How to cite this thesis / dissertation (APA referencing method):

Surname, Initial(s). (Date). *Title of doctoral thesis* (Doctoral thesis). Retrieved from http://scholar.ufs.ac.za/*rest of thesis URL on KovsieScholar*

Surname, Initial(s). (Date). *Title of master's dissertation* (Master's dissertation). Retrieved from http://scholar.ufs.ac.za/*rest of thesis URL on KovsieScholar*

Genetic diversity and performance trait analysis of the SA Boerperd

N. Breytenbach

Dissertation submitted in fulfilment of the requirements for the degree of

Magister Scientiae

in the Faculty of Natural and Agricultural Sciences,

Department of Genetics,

University of the Free State.

Supervisor:

H. Bindeman

Co-supervisor:

Prof. J.P. Grobler

2018

Genetic diversity and performance trait analysis of the SA Boerperd

(SA Boerperd Breeders Society)

Nadia Breytenbach

Department of Genetics

University of the Free State

2018

DECLARATION

I, Nadia Breytenbach, declare that the Master's Degree research dissertation that I herewith submit for the Master's Degree qualification in Genetics at the University of the Free State is my independent work, and that I have not previously submitted it for a qualification at another institution of higher education.

Nadia Breytenbach

2018

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to everyone who made the completion of this dissertation a reality:

My supervisors, Mrs. H. Bindeman and Prof J.P. Grobler, for your support and guidance not only throughout the duration of this study, but during my entire academic career. Thank you for allowing me to pursue my passion for horses.

The personnel and students of the Genetics Department for their continual encouragement during my project.

The SA Boerperd Breeders Society for providing support and necessary information about the SA Boerperd. Special thanks to Mrs. E. Prinsloo for her continual assistance and patience, as well as Dr P. Grové for her enthusiastic conversations and aid in obtaining hair samples.

All the breeders and owners who allowed me to sample their horses. Thank you for also sharing your insights and knowledge about the breed.

Unistel Medical Laboratories (Dr. Marx and H. Theron) for providing the STR marker data of the registered horses.

L.H.P. van de Goor (Dr. Van Haeringen Laboratorium, The Netherlands) and Dr K. Ehlers (University of the Free State) permitting the use of their equine DNA data published in previous works.

Lastly, Laetitia and Naas Breytenbach, thank you for always believing in me. To the rest of my family and friends, thank you for your continual support and interest in my project.

THANK YOU!

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

Symbols	**Meaning**
°C	Degrees Celsius
%	Percentage
ΔK	DeltaK
®	Registered trademark
™	Trademark

Abbreviation	**Meaning**
µl	Microliter
A	Adenine
ABI	Applied Biosystems
a.k.a.	Also known as
AND	Andalusian
AP-1	*Activator protein-1 transcription factor complex*
APP	Appaloosa
ARA	Arabian
BIEC2	Broad Institute *Equus caballus* version 2.0
BLUP	Best linear unbiased prediction
bp	Base-pair(s)
C	Cytosine
CBP	Cape Boerperd

cm	Centimetre
DMRT3	*Double-sex and mab-3-related transcription factor 3* gene
DNA	Deoxyribonucleic acid
EBVs	Estimated breeding values
ECA	*Equus caballus* autosome / horse chromosome
et al.	*et alli*: and others
f	Frequency
F	Inbreeding coefficient
F	Female
FEI	Fédération Equestre Internationale
F_{IS}	Population inbreeding coefficient
F_{IT}	Global inbreeding coefficient
FRI	Friesian
F_{ST}	Genetic differentiation
G	Guanine
g	Gravitational force
GBVs	Genomic estimated breeding values
H_2O	Water
HAC	Hackney
H_e	Expected heterozygosity
hh	Hands high
H_o	Observed heterozygosity
HMGA2	*High mobility group AT-hook 2* gene

HWE	Hardy-Weinberg equilibrium
H_Z	Unbiased heterozygosity
ICE	Icelandic Horse
ISAG	International Society for Animal Genetics
K	Number of clusters
kb	Kilo base-pair(s)
kg	kilogram
LASP1	*LIM and SH3 protein 1* gene
LCORL	*Ligand dependent nuclear receptor corepressor-like* gene
LD	Linkage disequilibrium
M	Male
MCMC	Markov Chain Monte Carlo
mg	Milligram
$MgCl_2$	Magnesium chloride
ml	Millilitre
mm	Millimetre
MSTN	Myostatin
mRNA	Mitochondrial ribonucleic acid
mtDNA	Mitochondrial deoxyribonucleic acid
n	Number of chromosomes
N	Sample size
N_a	Mean number of alleles
NA	Not applicable

N_e	Effective number of alleles
PA	Private alleles
PCR	Polymerase chain reaction
P_F	Forward primer
P_R	Reverse primer
QTL	Quantitative trait loci
rpm	Revolutions per minute
SA	South African
SAB	South African Boerperd
SD	Standard deviation
SNPs	Single nucleotide polymorphisms
ST	Stud
STA	Standardbred
STRs	Short tandem repeat markers
T	Thymine
TEN	Tennessee Walker
TFIID	*Transcription factor IID* gene
THO	Thoroughbred
UFS	University of the Free State
UML	Unistel Medical Laboratories
v	Version
ZFAT	*Zinc finger and AT hook domain containing* gene

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction to the equine industry

(Photos by Nadia Breytenbach)

1. Introduction

For millennia humans have been fascinated by the speed and grace by which horses move, thus it's not surprising that humans quickly domesticated these animals to utilize their mobility and endurance. The horse responded well to this partnership thanks to its herd mentality and acceptance of a hierarchy (Clark, 2011). Horses are capable of foraging for their own food, which make them easy to care for. They can also adapt to harsh environments and extreme temperatures (Brinkmann, Gerken & Riek, 2012).

Presently it is still unknown exactly when in history and at what location the domestication process of the horse took place, but it is estimated to have begun somewhere between 5 000 – 6 000 years ago in the Eurasian Steppe (Ludwig *et al.,* 2009; Outram *et al*., 2009; Lippold *et al.,* 2011). It is believed that the people of the Botai culture in this region used horses for harnessing, as well as for a source of milk and meat (Outram *et al.,* 2009). Other nomadic societies also made use of horses to help them expand their territories (Schubert *et al.,* 2014). The horse played a vital role in warfare, and consequently became woven into folklore where it came to represent death and rebirth (Kelekna, 2009). The utilization of horses in human society also facilitated the improvement of agriculture (Petersen *et al.,* 2013a; Petersen *et al.,* 2013b), and quickened the trade of both goods and information (Schubert *et al*., 2014). Finally, they assisted in the mining of metals (Kelekna, 2009), provided status to riders (Swart, 2010), and greatly influenced the development of different cultures and languages (Kelekna, 2009).

1.1. Equine breeding

Ever since the domestication of the horse, humans have primarily been responsible for the selection of equine reproductive partners. This selection has been aimed at choosing individuals displaying certain morphologies that can satisfy the needs of humans and aid in their advancement (Brooks *et al.,* 2010a). In the last few hundred years these needs have greatly changed to where the horse is now mainly used for sport and leisure purposes (Clark, 2011) in many parts of the world. Breeding has consequently become focused on improving and preserving certain traits that attribute to the appearance and performance of horses (Petersen *et al.,* 2013b). These two aspects encompass many traits that have their own array of variations; thus both artificial and natural selection have resulted in the formation of many

different breeds of horses. Today around 780 breeds exist, with a global horse population of more than 58 million individuals (FAOSTAT, 2014). Individuals belonging to a certain breed all display the same distinctive phenotype or fixed set of characteristics (Swart, 2010).

1.1.1. Traits under selection

Human preference has greatly influenced the aesthetics of the horse, especially since they have been associated with aristocracy and wealth for many years (Swart, 2010). For some time, work-horses, usually those with large heads and convex (Roman) noses, were regarded as inferior to riding horses that possess slender, dished heads, such as the Thoroughbred (Landry, 2008). This is thought to be caused by man's favour of animals displaying youthful characteristics like long legs and a small head-to-body ratio (Goodwin, Levine & McGreevy, 2008). This resulted in most equine breeds being of a 'light' classification, where their broad backs and slender build make them ideal for riding (Clark, 2011).

Equine coat colour has also been influenced by human preference, since prior to domestication horses mainly displayed the primitive coat colours of bay, black and dun (Ludwig *et al*., 2009; Imsland *et al.,* 2016). These basic colours most likely aided in camouflage and thermoregulation of ancestral horse populations (Protas & Patel, 2008). It was found that during the Bronze Age, coat colour variation in horses quickly increased due to human selection (Ludwig *et al*., 2009). The occurrence of interesting and rare coat colours or patterns contribute to the aesthetics of a horse, which greatly influences its financial value (Koenen, Aldridge & Philipsson, 2004). Coat colours are thus still being selected for today. In some instances, this selection has been implemented so intensely that it now defines certain breeds, such as the Norwegian Fjord, which solely displays the dun dilution colouration (Norwegian Fjord Horse Registry, 2016).

Speed and endurance have always been the trademark abilities of the horse. The acuteness of these abilities determines a horse's racing performance, and since some breeds like the Arabian and Thoroughbred are specifically bred for racing purposes, these abilities are of great economic importance (Petersen *et al*., 2013b). Physical aspects such as conformation and musculature enable racing performance, thus these features are focussed on in many breeding programs (Wood & Jackson, 2004). Racing breeds are bred to have lean muscles and

a fine-boned physique, whilst breeds bred for heavy draft work have a notably thick build to lend them strength and stamina (Brooks *et al.,* 2010a).

An array of height variations are present in the horse largely due to human selection, thus this attribute is an important standard for estimating the different breeds (Metzger *et al.,* 2013). The modern horse can range anything between 70 – 200 cm in height (Frischknecht *et al.,* 2015), but according to the Fédération Equestre Internationale (FEI) veterinary regulations an unshod horse of 148 cm or smaller is defined as a pony (Metzger *et al.,* 2013). Breeds such as the American Miniature and Shetland are classified as ponies, while the Clydesdale and Shire are two of the largest horse breeds wherein individuals can exceed 2 metres (Brooks *et al.,* 2010a). Figure 1.1 demonstrates the size difference between the American Miniature and Shire.

Figure 1.1 Size difference between a Shire and American Miniature (Therapy Horses, 2016).

Being a means of transportation has been the main function of horses for centuries. Horse breeds of 'light' classification can be ridden over long distances, and those that have ease of movement can be used for activities such as cattle herding and pleasure riding. In the past, when riders had to be in the saddle for hours at a time, it was important that the riding horse had a comfortable gait (Promerová *et al.,* 2014). A gait is a certain style of movement where a horse's hooves hit the ground in a particular sequence and time pattern (Kristjansson *et al.,* 2014). All breeds can perform the three common gaits, namely *walk*, *trot* and *gallop*, but some can perform additional gaits. These additional ambling or 'specialised' gaits are considered very comfortable for riders, and horses displaying them are dubbed as being gaited. Some well-known gaited breeds include the Icelandic Horse (Figure 1.2), Standardbred Trotter (North America) and Paso Fino (Spain) (Promerová *et al.,* 2014).

Figure 1.2 Icelandic Horse displaying the ambling gait, tölt (Hrísdalur Hestar, 2016).

The behaviour and general disposition of the horse has also been influenced by its domestication. Many breeds have been shown to display a certain typical temperament and personality (Lloyd *et al*., 2008), each of which can greatly influence a horse's trainability. A stubborn or anxious horse for example can be very challenging to train, thus generally inquisitiveness and docility are some of the traits selected for. The temperament and emotional state of a horse can also consequently influence its performance ability (McBride & Mills, 2012).

After many years of selection, the different extant horse breeds can thus now each be recognised by specific abilities and phenotypes. Icelandic horses for instance have been bred to be small and display powerful four- and five-gaited action, whilst draught breeds like Percherons and Shires are known for their efficient pulling power. Ponies on the other hand have primarily been shaped to be small and manageable by children, as well as have an amiable temperament (Arnason & Van Vleck, 2000).

1.1.2. Breed studbooks

Most breeds are governed by a breed association or studbook, which encourages its breeders to make use of certain breeding practices to ensure that individuals within the breed all conform to a certain 'type' (Clark, 2011). A studbook can be either closed or open, where the former does not allow the addition of horses from another breed or of mixed-breeding, but the latter does. Regardless of the type of association, great care is taken in selecting potential breeding stock that will produce foals representing the specific breed (Thomas, 2009).

Defining the different horse breeds is challenging. Most studbooks have been established within the last hundred years and are still developing their breeding goals. For the most part, each breed association's regulations governing its breed focus mainly on performance, conformation and gait characteristics (Koenen, Aldridge & Philipsson, 2004). Such characteristics can include leg action, height, ease of movement, coat colour (Clark, 2011) and temperament (Thomas, 2009). Due to the long maturation period of the horse its defining characteristics take a long time to fully develop (Clark, 2011), usually manifesting only from 3 – 4 years of age (Louw, 2008).

To evaluate the different breed-specific characteristics, some breeding associations use a point-based system during the registration process, whereby these characteristics are

subjectively scrutinized by a panel of breed selectors and are given points for their correctness (Koenen & Aldridge, 2002; Louw, 2008). If the individual horse is found to conform to the general 'type' of the specific breed, it is allowed into the registry. Unfortunately, equine traits are in truth difficult to physically measure, especially those described as making a "noble correct and beautiful horse" (Koenen & Aldridge, 2002). This consequently makes the process one of low objectivity (Koenen, Aldridge & Philipsson, 2004). Despite this drawback, the selection method still ensures that individuals displaying desired traits are allowed into the registry and permitted to be used for future breeding, yet there is no guarantee that the desired traits will then be passed onto progeny.

The genetics of a horse has been known to influence its phenotype since the mid-twentieth century through investigations of certain diseases (Dimock, 1950), equine physiology (Mathai, Ohno & Beutler, 1966) and coat colour (Castle, 1948). These investigations also revealed how genes allow such traits to be inherited. Understanding the underlying genetic patterns of inheritance in horses is thus vital to predict whether traits will be passed onto offspring. Since equine breeding requires time, hard work, and financial resources (Thomas, 2009), the assurance that the correct parents are mated to produce a foal with the best breed-specific potential is indispensable.

1.2. The South African Boerperd

South Africa has a national equine population of about 310 000 animals (FAOSTAT, 2014), which includes various breeds both registered and unregistered. The Thoroughbred is synonymous with the country's racing industry (Thoroughbred Breeders' Association, 2016), whilst other breeds, such as the Appaloosa and Quarter Horse, primarily contribute to the sporting industry (SA Stud Book, 2016). Of all the breeds present in the country only three had their origin in South Africa, namely the Boerperd, Basotho and Nooitgedacht, and all three can trace their ancestry back to the first European colonisation of the southern point of Africa (Swart, 2010). These breeds have thus been shaped over more than 350 years, but only one, the Boerperd, is commonly associated with the country.

1.2.1. Breed history

Unlike most continents where wild horses already inhabited the land before settlers invaded, Africa had no indigenous horses when Jan van Riebeeck set foot on the shores of the Cape of

Good Hope in 1652 (Swart, 2010). This is primarily due to disease barriers such as trypanosomiasis and the well-known African horse sickness, which prevented horses from migrating to the southern parts of the continent. Van Riebeeck quickly requested that horses be sent to the Cape, and in 1653 four Javanese ponies (Barb-Arab crosses) arrived in South Africa (Louw, 2008).

By 1662 the country had a herd of 40, but both inbreeding and African horse sickness weakened the genetic integrity of the stock, causing birth defects. To curb this problem some Arabian stallions were imported from Persia (now Iran), along with a few more Javanese ponies (Swart, 2010). An unexpected shipwreck also brought Andalusian and Isabella horses to the Cape (Nel, 2014). This new blood enriched the Cape's stock, making them healthier and taller. In the late eighteenth century horse racing became increasingly popular in the Cape, and Thoroughbreds were used to incorporate speed into the Cape horse (Du Toit, 2010). Horses not bred for racing purposes were used for transport, and these hardy, disease resistant horses were referred to as the Cape pony/horse or Hantam breed. These horses also had a naturally strong constitution and moved at a quick, comfortable pace (Swart, 2010). During the Great Trek (1835 – 1846) these horses were primarily used for transport between farms, as well as to scout new trails for ox-drawn wagons. These trusted mounts were consequently dubbed the "Boerperd" (Louw, 2008). This directly translates as 'farmer's horse', and Figure 1.3 illustrates one of these horses. The durability of these horses soon became known internationally through their export to countries like Australia and India (Louw, 2008; Swart, 2010).

Figure 1.3 Transvaal Burgher astride his Boerperd during the Anglo Boer War (WordPress, 2014).

At the end of the nineteenth century the quality of the Boerperd drastically declined again due to another bout of inbreeding, as well as due to the devastation brought by the Anglo Boer War (1899 – 1902) (Swart, 2010). Despite the death of countless horses during the battle, many survived and were diversified by the addition of breeds imported during the war. Unfortunately, it is still uncertain exactly which breeds influenced the Boerperd, but it is

thought that draft horses such as Percherons and Spanish Barbs sent from England, Ireland and Canada, played a minor role (Swart, 2010). Flemish and Friesian horses are also thought to have had a small contribution (Louw, 2008). Lighter breeds of smaller conformation, such as Hackneys, Cleveland Bays, Norfolk Trotters (Nel, 2014), Morgans and other American stock (Swart, 2010), would also have been involved in this period of the development of the Boerperd.

In 1923 the country's horse numbers were at its peak (Swart, 2010), but in the following years the golden age of mechanisation caused equine numbers to fluctuate once again (Louw, 2008). In the 1940s the allure of horse showing made its way from America to South Africa, and brought with it the American Saddlebred. This delicately-built and aristocratic breed quickly replaced the work horse to provide leisure in the show ring, and became the reason for the creation of two distinct breeds of Boerperd.

The Cape Boerperd was consequently developed through the addition of American Saddlebred blood to possess height and high-stepping leg action (Swart, 2010), and the Historical Boerperd denied this addition to preserve the stamina and conformation that was so renowned during the Boer War (Nel, 2014). After the division of the breed, the South African/Historical Boerperd Breeders Society was established in 1973, and the Historical Boerperd was officially awarded breed status by the Department of Agriculture in 1980. In 1996 the breed's name was changed to the South African (SA) Boerperd, and is still referred to as such today (Louw, 2008; Nel, 2014).

Since the Second World War drastically declined the number of horses in the country, all SA Boerperd horses can trace their ancestry back to at least one of only six bloodlines that survived the war's devastation. These include the Middleton, Hancke, Odendaal, Steenkamp, A2 and Cloete/Eggo bloodlines (Du Toit, 2010). Other bloodlines, namely the Namib, Sephton, Van der Wath, Vlampies and Streicher (Louw, 2008), were also added over the course of the last 50 years. Since so many bloodlines helped shape the SA Boerperd, each with their own unique phenotypical characteristics, a clear consensus had to be reached to define the 'model' Boerperd (Louw, 2008; Du Toit, 2010). A breed standard was thus drafted by the Breeders Society, which clearly defines the appearance, carriage and temperament of horses that qualify as a SA Boerperd.

Figure 1.4 Example of a SA Boerperd. Burgerstrots Simon, owned by Michiel Burger (Photo by Nadia Breytenbach).

1.2.2. Breed standard

Figure 1.4 illustrates an example of a modern SA Boerperd. The breed is defined as being symmetrical and well-balanced with a strong constitution. Its coat should be lustrous and thick with fine hair, and common colours include variations of black, brown, chestnut, palomino and roan (Du Toit, 2010). Concerning movement, horses can either be classified as being traditional or universal, where for the former they need to display a high knee action and for the latter a lower knee action (SA Boerperd Breeders Society, 2014). Regardless in what category they are placed, all horses of this breed must be able to cover ground with long, springy steps and should be well balanced on all four legs (Louw, 2008). All horses have the potential to perform five gaits, namely *walk*, *trot*, *canter*, *short-gait* and *rack*, and they possess considerable stamina. The hooves of these horses have thick, strong walls and are of average size compared to other breeds. The minimum height of stallions is 14.2 hands high (hh), whilst that of mares is 13.3hh. These horses have a calm, dependable nature and they are subjectively eager to please (Du Toit, 2010).

These traits, as well as the ratios of different conformational areas, are evaluated by breed selectors when horses are registered (Louw, 2008; Du Toit, 2010). Considerable attention is given to the formation and shape of the legs, since some faults can negatively affect the movement of an individual. The conformation of the hoof is also considered, and any potentially harmful defects, such as blindness or swayback, that can threaten the health of the animal is strongly selected against. As soon as a horse qualifies as a SA Boerperd it is branded on the right hindquarter with the breed trademark, a 'B' enveloped by a horseshoe (Figure 1.5), to display the individual's status (SA Boerperd Breeders Society, 2014).

Figure 1.5 SA Boerperd brand marking (Photo by Nadia Breytenbach).

1.2.3. Importance of the SA Boerperd breed

Today a wide variety of sport types exist that cater to the specific performance abilities of different horse breeds. Most breeds only excel in one riding discipline, but due to the versatility of the SA Boerperd it can successfully perform in a wide variety of disciplines including jumping, western riding, endurance riding, gymkhana, performance shows, and even Spanish classical riding (Louw, 2008; Du Toit, 2010). Since this breed was originally developed for the farmers of the country, it also serves as an excellent working horse. Labour intensive works such as herding livestock over difficult terrain is easily managed with the aid of this breed's power and stamina (Nel, 2014). The SA Boerperd also greatly contributes to tourism, taking both local and international vacationers on horseback safaris to experience Africa's wildlife up close (Equus Horse Safaris, 2016; Pakamisa, 2016). This breed is also a great companion. Its soft and even-tempered nature makes it an ideal horse for riders of any age, from novice to experienced, to enjoy recreationally in the ring or out in the veld.

As previously mentioned, the same foundation stock that shaped the Boerperd also lead to the creation and recognition of the Basotho pony as a breed in the 1800's (Swart, 2010). This breed, unique to Lesotho, was undoubtedly also influenced by the early Boerperd at the time of the Boer War due to the raids and trading of these horses (Swart, 2010; Clark, 2011). The Basotho breed is a source of pride to the nation's people, seeing as the horse is still their main means of transport (Swart, 2010). These horses also function as farm animals, ploughing and cultivating fields, and they are also used for trekking by tourists (Lekota, 2001). The Boerperd also lead to the establishment of the Nooitgedachter breed in the 1950's, which is a cross between Boerperd, Arabian and Basotho bloodlines (The Nooitgedachter Horse, 2018).

The extant SA Boerperd is also revered for its extraordinary fluid movement and unique gaits, thus it is not surprising that the breed is currently being used in a performance analysis study (Whitehead & Mansfield, 2013), whereby angular and linear conformation traits are being measured to see how they affect the movement of the breed. Furthermore, future studies focusing on the underlying genetic workings of this breed's gaits can help identify relevant genotypes, which in turn can greatly simplify the mate selection process when a foal with the potential to display alternate gaits is desired.

1.3. Equine performance

Today most horses kept for recreational purposes are bred to be athletes, thus their value greatly depends on how well they can perform at competitions. Equestrian sport consists of two divisions, namely racing and equestrian events (European Horse Network, 2012). The latter is divided further into numerous performance disciplines, each designed to test the athletic ability of a horse, as well as to display their aptness for certain kinds of work. This is done by expecting horses to execute specific tasks or routines that are representative of the horse's everyday use (Wood & Jackson, 2004).

Different breeds are used for different events, seeing as breeds with certain conformations and abilities are more adept for a specific form of work. For example, the long back of the American Standardbred allows it to cover ground quickly, which makes it specifically well-suited for the racetrack (Allfrey, 1980). The history, culture and environment of a breed also determines what type of task it must be able to execute. American breeds are mainly used for cattle ranching, thus sport events like cutting (The Horse, 2001) and reining (US Equestrian, 2018) developed to showcase the skills needed by these ranch horses. The rolling English countryside on the other hand, has led to the development of sport types such as cross-country and showjumping, where millions of people in the United Kingdom still annually attend the latter event (The British Horse Society, 2015).

South Africa also hosts a wide variety of equestrian sport disciplines. As many as fifteen different sport types are officially recognised in the country, and many are of both American and English influence (South African Equestrian Federation, 2016). Unlike most breeds, the SA Boerperd has proven itself to be highly adaptable and able to excel in many of these disciplines due to its physique. Both the national and regional SA Boerperd shows host many performance events that can include dressage, harness, jumping, working hunter, show-in-hand, and both three- and five-gaited classes (Louw, 2008). The many strengths presented by this breed leaves it up to the discretion of the breeders and trainers to decide for which discipline an individual will be used. In many cases this decision is driven by the specific performance-related traits displayed by the individual horse.

1.3.1. Performance traits in horses

The ability of a horse to perform is influenced by a combination of its physiology, conformation and disposition. A horse's function is largely determined by its conformation, or body shape, seeing as it defines the limits of a horse's movement and consequently its performance ability (Sánchez *et al.,* 2013). Characteristics like body proportions (e.g. chest width and height at croup), leg rotation and hoof structure greatly affect how freely a horse can move (Wood & Jackson, 2004). Smooth, long and deep tying muscles are also desired across a horse's entire body, and good health, evident by good quality skin and coat, ensures that a horse is in a condition to perform. A sound horse, which has no structural weakness or injuries that can interfere with its usefulness, is thus sought after so that it will be able to perform consistently (Thomas, 2009). Two similarly important traits that have been extensively selected for throughout the ages and shown to influence performance, are height and gaitedness (Petersen *et al.,* 2013b). As indicated by the breed standard of the SA Boerperd (Du Toit, 2010), these two traits play a significant role in the breed and would be helpful to keep in mind during mate selection.

1.3.1.1. Height in horses

The unit traditionally used to measure a horse is the hand, one of which is equivalent to four inches or 10.16 cm (Farmer's Weekly, 2015). To measure the full height of a horse, a measuring stick or tape is placed behind one of its forelegs and used to estimate the distance from the ground to the top of its withers (Figure 1.6). If, for example, a horse is found to be 15.3hh, then the horse would be approximately 160 cm tall. Horses and ponies that compete in FEI Championships or events must be re-measured by selected measuring veterinarians (Euro Dressage, 2016), but generally the studbook selectors of most breeds measure horses during the registration process.

Figure 1.6 Position for measuring equine height at withers. Goudhoek Ster, owned by Fritz Oosthuizen (Photo by Nadia Breytenbach).

Apart from being a means of breed classification, the height of a horse influences its performance by affecting its stride length (Baban *et al.,* 2009). Tall horses give long strides, which means they can cover ground much quicker than smaller breeds even though the strides per second of small horses (ponies) are more compared to that of bigger breeds. The height of a horse also influences its weight, which in turn also greatly affects its performance. The small American Miniature can weigh less than 133kg, whilst the large draft breeds can exceed even 907kg (Petersen *et al.,* 2013b). Draft breeds are therefore very bulky and slow-moving, and wouldn't be ideal to compete in high-speed events. Yet a significant weight, and therefore height, is needed for speed events such as jumping. The force exerted by a horse to get itself airborne from the ground needs to be greater than its bodyweight (Clayton & Barlow, 1991), thus a medium weight horse will be able to generate a lot of force and momentum to drive itself upward for a longer distance compared to a smaller horse.

Another aspect that can influence performance is a horse's trotting ability. The latter can be influenced by the height of the croup, which in turn impacts the extension and retraction angles of the hind legs (Sánchez *et al.,* 2013). These angles influences a horse's ease of movement. Height at croup and height at withers have been found to be strongly linked to one another (Saastamoinen & Barrey, 2000), which suggests a horse with a large height at croup measurement will also be tall measured at the withers. Ensuring that a horse is of an ideal height can thus indirectly improve the movement of its hind limbs and so its usefulness for a specific task. The SA Boerperd can range anywhere between 13.3 – 16hh (Du Toit, 2010), making it a medium-sized breed with an average stride length, weight, and thus agility that aid its versatility for different disciplines.

