary of performance traits and p-values according to PR3737 genotype (excluding CC genotype due to low representation (n=1). Summary is presented for total career, horses between the ages of 3 and 6 years and 7 to 10 years.

	TOTAL CAREER					3 TO 6 YEAR OLDS					7 TO 10 YEAR OLDS				
PR3737	**TT (n=176)**		**CT (n=45)**		**P-value**	**TT (n=176)**		**CT (n=45)**		**P-value**	**TT (n=176)**		**CT (n=45)**		**P-value**
Performance trait	**Median**	**Mean**	**Median**	**Mean**	**TT/CT**	**Median**	**Mean**	**Median**	**Mean**	**TT/CT**	**Median**	**Mean**	**Median**	**Mean**	**TT/CT**
Number of starts	44	60	39	56	0.61	8	9	7	9	0.33	13	14	9	12	0.07
Victories	4	7	2	4	0.53	1	1	0	1	0.14	1	1	0	1	0.49
Victories (frequency)	0.09	0.11	0.06	0.07	0.07	0.07	0.15	0.00	0.10	0.07	0.05	0.09	0.00	0.05	0.35
Placings	12	19	9	14	0.78	3	3	2	3	0.99	3	4	1	2	0.54
Placings (frequency)	0.28	0.29	0.25	0.26	0.50	0.33	0.34	0.28	0.32	0.98	0.25	0.27	0.14	0.17	0.08
Unplaced	30	41	27	42	0.72	5	6	5	6	0.97	8	10	8	10	0.05
Unplaced (frequency)	0.72	0.71	0.75	0.74	0.53	0.22	0.66	0.72	0.68	0.78	0.75	0.73	0.86	0.83	0.09
Earnings (Euro)	9388	36199	4284	17797	0.30	1655	5115	790	2920	0.06	2560	8088	820	2922	**<0.01**
Earnings per start	187	485	126	240	0.14	196	496	117	295	**0.04**	183	506	82	164	**<0.01**
Racing time auto (seconds)	86.3	87.3	89.5	88.9	0.07	89.0	90.0	90.7	91.1	**0.05**	86.6	87.1	89.7	89.5	**<0.01**
Racing time volt (seconds)	88.3	89.8	90.7	91.4	0.12	91.9	93.1	93.9	94.7	**<0.01**	87.9	88.8	90.7	90.9	**<0.01**
Disqualifications	7	9	7	10	0.24	1	2	1	2	0.11	1	2	2	2	**0.03**
Disqualifications (frequency	0.13	0.17	0.17	0.22	0.10	0.13	0.20	0.17	0.24	0.13	0.10	0.15	0.18	0.25	**<0.01**
EBV	115	114	109	109	**0.01**	-	-	-	-	-	-	-	-	-	-

Table 10. Summary of performance traits and p-values according to PR8604 genotype and grouped by age (excluding GG genotype due to low representation (n=1). Summary is presented for total career, horses between the ages of 3 and 6 years and 7 to 10 years.

	TOTAL CAREER					3 TO 6 YEAR OLDS					7 TO 10 YEAR OLDS				
PR8604	**TT (n=174)**		**TG (n=46)**		**P-value**	**TT (n=174)**		**TG (n=46)**		**P-value**	**TT (n=174)**		**TG (n=46)**		**P-value**
Performance trait	**Median**	**Mean**	**Median**	**Mean**	**TT/TG**	**Median**	**Mean**	**Median**	**Mean**	**TT/TG**	**Median**	**Mean**	**Median**	**Mean**	**TT/TG**
Number of starts	44	61	37	51	0.87	8	9	9	9	**0.04**	12	14	10	11	0.67
Victories	4	6	3	6	0.53	1	1	1	1	**0.06**	1	1	0	1	0.15
Victories (frequency)	0.08	0.10	0.07	0.12	**0.08**	0.07	0.14	0.05	0.14	0.78	0.03	0.08	0.00	0.11	0.12
Placings	12	19	9	15	**0.06**	3	3	3	3	0.25	3	4	2	3	0.94
Placings (frequency)	0.28	0.29	0.21	0.27	0.78	0.33	0.35	0.22	0.28	**0.02**	0.23	0.25	0.20	0.26	0.84
Unplaced	31	43	27	36	0.80	5	6	6	6	0.35	9	10	7	8	0.46
Unplaced (frequency)	0.72	0.71	0.79	0.74	0.60	0.67	0.65	0.79	0.72	**0.04**	0.77	0.75	0.80	0.75	0.90
Earnings (Euro)	8395	32976	5184	30683	0.75	1413	4579	1325	5206	**0.02**	2550	7148	1210	7036	**0.04**
Earnings per start	181	415	127	509	0.39	193	441	126	533	0.10	178	425	112	507	**0.020**
Racing time auto (seconds)	86.6	87.3	88.6	88.8	**0.08**	89.2	90.0	89.1	90.7	0.71	86.5	87.1	88.5	88.8	**<0.01**
Racing time volt (seconds)	88.4	90.1	89.7	90.5	0.45	92.2	93.3	92.4	94.0	0.54	88.0	89.1	88.9	89.7	0.110
Disqualifications	7	9	8	9	0.24	1	2	2	2	0.19	1	2	2	2	**0.040**
Disqualifications (frequency	0.14	0.18	0.17	0.21	0.17	0.13	0.21	0.15	0.24	0.32	0.09	0.16	0.15	0.20	0.060
EBV	114	114	109	111	**0.10**	-	-	-	-	-	-	-	-	-	-

5.1.3-Riding horses

5.1.3.1-Distributions and summary statistics

Distribution of birth year over the sample of riding horses was presented in Table 3. Summary statistics of performance traits for the sample of riding horses (n=72) can be found on Tables 11 and 12. Summary statistics with the different stratifications mentioned above for the sample of riding horses can be found in Appendix III.

Table 11. Summary statistics of performance traits for the sample of riding horses. Evaluation based on scale from 1 (poor) to 6 (perfect),

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Years of riding	0.5	2.3	4.0	5.7	8.0	19.0
Collected Canter Rythm	1.0	2.8	4.0	3.7	5.0	6.0
Collected Canter Balance	1.0	3.0	4.0	3.7	5.0	6.0
Collected Canter Transitions	1.0	3.0	4.0	3.5	4.0	6.0
Extended Canter Rythm	1.0	3.0	4.0	3.7	5.0	6.0
Extended Canter Balance	1.0	2.8	3.0	3.5	5.0	6.0
Extended Canter Transitions	1.0	2.8	3.0	3.4	4.0	6.0
Difficulty of achieving current gallop form	1.0	2.0	3.0	3.2	4.0	6.0
Collected Trot Rythm	1.0	3.0	4.0	3.8	5.0	6.0
Collected Trot Balance	1.0	3.0	4.0	3.8	5.0	6.0
Collected Trot Transitions	1.0	3.0	4.0	3.7	5.0	6.0
Extended Trot Rythm	1.0	3.0	4.0	3.6	4.0	6.0
Extended Trot Balance	1.0	3.0	4.0	3.5	4.0	6.0
Extended Trot Transitions	1.0	2.3	3.5	3.4	4.0	6.0
Walk Rythm	2.0	4.0	5.0	4.4	5.0	6.0
Walk Balance	2.0	4.0	5.0	4.4	5.0	6.0
Walk Transitions	2.0	3.0	4.0	4.2	5.0	6.0
Jumping evaluation	1.0	3.0	4.0	4.0	5.0	6.0

Table 12. Summary statistics concerning ability to produce special gaits and competition information of riding horses.

Other factors	"Yes"	"No"	NA	Lätt A	Lätt B	Lätt C
Was the horse a trotter previously?	23	48	2			
Can the horse tölt?	9	62	2			
Can the horse pace?	17	54	2			
Has the horse competed in dressage?	44	26	3			
If so, what is the highest class it competed in?			32	14	14	11

**Lätt A, B and C are three competition levels in dressage within the easy/light category, being A the most advanced level.*

5.1.3.2-Genotyping

Genotype distribution for the sample of riding horses can be found in Table 13.

Table 13. Genotype distribution for the sample of riding horses.

PR5826 (n=68)	AA	AG	GG
Stallion	0.89 (8)	0.11 (1)	0.00 (0)
Mare	0.97 (37)	0.03 (1)	0.00 (0)
Gelding	0.95 (20)	0.05 (1)	0.00 (0)
Total	0.96 (65)	0.04 (3)	0.00 (0)

PR3737 (n=72)	TT	CT	CC
Stallion	0.82 (9)	0.18 (2)	0.00 (0)
Mare	0.70 (28)	0.30 (12)	0.00 (0)
Gelding	0.81 (17)	0.14 (3)	0.05 (1)
Total	0.75 (54)	0.24 (17)	0.01 (1)

PR8604 (n=72)	TT	TG	GG
Stallion	0.64 (7)	0.36 (4)	0.00 (0)
Mare	0.58 (23)	0.42 (17)	0.00 (0)
Gelding	0.76 (16)	0.19 (4)	0.05 (1)
Total	0.64 (46)	0.35 (25)	0.01 (1)

5.1.4-Horses with no performance data

5.1.4.1-Distributions and summary statistics

Distribution of birth year over the sample of the horses with no performance data is presented in Table 4. Summary statistics concerning the different horse types stratified by their genotype can be found in Appendix IV.

5.1.4.2-Genotyping

Genotype distribution for the sample of horses with no performance data can be found in Table 14.

Table 14. Genotype distribution for the sample of horses with no performance data.

PR5826 (n=108)	AA	AG	GG
Stallion	1 (20)	0.00 (0)	0.00 (0)
Mare	0.93 (67)	0.097 (5)	0.00 (0)
Gelding	1 (16)	0.00 (0)	0.00 (0)
Total	0.95 (103)	0.05 (5)	0.00 (0)

PR3737 (n=109)	TT	CT	CC
Stallion	0.76 (16)	0.24 (5)	0.00 (0)
Mare	0.73 (53)	0.23 (17)	0.04 (3)
Gelding	0.60 (9)	0.33 (5)	0.07 (1)
Total	0.72 (78)	0.25 (27)	0.03 (4)

PR8604 (n=110)	TT	TG	GG
Stallion	0.81 (17)	0.19 (4)	0.00 (0)
Mare	0.86 (63)	0.14 (10)	0.00 (0)
Gelding	0.75 (12)	0.19 (3)	0.06 (1)
Total	0.84 (92)	0.15 (17	0.01 (1)

5.2.5-Non-raced horses

Genotype distribution and frequencies for all three SNPs for the sample of raced and non-raced horses can be found in Table 14. EBV and inbreeding coefficient means for both sample groups and are presented in Table 15.

Table 15. Genotype distribution and frequencies (in parenthesis) for the sample of raced and non-raced horses.

	Total (n)	MSTN.PR 3737	MSTN.PR8604	MSTN.PR5826
Raced	223	TT= 176 (0.79) CT= 45 (0.21) CC= 1 (0.004)	TT= 174 (0.78) TG= 46 (0.21) GG= 1 (0.004)	AA= 213 (0.96) AG= 7 (0.03) GG= 1 (0.004)
Unraced	172	TT= 125 (0.74) CT= 40 (0.24) CC= 5 (0.03)	TT= 131 (0.76) TG= 39 (0.22) GG= 1 (0.06)	AA= 157 (0.95) AG= 8 (0. 05) GG= 0
Total	395	P-value*= 0.105	P-value*= 0.857	P-value*= 0.588

*Fisher's Exact test was performed

Table 16. EBV and Inbreeding coefficients means and standard deviations for the sample of raced and non-raced horses.

