Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Is a potential regulatory region in the vicinity of *EDN3* linked to the performance in Standardbred horses?

Linda Nováčková
Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics
Uppsala

Swedish University of Agricultural Sciences
Department of Animal Breeding and Genetics Examensarbete /
SLU, Institutionen för husdjursgenetik no 542

Degree project 30 credits

Uppsala 2018

Cover: Trotter racer
(Picture: L. Nováčková)

Print: SLU Service/Repro, Uppsala 2018

Is a potential regulatory region in the vicinity of *EDN3* linked to the performance in Standardbred horses?

Je potencionální regulační oblast v blízkosti genu *EDN3* spojena s výkonností Amerického klusáckého koně?

Linda Nováčková

Main supervisor: Gabriella Lindgren, SLU, Department of Animal Breeding and Genetics

Assisting supervisors: Brandon Velie, SLU, Department of Animal Breeding and Genetics
Maria Rosengren, SLU, Department of Animal Breeding and Genetics

Examiner: Göran Andersson, SLU, Department of Animal Breeding and Genetics

Credits:	30 hp
Course title:	Degree project in Animal Science
Course code:	EX0556
Programme:	Master's Programme –Animal Science
Level:	Advanced, A2E
Place of publication:	Uppsala
Year of publication:	2018
Name of series:	Examensarbete/Swedish University of Agriculture Sciences, Department of Animal Breeding and Genetics, No 542
On-line publication:	http://epsilon.slu.se

Key words: Performance, Standardbred, *EDN3, endothelin-3*

Is a potential regulatory region in the vicinity of *EDN3* linked to the performance in Standardbred horses?

Abstract

Standardbred horses (SBs) are used in harness racing. Their common breeding history with Coldblooded trotters (CBTs) led us to the hypothesis, that the same region associated with race performance in CBTs, may be linked to the performance of SBs as well. EDN3 SNP (g.22:45748491C>T), was found to be associated ($p<0.05$) with performance traits in CBTs (Fegraeus, 2017). EDN3 SNP has a high frequency of the T-allele in breeds related to races and to the high performance (Fegraeus, 2017). EDN3 SNP is located in the vicinity of e*ndothelin-3* (*EDN3*) gene. Protein endothelin-3 (EDN3) is responsible for the development of the enteric nervous system and the melanin distribution in the foetus and it is involved in the blood regulation. Custom designed TaqMan SNP Genotyping Assays (Applied Biosystems StepOnePlusTM Instrument by Life Technologies) were used to genotype 550 SBs. Four hundred and fifty-seven genotypes were obtained and no statistically significant association between racing performance traits and genotype was found. Several associations (<0.05) between performance traits and other fixed effects – sex and year of birth - were found. The T-allele frequency of EDN3 SNP is 0.87 and the frequency of the C-allele is 0.13 for 467 genotyped SBs. Whole genome sequencing of homozygotes for the C-allele and the T-allele in trotter horses is planned. The follow-up research into the *EDN3* vicinity is recommended.

Keywords: Performance, Standardbred, *EDN3, endothelin-3*

Author's address: Linda Nováčková, SLU, Department of Animal Breeding and Genetics, P.O. Box 7023, 750 07 Uppsala, Sweden

Je potencionální regulační oblast v blízkosti genu *EDN3* spojena s výkonností Amerického klusáckého koně?

Abstrakt

Americký klusák je oblíbené dostihové plemeno. Jeho historie je spojena se Severským chladnokrevným klusákem (CBT), v jehož chovném programu byl Americký klusák využíván pro zlepšení dostihových vlastností. Tato skutečnost nás vedla k hypotéze, že stejná oblast genomu, která je spojena s výkonností CBTs, může být zodpovědná za výkonnosti Amerického klusáka. EDN3 SNP (g.22:45748491C>T) je zodpovědný ($p<0.05$) za výkonnost Severského chladnokrevného klusáka (Fegraeus, 2017). Frekvence T-alely v EDN3 SNP je vysoká v plemenech využívaných pro dostihový sport a náročnější jezdecký sport. EDN3 SNP se nachází v blízkosti genu *endothelin-3* (*EDN3)*. Protein endothelin-3 (EDN3) je zodpovědný za vývoj nervů trávicí soustavy a distribuci melaninu v plodu koně a také se podílí na regulaci krevního tlaku. Zakázkově navržený TaqMan SNP test genotypizace (Aplikované Biosystémy StepOnePlusTM nástrojem Life Technologies) byl použit ke genotypizaci 550 vzorků DNA Amerického klusáckého koně. Celkem šedesát sedm genotypů bylo získáno, a přitom nebyla nalezena statistický signifikantní asociace mezi žádnou proměnnou reprezentující výkonnost a genotypem. Bylo nalezeno několik asociací (<0.05) mezi výkonnostními proměnnými a fixními efekty – pohlaví a rok narození. Frekvence T-alely v EDN3 SNP je 0,8 a C-alely je 0,13. Sekvenování celého genomu je plánováno na homozygotech pro C a T alelu u vybraných koní klusáckých plemen. Navazující výzkum oblasti genu *EDN3* je doporučen.

Klíčová slova: Výkonnost, Americký klusák, *EDN3, endothelin-3*

Adresa autora: Linda Nováčková, SLU, Department of Animal Breeding and Genetics, P.O. Box 7023, 750 07 Uppsala, Sweden

Contents

Abbreviations

ATP	Adenosine triphosphate
BLUP	Best linear unbiased prediction
CBT	Coldblooded trotter
DNA	Deoxyribonucleic acid
EBV	Estimated breeding value
EDN3	Endothelin-3
EDNRB	Endothelin receptor type B gene
ENCC	Enteric neural crest cells
ET	Endothelin
ETA	Endothelin receptor A
ETB	Endothelin receptor B
G	Gelding
GWAS	Genome wide association study
HSCR	Hirschsprung disease
LD	Linkage disequilibrium
LWFS	Lethal white foal syndrome
M	Mare
NSDH	North-Swedish draught horse
PCR	Polymerase chain reaction
S	Stallion
SB	Standardbred
SNP	Single-nucleotide polymorphism
TB	Thoroughbred
USA	United States of America
VEGF	Vascular endothelial growth factor

1 Literature Review

1.1 Importance of horses in the life of mankind

During the last 5000 years, horses have been an integrated part of human life (Clutton-Brock, 1999; Outram *et al.*, 2009). The functional use of horses has changed during the ages. First horses were hunted for meat. Later, when the domestication process had started, the role of a horse expanded to a human assistant in many ways. For example, horses have made long-distance transport easier, being a carrier of possession, to ride on or for pulling a carriage. In some cultures, horse milk has been used for human diet enrichment (Levine, 1998). With agriculture development, literally horsepower has been used for cultivating the land. To own a horse has always been seen as a sign of a high social position. When the civilization developed the use of horses changed from only practical use more to the purpose of joy. Different types of competitions and races were developed in different cultures. In Ancient Greece, the Olympic Games had chariot horse-races with two and four horses (Olympics, 2017) and in Mongolia, the tribes have held different types of racing games.

In Great Britain, horse racing became a local, traditional, seasonal sport from the 12th century and onwards. It turned professional 500 years later, with the Jockey Club founded as an organization behind races all over the United Kingdom (Horse racing History, 2017).

1.2 History of harness racing

Harness racing is as old as 3000 years, the first of these races, where horses were put in straps and fitting, were held in Ancient Greece. Modern trot races became common in the United States of America (USA) at the end of the 18th century (Hendricks, 1995). In 1871, the Grant Circuit, a Quadrilateral Trotting Combination, was established and it has been one of the most famous annual harness races up until today. For this type of race, there was a demand for a suitable breed. To meet this demand a new breed was established based on cross-breeding between local mares and the main gene donator Messenger, a Thoroughbred stallion, imported from England. Messenger was bred directly with local mares and with pure Thoroughbred mares and his offspring became part of a new breed – the Standardbred.