1.3.1.2. Gait in horses

There are primarily four categories of gaits, namely regular rhythm ambling, diagonal ambling, lateral ambling, and pace (Andersson *et al*., 2012). Each category differs in tempo, footfall pattern and timing, and each encompasses many different types of gaits. Gaits can further be divided as being either symmetrical or asymmetrical (Robilliard, Pfau & Wilson*,* 2007). The former can include the *walk, trot*, *pace*, *short-gait* and *rack*, whilst the latter comprises mainly of the *canter* and *gallop*. The *pace* is a two-beat gait where the left fore- and hind leg strike the ground simultaneously, as do the legs on the right (Kidd, 1981). The stepping pace, or *short-gait*, differs from the *pace* by letting the front legs hit the ground just before the hind

Figure 1.7 SA Boerperd exhibiting the *rack* gait (SA Boerperd Breeders Society, 2016).

ones, making it a four-beat gait. The *rack* is similar to the *short-gait*, except that the interval before each footfall is a bit longer. Figure 1.7 illustrates a SA Boerperd performing the *rack*. Breeds that solely display the common gaits (*walk, trot* and *gallop*) are referred to as being three-gaited, whilst breeds with the potential for either one or two additional gaits are considered four and five-gaited respectively (Robilliard, Pfau & Wilson, 2007). The SA Boerperd is thus five-gaited.

Some gaits are also associated with a certain breed, though even within a breed it can happen that some individuals perform a certain gait naturally, whilst others cannot be trained to display the gait (Promerová *et al.,* 2014). This phenomenon usually occurs when a certain gait has not fully been selected for in a breed. Some individuals of the breed are thus only used for disciplines where gaitedness is not necessary, such as jumping or racing. Every individual of the Tennessee Walking Horse and Puerto Rican Paso Fino display alternate gaits, due to the intense selection for such gaits. In contrast to this, not all the SA Boerperd horses have the potential to exhibit the *short-gait* and *rack* (Bekker, 2012). The horses that can move in these specialized styles usually have excellent rhythm and balance, and most importantly receive intensive training to make these gaits come as second nature to them. Variables such as the environment, training method, fitness, rider and genetic characteristics greatly affect how quickly, and efficiently, a horse can display these gaits.

1.3.2. Genetic analysis of performance traits

Some studbooks make use of estimated breeding values (EBVs) to plan future matings and predict the possible phenotype of progeny (Arnason & Van Vleck, 2000). The EBVs that concern sport traits are often based on scores obtained from studbook recordings and performance tests, whilst those concerning conformation and movement utilise data collected during registration of certain sets of linear body measurements (Royal Dutch Sport Horse, 2016). Such measurements include body length, girth width and height at withers.

Breeding values are generally evaluated by a best linear unbiased prediction (BLUP) multiple trait animal model (Arnason & Van Vleck, 2000). The BLUP method assumes genetic parameters such as heritability and environmental correlations, and thus the breeding values it predicts are only approximations. It also does not provide information about performance-factors such as muscle strength and gaitedness. Moreover, this method is very complicated to utilise and is largely influenced by certain variables such as the age of the horse, and the EBVs of its parents and offspring (Royal Dutch Sport Horse, 2016). The use of this method is therefore not common practice for all breed associations.

With the advancement in genomics it is now much simpler to screen for the influencing genes and single nucleotide polymorphisms (SNPs) associated with performance traits. Due to the complexity of the equine genetic architecture, it is unfortunately not always so simple to determine what genotype causes a certain phenotype. Most traits and abilities are influenced by multiple genes and gene variants (Allendorf, Luikart & Aitken, 2013), and identification of these quantitative trait loci (QTL) are extremely difficult.

Genomic regions most often focused on are those shared between breeds of similar phenotype, and these are situated near genes with suspected or known functional effect (Petersen *et al.,* 2013b). Even if these regions (candidate genes) prove to be linked to traits significantly, the degree of association between genetic markers, or the linkage disequilibrium (LD), can differ between breeds (Finno & Bannasch, 2014). SNP detection can also be complicated when causative alleles are in low-frequency (Slatkin, 2008). Another factor to consider is that quantitative traits can also be influenced by environmental factors (Allendorf, Luikart & Aitken, 2013). Regarding horses, this could include aspects like diet, stabling conditions and the training regime of an individual.

Despite the abovementioned difficulties, the clinical mutations associated with 35 equine Mendelian traits and diseases have successfully been found, and 13 are currently being investigated (Finno & Bannasch, 2014). Such information can greatly aid mate selection for breeders, especially since the first fourteen ancestors, up until the great-grandparents, can influence the inherited traits of a horse (Thomas, 2009).

Most genetic studies concerning performance traits and their heritability have been investigated in Thoroughbreds (Gu *et al.,* 2009; Hill *et al.,* 2010a; Hill *et al.,* 2010b; Tozaki *et*

al., 2010) and Warmblood breeds (Stock & Distl, 2008; Viklund *et al.,* 2010; Schroder *et al.,* 2012). Investigations with regards to the latter sport breeds, have primarily been focused on conformation, limb health and showjumping ability (Stock & Distl, 2008). For racehorses, emphasis has been set on speed-affecting factors such as muscle strength and insulin signalling (Gu *et al.,* 2009). In recent years studies investigating traits such as height (Makvandi-Nejad *et al.,* 2012; Signer-Hasler *et al.,* 2012) and gait (Andersson *et al.,* 2012; Kristjansson *et al.,* 2014; Promerová *et al.,* 2014) have also been conducted, yet these haven't extensively been explored in many breeds. Nevertheless, these studies have shed light on the underlying genetics of these two traits, and have also paved the way for future research.

1.3.2.1. Genetics of height in horses

Signer-Hasler *et al.* (2012) found eight SNPs associated with height and determined that the trait's heritability is 72%. Two SNPs were found within a QTL region on horse chromosome (ECA) 3, and six were found on ECA 9. The first region is located 100 kb (kilo base-pairs) upstream of the *ligand dependent nuclear receptor corepressor-like* (*LCORL)* gene and the second is close to the *zinc finger and AT hook domain containing (ZFAT)* gene. Both these genes have large intergenic regions, and the ZFAT protein has been shown to be important for development during haematopoiesis (Tsunoda *et al.,* 2010). One SNP of each gene region was found to significantly affect height at the withers (Signer-Hasler *et al.,* 2012).

The alleles of SNP BIEC2_808543 (ECA3: 105 547 002) on *LCORL* can be either cytosine (C) or thymine (T), where the presence of the C-allele influences height with approximately 1 cm (Signer-Hasler *et al.,* 2012). This SNP is present in a TATA-box transcription factor binding site of the *transcription factor IID (TFIID)* complex that impacts mRNA transcription (Orphanides, Lagrange & Reinberg, 1996). The presence of the C-allele mutation removes this binding site by making it unrecognisable to the core promotor elements (Metzger *et al*., 2013). Ultimately this affects the expression of the *activator protein-1 transcription factor complex* (AP-1), which is central to bone cell development (Yang *et al.,* 2011). Unfortunately, the exact mechanism by which this mutation effects *LCORL* is still unknown.

The influencing SNP present on *ZFAT*, namely BIEC2_1105377 (ECA9: 74 798 143), can have either an adenine (A) or guanine (G) nucleotide, and increases height by 0.5 cm with the presence of one A-allele (Signer-Hasler *et al.,* 2012). Together the two minor alleles, C and A,

of these two SNPs can contribute 3 cm to a horse's height. These two QTL areas were calculated to ultimately explain 18.2% of the height at withers variation, and the remaining variance is thought to be influenced by multiple genes with small effects. In contrast to this, Makvandi-Nejad *et al.* (2012) believe that horse size is most likely affected by a few loci with large effects, since four loci on ECA 3, 6, 9 and 11 were found to explain 83% of size variation in 16 different breeds.

The loci investigated for ECA 3 and 9 by Makvandi-Nejad *et al.* (2012) are the same as that of Signer-Hasler *et al.* (2012), but significantly associated SNPs were also found on the *high mobility group AT-hook 2 (HMGA2)* gene on ECA 6. This gene is a transcription factor that controls gene expression, thus determining cell production, growth and differentiation (Cleynen & Van de Ven, 2008). One specific genotype within this gene, consisting of 9 SNPs, was shown to be very common in large breeds (Makvandi-Nejad *et al.*, 2012). Additionally, the variant c.83G>A was also shown to affect DNA binding of the HMGA2 protein and cause growth impairment in pony breeds (Frischknecht *et al.,* 2015). Other relevant SNPs were found in the first intron of the *LIM and SH3 protein 1 (LASP1)* gene on ECA 11, which facilitates cell survival and migration (Makvandi-Nejad *et al.*, 2012). Its locus is unfortunately very gene-dense, which makes the identification of causal size-related SNPs challenging.

Despite these difficulties, both previously discussed studies, as well as one conducted by He *et al.* (2015) on size in Yili horses, proved that ECA3: 105 547 002 on the *LCORL* gene significantly contributes to size. This has been found true for many horse breeds, and therefore this SNP is a very promising candidate marker for horse selection. An across-breed analysis of this gene revealed that the T-allele is significantly associated with small breeds (Metzger *et al*., 2013). Ponies and medium sized horses within the 130 – 160 cm size range are mainly of the TT-genotype, whilst larger breeds chiefly display the CC-genotype.

Signer-Hasler *et al.* (2012) adds that, excluding height, the region on ECA 3 is also linked to correctness of gait, leg conformation, mandible formation and head shape. Similarly, the region on ECA 9 also influences back and croup lengths. As previously discussed, these traits have been proven to have a positive genetic correlation with height at withers and can influence a horse's performance. Additionally, studies investigating the influence of the myostatin *(MSTN)* gene on the ratio of mass to height at withers in Thoroughbred horses,

revealed that sprinters possess a greater muscle mass and are normally shorter than the average race horse (Hill *et al.,* 2010a; Hill *et al.,* 2010b).

1.3.2.2. Genetics of gaitedness in horses

The ability of horses to move through different gaits has been shown to be influenced by a nonsense mutation in the *double-sex and mab-3-related transcription factor 3 (DMRT3)* gene on ECA 23, namely *DMRT3_Ser301STOP* (ECA23: 22 999 655) (Andersson *et al.,* 2012). The *DMRT3* gene is one of three genes that encode for different isoforms of the specialised DMRT transcription factor, and these genes consist of a dsx and mab-3 DNA-binding domain. When the mutation occurs, it leads to a premature stop at codon 301 and shortens the DMRT3 protein with 174 amino acid residues. The mutation causes an C to A substitution, and is more commonly known as the *gait keeper* mutation due to its strong influence on the pattern of equine limb movement (Andersson *et al.,* 2012; Promerová *et al.*, 2014). The genome-wide association study done by Andersson *et al.* (2012) also found that a linked SNP, BIEC2_620109 (ECA23: 22 967 656), located 32 kb upstream of the *gait keeper* mutation significantly effects the ability of horses to *pace*. This mutation causes an C to T change.

The *gait keeper* mutation allows horses to perform lateral gaits such as the *rack*, *pace* and *tölt*, where the latter is a fast four-beat gait unique to Icelandic horses (Kristjansson *et al.,* 2014). For horses to possess the potential to perform these alternative gaits they often must be homozygous for the mutation. The frequency of the mutated allele was found to be 100% in six gaited breeds (Andersson *et al.,* 2012). For eight non-gaited breeds, including the Przewalski's horse, the frequency was zero. In addition, it's believed that gaitedness is a derived trait in the modern horse, seeing as its wild relative, the Przewalski's Horse, does not possess the *DMRT3* mutation (Orlando *et al.,* 2013).

It was also noted in harness-racing Standardbreds that the *gait keeper* mutation increases the speed capability of the *trot* and prevents horses from making the alteration from *trot* to *gallop* (Andersson *et al.,* 2012). It thus seems that the mutation favours symmetrical gaits like *pace* and *trot* at high speeds, rather than asymmetric movement patterns (Kristjansson *et al.,* 2014; Promerová *et al.*, 2014). Further studies using mice as model organism, revealed that *DMRT3* is needed to enable the normal development of a coordinated locomotor network in the spinal cord (Andersson *et al.,* 2012). This shows that the g*ait keeper* mutation affects the

pattern of equine locomotion and overall coordination. This is supported by the findings of Kristjansson *et al.* (2014) where a significant difference between four- and five-gaited horses was observed. Of the tested homozygote Icelandic horses (AA), 45% were reported four-gaited and 96.8% were five-gaited. The AA-genotype thus allows the lateral leg movement of horses, whilst hindering the synchronised movements of legs moving diagonally during the *walk*, *trot* or *gallop*. Also of note is that Promerová *et al.* (2014) found that in a few rare instances some Icelandic horses homozygous for the C-allele could *tölt*, though not *pace*. This could either be due to other influencing mutations in or around the coding region, or due to a polygenic effect caused by the intense selection of gaitedness in this breed.

The worldwide frequency distribution study done by Promerová *et al.* (2014) used more than 4 000 horses from 141 breeds, and it was determined that 68 of those breeds possessed the *gait keeper* mutation. The frequency of the mutant A-allele ranged between 1.1% (Spanish Pure breed) to complete fixation (100%) for gaited breeds of the Americas. Due to the close linkage of *DMRT3_Ser301STOP* and BIEC2_620109, genotype frequencies were also calculated by Promerová *et al.* (2014) and the results are represented in Table 1.1. The frequency of the wild-type SNPs (C & C) occurring together was found to be much higher than that of the two mutant forms (T & A). Horses possessing the latter genotype are generally gaited, but horses that have a combination of the mutant allele at BIEC2_620109 (T) and the reference allele at *DMRT3_Ser301STOP* (C) are non-gaited. The frequency of this heterozygous genotype was shown to be only 2% (Promerová *et al.,* 2014).

Table 1.1 Estimated genotype frequencies for the *DMRT3_Ser301STOP* mutation (C>A) and the closely linked SNP BIEC2_620109 (C>T) in a total of 2,749 horses across different breeds (Promerová *et al.,* 2014).

Genotypes	BIEC2_ 620109	*DMRT3_ Ser301STOP*	*n*	*f*
Wild-type at both loci	C	C	3,237	0.589
Mutant 620109 and wild-type *DMRT3*	T	C	109	0.020
Mutant 620109 and mutant *DMRT3*	T	A	2,133	0.388
Wild-type 620109 and mutant *DMRT3*	C	A	19	0.003

n = number of chromosomes; *f* = frequency

It is assumed that the *DMRT3* nonsense mutation occurred in the presence of the mutant T-allele (BIEC2_620109) and in so doing has spread in multiple breeds across the world. Some non-gaited breeds, such as the New Forest and Welsh Ponies, contain individuals that possess

the *gait keeper* mutation at a very low frequency. The presence of the mutation in these breeds could be due to past cross-breeding with gaited breeds, or the mutation could have occurred in ancestral lines (Promerová *et al.,* 2014).

1.3.3. SA Boerperd trait investigations

Before the current study, no genetic study has been conducted to investigate height in the SA Boerperd, but the breed was included in the study of Promerová *et al.* (2014) to detect the presence of the *gait keeper* mutation. In the latter study the frequency of the A-allele for the *DMRT3* gene in the breed was found to be 15%, though only 20 individuals were tested. It was also found that none of the horses possessed the homozygous affected genotype. An additional study with a larger sample size could thus give a more accurate estimation of the presence of the *gait keeper* mutation in the SA Boerperd, as well as that of the other associated gait-mutation and size-related variants. Also of importance would be to investigate the overall genetic diversity of the breed, seeing as the selection for certain traits and their underlying mutations directly influence the rest of the equine genome.

1.4. Equine genetic diversity

For a species to progress both evolutionary and adaptively, adequate genetic diversity is generally needed. This diversity includes a variation of alleles, and consequently genotypes (Frankham *et al.,* 2002), which is governed by forces such as genetic drift, mutations, adaptations, and of course selective breeding (Groeneveld *et al.,* 2010). To this the horse is no exception. It is believed that early equine domestication included the continual genetic exchange between domestic and wild horses (Warmuth *et al.,* 2012), which would have caused domestic stock to retain a large amount of genetic diversity. This is confirmed by the observation that modern breeds were proven to possess high levels of diversity (Ludwig *et al.,* 2009; Lippold *et al.,* 2011; Achilli *et al.,* 2012). This is probably also due to the historic movement of the horse across the continents and its continued gene flow as mediated by humans (Petersen *et al*., 2013a).

Intense selection by humans have also caused the different breeds to have their own unique phenotypic and genetic homogeneity (Petersen *et al*., 2013a). As similar as individuals are within a certain breed, there can often be significant variation among the different breeds (Petersen *et al.,* 2013b), as well as between individuals of the same breed. This genetic

variation is believed to be affected by factors such as the status of a studbook (open/closed), the selective pressures of the breeding process, how much time has passed since breed establishment, and the diversity of the founder stock (Petersen *et al*., 2013a).

Equine breeds can genetically be placed into three different categories, namely landrace, native and feral (McClory & Kowalski, 2014). Those considered landrace receive little human interference and possess a high genetic diversity that allow them to survive under environmental pressures (McCue *et al.,* 2012), whilst native breeds are defined as those that have their origin in a region, such as the SA Boerperd. Lastly, feral breeds are those whose pedigree lead back to domestic horses that were either released or escaped at some point and successfully persist in the wild (Csurhes, Paroz & Markula, 2009). It was found that the highest within breed diversity are primarily in breeds that have recently been established, have a large population size, allow admixture with other populations/breeds, and are landrace (Petersen *et al.,* 2013a).

Since humans are actively involved in the reproduction of most horses, few landrace breeds exist today, which means that the genetic diversity of the horse is mainly controlled by humans. It is thus the responsibility of equine breeders to ensure that sustainable breeding practices are being implemented, lest the fitness of the horse decline. Genetic diversity is thus needed to ensure the continuation and adaptive potential of future generations (Groeneveld *et al.,* 2010). Conserving this diversity can be of great economic importance, and its continued existence can ensure opportunities for future selection and breed development.

1.4.1. Factors influencing loss of diversity

As previously mentioned, some breed associations do allow the addition of novel breeds if they meet certain requirements, but most studbooks are closed and forbid admixture to ensure the purity of their breed (Clark, 2011). Such studbooks often make use of inbreeding, where closely related individuals such as half-siblings or cousins are mated, or line-breeding, the mating of an individual with its descendants, to increase the chance of progeny displaying desired traits (Thomas, 2009). Selection pressures such as these often cause the genetic diversity of a breed to skew towards homogeneity through the fixation of an increasing number of loci (Lacy, 1987). This event greatly affects a breed's ability to persist after a

bottleneck event, such as caused by diseases. Environmental pressures can also decrease their fertility and adaptability (Juras, Cothran & Klimas, 2003).

This desire to 'breed true' for certain traits has also unintentionally led to the increased occurrence of detrimental alleles (McCue *et al.,* 2012), which can lead to the expression of genetic diseases such as severe combined immunodeficiency (Perryman, 2004) and lethal white foal syndrome (Brooks *et al*., 2010b). This increase of harmful alleles, as well as the possible decrease of beneficial heterozygous combinations, is known as inbreeding depression (Charlesworth & Charlesworth, 1999). Breeding programs should thus limit inbreeding, as well as ensure a large effective population size (Groeneveld *et al.,* 2010) and attempt to maintain high levels of genetic variation that will ensure the long-term survival of a breed (Bijma *et al.,* 2001). It is thus advised that closed studbooks must encourage stud owners to select mating partners in such a way that it will allow the exchange of genetic material between the different stud populations.

1.4.2. Measuring equine genetic diversity

The advancement of genetic technology and methodology has made it possible to investigate the underlying genetic characteristics of different species and breeds to determine aspects like the genetic variability of a population, as well as the phylogenetic relationships and co-ancestry between populations (Allendorf, Luikart & Aitken, 2013). Nowadays, DNA-based polymorphisms are the molecular markers of choice to investigate these genetic variations. Technologies such as the polymerase chain reaction (PCR) and sequencing have greatly simplified the screening process for such variations (Hanotte & Jianlin, 2005).

To study maternal inheritance, the *D-loop* and *cytochrome B* regions of the mitochondrial DNA (mtDNA) are usually examined, whilst Y-chromosome specific SNPs and microsatellites, also known as short tandem repeat markers (STRs), are used to study the paternal lineage (Avise, 2004). Also, bi-parental inheritance can be studied using autosomal microsatellites. Genetic diversity can further be measured and characterized through the sequencing of autosomal microsatellites, as well as through Y-chromosomal and mitochondrial genotyping (Rothschild, 2003).

Both STRs and SNPs have been found to function equally well in analysing population genetics (Coates *et al.,* 2009), but the common choice for determining equine breed diversity for years

has been STR markers. In addition to the use of molecular markers, other information such as knowledge of a breed's population size and genetic structure, its environment and geographical distribution, as well as its within- and between breed genetic diversity, is also needed to effectively manage genetic resources (Groeneveld *et al.,* 2010).

1.4.2.1. Microsatellite markers

Microsatellites or STRs are areas of the genome which display a repeat pattern of 1 – 6 bp (Tautz, 1993). These markers are generally highly polymorphic due to a high mutation rate, which in turn causes great levels of heterozygosity (Schlötterer & Tautz, 1992). Microsatellites are also present in great numbers within the nuclear genome of animals. These markers are neutral, meaning that they are mostly present in non-coding DNA regions generally not subject to natural selection (Ellegren, 2000).

By utilizing these markers it is possible to identify an individual, confirm parentage, determine the genetic structure of a population, and estimate the phylogenetic relationships among populations (Goldstein & Schlötterer, 1999; Avise, 2004). This is achieved by comparing the variation in allele frequencies between populations (Avise, 2004). The markers can also be used to determine inbreeding depression, seeing as they are extremely sensitive to genetic bottlenecks (Sunnucks, 2000). Microsatellites have contributed to analysing equine breed diversity (Marletta *et al.,* 2006; Leroy *et al.,* 2009; Khanshour *et al*., 2015), and have also investigated the possible origin of certain horse breeds (Groeneveld *et al.,* 2010). Additionally, these markers are a vital tool in forensic cases that involve horse theft, identity forgery (Van de Goor & van Haeringen, 2007), and doping (Chen *et al.,* 2014).

The International Society for Animal Genetics (ISAG) first proposed 9 equine STR markers to be investigated for routine kinship analyses and other comparisons during the 1990's, and in 2011 these were increased to twelve (Van de Goor, van Haeringen & Lenstra, 2011). To further improve genotyping, five additional loci were added, thus the internationally recommended set now consists of 17 specific STR markers (AHT4, AHT5, ASB2, ASB17, ASB23, CA425, HMS1, HMS2, HMS3, HMS6, HMS7, HTG4, HTG6, HTG7, HTG10, LEX3 and VHL20) (Van de Goor, Panneman & van Haeringen, 2009). This extensive panel of markers have been shown to strengthen the discriminating power of genetic studies compared to the panel consisting of fewer loci (Van de Goor, van Haeringen & Lenstra, 2011). This panel can also be

used to construct phylogenetic trees of the different equine breeds to study their relationships and ancestry.

1.4.2.2. Single nucleotide polymorphisms

Single nucleotide polymorphisms, better known as SNPs, are mutations that alter the DNA sequence of an individual by causing a single base change (Finno & Bannasch, 2014), thus creating the possibility that one of two nucleotides can occupy a certain position. Such a sequence alternative is generally only considered a SNP when the rarest allele has a frequency of 1% and higher (Vignal *et al.,* 2002). When the DNA sequence changes it also alters the amino acid chain, which in turn can affect protein expression and function (Finno & Bannasch, 2014). In some cases the phenotypic change caused by mutations can be neutral or harmless, but in others it can lead to serious illnesses like equine uveitis (Kulbrock *et al.,* 2013) and foal immunodeficiency syndrome (Fox-Clipsham *et al.,* 2011).

Unlike STR markers that are subject to small deletions or insertions that can alter the repeat number of a motif and complicate allele calling, the nomenclature of SNPs is based simply on whether one or two copies of the wild-type allele is present (Vignal *et al.,* 2002). This method of allele scoring thus simplifies and excels the rate of data analysis. The genetic diversity of a population generated by SNP data does not have to be compared to a base population, seeing as it is absolute, and it can be used to observe the Mendelian inheritance of alleles from parent to offspring (Engelsma, 2010). It also facilitates the identification of region-specific diversity, which can help identify low diversity areas on the genome that need to be conserved. Also, dense areas of SNPs occur over the entire genome, which enables the detailed estimation of individual genetic diversity (Vignal *et al.,* 2002).

The recent sequencing of a Quarter Horse revealed the presence of 3.1 million SNPs in the equine genome (Doan *et al.,* 2012). The discovery of these SNPs and the advancement in technology has led to the creation of a whole genome SNP array, or Equine SNP Beadchip (Illumina). The first-generation array could successfully detect around 53 000 SNPs to construct a reliable genotype for each tested horse (McCue *et al.,* 2012), and is being improved to detect nearly 2 million SNPs (Schaefer *et al.,* 2017). The SNP array has successfully identified gene variations between 33 breeds associated with traits such as

muscle fibre type, size determination and presence of alternate gaits (Petersen *et al.,* 2013b), as well as assessed the genetic diversity for 814 horses of 36 breeds (Petersen *et al.,* 2013a).

1.4.3. Genetic diversity of the SA Boerperd

Despite the Boerperd being South Africa's national equine breed, few studies have been conducted to determine its genetic diversity and relatedness to other breeds. Cothran and van Dyk (1998) found that despite the SA Boerperd's strict breeding program and avoidance of outcrossing, its observed heterozygosity (H_o = 0.385) was the highest compared to the Basotho and Nooitgedacht (Table 1.2). It was also shown that the breed is closely related to gaited North American breeds, such as the Standardbred Trotter, Morgan and Saddlebred, as well as to the Thoroughbred and its crosses. Its closest relatives are the Basotho and Nooitgedacht, seeing as all three have the Cape horse as a common ancestor. In the study of Cothran and van Dyk (1998) genetic diversity was determined with blood-typing, for which allele frequencies were determined for 17 loci. Seven of these were blood-group loci, whilst the rest were serum and rbc-lysate protein loci.

Table 1.2 Estimates of genetic variability for South African and selected other domestic horse breeds (Adapted from Cothran & van Dyk (1998)).

Breed	**N**	***Ho***	***He***	***Na***
Nooitgedacht	21	0.357	0.329	2.289
Boerperd	34	0.385	0.358	2.352
Basotho pony	34	0.347	0.336	2.607
American Saddlebred	259	0.404	0.409	2.625
Arabian	117	0.307	0.327	2.132
Bedouin Arabian	213	0.209	0.224	1.890
Chilean Criollo	173	0.375	0.383	2.919
Friesian	314	0.307	0.306	1.901
Peruvian Paso	141	0.451	0.445	2.761
Quarter Horse	168	0.396	0.393	2.653
Standardbred Pacer	341	0.401	0.395	2.142
Thoroughbred	265	0.294	0.288	2.009
Domestic horse mean	99	0.377	0.364	2.393
Standard deviation		0.051	0.044	0.250

N = sample size; N_a = mean number of alleles

In later years Botha (2002) looked at 12 microsatellite loci and compared the genetic diversity of seven breeds in South Africa, as shown in Table 1.3. According to Botha (2002), the SA Boerperd was found to have an H_o of 0.686, second only to that of the Welsh pony. Since the number of investigated SA Boerperd individuals were almost double the amount of the other breeds, it is possible that this value is more accurate than the others. Concerning population differentiation, F_{ST}-values indicated that the SA Boerperd is the least differentiated from the Welsh pony (0.024), and is also comparable to the Cape Boerperd with a value of 0.029. The phylogenetic tree constructed by this author supported that of Cothran and van Dyk (1998), corroborating that the SA Boerperd is closely related to the Cape horse and American breeds.

