Inaugural-Dissertation zur Erlangung der Doktorwürde
der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität
München

Mendelian traits in horse breeding

– studying the background of roan coat color and the distribution and origin of the allele causing Warmblood Fragile Foal Syndrome (WFFS)

von Simone Valerie Reiter

aus Berndorf bei Salzburg, Österreich

München 2022

Aus dem Veterinärwissenschaftlichen Department der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität München

Lehrstuhl für Molekulare Tierzucht und Biotechnologie

Arbeit angefertigt unter der Leitung von: Priv.-Doz. Dr. Ivica Međugorac

Angefertigt am Institut für Tierzucht und Genetik der Veterinärmedizinischen Universität Wien

Mentor: O.Univ.-Prof. Dr.med.vet. Gottfried Brem

Gedruckt mit Genehmigung der Tierärztlichen Fakultät

der Ludwig-Maximilians-Universität München

Dekan: Univ.-Prof. Dr. Reinhard K. Straubinger, Ph.D.

Berichterstatter: Priv.-Doz. Dr. Ivica Međugorac

Korreferent: Univ.-Prof. Dr. Anna May

Tag der Promotion: 30. Juli 2022

Die vorliegende Arbeit wurde gemäß § 6 Abs. 2 der Promotionsordnung für die Tierärztliche Fakultät der Ludwig-Maximilians-Universität München in kumulativer Form verfasst.

Folgende wissenschaftliche Arbeiten sind in dieser Dissertationsschrift enthalten:

1. Grilz-Seger G, Reiter S, Neuditschko M, Wallner B, Rieder S, Leeb T, Jagannathan V, Mesarič M, Cotman M, Pausch H, Lindgren G, Velie B, Horna M, Brem G, Druml T. **A Genome-Wide Association Analysis in Noriker Horses Identifies a SNP Associated With Roan Coat Color**, erschienen im Journal of Equine Veterinary Science 2020, online verfügbar unter doi: 10.1016/j.jevs.2020.102950.

2. Reiter S, Wallner B, Brem G, Haring E, Hoelzle L, Stefaniuk-Szmukier M, Długosz B, Piórkowska K, Ropka-Molik K, Malvick J, Penedo MCT, Bellone RR. **Distribution of the Warmblood Fragile Foal Syndrome Type 1 Mutation (PLOD1 c.2032G>A) in Different Horse Breeds from Europe and the United States**, erschienen in Genes 2020, online verfügbar unter. doi: 10.3390/genes11121518.

TABLE OF CONTENTS

I. INTRODUCTION **1**

II. LITERATURE SURVEY **3**

1. Genome research in horses **3**

2. Applied molecular genetics in horses **4**

3. Classification of genetically determined phenotypic traits **5**

3.1. Qualitative traits 6

3.2. Quantitative traits 7

4. Localising trait causative regions in the genome **7**

4.1. Genome wide association studies (GWAS) 7

4.2. Linkage analysis 8

4.3. Candidate gene approaches 9

5. Applied genetics in horse breeding **10**

6. Mendelian traits in horses **11**

6.1. Genetic diseases 11

6.1.1. Warmblood fragile foal syndrome (WFFS) 14

6.2. Coat color genetics 15

6.2.1. Roan 16

III. STUDY OBJECTIVES **18**

IV.RESULTS **19**

1. Publication I **20**

2. Publication II **28**

V. DISCUSSION **40**

VI.SUMMARY **47**

VII.ZUSAMMENFASSUNG **49**

VIII. REFERENCES **51**

IX.DANKSAGUNG **59**

I. INTRODUCTION

The special relationship between horses and humans dates back around 5000 years, where the first evidence of domestic horses was found in the Eneolithic Botai Culture in Kasachstan (Outram *et al.*, 2009). In the following millennia, horses spread across the world and revolutionized transport, communication, and warfare, and further had a significant impact on the development of agriculture (Kelekna, 2009; Anthony, 2010). Humans soon started to select horses for special traits, by preferring certain characteristics such as strength, aesthetics, racing performance, or endurance (Librado *et al.*, 2016). While coat colors were selected soon after domestication, speed was only likely to have gained importance in the last 1000 years (Ludwig *et al.*, 2009; Fages *et al.*, 2019).