	Raced	SD	Non-raced	SD
EBV	112.90	11.83	97.36*	13.20
Inbreeding coefficient	4.15	1.40	3.35	1.39

**Only 11 horses of the non-raced group had EBVs*

5.2-Shetland pony study

5.2.1-Distributions and summary statistics

Distribution of birth year over the sample is presented above in Table 17. HWE results with genotype and allele frequencies can be seen in Table 18. All SNPs deviated from HWE ($p>0.05$). Genotype distribution of all genotyped SNPs over the total Shetland sample can be found in Table 19.

Table 18. HWE for all genotyped SNPs.

SNP	PR5826			PR3737			PR8604		
	AA	AG	GG	TT	CT	CC	TT	TG	GG
Genotype frequency	122 (77.22%)	36 (22.79%)	0 (0%)	59 (37.34%)	77 (48.73%)	22 (13.92%)	103 (65.20%)	48 (30.38%)	7 (4.43%)
Allele Frequency	A= 280 (88.61%)	G= 36 (11.39%)		T=195 (61.71%)	C= 121 (38.29%)		T=254(80.38%)	G= 62 (19.62%)	
HWE p value	*p* = 0.22			*p* =0.74			*p* = 0.62		

Table 19. Genotype distribution for the Shetland sample

PR5826 (n=158)	AA	AG	GG
Stallion	0.77 (33)	0.23 (10)	0.00 (0)
Mare	0.76 (77)	0.24 (24)	0.00 (0)
Gelding	0.86 (12)	0.14 (2)	0.00 (0)
Total	0.77 (122)	0.23 (36)	0.00 (0)

PR3737 (n=158)	TT	CT	CC
Stallion	0.53 (23)	0.40 (17)	0.07 (3)
Mare	0.33 (33)	0.52 (52)	0.16 (16)
Gelding	0.21 (3)	0.57 (8)	0.21 (3)
Total	0.37 (59)	0.49 (77)	0.14 (22)

PR8604 (n=158)	TT	TG	GG
Stallion	0.67 (29)	0.30 (13)	0.03 (1)
Mare	0.65 (66)	0.30 (30)	0.05 (5)
Gelding	0.57 (8)	0.36 (5)	0.07 (1)
Total	0.66 (103)	0.30 (48)	0.04 (7)

5.2.2-Association analysis

Genotype distributions were calculated based on each pony type for all SNPs to observe possible differences between the groups. Additionally, Fisher's exact test was performed for all SNPs to determine statistically significant differences between the two types of ponies and their genotype. Three ponies owned by two different breeders, who used most of their ponies for showing, were classified as "Other" (for example, a stallion used only for reproductive purposes). These ponies were excluded from all analyses.

Table 20. Genotype distribution for the Heavy type and Light type ponies of the sample.

		PR5826				PR3737				PR8604			
Pony Type	*n*	AA	AG	GG	P-value*	TT	CT	CC	P-value*	TT	TG	GG	P-value*
Heavy type	105	79	26	0		45	47	13		67	32	6	
Light type	50	41	9	0		12	29	9		35	14	1	
Total	155	120	35	0	0.41	57	76	22	0.07	102	46	7	0.59

*Fisher's exact Test was performed.

Despite finding no significant difference in genotype distributions between Heavy and Light type ponies for any of the SNPs, three generalized linear models were built including only Sex, Age and each SNP genotype, respectively, as covariates. No significant associations were found in this analysis.

6-Discussion

Many studies have been conducted in the past on different horse breeds on MSTN sequence polymorphisms that affect performance and even morphology. Horses present striking variations in genotype frequencies both within and across breeds that make them suitable for different purposes. Different genotypes of a particular SNP, PR3737, have even been described to be powerful performance predictors suited to different competition methods. When studying the allele frequencies of *MSTN* SNPs in the Finnhorse, selection towards the T allele can be observed, where almost 80% of the sample present the TT genotype for both PR3737 and PR8604. When comparing these results to previous studies, where a similar selection is seen in both trotting breeds, draft horses and breeds used for endurance competitions (such as the Arabian horse), it is not surprising to find the same direction of selection in the Finnhorse, a breed that originates from heavy horses that has been bred over centuries for its stamina as transport animals (Hippos, 2017). It is interesting to compare the genotype frequencies of Finnhorses and Swedish-Norwegian Coldblooded trotters, where, in the latter, the TT genotype presented itself in almost 95% of the genotyped sample for PR3737 and 89 % for PR8604 in the study conducted in 2016 by Petäjistö. The frequency of heterozygous genotypes of both SNPs (CT for PR3737 and TG for PR8604) are of 21% in the genotyped sample of Finnhorses, while only of 4% and 11%, respectively, in the Swedish-Norwegian Coldblooded trotters. This difference could be explained by the fact that Finnhorses have been bred with European and Nordic warmbloods with the objective of producing a multi-purpose horse, while the Swedish-Norwegian Coldblooded trotter is bred to use in harness-racing. Genotype frequencies in these breeds vary drastically from frequencies in the Thoroughbred, a breed selected for its speed and sprinting ability. In the Thoroughbred, the C allele of PR3737 presents itself in high frequencies (0.51), while the T/T genotype frequency is much lower compared to these endurance breeds (0.35 for horses classified as long-distance racers; Bower et al., 2012). Similarly, PR8604 genotype frequencies were calculated for Thoroughbreds in a study presented by Tozaki et al. (2012) where the T/T genotype still presented a much lower frequency (0.34) than in these endurance breeds, being the heterozygous T/G the predominant one at a frequency of 0.50.

The results of the analysis in this study show that there is a significant association between MSTN polymorphism genotypes and harness racing performance. An overall analysis of the total career of the sample did not yield significant associations between *MSTN* genotypes and racing performance, only in their EBV for PR3737 (P=0.013). However, it can still be observed that TT horses for both PR3737 and PR8604 on average had more starts, a greater number of victories and placings, earned more money and had better racing times than their CT and TG counterparts (see Tables 9 and 10). An analysis stratified into different age groups did yield significant associations between the genotypes and certain performance traits. When looking at horses that competed between the ages of 3 and 6 years, we see that TT horses for PR3737 earned more money overall and per start (P=0.038) and were faster for both racing methods (P= 0.046 for auto start and P= 0.003 for volt start) than the CT horses. Concerning the genotypes of PR8604, a significant association was found for the number of starts (P=0.036). While the mean number of starts in this age group is the same for both TT and TG horses, the median showed a lower number for the TT horses, suggesting a detrimental effect of this genotype over the number of starts. However, considering the proximity of the means and medians, it is doubtful whether this information is relevant; perhaps a larger sample number would be useful to determine the validity of this association. Concerning other performance traits, the TT genotype showed to have a positive effect on the placing frequency (P=0.024), the unplaced frequency (P=0.040) and the overall earnings (P=0.016) over the TG horses.

When analyzing the careers of horses between the ages of 7 and 10, an even greater effect of PR3737 could be observed, where once more the TT horses earned more money per year (P=0.0001) and per race start (P<0.0001), were faster for both race methods (P<0.001 for both autostart and voltstart) and additionally presented a lower number of disqualifications (P<0.001) and disqualification frequency (P<0.001) than the CT horses. Similarly, TT horses for PR8604 also earned more money per year (P=0.043) and per race start (P=0.016), were faster in auto start racing method (P=0.0006) and presented a lower number of disqualifications (P<0.001) than their TG counterparts. However, the advantage of PR8604 TT horses in this age group was not as accented as of those PR3737 TT horses, as observed in the means of their racing performance traits. It is interesting to observe that, while PR3737 is significantly associated with race times for both methods in this age group, it is not the case for PR8604, where an association was only found for autostart mode. Out of all 223 horses in the sample, 220 have a registered time record for volt start, while there are only 186 observations for auto start. The fact, then, that PR8604 is significantly associated with the race method with a lower number of observations might mean that the number of observations itself is too low and there are some individuals in the group that are producing a biased analysis (for example, the slowest time registered for this group is of 113,4 seconds for volt start and only 104 seconds for auto start, which is a great difference). Once again, a larger sample number might clarify whether PR8604 genotype truly only has an effect on one of the racing methods and not the other.

Further studies should be conducted to better understand the effect these genotypes may have on the different performance traits. For example, an analysis of best racing distance, as was performed in the Thoroughbreds, would be interesting to assess if the T and C alleles are having a similar effect over speed as they do in the Thoroughbred. It would also be interesting to analyze the haplotypes to assess if horses who have the TT genotype for both SNPs simultaneously are exceedingly better than those with a homozygous and heterozygous combination. It is also noteworthy to take into account the known effect of the *DMRT3* genotype on Finnhorses, as seen in the study conducted by Jäderkvist et al. in 2015. Analysis of p-values and coefficient estimates of the different DMRT3 genotypes in this study confer the same results obtained by Jäderkvist et al., where AA horses performed better than CA or CC horses. A study that investigates the possible relationship between *MSTN* genotype and *DMRT3* genotype might be necessary to better understand the influence of *MSTN* over racing performance in Finnhorses and whether it is powerful enough on its own to be considered a predictor of athletic phenotype as it is with Thoroughbreds. Additionally, it would be of importance to the horse racing industry in Nordic countries and to breeders to better understand how *MSTN* genotypes affect racing performance and in what degree. As of today, Swedish-Norwegian Coldblooded trotters and Finnhorses may race against each other, even if it is in very few races, since they are considered to be very similar breeds (The Swedish Trotting Association, 2017). However, seeing as *MSTN* has no known effects on the Swedish-Norwegian Coldblooded trotter and taking into account the difference in optimum DMRT3 genotypes, it would be of interest to further analyze the differences between these breeds. It might be so that they are truly not a single group of Swedish-Norwegian Coldblooded trotters and should compete separately, in the same manner as Standardbreds do not compete against Swedish-Norwegian Coldblooded trotters.

Concerning the Shetland pony sample, the genotype distributions suggested a direction of selection towards the A allele in PR5826 and the T allele for PR8604, although this selection is not as pronounced as in other breeds. As for PR3737, the allele frequencies are more equal between each other, having 68% of T alleles and 38% of C alleles in the genotyped sample. These results differ somewhat from those presented by Bower et al. (2012), where the allele frequencies were equally distributed (50% of each). These differences between studies could be due to different breeding goals between Shetland ponies in the United Kingdom and in Sweden, seeing as in Sweden the Shetland pony is a very popular breed for harness racing, which is not the case in most countries. Considering that the T allele might have a positive effect in harness racing, as demonstrated by the Finnhorse results, this might be a possible explanation for this direction of selection.

No significant differences in genotype distribution were found between Heavy and Light type ponies and MSTN genotypes. Additionally, the association analysis did not yield any significant associations with SNP genotypes and pony type. This is most likely due to the limitations of the study and the data available, as there was an unequal proportion of pony types. Additionally, the classification method could be improved, seeing as classification was purely subjective by the owner's own evaluation and there was no other, more objective, information available. It would be interesting to conduct a sounder study including, for example, morphological data such as height, mass and height/mass ratio, just as Hill et al. (2010a) analysed in their Thoroughbred population to associate the C allele with muscularity. A larger sample number with equal proportions of both types would also be necessary, along with an improved classification system that takes into account morphological measurements, sex, age and time of the year, seeing as the weight of the ponies usually varies between seasons. In any case, further investigation could lead to interesting results if one takes into account that the Shetland pony may have contributed to the introduction of the C allele in the Thoroughbred population. An association analysis on harness racing performance on Shetland ponies, as was done with the Finnhorses, could explain the different genotype frequencies seen in the Swedish population versus the British one, if ponies have been selected for endurance and slimmer body types versus faster, more muscular sprinters for this competition method.