Table 1.3 Genetic variability at 12 loci in seven horse populations (Adapted from Botha (2002)).

Population	**N**	N_a	H_o	H_e
Arabian	80	5.6	0.575	0.640
Quarter Horse	80	7.3	0.657	0.699
American Saddlebred	80	7.0	0.661	0.695
Cape Boerperd	77	5.8	0.610	0.634
SA Boerperd	128	7.7	0.686	0.744
Welsh pony	81	7.7	0.707	0.766
Friesian	39	5.3	0.534	0.576

N = sample size; N_a = mean number of alleles

Since the last diversity study was performed in 2002, it is possible that in the last fifteen years inbreeding could have influenced the allele frequencies of the breed due to the stringent selection of the breed's registry. Selection for a certain conformational 'type', as well as for certain physical abilities, can often lead to homogeneity both in phenotype and genotype. The registry has only been closed since 1998, thus the breed had recently been outcrossed with other breeds at the time of Botha's study (2002). More research is thus needed to ensure that the breeding practices implemented on the SA Boerperd at present are not genetically degrading the country's stock.

1.5. Aims of the current study

The harsh landscape of Southern Africa created the SA Boerperd and today this breed still has value within the country. Annually large amounts of time and money are spent on breeding horses that possess certain desired traits, thus presenting a method able to aid the selection of potential mates can greatly simplify selection. One such method is the screening of individuals for the presence of genetic markers associated with physical traits, such as height and gaitedness. Unfortunately, with the selection for traits comes the possibility of unintentionally causing the genetic diversity of the breed to decline, through inbreeding. This can prove hazardous to the health and future continuation of the breed, thus the levels of diversity within the breed must be monitored to determine whether more sustainable breeding practices must be implemented.

Based on the above the specific aims of this study were:

i. To investigate the presence of both the height and gait associative SNPs in the SA Boerperd.
ii. To determine the presence of substructure formation within the SA Boerperd using 17 microsatellite markers.
iii. To compare the current level of diversity and inbreeding to that of the breed's diversity determined by Botha (2002).
iv. To compare the level of genetic diversity of the SA Boerperd with that of other breeds that may have influenced it.

REFERENCES

Achilli, A., Olivieri, A., Soares, P., Lancioni, H., Kashani, B.H., Perego, U.A., Nergadze, S.G., Carossa, V., Santagostino, M., Capomaccio, S., Felicetti, M., Al-Achkar, W., Penedo, M.C.T., Verini-Supplizi, A., Houshmund, M., Woodward, S.R., Semino, O., Silvestrelli, M., Giulotto, E., Pereira, L., Bandelt, H. and Torroni, A. (2012). Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proceedings of the National Academy of Sciences of the United States of America,* 109: 2449 – 2454.

Allendorf, F.W., Luikart, G. and Aitken, S.N. (2013). Conservation and the genetics of populations. Chicester: Wiley-Blackwell.

Allfrey, S. (1980). *Horses and foals*. London: Ward Lock Limited.

Andersson, L.S., Larhammar, M., Memic, F., Wootz, H., Schwochow, D., Rubin, C., 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 in mice. *Nature,* 488: 642 – 646. doi: 10.1038/nature11399

Arnason, T. and Van Vleck, L.D. (2000). Genetic improvement of the horse. In: *The Genetics of the Horse:* 473 – 497. Edited by A.T. Bowling and A. Ruvinsky. New York: CABI Publishing.

Avise, J.C. (2004). *Molecular markers, natural history and evolution*. 2nd edition. Massachusetts: Sinauer Associates.

Baban, M., Curik, I., Antunovic, B., Cacic, M., Korabi, N. and Mijic, P. (2009). Phenotypic correlations of stride traits and body measurements in Lipizzaner stallions and mares. *Journal of Equine Veterinary Science,* 29(6): 513 – 518.

Bekker, L. (2012). Vyfgang. In: *SA Boerperd: Die perd vir die toekoms:* 31 - 34. Edited by S. Louw. Bloemfontein: SA Boerperdtelersgenootskap.

Bijma, P., Van Arendonk, J.A.M. and Woolliams, J.A. (2001). Predicting rates of inbreeding for livestock improvement schemes. *Journal of Animal Science*, 79: 840 – 853.

Botha, K. (2002). *Genetic relationships of seven horse breeds in South Africa based on DNA markers*. Dissertation (M.Sc.). Bloemfontein: University of the Free State.

Brinkmann, L., Gerken, M. and Riek, A. (2012). Adaptation strategies to seasonal changes in environmental conditions of a domesticated horse breed, the Shetland pony (*Equus ferus caballus*). *Journal of Experimental Biology*, 215: 1061 – 1068. doi: 10.1242/jeb.064832

Brooks, S.A., Makvandi-Nejad, S., Chu, E., Allen, J.J., Streeter, C., Gu, E., McCleery, B., Murphy, B.A., Bellone, R. and Sutter, N.B. (2010a). Morphological variation in the horse: Defining complex traits of body size and shape. *Animal Genetics,* 41: 159 – 165.

Brooks, S. A., Gabreski, N., Miller, D., Brisbin, A., Brown, H. E., Streeter, C., Mezey, J., Cook, D. and Antczak, D. F. (2010b). Whole-genome SNP association in the horse: Identification of a deletion in Myosin Va responsible for Lavender Foal Syndrome. *PLOS Genetics,* 6 (4): e1000909. doi: 10.1371/journal.pgen.1000909

Castle, W.E. (1948). The ABC of color inheritance in horses. *Genetics,* 33: 22 – 35.

Charlesworth, B. and Charlesworth, D. (1999). The genetic basis of inbreeding depression. *Genetical Research*, 74: 329 – 340.

Chen, J.W., Uboh, C.E., Soma, L.R., You, Y., Jiang, Z., Guan, F. and Liu, Y. (2014). Identification of sample donor by 24-plex short tandem repeat in a post-race equine plasma containing dexamethasone. *SpringerPlus,* 3: 94.

Clark, C. (2011). *The ultimate guide to horses*. Bath: Parragon.

Clayton, H.M. and Barlow, D.A. (1991). Stride characteristics of four Grand Prix jumping horses. *Equine Exercise Physiology,* 3: 151 – 157.

Cleynen, I. and Van de Ven, W.J. (2008). The HMGA proteins: A myriad of functions (Review). *International Journal of Oncology,* 32: 289 – 305.

Coates, B.S., Sumerford, D.V., Miller, N.J., Kim, K.S., Sappington, T.W., Siegfried, B.D. and Lewis, L.C. (2009). Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. *Journal of Heredity,* 100: 556 – 564.

Cothran, E.G. and Van Dyk, E. (1998). Genetic analysis of three South African horse breeds. *Journal of the South African Veterinary Association,* 69(4): 120 – 125.

Csurhes, S., Paroz, G. and Markula, A. (2009). Pest Animal Risk Assessment: Feral Horse. *Queensland Primary Industries and Fisheries,* 1: 1 – 27.

Dimock, W.W. (1950). "Wobbles": An hereditary disease in horses. *Journal of Heredity,* 41: 319 – 323.

Doan, R., Cohen, N.D., Sawyer, J., Ghaffari, N., Johnson, C.D. and Dindot, S.V. (2012). Whole-genome sequencing and genetic variant analysis of a Quarter Horse mare. *BMC Genomics*, 13: 78.

Du Toit, K. (2010). *Boerperd: Die geskiedenis van die S.A. Boerperd*. Pretoria: West Publishers.

Ellegren, H. (2000). Heterogeneous mutation processes in human microsatellite DNA sequences. *Nature Genetics*, 24(4): 400 – 402.

Engelsma, K.A. (2010). *Use of SNP markers to conserve genome-wide genetic diversity in livestock*. Thesis (Ph.D.). Wageningen: Wageningen University.

Equus Horse Safaris. (2016). *Horse Safaris.* Available at: http://www.equus.co.za/safaris.html [Accessed 15 Sept. 2016].

Euro Dressage. (2016). *FEI pony measurement rules fine-tuned and regulated*. Available at: http://www.eurodressage.com/equestrian/2008/04/17/fei-pony-measurement-rules-fine-tuned-and-regulated [Accessed 12 Oct. 2016].

European Horse Network. (2012). *Equine Sport*. Available at: www.europeanhorsenetwork.eu/app/download/.../Equine+Sport.pdf [Accessed 11 Oct. 2016].

FAOSTAT. 2014. Available at: http://faostat.fao.org/ [Accessed 15 Sept. 2016].

Farmer's Weekly. (2015). *Just how high is your horse?* Available at: http://www.farmersweekly.co.za/article.aspx?id=76183&h=Just-how-high-is-your-horse [Accessed 12 Oct. 16].

Finno, C.J. and Bannasch, D.L. (2014). Applied equine genetics. *Equine Veterinary Journal,* 46(5): 538 – 544.

Fox-Clipsham, L.Y., Carter, S.D., Goodhead, I., Hall, N., Knottenbelt, D.C., May, P.D., Ollier, W.E. and Swinburne, J.E. (2011). Identification of a mutation associated with fatal foal immunodeficiency syndrome in the Fell and Dales pony. *PLoS Genetics,* 7: e1002133. doi: 10.1371/journal.pgen.1002133

Frankham, R., Ballou, J.D. and Briscoe, D.A. (2002). *Introduction to Conservation Genetics*. Cambridge: Cambridge University Press.

Frischknecht, M., Jagannathan, V., Plattet, P., Neuditschko, M., Signer-Hasler, H., Bachmann, I., Pacholewska, A., Drögemüller, C., Dietschi, E., Flury, C., Rieder, S. and Leeb, T. (2015). A non-synonymous *HMGA2* variant decreases height in Shetland Ponies and other small horses. *PLoS ONE,* 10(10): e0140749. doi: 10.1371/journal.pone.0140749.

Goldstein, D. and Schlötterer, C. (1999). *Microsatellites: Evolution and applications*. New York: Oxford University Press.

Goodwin, D., Levine, M. and McGreevy, P.D. (2008). Preliminary investigation of morphological differences between ten breeds of horses suggests selection for paedomorphosis. *Journal of Applied Animal Welfare Science,* 11: 204 – 212.

Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R., Finlay, E.K., Jianlin, H., Groeneveld, E., Weigend, S. and the GLOBALDIV Consortium. (2010). Genetic diversity in farm animals – a review. *Animal Genetics*, 41: 6 – 31.

Gu, J., Orr, N., Park, S.D., Katz, L.M., Sulimova, G., MacHugh, D.E. and Hill, E.W. (2009). A genome scan for positive selection in thoroughbred horses. *PLoS One,* 4: e5767. doi: 10.1371/journal.pone.0005767

Hanotte, O. and Jianlin, H. (2005). Genetic characterization of livestock populations and its use in conservation decision-making. *The role of Biotechnology*, Villa Gualino, Turin, Italy: 5-7 March, 2005, 131 – 136.

He, S., Zhang, L., Li, W. and Liu, M. (2015). BIEC2-808543 SNP in the *LCORL* gene is associated with body conformation in the Yili horse. *Animal Biotechnology,* 26: 289 – 291.

Hill, E.W., Gu, J., Eivers, S.S., Fonseca, R.G., McGivney, B.A., Govindarajan, P., Orr, N., Katz, M. and MacHugh, D. (2010a). A sequence polymorphism in MSTN predicts sprinting ability and racing stamina in thoroughbred horses. *PLoS One,* 5: e8645. doi: 10.1371/journal.pone.0008645.

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,* 11: 552.

Hrísdalur Hestar. (2016). *Horse Shows*. [image] Available at: http://www.hrisdalur.is/photogallery/horseshows [Accessed 3 Oct. 2016].

Imsland, F., McGowan, K., Rubin, C., Henegar, C., Sundström, E., Berglund, J., Schwochow, D., Gustafson, U., Imsland, P., Lindblad-Toh, K., Lindgren, G., Mikko, S., Millon, L., Wade, C., Schubert, M., Orlando, L., Penedo, M.C.T., Barsh, G.S. and Andersson, L. (2016). Regulatory mutations in *TBX3* disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses. *Nature Genetics,* 48: 152 – 158.

Juras, R., Cothran, E.G. and Klimas, R. (2003). Genetic analysis of three Lithuanian native horse breeds. *Acta Agriculturae Scandinavica, Section A – Animal Science,* 53(4): 180 – 185.

Kelekna, P. (2009). *The horse in human history*. New York: Cambridge University Press.

Khanshour, A., Juras, R., Blackburn, R. and Cothran, E.G. (2015). The legend of the Canadian Horse: Genetic diversity and breed origin. *Journal of Heredity,* 106(1): 37 – 44.

Kidd, J. (1981). *An illustrated guide to horse and pony care*. London: Salamander Books.

Koenen, E.P.C. and Aldridge, L.I. (2002). *Testing and genetic evaluation of sport horses in an international perspective.* Paper presented at the 7th World Congress Applied to Livestock Production: Montpellier, France.

Koenen, E.P.C., Aldridge, L.I. and Philipsson, J. (2004). An overview of breeding objectives for warmblood sport horses. *Livestock Production Science,* 88: 77 – 84.

Kristjansson, T., Bjornsdottir, S., Sigudsson, A., Andersson, L.S., Lindgren, G., Helyar, S.J., Klonowski, A.M. and Arnason, T. (2014). The effect of the 'gait keeper' mutation in the *DMRT3* gene on gaiting ability in Icelandic horses. *Journal of Animal Breeding and Genetics,* 131: 415 – 425.

Kulbrock, M., Lehner, S., Metzger, J., Ohnesorge, B. and Distl, O. (2013). A genome-wide association study identifies risk loci to equine recurrent uveitis in German warmblood horses. *PLoS One,* 8: e71619. doi: 10.1371/journal.pone.0071619

Lacy, R.C. (1987). Loss of genetic diversity from managed populations: Interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology*, 1: 143 – 158.

Landry, D. (2008). *Noble brutes: How Eastern horses transformed English culture*. Baltimore: Johns Hopkins University Press.

Lekota, T.A. (2001). *The state of the Basotho Pony in Lesotho* (Community-based management of animal genetic resources). Proceedings of the workshop held in Mbabane, Swaziland.

Leroy, G., Callède, L., Verrier, E., Mériaux, C., Ricard, A., Danchin-Burge, C. and Rognon, X. (2009). Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. *Genetics Selection Evolution,* 41: 5.

Lippold, S., Matzke, N.J., Reissmann, M. and Hofreiter, M. (2011). Whole mitochondrial genome sequencing of domestic horses reveals incorporation of extensive wild horse diversity during domestication. *BMC Evolutionary Biology,* 11: 328. doi: 10.1186/1471-2148-11-328

Lloyd, A.S., Martin, J.E., Bornett-Gauci, H.L.I. and Wilkinson, R.G. (2008). Horse personality: Variation between breeds. *Applied Animal Behaviour Science,* 112: 369 – 383.

Louw, S. (2008). SA Boerperd –Indigenous breed to southern Africa. In: *SA Boerperd: The ideal sport and pleasure horse:* 22 - 29. Edited by S. Louw. Maanhaarrand: Kejafa Knowledge Works.

Ludwig, A., Pruvost, M., Reissmann, M., Benecke, N., Brockmann, G.A., Castaños, P., Cieslak, M., Lippold, S., Llorente, L., Malaspinas, A., Slatkin, M. and Hofreiter, M. (2009). Coat color variation at the beginning of horse domestication. *Science,* 324: 485. doi: 10.1126/science.1172750

Makvandi-Nejad, S., Hoffman, G.E., Allen, J.J., Chu, E., Gu, E., Chandler, A.M., Loredo, A.I., Bellone, R.R., Mezey, J.G., Brooks, S.A. and Sutter, N.B. (2012). Four loci explain 83% of size variation in the horse. *PLoS ONE,* 7(7): e39929. doi: 10.1371/journal.pone.0039929

Marletta, D., Tupac-Yupanqui, I., Bordonaro, S., Garcia, D., Guastella, A.M., Criscione, A., Cañón, J. and Dunner, S. (2006). Analysis of genetic diversity and the determination of relationships among western Mediterranean horse breeds using microsatellite markers. *Journal of Animal Breeding Genetics*, 123: 315 – 325.

Mathai, C.K., Ohno, S. and Beutler, E. (1966). Sex-linkage of the glucose-6-phosphate dehydrogenase gene in equidae. *Nature,* 210: 115 – 116.

McBride, S.D. and Mills, D.S. (2012). Psychological factors affecting equine performance. *BMC Veterinary Research,* 8: 180 – 191.

McClory, H. and Kowalski, S.P. (2014). Horses as sources of proprietary information: Commercialization, conservation, and compensation pursuant to the convention on biological diversity. *Journal of Agrobioechnology Management & Economics,* 17(2): 141 – 155.

McCue, M.E., Bannasch, D.L., Petersen, J.L., Gurr, J., Bailey, E., Binns, M.M., Distl, O., Guérin, G., Hasegawa, T., Hill, E.W., Leeb, T., Lindgren, G., Penedo, M.C.T., Røed, K.H., Ryder, O.A., Swinburne, J.E., Tozaki, T., Valberg, S.J., Vaudin, M., Lindblad-Toh, K., Wade, C.M. and Mickelson, J.R. (2012). A high density SNP array for the domestic horse and extant perissodactyla: Utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genetics,* 8(1): e1002451. doi: 10.1371/journal.pgen.1002451

Metzger, J., Schrimpf, R., Philipp, U. and Distl, O. (2013). Expression levels of *LCORL* are associated with body size in horses. *PLoS ONE,* 8(2): e56497. doi: 10.1371/journal.pone.0056497

Nel, C. (2014). *SA Boerperd Breed History*. SA Boerperd Breeders Society. Available at: http://www.saboerperd.com/p2/history/sa-boerperd-breed-history.html [Accessed 5 July 2016].

Norwegian Fjord Horse Registry. (2016). *Norwegian Fjord Horse colors*. Available at: http://www.nfhr.com/catalog/index.php?colors=1 [Accessed 12 Sept. 2016].

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P.L., Fumagalli, M., Vilstrup, J.T., Raghavan, M., Korneliussen, T., Malaspinas, A.S., Vogt, J., Szklarczyk, D., Kelstrup, C.D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A.M., Cahill, J., Rasmussen, M., Wang, X., Min, J., Zazula, G.D., Seguin-Orlando, A., Mortensen, C., Magnussen, K., Thompson, J.F., Weinstock, J., Gregersen, K., Røed, K.H., Eisenmann, V., Rubin, C.J., Miller, D.C., Antczak, D.F., Bertelsen, M.F., Brunak, S., Al-Rasheid, K.A., Ryder, O., Andersson, L., Mundy, J., Krogh, A., Gilbert, M.T., Kjær, K., Sicheritz-Ponten, T., Jensen, L.J., Olsen, J.V., Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J. and Willerslev, E. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature,* 499: 74 – 78.

Orphanides, G., Lagrange, T. and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes & Development,* 10(21): 2657 – 2683.

Outram, A.K., Stear, N.A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., Thorpe, N. and Evershed, R.P. (2009). The earliest horse harnessing and milking. *Science,* 323(5919): 1332 – 1335. doi: 10.1126/science.1168594

Pakamisa. (2016). *Horseback Safaris*. Available at: http://www.pakamisa.co.za/horseback-safaris.php [Accessed 15 Sept. 2016].

Perryman, L. E. (2004). Molecular pathology of severe combined immunodeficiency in mice, horses, and dogs. *Veterinary Pathology,* 41: 95 – 100.

Petersen, J.L., Mickelson, J.R., Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., Da Câmara Machado, A., Distl, O., Felicetti, O., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C.T., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Silvestrelli, M., Swinburne, J., Tozaki, T., Vaudin, M., Wade, C.M. and McCue, M.E. (2013a). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *Plos ONE,* 8(1): e54997. doi: 10.1371/journal.pone.0054997

Petersen, J.L., Mickelson, J.R., Rendahl, A.K., Valberg, S.J., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., Da Câmara Machado, A., Capomaccio, S., Cappelli, K., Cothran, E.G., Distl, O., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C.T., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Swinburne, J., Tozaki, T.,

Vaudin, M., Wade, C.M. and McCue, M.E. (2013b). Genome-wide analysis reveals selection for important traits in domestic horse breeds. *Plos Genetics,* 9(1): e1003211. doi: 10.1371/journal.pgen.1003211

Promerová, M., Andersson, L.S., Juras, R., Penedo, M.C.T., Reissmann, M., Tozaki, T., Bellone, R., Dunner, S., Hořín, P., Imsland, F., Imsland, P., Mikko, S., Modrý, D., Roed, K.H., Schwochow, D., Vega-Pla, J.L., Mehrabani-Yeganeh, H., Yousefi-Mashouf, N., Cothran, E.G., Lindgren, G. and Andersson, L. (2014). Worldwide frequency distribution of the '*Gait keeper*' mutation in the *DMRT3* gene. *Stichting International Foundation for Animal Genetics,* 45: 274 – 282.

Protas, M.E. and Patel, N.H. (2008). Evolution of coloration patterns. *Annual Reviews of Cell and Developmental Biology,* 24: 425–46.

Robilliard, J.J., Pfau, T. and Wilson, A.M. (2007). Gait characterization and classification in horses. *Journal of Experimental Biology,* 210: 187 – 197.

Rothschild, M.F. (2003). Approaches and challenges in measuring genetic diversity in pigs. *Archivos de Zootecnia,* 52: 129 – 135.

Royal Dutch Sport Horse. (2016). *Breeding values*. Available at: http://www.kwpn.org/kwpn/breeding/breeding-values/ [Accessed 30 Oct. 2016].

Saastamoinen, M.T. and Barrey, E. (2000). Genetics of conformation, locomotion and physiological traits. In: *The Genetics of the Horse:* 473 – 497. Edited by A.T. Bowling and A. Ruvinsky. New York: CABI Publishing.

SA Boerperd Breeders Society. (2014). *Handleiding vir rasstandaard en bouvorm*. Available at: http://www.saboerperd.com/images/photos/rasstandaard_saboerperd.pdf [Accessed 26 May 2016].

SA Boerperd Breeders Society. (2016). *Showing – Traditional 5 gaited classes*. [image] Available at: http://www.saboerperd.com/g4/v1/sa-boerperd/sa-boerperd-gallery/traditional-5-gaited/showing---traditional-5-gaited-classes.html [Accessed 22 Oct. 2016].

Sánchez, M.J., Gómez, M.D., Peña, F., Monterde, J.G., Morales, J.L., Molina, A. and Valera, M. (2013). Relationship between conformation traits and gait characteristics in Pura Raza Español horses. *Archiv Tierzucht,* 56: 137 – 148.

SA Stud Book. (2016). *Perdetelersgenootskappe*. Available at: http://www.sastudbook.co.za/dc3/sa-stamboek/telersgenootskappe/perde/perdetelersgenootskappe.html [Accessed 2 Oct. 2016].

Schaefer, R.J., Schubert, M., Bailey, E., Bannasch, D.L., Barrey, E., Bar-Gal, G.K., Brem, G., Brooks, S.A., Distl, O., Fries, R., Finno, C.J., Gerber, V., Haase, B., Jagannathan, V., Kalbfleisch, T., Leeb, T, Lindgren, G., Lopes, M.S., Mach, N., Da Câmara Machado, A., MacLeod, J.N., McCoy, A., Metzger, J., Penedo, C., Polani, S., Rieder, S., Tammen, I., Tetens, J., Thaller, G., Verini-Supplizi, A., Wade, C.M., Wallner, B., Orlando, L., Mickelson, J.R. and McCue, M.E. (2017). Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds. *BioMed Central Genomics*, 18: 565. doi: 10.1186/s12864-017-3943-8

Schlötterer, C. and Tautz, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*, 20: 211 – 215.

Schroder, W., Klostermann, A., Stock, K.F. and Distl, O. (2012). A genome-wide association study for quantitative trait loci of show-jumping in Hanoverian warmblood horses. *Animal Genetics,* 43: 392 – 400.

Schubert, M., Jónsson, H., Chang, D., Sarkissian, C.D., Ermini, L., Ginolhac, A., Albrechtsen, A., Dupanloup, I., Foucal, A., Petersen, B., Fumagalli, M., Raghavan, M., Seguin-Orlando, A., Korneliussen, T.S., Velazquez, A.M.V., Stenderup, J., Hoover, C.A., Rubin, C., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., MacHugh, D.E., Kalbfleisch, T., MacLeod, J.N., Rubin, E.M., Sicheritz-Ponten, T., Andersson, L., Hofreiter, M., Marques-Bonet, T., Gilbert, M.T.P., Nielsen, R., Excoffier, L., Willerslev, E., Shapiro, B. and Orlando, L. (2014). Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *Proceedings of the National Academy of Science of the United States of America,* 111(52): 5661 – 5669.

Signer-Hasler, H., Flury, C., Haase, B., Burger, D., Simianer, H., Leeb, T. and Rieder, S. (2012). A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PLoS ONE,* 7(5): e37282. doi: 10.1371/journal.pone.0037282

Slatkin, M. (2008). Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. *Nature,* 9: 477 – 485.

South African Equestrian Federation. (2016). *Welcome to the SAEF*. Available at: http://www.saef.org.za/EN/Content/Pages/Home [Accessed 11 Oct. 2016].

Stock, K.F. and Distl, O. (2008). Multiple-trait selection for radiographic health of the limbs, conformation and performance in Warmblood riding horses. *Animal,* 2: 1724 – 1732.

Sunnucks, P. (2000). Efficient genetic markers from population biology. *Trends of Ecology and Evolution,* 15: 199 – 203.

Swart, S. (2010). *Riding High: Horses, Humans and History in South Africa*. Johannesburg: Wits University Press.

Tautz, D. (1993). Notes on the defunction and nomenclature of tandemly repetitive DNA sequences. In: *DNA Fingerprinting: State of the Science:* 21–28. Edited by S.D.J. Pena. Basel, Switzerland: Birkhauser Verlag.

The British Horse Society. (2015). *Equestrian Statistics*. Available at: http://www.bhs.org.uk/our-charity/press-centre/equestrian-statistics [Accessed 11 Oct. 2016].

The Horse. (2001). *Focus on discipline: Cutting horses.* Available at: http://www.thehorse.com/articles/12445/focuson-discipline-cutting-horses [Accessed 24 Jan. 2018].

The Nooitgedachter Horse. (2018). *History of the Nooitgedachter*. Available at: http://nooitgedachter.co.za/index.php?option=com_content&view=article&id=107&Itemid=582 [Accessed 14 Jan. 2018].

Therapy Horses. (2016). *Gentle Carousel Miniature Therapy Horses*. Facebook. [image] Available at: https://web.facebook.com/TherapyHorses/ [Accessed 6 Oct. 2016].

Thomas, H.S. (2009). *Storey's guide to raising horses*. North Adams: Storey Publishing.

Thoroughbred Breeders' Association. (2016). *Who we are*. Available at: http://www.tba.co.za/tba/index.php [Accessed 2 Oct. 2016].

Tozaki, T., Miyake, T., Kakoi, H., Gawahara, H., Sugita, S., Hasegawa, T., Ishida, N., Hirota, K. and Nakano, Y. (2010). A genome-wide association study for racing performances in Thoroughbreds clarifies a candidate region near the MSTN gene. *Animal Genetics,* 41: 28 – 35.

Tsunoda, T., Takashima, Y., Tanaka, Y., Fujimoto, T., Doi, K., Hirose, Y., Koyanagi, M., Yoshida, Y., Okamura, T., Kuroki, M., Sasazuki, T. and Shirasawa, S. (2010). Immune-related zinc finger gene ZFAT is an essential transcriptional regulator for hematopoietic differentiation in blood islands. *Proceedings of the National Academy of Sciences of the United States of America,* 107: 14199 – 14204.