Active selection and human-driven management in recent centuries led to the more than 500 breeds registered today. Significant phenotypic variation is observed among breeds (Rosseau, 2017). For example, an American Miniature horse measures just 74 cm at the withers, while a Shire horse can reach up to 200 cm (Brooks *et al.*, 2010). Additionally, horses' performance have been adopted to certain extremes. Yakutsk horses endure temperatures of –50°C during winter. Arabian horses travel distances up to 160 km in one day, and Thoroughbreds gallop at a speed faster than 60 km/h.

However, the intense selection also has drawbacks. It led to a reduction in genetic diversity and the accumulation of various hereditary diseases in several horse breeds (Fages *et al.*, 2019; Orlando and Librado, 2019). Monogenetically determined (so-called 'Mendelian') genetic defects that are inherited as a recessive trait can spread throughout a population and rise to a high frequency without being noticed.

Today, molecular biological and statistical methods have been established to disentangle the genetic architecture of Mendelian traits in horses. The causal variant in the genome has already been determined for numerous Mendelian traits in horses, facilitating genetic testing and enabling the monitoring of allele distribution and strategic mating. Hence, defining the location of causal variants responsible for Mendelian traits can contribute to avoiding affected animals and maintaining a high genetic diversity within breeds.

This doctoral thesis deals with important aspects of Mendelian traits in horse breeding, spanning the localization of causative regions and defining the causative changes to the development of a test system. Further questions and potentials, that arise for the breeding community after the resolution of the genetic background are discussed.

After a general literature survey, the challenges to localize a causal variation for a Mendelian trait in the genome are shown and discussed for roan coat color in Noriker horses. Roan is a dominant trait and occurs in a wide range of different breeds around the world. The questions that arise after the determination of a recessive deleterious variant are shown and discussed based on the Warmblood Fragile Foal Syndrome (WFFS). WFFS is a lethal disorder, and carrier animals are widespread in diverse warmblood breeds. Two publications form the scientific basis of this thesis.

II. LITERATURE SURVEY

1. Genome research in horses

Almost 30 years have passed since the first time the molecular basis of a genetic disorder was determined in horses. A single base pair change, resulting in an amino acid substitution from phenylalanine to leucine in the *SCN4A* gene, was identified in 1992 as causal for Hyperkalaemic Periodic Paralysis (HYPP) in Quarter Horses (Rudolph *et al.*, 1992). Four years later the genetic background of the coat color chestnut was deciphered (Marklund *et al.*, 1996), and thereafter, a steady increase in resolved causal mutations can be recognized.

In 1995, the Horse Genome Workshop, an international cooperation of more than 100 scientists from 25 countries, was founded with the aim to study and characterize the domestic horse's genome (*Horse Genome Workshop*, 2021). Horse genomic research accelerated in 2005, with the introduction of Next Generation Sequencing (NGS). With this technique, millions of small DNA fragments can be sequenced in parallel (Behjati and Tarpey, 2013; van Dijk *et al.*, 2014).

The first high-quality reference sequence (EquCab2.0) for the domestic horse was generated from the Thoroughbred mare 'Twilight' in 2009, and it incorporated the 31 horse autosomes, the X chromosome, and the mitochondrial genome (Wade *et al.*, 2009). With an approximate size of 2.7 Gb, the horse genome is in the dimension of the dog, bovine, and human genome (Lander *et al.*, 2001; Lindblad-Toh *et al.*, 2005; Elsik *et al.*, 2009).

NGS and additional new genomic technologies have been used to resequence and reassemble the genome of Twilight. This resulted in an improved reference genome, which was published in 2019, EquCab3.0 with 20,955 coding genes annotated to date (Kalbfleisch *et al.*, 2019; *Ensembl genome browser*, 2021).

The availability of a reference genome built the basis for resequencing projects, which have now become effectively feasible. In 2019, a large-scale study entitled 'Comprehensive characterisation of horse genome variation by whole-genome sequencing of 88 horses' was published. In this study, the genomes of 88 horses representing 25 breeds were resequenced, using short-read NGS Illumina technology with an average coverage of 19.1×. In this work, the authors focused on screening the resequenced horses for base substitutions of a single nucleotide at a specific position in the DNA (Single nucleotide variants–SNVs) and short insertions and deletions (Indels). Among the 88 horses sequenced, 23.5 million SNVs and 2.3 million short indels were detected (Jagannathan *et al.*, 2019).