7-Conclusions

This study analyzed the effect of three known MSTN polymorphisms on racing performance and body conformation in two different breeds. In the Finnhorse, a significant association was found between two known variants of *MSTN* and harness racing performance, where individuals homozygous for the T allele of SNPs PR3737 and PR8604 performed better than the heterozygous individuals in many of the analyzed traits. Further studies are required to better understand the exact nature of the effect these genotypes have on racing performance to determine whether *MSTN* can be considered a valid predictor of athletic phenotype in this breed.

In the Shetland pony, no significant associations were found between these polymorphisms and Shetland pony types. The limitations in the study caused by a subjective classification of the animals and a low sample number warrant a sounder study to be performed. This could include exact morphological measurements, for example, to create a more objective classification system in order to assess the validity of this significant difference between the types and to assess whether *MSTN* polymorphisms have an effect or not on their conformation.

This study portrays the importance of studying the effects a single gene may have in different breeds no matter how similar the breeds may be, and even within breeds to explain differences amongst individuals.

8- 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 work in her research group and carry out this project. Special thanks as well to Brandon Velie and Kim Jaderkvist Fegraeus, also supervisors, for all their teachings, patience and guidance during the project. To Sofia Mikko for being my examiner and to Maria Rosengren and Miguel Novoa for their input and support during the project. Thank you to all the horse owners and breeders who shared their interest in the project and contributed with samples-without you, the Shetland pony study would not have been possible. Also, many thanks to Team Rembacken for giving me the opportunity to begin my pursuit of a career in science and to Mathias Holmqvist for supporting me all this time in all my endeavours.

9-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. Results on the whole racing careers of harness-racing Finnhorses were presented in a conference proceeding submitted to the World Congress on Genetics Applied to Livestock Production, Aotea Centre Auckland, New Zealand, 11-16 February, 2018. Gabriella Lindgren, Brandon Velie and Kim Jaderkvist Fegraeus, scientists at the department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Sweden, conceived and designed the experiment for this study. Kim Jäderkvist Fegraeus contributed with Finnhorse DNA, while the Animal Genetics Laboratory, Swedish University of Agricultural Sciences, Uppsala, Sweden contributed with materials and analysis tools. Gabriella Lindgren, Brandon Velie and Kim Jaderkvist Fegraeus supervised the project and contributed with input on my thesis and guidance throughout the whole project.

10-References

Andersson, L. S, Larhammar, M., Memic, F., Wootz, H., Schwochow, D., Rubin, C-J., Patra, K., Arnason, T., Wellbring, L., Hjälm, G., Imsland, F., Petersen, J. L., McCue, M. E., Mickelson, J. R., 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. 1994. The importance of different traits in genetic improvement of trotters. Proceedings of World Congress on Genetics Applied to Livestock Production, August 7–12, 1994. Vol. 17. p. 462–469. Guelph (Canada): University of Guelph.

Artaza JN, Bhasin S, Magee TR, Reisz-Porszasz S, Shen R, Groome NP, et al. Myostatin inhibits myogenesis and promotes adipogenesis in C3H 10T(1/2) mesenchymal multipotent cells. Endocrinology. 2005 Aug; 146 (8): 3547–3557.

Arthur, P. F. 1995. Double muscling in cattle: A review. Aust. J. Agric. Res. 46:1493-1515.

Binns, M. M., Boehler, D. A. and Lambert, D. H. (2010) Identification of the myostatin locus (MSTN) as having a major effect on optimum racing distance in the thoroughbred horse in the USA. *Animal Genetics*, vol. 41 (Suppl. 2), pp. 154-158.

Bower, M. A., McGivney, B. A., Campana, M. G., Gu, J., Andersson, L. S., Barrett, E., Davis, C. R., Mikko, S., Stock, F., Voronkova, V., Bradley, D. G., Fahey, A. G., Lindgren, G., MavHugh, D. E., 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.

Casas, E., Keele, J. W., Shackelford, S. D., Koohmarate, M., Sonstegard, T. S., Smith, T. P. L., Kapes, S. M. and Stone, R. T. (1998) Association of the muscle hypertrophy locus with carcass traits in beef cattle, *J. Anim. Sci*, vol. 76 (2), pp. 468-473.

Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibé, B., Bouix, J., Caiment, F., Elsen, J-M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C. and Georges, M. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genetics*, vol. 38, pp. 813-818.

Culley, G. (1807) *Observations on Livestock: Containing Hints for Choosing and Improving the Best Breeds of the Most Useful Kinds of Domestic Animals*, 4th ed. London: G. Woodfall.

Dall'Olio, S., Fontanesi, L., Nanni Costa, L., Tassinari, M., Minieri, L. and Falaschini, A. (2010). Analysis of Horse Myostatin Gene and Identification of Single Nucleotide Polymorphisms in Breeds of Different Morphological Types. *Journal of Biomedicine and Biotechnology*, doi:10.1155/2010/542945.

Elkasrawy MN, Hamrick MW. 2010. Myostatin (GDF-8) as a key factor linking muscle mass and bone structure. *J Musculoskelet Neuronal Interact*. 10:56–63.

François, L., Jäderkvist Fegraeus, K., Eriksson, S., Andersson, L. S., Tesfayonas, G. Y., Viluma, A., Imsland, F., Buys, N., Mikko, S., Lindgren, G. and Velie, B. D. (2016) Conformation Traits and Gaits in the Icelandic Horse are Associated with Genetics Variants in Myostatin (MSTN). *Journal of Heredity*, vol. 00, pp. 1-7.

González, J. R., Armengol, L., Solé, X., Guíno, E., Mercader, J. M., 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 [2017-10-24]

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.

Guo W, Flanagan J, Jasuja R, Kirkland J, Jiang L, Bhasin S. The effects of myostatin on adipogenic differentiation of human bone marrow-derived mesenchymal stem cells are mediated through cross-communication between Smad3 and Wnt/beta-catenin signaling pathways. J Biol Chem. 2008 Apr 4; 283 (14) : 9136–9145.

Hendricks, B. L. (2007*) International Encyclopedia of Horse Breeds*, Norman: University of Oklahoma Press.

Hill, E. W., Fonseca, R. G., McGivney, B. A., Gu, J., MacHugh, D. E. and Katz, L. M. (2011) MSTN genotype (g.66493737C/T) association with speed indices in Thoroughbred racehorses. *J Appl Physiol*, doi: 10.1152/japplphysiol.00793.2011.

Hill, E. W., Gu. J., Eivers. S. S., Fonseca. R. G., McGivney. B. A., Govindarajan. P., Orr. N., Katz. L. M., MacHugh. D. (2010a) A Sequence Polymorphism in MSTN Predicts Sprinting Ability and Racing Stamina in Thoroughbred Horses. PLOS one, vol. 5(1), Available: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0008645 [2017-10-24]

Hill, E. W., McGivney, B. A., Gu, J., Whiston, R. and MacHugh, D. E. (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.

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

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

Jäderkvist, K., Johansson, L., Mäenpää, M., Mykkänen, A., Andersson, L. S., Velie, B. D., Andersson, L., Árnason, T. and Lindgren, G. (2015) Different DMRT3 Genotypes Are Best Adapted for Harness Racing and Riding in Finnhorses, *Journal of Heredity*, vol. 2015, pp. 734-740.

Jäderkvist, K., Holm, N., Imsland, F., Árnason, T., Andersson, L., Andersson, L.S. and Lindgren, G (2015). The importance of the *DMRT3* 'Gait keeper' mutation on riding traits and gaits in Standardbred and Icelandic horses. *Livestock Science*, vol. 176, pp 33-39.

Jäderkvist Fegraeus, K., Lawrence, C., Petäjistö, K., Johansson, M.K., Wiklund, M., Olsson, C., Andersson, L., Andersson, L.S., Røed, K. H., Ihler, C-F., Strand, E., Lindgren, G. and Velie, B. D. (2017). Lack of significant associations with early career performance suggest no link between the DMRT3 ´´Gait Keeper´´ mutation and precocity in Coldblooded trotters. *PLoS ONE*, vol. 12(5): e0177351.https://doi.org/10.1371/journal.pone.0177351

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

McPherron, A. C., Lawler, A. M., Lee, S-J. (1997) Regulation of skeletal muscle mass in mice by a new TGF-β superfamily member, *Letters to nature*, vol. 387, pp. 83-90.

Mitchell AD, Wall RJ. In vivo evaluation of changes in body composition of transgenic mice expressing the myostatin pro domain using dual energy X-ray absorptiometry.

Growth Dev Aging. 2007 Summer; 70(1): 25–37.

Mosher, D. S., Quignon, P., Bustamante, C. D., Sutter, N. B., Mellersh, C. S., Parker, H. G. and Ostrander, E. A. (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.

Nagamine T, Imamura T, Ishidou Y, Kato M, Murata F, ten Dijke P, et al. Immunohistochemical detection of activin A, follistatin, and activin receptors during fracture healing in the rat. J Orthop Res. 1998 May; 16(3): 314–321.

Petersen, J. L., Valberg, S. J., Mickelson, J. R. And McCue, M. E. (2014) Haplotype diversity in the equine myostatin gene with focus on variants associated with race distance propensity and muscle fiber type proportions. *Animal Genetics*, vol. 45, pp. 827-835

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/

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

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

Schuelke, M., Wagner, K. R., Stolz, L. E., Hübner, C., Riebel, T., Kömen, W., Braun, T., Tobin, J. F. and Lee,.S-J. (2004) Myostatin Mutation Associated with Cross Muscle Hypertrophy in a Child. *The New England Journal of Medicine*, vol. 350, pp. 2682-2688.

Stefaniuk, M., Ropka-Molik, K., Piórkowska, K., Kulisa, M., Podstawski, Z (2016). Analysis of polymorphisms in the equine MSTN gene in Polish populations of horse breeds. *Livestock Science* , vol.187, pp 151-157

The Finnish Trotting and Breeding Association [2017-10-24]. Available: http://www.hippos.fi/

The Finnish Trotting and Breeding Association [2017-10-24]. *Get to know the Finnhorse.* Available: http://www.hippos.fi/files/12598/SHjulkaisu_englanti_nettiin.pdf

The Finnish Trotting and Breeding Association [2017-10-24]. *Reglamente.* Available: http://www.hippos.fi/files/17672/reglemente_170101.pdf

The Shetland Pony Studbook Society [2017-10-24]. *Breed History*. Available: http://www.shetlandponystudbooksociety.co.uk/about-the-breed

The Swedish Horse Racing Authority [2017-10-24]. Hästen – Det engelska fullblodet. Available: https://www.svenskgalopp.se/artikel/hasten?defaultMenuId=true

The Swedish Trotting Association [2017-10-24]. Available: https://travsport.se/

Thiruvenkadan, A. K., Kandasamy, N., Panneerselvam, S. (2009) Inheritance of racing performance of trotter horses: An overview. Livestock Science, vol. 124, pp. 163-181.

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.

T. Tozaki, T. Miyake, H. Kakoi, H. Gawahara, S. Sugita, T. Hasegawa, N. Ishida, K. Hirota and Y. Nakano (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, E. W., 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.

Velie, B. D., Jäderkvist, K., Imsland, F., Viluma, V., Andersson, L. S., Mikko, S., Eriksson, S. and Lindgren, G. (2015) Frequencies of polymorphisms in myostatin vary in Icelandic horses according to the use of the horses, *Animal Genetics*, vol. 46 (4).