US Equestrian. (2018). *Reining*. Available at: http://www.usef.org/compete/disciplines/reining [Accessed 24 Jan. 2018].

Van de Goor, L.H.P. and van Haeringen, W.A. (2007). Identification of horses sold for slaughter using 17 genetic markers. In: *Nonhuman DNA Typing, Theory and Casework Applications*: 120 – 121. Edited by H.M. Coyle. Boca Raton: CRC Press.

Van de Goor, L.H.P., Panneman, H. and van Haeringen, W.A. (2009). A proposal for standardization in forensic equine DNA typing: Allele nomenclature of 17 equine-specific STR loci. *Animal Genetics,* 41: 122 – 127.

Van de Goor, L.H.P., van Haeringen, W.A. and Lenstra, J.A. (2011). Population studies of 17 equine STR for forensic and phylogenetic analysis. *Stichting International Foundation for Animal Genetics,* 42: 627 – 633.

Vignal, A., Milan, D., SanCristobal, M. and Eggen, A. (20020. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, 34: 275 – 305.

Viklund, A., Nasholm, A., Strandberg, E. and Philipsson, J. (2010). Effects of long-time series of data on genetic evaluations for performance of Swedish Warmblood riding horses. *Animal*, 4: 1823 – 1831.

Warmuth, V., Eriksson, A., Bower, M.A., Barker, G., Barrett, E., Hanks, B.K., Li, S., Lomitashvili, D., Ochir-Goryaeva, M., Sizonov, G.V., Soyonov, V. and Manica, A. (2012). Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proceedings of the National Academy of Sciences of the United States of America,* 109: 8202 – 8206.

Whitehead, L. and Mansfield, D. (2013). Performance analysis: The SA Boerperd takes hands with an exciting new research project. *HQ,* March, pp. 50 – 53.

Wood, C.H. and Jackson, S.G. (2004). *Horse judging manual*. UK Cooperative Extension Service: University of Kentucky. Available at: http://www2.ca.uky.edu/agcomm/pubs/asc/asc118/asc118.pdf [Accessed 6 Sept. 2016].

WordPress. (2014). *The horse in the Anglo-Boer War*. [image] Available at: https://grahamlesliemccallum.wordpress.com/2014/05/07/1523/ [Accessed 4 Oct. 2016].

Yang, M.Q., Laflamme, K., Gotea, V., Joiner, C.H., Seidel, N.E., Wong, C., Petrykowska, H.M., Lichtenberg, J., Lee, S., Welch, L., Gallagher, P.G., Bodine, D.M. and Elnitski, L. (2011). Genome-wide detection of a *TFIID* localization element from an initial human disease mutation. *Nucleic Acids Research,* 39: 2175 – 2187.

Chapter 2: Performance analysis of the SA Boerperd

(Pictures obtained from SA Boerperd Breeders Society; www.saboerperd.com)

2.1. Introduction

The South African (SA) Boerperd showing industry is extremely competitive. For optimal performance, each horse must possess certain performance-based characteristics to be able to execute a specific task. Usually the unique conformation of a breed makes it proficient for a certain event, but the SA Boerperd is one of few breeds able to excel in many disciplines (Louw, 2008; Du Toit, 2010). One of these disciplines, the five-gaited class, requires participating horses to showcase five varying movement styles. It should be noted that although the breed standard of the SA Boerperd states that the breed is known to be five-gaited (Du Toit, 2010), not all individuals of the population possess this ability (Promerová *et al.*, 2014).

In recent years, research has shown that the display of additional gaits, also referred to as gaitedness, is largely subject to the presence of specific polymorphisms in the genetic code of the horse (Andersson *et al.,* 2012). These mutations, also known as single nucleotide polymorphisms (SNPs), have been shown to give horses the potential to display certain gaits uncommon to most equine breeds. Two such gaits, namely the *rack* and *short-gait*, are required of SA Boerperd individuals that partake in five-gaited classes (Louw, 2008).

To ensure that show horses have the potential to perform these gaits, it would be prudent to screen individuals for the presence of the relevant associative SNPs. Even with genetic screening, it should however be noted that the presence of these SNPs alone may not be able to guarantee gaitedness in the individual due to variables like individual fitness and the training method used (Bekker, 2012). The genetic result can nevertheless inform trainers on whether the tested individual will be easily trainable or not. This is because individuals who are homozygous for the *gait keeper* mutation have a higher affinity for lateral gaits, such as *tölt* (Andersson *et al*., 2012; Kristjansson *et al.,* 2014) and possibly the *rack*. In addition, individuals that have the relevant SNPs can potentially be used for breeding to ensure the continuance of the mutations in the offspring.

In addition to gaitedness, the other trait-related polymorphisms worth screening for are those influencing size. This performance trait is subject to multiple major and minor contributing SNPs (Makvandi-Nejad *et al.,* 2012; Signer-Hasler *et al.,* 2012), and can influence how well an individual performs in a specific event. The height of a horse has been linked to its execution

of certain gaits (Signer-Hasler *et al.,* 2012), which can affect how it performs in five-gaited classes. It can also influence performance in events such as jumping. For this event, factors such as the length of strides taken in between jumps, the ability to clear a jump of a certain height, and the speed by which the course can be completed need to be considered. The physical height of a horse has been found to influence such performance factors. A tall horse gives long strides (Baban *et al.,* 2009) and is high in relation to a jump obstacle. Both these facts are advantageous in aiding the horse to complete a fast, clear round. The SA Boerperd is used to compete in show jumping and eventing courses (SA Boerperd Breeders Society, 2016), and therefore it would be beneficial to test for the mutations associated with height in this breed.

Very few studies have focused on the genetic composition of the SA Boerperd, and only one has genetically investigated a performance-based trait (Promerová *et al.*, 2014). This lack of investigation into the breed has gone unnoticed by most breeders, mainly since the traditional phenotype-based way of selecting horses and ensuring that characteristics are passed on to progeny has served them well over the years (Du Toit, 2010; Swart, 2010). Yet with jumping and equitation being two of the most popular equestrian disciplines in South Africa (Equestrian, 2017), the demand for equine performers can only increase. A cost- and time-effective method would be ideal to meet this demand, thus screening for performance-influencing SNPs is a viable venture. Breeders generally tend to shy away from genetic testing due to its high cost, but in the long run it will save both time and money when hours of training are only spent on individuals proven to have the genetic predisposition for a certain trait.

This study was conducted to generate data to inform breeders of the benefit of genetic screening and to aid future breeding of the SA Boerperd. The specific aims of this study were to verify the occurrence of choice SNPs associated with both height and gaitedness in the SA Boerperd, and to determine the frequency of these SNPs. The occurrence of the height-SNPs were also compared to the physical height of the horse to determine at what height the mutations start playing a role. Specific attention was also given to whether reported five-gaited individuals possessed the associative gait-SNPs. The frequencies of all these gene variants were also compared to that obtained for other breeds.

2.2. Materials and methods

2.2.1. Ethics statement

This study was conducted with approval from the Interfaculty Animals Ethics Committee of the University of the Free State (UFS) and the Department of Agriculture, Forestry and Fisheries under Section 20 of the Animal Disease Act 1984 (Act number 35 of 84). Ethical approval number UFS-AED2016/0033. Permission was also granted by the SA Boerperd Breeders Society of South Africa. All horses were sampled in the presence and with the consent of owners. The purpose of the project and the sampling procedure were also explained to the owners. The involvement of all horses and owners were kept confidential.

2.2.2. Sample collection

A total of 100 horses were sampled in various localities across the country. Care was taken to sample as much unrelated individuals as possible, but a few were found to have a tested sire in common. Within the Cape provinces these included the Augrabies, Upington, Jan Kempdorp, Swellendam, Buffeljagsrivier and Barkly East areas. In the Free State samples were taken from Bloemfontein, Senekal, Kroonstad and Bethlehem, and the North-West included Vryburg and Rustenburg. Gauteng localities included Johannesburg and Pretoria. Samples were also taken from Bethal, Ermelo and Nelspruit in Mpumalanga, and from Vryheid in Kwazulu-Natal. The sample areas are illustrated in Figure 2.1.

Either mane or tail hair were pulled from each horse and temporarily stored in a paper envelope until DNA analysis. This method seemed the best non-invasive approach that would yield good quality DNA (Taberlet, Waits & Luikart, 1999). This type of sampling method is also cost-effective, simple and quick. In addition to sampling, the owner of each horse was also asked to supply the individual's gender, height and gait status (three- or five-gaited). Additionally, the hair samples of two Cape Boerperd individuals were included in the study to optimise the primer-pairs of each investigated SNP. These two individuals were from the researcher's personal stock.

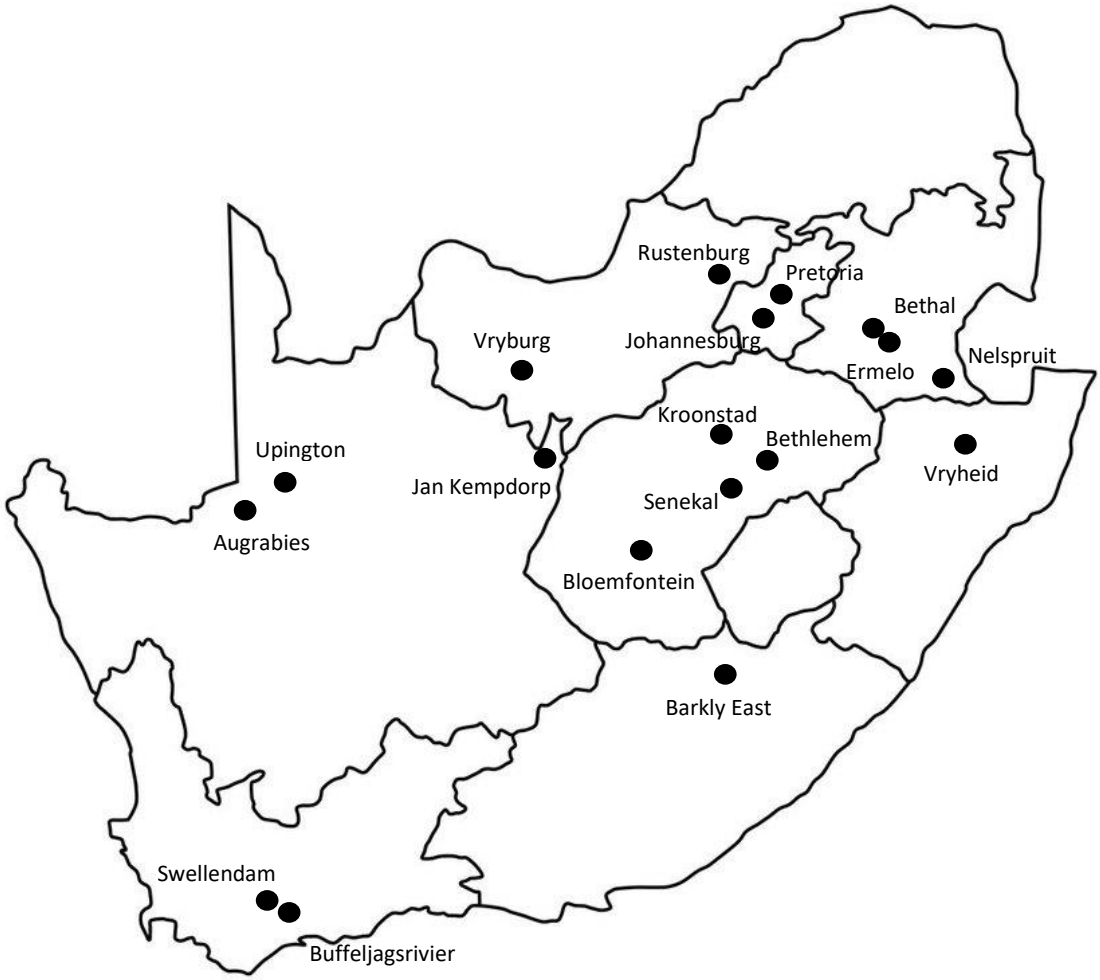


Figure 2.1 South African map indicating sampling locations of SA Boerperd individuals.

2.2.3. SNP markers used

Of the multiple gene variations that have been shown to influence both height and gait, four were chosen for investigation in this study. Firstly, the height-influencing SNP BIEC2_808543 was selected, since it was shown to have the largest phenotypical effect on the trait (Siger-Hasler *et al.,* 2012). The SNP shown to contribute the second most to height, BIEC2_1105377, was also chosen. For the gait investigation both associative SNPs, *DMRT3_Ser301STOP* and BIEC2_620109 (Andersson *et al.,* 2012), were selected. The primer sequences of all four SNPs are listed in Table 2.1. The reference sequences of the SNPs were extracted from EquCab2.0 (Broad Institute, 2008), and Primer3Plus v.4.0.0 (Untergasser *et al.,* 2012) was used to design the primer-pairs of BIEC2_1105377 and BIEC2_620109. These pairs were then subjected to the online Sequence Manipulation Suite (Stothard, 2000) to evaluate their PCR suitability before selection.

Table 2.1 Investigated SNPs, their chromosomal positions, primer sequences, annealing temperatures and references.

SNP	Position	Primer sequences	Annealing temp. (°C)	References
BIEC2_808543	ECA3: 105 547 002	P_F: 5'-TGG AGT CAG TTG GGT TTA ATG-3' P_R: 5'-GAC CGG ATA GCA TAG AGA GAG-3'	56	He *et al.* (2015)
BIEC2_1105377	ECA9: 74 798 143	P_F: 5'-CAG GAC AAC CTC CCT CAC CA-3' P_R: 5'- TAC GTT GGC TGT CTT GGG TG-3'	62	Primer3Plus
DMRT3_Ser301STOP	ECA23: 22 999 655	P_F: 5'-CGA CAA AGA CAC CGA CCA GA-3' P_R: 5'-CCG ATC CCA CGG ACC ATT-3'	62	Han *et al.* (2015)
BIEC2_620109	ECA23: 22 967 656	P_F: 5'- GCC TCT CAC CCA GAC ACC AT-3' P_R: 5'- ATC TTA TGG CAC ACG GCA CC-3'	60	Primer3Plus

P_F = forward primer; P_R = reverse primer; Primer3Plus (Untergasser *et al.,* 2012)

2.2.4. Molecular techniques

DNA from hair bulbs were used in a direct polymerase chain reaction (PCR), since this method yielded good quality products. The KAPA HiFi HotStart ReadyMix PCR Kit (KAPA Biosystems) was used for direct PCR. This kit has a low error rate with increased sensitivity, short reaction times, and is ideal for amplifying an assortment of targets and fragment sizes (KAPA Biosystems, 2016a). In addition, the KAPA2G Robust HotStart ReadyMix PCR Kit (KAPA Biosystems) was used for certain samples that would not amplify with the use of HiFi. This kit improves inhibitor tolerance and the processing of difficult assays (Kapa Biosystems, 2016b).

A number of modifications were made to the protocol provided by the manufacturer of these kits. To each tube 4.6 µl deionized H_2O, 7 µl HiFi/Robust Ready Mix, and 0.7 µl of each forward and reverse primer were added, along with 2 – 4 hair bulbs of a specific horse. This procedure was repeated for the four investigated SNPs and all 100 horses. The PCR protocol described by He *et al.* (2015) was used, though with varying annealing temperatures for each primer-pair. The protocol consisted of an initial denaturing step of 94°C for 5 minutes. A cycle consisting of 94°C for 30 seconds, 56/60/62°C (depending on the primer) for 30 seconds, and

72°C for 30 seconds was repeated 35 times. A final extension step was run at 72°C for 5 minutes.

To verify amplification success the PCR products were run on an 1% agarose electrophoresis gel for 30 minutes. If a clear fragment was visible the product was then used in the next step, but if a smear was seen the PCR product of the sample was discarded and the entire process was repeated with new hair bulbs. In some instances, especially concerning both gait-SNPs, 0.7 µl $MgCl_2$ was added to the direct PCR master mix to improve amplification.

To remove the forward and reverse primers present in the PCR product, the Bio-Spin® PCR Purification Kit (BIOER Technology) was used according to the specifications of the manufacturer. This kit was selected since it is simple to use, fast acting, and has a high yield of up to 95% (BIOER Technology, 2015). The protocol entailed adding two volumes of Binding Buffer to one volume of PCR product. After vortexing, the mixture was pipetted into a spin column. The column was then centrifuged at 6,000xg for 1 minute and the flow-through discarded. To wash the column, 650 µl Wash Buffer was added, after which it was centrifuged for 1 minute at 12,000xg and the flow-through discarded. The latter wash step was repeated for a second time and the column was centrifuged for an additional minute at 12,000xg to ensure all liquid was out of the column. The spin column was transferred to a sterile 1.5 ml micro-centrifuge tube, and 50 µl Elution Buffer was added. One of the modifications made was to add 25 µl instead of 50 µl Elution Buffer if the fragment visible on the electrophoresis gel was indistinct. After each addition of Elution Buffer the column was allowed to stand for 5 minutes instead of the recommended 1 minute at room temperature, to allow the alcohol to evaporate. Lastly, the column was centrifuged again at 12,000xg for 1 minute, with the DNA solution in the tube then kept whilst the column was discarded. This purified DNA was stored at -20°C.

The nucleotide sequence of each sampled horse was obtained using the BigDye® Terminator Cycle Sequencing Kit (Applied Biosystems), which adds different fluorescent dyes to the target DNA. This kit was selected since it can read GT-rich areas, ensures the production of high and uniform peak heights, has good dye mobility characteristics, and is cost effective (Thermo Fischer Scientific, 2016). A number of modifications were made to the protocol, with 7.3 µl of deionized H_2O, 0.5 µl buffer, 1 µl BigDye® and 0.2 µl of either the forward or reverse primer added to a tube. Lastly, 1 µl of purified PCR product was added. The sequencing PCR protocol

entailed heating to 94°C for 2 minutes, after which a cycle consisting of 94°C for 15 seconds, 53°C for 10 seconds, and 60°C for 3 minutes ran for 25 cycles.

The pre-sequencing protocol of the ZR DNA Sequencing Clean-up Kit™ (Zymo Research) was used after amplification per the specifications of the manufacturer. This kit removes contaminants such as dyes, salts and primers from the DNA extension products, since the latter can inhibit the signal strength and quality of the sequencing data (Zymo Research, 2016). The protocol involved adding 240 µl Sequencing Binding Buffer to a reaction tube, and after the mixture was vortexed it was transferred to a Zymo-Spin column in its collection tube. The column was then centrifuged for 1 minute at 13,000 rpm, after which 300 µl Sequencing Wash Buffer was added. It was once again centrifuged at 13,000 rpm, for 2 minutes. After the column was reassigned to a sterile micro-centrifuge tube 10 µl HiDi was added, and the column was centrifuged for the last time at 13,000 rpm for 15 seconds to elute the DNA. The eluted DNA was then sequenced by use of the 3130xl Genetic Analyzer (Life Technologies) to determine the genotypes of each sampled horse.

2.2.5. Data analysis

The DNA sequence of each horse was organised and analysed in the bio-informatic software, GENEIOUS v.8.1.7 (Kearse *et al.,* 2012). Concerning sequence analysis, this program enables sequence assembly, correct base calling, and automatically annotates gene variants and translation (Geneious, 2017). The average height of horses associated with a specific genotype was calculated in Microsoft Office Excel. This program was also used to count the number of gaited individuals for SNP *DMRT3_Ser301STOP*, as well as to determine the allele and genotype frequencies of all four SNPs within the SA Boerperd. Additionally, the *r* x *c* Fisher's exact test (Mehta & Patel, 1983) was performed on each SNP. This statistical test evaluates whether the proportions of two variables are independently different when sample sizes are small (McDonald, 2014).

2.3. Results

The sequences generated by the Genetic Analyzer were viewed in GENEIOUS (Kearse *et al.,* 2012) and are available in Appendix A. The SNP data obtained for each of the sampled individuals are presented in Table 2.2, along with the gender, height and gait-ability of each horse. Most of the horses sampled were mares (70%) and showed the three-gaited

phenotype (87%). The sampled group broadly ranged in height from 140 – 167.6 cm. The lack of SNP data for some samples were due to failure to amplify during PCR for the primer-pair of BIEC2_620109.

Table 2.2 Gender, height, gait-ability and diversity at performance-related SNPs in 100 SA Boerperd horses.

				Performance-related mutations (SNPs)			
				Height		Gait	
Sample	Gender	Height (cm)	Gait-ability	BIEC2_808543 (T > C)	BIEC2_1105377 (G > A)	*DMRT3_Ser301STOP* (C > A)	BIEC2_620109 (C > T)
CBP1	F	155	3	TT	GA	CC	CC
CBP2	F	157.5	3	TT	GG	CA	CT
SAB1	M	155	3	TT	GA	CC	CC
SAB2	M	156	3	TT	GA	CC	CC
SAB3	F	153	3	TT	GG	CC	CC
SAB4	F	155	3	TT	GG	CC	CC
SAB5	M	157.5	5	TT	GG	CA	
SAB6	M	157	3	TT	GG	CC	CC
SAB7	M	167.6	3	TC	GA	CC	CC
SAB8	M	155	3	TT	GA	CC	CC
SAB9	M	150	5	TT	GA	CC	
SAB10	F	143	3	TT	GG	CC	CC
SAB11	F	151	3	TC	GA	CC	CC
SAB12	F	144	3	TT	GG	CC	CC
SAB13	F	141	3	TT	AA	CC	CC
SAB14	F	146	3	TT	GA	CA	CT
SAB15	F	148	3	TT	GG	CC	CC
SAB16	M	157.5	3	TT	GA	CC	
SAB17	F	151	3	TT	AA	CA	
SAB18	F	155	3	TT	GA	CA	CT
SAB19	F	149	3	TT	GG	CC	CC
SAB20	F	146	3	TT	GA	CC	CC
SAB21	F	148	3	TT	GG	CC	CC
SAB22	M	149	3	TT	GA	CC	CC
SAB23	M	150	3	TT	GG	CC	CC
SAB24	F	147.5	3	TT	AA	CA	CT
SAB25	M	143	3	TT	GA	CC	CC
SAB26	M	150	3	TT	GA	CC	CC
SAB27	F	146	3	TT	GA	CC	CC
SAB28	F	152	3	TT	GA	CC	CC
SAB29	M	142	3	TT	GG	CC	
SAB30	M	150	3	TT	GG	CC	CC
SAB31	M	152	3	TT	GG	CC	CC
SAB32	M	147	3	TT	GG	CC	CC
SAB33	M	150	5	TT	GG	CA	CT
SAB34	M	155	5	TT	GA	CA	CT
SAB35	M	157.5	3	TT	GA	CC	CC

				Performance-related mutations (SNPs)			
				Height		Gait	
Sample	Gender	Height (cm)	Gait-ability	BIEC2_808543 (T > C)	BIEC2_1105377 (G > A)	*DMRT3_Ser301STOP* (C > A)	BIEC2_620109 (C > T)
SAB36	F	145	3	TT	GG	CC	CC
SAB37	M	147	3	TT	GA	CC	CC
SAB38	M	155	5	TT	GG	CA	CT
SAB39	M	155	3	TT	GG	CC	
SAB40	F	150	5	TT	GA	CC	CC
SAB41	F	145	3	TT	GA	CA	CT
SAB42	M	157.5	3	TT	GG	CA	
SAB43	F	150	3	TT	GG	CC	CC
SAB44	M	155	5	TT	GA	CA	
SAB45	M	150	5	TT	GA	CA	
SAB46	M	147	3	TT	GA	CC	CC
SAB47	M	162.5	3	TT	GG	CC	CC
SAB48	M	155	3	TT	GG	CA	CT
SAB49	M	152	3	TT	GG	CC	CC
SAB50	F	150	3	TT	GG	CC	CC
SAB51	F	153	3	TT	GG	CC	CC
SAB52	M	156	3	TT	GG	CC	CC
SAB53	F	152	3	TT	GA	CC	CC
SAB54	F	148	3	TT	GG	CC	CC
SAB55	M	155	3	TT	GG	CC	CC
SAB56	M	147	3	TT	GG	CA	CT
SAB57	F	146	3	TT	GG	CC	CC
SAB58	F	156	3	TT	GG	CA	CT
SAB59	F	157.5	5	TT	GG	CA	CT
SAB60	M	150	5	TT	GG	CA	CT
SAB61	F	150	3	TT	GG	CC	CC
SAB62	M	157.5	5	TT	GA	CC	
SAB63	F	145	5	TT	GG	CC	CC
SAB64	F	162.5	5	TT	GA	CA	CT
SAB65	M	150	3	TT	GA	CC	CC
SAB66	M	157.5	3	TT	GG	CA	CT
SAB67	M	165	5	TT	GA	CC	CC
SAB68	F	165	3	TT	GG	CC	CC
SAB69	M	162.5	3	TT	AA	CC	CC
SAB70	M	162.5	3	TT	GA	CC	CC
SAB71	M	151	3	TT	GG	CA	CT
SAB72	F	157.5	3	TC	GA	CC	CC
SAB73	F	157.5	3	TT	GA	CC	CC
SAB74	F	160	3	TT	GA	CC	CC
SAB75	F	165	3	TC	GA	CC	CC
SAB76	M	157.5	3	TT	AA	CA	CT
SAB77	M	165	3	TT	GA	CC	CC
SAB78	M	165	3	TT	AA	CC	CC
SAB79	M	160	3	TT	GA	CC	

				Performance-related mutations (SNPs)			
				Height		Gait	
Sample	Gender	Height (cm)	Gait-ability	BIEC2_808543 (T > C)	BIEC2_1105377 (G > A)	*DMRT3_Ser301STOP* (C > A)	BIEC2_620109 (C > T)
SAB80	F	157.5	3	TT	GG	CC	CC
SAB81	F	155	3	TT	GA	CC	CC
SAB82	F	152.5	3	TT	AA	CC	CC
SAB83	F	155	3	TT	GA	CC	CC
SAB84	F	152.5	3	TT	GA	CC	CC
SAB85	F	157.5	3	TT	GG	CC	CC
SAB86	F	157.5	3	TT	GA	CC	CC
SAB87	F	155	3	TT	GA	CC	CC
SAB88	F	160	3	TT	GA	CC	CC
SAB89	M	157.5	3	TT	GA	CC	CC
SAB90	M	157.5	3	TT	GA	CC	CC
SAB91	F	160	3	TC	AA	CC	CC
SAB92	F	155	3	TT	GG	CC	CC
SAB93	F	155	3	TT	GA	CC	CC
SAB94	F	157.5	3	TT	GG	CC	
SAB95	M	153	3	TT	GA	CC	CC
SAB96	M	151.5	3	TT	AA	CC	CC
SAB97	M	160	3	TT	GG	CC	CC
SAB98	M	157	3	TT	GA	CC	CC
SAB99	M	154.5	3	TT	GA	CC	CC
SAB100	F	153	3	TT	GA	CC	CC

CBP = Cape Boerperd; SAB = SA Boerperd; 3 = three-gaited; 5 = five-gaited; F = female; M = male

Table 2.3 indicates the average height associated with each size-associated genotype present within the breed, to evaluate the contribution of each SNP. Table 2.4 displays the number of gaited individuals with and without the *DMRT3_Ser301STOP* SNP.

Table 2.3 Average height (cm) of each size-associated genotype within the SA Boerperd.

Genotype	TT/GG	TT/GA	TT/AA	TC/GA	TC/AA
Individuals	43	44	8	4	1
Mean height (cm)	151.06	152.94	152.33	158	160
Standard deviation (cm)	4.94	5.43	12.06	9.90	0.00
Height range (cm)	142 - 160	143 - 165	141 - 165	151 – 167.5	NA

NA = not applicable

Table 2.4 Proportions of 3/5-gaited individuals for the SNP *DMRT3_Ser301STOP* (C > A).