2. Applied molecular genetics in horses

Apart from NGS, Polymorphism screening based on a certain set of preselected markers is a widely used and well-established system to study variations between whole genomes.

Genetic markers are defined polymorphic regions in the genome. The genomic position of genetic markers should be unambiguously defined, and their allelic state can be determined by molecular biological methods. Markers are often located in intergenic or non-coding genic regions, as those regions are presumed to be selectively neutral and make up most of the genome. There are different marker types, such as microsatellites or single nucleotide polymorphisms (SNPs).

Microsatellite markers are based on short repetitive elements in the genome (e.g., TATATATA) that are of great instability; thus, the length is highly variable. The alleles of a microsatellite marker differ in the number of repetitions. Therefore, they reach higher mutation rates, resulting in multiple alleles per locus (microsatellites considered 'multiallelic') (Willam and Simianer, 2011). Microsatellites have played a major role in parentage testing in horses for more than 20 years. A panel of 12 markers, known as the ISAG-panel, was established by the International Society of Animal Genetics (*ISAG*, 2022). Although the ISAG-panel comprises 12 markers, additional markers are included in routine parentage testing across laboratories (Bellone and Avila, 2020).

SNPs are base substitutions of a single nucleotide (SNVs) with the minor allele segregating in a frequency of more than one percent in a population. Most SNPs show two alleles, so they are termed biallelic. The allelic states for SNPs can be determined by different molecular genetic screening approaches. A widely used screening method is the so-called 'SNP array.' This is a fully automatic array technology where the allelic states at up hundred thousand preselected SNPs, usually spread across the entire genome, are determined simultaneously. Today, SNP array genotyping data are the basis for determining regions in the genome that harbor alleles associated with traits of interest. This also involves their application for genome-supported selection in livestock animals (Willam and Simianer, 2011).

Different SNP arrays are available to date, differing in the selection and number of variants they contain and therefore in price. The composition of SNP arrays depends, among other things, on the current knowledge of variation in the genome of a species. This information is regularly supplemented through resequencing studies. To date, three different SNP arrays are available for the horse, whereby the first two arrays, containing 54,602 SNVs (EquineSNP50 BeadChip, Illumina) and 74,500 SNVs (EquineSNP70 Genotyping BeadChip, Illumina),

became available in 2011. The latest array, containing 670,805 SNVs (MNEc670k array, Affymetrix), has been available since 2017 and was developed using whole genome data from 156 horses representing 24 different breeds (Schaefer *et al.*, 2017; Schaefer and McCue, 2020).

3. Classification of genetically determined phenotypic traits

A phenotypic trait is a specific characteristic of an organism that can be a disease, as well as a non-pathogenic attribute. The visible or measurable expression of such a trait is termed a phenotype. The phenotype of an individual is determined by its genotype and the environment (Willam and Simianer, 2011). The genotype of an individual is composed of thousands of variants spread across the genome. In a narrower sense, the term genotype is also used for the allele combination at a defined locus. Among the different types of variants in the genome, SNVs are the most common (Jagannathan *et al.*, 2019). Variants can be located in coding or non-coding regions of genes, as well as in intergenic regions between genes. Exonic SNVs can be either synonymous (also termed silent) or nonsynonymous, depending on the effect the variant has on the protein. Synonymous changes do not affect the protein, and the amino acid is unaltered. Nonsynonymous SNVs can either be missense or nonsense, whereas a missense change results in an alteration of an amino acid, while a nonsense change results in a premature stop codon, which leads to an incomplete protein. The classification of SNVs is illustrated in Figure 1.