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

Zhang ZL, He JW, Qin YJ, Hu YQ, Li M, Zhang H, et al. Association between myostatin gene polymorphisms and peak BMD variation in Chinese nuclear families. *Osteoporos Int*. 2008 Jan;19(1):39–47.

APPENDIX I

I.1-Summary statistics of performance traits for the sample of raced Finnhorses stratified by sex and for their total career.

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Starts							
Stallions	4	34	68	82	124	232	54
Mares	1	19	36	50	63	329	52
Geldings	1	21	53	58	82	200	47
Earnings							
Stallions	0	18050	48395	80917	75230	516230	109268
Mares	0	1275	4475	17396	12168	362905	41193
Geldings	0	2049	7220	20678	25258	113890	28936
EPS							
Stallions	0	270	536	1033	1151	7309	1334
Mares	0	56	117	245	256	4886	508
Geldings	0	88	154	298	379	2423	403
Victories							
Stallions	0	7	12	13	17	38	9
Mares	0	1	2	4	5	56	6
Geldings	0	1	4	5	8	18	5
Victories (freq)							
Stallions	0.00	0.10	0.14	0.19	0.21	0.63	0.14
Mares	0.00	0.02	0.05	0.07	0.10	0.55	0.08
Geldings	0.00	0.04	0.08	0.09	0.13	0.56	0.10
Race time auto							
Stallions	78.6	81.7	83.7	84.0	85.3	94.4	3.0
Mares	81.5	85.5	88.9	89.1	92.2	104.0	4.7
Geldings	79.9	84.9	87.0	88.2	91.0	105.3	4.8
Race time volt							
Stallions	80.3	84.0	85.0	86.0	87.1	104.6	4.0
Mares	83.3	87.3	90.4	91.8	93.3	126.6	7.1
Geldings	83.8	86.0	89.2	90.6	92.7	117.1	6.5
Placings							
Stallions	0	19	27	32	45	95	19
Mares	0	3	8	13	16	123	16
Geldings	0	4	12	17	25	58	18
Placings (freq)							
Stallions	0.00	0.28	0.37	0.42	0.50	0.78	0.17
Mares	0.00	0.14	0.21	0.24	0.32	1.00	0.16
Geldings	0.00	0.17	0.29	0.28	0.35	0.87	0.18
Unplaced							
Stallions	4	19	37	50	78	160	40
Mares	0	14	26	37	48	269	39
Geldings	1	15	37	42	66	162	35
Unplaced (freq)							
Stallions	0.22	0.50	0.63	0.59	0.72	1.00	0.17
Mares	0.00	0.68	0.79	0.76	0.86	1.00	0.16
Geldings	0.13	0.65	0.71	0.73	0.84	1.00	0.18
Disqualifications							
Stallions	0	5	7	10	12	39	9
Mares	0	4	7	8	11	33	6
Geldings	0	2	7	10	15	42	10
Disqualifications (freq)							
Stallions	0.00	0.06	0.10	0.13	0.19	0.50	0.09
Mares	0.00	0.08	0.16	0.21	0.30	1.00	0.18
Geldings	0.00	0.07	0.15	0.17	0.21	0.51	0.15
EBV							
Stallions	82	115	122	119	126	148	13
Mares	68	106	113	111	119	131	11
Geldings	72	106	113	111	116	138	12

I.2-Summary statistics of performance traits of the sample of raced Finnhorses at 3 years of age.

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Starts							
Stallions	1	2	2	2	2	6	2
Mares	1	1	1	2	3	4	1
Geldings	1	2	3	2	3	3	1
Total	1	1	2	2	3	6	1
Earnings							
Stallions	0	1000	1200	2611	3400	8430	2791
Mares	0	0	120	513	575	2530	802
Geldings	0	240	300	804	330	3150	1318
Total	0	23	315	1239	1395	8430	1952
EPS							
Stallions	0	600	850	954	1200	2800	830
Mares	0	0	120	183	275	632	217
Geldings	0	100	110	300	240	1050	428
Total	0	23	183	452	608	2800	616
Victories							
Stallions	0	1	1	1	2	3	1
Mares	0	0	0	0	0	2	1
Geldings	0	0	0	0	0	2	1
Total	0	0	0	1	1	3	1
Victories (freq)							
Stallions	0.00	0.33	0.50	0.51	0.75	1.00	0.37
Mares	0.00	0.00	0.00	0.04	0.00	0.50	0.13
Geldings	0.00	0.00	0.00	0.13	0.00	0.67	0.30
Total	0.00	0.00	0.00	0.21	0.50	1.00	0.33
Race time auto							
Stallions	90.1	90.8	91.5	92.2	92.9	95.9	2.6
Mares	94.8	96.9	99.2	100.3	102.5	107.8	5.6
Geldings	-	-	-	-	-	-	-
Total	90.1	91.7	95.4	96.2	98.4	107.8	5.9
Race time volt							
Stallions	91.8	93.4	95.8	96.3	98.8	101.9	3.7
Mares	94.0	97.4	102.7	102.4	105.8	116.2	6.7
Geldings	93.7	101.2	102.2	101.3	104.0	105.3	4.5
Total	91.8	94.3	101.2	100.3	103.6	116.2	6.0
Placings							
Stallions	0	1	2	2	2	6	2
Mares	0	0	1	1	2	3	1
Geldings	0	0	1	1	1	3	1
Total	0	0	1	1	2	6	1
Placings (freq)							
Stallions	0.00	0.75	1.00	0.81	1.00	1.00	0.35
Mares	0.00	0.00	0.17	0.39	0.75	1.00	0.44
Geldings	0.00	0.00	0.33	0.33	0.33	1.00	0.41
Total	0.00	0.00	0.58	0.52	1.00	1.00	0.44
Unplaced							
Stallions	0	0	0	0	1	1	1
Mares	0	1	1	1	1	2	1
Geldings	0	1	2	1	2	2	1
Total	0	0	1	1	1	2	1
Unplaced (freq)							
Stallions	0.00	0.00	0.00	0.19	0.25	1.00	0.35
Mares	0.00	0.25	0.83	0.61	1.00	1.00	0.44
Geldings	0.00	0.67	0.67	0.67	1.00	1.00	0.41
Total	0.00	0.00	0.42	0.49	1.00	1.00	0.44
Disqualifications							
Stallions	0	0	0	0	0	1	0
Mares	0	0	0	0	1	3	1
Geldings	0	0	0	0	1	1	1
Total	0	0	0	0	1	3	1
Disqualifications (freq)							
Stallions	0.00	0.00	0.00	0.14	0.00	1.00	0.33
Mares	0.00	0.00	0.00	0.18	0.25	1.00	0.32
Geldings	0.00	0.00	0.00	0.17	0.33	0.50	0.24
Total	0.00	0.00	0.00	0.16	0.27	1.00	0.30

I.3-Summary statistics for the performance traits of the sample of raced Finnhorses between the ages 3 and 6 years.

Performance trait	Min	st Quantil	Median	Mean	3rd Quantil	Max	SD
Starts							
Stallions	1	6	10	10	14	31	6
Mares	1	3	8	9	13	32	6
Geldings	1	3	8	10	14	42	7
Total	1	4	8	9	14	42	7
Earnings							
Stallions	0	1720	5100	9745	10180	66180	12403
Mares	0	148	703	2758	3276	47250	5023
Geldings	0	250	1125	2967	2650	56400	6619
Total	0	290	1412	4665	5200	66180	8517
EPS							
Stallions	0	240	517	963	1133	6843	1267
Mares	0	28	106	254	300	5250	473
Geldings	0	45	120	313	275	8057	875
Total	0	50	175	455	498	8057	893
Victories							
Stallions	0	1	2	2	3	10	2
Mares	0	0	0	1	1	9	1
Geldings	0	0	0	1	2	6	1
Total	0	0	1	1	2	10	2
Victories (freq)							
Stallions	0.00	0.07	0.20	0.24	0.33	1.00	0.22
Mares	0.00	0.00	0.00	0.09	0.13	1.00	0.16
Geldings	0.00	0.00	0.00	0.10	0.14	1.00	0.18
Total	0.00	0.00	0.06	0.14	0.20	1.00	0.20
Race time auto							
Stallions	82.3	85.6	87.5	88.0	89.8	96.9	3.2
Mares	83.0	88.3	90.5	91.3	94.5	107.8	4.4
Geldings	83.1	88.2	91.0	91.9	93.6	115.4	5.7
Total	82.3	87.0	89.2	90.2	92.3	115.4	4.6
Race time volt							
Stallions	82.1	87.4	89.7	90.4	92.4	104.6	4.2
Mares	85.0	90.6	93.8	94.7	97.1	126.6	6.4
Geldings	84.1	90.5	93.1	94.2	96.6	117.1	5.7
Total	82.1	89.3	92.3	93.4	95.9	126.6	6.0
Placings							
Stallions	0	2	5	5	7	15	3
Mares	0	0	2	3	4	18	3
Geldings	0	1	3	3	5	14	3
Total	0	1	3	3	5	18	3
Placings (freq)							
Stallions	0.00	0.33	0.50	0.50	0.67	1.00	0.25
Mares	0.00	0.00	0.22	0.26	0.40	1.00	0.25
Geldings	0.00	0.08	0.28	0.31	0.45	1.00	0.26
Total	0.00	0.13	0.33	0.34	0.50	1.00	0.27
Unplaced							
Stallions	0	2	4	5	7	20	4
Mares	0	2	5	6	9	24	5
Geldings	0	2	6	7	10	30	5
Total	0	2	5	6	9	30	5
Unplaced (freq)							
Stallions	0.00	0.33	0.50	0.49	0.67	1.00	0.25
Mares	0.00	0.60	0.77	0.74	1.00	1.00	0.25
Geldings	0.00	0.55	0.73	0.69	0.92	1.00	0.26
Total	0.00	0.50	0.67	0.66	0.87	1.00	0.27
Disqualifications							
Stallions	0	0	1	2	2	7	2
Mares	0	0	1	2	3	8	2
Geldings	0	0	1	2	3	8	2
Total	0	0	1	2	3	8	2
Disqualifications (freq)							
Stallions	0.00	0.00	0.11	0.15	0.22	1.00	0.18
Mares	0.00	0.00	0.15	0.23	0.33	1.00	0.26
Geldings	0.00	0.00	0.17	0.24	0.33	1.00	0.26
Total	0.00	0.00	0.14	0.21	0.33	1.00	0.24

I.4-Summary statistics for the performance traits of the sample of raced Finnhorses between ages 7 and 10 years.