Gait ability	3-gaited		5-gaited	
Genotype	CC	CA	CC	CA
Individual count	74	12	5	9

Additionally, the *r* x *c* Fisher's exact test (Mehta & Patel, 1983) was performed on three of the investigated SNPs, represented in Table 2.5. Due to its incomplete SNP data, SNP BIEC2_620109 was not included in the test. The p-values of the three SNPs were significantly different.

Table 2.5 Contingency table input of the SNP data used to perform the Fisher's exact test (www.in-silico.net). Significant values are associated with $p < 0.05$.

BIEC2_808543 (T > C)			BIEC2_1105377 (G > A)			*DMRT3_Ser301STOP* (C > A)		
	<160cm	>160cm		<160cm	>160cm		3-gaited	5-gaited
TT	85	13	GG	42	3	CC	74	5
TC	2	3	GA	39	9	CA	12	9
CC	0	0	AA	6	4	AA	0	0
	p = 0.0497			p = 0.0247			p = 0.0003	

The allele and genotype frequencies for both size- and gait-associative SNPs are shown in Table 2.6. The wild-type alleles of all four SNPs (indicated with *) were found to be predominant. Also, calculations made regarding SNP BIEC2_620109 were done for the 88 horses SNP data could be obtained.

Table 2.6 Distributions of the genotype and allelic frequencies of the four investigated SNPs in the SA Boerperd.

SNP	Genotype	N	Genotype *f*	Allele	Allele *f*
BIEC2_808543	TT	95	0.950	T*	0.975
	TC	5	0.050	C	0.025
	CC	0	0.000		
BIEC2_1105377	GG	43	0.430	G*	0.670
	GA	48	0.480	A	0.330
	AA	9	0.090		
DMRT3_Ser301STOP	CC	79	0.790	C*	0.895
	CA	21	0.210	A	0.105
	AA	0	0.000		
BIEC2_620109	CC	72	0.818	C*	0.909
	CT	16	0.182	T	0.091
	TT	0	0.000		

*Indicate wild-type allele

2.4. Discussion & conclusion

2.4.1. Height at withers

Metzger *et al.* (2013) found that horses across many breeds measuring 160 cm and less at the withers (ridge between the shoulder blades), displayed the highest frequency for the wild-type T-allele when genotyped for the SNP BIEC2_808543. Individuals higher than 160 cm tended to have a higher frequency for the mutated C-allele. Similar results were obtained for the SA Boerperd. The T-allele was shown to be predominant since the breed falls within the medium height range (140 – 160 cm). Those individuals that possessed the mutated allele were found to be of average height, the tallest being 167.6 cm (SAB7). It should however be noted that not all the horses that were above 160 cm had the minor C-allele.

He *et al.* (2015) reported the presence of one C-allele in the Yili breed from a mean height of 142 cm and the CC-genotype in horses of about 158 cm. The height at which the C-allele is present in the latter breed is thus considerably lower than that documented for the SA Boerperd. The height at which the C-allele starts playing a role therefore seems to differ between breeds. Supporting evidence of this was found by Metzger *et al.* (2013), where CT-individuals were present within the Dülmener breed, which had an average height of 130 cm. The same study found the Arabian breed, with an average of 148 cm, fixated for the TT-genotype.

On a molecular level, the presence of the minor C-allele decreases the expression levels of the *LCORL* gene, and thus the transcriptional regulation of the *transcription factor IID* complex that influences size-related processes (Metzger *et al.*, 2013). The mechanism by which BIEC2_808543 effects *LCORL* is still unknown, but it is possible that the exact *LCORL*-level of a heterozygous or CC-individual may differ within and amongst breeds. This could explain why some individuals shorter than 160 cm displayed the C-allele, whilst some taller did not.

The Fisher's exact test revealed that there is a significant difference between the two size ranges, namely taller and shorter than 160 cm. Most horses were smaller than 160 cm in the presence of both height-SNPs (BIEC2_808543 & BIEC2_1105377). In addition, the average height determined for each size-associated genotype revealed that horses with or without the mutated A-allele at SNP BIEC2_1105377 (TT/GA & TT/AA) averaged about 152 cm, and horses with the mutated C-allele at SNP BIEC2_808543 were much closer to 160 cm (TC/GA

& TC/AA). Since only five individuals possessed the latter mutation and three of those were taller than 160 cm, it would not be wise to accept that BIEC2_808543 attributes significantly to size without testing additional tall horses. Furthermore, it could not be established whether BIEC2_808543 attributes 1 cm and BIEC2_1105377 0.5 cm to height, as was found by Signer-Hasler *et al.* (2012).

It should be noted that most of the sampled horses weren't personally measured by the researcher, with the heights mainly provided by the stud owners. It is possible that some of these heights could have been inaccurate due to differing measuring techniques, as well as the fact that some measurements had to be converted from hands to centimetres. Nevertheless, even a possible difference of 2 cm does not explain why SAB11 (151 cm) possessed SNP BIEC2_808543, whilst SAB97 (160 cm) did not.

The effect external factors have on the phenotype of horses should also be considered. Factors such as differing hoof lengths or the presence of horseshoes can affect height at withers, especially since the maximum permitted hoof length for showing purposes is 11 cm, with an additional 10 mm thick horseshoe (SA Boerperd Breeders Society, 2017a). The heights of the sampled horses were not adjusted to compensate for hoof and shoe length, due to the possible inaccuracy of the measurements provided by the owners. Also, the thickness and shape of a horse's hoof can contribute to its height. Other environmental factors such as the quality and type of feed, as well as the hardness of the ground, can influence hoof shape (Life Data Labs, Inc., 2017). It could happen that a horse may possess the SNP BIEC2_808543, but due to lack of exercise or bad stabling conditions may be under-developed, and thus shorter than expected.

The frequency of the mutant BIEC2_1105377 allele (A) was found to be much higher than the minor BIEC2_808543 allele (C). It also did not seem as if the presence of one or two of the A-alleles within an individual was dependant on the presence of the minor C-allele (BIEC2_808543) in the same individual. This indicates that the SNPs are inherited independently of one another, most likely since they occur on different chromosomes. It was interesting to note that the T-allele (BIEC2_808543) was always accompanied by at least one A-allele (BIEC2_1105377). Though this may purely be coincidence. Also, the A-allele did not indicate any sole association with heights larger than 160 cm, since the shortest individual

with the SNP was a medium height of 141 cm (SAB13). Some horses measuring above 160 cm for this SNP did not display the minor allele, as was similarly found with BIEC2_808543.

According to Signer-Hasler *et al.* (2012) 31 autosomes (38 124 SNPs) explain 70.2% of estimated breeding values (EBVs), including size. Of these autosomes, ECA 3 (11.6%) and ECA 9 (7.4%) were shown to have the largest effect on size, though it is unclear which chromosomes and SNPs explain the additional 81%. The latter study also found that ECA 1 has a strong association with size and may explain some of the variation. Unfortunately, specific SNPs could not be identified on the latter chromosome. Since these findings were made solely within the Franches-Montagnes breed, it is possible that some of the other autosomes and associative SNPs, or even particular SNPs on ECA 1, may attribute the most to size in other horse breeds.

The study of Makvandi-Nejad *et al.* (2012) was particularly thorough by investigating size in sixteen equine breeds, though less than five individuals were tested within most breeds. In contrast to Signer-Hasler *et al.* (2012), these authors found that the BIEC2_808543 SNP explained 68.5% of size variance, although this SNP was solely found within heavy, cold-blooded breeds like the Percheron and Clydesdale. Four SNPs, one each on ECA 3, 6, 9 and 11, were ultimately shown to attribute 83.5% to size. One of the latter loci also occur within the *ZFAT* gene (ECA 9), though the SNP is not the same as that investigated by Signer-Hasler *et al.* (2012) and in the current study.

Overall, it is thus uncertain what the exact level of influence both BIEC2_808543 and BIEC2_1105377 have on height at withers on their own. Based on the findings of the current study, these two SNPs alone appear not to be definitive predictors of height in the SA Boerperd due to the lack of a clear correlation. It is thus reasonable to suspect that combinations of strongly associated and linked SNPs, which form different haplotypes, may contribute more to height and be a better indication of the size of a horse. Such haplotypes were identified in both the *LCORL* and *HMGA2* genes (Makvandi-Nejad *et al,* 2012), though the exact measurements of the studied individuals weren't documented to determine the theory's validity.

A recent study investigating the skeletal variation within the Tennessee Walking Horse found 17 additional SNPs and an indel associated with size on ECA 3 (Staiger *et al.,* 2016), and

proposed that the causative SNP may in truth be downstream of the SNP BIEC2_808543, within the coding region of *LCORL*. Yet it appears that ECA 9 contains a homologous 5kb retrogene copy of *LCORL's* coding sequence, which may complicate the identification of the causal SNP and other possible associative SNPs. Future studies will thus need to be conducted to locate this causal SNP, as well as conclusively determine which genes have the largest effect on size within the SA Boerperd. Such studies can benefit from accurate conformation measurements, like the height at withers and body length of each individual horse, as was done by Staiger *et al.* (2016).

2.4.2. Gaitedness

Most gaited breeds are fixed for the A-allele of *DMRT3_Ser301STOP* (Andersson *et al.,* 2012; Promerová *et al.,* 2014). The SA Boerperd however was found to have a very low frequency of 10.5% for this allele. A similar value was calculated by Promerová *et al.* (2014) for the Boer Pony, which possibly includes both the SA and Cape Boerperd. Also, none of the tested individuals were found to possess the AA-genotype, as was also the case within the current study. Furthermore, the Fisher's exact test also indicated a significant difference between the two gait categories, where most horses were three-gaited. It would thus seem that the SA Boerperd is not a true gaited breed due to the low frequency of the A-allele and the breed's bias towards being three-gaited.

The partiality of three-gaited SA Boerperd individuals can be explained with selective breeding. Presently there is a very small market for gaited individuals within the SA Boerperd breeders community, since fewer five-gaited classes are hosted per show than three-gaited ones (SA Boerperd Breeders Society, 2017b). Generally, hosted events such as jumping, endurance riding and harness classes require three-gaited individuals. A lot of time and money are thus invested in creating five-gaited horses, when most buyers are interested in three-gaited eventers. Furthermore, individuals selected to be five-gaited must undergo a specialised training regime to ensure the *rack* and *short-gait* become natural, flawless movements, which few trainers can achieve (Louw, 2008). It has also been mentioned that gaited individuals seldom choose to trot of their own volition, preferring the less exertive *rack* (Joubert, personal communication). When these gaited horses do trot at liberty, the movement seems uneven and forced which creates an unpleasant display for judges. They

are thus required to be ridden in such a way that the *trot* displays fast with longer and more forceful strides (SA Boerperd Breeders Society, 2017c).

Apart from being strongly associated with ambling gaits, Andersson *et al.* (2012) also showed that *DMRT3_Ser301STOP* occurs in pacer and trotter breeds used for racing, such as the Standardbred and French Trotter. This suggests that the *gait keeper* mutation positively affects the speed of symmetrical gaits like the *trot,* and discourages the switch to the general asymmetric high-speed *gallop* (Promerová *et al.*, 2014). The SA Boerperd can achieve a naturally quick pace during the *trot* (personal observation), making it reasonable to suspect that the *gait keeper* mutation benefits the *trot* when such individuals are not trained to display the *rack* or *short-gait*. This, however, should be proven by monitoring the trotting ability of the three-gaited heterozygous individuals. Such individuals may also make excellent harness horses.

One or both gait-associative SNPs may also allow an individual to become disjointed or disunited during the *canter,* which refers to the desynchronization of the hindquarters with the forequarters (De Jong, 2017). This desynchronization can occur when an individual is imbalanced due to its body's natural asymmetry or muscle soreness (De Jong, 2012). The association of the two SNPs with this phenomenon is grounded on the researcher's own experience with CBP2, which was shown to possess both mutations and regularly become disunited whilst cantering. This seems plausible given the neurons controlled by *DMRT3* coordinate the movement of the fore- and hind limbs (Andersson *et al.,* 2012). Individuals that can disunite often favour leading the canter with their stronger side, which complicates training them to lead on both sides. Such individuals may also fail as dressage horses, since this discipline requires controlled and well-timed lead changes (De Ruffieu, 2015). More research is however needed to verify the association of this phenomenon with either gait-SNP.

Based on the findings of this study, it seems that *DMRT3_Ser301STOP* and BIEC2_620109 are co-inherited, which is plausible since the two SNPs are significantly linked (Promerová *et al.,* 2014). The frequencies for both SNPs would have been identical in the current results, if not for the lack of some data of BIEC2_620109, and genotypes consisting of both wild-type alleles (C-C) and both mutant alleles (A-T) were present. Promerová *et al.* (2014) found that the mutant allele of either SNP did however occur along with the wild-type allele of the other,

though at very low frequencies. Only 0.3% of individuals across various breeds had the SNP *DMRT3_Ser301STOP* (A) and the wild-type C-allele at BIEC2_620109.

Also of significance was that some five-gaited individuals (SAB9, SAB40, SAB63, SAB64 & SAB68) tested negative for the *gait keeper* mutation. Similar findings were made by Promerová *et al.* (2014), who theorised that either another SNP greatly affects gaitedness, or that multiple genes affect the trait. It was also speculated that the strong selection for gaitedness over many generations may have increased the occurrence of mutations that possibly enable the ability in horses homozygous for the wild-type *DMRT3_Ser301STOP* allele. This may be true for the Icelandic Horse, which has intensively been bred for gaitedness (Kristjansson *et al.,* 2014), but this is not the case in the SA Boerperd. Despite the latter inconsistency, it does however seem that either *DMRT3_Ser301STOP* or BIEC2_620109 can be used to determine the five-gaited potential of a SA Boerperd individual. Horses possessing these SNPs can be bred to produce gaited offspring, based on the observation that SAB49 (CA) is the sire of SAB71 (CA). Though the inheritance of these SNPs should further be investigated due to the absence of these mutations in a few gaited individuals.

Interestingly, none of the horses possessing the *gait keeper* mutation had the size-related BIEC2_808543 SNP, nor vice versa. It is thus possible that gaitedness is influenced by size. Similar findings were made within the Tennessee Walking Horse (Staiger *et al.*, 2016), where lateral-gaited individuals had a larger build than multi-gaited ones. Individuals thus able to perform additional gaits were shown to be of moderate height with long front pasterns. If this is similarly true for the SA Boerperd, then the breed can simply be screened for the *gait keeper* mutation in future where its presence may be indicative of a foal's future medium height.

REFERENCES

Andersson, L.S., Larhammar, M., Memic, F., Wootz, H., Schwochow, D., Rubin, C., 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 in mice. *Nature,* 488: 642 – 646.

Baban, M., Curik, I., Antunovic, B., Cacic, M., Korabi, N. and Mijic, P. (2009). Phenotypic correlations of stride traits and body measurements in Lipizzaner stallions and mares. *Journal of Equine Veterinary Science,* 29(6): 513 – 518.

Bekker, L. (2012). Vyfgang. In: *SA Boerperd: Die perd vir die toekoms:* 31 - 34. Edited by S. Louw. Bloemfontein: SA Boerperdtelersgenootskap.

BIOER Technology. (2015). *Spin column extraction kits – Biospin PCR Purification Kit*. Available from: http://www.bioer.com.cn/en/html/productcenter/reagents/centrifugal/142.html#ad-image-0 [Accessed 24 Nov. 2016].

Broad Institute. (2008). FTP Directory: Distribution horse SNP release. Available from: ftp://ftp.broadinstitute.org/distribution/horse_snp_release/v2/ [Accessed 4 June 2016].

De Jong, M. (2012). *An introduction to straightness training*. 2nd edition. Straightness Training. Available at: http://straightnesstraining.com/wp-content/uploads/2013/02/EBRR-free-eBook-EN-20.pdf [Accessed 14 Sept. 2017].

De Jong, M. (2017). *Straightness Training: Horse gaits & gaited horses*. Straightness Training. Available at: http://straightnesstraining.com/the-horse/the-body-of-the-horse/horse-gaits/ [Accessed 13 Sept. 2017].

De Ruffieu, F.L. (2015). *The handbook of riding essentials*. Franktown: Xenophon Press.

Du Toit, K. (2010). *Boerperd: Die geskiedenis van die S.A. Boerperd*. Pretoria: West Publishers.

Equestrian. (2017). *Five popular disciplines of the South African Equestrian Federation – Dressage, show jumping, endurance ride, eventing and equitation.* Available at: https://www.equestrian.co.za/116-equestrian-horse-disciplines/19316-five-popular-disciplines-of-the-south-african-equestrian-federation-dressage-show-jumping-endurance-ride-eventing-and-equitation [Accessed 7 Feb. 2017].

Geneious. (2017). *Features*. Available at: https://www.geneious.com/features/ [Accessed 22 Nov. 2017].

Han, H., Zeng, L., Dang, R., Lan, X., Chen, H. and Lei, C. (2015). The DMRT3 gene in Chinese horse breeds. *Stichting International Foundation for Animal Genetics*, 46: 340 – 342.

He, S., Zhang, L., Li, W. and Liu, M. (2015). BIEC2-808543 SNP in the *LCORL* gene is associated with body conformation in the Yili horse. *Animal Biotechnology,* 26: 289 – 291.

KAPA Biosystems. (2016a). *KAPA HiFi PCR Kits*. Available from: https://www.kapabiosystems.com/product-applications/products/pcr-2/kapa-hifi-pcr-kits/ [Accessed 23 Nov. 2016].

KAPA Biosystems. (2016b). *KAPA2G Robust PCR Kits*. Available from: https://www.kapabiosystems.com/product-applications/products/pcr-2/kapa2g-robust-pcr-kits/ [Accessed 28 July 2017].

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P. and Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12): 1647 – 1649.

Kristjansson, T., Bjornsdottir, S., Sigudsson, A., Andersson, L.S., Lindgren, G., Helyar, S.J., Klonowski, A.M. and Arnason, T. (2014). The effect of the 'gait keeper' mutation in the *DMRT3* gene on gaiting ability in Icelandic horses. *Journal of Animal Breeding and Genetics,* 131: 415 – 425.

Life Data Labs, Inc. (2017). *Factors that affect hoof quality*. Available at: https://www.lifedatalabs.com/articles/18-nutritional-articles/177-factors-that-affect-hoof-quality-extended [Accessed 13 Sept. 2017].

Louw, S. (2008). SA Boerperd – Indigenous breed to southern Africa. In: *SA Boerperd: The ideal sport and pleasure horse:* 22 - 29*.* Edited by S. Louw. Maanhaarrand: Kejafa Knowledge Works.

Makvandi-Nejad, S., Hoffman, G.E., Allen, J.J., Chu, E., Gu, E., Chandler, A.M., Loredo, A.I., Bellone, R.R., Mezey, J.G., Brooks, S.A. and Sutter, N.B. (2012). Four loci explain 83% of size variation in the horse. *PLoS ONE,* 7(7): e39929. doi: 10.1371/journal.pone.0039929

McDonald, J.H. (2014). Handbook of Biological Statistics, 3rd edition. Maryland: Sparky House Publishing.

Mehta, C.R. and Patel, N.R. (1983). A network algorithm for performing Fisher's exact test in *r* x *c* contingency tables. *Journal of the American Statistical Association*, 78(382): 427 – 434.

Metzger, J., Schrimpf, R., Philipp, U. and Distl, O. (2013). Expression levels of *LCORL* are associated with body size in horses. *PLoS ONE,* 8(2): e56497. doi: 10.1371/journal.pone.0056497

Promerová, M., Andersson, L.S., Juras, R., Penedo, M.C.T., Reissmann, M., Tozaki, T., Bellone, R., Dunner, S., Hořín, P., Imsland, F., Imsland, P., Mikko, S., Modrý, D., Roed, K.H., Schwochow, D., Vega-Pla, J.L., Mehrabani-Yeganeh, H., Yousefi-Mashouf, N., Cothran, E.G., Lindgren, G. and Andersson, L. (2014). Worldwide frequency distribution of the '*Gait keeper*' mutation in the *DMRT3* gene. *Stichting International Foundation for Animal Genetics,* 45: 274 – 282.

SA Boerperd Breeders Society. (2016). *SA Boerperd as show horse and sport horse*. Available at: http://www.saboerperd.com/p26/horses/sa-boerperd-as-show-and-sport-horse.html [Accessed 25 Jan. 2017].

SA Boerperd Breeders Society. (2017a). *General Show Practice & Etiquette*. Available at: http://www.saboerperd.com/images/photos/2017-C5-GeneralShowPracticeAndEtiquette.pdf [Accessed 13 Sept. 2017].

SA Boerperd Breeders Society. (2017b). *Show Classes (Comprehensive & Minimum) June 2017,* 2nd ed. [pdf] Available at: http://www.saboerperd.com/images/photos/2017-C4-2-Meesterklaslys-Junie2017a.pdf [Accessed 13 Sept. 2017].

SA Boerperd Breeders Society. (2017c). *Manual: Traditional 3G, 5G & Show Horse Classes,* 2nd edition. Available at: http://www.saboerperd.com/images/photos/2017-C11-Manual-Traditional3G5HShowHorseClasses.pdf [Accessed 13 Sept. 2017].

Signer-Hasler, H., Flury, C., Haase, B., Burger, D., Simianer, H., Leeb, T. and Rieder, S. (2012). A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PLoS ONE,* 7(5): e37282. doi: 10.1371/journal.pone.0037282

Staiger, E.A., Al Abri, M.A., Pflug, K.M., Kalla, S.E., Ainsworth, D.M., Miller, D., Raudsepp, T., Sutter, N.B. and Brooks, S.A. (2016). Skeletal variation in Tennessee Walking Horses maps to the *LCORL/NCAPG* gene region. *Physiological Genomics,* 48: 325 – 335.

Stothard, P. (2000). The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques*, 28: 1102 – 1104.

Swart, S. (2010). *Riding High: Horses, Humans and History in South Africa*. Johannesburg: Wits University Press.

Taberlet, P., Waits, L.P. and Luikart, G. (1999). Noninvasive genetic sampling: Look before you leap. *Trends in Ecology & Evolution*, 14(8): 323 – 327.

Thermo Fischer Scientific. (2016). *BigDye® Terminator v3.1 Cycle Sequencing Kit.* Available at: https://www.thermofisher.com/order/catalog/product/4337455 [Accessed 24 Nov. 2016].

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012). Primer3 – New capabilities and interfaces*. Nucleic Acids Research*, 40(15): 1 – 12.

Zymo Research. (2016). *ZR DNA Sequencing Clean-up Kit™.* Available from: https://www.zymoresearch.com/dna/dna-clean-up/dna-sequencing-clean-up/zr-dna-sequencing-clean-up-kit [Accessed 24 Nov. 2016].

Personal Communication:

Jan Joubert, Vice President of the SA Boerperd Breeders Association.

Chapter 3: Genetic diversity of the SA Boerperd

(Picture obtained from the SA Boerperd Breeders Society; www.saboerperd.com)

(Photo by Nadia Breytenbach)

3.1. Introduction

Like all species, horses need to be genetically diverse to ensure their overall fitness and continued existence. Genetic diversity is construed as the presence of different alleles and genotypes in a population (Frankham *et al.,* 2002), and a great variation of these safeguard a population's ability to survive harmful environmental changes (Juras, Cothran & Klimas, 2003).

In South Africa, climate change poses a danger to animals (Farmer's Weekly, 2017). This phenomenon brings floods and drought, which threaten forage crops and cause the increase of vector-borne diseases. All horses that exist in the country, including the SA Boerperd, are exposed to such threats. Conserving the genetic diversity of this breed is not only important for the benefit of the equine population, but also because it benefits the country. The SA Boerperd fulfils a notable economic role by contributing to agriculture (Nel, 2014), tourism (Equus Horse Safaris, 2016; Pakamisa, 2016), and the showing industry (Louw, 2008; Du Toit, 2010).

The breeding books of the SA Boerperd have been officially closed since 1998 (Swart, 2010; Nel, 2014) and all existing individuals of the breed are the progeny of 11 founding bloodlines (Louw, 2008; Du Toit, 2010). No new genetic material has thus been added to the breed in roughly twenty years. This is not necessarily disadvantageous, but it can increase the probability of inbreeding and the loss of allelic diversity if the breeding of such a closed population is not well managed (Lacy, 1987). By genetically characterizing a breed, its allelic richness, among other essential information, can be determined. This can aid the conservation of the breed (Bjørnstad & Røed, 2002) by helping improve breeding strategies.

Microsatellite markers, or short tandem repeat markers (STRs), have been proven to be effective when evaluating genetic diversity. These markers are less likely to misalign due to their repetitive nature (Goldstein & Schlötterer, 1999) and are thus reliable when determining genotypes. They are also highly polymorphic and are inherited in a co-dominant fashion, which contribute to a high heterozygosity (Takezaki & Nei, 1996). Some equine breeds that have successfully been investigated by these markers include Spanish Celtic breeds (Cañon *et al.,* 2000), the Ruertas horse (Sereno *et al.,* 2008), and the Canadian horse (Khanshour *et al.,* 2015).

Only one study to date has determined the diversity of the SA Boerperd by using microsatellites (Botha, 2002), and the latter study revealed that the breed is highly diverse. This is generally expected from a newly established breed with a large founder population (Allendorf, Luikart & Aitken, 2013). After years of diligent breeding certain phenotypic characteristics for the admission to the studbook (Louw, 2008), it is possible that the level of diversity of the SA Boerperd may have decreased. To address the latter concern, the services of stud stallions are available to breeders at annual auctions (Langenhoven, personal communication). This exchange of genetic material may keep the gene pool diverse, and avoid the formation of subpopulations within the larger population (Allendorf, Luikart & Aitken, 2013). Nevertheless, breeder bias in selecting stud stallions with certain phenotypes may very well be attributing to subpopulation formation (Joubert, personal communication).

The aim of this study was to determine the level of genetic diversity, and screen for possible inbreeding, within the SA Boerperd, as well as to evaluate whether subpopulations have formed within the breed. Additionally, the genetic markers of some of the Boerperd's ancestral breeds were compared to the SA Boerperd to evaluate their relatedness and compare their levels of genetic diversity. With such knowledge, the sustainability of the SA Boerperd can be assessed, and well-informed decisions can be made concerning its breeding program.

3.2. Materials and methods

3.2.1. Data collection

Microsatellite data of 363 horses were obtained from Unistel Medical Laboratories (UML), which does routine identification and parentage verification tests for the SA Boerperd Breeder's Association. The studs that gave permission for the utilization of their horses' DNA data are indicated in Figure 3.1. The microsatellite panel used by UML consists of 17 loci (AHT4, AHT5, ASB2, ASB17, ASB23, CA425, HMS1, HMS2, HMS3, HMS6, HMS7, HTG4, HTG6, HTG7, HTG10, LEX3 & VHL20). This panel is commercially available through the StockMarks® for Horses Equine 17-plex Genotyping Kit (Applied Biosystems), as well as the EquineGenotypes™ Panel 1.1 (Finnzymes Diagnostics) (Van de Goor, Panneman and van Haeringen, 2009).

Data obtained by Botha (2002) for 81 SA Boerperd horses were also included in this study for comparison with the present data. The latter author investigated 12 microsatellite markers during her study, and of those only nine were found to be significantly comparable with the data obtained from UML. Additionally, the 17-loci panel data used by Van de Goor, van Haeringen & Lenstra (2011) for 9 of their 35 studied equine populations were also included for diversity investigations.