1.3 History of Standardbreds

In 1879 the Standardbred was officially established as a new breed in the USA. To be recognized as a Standardbred horse, the horse had to make a one-mile track in 2 minutes and 30 seconds (Parker, 2008; Jill, 2018).

The Standardbred has been very popular all over the world, including in the Nordic countries. Standardbreds were used as riding horses in different categories including dressage, show jumping and endurance but mainly they were used as trot racing horses (United States Trotting Association, 2017). Over the course of years, Standardbreds have been selected to improve the horse's racing performance (Fegraeus, 2017). The annual genetic performance progress of Standardbreds corresponds with 5% of the phenotypic standard deviation (Thiruvenkadan, 2009).

1.4 Endothelin-3 function

The *EDN3* gene is located on chromosome 22 (Ensembl, 2018a) and encodes the protein endothelin-3 (EDN3), the potent vasoactive peptide (Peroutka, 2007). Peptides are used mainly for cell-cell communication and they play the important role as transmitters in the nervous system (Reid *et al.*, 2018). Vasoactive peptides influence directly vascular and other smooth muscles (Reid *et al.*, 2018). EDN3 protein is produced in various tissues as muscular and nervous peripherals (Genetics Home Reference, 2018). EDN3 is known to influence the prenatal development of enteric nerves (Gazques *et al.*, 2016). Another known function of endothelin-3 is its role in the production of melanocytes during a prenatal development (Genetics Home Reference, 2018). The endothelin (ET) family contains three isoforms – ET-1, ET-2 and ET-3. There are two receptors for ET, endothelin receptor A (ETA) and endothelin receptor B (ETB) (Arinami, 1991; Davenport, 2002). The signalling pathway for these

receptors uses G-protein signalling cascade (Human Protein Reference Database, 2018).

G-protein coupled receptors receive their signal by binding to extracellular ligands. Then the G-protein coupled receptor activates the G-protein by stimulating the exchange of the guanosine diphosphate to the guanosine triphosphate. In active form, the G-protein alpha subunit dissociates from the beta-gamma complex. Effectors are released from subunits to continue the signalling pathway. After the signal is sent from the G-protein, the bound guanosine triphosphate is hydrolysed back to the guanosine diphosphate. The G-protein is deactivated (Tuteja, 2009).

The endothelin-3 leads to the placement of enteric neural crest cells (ENCC), which creates the enteric nervous system. The role of endothelin-3 in this process is three-fold. Firstly, it is functioning in a signalling pathway with β1-integrins subfamily. Integrins are cell surface receptors that are important for cell migration and survival. These receptors have one α and one β chain with several different subunits. The largest family is the β1-integrin. The β1-integrin has a task to regulate the colonization of ENCC and controls the general neurological network organization in the gut. Secondly, EDN3 influences directly ENCC. Endothelin-3 increases the adhesion of the enteric neural crest cells during their colonization of the embryonic gut. Thirdly, the protein is part of the maintenance and the recovery of enteric progenitors (Gazquez *et al.*, 2016). In prenatal development, the manifestation of the gene is restricted to mesenchymal cells of the caecum before and after neural crest cells arrived (Leibl *et al.*, 1999).

For their work, muscles depend on oxygen distribution, as oxygen is essential for aerobic adenosine triphosphate (ATP) production (Shubhangi, 2018). ATP is a basic energy unit in every mammal's body. Physick-Sheard (1985) described how greater flow capacity of oxygen in muscles influences the body exercise and that this capacity is a limiting factor in the performance. Muscle capillarity is an effect of intense training. The blood's capacity to carry oxygen can be the most important component of the horse's workability in the aerobic environment (Physick-Sheard, 1985) (Figure 1).

ET-1 and ET-2 are long-lasting constrictors of vessels (Davenport, 2002, Moore *et al.*, 1992). ET-1 and ET-3 were found to be stimulators of the vascular endothelial growth factor (VEGF) protein synthesis. VEGF is responsible for angiogenesis, the formation of new vessels from existing ones. This stimulation is performed by a signalling pathway including endothelin receptors ETA and ETB and protein kinase C (Pedram *et al.*, 1997). It was described in pigs how ET-1 and ET-3 regulate vasoconstriction (Ushio-Fukai *et al.*, 1995). The process is slightly different for the two endothelins ET-1 and ET-3, but both influence the level of plasmatic calcium (Ushio-Fukai *et al.*, 1995) (Figure 1).

Plasma concentrations of ET-1 and ET-3 proved that these vasoactive peptides have a role in the pathogenesis of the hepatorenal syndrome and liver disease (Moore *et al.*, 1992). Moore and his team (1992) compared plasma concentration of ET-1 and ET-2 in subjects with several categories of liver and renal diseases in comparison to healthy individuals.

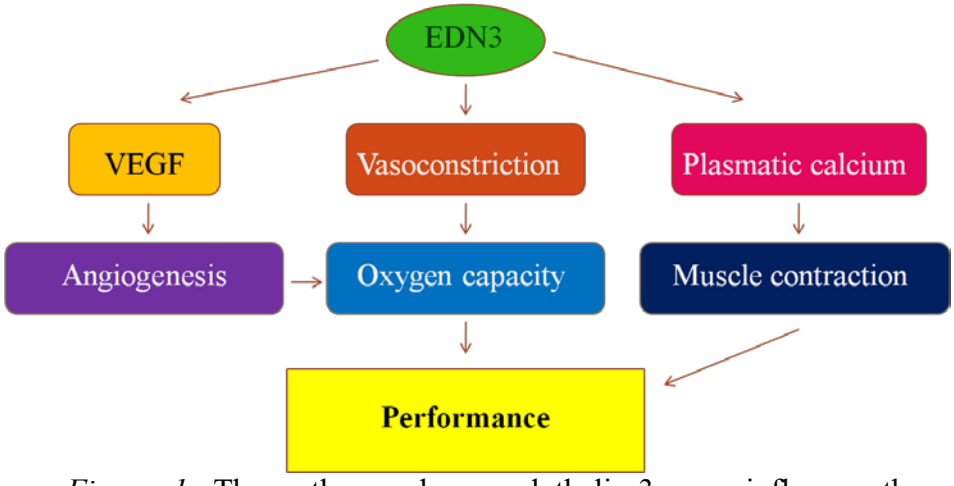

Figure 1. The pathways how endothelin-3 may influence the performance of horses.

The *EDN3* gene interacts with the *endothelin receptor type B* (*EDNRB)* gene, the latter being, responsible for the production of the endothelin receptor B (ETB). *EDNRB* is associated with the frame overo coat colour and Lethal white foal syndrome (LWFS). LWFS is manifesting by an intestinal aganglionosis (Ensembl, 2018b).

Different mutations of the *endothelin 3* gene cause terminal aganglionosis, known as the Hirschsprung's disease (Leibl *et al.*, 1999) and Waardenburg syndrome in humans (Touraine *et al.*, 2000). This syndrome is inherited in an autosomal recessive manner, where melanocytes and intrinsic ganglion cells of the terminal hindgut are absent (Touraine *et al.*, 2000). The role of *EDN3* in the development of Hirschsprung (HSCR) disease in a Chinese population is described by Garcia-Barceló *et al.* (2004). Two mutations of *EDN3* were described as additional HSCR disease-causing mutations. Since they were found only in a small part of the investigated sample, Garcia-

Barceló *et al.* (2004) added that these mutations can be either specific for a subpopulation or for certain individuals.