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Starts							
Stallions	1	6	14	14	22	36	9
Mares	1	5	10	13	18	45	9
Geldings	1	8	12	14	19	34	8
Total	1	6	12	13	19	45	9
Earnings							
Stallions	0	1560	7020	14670	17180	187350	24866
Mares	0	200	1000	3632	3370	52760	6749
Geldings	0	578	1905	4832	5155	33500	6975
Total	0	425	2020	7056	7930	187350	15175
EPS							
Stallions	0	145	489	961	1032	12490	1620
Mares	0	31	102	201	231	1884	301
Geldings	0	61	168	289	330	1955	362
Total	0	54	167	438	441	12490	960
Victories							
Stallions	0	0	1	2	3	14	2
Mares	0	0	0	1	1	8	1
Geldings	0	0	1	1	2	8	2
Total	0	0	1	1	2	14	2
Victories (freq)							
Stallions	0.00	0.00	0.09	0.15	0.20	1.00	0.19
Mares	0.00	0.00	0.00	0.05	0.08	0.63	0.09
Geldings	0.00	0.00	0.05	0.08	0.14	0.47	0.10
Total	0.00	0.00	0.03	0.09	0.14	1.00	0.13
Race time auto							
Stallions	78.6	82.9	95.0	95.0	87.1	95.7	3.0
Mares	82.0	85.4	88.5	88.7	91.2	104.0	4.2
Geldings	79.9	85.5	87.7	88.5	90.2	99.0	4.1
Total	78.6	84.6	86.9	87.5	89.7	104.0	4.2
Race time volt							
Stallions	80.3	84.6	86.1	86.9	88.0	101.2	3.6
Mares	83.9	87.1	89.9	90.6	93.1	113.4	4.8
Geldings	83.8	86.8	88.6	89.7	92.2	105.3	3.9
Total	80.3	86.0	88.2	89.3	91.8	113.4	4.5
Placings							
Stallions	0	1	4	5	8	17	4
Mares	0	0	2	3	4	19	3
Geldings	0	1	3	4	6	15	3
Total	0	1	3	4	6	19	4
Placings (freq)							
Stallions	0.00	0.17	0.32	0.34	0.50	1.00	0.25
Mares	0.00	0.00	0.18	0.20	0.29	1.00	0.19
Geldings	0.00	0.07	0.22	0.24	0.40	0.88	0.20
Total	0.00	0.08	0.22	0.25	0.37	1.00	0.22
Unplaced							
Stallions	0	4	8	10	14	30	7
Mares	0	4	8	10	14	40	8
Geldings	1	5	9	10	14	28	6
Total	0	4	8	10	14	40	7
Unplaced (freq)							
Stallions	0.00	0.50	0.68	0.66	0.83	1.00	0.34
Mares	0.00	0.71	0.81	0.80	1.00	1.00	0.19
Geldings	0.11	0.60	0.78	0.76	0.93	1.00	0.20
Total	0.00	0.63	0.78	0.75	0.92	1.00	0.22
Disqualifications							
Stallions	0	0	1	2	3	11	2
Mares	0	0	2	2	3	13	2
Geldings	0	0	2	2	4	10	2
Total	0	0	1	2	3	13	2
Disqualifications (freq)							
Stallions	0.00	0.00	0.07	0.12	0.19	1.00	0.16
Mares	0.00	0.00	0.13	0.20	0.25	1.00	0.24
Geldings	0.00	0.00	0.11	0.19	0.25	1.00	0.22
Total	0.00	0.00	0.11	0.17	0.25	1.00	0.22

I.5-Summary statistics of performance traits stratified by AA genotype of PR5826

Performance trait (AA)	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Number of starts	1	22	43	59	80	232	0
Victories	0	1	4	6	10	56	8
Victories (frequency)	0.00	0.04	0.08	0.11	0.14	0.63	0.11
Placings	0	4	12	18	26	123	18
Placings (frequency)	0.00	0.17	0.27	0.29	0.37	1.00	0.18
Unplaced	0	15	29	41	57	165	36
Unplaced (frequency)	0.00	0.63	0.73	0.71	0.83	1.00	0.18
Earnings (Euro)	0	1840	8430	32903	39975	516230	67001
Earnings per start	0	76	175	444	472	7309	831
Racing time auto (seconds)	78.6	83.7	86.6	87.4	89.9	105.3	4.7
Racing time volt (seconds)	80.3	85.6	88.5	90.0	92.4	121.2	6.4
Disqualifications	0	3	7	9	12	42	8
Disqualifications (frequency)	0.00	0.07	0.14	0.18	0.25	1.00	0.16
EBV	68	107	114	113	120	148	12

I.6-Summary statistics of performance traits stratified by AG genotype of PR5826

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Number of starts	1	16	20	66	39	329	117
Victories	0	1	1	4	5	17	6
Victories (frequency)	0.00	0.02	0.05	0.07	0.07	0.26	0.09
Placings	0	3	4	13	11	60	21
Placings (frequency)	0.00	0.11	0.18	0.20	0.24	0.53	0.17
Unplaced	1	10	18	53	31	269	96
Unplaced (frequency)	0.47	0.76	0.82	0.80	0.89	1.00	0.17
Earnings (Euro)	0	1092	1300	20803	7310	127515	47159
Earnings per start	0	49	84	165	267	440	177
Racing time auto (seconds)	82.5	89.7	93.7	92.2	95.3	99.2	5.9
Racing time volt (seconds)	84.1	90.3	91.3	95.7	93.6	126.6	14.0
Disqualifications	0	4	7	9	11	27	9
Disqualifications (frequency)	0.00	0.09	0.21	0.20	0.33	0.38	0.15
EBV	81	108	109	109	116	126	14

I.7-Summary statistics of performance traits stratified by TT genotype of PR3737

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Number of starts	1	24	44	60	80	232	51
Victories	0	1	4	7	10	56	8
Victories (frequency)	0.00	0.04	0.09	0.11	0.14	0.63	0.12
Placings	0	5	12	19	28	123	19
Placings (frequency)	0.00	0.18	0.28	0.29	0.40	1.00	0.18
Unplaced	0	17	30	41	56	165	36
Unplaced (frequency)	0.00	0.61	0.72	0.71	0.82	1.00	0.18
Earnings (Euro)	0	1868	9388	36199	41035	516230	72320
Earnings per start	0	82	187	485	494	7309	864
Racing time auto (seconds)	78.6	83.7	86.3	87.3	89.7	105.3	5.0
Racing time volt (seconds)	80.3	85.4	88.3	89.8	92.1	126.6	6.5
Disqualifications	0	4	7	9	12	42	8
Disqualifications (frequency)	0.00	0.07	0.13	0.17	0.22	1.00	0.15
EBV	68	107	115	114	122	148	12

I.8-Summary statistics of performance traits stratified by CT genotype of PR3737

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Number of starts	1	18	39	56	71	329	61
Victories	0	1	2	4	5	18	5
Victories (frequency)	0.00	0.03	0.06	0.07	0.09	0.41	0.08
Placings	0	4	9	14	20	60	14
Placings (frequency)	0.00	0.16	0.25	0.26	0.33	0.83	0.17
Unplaced	1	12	27	42	57	269	49
Unplaced (frequency)	0.17	0.67	0.75	0.74	0.84	1.00	0.17
Earnings (Euro)	0	1815	4284	17797	12235	127515	29673
Earnings per start	0	72	126	240	300	1641	338
Racing time auto (seconds)	82.5	85.0	89.5	88.9	91.7	96.5	4.2
Racing time volt (seconds)	84.0	87.4	90.7	91.4	93.3	121.2	7.4
Disqualifications	0	2	7	10	15	34	9
Disqualifications (frequency)	0.00	0.07	0.17	0.22	0.32	1.00	0.19
EBV	84	104	109	109	117	134	11

I.9-Summary statistics of performance traits stratified by TT genotype of PR8604

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Number of starts	1	22	44	61	80	329	56
Victories	0	1	4	6	10	56	8
Victories (frequency)	0.00	0.04	0.08	0.10	0.14	0.57	0.10
Placings	0	4	12	19	27	123	19
Placings (frequency)	0.00	0.17	0.28	0.29	0.37	1.00	0.18
Unplaced	0	15	31	43	57	269	41
Unplaced (frequency)	0.00	0.63	0.72	0.71	0.83	1.00	0.18
Earnings (Euro)	0	1759	8395	32976	41556	516230	63311
Earnings per start	0	78	181	415	467	4886	683
Racing time auto (seconds)	79.9	83.7	86.6	87.3	89.7	104.0	4.6
Racing time volt (seconds)	81.4	85.6	88.4	90.1	92.4	126.6	6.9
Disqualifications	0	3	7	9	12	42	8
Disqualifications (frequency)	0.00	0.06	0.14	0.18	0.23	1.00	0.15
EBV	68	107	114	114	121	138	11

I.10-Summary statistics of performance traits stratified by TG genotype of PR8604

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
Number of starts	1	21	37	51	65	162	42
Victories	0	1	3	6	6	32	8
Victories (frequency)	0.00	0.03	0.07	0.12	0.14	0.63	0.14
Placings	0	4	9	15	20	63	16
Placings (frequency)	0.00	0.14	0.21	0.27	0.38	0.73	0.19
Unplaced	1	16	27	36	51	123	30
Unplaced (frequency)	0.27	0.62	0.79	0.74	0.86	1.00	0.19
Earnings (Euro)	0	1752	5184	30683	26244	460495	77850
Earnings per start	0	58	127	509	454	7309	1213
Racing time auto (seconds)	78.6	84.7	88.6	88.8	92.7	105.3	5.9
Racing time volt (seconds)	80.3	86.1	89.7	90.5	92.7	111.2	6.4
Disqualifications	0	4	8	9	13	36	8
Disqualifications (frequency)	0.00	0.10	0.17	0.21	0.28	1.00	0.18
EBV	72	103	109	111	118	148	13

I.11-Summary statistics of performance traits according to AA genotype of PR5826 and stratified by age (3 years, 3-6 years and 7-10 years).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
3 YEAR OLDS							
Number of starts	1	1	2	2	3	6	1
Victories	0	0	0	1	1	3	1
Victories (frequency)	0.00	0.00	1.00	0.21	0.50	1.00	0.34
Placings	0	0	1	1	2	6	1
Placings (frequency)	0.00	0.00	1.00	0.54	1.00	1.00	0.46
Unplaced	0	0	25	1	1	2	1
Unplaced (frequency)	0.00	0.00	300.00	0.46	1.00	1.00	0.46
Earnings (Euro)	0	30	200	1320	1980	8430	2048
Earnings per start	0	30	95	476	633	2800	644
Racing time auto (seconds)	90.1	91.5	101.2	95.6	96.8	107.8	6.0
Racing time volt (seconds)	91.8	94.2	0.0	100.4	104.0	116.2	6.2
Disqualifications	0	0	0	0	0	1	0
Disqualifications (frequency)	0.00	0.00	0.00	0.14	0.00	1.00	0.29
3 TO 6 YEAR OLDS							
Number of starts	1	4	8	9	14	42	7
Victories	0	0	0	1	2	10	2
Victories (frequency)	0.00	0.00	3.00	0.14	0.20	1.00	0.20
Placings	0	1	0	3	5	18	3
Placings (frequency)	0.00	0.13	5.00	0.34	0.50	1.00	0.27
Unplaced	0	2	1	6	9	30	5
Unplaced (frequency)	0.00	0.50	1430.00	0.67	0.87	1.00	0.27
Earnings (Euro)	0	290	177	4809	5375	66180	8678
Earnings per start	0	51	89	469	506	8057	911
Racing time auto (seconds)	82.3	87.0	92.1	90.1	92.2	115.4	4.6
Racing time volt (seconds)	82.1	89.2	1.0	93.3	95.9	121.2	5.8
Disqualifications	0	0	0	2	3	8	2
Disqualifications (frequency)	0.00	0.00	0.13	0.21	0.33	1.00	0.25
7 TO 10 YEAR OLDS							
Number of starts	1	6	12	13	19	39	9
Victories	0	0	0	1	2	14	2
Victories (frequency)	0.00	0.00	3.00	0.09	0.14	1.00	0.14
Placings	0	1	0	4	6	19	4
Placings (frequency)	0.00	0.08	8.00	0.26	0.38	1.00	0.22
Unplaced	0	4	1	10	14	35	7
Unplaced (frequency)	0.00	0.63	2035.00	0.75	0.92	1.00	0.22
Earnings (Euro)	0	433	168	7196	7990	187350	15390
Earnings per start	0	58	87	449	457	12490	974
Racing time auto (seconds)	78.6	84.6	88.1	87.4	89.7	104.0	4.0
Racing time volt (seconds)	80.3	86.0	1.0	89.2	91.8	113.4	4.5
Disqualifications	0	0	0	2	3	13	2
Disqualifications (frequency)	0.00	0.00	0.10	0.17	0.25	1.00	0.22