All allele data used in the current study were converted from the International Society for Animal Genetics (ISAG) alphabetic nomenclature to the accepted repeat-based nomenclature, as suggested by Van de Goor, Panneman & van Haeringen (2009). This was to ensure the uniformity of the data, as well as to make the data compatible with the statistical programs used. The conversion between the two nomenclatures for the 17-panel microsatellites are shown in Table 3.1

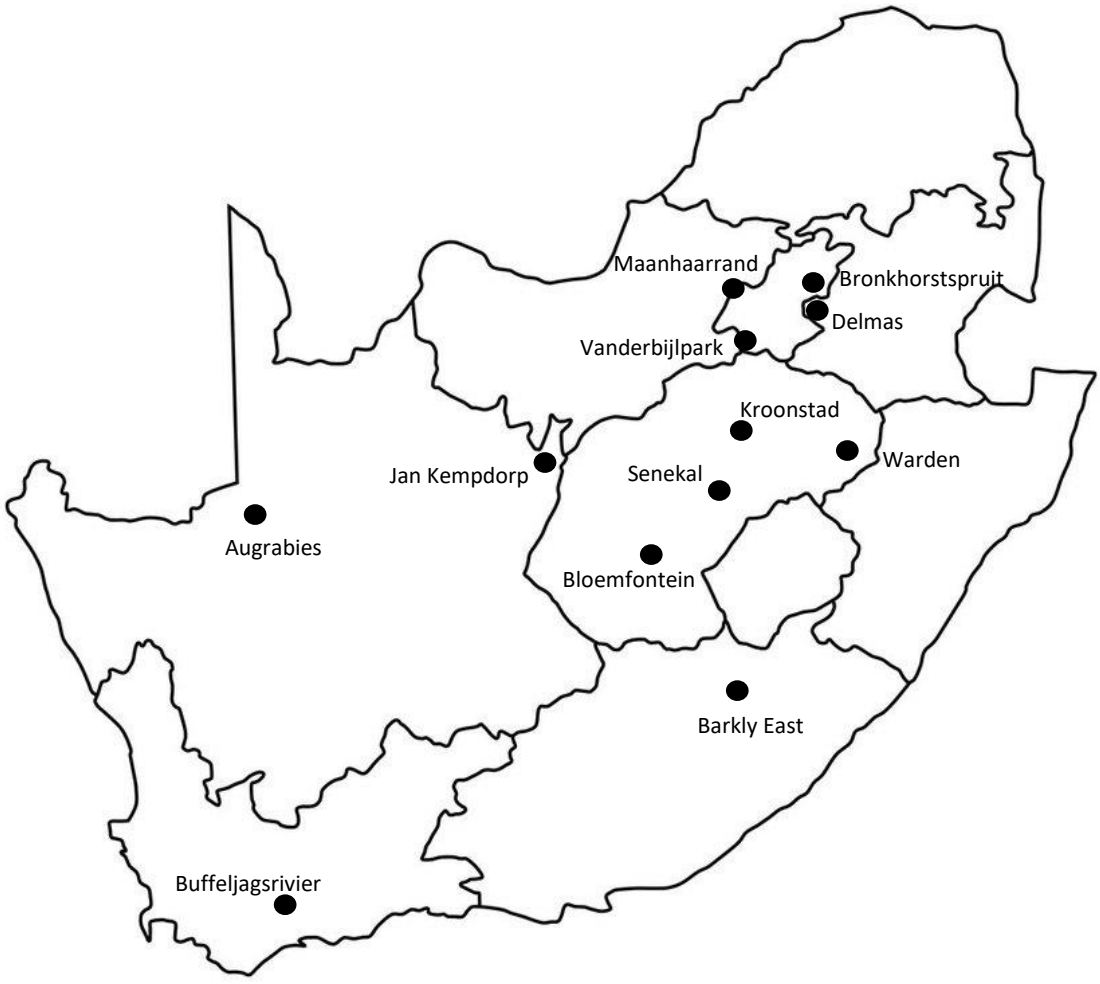


Figure 3.1 South African map indicating the 12 stud localities from which microsatellite data were obtained.

Table 3.1 The conversion of ISAG (alphabetical) nomenclature to the number of repeat nomenclature (Adapted from Van de Goor, Panneman & van Haeringen (2009)).

Marker	AHT4	AHT5	ASB2	ASB17	ASB23	CA425	HMS1	HMS2	HMS3	HMS6	HMS7	HTG4	HTG6	HTG7	HTG10	LEX3	VHL20
B			9														
C			10														
D			11	11													
E			12	12													
F			13	13		12										13	
G			14	14	15	13							12			14	
H	25	14	15	15	16	14		15	20				13		16	15	
I	26	15	16	16	17	15	14	16	21				14		17	16	13
J	27	16	17	17	18	16	15	17	22		16		15		18	17	14
K	28	17	18	18	19	17	16	18	23	13	17	30	16	15	19	18	15
L	29	18	19	19	20	18	17	19	24	14	18	31	17	16	20	19	16
M	30	19	20	20	21	19	18	20	25	15	19	32	18	17	21	20	17
N	31	20	21	21	22	20	19	21	26	16	20	33	19	18	22	21	18
O	32	21	22	22	23	21	20	22	27	17	21	34	20	19	23	22	19
P	33	22	23	23	24	22	21	23	28	18	22	35	21	20	24	23	20
Q	34	23	24	24	25		22	24	29	19	23	36	22		25	24	21
R	35		25	25	26			25	30						26		22
S			26	26	27			26	31						27		
T				27	28			27							28		
U				28	29			28									
V				29	30												
W				30													

3.2.2. Statistical analysis

3.2.2.1. Data organization

Microsatellite marker data for 10 equine breeds, which include the SA Boerperd, were entered into Microsoft Office Excel and analysed by the GenAlEx v.6.503 (Peakall & Smouse, 2006; 2012) add-in. This software was used to determine allelic variability and frequencies, as well as to calculate other measures of genetic diversity (see below). Input files were also created with this program, along with CONVERT v.1.31 (Glaubitz, 2004), for use in other statistical programs.

3.2.2.2. Genetic diversity

Genetic diversity is most often defined as the occurrence of multiple alleles at a given locus (Frankham *et al.,* 2002), which enables populations to adapt to environmental changes and selective pressures. A variety of alleles also ensures that this adaptive ability gets passed on to future generations (Groeneveld *et al.,* 2010). The loss of alleles, and thus diversity, is primarily caused by inbreeding. This phenomenon can occur within small populations, as well as within breeds with a closed register, such as the SA Boerperd. Monitoring the genetic diversity of this breed is thus vital to ensure its continued existence.

Genetic diversity can be measured by determining multi-locus heterozygosity and allelic richness (Allendorf, Luikart & Aitken, 2013). The former refers to the number of heterozygotes in the sampled population, whilst the latter indicates the number of different alleles present in the same group. High values for both these measures indicate high genetic diversity. In addition, heterozygosity can be classified as the number of heterozygote individuals observed (H_o), expected (H_e), and unbiased by taking the sample size into consideration (H_z). Also of importance is determining the mean number of alleles (N_a) per locus and the allelic frequency at each locus (Ojango *et al.,* 2011). Seeing as sample sizes are rarely the same for each population, and the distribution of alleles also tend to differ, the effective number of alleles (N_e) should also be calculated since it compensates for these biases. The N_a, N_e and heterozygosity of the SA Boerperd were all calculated using GenAlEx.

3.2.2.3. Hardy-Weinberg Equilibrium

The Hardy-Weinberg Equilibrium (HWE) principle calculates the relationship between gene and genotype frequencies. This principle assumes that both observed heterozygosity (H_o) and expected heterozygosity (H_e) will remain constant between generations in the absence of genetic drift, mutations or immigration of individuals into the population (Allendorf, Luikart & Aitken, 2013). Deviations from HWE could be due to inbreeding or population stratification when homozygotes are in excess, or due to genotyping errors if the proportion of heterozygotes are high (Wigginton, Cutler & Abecasis, 2005). GenAlEx was also used to test whether the different populations of the SA Boerperd were in HWE.

3.2.2.4. Genetic differentiation and drift

Genetic differentiation refers to the difference in allele frequencies between subpopulations due to the subdivision of the larger population. This subdivision often leads to the decrease in heterozygosity and fixation of certain alleles, and can be measured with Wright's F-statistic (Wright, 1965). This statistic determines the mean heterozygosity of each population through direct count, as well as the potential heterozygote frequency if all individuals in a subpopulation assorts independently (Hartl & Clark, 1997).

The F-statistic encompasses three inbreeding coefficients, namely F_{ST}, F_{IS} and F_{IT}. These statistic measures determine the divergence of allele frequencies among subpopulations (F_{ST}), within subpopulations (F_{IS}) and within the total population (F_{IT}) (Allendorf, Luikart & Aitken, 2013). F_{ST} is used most often since it can highlight the occurrence of drift and inbreeding, but it can also indicate which loci are under selection (Beaumont & Balding, 2004). Values of F_{ST} are generally bounded between zero and one (Meirmans & Hedrick, 2011). High F_{ST} values can indicate loci that are under divergent selection, whilst those under balancing selection produce low values (Lewontin & Krakauer, 1973). Wright's F-statistic values were calculated with ARLEQUIN v.3.5.2.2 (Excoffier, Laval & Schneider, 2005).

3.2.2.5. Individual assignment

The highly diverse nature of microsatellite markers makes them ideal for investigating the differentiation of individuals, which is accomplished by comparing the variance in allele frequencies at a locus (Avise, 2004). The program STRUCTURE v.2.3.4 (Pritchard *et al.,* 2000) is often used to determine such population differentiation or structure. This software calculates and assigns probabilities for an individual to have originated in a particular group by using genotypic data. The program implements a Bayesian-based assignment approach, whereby individuals are divided into theoretical clusters (K), and also depicts the admixture proportions (if any) within individuals.

STRUCTURE was thus run to determine whether subpopulations had formed within the SA Boerperd. Possible K-values were set between 1 and 12, seeing as the genomic data of twelve studs were available. Five independent runs were repeated for each K-value, and a burn-in period of 20,000 steps followed by 100,000 MCMC (Markov Chain Monte Carlo) iterations were used. The MCMC process estimates variant frequencies for each possible

number of clusters, and then individuals are re-assigned to another cluster/group based on the frequency estimates (Porras-Hurtado *et al.,* 2013). The output of this software was then entered into Structure Harvester v.0.6.94 (Earl & von Holdt, 2012), so as to determine the true value of K by measuring deltaK (ΔK; Evanno *et al.,* 2005) from -Ln probability values.

3.3. Results

3.3.1. Within-breed analysis of the SA Boerperd

To protect the anonymity of the breeders and horses involved in this study each stud was given a new reference, namely Stud1 – Stud12. Table 3.2 shows the genetic diversity of these studs. The mean number of alleles (N_a) over all 12 studs was found to range from 3.471 (Stud8) to 6.529 (Stud7). It should be noted that Stud8 consisted of only 3 individuals, thus the effective number of alleles (N_e) was a much more accurate indicator for assessing the allelic richness. Taking sample size into consideration, the N_e-values showed that Stud2 had the most alleles per locus (3.937), whilst Stud9 had the least (2.973). Of all the studs, Stud12 was shown to have the highest observed heterozygosity (H_o) at 0.735, and the rest of the studs all had values higher than 0.640. Unbiased expected heterozygosity (H_z) was highest for Stud2, which correlates with its high N_e-value. Despite the small sample sizes, it seems that all the studs are exceptionally diverse.

The inbreeding coefficient that measures diversity within subpopulations (F_{IS}) can have values that range between -1 and 1. Negative values are indicative of a heterozygote excess and positive values indicate a deficiency (Allendorf, Luikart & Aitken, 2013). It was found that Stud12 was the most outbred stud with a value of -0.118. The even lower value of Stud8 was ignored due to its small sample size and obvious bias. The stud proven to be most inbred was Stud6 (0.074) with an inbreeding coefficient of 7.4%. The values of the other studs were all close to zero, indicating an almost equal ratio of heterozygotes to homozygotes.

Table 3.2 Genetic diversity of 12 SA Boerperd populations based on 17 microsatellite markers. The calculated parameters were number of alleles (N_a), number of effective alleles (N_e), observed heterozygosity (H_o), unbiased heterozygosity (H_z) and inbreeding coefficient (F_{IS}). The highest value of each parameter is in bold.

Stud	N	N_a (SD)	N_e (SD)	H_o (SD)	H_z (SD)	F_{IS}
ST1	33	5.529 (0.286)	3.116 (0.230)	0.689 (0.040)	0.658 (0.028)	-0.061
ST2	23	6.176 (0.395)	**3.937 (0.372)**	0.680 (0.043)	**0.727 (0.026)**	0.050
ST3	31	5.353 (0.383)	3.141 (0.252)	0.649 (0.040)	0.654 (0.032)	-0.003
ST4	35	5.765 (0.338)	3.487 (0.311)	0.643 (0.034)	0.686 (0.028)	0.043
ST5	14	4.706 (0.329)	3.063 (0.260)	0.683 (0.053)	0.652 (0.039)	-0.082
ST6	61	6.471 (0.463)	3.876 (0.320)	0.655 (0.036)	0.716 (0.026)	0.074
ST7	29	**6.529 (0.385)**	3.643 (0.316)	0.713 (0.039)	0.696 (0.036)	-0.048
ST8	3	3.471 (0.286)	2.977 (0.302)	0.725 (0.065)	0.722 (0.052)	-0.215
ST9	12	4.588 (0.243)	2.973 (0.188)	0.681 (0.045)	0.667 (0.028)	-0.063
ST10	26	5.706 (0.460)	3.511 (0.337)	0.646 (0.040)	0.682 (0.033)	0.023
ST11	16	4.941 (0.369)	3.317 (0.243)	0.685 (0.034)	0.691 (0.027)	-0.028
ST12	80	6.353 (0.461)	3.276 (0.276)	**0.735 (0.039)**	0.662 (0.031)	-0.118

ST = stud; SD = standard deviation

Significant deviations from the expected HWE were found in eleven of the seventeen tested loci (Table 3.3), and most loci only deviated for one or two populations. HTG6 and LEX3 were the only two loci to deviate for three or more populations. Overall, two studs deviated significantly from HWE, where Stud7 deviated at four loci and Stud12 at seven. This could be due to genotyping errors, seeing as a large amount of data was missing at LEX3 for Stud12. It is also probable that these deviations are due to outbreeding, seeing as the H_o of each stud was higher than the H_e. In addition, the highly diverse N_a and F_{IS}-values of these studs also supported the deviations.

Table 3.3 Conformation to expected Hardy-Weinberg Equilibrium of genotypes for 17 loci within 12 SA Boerperd studs. Loci deviate from HWE when p < 0.05, and these values are emphasised in bold.

Stud / Locus	ST1	ST2	ST3	ST4	ST5	ST6	ST7	ST8	ST9	ST10	ST11	ST12
AHT4	0.982	0.888	0.213	0.198	0.115	0.923	0.396	0.392	0.473	0.074	0.233	0.872
AHT5	0.523	0.507	0.409	0.314	0.221	0.813	0.989	0.386	0.154	0.872	0.586	0.497
ASB2	0.379	0.923	0.789	0.095	0.589	0.829	0.601	0.609	0.801	0.114	0.314	0.083
ASB17	0.863	0.638	0.863	0.186	0.915	0.744	0.228	0.451	0.739	0.431	0.244	**0.030**
ASB23	0.753	0.700	0.779	0.546	0.826	0.343	0.584	0.387	0.772	0.447	**0.032**	**0.000**
CA425	0.285	0.128	0.269	0.387	0.808	0.500	0.693	0.532	0.575	0.707	0.484	0.477
HMS1	0.973	0.961	0.489	0.167	0.515	0.787	**0.017**	0.506	0.061	0.430	0.227	**0.022**
HMS2	0.835	0.807	0.966	0.411	0.772	0.207	**0.003**	0.387	0.883	0.539	0.789	**0.023**
HMS3	0.290	0.825	0.760	0.127	0.387	0.400	0.560	0.609	0.813	0.985	0.309	**0.014**
HMS6	0.689	0.761	0.138	0.907	0.761	0.887	0.643	0.729	0.658	0.966	0.740	0.837
HMS7	0.467	**0.000**	0.797	**0.000**	0.484	0.166	0.791	0.729	0.172	0.929	0.204	0.970
HTG4	0.891	0.708	**0.047**	0.820	0.698	0.955	0.919	0.392	0.825	0.095	0.815	0.196
HTG6	0.612	**0.000**	0.346	0.931	0.451	0.135	0.893	0.506	0.850	**0.003**	0.573	**0.001**
HTG7	0.833	0.176	0.675	0.353	1.000	0.253	0.909	0.729	0.228	0.266	0.521	0.840
HTG10	0.990	0.812	0.646	0.609	0.496	0.086	0.316	0.532	0.877	**0.000**	0.551	**0.018**
LEX3	**0.001**	0.107	0.251	**0.000**	0.109	**0.000**	**0.023**	0.174	0.583	**0.000**	0.184	0.923
VHL20	0.574	0.903	0.766	0.388	0.673	0.948	**0.031**	0.506	0.975	0.255	**0.021**	0.402

ST = stud

The pairwise population F_{ST}-values and p-values for the twelve studs are presented in Table 3.4. The overall genetic differentiation of the SA Boerperd studs were calculated as F_{ST} = 0.061 with a Bonferroni correction of p = 0.001. Significant differentiation would then be defined as p ≤ 0.001. All pair-wise combinations between Stud8 and the rest of the studs, with the exception of Stud1, show F_{ST}-values and p-values that do not support the hypothesis of significant differentiation. This could however be influenced by the small sample size of Stud8. Based on the F_{ST}- and p-values most of the studs were significantly differentiated from one another. The largest genetic difference based on F_{ST}-values occurred between Stud11 and Stud12 at 0.119, and the second largest between Stud1 and Stud11 at 0.113.

Table 3.4 Pairwise population genetic differentiation (F_{ST}-values) between 12 SA Boerperd studs. P-values are indicated within brackets. Significant differentiation is emphasised in bold.

Stud	Stud1	Stud2	Stud3	Stud4	Stud5	Stud6	Stud7	Stud8	Stud9	Stud10	Stud11	Stud12
Stud1	*											
Stud2	**0.038**	*										
	(0.000)											
Stud3	**0.094**	**0.040**	*									
	(0.000)	**(0.000)**										
Stud4	**0.069**	**0.034**	**0.068**	*								
	(0.000)	**(0.000)**	**(0.000)**									
Stud5	**0.095**	**0.041**	**0.073**	**0.065**	*							
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**								
Stud6	**0.046**	0.014	**0.055**	**0.034**	**0.038**	*						
	(0.000)	(0.005)	**(0.000)**	**(0.000)**	**(0.000)**							
Stud7	**0.043**	**0.024**	**0.054**	**0.046**	**0.030**	**0.020**	*					
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**						
Stud8	**0.069**	0.015	0.064	0.005	0.054	0.013	0.009	*				
	(0.001)	(0.301)	(0.007)	(0.463)	(0.031)	(0.333)	(0.250)					
Stud9	**0.086**	**0.058**	**0.100**	**0.057**	**0.057**	**0.049**	**0.048**	0.055	*			
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	(0.026)				
Stud10	**0.074**	**0.026**	**0.052**	**0.059**	**0.047**	**0.042**	**0.043**	0.036	**0.065**	*		
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	(0.116)	**(0.000)**			
Stud11	**0.113**	**0.055**	**0.096**	**0.053**	**0.095**	**0.063**	**0.064**	0.031	**0.097**	**0.083**	*	
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	(0.135)	**(0.000)**	**(0.000)**		
Stud12	**0.087**	**0.055**	**0.089**	**0.074**	**0.066**	**0.054**	**0.043**	0.051	**0.075**	**0.062**	**0.119**	*
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	(0.009)	**(0.000)**	**(0.000)**	**(0.000)**	

The Bayesian assignment approach utilised by STRUCTURE revealed that the data obtained from twelve SA Boerperd studs had the highest probability of representing two clusters, with this arrangement shown in Figure 3.2.

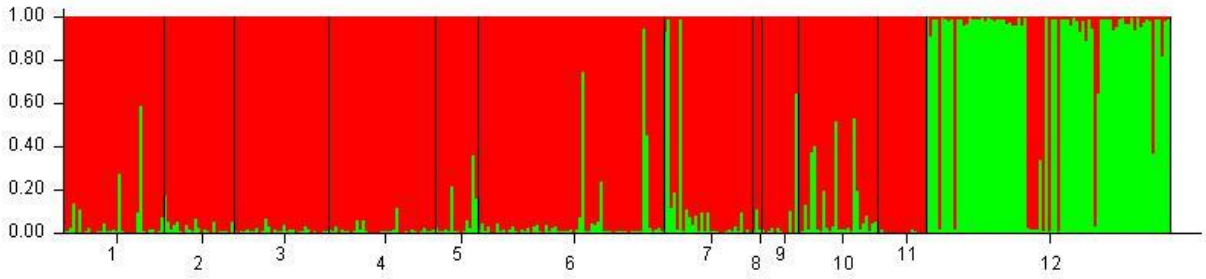


Figure 3.2 Graphical representation of the population structure of 363 SA Boerperd horses from 12 studs (K=2). Each vertical line with its colour segments indicate the inferred membership of an individual to a specific cluster. Generated by STRUCTURE.

The presence of two main clusters is based on the ΔK and mean likelihood values generated by Structure Harvester, as shown in Figures 3.3 and 3.4. Estimates of K that are below the proper estimate of populations produce small -LnPr values, but increases and plateaus as the K-values increase (Earl & von Holdt, 2012). The appropriate K-value is then the smallest plateau value. Furthermore, the Evanno *et al.* (2005) method establishes K at the maximum ΔK-value. Based on the composition of the bar plot it seemed that Cluster1 consisted primarily of the individuals from Stud1 – Stud11, and Cluster2 of horses from Stud12.

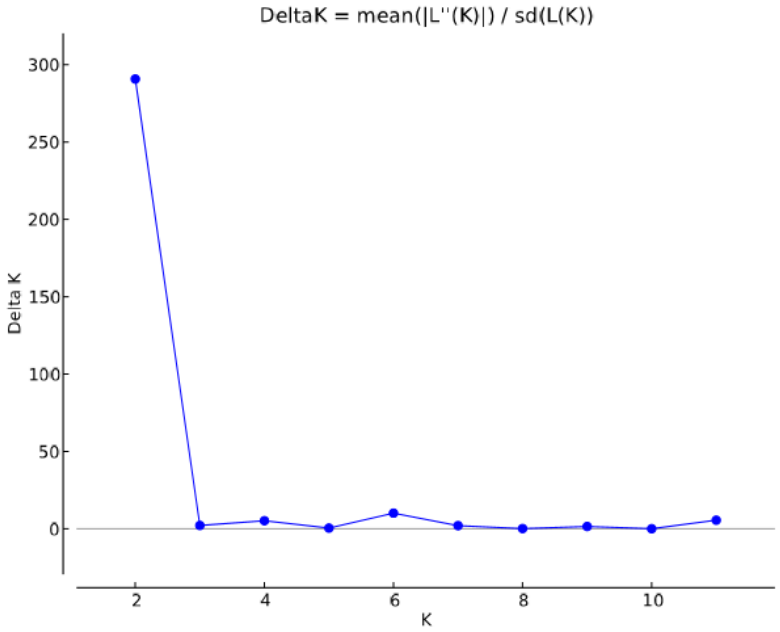


Figure 3.3 Plot indicating the number of clusters (K) that best fit the genomic data of 12 SA Boerperd studs. Generated by Structure Harvester.

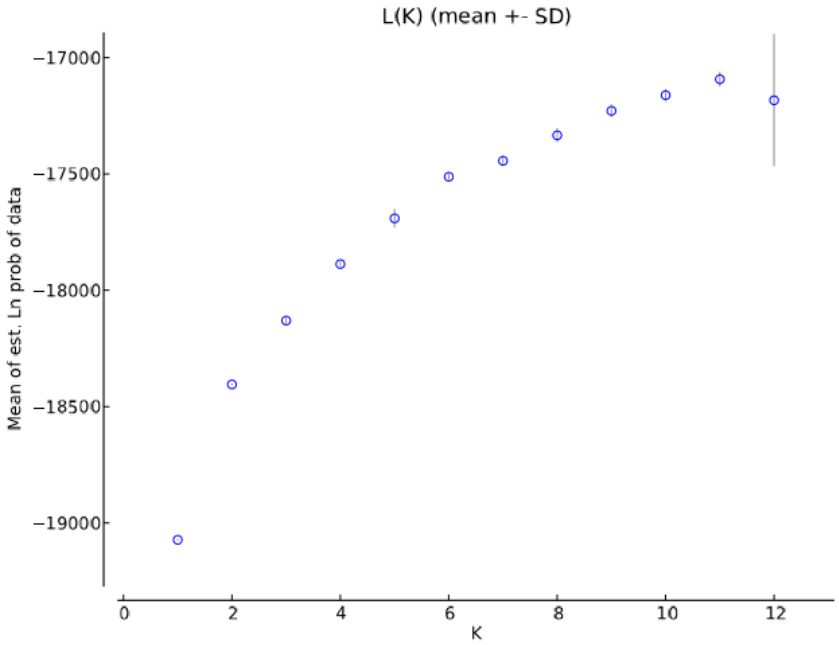


Figure 3.4 Plot indicating the mean likelihood L(K) and variance per K-value for 12 SA Boerperd studs obtained from STRUCTURE. Generated by Structure Harvester.

The membership likelihood of the twelve studs to each of the two inferred clusters are presented in Table 3.5. Additionally, past ownership detail obtained from the SA Boerperd Breeders Society revealed that a few individuals of Stud7 and Stud10 were obtained from Stud12, and some within Stud12 also originate from other studs.

Table 3.5 Proportion of membership of each SA Boerperd stud to one of the two inferred clusters (K=2). Generated by STRUCTURE.

Stud	N	Inferred clusters	
		Cluster1	Cluster2
Stud1	33	0.950	0.050
Stud2	23	0.968	0.032
Stud3	31	0.984	0.016
Stud4	35	0.983	0.017
Stud5	14	0.932	0.068
Stud6	61	0.942	0.058
Stud7	29	0.859	0.141
Stud8	3	0.951	0.049
Stud9	12	0.928	0.072
Stud10	26	0.887	0.113
Stud11	16	0.993	0.007
Stud12	80	0.164	0.836

The current (2017) genetic diversity of the SA Boerperd was compared to its diversity in the year 2002. Some of the investigated loci proved to be unsuitable for comparison, thus only 9 loci (AHT4, AHT5, ASB2, HMS3, HMS6, HMS7, HTG4, HTG7, VHL20) were used. Even though the 2002 population had considerably less individuals than 2017, its N_e was slightly higher than that of the present year (Table 3.6). Both populations' H_o and H_z had stayed relatively constant though. A significant change could be seen regarding the breed's inbreeding coefficient. The number of heterozygotes seems to have increased over the years, though it is probable that the larger sample size of 2017 may have influenced this value.

Table 3.6 Genetic diversity of two SA Boerperd populations sampled in the years 2002 and 2017. The calculated parameters were based on 9 microsatellite markers.

Population	N	N_a (SD)	N_e (SD)	H_o (SD)	H_z (SD)	F_{IS}
2002	81	7.556 (0.626)	4.212 (0.474)	0.672 (0.037)	0.739 (0.033)	0.084
2017	363	7.444 (0.626)	3.804 (0.461)	0.679 (0.044)	0.701 (0.041)	0.032

The 2002 population deviated from HWE at five loci, whereas the 2017 population only deviated at four (Table 3.7). Both populations deviated significantly at HMS7 and HTG7. This could potentially be an indication of linkage disequilibrium at the two unlinked loci. Additionally, the HWE changes displayed for most of the other loci (AHT5, ASB2, HMS3, HTG4 & VHL20) could be due to the loss and addition of novel alleles within the breed, as was found with GenAlEx.

Table 3.7 Conformation to expected Hardy-Weinberg Equilibrium of genotypes for 9 loci within two SA Boerperd populations sampled in the years 2002 and 2017. Values of $p < 0.05$ indicate significant deviation from expected HWE and are emphasised in bold.