1.5 Performance trait of trotters

Performance in trotters is a complex trait which can have many characteristics. Fegraeus (2017) used number of starts, number of wins, number of placings (1-3), frequency of wins, frequency of placings, earnings, earnings per start, time-record-volt start per kilometre and time-record-auto start per kilometre as representative traits for performance. Volt starts mean horses are walking in a circle over a specific distance and starting a race from that formation. During auto start, a car with wings is launching the race. Statistically significant results with genotypes were found for the variables number of starts, number of wins, number of placings, earnings, earnings per start and time record volt start per race (Fegraeus, 2017).

Tanner *et al.* (2011) used the performance variables earnings and number of starts. In Tanner´s statistical model (2011), estimated breeding value (EBV) was included. The EBV for Swedish Standardbred trotters was based on these traits: number of races (start variable), percentage of races with first, second or third place, earnings (for every race and sum per career), record time per 1 km and racing (or start) status – if the horse started a race (1) or not (0) (Árnason, 1999). The best linear unbiased prediction (BLUP) evaluation has been used for Standardbreds in Sweden since 1989 and since 1992 the Animal model has replaced the previously used Sire model and has brought more accurate values (Árnason, 1999).

Cheetham *et al.* (2010) recommended using number of starts and earnings for performance variables, as well as age, sex, breed, surface of racetrack and gait. Cheetham *et al.* (2010) mentioned in their study, that castration influences the performance of a horse, especially if it was done during its racing career.

Suontama *et al.* (2013) described that the correlation in Standardbred foals was 0.54 for trot and earnings variables. In studbooks of Standardbreds the correlation between earnings and character was 0.70 and between earnings and movement was 0.88. These correlations were used in the EBV. Because these correlations have high values, the EBV may be used as a representative parameter of the performance trait.

The linear regression model was used in previous studies to compare genotype and phenotype data (Árnason, 1999; Cheetham *et al.*, 2010).

2 Introduction

2.1 Previous research

Investigation of genetic markers in the vicinity of *EDN3* was based on the results from previous research done by Fegraeus *et al.* (2018). Fegraeus *et al.* (2018) found several potential regions linked to the horse's performance in harness racing, by comparing a fixation index analysis of Standardbred horses, Coldblooded trotters and North-Swedish draught horses (NSDH). One statistically significant associated single-nucleotide polymorphism (SNP) located in the vicinity of *EDN3* linked to the performance in harness racing was found – SNP g.22: 45 748 491 C>T (EDN3 SNP). In genotyping of Coldblooded trotters, SNPs with high linkage disequilibrium (LD) to EDN3 SNP were chosen to further analyse - SNP g.22:45748082T>G, SNP g.22:45748586G>A, SNP g.22:45749526G>A and SNP g.22:45749595G>A. These five SNPs were significantly associated with performance traits in CBT. According to Fegraeus *et al.* (2018), the variability of EDN3 SNP alleles' variants within 14 breeds have an obvious pattern, where the T-allele had a higher frequency in highly

performing breeds such as the Thoroughbred, the SB, the Swedish Warmblood and the American Curly. On the other side, the C-allele is most often present in breeds, which doesn't race, as the NSDH and the Exmoor. Based on this theory, future research of EDN3 SNP was recommended in other racing breeds (Fegraeus *et al.,* 2018).

2.2 *EDN3* vicinity description

The location of EDN3 SNP is shown in Figure 2. Because of a recent update of the reference genome from EquCab 2.0 to EquCab 3.0 a new location of EDN3 SNP has been defined (Table 1). The distance of chosen SNP from the *EDN3* gene was counted and compared for both genomes (Table 1).

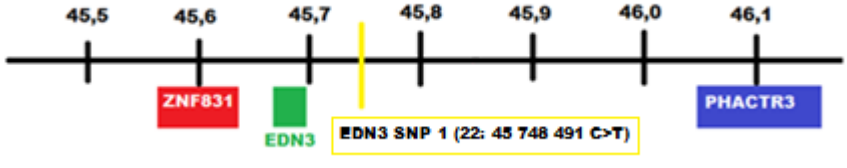

Figure 2. Location of EDN3 SNP towards surroundings genes on a chromosome 22, the forward strand (EquCab 2.0)

Table 1. The l*ocation and the distance between EDN3 SNP and EDN3 on a horse genome EquCab 2.0 and EquCab 3.0*

Chromosome 22	***EquCab 2.0***	***EquCab 3.0***
EDN3 SNP	45 748 491	46 717 861
EDN3 gene	45 674 735 - 45 696 542	46 643 067 – 46 665 984
Distance of EDN3 SNP from EDN3	51949 bp	51877 bp

Zinc Finger Protein 831 (ZNF831) is a protein-coding gene. This protein is metal ion binder (NCBI, 2018a). Other known function is transcriptional repression activity (UniProt, 2018). *Phosphatase and Actin regulator 3 (PHACTR3)* is also a protein-coding gene. Its suggested function is a regulatory subunit of protein phosphatase-1 (Gene cards, 2018b).

The EDN3 SNP wasn't found by using the BLAST method in the human genome. There are no transcription factors binding sites known in the location of EDN3 SNP on a horse genome (NCBI, 2018b).

In a human genome, the *EDN3* has 24 regulatory elements, one promoter and 24 enhancers, where the promoter is also behaving as an enhancer (Gene cards, 2018a). Regulatory elements influencing *EDN3* are located in both directions from the gene. *ZNF831 h*as three promoters which have also a role of enhancers plus 28 other enhancers (Gene cards, 2018c). These regulatory elements are located in both directions, upstream and downstream. *PHACTR3* has three promoters and two of them have a role of enhancer as well next to other 40 enhancers, they are located in both directions from gene (Gene cards, 2018b).

2.3 Aim of this study

The aim of this study was to investigate the potential association between EDN3 SNP and selected performance traits in Standardbred horses.

3 Material and Methods

3.1 Sample description

3.1.1 Population sample

The racing and pedigree data of Standardbred horses born between 2003 and 2014 were provided by the Swedish Trotting Association. Deoxyribonucleic acid (DNA) samples were provided by the Animal Genetics Laboratory, Swedish University of Agricultural Sciences (SLU). The cleaning of data was performed in SAS software 9.4 (SAS Institute Inc, 2013). Race data described information of 28 753 horses and pedigree data contained information of 44 780 individual horses.

3.1.2 Defining phenotypes

Race data contained 867 453 observations. This dataset described characteristics of individual races for each racing horse.

Raw race data were cleaned and variables representing different performance traits were estimated. First of all, only races with earning prices were chosen. For every variable, races data were individually

grouped for each horse by ID of the horse. New variables were created. The number of starts based on information of galloping and the average earnings per race by dividing total earnings with the number of starts. Performance sample data was controlled and duplicates were removed.

The dataset containing information about the pedigree was used to gain fixed effect variables for a statistical analysis. Extracting the first numbers of the dates of birth created the variable birth year. To avoid bias from unequal values, birth years were grouped into 3 groups - 2003-2006 (old age group), 2007-2010 (middle age group) and 2011-2014 (young age group).

A dataset with the ID of horses, for which DNA is stored in the Animal Genetics Laboratory, had 74 354 hair samples of Standardbreds.

3.1.3 Performance traits definitions

The performance of trotter horses is a complex trait and can be approach from different perspectives. Ten traits were chosen to represent the performance in this study. These are – the number of starts, total earnings, average earnings per race, gallops, disqualifications, wins, placings, unplacings, the best time per kilometre and the estimated breeding value.

Total earnings were calculated as the total number of all earnings per career of each horse.

The number of **starts** was calculated as the total number of times a horse entered a race.

Average earnings per race were calculated from total earnings divided by the number of starts for each horse.

The number of **gallops** was calculated as the total number of times a horse was galloping during a race.

The number of **wins** was calculated as the total number of times a horse finished a race in first place.

The number of **placings** was calculated as the total number of times a horse finished a race in first, second or third place.