I.10-Summary statistics of performance traits according to AG genotype of PR5826 and stratified by age (3 years, 3-6 years and 7-10 years).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
3 YEAR OLDS							
Number of starts	2	2	3	3	3	3	1
Victories	0	0	0	0	0	0	0
Victories (frequency)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Placings	0	0	1	1	1	1	1
Placings (frequency)	0.00	0.08	0.17	0.17	0.25	0.33	0.24
Unplaced	2	2	2	2	2	2	2
Unplaced (frequency)	0.67	0.75	0.83	0.83	0.92	1.00	0.24
Earnings (Euro)	0	125	250	250	375	500	354
Earnings per start	0	42	83	83	125	167	118
Racing time auto (seconds)	100.8	100.8	100.8	100.8	100.8	100.8	100.8
Racing time volt (seconds)	103.2	103.2	103.2	103.2	103.2	103.2	103.2
Disqualifications	1	2	2	2	3	3	1
Disqualifications (frequency)	0.50	0.54	0.59	0.59	0.63	0.67	0.12
3 TO 6 YEAR OLDS							
Number of starts	1	3	6	8	10	20	7
Victories	0	0	0	1	1	2	1
Victories (frequency)	0.00	0.00	0.00	0.07	0.13	0.40	0.11
Placings	0	1	1	2	3	5	2
Placings (frequency)	0.00	0.09	0.22	0.23	0.37	0.50	0.18
Unplaced	1	2	3	6	7	19	6
Unplaced (frequency)	0.50	0.63	0.78	0.77	0.90	1.00	0.18
Earnings (Euro)	0	240	530	1363	2370	5470	1687
Earnings per start	0	20	96	153	259	507	127
Racing time auto (seconds)	90.4	91.0	93.7	94.3	96.4	100.8	3.8
Racing time volt (seconds)	91.0	92.4	93.2	96.6	95.2	126.6	9.6
Disqualifications	0	1	2	2	4	4	2
Disqualifications (frequency)	0.00	0.08	0.22	0.27	0.37	1.00	0.28
7 TO 10 YEAR OLDS							
Number of starts	1	7	15	20	36	45	16
Victories	0	0	1	1	2	3	1
Victories (frequency)	0.00	0.00	0.03	0.04	0.06	0.15	0.05
Placings	0	1	4	4	6	9	3
Placings (frequency)	0.00	0.05	0.20	0.15	0.23	0.35	0.12
Unplaced	1	7	11	17	29	40	14
Unplaced (frequency)	0.65	0.77	0.80	0.85	0.95	1.00	0.12
Earnings (Euro)	0	315	2680	3924	7760	12110	4285
Earnings per start	0	34	168	132	184	311	102
Racing time auto (seconds)	83.7	84.7	88.9	90.0	95.3	99.2	5.8
Racing time volt (seconds)	85.1	87.3	90.2	89.7	92.4	94.0	3.3
Disqualifications	1	2	2	3	3	8	2
Disqualifications (frequency)	0.02	0.05	0.20	0.26	0.41	1.00	0.29

I.11-Summary statistics of performance traits according to TT genotype of PR3737 and stratified by age (3 years, 3-6 years and 7-10 years).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
3 YEAR OLDS							
Number of starts	1	1	2	2	2	4	1
Victories	0	0	0	0	1	2	1
Victories (frequency)	0.00	0.00	1.00	0.19	0.38	1.00	0.34
Placings	0	0	0	1	2	3	1
Placings (frequency)	0.00	0.00	1.00	0.49	1.00	1.00	0.46
Unplaced	0	0	1	1	1	2	1
Unplaced (frequency)	0.00	0.00	300.00	0.52	1.00	1.00	0.46
Earnings (Euro)	0	8	190	977	1200	5600	1408
Earnings per start	0	8	97	450	600	2800	645
Racing time auto (seconds)	91.0	95.1	101.2	98.0	100.0	107.8	5.8
Racing time volt (seconds)	93.7	94.3	0.0	100.8	105.1	116.2	6.2
Disqualifications	0	0	0	0	0	1	0
Disqualifications (frequency)	0.00	0.00	0.00	0.13	0.00	1.00	0.31
3 TO 6 YEAR OLDS							
Number of starts	1	4	8	9	14	42	7
Victories	0	0	0	1	2	10	2
Victories (frequency)	0.00	0.00	3.00	0.15	0.23	1.00	0.20
Placings	0	1	0	3	5	18	3
Placings (frequency)	0.00	0.14	5.00	0.34	0.50	1.00	0.27
Unplaced	0	2	1	6	9	30	5
Unplaced (frequency)	0.00	0.50	1655.00	0.66	0.86	1.00	0.27
Earnings (Euro)	0	300	196	5115	5610	66180	9105
Earnings per start	0	50	89	496	517	8057	963
Racing time auto (seconds)	82.3	86.8	91.9	90.0	92.2	115.4	4.7
Racing time volt (seconds)	82.1	89.0	1.0	93.1	95.6	126.6	5.9
Disqualifications	0	0	0	2	3	8	2
Disqualifications (frequency)	0.00	0.00	0.13	0.20	0.30	1.00	0.24
7-10YO							
Number of starts	1	7	13	14	19	39	9
Victories	0	0	0	1	2	14	2
Victories (frequency)	0.00	0.00	3.00	0.09	0.15	1.00	0.14
Placings	0	1	0	4	6	19	4
Placings (frequency)	0.00	0.11	8.00	0.27	0.40	1.00	0.22
Unplaced	0	4	1	10	14	35	7
Unplaced (frequency)	0.00	0.60	2560.00	0.73	0.89	1.00	0.22
Earnings (Euro)	0	630	183	8088	8870	187350	16626
Earnings per start	0	67	87	506	516	12490	1054
Racing time auto (seconds)	78.6	84.3	87.9	87.1	89.0	104.0	4.0
Racing time volt (seconds)	80.3	85.8	1.0	88.8	90.9	113.4	4.2
Disqualifications	0	0	0	2	3	13	2
Disqualifications (frequency)	0.00	0.00	0.10	0.15	0.23	1.00	0.20

I.12-Summary statistics of performance traits according to CT genotype of PR3737 and stratified by age (3 years, 3-6 years and 7-10 years).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
3 YEAR OLDS							
Number of starts	2	2	3	3	4	6	2
Victories	0	0	1	1	2	3	1
Victories (frequency)	0.00	0.00	0.17	0.26	0.46	0.75	0.32
Placings	0	1	2	2	3	6	2
Placings (frequency)	0.00	0.42	0.71	0.63	0.94	1.00	0.39
Unplaced	0	0	1	1	2	2	1
Unplaced (frequency)	0.00	0.06	0.29	0.38	0.58	1.00	0.39
Earnings (Euro)	0	328	550	2200	2700	8430	3294
Earnings per start	0	143	183	549	688	1405	551
Racing time auto (seconds)	90.1	90.6	91.0	91.0	91.5	91.9	1.3
Racing time volt (seconds)	91.8	94.0	100.7	98.8	103.1	104.0	5.5
Disqualifications	0	0	1	1	1	3	1
Disqualifications (frequency)	0.00	0.06	0.29	0.29	0.46	0.67	0.27
3 TO 6 YEAR OLDS							
Number of starts	1	3	7	9	13	31	6
Victories	0	0	0	1	1	8	1
Victories (frequency)	0.00	0.00	0.00	0.10	0.13	0.75	0.16
Placings	0	1	2	3	4	11	3
Placings (frequency)	0.00	0.14	0.28	0.32	0.47	1.00	0.27
Unplaced	0	2	5	6	8	24	5
Unplaced (frequency)	0.00	0.53	0.72	0.68	0.85	1.00	0.27
Earnings (Euro)	0	245	790	2920	3900	32270	5284
Earnings per start	0	53	117	295	304	3227	498
Racing time auto (seconds)	83.7	88.6	90.7	91.1	94.0	104.2	4.2
Racing time volt (seconds)	84.0	91.0	93.9	94.7	96.7	121.2	6.4
Disqualifications	0	0	1	2	3	8	2
Disqualifications (frequency)	0.00	0.00	0.17	0.24	0.33	1.00	0.26
7 TO 10 YEAR OLDS							
Number of starts	1	4	9	12	19	45	11
Victories	0	0	0	1	1	5	1
Victories (frequency)	0.00	0.00	0.00	0.05	0.06	0.50	0.10
Placings	0	0	1	2	4	12	3
Placings (frequency)	0.00	0.00	0.14	0.17	0.27	1.00	0.19
Unplaced	0	4	8	10	14	40	9
Unplaced (frequency)	0.00	0.73	0.86	0.83	1.00	1.00	0.19
Earnings (Euro)	0	70	820	2922	3105	18100	4516
Earnings per start	0	19	82	164	207	1658	239
Racing time auto (seconds)	82.9	85.6	89.7	89.5	92.7	98.9	4.3
Racing time volt (seconds)	84.3	87.1	90.7	90.9	93.6	104.1	4.6
Disqualifications	0	1	2	2	3	10	2
Disqualifications (frequency)	0.00	0.05	0.18	0.25	0.37	1.00	0.27

I.13-Summary statistics of performance traits according to TT genotype of PR8604 and stratified by age (3 years, 3-6 years and 7-10 years).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
3 YEAR OLDS							
Number of starts	1	1	2	2	3	6	1
Victories	0	0	0	1	1	3	1
Victories (frequency)	0.00	0.00	1.00	0.23	0.50	1.00	0.35
Placings	0	0	1	1	2	6	1
Placings (frequency)	0.00	0.00	1.00	0.56	1.00	1.00	0.44
Unplaced	0	0	0	1	1	2	1
Unplaced (frequency)	0.00	0.00	330.00	0.44	1.00	1.00	0.44
Earnings (Euro)	0	65	200	1173	1200	8430	1895
Earnings per start	0	65	97	401	616	1405	401
Racing time auto (seconds)	90.1	92.9	101.2	97.4	100.0	107.8	6.4
Racing time volt (seconds)	91.8	94.2	0.0	99.4	103.0	107.6	4.9
Disqualifications	0	0	0	0	1	3	1
Disqualifications (frequency)	0.00	0.00	0.00	0.20	0.33	1.00	0.32
3 TO 6 YEAR OLDS							
Number of starts	1	4	8	9	14	42	7
Victories	0	0	0	1	2	10	2
Victories (frequency)	0.00	0.00	3.00	0.14	0.20	1.00	0.20
Placings	0	1	0	3	5	18	3
Placings (frequency)	0.00	0.14	5.00	0.35	0.50	1.00	0.27
Unplaced	0	2	1	6	8	30	5
Unplaced (frequency)	0.00	0.50	1413.00	0.65	0.86	1.00	0.27
Earnings (Euro)	0	310	193	4579	5410	56400	7897
Earnings per start	0	54	89	441	501	8057	828
Racing time auto (seconds)	83.0	86.9	92.2	90.0	92.2	107.8	4.2
Racing time volt (seconds)	82.1	89.1	1.0	93.3	95.9	126.6	5.8
Disqualifications	0	0	0	2	3	8	2
Disqualifications (frequency)	0.00	0.00	0.13	0.21	0.33	1.00	0.24
7-10YO							
Number of starts	1	7	12	14	20	45	9
Victories	0	0	0	1	2	14	2
Victories (frequency)	0.00	0.00	3.00	0.08	0.14	0.86	0.12
Placings	0	1	0	4	6	19	4
Placings (frequency)	0.00	0.09	9.00	0.25	0.36	1.00	0.21
Unplaced	0	4	1	10	15	40	8
Unplaced (frequency)	0.00	0.64	2550.00	0.75	0.91	1.00	0.21
Earnings (Euro)	0	563	178	7148	8340	82800	12039
Earnings per start	0	60	87	425	455	5669	712
Racing time auto (seconds)	79.9	84.3	88.0	87.1	89.2	104.0	3.9
Racing time volt (seconds)	81.4	85.9	1.0	89.1	91.5	113.4	4.5
Disqualifications	0	0	0	2	3	13	2
Disqualifications (frequency)	0.00	0.00	0.09	0.16	0.25	1.00	0.21