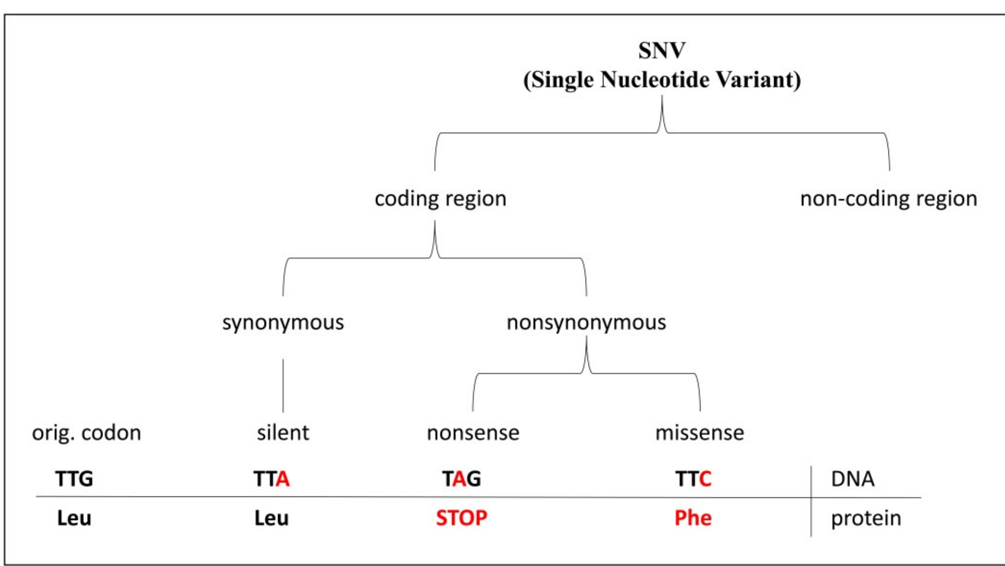

Figure 1: Classification of SNVs

Even SNVs in non-coding regions can influence protein biosynthesis, for example by altering RNA splicing or transcription factor binding efficiency (Abramowicz and Gos, 2018). In

addition to SNVs, there are several other types of variations, such as insertions, deletions, duplications, copy number variants, inversions, and translocations that range from a few up to several thousand base pairs. Variants larger than 1 kb in size are primarily structural variation. Structural variants can also affect the phenotype, for example, the tobiano spotting pattern in horses is caused by a large chromosomal inversion (Brooks *et al.*, 2007).

Depending on the number of loci that contribute to a phenotypic trait, a distinction is made between qualitative and quantitative traits.

3.1. Qualitative traits

Qualitative traits are generally expressed in the form of distinct phenotypes, that can be clearly assigned to a category. In most cases, qualitative traits are regulated by a single variant in the genome and their inheritance follows Mendel`s rules. Those traits are also called 'Mendelian traits.'

For Mendelian traits, one can discern the mode of inheritance by the manifestation of alleles in heterozygous individuals. In a dominant–recessive situation, one allele dominates the effect of the other allele. Hence, the expressed phenotype of one allele (the recessive allele) is masked by the phenotype of another allele, what is called dominance. Dominance is also observed as a continuous scale that ranges from 'complete' dominance through 'incomplete' dominance to 'codominance.' In the case of complete dominance one allele completely masks the phenotype of another allele, while incomplete dominance leads to an incomplete masking effect of the dominant allele, resulting in a phenotype intermediate to both homozygotes. Codominance occurs, when both alleles are phenotypically expressed (Hamilton, 2009). In addition, 'overdominance' can occur when heterozygous individuals show increased fitness compared to individuals that are homozygous for either allele. Therefore, the heterozygous phenotype lies outside the range of both homozygous parents.

Especially in the case of incomplete dominance, the phenotypic manifestation of a known variant can vary. The variation can be due to differences in the genetic background of an individual, thus there may be differences in the phenotypic manifestation of an allele among breeds. An example is Polysaccharide Storage Myopathy (PSSM) Type 1, caused by an incomplete dominant inherited mutation in the *GYS1* gene (McCue *et al.*, 2008). While quarter horses carrying the mutation tend to have severe to lethal forms of PSSM, weaker symptoms have been described in coldbloods. Remarkably, even extremely high frequencies of the mutation, up to 90%, have been observed in some coldblood breeds (Firshman *et al.*, 2003, 2005).

Depending on the genomic location of the causative gene, this can be an autosomal chromosome, a sex chromosome, or mitochondrial DNA, the inheritance is said to be either autosomal, gonosomal or mitochondrial.

Many coat colors and genetic diseases in horses are determined by a single variant in a particular gene and can be categorized as qualitative traits.