The number of **unplacings** was calculated as the total number of times a horse finished a race in fourth or lower place.

The number of **disqualifications** was calculated as the total number of times a horse failed for different reasons, as a gallop, health problems or an involved person's failure.

The **best times per race** for each horse were defined as the lowest average time per kilometre in minutes, seconds and milliseconds.

The **estimated breeding value** was an individual value combining several characteristics as sex, age, the number of starts, earnings and it considered results of horse´s offspring as well.

3.1.4 Genotyped sample

Five hundred and fifty horses were randomly selected from the population sample, what corresponds to slightly less than 2% of population sample. More than 10 % of the genotyped sample dataset were double checked with the online database of the Swedish trotting association.

3.1.5 Summary statistics

The summary statistics of the Standardbred population sample and the genotyped sample were analysed using R software 3.4.4 (R core team, 2018).

3.2 Genotyping for *EDN3*

3.2.1 DNA isolation from hair samples

DNA was isolated from hair samples of Standardbreds, with the material being provided by the Animal Genetics Laboratory at the Department of Animal Breeding and Genetics, SLU. From each sample five to seven hairs were collected, hair bulbs were cut off and located randomly on plates. The DNA was isolated by adding 100µl Chelex (5%) and 7 µl proteinase K (20 mg/ml) in a 100 µl:7 µl ratio. The plates were heated for 60 min at 56°C and 10 min at 95°C. After cooling down, the plates were frozen. A total of 548 DNA samples were isolated.

3.2.2 StepOnePlus Real-Time PCR system (TaqMan)

Custom designed TaqMan SNP Genotyping Assays (Applied Biosystems StepOnePlusTM Instrument by Life Technologies) were used to genotype 548 DNA samples of Standardbreds. Analyses were performed with one negative and two positive controls of each genotype, 7 controls in total. TaqMan uses a double-stranded template and thanks to the fluorescence of the assay it detects the genotype of the desired SNP. 467 genotypes were obtained.

3.3 Association analysis

Statistical analyses were performed using the software program for statistical computing R (R Development Core Team, 2018). All dependent variables were tested for the normal distribution with the Anderson-Darling´s normality test and the Quantile-Quantile plot. The variables were not normally distributed and had to be transformed by adding formula “log (variant + 1)”. A linear regression model was used to investigate the links between genotypes and traits representing the performance. The sex, the country of birth, the year of birth and the genotype were set as fixed effects.

4 Results

4.1 Summary statistics comparing the population sample and the genotyped sample

4.1.1 Sex, country of birth and year of birth

In both datasets, the most represented sex was mares, followed by geldings and - on 3rd place - stallions (Figure 3).

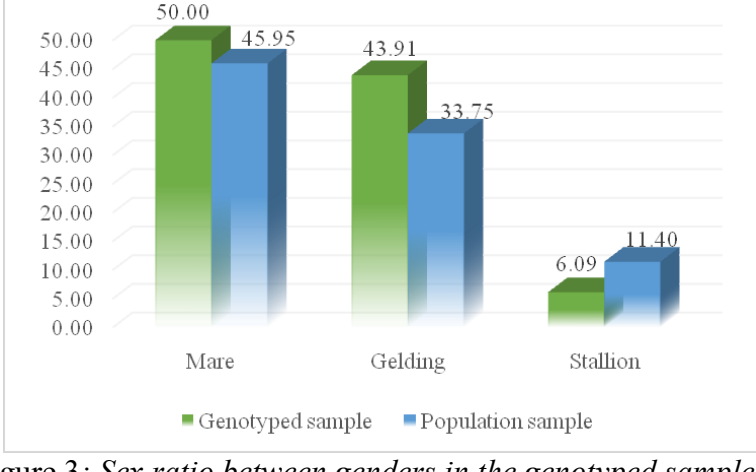

Figure 3*: Sex ratio between genders in the genotyped sample and the population sample for SB born 2003-2014*

Country of birth was divided into two levels – Sweden and other. "Other" includes for example: Germany, France, the USA, Finland and Norway. The proportions of these two levels in datasets can be seen in Figure 4. There was a higher ratio of Swedish horses in the genotyped sample than in the population sample.

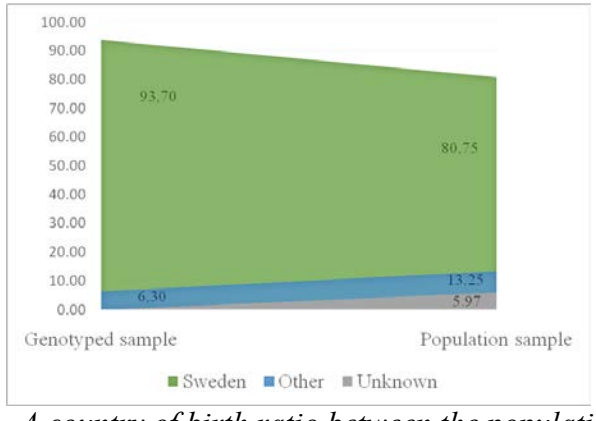

Figure 4. *A country of birth ratio between the population sample and the genotyped sample in SB born 2003-2014*

The horses used for analyses, as well as those in the population sample, were born between the years 2003 and 2014. There was a higher frequency of horses born 2013 in the genotyped sample than in the population sample. Horses born in the years 2014 and 2009 were less common in the genotyped sample than in the population sample. Other years were fairly equally present in both the population and the genotyped sample. Data about the birth year variable can be found in the Supplementary Table 8. Three age categories: Young age group (3-7 years old), Middle age group (8-10 years old) and Old age group (11 to 14 years old) – were created

to see how genetics influences the performance in young and older horses.

4.1.2 Performance traits

In the Supplementary Table 7, the summary statistics of traits representing the performance can be found. Minimum, mean, median and maximum were presented for the genotyped sample dataset and the population sample.

For all variables representing the performance traits minimum, median and mean were similar in both datasets; maximum was different. EBV was higher for the genotyped sample than for the population sample.

The maximum values in the summary statistics of performance traits may differ, because of a low frequency of high extremes in the population sample and using random sampling. The number of starts in the genotyped sample started with a value 1 because only racing horses were used in the analyses.

4.2 Genotyping EDN3 SNP

Five hundred forty-eight horses were genotyped and 467 genotypes were obtained. Some DNA samples had to low DNA concentration to be genotyped. The ratio of genotypes' frequency can be found in Figure 5. 75% of samples were homozygous for T-allele, 23% were heterozygous and only 2% were homozygous for C-allele. The alleles' frequency can be found in table 2.

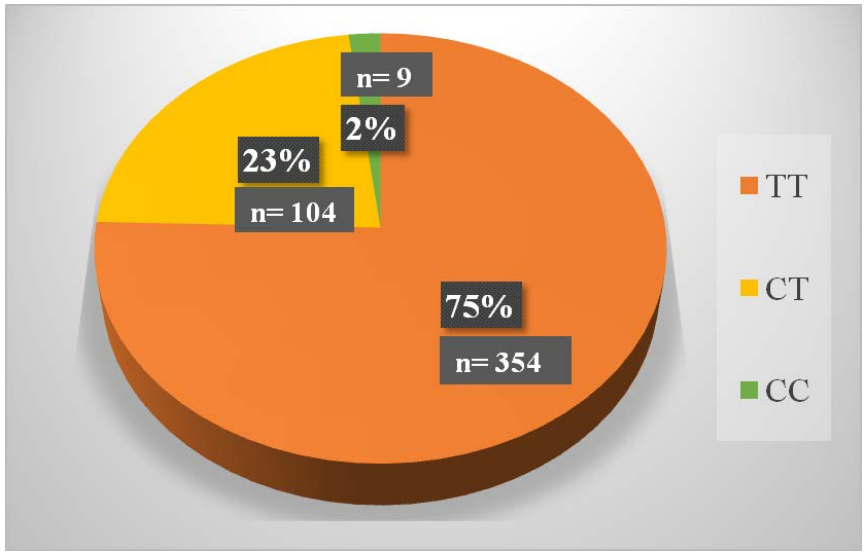

Figure 5. The frequency of genotypes for EDN3 SNP, n=467 Standardbreds born between 2003 and 2014

Table 2. *The alleles' frequency for EDN3 SNP in the genotyped sample of SB horses born 2003-2014, n=467*

SNP	T-allele	C-allele
EDN3 SNP	0.87	0.13

4.3 Association study

The main result of the study was that EDN3 SNP is not linked to performance in SBs.