I.14-Summary statistics of performance traits according to TG genotype of PR8604 and stratified by age (3 years, 3-6 years and 7-10 years).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
3 YEAR OLDS							
Number of starts	1	1	1	2	2	4	1
Victories	0	0	0	0	0	1	0
Victories (frequency)	0.00	0.00	0.00	0.10	0.00	0.50	0.22
Placings	0	0	0	1	2	3	1
Placings (frequency)	0.00	0.00	0.00	0.35	0.75	1.00	0.49
Unplaced	0	1	1	1	1	1	0
Unplaced (frequency)	0.00	0.25	1.00	0.65	1.00	1.00	0.49
Earnings (Euro)	0	0	140	1544	1980	5600	2418
Earnings per start	0	0	140	687	495	2800	1198
Racing time auto (seconds)	91.0	92.0	92.9	92.9	93.9	94.8	2.7
Racing time volt (seconds)	94.2	99.0	103.5	104.3	108.8	116.2	9.3
Disqualifications	0	0	0	0	0	0	0
Disqualifications (frequency)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3 TO 6 YEAR OLDS							
Number of starts	1	4	9	9	13	23	6
Victories	0	0	1	1	2	8	2
Victories (frequency)	0.00	0.00	0.05	0.14	0.20	0.86	0.19
Placings	0	0	3	3	4	14	3
Placings (frequency)	0.00	0.00	0.22	0.28	0.47	1.00	0.27
Unplaced	0	3	6	6	9	20	5
Unplaced (frequency)	0.00	0.53	0.79	0.72	1.00	1.00	0.27
Earnings (Euro)	0	148	1325	5206	4435	66180	10912
Earnings per start	0	25	126	533	493	6843	1141
Racing time auto (seconds)	82.3	87.1	89.1	90.7	91.9	115.4	6.2
Racing time volt (seconds)	82.9	89.9	92.4	94.0	96.2	116.2	6.7
Disqualifications	0	0	2	2	3	8	2
Disqualifications (frequency)	0.00	0.00	0.15	0.24	0.33	1.00	0.26
7 TO 10 YEAR OLDS							
Number of starts	1	5	10	11	16	30	7
Victories	0	0	0	1	2	12	2
Victories (frequency)	0.00	0.00	0.00	0.11	0.16	1.00	0.18
Placings	0	0	2	3	5	14	4
Placings (frequency)	0.00	0.00	0.20	0.26	0.40	1.00	0.25
Unplaced	0	4	7	8	12	25	6
Unplaced (frequency)	0.00	0.60	0.80	0.75	1.00	1.00	0.25
Earnings (Euro)	0	265	1210	7036	3440	187350	23830
Earnings per start	0	49	112	507	386	12490	1598
Racing time auto (seconds)	78.6	85.4	88.5	88.8	92.3	99.2	4.7
Racing time volt (seconds)	80.3	86.4	88.9	89.7	93.0	105.3	4.5
Disqualifications	0	0	2	2	3	10	2
Disqualifications (frequency)	0.00	0.00	0.15	0.20	0.28	1.00	0.23

APPENDIX II

II.1- P values for fixed effects and covariates included in the models of racing performance traits in Finnhorses for their total career. Significant p-values appear in bold.

Trait	Fixed effect					Covariates			
	Sex	Birth year	DMRT3	MSTN.PR3737	MSTN.PR8604	Starts	Earnings	Race time (auto)	Race time (volt)
Starts	0.154	**0.005**	0.110	*	*	-	**<0.0001**	-	-
Wins	**0.015**	**<0.001**	0.466	*	*	*	-	**0.0024**	**<0.0001**
Wins frequency	**0.002**	0.797	**<0.0001**	*	*	-	-	**0.011**	**<0.0001**
Placings	**0.020**	**0.002**	**0.033**	*	*	**<0.0001**	-	**<0.0001**	**0.002**
Placings frequency	**0.0005**	0.642	**<0.0001**	*	*	-	-	**0.001**	*
Unplaced	**0.005**	**<0.001**	0.365	*	*	**<0.001**	-	**0.002**	**<0.0001**
Unplaced frequency	**0.002**	**<0.0001**	0.175	*	*	-	-	**<0.0001**	**<0.0001**
Earnings	**0.0001**	**<0.0001**	0.074	*	*	**<0.0001**	-	-	-
EPS	**<0.0001**	**<0.0001**	**0.038**	*	*	-	-	-	-
Race time (auto)	**<0.0001**	0.145	**0.041**	*	*	-	-	-	-
Race time (volt)	**<0.0001**	**<0.0001**	**0.044**	*	*	-	-	-	-
Disqualifications	0.314	0.086	**0.0001**	*	*	**<0.0001**	-	-	-
Disqualification frequency	**0.025**	0.091	0.246	*	*	-	-	-	-
EBV	-	-	-	**0.003**	0.227	-	-	-	-

-, trait that was never included in the model

*, trait that was dropped from the model (p>0.05)

II.2. P values for fixed effects and covariates included in the models of racing performance traits in Finnhorses for their careers between the ages of 3 and 6 years. Significant p-values appear in bold.

Trait	Fixed effect					Covariates			
	Sex	Birth year	DMRT3	MSTN.PR3737	MSTN.PR8604	Starts	Earnings	Race time (auto)	Race time (volt)
Starts	**0.002**	**<0.0001**	**<0.0001**	*	**0.036**	-	**<0.0001**	-	-
Wins	**<0.0001**	**0.0006**	0.095	*	*	**<0.0001**	-	*	**<0.0001**
Wins frequency	**<0.0001**	**<0.0001**	**0.002**	*	*	-	-	*	**<0.0001**
Placings	**0.004**	**<0.0001**	0.817	*	*	**<0.0001**	-	**0.024**	**0.0008**
Placings frequency	**<0.0001**	**<0.0001**	**0.022**	*	**0.024**	-	-	*	**<0.0001**
Unplaced	**0.0002**	**<0.0001**	**0.020**	*	*	**<0.0001**	-	*	**<0.0001**
Unplaced frequency	**<0.0001**	**<0.0001**	**0.002**	*	**0.040**	-	-	*	**<0.0001**
Earnings	**<0.0001**	**0.0005**	**<0.0001**	*	**0.016**	**<0.0001**	-	-	-
Earnings per start	**<0.0001**	0.140	**<0.0001**	**0.038**	*	-	-	-	-
Race time (auto)	**<0.0001**	**<0.0001**	**2.00E-04**	**0.049**	*	-	-	-	-
Race time (volt)	**<0.0001**	**<0.0001**	**<0.0001**	**0.002**	*	-	-	-	-
Disqualifications	**0.005**	0.497	**0.002**	*	*	**<0.0001**	-	-	-
Disqualification frequency	**0.008**	**0.003**	0.146	*	*	-	-	-	-

-, trait that was never included in the model

*, trait that was dropped from the model (p>0.05)

II.3- P values for fixed effects and covariates included in the models of racing performance traits in Finnhorses for their careers between the ages of 7 and 10 years. Significant p-values appear in bold.

Trait	Fixed effect					Covariates			
	Sex	Age	DMRT3	MSTN.PR3737	MSTN.PR8604	Starts	Earnings	Race time (auto)	Race time (volt)
Starts	**<0.0001**	0.054	**0.006**	*	*	-	**<0.0001**	-	-
Wins	**0.001**	**0.0002**	**0.0040**	*	*	**<0.0001**	-	**<0.0001**	*
Wins frequency	**<0.0001**	**0.0005**	**<0.0001**	*	*	-	-	**<0.0001**	*
Placings	0.0840	**<0.0001**	0.598	*	*	-	-	**<0.0001**	**<0.0001**
Placings frequency	**0.0140**	**<0.0001**	**0.0180**	*	*	-	-	*	**<0.0001**
Unplaced	**0.014**	**<0.0001**	**0.002**	*	*	**<0.0001**	-	*	**<0.0001**
Unplaced frequency	**0.0050**	**<0.0001**	**0.012**	*	*	-	-	*	**<0.0001**
Earnings	**<0.0001**	**0.036**	**0.0003**	**0.000**	**0.043**	**<0.0001**	-	-	-
Earnings per start	**<0.0001**	0.176	**<0.0001**	**<0.0001**	**0.016**	-	-	-	-
Race time (auto)	**<0.0001**	**0.012**	0.116	**<0.0001**	**0.0006**	-	-	-	-
Race time (volt)	**<0.0001**	**0.0010**	**<0.0001**	**<0.0001**	*	-	-	-	-
Disqualifications	0.370	0.111	**0.0007**	*	**0.031**	**<0.0001**	-	-	-
Disqualification frequency	**0.0250**	0.740	**0.0180**	**0.0040**	*	-	-	-	-

-, trait that was never included in the model

*, trait that was dropped from the model (p>0.05)

APPENDIX III

III.1-Summary statistics of performance traits of riding Finnhorses by AA genotype of PR5826. Scale from 1 (poor) to 6 (perfect).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Years of riding	0.50	3.00	5.00	5.84	8.00	19.00
Canter Rythm	1.00	2.00	4.00	3.66	5.00	6.00
Canter Balance	1.00	2.00	4.00	3.60	5.00	6.00
Canter Transitions	1.00	3.00	3.00	3.43	4.00	6.00
Galop Rythm	1.00	3.00	4.00	3.66	5.00	6.00
Galop Balance	1.00	2.00	3.00	3.51	5.00	6.00
Galop Transitions	1.00	2.00	3.00	3.34	4.00	6.00
Difficulty of achieving current galop form	1.00	2.00	3.00	3.24	4.50	6.00
Gathered Trot Rythm	1.00	3.00	4.00	3.77	5.00	6.00
Gathered Trot Balance	1.00	3.00	4.00	3.75	5.00	6.00
Gathered Trot Transitions	1.00	3.00	4.00	3.74	5.00	6.00
Increased Trot Rythm	1.00	3.00	4.00	3.62	4.00	6.00
Increased Trot Balance	1.00	3.00	3.00	3.51	4.50	6.00
Increased Trot Transitions	1.00	2.00	4.00	3.43	4.00	6.00
Walk Rythm	2.00	4.00	4.00	4.33	5.00	6.00
Walk Balance	2.00	3.50	5.00	4.37	5.00	6.00
Walk Transitions	2.00	3.00	4.00	4.19	5.00	6.00
Jumping evaluation	1.00	3.00	4.00	3.91	5.00	6.00