Marker / Population	AHT4	AHT5	ASB2	HMS3	HMS6	HMS7	HTG4	HTG7	VHL20
2002	0.258	**0.030**	0.053	0.163	0.385	**0.000**	**0.039**	**0.028**	**0.014**
2017	0.254	0.153	**0.009**	**0.000**	0.693	**0.044**	0.404	**0.001**	0.155

3.3.2. Among-breed analysis

The microsatellite data of 35 breeds used by Van de Goor, van Haeringen & Lenstra (2011) were reduced to 9 breeds. These breeds were specifically chosen to include the breeds thought to have influenced the SA Boerperd, as discussed in various resources (Louw, 2008;

Du Toit, 2010; Swart, 2010; Nel, 2014). The ancestral breeds thus encompassed the Tennessee Walker, Andalusian, Standardbred, Appaloosa, Arabian, Friesian, Hackney and Thoroughbred. The Icelandic Horse was also included, seeing as it is a gaited breed like the SA Boerperd.

It was found that the SA Boerperd had the highest N_a (8.000) and the Thoroughbred the lowest at 5.118 (Table 3.8). Yet when taking sample size into consideration, the Appaloosa had the highest N_e of 4.499 and the Friesian the lowest (2.163). Both H_o and H_z were also highest for the Appaloosa, and lowest for the Friesian. The SA Boerperd had average heterozygosity values of H_e = 0.682 and H_z = 0.722.

Table 3.8 Genetic diversity of 10 horse breeds based on 17 microsatellite markers. The highest values of each parameter are highlighted in bold.

Breed (Abbreviation)	**N**	N_a **(SD)**	N_e **(SD)**	H_o **(SD)**	H_z **(SD)**	**PA (SD)**	F_{IS}
Tennessee Walker (TEN)	23	5.176	3.204	0.670	0.674	0.000	-0.019
		(0.312)	(0.237)	(0.033)	(0.026)	(0.000)	
Andalusian (AND)	67	6.824	3.405	0.658	0.670	0.059	0.000
		(0.456)	(0.292)	(0.037)	(0.033)	(0.059)	
Standardbred (STA)	997	7.588	3.764	0.690	0.710	0.235	0.024
		(0.500)	(0.254)	(0.032)	(0.025)	(0.136)	
Appaloosa (APP)	99	7.706	**4.499**	**0.717**	**0.765**	0.118	0.058
		(0.460)	**(0.267)**	**(0.034)**	**(0.019)**	(0.081)	
Arabian (ARA)	615	7.235	3.412	0.645	0.678	0.235	0.043
		(0.497)	(0.242)	(0.030)	(0.028)	(0.106)	
Friesian (FRI)	781	5.235	2.163	0.439	0.447	0.059	0.023
		(0.450)	(0.262)	(0.053)	(0.051)	(0.059)	
Hackney (HAC)	141	6.118	3.210	0.618	0.661	0.059	0.060
		(0.453)	(0.231)	(0.030)	(0.027)	(0.059)	
Thoroughbred (THO)	55	5.118	3.542	0.667	0.695	0.000	0.026
		(0.331)	(0.276)	(0.031)	(0.025)	(0.000)	
Icelandic (ICE)	134	7.176	3.945	0.659	0.705	0.235	0.049
		(0.516)	(0.383)	(0.033)	(0.032)	(0.106)	
SA Boerperd (SAB)	363	**8.000**	4.123	0.682	0.722	0.176	0.053
		(0.536)	(0.378)	(0.030)	(0.027)	(0.095)	

PA = private alleles

The number of private alleles within each breed were also calculated. It was found that the Standardbred, Arabian and Icelandic shared the highest value at 23.5%. The SA Boerperd had

the second highest proportion at 17.6%. The least inbred breed was shown to be the Tennessee Walker with a value of -0.019, whilst the breed with the highest inbreeding coefficient was the Hackney at 6%.

Table 3.9 presents the HWE values of the SA Boerperd and accompanying breeds. Thirteen of the seventeen investigated loci deviated from the expected HWE, and most deviated for one or two breeds. ASB2, HTG10, LEX3 and VHL20 departed from HWE for four or more breeds. Of most significance was LEX3, which deviated for all 10 breeds. This locus is X-linked and would thus deviate, seeing as male individuals only have one allele at this locus (Van de Goor, van Haeringen & Lenstra, 2011). Each breed consequently had at least one locus that deviated from HWE. Interestingly, the SA Boerperd had the most loci (8) that did not conform to expected HWE, and the Friesian had the second most at 7 loci.

Table 3.9 Conformation to expected Hardy-Weinberg Equilibrium of genotypes for 17 loci within 10 horse breeds. Values of $p < 0.05$ indicate significant deviation from expected HWE and are highlighted in bold.

Breed / Locus	TEN	AND	STA	APP	ARA	FRI	HAC	THO	ICE	SAB
AHT4	0.875	0.993	0.573	0.380	**0.001**	**0.000**	0.853	0.730	0.955	0.254
AHT5	0.217	**0.002**	0.329	0.366	0.194	0.472	**0.020**	0.083	0.738	0.153
ASB2	0.572	0.194	0.988	0.958	**0.000**	**0.000**	**0.000**	0.955	**0.001**	**0.009**
ASB17	0.590	0.074	0.051	0.455	0.787	**0.000**	0.070	0.817	0.982	**0.000**
ASB23	0.314	0.310	0.530	0.258	**0.000**	0.205	0.212	0.237	**0.032**	0.113
CA425	0.556	0.211	0.953	0.073	0.087	0.952	0.452	0.991	0.070	0.283
HMS1	0.665	0.846	0.901	0.167	0.519	0.925	0.534	0.709	0.994	0.133
HMS2	0.566	0.459	0.110	0.618	0.189	**0.000**	0.890	0.587	0.786	0.093
HMS3	0.317	0.540	0.274	**0.000**	0.468	1.000	0.336	0.922	0.163	**0.000**
HMS6	0.326	0.837	0.329	0.786	0.118	0.202	0.058	0.496	0.636	0.693
HMS7	0.061	0.955	0.706	0.785	0.665	0.451	0.248	0.882	0.966	**0.044**
HTG4	0.466	0.307	0.360	0.736	0.802	0.711	0.068	0.421	0.303	0.404
HTG6	0.917	0.986	0.694	0.603	0.932	0.982	0.286	0.750	0.392	**0.037**
HTG7	0.768	0.889	0.925	**0.002**	0.059	0.512	0.877	0.702	0.675	**0.001**
HTG10	0.621	**0.027**	0.129	0.677	**0.003**	**0.007**	0.999	0.297	0.223	**0.000**
LEX3	**0.028**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
VHL20	0.449	**0.001**	**0.041**	**0.033**	0.057	**0.001**	0.643	0.422	0.535	0.155

TEN = Tennessee Walker; AND = Andalusian; STA = Standardbred; APP = Appaloosa; FRI = Friesian; HAC = Hackney; THO = Thoroughbred; ICE = Icelandic Horse; SAB = SA Boerperd

The pair-wise population F_{ST}-values and p-values for the ten breeds are presented in Table 3.10. The Bonferroni correction was calculated as p = 0.001. The pair-wise combinations between all the breeds show F_{ST}-values and p-values that support the hypothesis of significant differentiation. The Friesian was consistently shown to be the furthest related to all the breeds. The closest associations however were found between the Appaloosa and Hackney, as well as the Appaloosa and SA Boerperd, with an F_{ST}-value of 0.070. In addition to the Appaloosa, the SA Boerperd also showed a close association with the Andalusian and Standardbred. Compared to the other breed's associations, the SA Boerperd is the closest related to most of the breeds, which could indicate the influence some of the breeds had in the recent development of the SA Boerperd.

Table 3.10 Pairwise F_{ST}-values between 10 horse breeds. P-values are indicated within brackets.

	TEN	**AND**	**STA**	**APP**	**ARA**	**FRI**	**HAC**	**THO**	**ICE**	**SAB**
TEN	*									
AND	0.105	*								
	(0.000)									
STA	0.124	0.112	*							
	(0.000)	(0.000)								
APP	0.071	0.095	0.107	*						
	(0.000)	(0.000)	(0.000)							
ARA	0.126	0.099	0.097	0.106	*					
	(0.000)	(0.000)	(0.000)	(0.000)						
FRI	0.373	0.352	0.314	0.308	0.327	*				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
HAC	0.117	0.138	0.153	0.070	0.151	0.318	*			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
THO	0.118	0.116	0.084	0.112	0.099	0.403	0.166	*		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
ICE	0.124	0.157	0.167	0.065	0.174	0.338	0.099	0.180	*	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
SAB	0.086	0.078	0.083	0.070	0.093	0.308	0.100	0.087	0.126	*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

TEN = Tennessee Walker; AND = Andalusian; STA = Standardbred; APP = Appaloosa; FRI = Friesian; HAC = Hackney; THO = Thoroughbred; ICE = Icelandic Horse; SAB = SA Boerperd

Based on the ΔK (Figure 3.5) and mean likelihood values (Figure 3.6) for various values of K in STRUCTURE, it was found that the ten equine breeds had the highest probability for belonging to two clusters. On closer inspection, this phenomenon appeared to be the separation of the cold-blooded Friesian from the more hot-blooded breeds, as seen in Figure 3.7. Both the clustering output of K=4 and K=10 were examined additionally (Figure 3.7). These clustering results were included seeing as the likelihood scores at K=4 also showed additional plateauing and an elevated ΔK-value, and since it was known that the genomic data used originated from ten breeds.

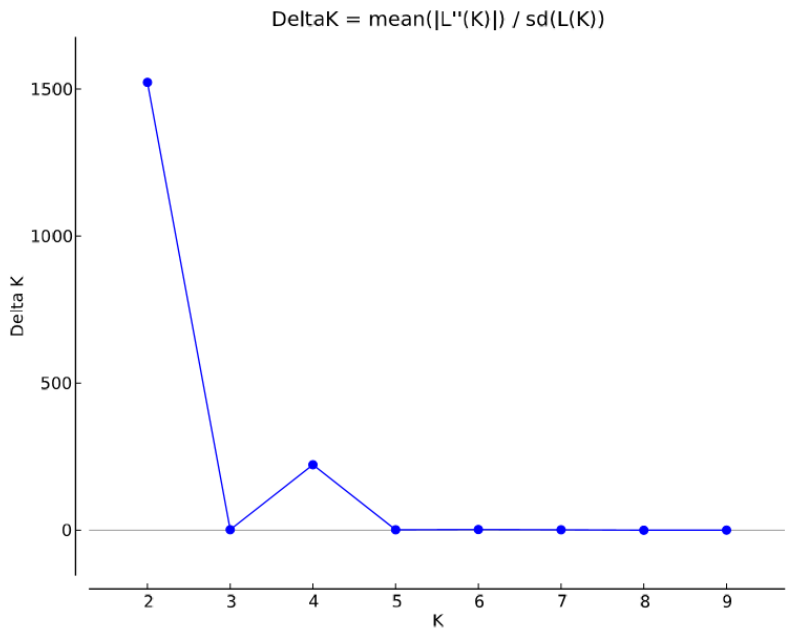


Figure 3.5 Number of clusters (K) that best fit the genomic data of 10 equine breeds. Generated by Structure Harvester.

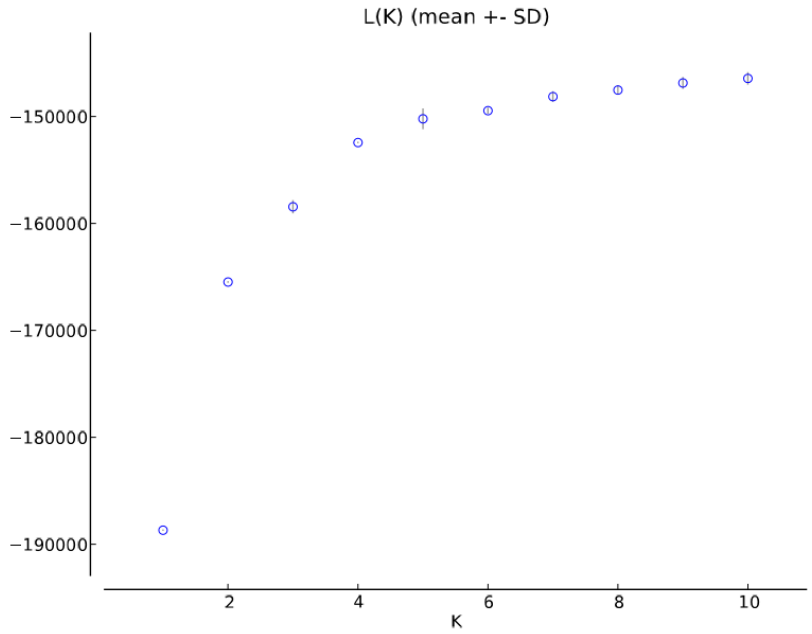


Figure 3.6 Mean likelihood L(K) and variance per K-value for 10 equine breeds. Generated by Structure Harvester.

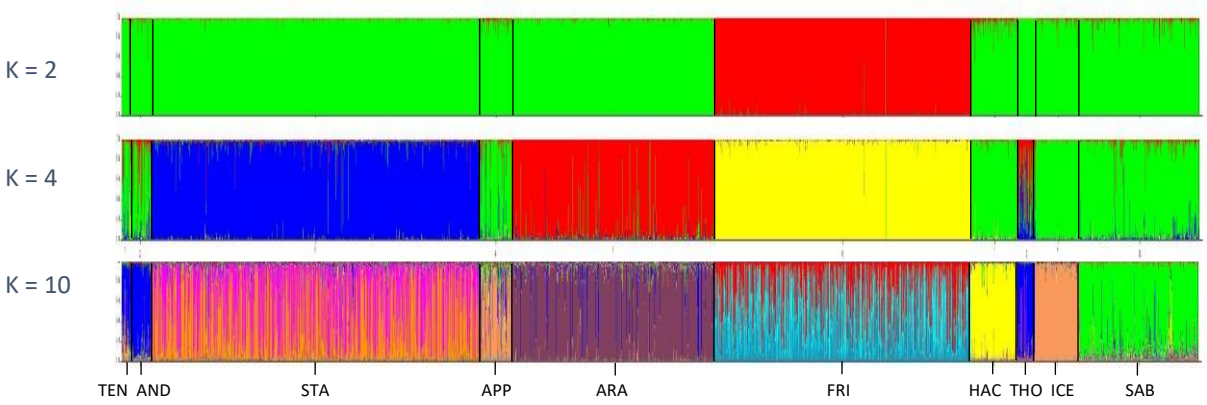


Figure 3.7 Clustering output for three values of K in 3,275 horses belonging to 10 breeds. An individual is represented by one vertical line, which indicates the proportion of assignment to one or more clusters. Generated by STRUCTURE. Abbreviations: TEN = Tennessee Walker; AND = Andalusian; STA = Standardbred; APP = Appaloosa; FRI = Friesian; HAC = Hackney; THO = Thoroughbred; ICE = Icelandic Horse; SAB = SA Boerperd.

The proportion of membership of each breed to two and ten inferred clusters are presented in Tables 3.11 and 3.12 respectively. The Friesian was the only breed significantly assigned to its own cluster at K=2. When ten clusters were assumed, only two breeds (Hackney & Icelandic) showed more than 90% identity or assignment to their own unique cluster. The Arabian (84.9%) and SA Boerperd (85%) were also assigned to their own clusters.

Table 3.11 Proportion of membership of each breed to one of the two inferred clusters (K=2). Generated by STRUCTURE.

Breed	N	Inferred clusters	
		Cluster1	Cluster2
TEN	23	0.011	0.989
AND	67	0.038	0.962
STA	997	0.006	0.994
APP	99	0.032	0.968
ARA	615	0.007	0.993
FRI	781	0.994	0.006
HAC	141	0.050	0.950
THO	55	0.002	0.998
ICE	134	0.024	0.976
SAB	363	0.026	0.974

TEN = Tennessee Walker; AND = Andalusian; STA = Standardbred; APP = Appaloosa; FRI = Friesian; HAC = Hackney; THO = Thoroughbred; ICE = Icelandic Horse; SAB = SA Boerperd

Table 3.12 Proportion of membership of each breed to one of ten inferred clusters (K=10). Generated by STRUCTURE.

Breed	N	Inferred clusters									
		1	2	3	4	5	6	7	8	9	10
TEN	23	0.007	0.029	0.650	0.056	0.035	0.012	0.060	0.020	0.119	0.010
AND	67	0.020	0.017	0.844	0.021	0.011	0.018	0.011	0.027	0.012	0.019
STA	997	0.006	0.014	0.019	0.010	0.432	0.006	0.488	0.012	0.008	0.006
APP	99	0.016	0.051	0.129	0.061	0.027	0.016	0.040	0.050	0.593	0.017
ARA	615	0.005	0.013	0.077	0.012	0.012	0.005	0.010	0.849	0.012	0.005
FRI	781	0.342	0.003	0.003	0.004	0.003	0.312	0.003	0.003	0.004	0.323
HAC	141	0.006	0.018	0.011	0.919	0.007	0.005	0.006	0.009	0.014	0.006
THO	55	0.003	0.011	0.819	0.009	0.076	0.003	0.031	0.042	0.005	0.003
ICE	134	0.006	0.006	0.007	0.011	0.006	0.006	0.005	0.005	0.943	0.005
SAB	363	0.009	0.850	0.037	0.028	0.015	0.009	0.013	0.019	0.011	0.008

TEN = Tennessee Walker; AND = Andalusian; STA = Standardbred; APP = Appaloosa; FRI = Friesian; HAC = Hackney; THO = Thoroughbred; ICE = Icelandic Horse; SAB = SA Boerperd

3.4. Discussion & conclusion

3.4.1. Within-breed diversity and differentiation

The genetic diversity of all the SA Boerperd studs were surprisingly similar, despite the varying sample sizes. A total of 136 alleles were detected across the 17 loci for the breed, and roughly three alleles averaged per locus for each stud. The observed number of heterozygotes within each stud did not deviate too much from that having been expected to be present. Higher inbreeding levels were expected for the SA Boerperd, yet sufficiently low F_{IS}-values and a surplus of heterozygotes were found for most studs. Of significance was Stud12, which had the most heterozygotes and lowest inbreeding coefficient.

Population differentiation is influenced by gene flow, or the exchange of genetic material between populations (Allendorf, Luikart & Aitken, 2013). The differentiation estimates in this study revealed that most studs exchanged genes, yet Stud11 and Stud12 exchanged less by differing significantly from one another. As did Stud1 from Stud11. This could be due to the geographical localities of these studs, which are in different provinces, as well as the fact that the stud owners probably do not obtain fresh stock from one another.

Further structural analysis of the breed revealed that Stud12 was significantly different from the overall population. This population is situated closely to some of the other studs, thus the differentiation is most likely due to selective breeding. It is possible that the owner of this stud is indeed selecting individuals of certain phenotypes to help conform the rest of the stock to this standard. This practice has thus lead to the creation of a genetically distinct subpopulation within the breed. Although the creation of such a homogenous group can decrease fertility and sustainability (Lacy, 1987), it seems that the selection process of this particular stud has ensured its heightened genetic diversity when considering the negative inbreeding value.

The comparison of the genetic data of the SA Boerperd in 2017 with that of 2002 revealed that the number of heterozygotes had stayed constant, despite the significantly larger sample size of 2017. Of concern is the slight decrease in the breed's allelic richness, which was unexpected given the increased sample size. This could be due to sampling error, but since the data of entire reproductive studs were used, it's more likely that certain alleles have been lost due to selective breeding. Despite the latter concern, the breed's inbreeding coefficient

showed a promising decline. Yet this could be due to the larger number of tested individuals. Even if this is the case, the overall diversity estimates of the SA Boerperd demonstrate its variability and continual sustainability.

Something to keep in mind is that the microsatellite markers used within this study are genetically neutral. Such loci are ideal for investigating population relationships, but loci related to disease resistance and immune responses should be investigated to further potential conservation efforts. These investigations can help shed light on the impact which the loss of alleles linked to adaptability can have on a specific breed, as was suggested by Marletta *et al.* (2006).

3.4.2. Differences between breeds

The comparison of the genetic diversity of the SA Boerperd to that of other breeds revealed it to be at a healthy, sustainable level. It possessed the highest mean number of alleles per locus, proving its long-term evolutionary potential (Allendorf, Luikart & Aitken, 2013) and suggesting that selective pressures must have had little effect on the breed over the years. It also had a significant number of private alleles, which makes it a distinct and unique breed (Allendorf, Luikart & Aitken, 2013). The 5.3% inbreeding value of the breed could be due to inbreeding or selection, but is more likely caused by the substructure found in the breed. Significantly positive values of F_{IS} were also found within Standardbreds due to substructure (Cothran *et al.,* 1987).

The large number of loci that deviated from HWE within the SA Boerperd could possibly be due to the stratification of the overall population, seeing as similar findings were made by Cañon *et al.* (2000) within the Caballo Gallego breed. Pairwise differentiation values showed that the SA Boerperd is separated the furthest from the Friesian, and Botha (2002) had similar findings. Additionally, the close association of the breed to that of the Appaloosa and Standardbred suggests that they may indeed be some of the American breeds that had an influence in the Boerperd. Though for more conclusive results the SA Boerperd should be compared to other American breeds as well.

As within the current study, Van de Goor, van Haeringen and Lenstra (2011) could not determine a 'true' value for K when they investigated the genetic relationships between thirty-five different horse breeds. At K=4 these authors did however find that the cluster

pattern represented the grouping of related breeds into the three main equine classes, namely pony, hot-blooded, and cold-blooded breeds. Similar results were found in this study, where the cold-blooded Friesian was already distinctly separated from the rest of the riding breeds at K=2. In contrast to Van de Goor, van Haeringen and Lenstra (2011), the Appaloosa and Icelandic (pony breeds) didn't exclusively cluster together until K=8. Despite having used SNP data instead of microsatellites, Petersen *et al.* (2013) also could not establish a value for K when 38 breeds were investigated. As within the latter paper, it was similarly found within the current study that the Standardbred, Arabian and Friesian were the first to be assigned to their own clusters (K=4). Factors such as allele frequency correlations, sample sizes, the number of assumed clusters, and number of typed markers can all influence clustering and should be considered when estimating K (Rosenberg *et al.,* 2005).

The SA Boerperd was continually grouped together with the Appaloosa, Hackney and Icelandic until it was assigned to its own cluster at K=7. This could hint at its relatedness to these pony breeds. Yet based on F_{ST}-values, the SA Boerperd is genetically closer to the Appaloosa, Standardbred and Andalusian. At lower values of K, it was shown that the SA Boerperd also clustered together with the Tennessee Walker, Andalusian and Standardbred. Like the SA Boerperd, the latter three breeds are also gaited. Therefore, considering that all modern gaited breeds share a common ancestor (Bowling & Ruvinsky, 2000), this clustering pattern is plausible.

The study of Cothran and van Dyk (1998) showed that the Boerperd, consisting of both the SA Boerperd and Cape Boerperd, were the closest associated with the American Saddlebred and Rocky Mountain Horse. Botha (2002) also found the Saddlebred to be closely associated with both Boerperd breeds. Unfortunately, these American breeds were not investigated by Van de Goor, van Haeringen and Lenstra (2011), thus STR marker data were not available for comparative purposes during this study. It would thus be beneficial to investigate the exact relationship of these breeds specifically to the SA Boerperd in future.

REFERENCES

Allendorf, F.W., Luikart, G. and Aitken, S.N. (2013). Conservation and the genetics of populations. Chicester: Wiley-Blackwell.

Avise, J.C. (2004). *Molecular markers, natural history and evolution*. 2nd ed. Massachusetts: Sinauer Associates.

Beaumont, M.A. and Balding, D. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, 13: 969 – 980.

BJØRNSTAD, G. and RØED, K.H. (2002). Evaluation of factors affecting individual assignment precision using microsatellite data from horse breeds and simulated breed crosses. *Animal Genetics,* 33: 264 – 270.

Botha, K. (2002). *Genetic relationships of seven horse breeds in South Africa based on DNA markers*. Dissertation (M.Sc.). Bloemfontein: University of the Free State.

Bowling, A.T. and Ruvinsky, A. (2000). *The genetics of the horse*. Wallingford: CABI Publishing.

Cañon, J., Checa, M.L., Carleos, C., Vega-Pla, J.L., Vallejo, M. and Dunner, S. (2000). The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. *Animal Genetics*, 31: 39 – 48.

Cothran, E.G., MacCluer, J.W., Weitkamp, L.R. and Bailey, E. (1987). Genetic differentiation associated with gait within American Standardbred horses. *Animal Genetics*, 18: 285 – 296.

Cothran, E.G. and Van Dyk, E. (1998). Genetic analysis of three South African horse breeds. *Journal of the South African Veterinary Association,* 69(4): 120 – 125.

Du Toit, K. (2010). *Boerperd: Die geskiedenis van die S.A. Boerperd*. Pretoria: West Publishers.

Earl, D.A. and von Holdt, B.M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources,* 4(2): 359 – 361.

Equus Horse Safaris. (2016). *Horse Safaris.* Available at: http://www.equus.co.za/safaris.html [Accessed 15 Sept. 2016].

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14: 2611 – 2620.

Excoffier, L., Laval, G. and Schneider, S. (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1: 4750.

Farmer's Weekly. (2017). *Why climate change is a threat to livestock production*. Available at: http://www.farmersweekly.co.za/animals/sheep-goats/climate-change-threat-livestock-production/ [Accessed 1 Feb. 2017].

Frankham, R., Ballou, J.D. and Briscoe, D.A. (2002). *Introduction to Conservation Genetics*. Cambridge: Cambridge University Press.

Glaubitz, J.C. (2004). CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes,* 4: 309 – 310.

Goldstein, D. and Schlötterer, C. (1999). *Microsatellites: Evolution and applications*. New York: Oxford University Press.

Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R., Finlay, E.K., Jianlin, H., Groeneveld, E., Weigend, S. and the GLOBALDIV Consortium. (2010). Genetic diversity in farm animals – a review. *Animal Genetics*, 41: 6 – 31.

Hartl, D.L. and Clark, A.G. (1997). *Principles of population genetics*. Sunderland, MA: Sinauer Associates.

Juras, R., Cothran, E.G. and Klimas, R. (2003). Genetic analysis of three Lithuanian native horse breeds. *Acta Agriculturae Scandinavica, Section A – Animal Science,* 53(4): 180 – 185.

Khanshour, A., Juras, R., Blackburn, R. and Cothran, E.G. (2015). The legend of the Canadian Horse: Genetic diversity and breed origin. *Journal of Heredity,* 106(1): 37 – 44.

Lacy, R.C. (1987). Loss of genetic diversity from managed populations: Interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology*, 1: 143 – 158.

Lewontin, R. and Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74: 175.

Louw, S. (2008). SA Boerperd –Indigenous breed to southern Africa. In: *SA Boerperd: The ideal sport and pleasure horse:* 22 - 29*.* Edited by S. Louw. Maanhaarrand: Kejafa Knowledge Works.

Marletta, D., Tupac-Yupanqui, I., Bordonaro, S., Garcia, D., Guastella, A.M., Criscione, A., Cañón, J. and Dunner, S. (2006). Analysis of genetic diversity and the determination of relationships among western Mediterranean horse breeds using microsatellite markers. *Journal of Animal Breeding Genetics*, 123: 315 – 325.

Meirmans, P.G. and Hedrick, P.W. (2011). Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources,* 11: 5 – 18.

Nel, C. (2014). *SA Boerperd Breed History*. SA Boerperd Breeders Society. Available at: http://www.saboerperd.com/p2/history/sa-boerperd-breed-history.html [Accessed 5 July 2016].