The p-values of the performed statistical analysis for all traits are presented in Table 3. All p-values of association study between fixed effect genotype and performance traits were > 0.05. Analyses were done for several models using fixed effects sex, genotype, birth year and country of birth. Final models used for individual analyses are presented in Table 4.

Table 3. *Linear regression analyses performed in R for SB's performance traits and EDN3 SNP, n=467 born 2003-2014 Significant results p≤0.05 in bold*

Total earnings

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	-1.3081	0.5570	**0.0193**
Sex G vs. S	-1.1889	0.5607	**0.0345**
Sex G vs. M	0.1192	0.2669	0.6555
Birth year 07-10 vs. 03-06	0.5633	0.3638	0.1223
Birth year 11-14 vs. 03-06	-0.1033	0.3170	0.7447
Birth year 07-10 vs. 11-14	0.6665	0.3318	**0.0451**
Genotype CT vs. CC	-0.4191	0.9586	0.6622
Genotype TT vs. CC	-0.5132	0.9324	0.5823
Genotype CT vs. TT	0.0942	0.3078	0.7599

Starts

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	-0.2527	0.1851	0.1729
Sex G vs. S	-0.0924	0.1863	0.6201
Sex G vs. M	0.1603	0.0887	0.0715
Birth year 07-10 vs. 03-06	0.0556	0.1194	0.6420
Birth year 11-14 vs. 03-06	0.4095	0. 1041	**<0.0001**
Birth year 07-10 vs. 11-14	0.4651	0.1089	**<0.0001**
Genotype CT vs. CC	-0.0019	0.3186	0.9951
Genotype TT vs. CC	0.0576	0.3098	0.8526
Genotype CT vs. TT	-0.0596	0.1023	0.5607

Gallops

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	0.0225	0.1720	0.8960
Sex G vs. S	0.2321	0.1731	0.1803
Sex G vs. M	0.2096	0.0824	**0.0114**
Birth year 07-10 vs. 03-06	0.0465	0.1099	0.6724
Birth year 11-14 vs. 03-06	-0.2979	0.0958	**0.0019**
Birth year 07-10 vs. 11-14	0.3444	0.1003	**0.0006**
Genotype CT vs. CC	-0.0684	0.2960	0.8173
Genotype TT vs. CC	0.0809	0.2879	0.7786
Genotype CT vs. TT	-0.1494	0.0951	0.1168

Average earning per race

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	-1.0157	0.3999	**0.0114**
Sex G vs. S	-0.0522	0.4025	**0.0092**
Sex G vs. M	-0.0365	0.1916	0.8489
Birth year 07-10 vs. 03-06	0.4685	0.2597	0.0719
Birth year 11-14 vs. 03-06	0.2875	0.2263	0.2046
Birth year 07-10 vs. 11-14	0.1810	0.2368	0.4452
Genotype CT vs. CC	-0.4270	0.6882	0.5352
Genotype TT vs. CC	-0.6048	0.6693	0.3667
Genotype CT vs. TT	0.1778	0.2210	0.4216

Placings

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	-0.5514	0.1988	**0.0058**
Sex G vs. S	-0.3245	0.2007	0.1056
Sex G vs. M	0.2269	0.0953	**0.0176**
Birth year 07-10 vs. 03-06	0.1259	0.1269	0.3217
Birth year 11-14 vs. 03-06	-0.3366	0.1106	**0.0025**
Birth year 07-10 vs. 11-14	0.4626	0.1158	**<0.0001**
Genotype CT vs. CC	-0.0221	0.3421	0.9485
Genotype TT vs. CC	-0.0229	0.3327	0.9452
Genotype CT vs. TT	0.0008	0.1098	0.9944

Disqualifications

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	-0.1409	0.1685	0.4034
Sex G vs. S	0.0515	0.1696	0.7616
Sex G vs. M	0.16238	0.0808	**0.0176**
Birth year 07-10 vs. 03-06	0.1102	0.1080	0.3083
Birth year 11-14 vs. 03-06	-0.2869	0.0941	**0.0024**
Birth year 07-10 vs. 11-14	0.3981	0.0985	**<0.0001**
Genotype CT vs. CC	0.0447	0.2899	0.8774
Genotype TT vs. CC	0.0556	0.2820	0.8637
Genotype CT vs. TT	-0.0109	0.0931	0.9071

Unplacings

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	-0.0737	0.1812	0.6846
Sex G vs. S	0.0475	0.1824	0.7945
Sex G vs. M	0.1212	0.0868	0.1636
Birth year 07-10 vs. 03-06	0.0303	0.1163	0.0795
Birth year 11-14 vs. 03-06	-0.4268	0.1014	**<0.0001**
Birth year 07-10 vs. 11-14	0.4571	0.1061	**<0.0001**
Genotype CT vs. CC	-0.0826	0.3118	0.7912
Genotype TT vs. CC	0.0300	0.3033	0.9212
Genotype CT vs. TT	-0.1126	0.1001	0.2613

Wins

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	-0.4855	0.1662	**0.0037**
Sex G vs. S	-0.1992	0.1672	0.2343
Sex G vs. M	0.2863	0.0796	**0.0004**
Birth year 07-10 vs. 03-06	0.1012	0.1062	0.3412
Birth year 11-14 vs. 03-06	-0.2367	0.0925	**0.0108**
Birth year 07-10 vs. 11-14	0.3379	0.0968	**0.0005**
Genotype CT vs. CC	-0.1352	0.2859	0.6366
Genotype TT vs. CC	-0.1289	0.2781	0.6430
Genotype CT vs. TT	-0.0062	0.0918	0.9462

Best time per kilometre

Coefficients	Estimate	Standard error	p-value
Sex M vs. S	0.0106	0.0127	0.4017
Sex G vs. S	0.0014	0.0127	0.9132
Sex G vs. M	-0.0092	0.0061	0.1283
Birth year 07-10 vs. 03-06	-0.0109	0.0080	0.3173
Birth year 11-14 vs. 03-06	0.0062	0.0069	0.3710
Birth year 07-10 vs. 11-14	-0.0171	0.0073	**0.0196**
Genotype CT vs. CC	0.0115	0.0232	0.6225
Genotype TT vs. CC	0.0117	0.0227	0.6071
Genotype CT vs. TT	-0.0002	0.0070	0.9843

EBVs

Coefficients	Estimate	Standard error	p-value
Genotype CT vs. CC	-0.0271	0.0278	0.3300
Genotype TT vs. CC	-0.0312	0.0269	0.2500
Genotype CT vs. TT	0.0039	0.0090	0.6580

Table 4. *Final models of linear regression used for individual variables describing performance in SB*

Variable (Y)	**Model**
Total earnings	Y=Sex + Genotype + Birth year
Starts	Y=Sex + Genotype + Birth year
Gallops	Y=Sex + Genotype + Birth year
Average earnings per race	Y=Sex + Genotype + Birth year
Placings	Y=Sex + Genotype + Birth year
Disqualifications	Y=Sex + Genotype + Birth year
Unplacings	Y=Sex + Genotype + Birth year
Wins	Y=Sex + Genotype + Birth year
Best time per kilometre	Y=Sex + Genotype + Birth year
Estimated Breeding Values	Y= Genotype

There were statistically significant results for other fixed effects beside the genotype. According to the association analyses, the sex was associated with performance in several ways. In general, stallions had a probability to earn more money than geldings (p=0.0092) and mares (p=0.0114). Males had better predisposition to be placed on 1st, 2nd or 3rd place than females (p=0.0004). Mares had a lower risk to be disqualified (p=0.0176) or switch to the gallop (p=0.0114) during a race than geldings.