3.2. Quantitative traits

Traits that show a continuous phenotypic distribution are defined as quantitative or complex traits. Continuous variation in the expression of the phenotype is due to the interaction of many genes and non-genetic factors. Thus, the underlying genetic architecture of such traits is mostly polygenic. Most performance traits in animals, such as milk yield in cattle, carcass quality in pigs, and height at the withers in horses, fall into this category (Signer-Hasler *et al.*, 2012; Miglior *et al.*, 2017; Bergamaschi *et al.*, 2020). Additionally, some diseases are classified as 'quantitative traits' or 'complex diseases', e.g., equine osteochondrosis (Distl, 2013).

Loci in the genome, that contribute to a certain quantitative trait, are termed as 'Quantitative Trait Loci' (QTL) for the respective trait. QTL have a significant impact on animal breeding, as they are used to estimate genomic breeding values. The interest in genetic markers for various performance traits, such as racing or jumping ability, is also growing in horses, and there have been several studies carried out on this topic. For example, 5 QTL for endurance exercise ability were published in 2017 (Ricard *et al.*, 2017).

4. Localising trait causative regions in the genome

To identify the genetic background of both, qualitative and quantitative traits, two major approaches, genome-wide association studies (GWAS) and linkage analyses, are used. Both are no candidate-driven approaches. For qualitative traits also candidate gene approaches are used.

4.1. Genome wide association studies (GWAS)

In domestic animal genetics, GWAS are mainly based on a set of neutral markers, spread throughout the genome (hence 'genome wide'). Typically, SNP-array data are analysed in a case–control setup, in which individuals are divided into two groups based on their phenotype: one cohort expressing the trait versus a control group that does not. Independent single-locus tests (mostly chi-square tests) are carried out to determine the association of marker alleles with the phenotype (Bush and Moore, 2012). The significance levels of association with the trait, for the variants tested, are usually shown in a so-called 'Manhattan plot.' As shown in Figure 2, the markers are ordered upon their genomic coordinates on the X-axis with the negative

logarithm of the association *P*-value on the Y-axis. As a strong association results in a small *P*-value, the higher the value on the y-axis (negative logarithm of the *P*-value), the greater the probability that an allele of a particular SNP is associated with the phenotype of interest. GWAS has been applied to successfully locate the genomic region harboring the causative allele for numerous Mendelian traits in horses. For example, the causative region of mushroom coat color in Shetland Ponies was identified on chromosome 7 (Tanaka *et al.*, 2019). As shown in Figure 2, several SNPs on chromosome 7 show a more or less strong association with the mushroom phenotype, which is due to genetic linkage. Genetic linkage is the tendency of DNA segments, located in proximity on a chromosome, to be inherited together from a single parent. The closer two loci are on a chromosome, the lower the probability of recombination is between them and the more likely they are to be inherited together through genetic linkage. Linked loci form a so-called haplotype.

GWAS has also been used in horses to determine the genetic variation contributing to complex traits. An example, therefore, is given in Makvandi-Nejad *et al.*, (2012), where 4 loci were determined, explaining 83% of size variation in the horse.

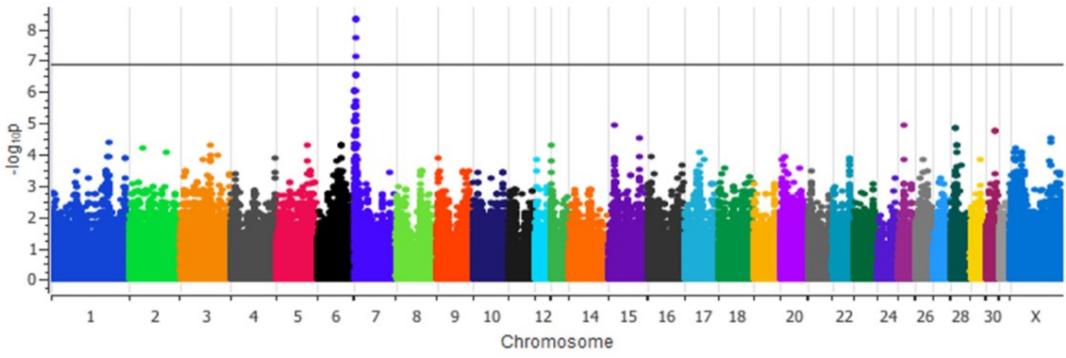

Figure 2: Manhatton plot of a GWAS for mushroom coat color in Shetlandponies. Highly associated loci were identified on chromosome 7. Figure taken from Tanaka et al. (2019).