Ojango, J.M., Mpofu, N., Marshall, K. and Andersson-Eklund, L. (2011). Quantitative methods to improve the understanding and utilisation of animal genetic resources. In: *Animal Genetics Training Resource*, 3rd edition. Edited by J.M. Ojango, B. Malmfors, and A.M. Okeyo. International Livestock Research Institute, Nairobi, Kenya, and Swedish University of Agricultural Sciences, Uppsala, Sweden.

Pakamisa. (2016). *Horseback Safaris*. Available at: http://www.pakamisa.co.za/horseback-safaris.php [Accessed 15 Sept. 2016].

Peakall, R. and Smouse P.E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* 6: 288 – 295.

Peakall, R. and Smouse P.E. (2012). GENALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, 28: 2537 – 2539.

Petersen, J.L., Mickelson, J.R., Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., Da Câmara Machado, A., Distl, O., Felicetti, O., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C.T., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Silvestrelli, M., Swinburne, J., Tozaki, T., Vaudin, M., Wade, C.M. and McCue, M.E. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *Plos ONE,* 8(1): e54997. doi: 10.1371/journal.pone.0054997

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, A. and Lareu, M.V. (2013). An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Frontiers in Genetics*, 4: 1 – 13.

Pritchard, J.K., Stephens, M. and Donnelly, P. (2000). Inference of population structure using multi-locus genotype data. *Genetics*, 155: 945 – 959.

Rosenberg, N.A., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J.K. and Feldman, M.W. (2005). Clines, clusters, and the effect of study design on the inference of human population structure. *PloS Genetics,* 1(6): e70. doi: 10.1371/journal.pgen.0010070

Sereno, F.T.P.S., Sereno, J.R.B., Vega-Pla, J.L., Kelly, L. and Bermejo, J.V.D. (2008). Genetic diversity of Brazialian Pantaneiro horse and relationships among horse breeds. *Pesquisa Agropecuária Brasileira,* 43: 595 – 604.

Swart, S. (2010). *Riding High: Horses, Humans and History in South Africa*. Johannesburg: Wits University Press.

Takezaki, N. and Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, 144: 389 – 399.

Van de Goor, L.H.P., Panneman, H. and van Haeringen, W.A. (2009). A proposal for standardization in forensic equine DNA typing: Allele nomenclature of 17 equine-specific STR loci. *Animal Genetics,* 41: 122 – 127.

Van de Goor, L.H.P., van Haeringen, W.A. and Lenstra, J.A. (2011). Population studies of 17 equine STR for forensic and phylogenetic analysis. *Stichting International Foundation for Animal Genetics,* 42: 627 – 633.

Wigginton, J.E., Cutler, D.J. and Abecasis, G.R. (2005). A note on exact tests of Hardy-Weinberg Equilibrium. *American Journal of Human Genetics*, 76: 887 – 893.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution,* 19: 395 – 420.

Personal Communication:

Jan Joubert, Vice President of the SA Boerperd Breeders Association.

Johannes Langenhoven, President of the SA Boerperd Breeders Association.

Chapter 4: General discussion, conclusion & recommendations

(Picture obtained from SA Boerperd Breeders Society; www.saboerperd.com)

This study was the second to investigate the genetic influence on performance traits in the SA Boerperd, the first having been done by Promerová *et al.* (2014). The objectives of the current study were to screen for the occurrence of two height-associated SNPs and two gait-associated SNPs in the breed; and to study the genetic diversity and differentiation of the breed using 17 microsatellite markers. The latter aim included a number of specific objectives, specifically to (i) determine the presence of subpopulations, (ii) evaluate the present diversity levels with that of previous years, and (iii) compare the breed's diversity to that of other equine breeds to determine its sustainability. These objectives were met as follows:

4.1. Performance trait analysis

The SA Boerperd is well-known within the country's equestrian circle as a willing performer able to clear obstacle courses and known to display five alternating gaits. The continuation of such abilities are important to breeders; thus determining and maintaining the factors influencing these performance traits in the breed is invaluable. Genetics have been found to play a significant role in traits like height and gaitedness, which influences performance. The mutations thought to contribute the most to height and gait-ability were investigated in the SA Boerperd to assist in the breeding of performers.

The height-associated SNPs BIEC2_808543 and BIEC2_1105377 were found present within the breed at low frequencies. Studies have indicated that BIEC2_808543 attribute 1 cm and BIEC2_1105377 0.5 cm to height (Signer-Hasler *et al.,* 2012), though this could not be established for the SA Boerperd. Furthermore, only five individuals taller than 160 cm possessed BIEC2_808543, and the average height of horses stayed constant whether BIEC2_1105377 was present or not. It is thus doubtful that these SNPs have a truly significant impact on height within the breed.

Only 10.5% of the tested SA Boerperd individuals displayed the *gait keeper* mutation (*DMRT3_Ser301STOP*), which is evidently low for an acclaimed gaited breed. Also of interest was that five of the reported five-gaited individuals were homozygous normal for the *gait keeper* mutation. This phenomenon was also previously observed within the Icelandic Horse (Promerová *et al.,* 2014), though as to how this is possible is still unclear. In contrast, a few three-gaited individuals possessed the mutation, indicating their potential to be trained for the additional gaits. The other gait-associated SNP, BIEC2_620109, was also found present

within the breed, and is most likely co-inherited with *DMRT3_Ser301STOP* due to its similar expression. SA Boerperd individuals need thus only be screened for the *gait keeper* mutation.

4.2. Genetic diversity

The continued existence of breeds is subject to the presence of alternate alleles and genotypes (Frankham *et al.,* 2002), and this applies equally to the SA Boerperd. This genetic diversity can help the breed adapt when changes occur within its environment (Juras, Cothran & Klimas, 2003). Such changes are unavoidable, thus monitoring the level of diversity within the SA Boerperd is essential. The current diversity of the breed was thus compared to that obtained in previous years (Botha, 2002), and it seems that the level has slightly declined. The allelic richness (N_e) weakened from 4.212 to 3.804. At present this decline is moderate and probably of no concern, but if it continues then alternative breeding strategies should be put into place to preserve the remaining diversity.

Another possible consequence of selective breeding was the occurrence of substructure found within the SA Boerperd. Horses originating from a sole stud showed an extremely low inbreeding value (F_{IS} = -0.118) and the highest number of heterozygotes (H_o = 0.735). Furthermore, structural analysis of the breed revealed the latter stud to be genetically distinct from the rest of the tested populations. It is known that this stud engages in a breeding program selecting for a specific conformational 'type' whilst still adhering to the breed standard. This selective breeding could explain the substructure found within the breed. This outbred population in truth could be advantageous to the breed as a whole if it is allowed to interbreed with the rest of the country's stock and distribute its unique alleles.

The diversity of the SA Boerperd was also compared to that of nine additional equine breeds. The SA Boerperd had the highest number of alleles (N_a = 8.000) and an above average heterozygote count of H_o = 0.682. It also had the second highest number of private alleles (17.6%), which establishes the breed as a distinct population. Since the breed was possibly influenced by American breeds not included in this study, it would be beneficial to also compare the diversity level of the SA Boerperd with such breeds to achieve a more accurate diversity estimation. Nevertheless, the SA Boerperd was found to be a highly diverse breed and is most likely at an acceptably sustainable level.

4.3. Recommendations

The above-mentioned observations revealed that BIEC2_808543 and BIEC2_1105377 do not contribute significantly to height in the SA Boerperd. Screening horses for these mutations will thus not be meaningful for breeders, unless future research involving a much larger sample population find evidence to the contrary. The same can be said about both gait-associated SNPs (*DMRT3_Ser301STOP* & BIEC2_620109). Although most five-gaited individuals were shown to possess the investigated SNPs, roughly a third did not. Since the frequencies of the mutated alleles were very low within the breed, it's difficult to determine whether this is a random phenomenon or not. Future research is thus needed to investigate the influence other mutations may have on the expression of gaitedness.

Performance traits are often complex by nature and thus difficult to predict when focusing solely on one or two genetic markers of large effect. It has been suggested that the selection of stud animals be based on genomic information that encompasses loci that affect specific traits on both a large and small scale (Meuwissen *et al.,* 2001). This concept is referred to as genomic selection and has been found to support accurate selection decisions based on the genomic estimated breeding values (GEBVs) of parent individuals (Stock *et al.,* 2016). A significant benefit of this process is that it can be used to improve traits that are difficult to measure or physically manifest late during an individuals' development (Wolc, 2014).

The genomic selection process does however require a large reference population of which both the phenotypes and genotypes of all individuals are known to estimate the effect of specific markers on certain traits (Stock *et al.,* 2016). Since this selection process require the constant record taking of high quality trait data, it is at present not being implemented regarding equine breeds. Despite this drawback, Fevrelle (2017) extensively studied the effects of genomic selection in the Danish Warmblood horse and recommended its use in the breed's breeding systems. This selection process can similarly benefit the SA Boerperd, where the future phenotype of an individual can then potentially be established from birth.

If such a selection process is ever considered to be utilised for the SA Boerperd, then accurate records of phenotypical traits like height should be kept by the Breeders Association to correctly monitor the effect of genes on estimated breeding values (EBVs). The use of high-density equine SNP arrays should also be considered to genotype horses in aid of this

selection process. Such arrays are known to enable whole genome mapping of individuals by sequencing thousands of genetic variants (Petersen *et al.,* 2013a). This technology can simultaneously be used to monitor the SA Boerperd's level of diversity when selecting for choice genetic markers. Petersen *et al.* (2013b) demonstrated that the genetic diversity observed by SNP arrays for one horse breed can also be compared to that of other equine breeds determined in the same fashion. Such comparisons can better aid investigations into the relatedness of the SA Boerperd to other breeds. However, it should be noted that the use of SNP array technology concerning the selection of traits is at present still highly experimental and should therefore not be implemented by the Association until such time when more extensive research has been done.

REFERENCES

Botha, K. (2002). *Genetic relationships of seven horse breeds in South Africa based on DNA markers*. Dissertation (M.Sc.). Bloemfontein: University of the Free State.

Fevrelle, S.A.G. (2017). *Implementing genomic information in breeding schemes of Danish Warmblood horses*. Dissertation (M.Sc.). Aarhus: Aarhus University.

Frankham, R., Ballou, J.D. and Briscoe, D.A. (2002). *Introduction to Conservation Genetics*. Cambridge: Cambridge University Press.

Juras, R., Cothran, E.G. and Klimas, R. (2003). Genetic analysis of three Lithuanian native horse breeds. *Acta Agriculturae Scandinavica, Section A – Animal Science,* 53(4): 180 – 185.

Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157: 1819 – 1829.

Petersen, J.L., Mickelson, J.R., Rendahl, A.K., Valberg, S.J., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., Da Câmara Machado, A., Capomaccio, S., Cappelli, K., Cothran, E.G., Distl, O., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C.T., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Swinburne, J., Tozaki, T., Vaudin, M., Wade, C.M. and McCue, M.E. (2013a). Genome-wide analysis reveals selection for important traits in domestic horse breeds. *Plos Genetics,* 9(1): 1 – 17.

Petersen, J.L., Mickelson, J.R., Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., Da Câmara Machado, A., Distl, O., Felicetti, O., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C.T., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Silvestrelli, M., Swinburne, J., Tozaki, T., Vaudin, M., Wade, C.M. and McCue, M.E. (2013b). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *Plos ONE,* 8(1): e54997. doi: 10.1371/journal.pone.0054997

Promerová, M., Andersson, L.S., Juras, R., Penedo, M.C.T., Reissmann, M., Tozaki, T., Bellone, R., Dunner, S., Hořín, P., Imsland, F., Imsland, P., Mikko, S., Modrý, D., Roed, K.H., Schwochow, D., Vega-Pla, J.L., Mehrabani-Yeganeh, H., Yousefi-Mashouf, N., Cothran, E.G., Lindgren, G. and Andersson, L. (2014). Worldwide frequency distribution of the '*Gait keeper*' mutation in the *DMRT3* gene. *Stichting International Foundation for Animal Genetics,* 45: 274 – 282.

Signer-Hasler, H., Flury, C., Haase, B., Burger, D., Simianer, H., Leeb, T. and Rieder, S. (2012). A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PLoS ONE,* 7(5): e37282. doi: 10.1371/journal.pone.0037282

Stock, K.F., Jönsson, L., Ricard, A. and Mark, T. (2016). Genomic applications in horse breeding. *Animal Frontiers*, 6: 45 – 52.

Wolc, A. (2014). Understanding genomic selection in poultry breeding. *World's Poultry Science Journal,* 70: 309 – 314.

SUMMARY

This study aimed to determine the occurrence and frequency of mutations that influence certain performance-associated traits within the indigenous SA Boerperd, to determine the level of diversity in the breed, and to investigate its relationship with other equine breeds.

The occurrence of four single nucleotide polymorphisms (SNPs) were investigated within the SA Boerperd. Two of these SNPs attribute to height (BIEC2_808543 & BIEC2_1105377) and two influence alternative gaits (*DMRT3_Ser301STOP* & BIEC2_620109). A total of 100 horses originating from different geographical areas within South Africa were included in the study for DNA sequencing. The obtained SNP data revealed that the mutated C-allele of BIEC2_808543 was present within 5% of horses, whilst 57% of the population possessed the A-allele of BIEC2_1105377. It was not possible to assess whether either of these height-associated SNPs cause the formation of tall individuals. The majority of the studied population were three-gaited, whilst 14% of horses were five-gaited. Both gait-associated SNPs occurred at low frequencies within the breed. The *gait keeper* mutation (*DMRT3_Ser301STOP*) had an allelic frequency of 0.105 and the SNP BIEC2_620109 a frequency of 0.091. It was established that neither height-associated SNP are informative within the SA Boerperd breed. Either one of the two gait-associated SNPs can however be screened for, but it is not guaranteed that an individual possessing these mutations will be able to showcase five gaits.

Genotypic data for 363 horses, consisting of 17 microsatellite markers, were obtained from Unistel Medical Laboratories to determine the breed's genetic diversity. The current heterozygote estimate (H_o = 0.679) of the breed was similar to that determined by a study conducted 15 years ago, and the level of inbreeding had decreased to an acceptable 3.2%. The breed's allelic richness had however declined from 4.212 to 3.804. Analysis of the population structure revealed that two distinct subpopulations (K=2) are present within the breed (based on STRUCTURE results). One of the 12 studs had 83.6% of its horses placed in a separate cluster. It was established that a high degree of variation is still present within the SA Boerperd, despite the influence of selective breeding. Levels of inbreeding are at present still manageable.

The genotypic data of the SA Boerperd was compared to that of nine other equine breeds. These consisted of the Tennessee Walker, Andalusian, Standardbred, Appaloosa, Friesian, Hackney, Thoroughbred and Icelandic. Compared to these breeds, the SA Boerperd had above average levels of heterozygosity and a high number of private alleles (17.6%). Both F_{ST}-values and structure analysis suggested that the SA Boerperd is closely related to the Andalusian and Standardbred. Other breeds, specifically of American origin, are thought to have influenced the SA Boerperd and should in future also be investigated.

Results obtained within this study can be used by SA Boerperd breeders to further the development of the breed, whilst simultaneously conserving its genetic potential. Similar studies can also be conducted in other equine breeds to further their conservation and understand their relationship to the SA Boerperd.

Key words: SA Boerperd, performance-associated traits, genetic diversity, microsatellite markers, selective breeding, single nucleotide polymorphisms, subpopulations, inbreeding, height, alternative gaits.

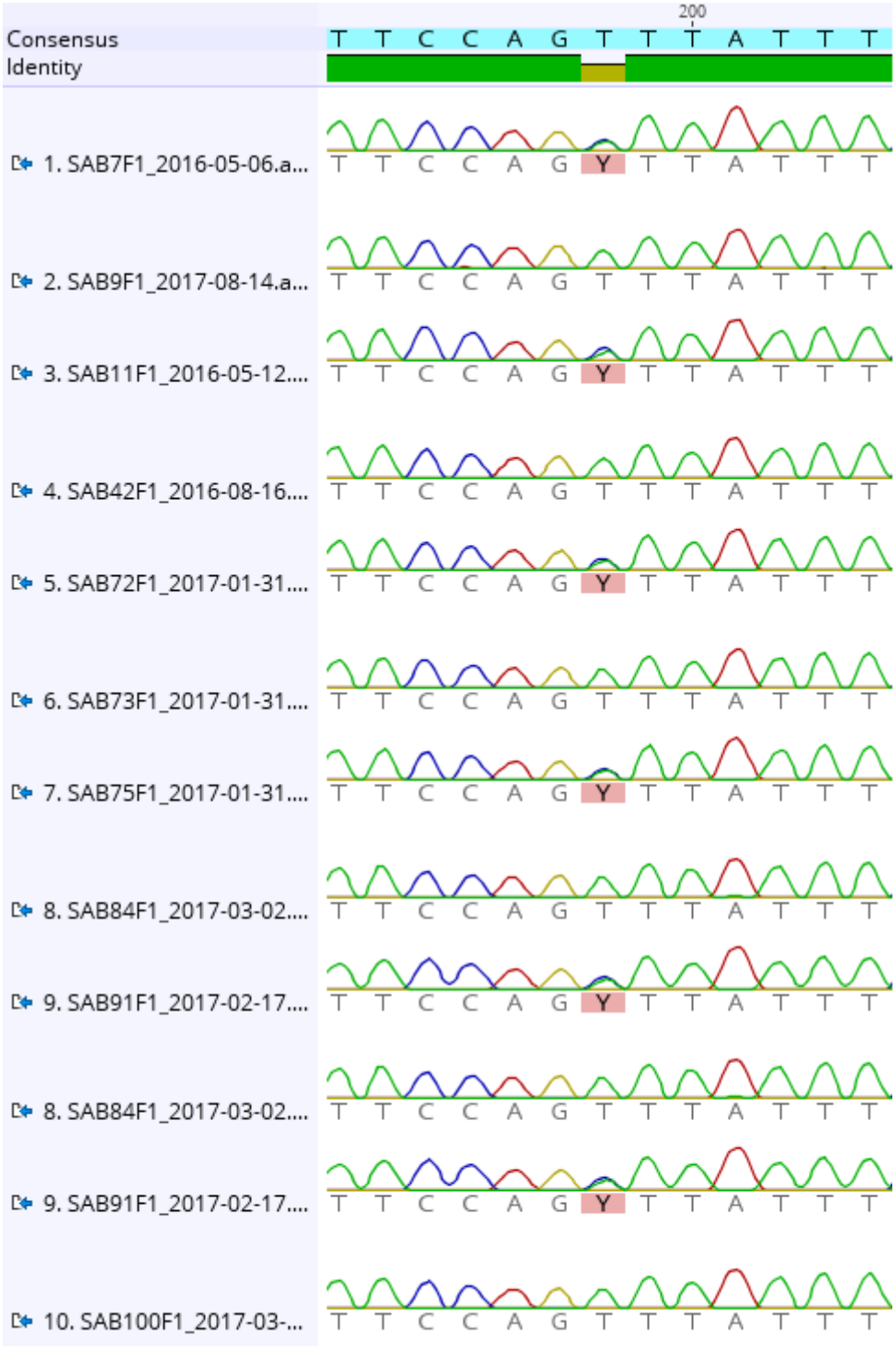
APPENDIX A

SA Boerperd DNA sequence data

SNP-containing sequences for the four performance-associated mutations, viewed in GENEIOUS v. 8.1.7 (Kearse *et al.,* 2012).

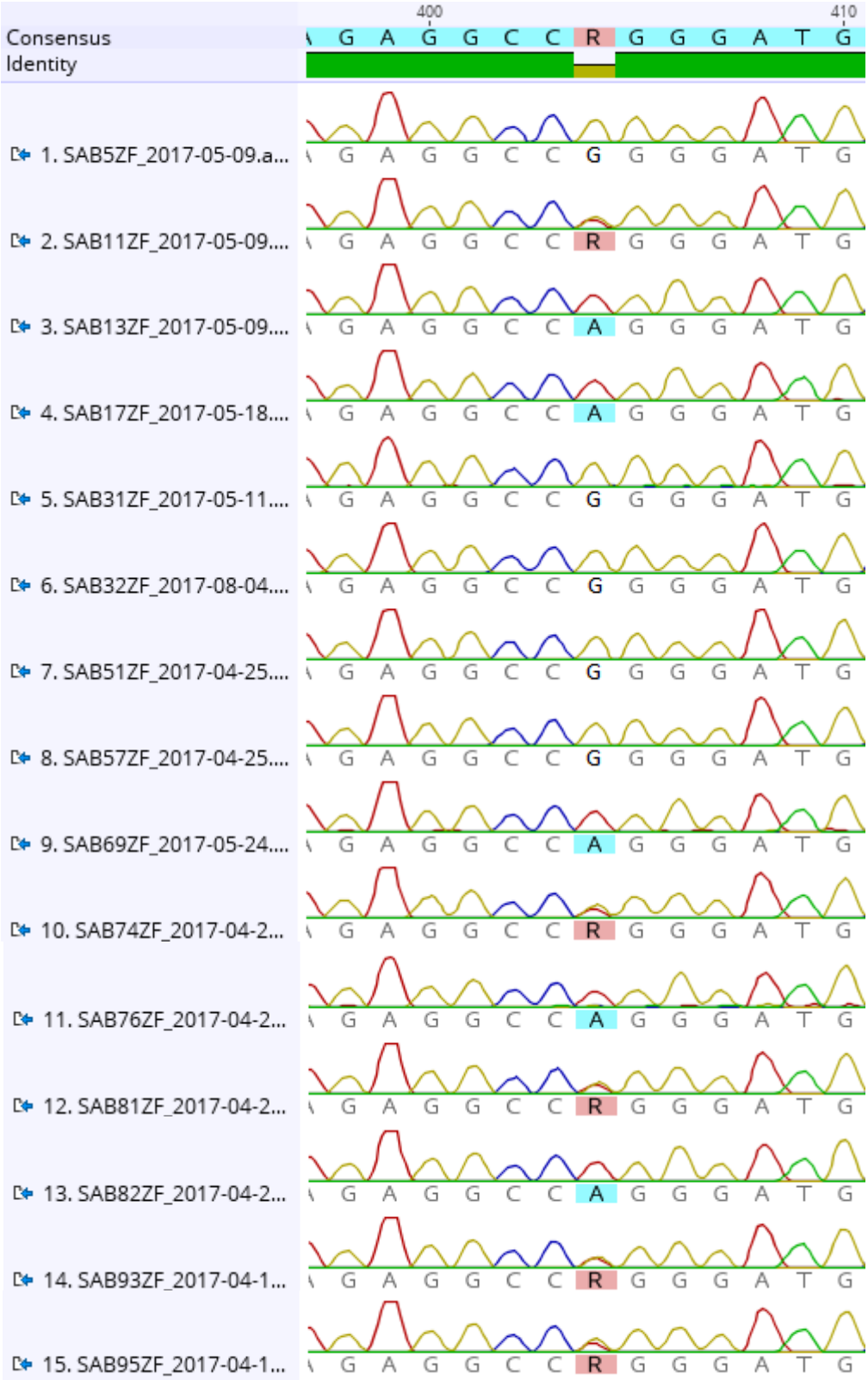
HEIGHT: BIEC2_808543/ECA3: 105 547 002

Partial sequences of samples representative of the identified genotypes are presented. Five of the TT-genotype and five of the TC-genotype. (Note that GENEIOUS labels TC-genotypes as Y). None of the sampled horses possessed the CC-genotype.



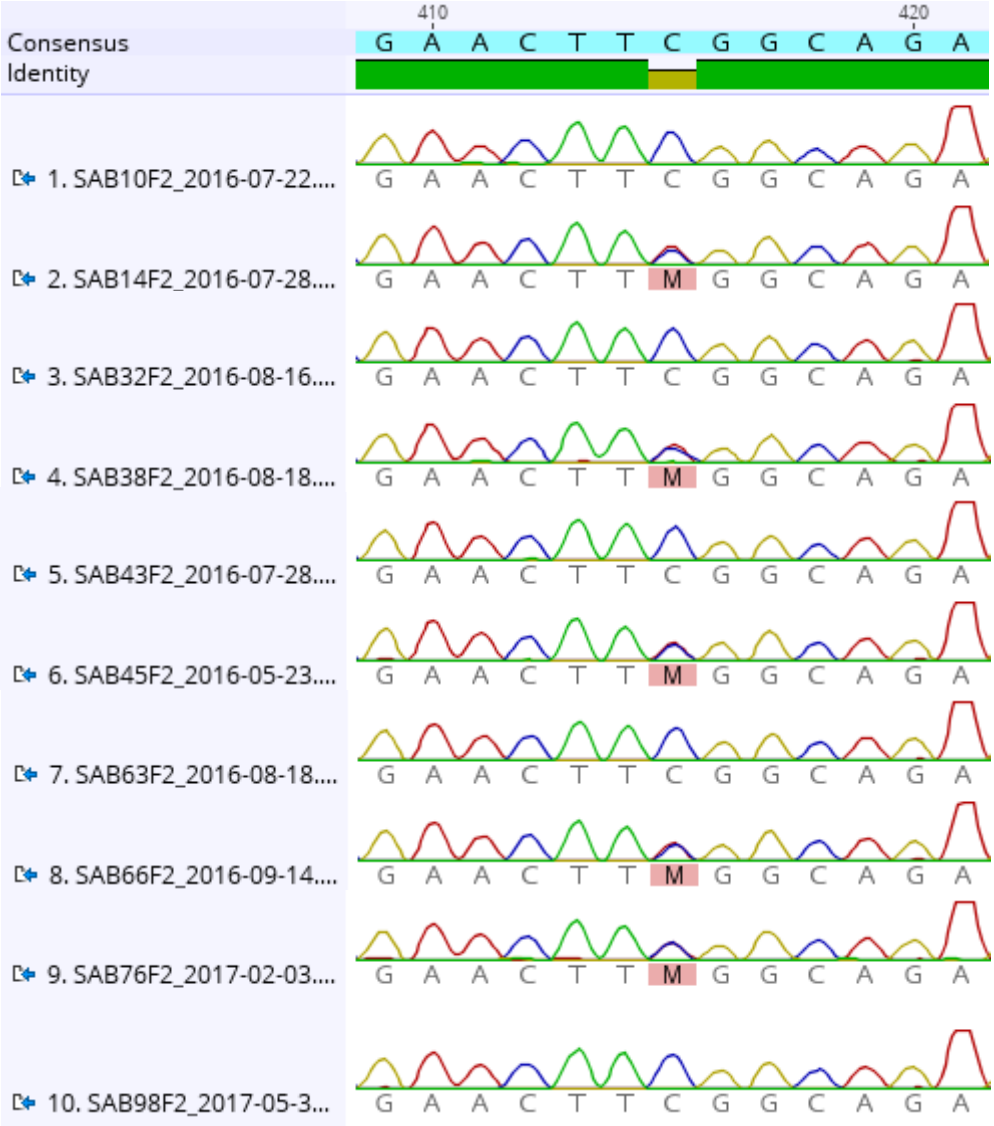
HEIGHT: BIEC2_1105377/ECA9: 74 798 143

Partial sequences of samples representative of the identified genotypes are presented. Five of the GG-genotype, five of the GA-genotype and five of the AA-genotype. (Note that GENEIOUS labels GA-genotypes as R).



GAIT: *DMRT3_Ser301STOP*/ECA23: 22 999 655

Partial sequences of samples representative of the identified genotypes are presented. Five of the CC-genotype and five of the CA-genotype. (Note that GENEIOUS labels CA-genotypes as M). None of the sampled horses possessed the AA-genotype.



GAIT: BIEC2_620109/ECA23: 22 967 656

Partial sequences of samples representative of the identified genotypes are presented. Five of the CC-genotype and five of the CT-genotype. (Note that GENEIOUS labels CT-genotypes as Y). None of the sampled horses possessed the TT-genotype.