Another investigated fixed effect with a significant association with performance was the year of birth. It had a significant effect on all performance traits except average earnings per race and the EBV. For example: for total earnings, horses in the middle age group had a chance to earn 0.7 times more money than horses in the young age group (p=0.0451). The young age group horses had 0.3 times lower probability to switch to gallop during a race than the old age group horses (p=0.0006). Horses in the middle age group had 0.4 times

higher probability to be disqualified than horses from the young age group ($p<0.0001$) and the old age group horses had 0.3 times higher chance to be disqualified than rivals from the young age group ($p=0.0024$).

No significant effect was found for the country of birth in performance traits, therefore this fixed effect wasn't used in the final models.

5 Discussion

Endothelin-3 is mainly associated with horse´s prenatal development of the enteric nervous system and melanin distribution (Ensembl, 2018a). In other mammals, the link between *EDN3* and a blood pressure regulation (Ushio-Fukai *et al.*, 1995), an aganglionosis (Leibl *et al.*, 1999), a hepatorenal syndrome and liver diseases (Moore *et al.*, 1992) has been proven. The *EDN3* gene is located on equine chromosome 22 (Ensembl, 2018a). *Endothelin-3* may be linked to the performance thanks to its effect on the vasoconstriction. Blood pressure is an important factor during physical activity and oxygen capacity in muscles is a limiting factor in performance (Physick-Sheard, 1985). ET-1 and ET-3 are stimulating factors for VEGF and VEGF is important for control of angiogenesis. The number of muscle blood vessels is an essential factor for oxygen distribution and sufficient production of ATP in the exercising horse (Pedram *et al.*, 1997). Level of calcium (Ca^{2+}) is also an essential factor for a sufficient muscle contraction (Szent-Györgyi, 1975). ET-1 and ET-3 influence the Ca^{2+} level (Ushio-Fukai *et al.*, 1995).

The performance of trotters is a complex trait which can be approach from several different angles. Performance traits summary

statistics was performed to prove that the genotyped sample is representative of the population sample. Because the variables were not normally distributed, median should be used for the best comparison. Medians between the genotyped sample and the population sample were very similar and this showed that the genotyped sample is a good representation of the population sample.

As a fixed effect, sex, genotype, year and country of birth were considered. There are several environmental variables that could substantially influence the performance of a horse such as diet, climate and for example how many times a trainer was changed. Social environment and welfare status of the horse may be considered as well for evaluating the performance trait.

Cheetham *et al.* (2010) used as fixed effects age, sex, breed, surface of the racetrack and gait. Since in this study, only Standardbreds were used and only trot gait was recorded, these two attributes can be excluded. The surface of the racetrack was not included in provided data set by the trotting association. Cheetham *et al.* (2010) described how castration influences a horse's performance, mostly if it was done during its racing carrier. The Swedish trotting association recorded the castration. The most often used gender in harness races in Sweden was mares. On a second place were geldings. Very few stallions were used for racing. When a male horse proves it-self in racing at a young age, it goes to breeding and does not race anymore. The stallions, which were not suitable for breeding, were castrated and left in races as geldings. The fixed effect, country of birth was categorized into two levels. This was done because of the few horses born in a foreign

country in the population sample that was used for this study. Only horses, that were considered to be very good, were imported. Swedish horse owners probably choose horses bred in Sweden because of financial reasons. This situation can affect the statistical analysis, but the genotyped sample, in this case, represents the trend of the population sample. According to summary statistics, the frequencies of particular years of birth differ the most in 2013. This difference may be an effect of the random sampling.

Saastamoinen *et al.* (2009) found a link between young age at the start of the first race and the superior performance. A high correlation between the first season's performance and the performance of the following three years of the horses´ carrier was found (Saastamoinen *et al.* 2009). In this analysis, horses were divided into three categories of four years of birth.

Health is an important trait for every horse. It was found that straight hocks and curby hocks were negatively associated with the number of starts and earnings and cow hocks had a negative effect on a start status (Dolvik *et al.*, 2010). No health status was included in the performed study.

Only horses, which have passed the qualifying races, were included in the genotyped sample. It is possible that horses, which did not participate in any money earning races, homozygous for the non-performing allele variant (CC). If the genotyped sample includes these horses as well, the link between the phenotype and the genotype may be found. How horses are selected for racing career can depend on unwanted characteristics, such as unfitting personality, conformation

or health. Horses´ personality not suiting for racing can be shown in lack of motivation to compete, lack of a determination to win or whether a horse had a general problem to cope with the racing or the training environment. Horses excluded from the population sample, horses without racing career, from these reasons may carry the genotype what differs from highly performing horses and when such genotypes would be compared, several regions linked to performance may be found.

EDN3 SNP was proven to be associated with several performance traits in Coldblooded trotters (Fegraeus, 2017). In contrast, results of the performed analyses showed that the EDN3 SNP is not linked to any of the chosen representative traits of performance in Standardbreds.

In this study, only the one gene variant was examined. The performance trait is expected to have more quantitative trait loci as it is a complex trait.

The alleles' frequencies for EDN3 SNP in SBs (n=467) were found to be 0.87 for the T-allele and 0.13 for the C-allele. Fegraeus (2017) found frequencies of T- allele 0.86 and C-allele 0.14 in SBs (n=250).

To analyse for the potential association between the performance of SBs and EDN3 SNP was based on this association in CBTs. Before paternity tests were introduced in Coldblooded trotter breeding, CBTs were crossbred with Standardbreds to increase their racing performance (Fegraeus *et al.*, 2017). Both breeds have been bred for high racing performances. The T-allele is more often present in breeds used in races and riding sports (Fegraeus, 2017).

The expectation that EDN3 SNP may be associated with the performance of SBs was assumed on *EDN3* regulation of blood pressure through vasoconstriction, angiogenesis and also its influence of Ca^{2+} level in muscles vessels.

The hypothesis, that the CBT's and the SB's performance is associated with the same genotype, was tested. It is possible that the CBT and the SB share the haplotype containing gene variants associated with the performance and that the haplotypes don't have the same length and then the EDN3 SNP doesn't tag the gene variant in SBs. Obtained result from the performed association study may be caused by different in pronouncing of the EDN3 SNP´s effect in these two breeds. Or it is possible that the gene variants underlying the performance in CBTs an SBs are different. Future research is needed to find the right marker linked to the performance in trotter horses.

6 Conclusions

EDN3 SNP is not tagging the gene variant associated with the chosen performance traits in SBs. The alleles' frequencies for EDN3 SNP in SB (n=467) are for the T-allele 0.87 and for the C-allele 0.13. There is an association ($p<0.05$) between the performance and the sex plus the performance and the birth year. The haplotype tagging the associated gene variant for the performance in CBTs may have a different length than the haplotype of SBs related to the performance or the gene variant of the performance is different one in CBTs and SBs. Identity by descent mapping method, haplotype genotyping and whole genome sequencing may be used to compare different breeds divided in high performing and non-performing breeds. Statistical analysis within chosen family can show higher association for a chosen genotype. Because *EDN3* is involved in melanin distribution and coat colour pattern, mealy and non-mealy fixed coat colour in breeds may be used to analyse if the same mutation is responsible for mealy coat pattern and performance of horses. Further research is needed.