4.2. Linkage analysis

In contrast to GWAS, where the relatedness among individuals is not considered, linkage analyses deal with the segregation of a phenotype within known families. The most common approach to investigate Mendelian traits is a parametric linkage analysis, that builds upon the specification of an inheritance model for the trait locus. Two genes, MITF and PAX3 were identified as responsible for the splashed white phenotype in a Quarter Horse family based on parametric linkage analysis (Hauswirth *et al.*, 2012).

There are numerous other mapping methods, that use the principle of genetic linkage. One

example is homozygosity mapping, which is based on searching for shared segments of homozygosity that are not identical by state (IBS). Homozygosity mapping is highly effective in the setting of consanguineous parents (Lander and Botstein, 1987; Charlier *et al.*, 2008).

4.3. Candidate gene approaches

Unlike GWAS and linkage analysis, where markers spread throughout the genome are investigated, a focus on particular pre-selected genes is made in the candidate gene approach. Therefore, it is necessary to understand the biological mechanisms underlying the phenotype; otherwise, there is no basis for the selection of a candidate gene (Kwon and Goate, 2000). Once candidate genes are determined, sequence variation within this region is investigated by comparatively sequencing cases. A typical candidate gene for white markings in horses is the *KIT* gene, as several white marking phenotypes are caused by variants in this gene. In Haase *et al.*, (2013), the two candidate genes *KIT* and *MITF* were investigated, to identify alleles causative for white markings in Franches-Montagnes Horses.

GWAS and linkage analysis in domestic animals are usually conducted based on SNP-array data. Accordingly, only the genomic region and the markers associated with the causal variant (the haplotype), and not the causal variant itself, can be identified. Therefore, after the regions are defined by GWAS or linkage analyses, additional analysis steps are needed to identify the causal variant itself. This is usually done by sequencing the associated haplotype, as this approach (today, mostly NGS technologies are used) discloses all variants in the region of interest. Defining the causative variant within the proposed region is often challenging, and success depends on the type of variant. Typically, researchers determine whether there are genes in the predicted region, whose function has already been shown to be related to the observed phenotype in the literature. Sequence variation is screened concisely in those genes and the first focus is on nonsynonymous variants, as for those variants the consequence on the phenotype is easy to predict. A practical example for the detection of a causal variant, located in an exon of a candidate gene, is given in Drögemüller *et al.*, (2009). In this work, a missense mutation in the *SERPINH1* Gene leading to Osteogenesis Imperfecta in Dachshunds was identified. If no exonic variants are observed, intronic and promotor regions are investigated more closely. Intronic variants can affect splicing and variants in promotors can result in altered gene expression levels, leading to alteration in protein integrity and volume (Abramowicz and Gos, 2018). Predicting or even proving the causality of noncoding variants is challenging. Regardless of whether a variant is exonic or intronic, it is difficult to fully prove causality in non-model organisms.

5. Applied genetics in horse breeding

Intense selection within horse breeding led to a drop in diversity, especially in recent centuries. While early horse breeders conserved the level of genetic diversity for several millennia, it has decreased by 16% in the last 200 years (Fages *et al.*, 2019). This decrease in diversity was accompanied by a significant increase of deleterious alleles (Orlando and Librado, 2019). The total genetic burden in a population resulting from the accumulation of deleterious alleles is defined as mutational load. The recent increase in the mutational load is due to a severe drop in effective population size in modern horse breeds due to intensive selection, even involving the mating of related animals termed as 'inbreeding.' Consequently, there has been an accumulation of homozygous regions in the genome in modern horse breeds, also known as extended runs of homozygosity (ROH) (Ceballos *et al.*, 2018; Grilz-Seger, Druml, *et al.*, 2019). The effect of recessive mutations is hidden in the heterozygous state. When recessive alleles are homozygous, the accumulation of ROHs within the genome leads to the expression of recessive disorders and a reduction in fitness, which has been designated as inbreeding depression. The accumulation of deleterious alleles is an undesirable side effect of intensive selection in horse breeding, which we must deal with responsibly. Therefore, it is necessary to study the genetic background of deleterious traits, as this knowledge is fundamental for the development of genetic testing. Genetic testing can be confirmed either directly when the causative allele is determined, or indirectly, in cases where the causal variant could not be explicitly determined, and a marker linked to the causative allele is tested instead (also called a 'marker test'). Genetic testing enables the determination of carrier animals and thus represents the basis for responsible breeding management. By strategically mating carrier animals only with free animals, genetic diversity in a breed can be preserved, as no animals need to be excluded from breeding. At the same time, the birth of affected animals is prevented. Furthermore, the determination of color loci enables strategic mating to obtain offspring of the desired color. In addition to routine examinations as part of breeding, genetic tests are also valuable for individual animals. In the case of disorders, genetic testing can support veterinarian diagnostics and create clarity about further treatment and prognosis.