Other factors	"Yes"	"No"	NA	Lätt A	Lätt B	Lätt C
Was the horse a trotter previously?	22	42	6			
Can the horse tölt?	9	55	6			
Can the horse pace?	17	47	6			
Has the horse competed in dressage?	40	23	7			
If so, what is the highest class it competed in?			34	12	13	10

III.2- Summary statistics of performance traits of riding Finnhorses by AG genotype of PR5826. Scale from 1 (poor) to 6 (perfect).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Years of riding	1.00	1.50	2.00	2.00	2.50	3.00
Canter Rythm	4.00	4.50	5.00	4.67	5.00	5.00
Canter Balance	4.00	4.00	4.00	4.33	4.50	5.00
Canter Transitions	4.00	4.00	4.00	4.00	4.00	4.00
Galop Rythm	3.00	3.50	4.00	4.00	4.50	5.00
Galop Balance	3.00	3.50	4.00	3.67	4.00	4.00
Galop Transitions	2.00	2.50	3.00	3.00	3.50	4.00
Difficulty of achieving current galop form	1.00	1.50	2.00	2.30	3.00	4.00
Gathered Trot Rythm	3.00	3.50	4.00	4.00	4.50	5.00
Gathered Trot Balance	3.00	4.00	5.00	4.33	5.00	5.00
Gathered Trot Transitions	2.00	3.00	4.00	3.67	4.50	5.00
Increased Trot Rythm	2.00	2.50	3.00	3.33	4.00	5.00
Increased Trot Balance	3.00	3.00	3.00	3.33	3.50	4.00
Increased Trot Transitions	2.00	2.50	3.00	2.67	3.00	3.00
Walk Rythm	5.00	5.00	5.00	5.33	5.50	6.00
Walk Balance	5.00	5.00	5.00	5.33	5.50	6.00
Walk Transitions	5.00	5.00	5.00	5.33	5.50	6.00
Jumping evaluation	3.00	3.50	4.00	4.33	5.00	6.00

Other factors	"Yes"	"No"	NA	Lätt A	Lätt B	Lätt C
Was the horse a trotter previously?	0	3	4			
Can the horse tölt?	0	3	4			
Can the horse pace?	0	3	4			
Has the horse competed in dressage?	2	1	4			
If so, what is the highest class it competed in?				1	1	

III.3- Summary statistics of performance traits of riding Finnhorses by TT genotype of PR3737. Scale from 1 (poor) to 6 (perfect).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Years of riding	0.50	3.00	4.00	5.89	8.00	19.00
Canter Rythm	1.00	3.00	4.00	3.85	5.00	6.00
Canter Balance	1.00	3.00	4.00	3.76	5.00	6.00
Canter Transitions	1.00	3.00	4.00	3.65	4.75	6.00
Galop Rythm	1.00	3.00	4.00	3.76	5.00	6.00
Galop Balance	1.00	3.00	3.50	3.65	5.00	6.00
Galop Transitions	1.00	3.00	4.00	3.52	4.00	6.00
Difficulty of achieving current galop form	1.00	2.00	4.00	3.36	5.00	6.00
Gathered Trot Rythm	1.00	3.00	4.00	3.78	5.00	6.00
Gathered Trot Balance	1.00	3.00	4.00	3.76	5.00	6.00
Gathered Trot Transitions	1.00	3.00	4.00	3.71	5.00	6.00
Increased Trot Rythm	2.00	3.00	4.00	3.62	4.00	6.00
Increased Trot Balance	1.00	3.00	3.00	3.51	4.00	6.00
Increased Trot Transitions	1.00	3.00	3.00	3.45	4.00	6.00
Walk Rythm	2.00	4.00	5.00	4.43	5.00	6.00
Walk Balance	2.00	4.00	5.00	4.41	5.00	6.00
Walk Transitions	2.00	3.00	4.00	4.25	5.00	6.00
Jumping evaluation	1.00	3.00	4.00	4.04	5.00	6.00

Other factors	"Yes"	"No"	NA	Lätt A	Lätt B	Lätt C
Was the horse a trotter previously?	16	37	2			
Can the horse tölt?	5	49	1			
Can the horse pace?	12	42	1			
Has the horse competed in dressage?	34	19	2			
If so, what is the highest class it competed in?				12	10	7

III.4- Summary statistics of performance traits of riding Finnhorses by CT genotype of PR3737. Scale from 1 (poor) to 6 (perfect).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Years of riding	0.50	2.00	4.00	5.21	6.00	15.00
Canter Rythm	1.00	2.00	4.00	3.47	4.00	6.00
Canter Balance	1.00	2.00	4.00	3.47	4.00	6.00
Canter Transitions	1.00	2.00	3.00	3.12	4.00	6.00
Galop Rythm	1.00	3.00	4.00	3.59	5.00	5.00
Galop Balance	1.00	2.00	3.00	3.18	4.00	5.00
Galop Transitions	1.00	2.00	3.00	2.82	3.00	5.00
Difficulty of achieving current galop form	1.00	1.75	3.00	2.81	4.00	5.00
Gathered Trot Rythm	1.00	2.75	4.00	3.69	5.00	6.00
Gathered Trot Balance	1.00	2.75	4.00	3.81	5.00	6.00
Gathered Trot Transitions	1.00	2.75	4.00	3.69	5.00	6.00
Increased Trot Rythm	1.00	2.75	3.00	3.44	5.00	5.00
Increased Trot Balance	1.00	2.75	3.00	3.50	5.00	5.00
Increased Trot Transitions	1.00	2.00	3.50	3.19	4.00	5.00
Walk Rythm	2.00	4.00	4.00	4.31	5.00	6.00
Walk Balance	2.00	4.00	4.50	4.44	5.00	6.00
Walk Transitions	2.00	3.00	4.00	4.19	5.00	6.00
Jumping evaluation	1.00	3.00	4.00	3.69	5.00	6.00

Other factors	"Yes"	"No"	NA	Lätt A	Lätt B	Lätt C
Was the horse a trotter previously?	7	10				
Can the horse tölt?	3	13	1			
Can the horse pace?	4	12	1			
Has the horse competed in dressage?	10	6	1			
If so, what is the highest class it competed in?				4	2	3

III.5- Summary statistics of performance traits of riding Finnhorses by TT genotype of PR8604. Scale from 1 (poor) to 6 (perfect).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Years of riding	0.50	2.00	5.00	5.42	8.00	18.00
Canter Rythm	1.00	2.25	4.00	3.63	5.00	6.00
Canter Balance	1.00	2.25	4.00	3.69	5.00	6.00
Canter Transitions	1.00	3.00	4.00	3.44	4.00	6.00
Galop Rythm	1.00	3.00	4.00	3.57	5.00	6.00
Galop Balance	1.00	2.00	3.00	3.50	5.00	6.00
Galop Transitions	1.00	2.00	3.00	3.28	4.00	6.00
Difficulty of achieving current galop form	1.00	2.00	3.00	3.05	4.00	5.00
Gathered Trot Rythm	1.00	3.00	3.50	3.56	5.00	6.00
Gathered Trot Balance	1.00	3.00	4.00	3.67	5.00	6.00
Gathered Trot Transitions	1.00	3.00	4.00	3.56	4.20	6.00
Increased Trot Rythm	1.00	3.00	3.50	3.61	5.00	6.00
Increased Trot Balance	1.00	3.00	3.00	3.46	4.00	6.00
Increased Trot Transitions	1.00	2.00	3.00	3.36	4.00	6.00
Walk Rythm	2.00	4.00	5.00	4.40	5.00	6.00
Walk Balance	2.00	3.00	5.00	4.41	5.00	6.00
Walk Transitions	2.00	3.00	4.00	4.23	5.00	6.00
Jumping evaluation	1.00	3.00	4.00	4.00	5.00	6.00

Other factors	"Yes"	"No"	NA	Lätt A	Lätt B	Lätt C
Was the horse a trotter previously?	13	32	2			
Can the horse tölt?	6	39	2			
Can the horse pace?	13	32	2			
Has the horse competed in dressage?	27	18	2			
If so, what is the highest class it competed in?				9	7	7

III.6- Summary statistics of performance traits of riding Finnhorses by TG genotype of PR8604. Scale from 1 (poor) to 6 (perfect).

Performance trait	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Years of riding	1.00	3.00	4.00	6.28	11.00	19.00
Canter Rythm	2.00	3.00	4.00	4.00	5.00	6.00
Canter Balance	1.00	3.00	4.00	3.68	4.00	6.00
Canter Transitions	1.00	3.00	4.00	3.64	5.00	6.00
Galop Rythm	2.00	3.00	4.00	4.00	5.00	6.00
Galop Balance	1.00	3.00	3.00	3.68	5.00	6.00
Galop Transitions	1.00	3.00	4.00	3.56	4.00	5.00
Difficulty of achieving current galop form	1.00	3.00	3.00	3.44	4.00	5.00
Gathered Trot Rythm	2.00	3.00	4.00	4.16	5.00	6.00
Gathered Trot Balance	2.00	3.00	4.00	4.04	5.00	6.00
Gathered Trot Transitions	2.00	3.00	4.00	4.00	5.00	6.00
Increased Trot Rythm	2.00	3.00	4.00	3.60	4.00	6.00
Increased Trot Balance	1.00	3.00	4.00	3.68	5.00	6.00
Increased Trot Transitions	1.00	3.00	4.00	3.52	4.00	5.00
Walk Rythm	2.00	4.00	4.00	4.32	5.00	6.00
Walk Balance	2.00	4.00	4.00	4.44	5.00	6.00
Walk Transitions	2.00	4.00	4.00	4.28	5.00	6.00
Jumping evaluation	1.00	3.00	4.00	4.04	5.00	6.00

Other factors	"Yes"	"No"	NA	Lätt A	Lätt B	Lätt C
Was the horse a trotter previously?	9	16				
Can the horse tölt?	2	23				
Can the horse pace?	4	21				
Has the horse competed in dressage?	16	8	1			
If so, what is the highest class it competed in?				5	6	3

APPENDIX IV

IV.1-Summary of horse types in category of Finnhorses with no performance data

Type	
Riding/trotter	60
Riding/trotter (tryouts)	4
Riding/trotter (banned)	1
Riding/trotter (in training)	2
Trotter started 2017	1
Riding	4
Mini	19
Work	4
Trotter	3
NA	13

IV.2-Type distribution of Finnhorses with no performance data according to PR5826 genotype

Type	AA	AG	GG
Riding/trotter	58	1	-
Riding/trotter (tryouts)	4	0	-
Riding/trotter (banned)	1	0	-
Riding/trotter (in training)	1	0	-
Trotter started 2017	1	0	-
Riding	2	2	-
Mini	17	2	-
Work	4	0	-
Trotter	3	0	-
Total	91	5	0

IV.3-Type distribution of Finnhorses with no performance data according to PR3737 genotype

Type	TT	CT	CC
Riding/trotter	39	16	4
Riding/trotter (tryouts)	3	1	0
Riding/trotter (banned)	1	0	0
Riding/trotter (in training)	2	0	0
Trotter started 2017	1	0	0
Riding	3	1	0
Mini	11	7	0
Work	3	1	0
Trotter	3	0	0
Total	66	26	4

IV.4-Type distribution of Finnhorses with no performance data according to PR8604 genotype

Type	TT	TG	GG
Riding/trotter	50	8	1
Riding/trotter (tryouts)	3	1	0
Riding/trotter (banned)	1	0	0
Riding/trotter (in training)	1	1	0
Trotter started 2017	1	0	0
Riding	3	1	0
Mini	18	1	0
Work	3	1	0
Trotter	2	1	0
Total	82	14	1