Acknowledgements

I would like to thank Gabriella Lindgren, my supervisor, for the possibility to run this study and for her enthusiastic approach that has motivated me in every step. Big thanks belong to Maria Rosengren and Brandon Velie for all their help, guidance and advice. Great thanks also to the whole EDN3 team, including Marina Solé Berga and Kim Jäderkvist Fegraeus. I would also like to thank all Animal Genetics Laboratory, SLU for providing DNA and hair samples of horses and their help with solving problems. To my examiner, Göran Andersson, thank you for your feedback. And finally I would like to thank Marcel Burger for giving spelling and grammar suggestions and Niklas Jannborg for his infinite support, the energy and ideas.

Uppsala, June 2018

Study contributions

This experiment was designed by Gabriella Lindgren and her team of researchers, Department of Animal Breeding and Genetics, SLU. Material and analysis tools were provided by the Animal Genetics Laboratory, SLU in Uppsala, Sweden. Race data were provided by the Swedish trotting association.

References

Arinami, T., Ishikawa, M., Inoue, A., Yanagisawa, M., Masaki, T., Yoshida, M.C., Hamaguchi, H. (1991). Chromosomal assignments of the human endothelin family genes: the endothelin-1 gene to 6p23-p24, the endothelin-2 gene to 1p34, and the endothelin -3 gene to 20q13.3-q13.3.*Am. J. Hum*. Genet.48(5):990-6. Available at: https://www.ncbi.nlm.nih.gov/pubmed/2018043?dopt

Árnason, T. (1999). Genetic evaluation of Swedish standard-bred trotters for racing performance traits and racing status. *Journal of Animal Breeding and Genetics,* 116: 387–398. doi:10.1046/j.1439-0388.1999.00202.x

Cheetham, J., Riordan, A.S., Mohammed, H.O., McIlwraith, C. W., Fortier, L. A. (2010). Relationships between race earnings and horse age, sex, gait, track surface and number of race starts for Thoroughbred and Standardbred racehorses in North America*. Equine Veterinary Journal* [online].**42**(4), 346-350 [2017-12-12]. DOI: 10.1111/j.2042-3306.2010.00032.x. ISSN 04251644. Available at: http://doi.wiley.com/10.1111/j.2042-3306.2010.00032.x

Clutton-Brock, J. (1999). A Natural History of Domesticated Mammals. *Cambridge University Press*, Cambridge, Massachusetts.

Davenport, A.P. (2002*).* International Union of Pharmacology. XXIX. Update on endothelin receptor nomenclature. *Pharmacol. Rev.* 54(2):219-26. Available at: https://www.ncbi.nlm.nih.gov/pubmed/12037137?dopt

Dolvik, N. I., Klemetsdal, G. (2010). Conformational Traits of Norwegian Cold-blooded Trotters: Heritability and the Relationship with Performance, *Acta Agriculturae Scandinavica, Section A — Animal Science*, 49:3,156-162, DOI: 10.1080/090647099424060

Ensembl. (2018a). EDN3 - Equus caballus. [online] *The European Bioinformatics Institute* [cit. 2018-05-09]. Available at: http://www.ensembl.org/Equus_caballus/Marker/Details?m=EDN3

Ensembl. (2018b). EDNRB - Equus caballus. [online]. *The European Bioinformatics Institute* [cit. 2018-05-09]. Available at: http://www.ensembl.org/Equus_caballus/Gene/Summary?db=core;g=ENSECAG00000024851;ph=22;r=17:50604167-50625930;t=ENSECAT00000026836

Fegraeus, K.J. (2017). Exploring the Horse Genome to Elucidate the Genetics of Gaits and Athletic Performance*. Acta Universitasis agriculturae Sueciae.* SLU, Sweden.

Fegraeus, K.J., Velie, B.D., Meadows, J.R.S., Lindgren, G. (2017). Using the unique breeding history of Coldblooded trotters to identify genes that influence athletic performance. *Journal of Equine Veterinary Science*. 52:73

Fegraeus, K.J, Velie, B.D., Axelsson, J., Ang, R., Hamilton, N.A., Andersson, L., Meadows, J.R.S., Lindgren, G. (2018). A potential regulatory region near the EDN3 gene may control both harness racing performance and coat color variation in horses. *Physiological Reports.* 6(10)

Garcia-Barceló, M., Sham, M.H., Lee, W.S., Lui, V.C., Chen, B.L., Wong, K.K., Wong, J.S., Tam, P.K. (2004). Highly recurrent RET

mutations and novel mutations in genes of the receptor tyrosine kinase and endothelin receptor B pathway in Chinese patients with sporadic Hirschsprung disease. *Clin. Chem.*10.1373/clinchem.2003.022061. Available at: http://clinchem.aaccjnls.org/content/50/1/93.long

Gazquez, E., Watanabe, Y., Broders-Bondon, F., Paul-Gilloteaux, P., Heysch, J., Baral, V., Bondurand, N. and Dufour, S. (2016*).* Endothelin-3 stimulates cell adhesion and cooperates with β1-integrins during enteric nervous system ontogenesis. *Scientific Reports,* 6(1).

Gene cards, (2018a). *Weizmann Institute of Science* . [cit. 2018-08-30]. Available at: https://www.genecards.org/cgi-bin/carddisp.pl?gene=EDN3

Gene cards, (2018b). *Weizmann Institute of Science* . [cit. 2018-08-30]. Available at: https://www.genecards.org/cgi-bin/carddisp.pl?gene=Phactr3

Gene cards, (2018c). *Weizmann Institute of Science* . [cit. 2018-08-30]. Available at: https://www.genecards.org/cgi-bin/carddisp.pl?gene=Znf831

Genetics Home Reference. (2018). EDN3 gene. [online] *U.S. National Library of Medicin*. Available at: https://ghr.nlm.nih.gov/gene/EDN3

Hendricks, B. L. (1995). International encyclopedia of horse breeds. *Norman: University of Oklahoma Press*. ISBN 0806127538.

Horse racing History (2017). *Equine World UK* [online]. UK. Available at: http://www.equine-world.co.uk/horse_sports/horse_racing_history.asp [2017-12-08]

Human protein reference database, (2018). [online] *Johns Hopkins University and the Institute of Bioinformatics.* [cit 2018-06-19]. Available at: http://www.hprd.org/interactions?hprd_id=00571&isoform_id=00571_1&isoform_name=Isoform_1

Jill, (2018). Standardbred Horse Breed Information. *The Stallion Place* [online]. Illinois: Exacta Design. Available at: https://web.archive.org/web/20100613183631/http:/www.thestallionplace.com/breed_info/standardbreds.htm [2018-01-04]

Leibl, M.A., Ota, T., Woodward, M.N., Kenny, S.E., Lloyd, D.A., Vaillant, C.R., Edgar, D.H. (1999*).* Expression of endothelin 3 by mesenchymal cells of embryonic mouse caecum. *Gut.* 44(2):246-52. Available at: https://www.ncbi.nlm.nih.gov/pubmed/9895385?dopt

Levine, M. (1998). Eating horses: The evolutionary significance of hippophagy.*Antiquity,*72(275), 90-100. Doi:10.1017/S0003598X00086300. Available at: https://www.cambridge.org/core/journals/antiquity/article/eating-horses-the-evolutionary-significance-of-hippophagy/632397D3F96FBA177EB0D47335107636

Moore, K., Wendon, J., Frazer, M., Karani, J., Williams, R., Badr, K. (1992). Plasma endothelin immunoreactivity in liver disease and the hepatorenal syndrome. *N. Engl. J. Med.* 327(25):1774-8. Available at: https://www.ncbi.nlm.nih.gov/pubmed/1435931?dopt

NCBI (2018a). *U.S. National Library of Medicine*. [cit. 2018-06-20]. Available at: https://www.ncbi.nlm.nih.gov/gene/128611

NCBI (2018b). *U.S. National Library of Medicine* [cit. 2018-08-30]. Available at: https://www.ncbi.nlm.nih.gov/genome/gdv/browser/?context=genome&acc=GCF_002863925.1

Olympics. (2017). Ancient Greece. Olympics. *University Press Inc.* Available at: http://www.ancientgreece.com/s/Olympics/ [2017-12-08]

Outram, A.K, Stear, N.A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert. V., Thorpe, N., Evershed, R.P. (2009). The Earliest Horse Harnessing and

Milking. *Science:* 1332-1335. Available at: http://science.sciencemag.org/content/323/5919/1332

Parker, M., (2008). The history of horse racing. *Real Estate, Inc.* Available at: http://www.mrmike.com/Explore/hrhist.htm [2017-12-08]

Pedram, A., Razandi, M., Hu, R.M., Levin, E.R. (1997). Vasoactive peptides modulate vascular endothelial cell growth factor production and endothelial cell proliferation and invasion. *J. Biol. Chem.*272(27):17097-103.Available at: https://www.ncbi.nlm.nih.gov/pubmed/9202027?dopt

Peroutka, S.J. (2007). Genetic Disorders of the Autonomic Nervous System. *Molecular Neurology*, Academic Press, San Diego. 321-332. ISBN 9780123695093. Available at: https://doi.org/10.1016/B978-012369509-3.50023-8

Physick-Sheard, P.W., (1985). Cardiovascular response to exercise and training in the horse. *Vet Clin North Am Equine Pract.* 1(2):383-417. Available at: https://www.ncbi.nlm.nih.gov/pubmed/3878196

R Core Team (2018). *R:* A language and environment for statistical computing. *R Foundation for Statistical Computing,* Vienna, Austria. Available at: https://www.R-project.org/

Reid, I.A., Katzung, B.A., Trevor, A.J., (2018). Chapter 17: Vasoactive Peptides. Basic & Clinical Pharmacology, 13th edition. Access Medicine. [cit. 2018-06-19]. Available at: https://accessmedicine.mnmedical.com/content.aspx?sectionid=69106039&bookid=1193

Saastamoinen, M. T., Ojala, M.J. (2009)Estimates of Genetic and Phenotypic Parameters for Racing Performance in Young Trotters,*Acta Agriculturae Scandinavica*,41:4,427-436,DOI: 10.1080/00015129109439925

SAS Institute Inc (2013). SAS/ACCESS® 9.4 Interface to ADABAS: Reference. *Cary, NC: SAS Institute Inc.*

Shubhangi, L., (2018). Inhibition of Glycolysis by Oxygen – Carbohydrates. *Biology discussion.* Available at: http://www.biologydiscussion.com/biochemistry/carbohydrates-biochemistry/inhibition-of-glycolysis-by-oxygen-carbohydrates/72573 [2018-05-27]

Suontama, M., van der Werf, J.H.J., Juga, J., Ojala, M. (2013). Genetic correlations for foal and studbook traits with racing traits and implications for selection strategies in the Finnhorse and Standardbred trotter. *Journal of Animal Breeding and Genetics*, 130: 178–189. doi:10.1111/j.1439-0388.2012.01011.x

Szent-Györgyi, A. G. (1975). Calcium regulation of muscle contraction. *Biophysical Journal*, *15*(7), 707–723. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1334730/

Tanner, J.C., Rogers, C.W., Firth, E.C. (2011). The relationship of training milestones with racing success in a population of Standardbred horses in New Zealand, *New Zealand Veterinary Journal*, 59:6, 323-327, DOI: 10.1080/00480169.2011.617029. Available at: https://doi.org/10.1080/00480169.2011.617029

Thiruvenkadan, A.K., Kandasamy, N., Panneerselvam, S. (2009). Inheritance of racing performance of trotter horses: an overview. Livestock Science; 124(1):163–181

Touraine, RL., Attié-Bitach, T., Manceau, E., Korsch, E., Sarda, P., Pingault, V., Encha-Razavi, F., Pelet, A., Augé, J., Nivelon-Chevallier, A., Holschneider, A.M., Munnes, M., Doerfler, W., Goossens, M., Munnich, A., Vekemans, M., Lyonnet, S. (2000*).* Neurological phenotype in Waardenburg syndrome type 4 correlates with novel SOX10 truncating mutations and expression in a developing brain. *Am. J. Hum. Genet.* ; 66(5):1496-503. Available at: https://www.ncbi.nlm.nih.gov/pubmed/10762540?dopt

Tuteja, N. (2009). Signalling through G protein coupled receptors. *Plant signalling behaviour.* 4(10): 942-947. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2801357/#!po=97.4576

UniProt, (2018). *UniProt Consortium.* [cit. 2018-06-20]. Available at: http://www.uniprot.org/uniprot/Q5JPB2

Ushio-Fukai, M., Nishimura, J., Kobayashi, S., Kanaide. (1995). Endothelin-1 and endothelin-3 regulate differently vasoconstrictor responses of the smooth muscle of the porcine coronary artery. H. *Br. J. Pharmacol.* 114(1):171-9. Available at: https://www.ncbi.nlm.nih.gov/pubmed/7712014?dopt

United States Trotting Association (*USTA)* [online] (2017). Ohio. Available at: http://www.ustrotting.com/ [2017-12-06]

The supplementary tables

The supplementary table 7. *Summary statistics (minimum, mean, median and maximum) of variables representing the performance trait in the genotyped sample and the population sample for SB born 2003-2014*

Total Earnings (Swedish crown)

variable	Genotyped sample	Population sample
Min	0	0
Median	93 204	94 150
Mean	269 974	244 708
Max	7 290 167	34 345 097

Average earnings per race (Swedish crown)

variable	Genotyped sample	Population sample
Min	0	0
Median	4 474	3 979
Mean	8 084	7 053
Max	132 010	731 730

Starts (sum)

variable	Genotyped sample	Population sample
Min	1	0
Median	19	23
Mean	26.28	30.17
Max	179	679

Gallop (sum)

variable	Genotyped sample	Population sample
Min	0	0
Median	4	5
Mean	6.05	7.03

Max	42	148

Placing (sum)

variable	Genotyped sample	Population sample
Min	0	0
Median	5	6
Mean	7.89	8.54
Max	59	217

Disqualification (sum)

variable	Genotyped sample	Population sample
Min	0	0
Median	4	4
Mean	4.92	5.54
Max	44	119

Unplacing (sum)

variable	Genotyped sample	Population sample
Min	0	0
Median	13	16
Mean	18.36	21.63
Max	128	460

Wins (sum)

variable	Genotyped sample	Population sample
Min	0	0
Median	2	2
Mean	2.87	2.95
Max	23	95

Best Time per km (minutes : seconds : milliseconds)

variable	Genotyped sample	Population sample
Min	1:00:0	1:00:0
Median	1:14:2	1:14:1
Mean	1:12:5	1:14:7
Max	1:25:7	9:00:0
NA's	45	3 190

Estimated breeding value

variable	Genotyped sample	Population sample
Min	76	64
Median	104	99
Mean	103.9	99.08
Max	124	135

NA's	7	4 601

The supplementary table 8. *The frequency of a birth year in the population sample and the genotyped sample for SBs born 2003-2014*

Birth year	Groups	Genotyped Sample frequency (%)	Population Sample frequency (%)	Age (years)
2003	Old age group	30	37.24	11-14
2004				
2005				
2006				
2007	Middle age group	25	33.71	8-10
2008				
2009				
2010				
2011	Young age group	45	29.04	3-7
2012				
2013				
2014				