The relevance of genome research in horses is not limited to the species itself, as horses can be used as a model for human diseases. For example, the horse is an important cancer model, because metastases of tumors occur less frequently in horses than in humans. In 2017, a missense mutation in the *DDB2* gene, associated with ocular squamous cell carcinoma (SCC), was identified in horses. Mutations within the *DDB2* gene have also been linked to a similar disorder in humans, allowing interspecies comparisons (Bellone *et al.*, 2017).

6. Mendelian traits in horses

The website OMIA (*OMIA - Online Mendelian Inheritance in Animals*, 2022) is a compendium of inherited disorders, traits, and genes in many different species. In February 2022, 48 genetic traits of horses with known causal variants were listed in this database. This number appears rather low when compared to other domestic species. By far the most causal variants were clarified in dogs (319), followed by cattle (186) and cats (96). Pigs (58), sheep (49) and chickens (46) are roughly on the same scale as horses, while fewer causal variants for Mendelian traits have been identified in goats (17) and rabbits (13). However, when considering only the identified coat color variants, the horse is placed further ahead (*OMIA - Online Mendelian Inheritance in Animals*, 2022).

A brief overview of the most common monogenetically determined diseases and coat colors in horses, with their genetic background resolved, is given in the following sub-chapters.

6.1. Genetic diseases

To date, 28 monogenic diseases in horses have been clarified and genetic testing is commercially available in Austria and Germany (listed in Table 1). Most of them, namely 21 out of 28, are recessively inherited.

Notably, some genetic disorders or risk factors for diseases are linked to coat color mutations. A prominent example is a di-nucleotide substitution in the EDNRB gene, responsible for the overo spotting pattern. While heterozygous horses show the desired phenotype, horses that are homozygous for this variant are born white or nearly white. Homozygous foals die within a few days after birth through intestinal obstructions, resulting from a lack of nerve cells in the intestine, also referred to as Overo lethal white syndrome (Metallinos et al., 1998). Further examples are Congenital stationary night blindness, which is observed in homozygous carriers of a retroviral long terminal repeat insertion in the *TRPM1* gene that is responsible for the leopard spotting pattern, and the augmented incidence of melanoma in Gray horses (Rosengren Pielberg et al., 2008; Bellone et al., 2013).

Table 1: Disorders for which a genetic test is offered through common providers in Austria and Germany. Additional information was taken from OMIA in November 2021.

Disease	**OMIA ID**	**Gene**	**Type of variant**	**Breeds**
Androgen insensitivity syndrome (AIS)	000991-9796	*AR*	1x regulatory, 2x missense, 2x deletion	Quarter Horse, Tennessee Walking Horse, Thoroughbred, Warmbloods
Cerebellar abiotrophy (CA)	000175-9796	*MUTYH, TOE1*	not definitely clarified	Arabian, Bashkir Curly Horse, Trakehner, Welsh Pony
Dwarfism	002068-9796	*B4GALT7*	splicing	Friesian horse
Dwarfism, ACAN-related	001271-9796	*ACAN*	2x missense, 2x deletion	Miniature Horse, Shetland Pony
Foal immunodeficiency syndrome (FIS)	001578-9796	*SLC5A3*	missense	Dales Pony, Fell Pony
Glycogen branching enzyme deficiency (GBED)	000420-9796	*GBE1*	nonsense	American Quarter Horse
Hereditary equine regional dermal asthenia (HERDA)	000327-9796	*PPIB*	missense	American Quarter Horse
Hoof wall separation disease (HWSD)	001897-9796	*SERPINB11*	insertion (1 bp)	Connemara pony
Hydrocephalus	000487-9796	*B3GALT2*	nonsense	Belgian draft horse, Friesian horse
Hyperkalemic periodic paralysis (HYPP)	000785-9796	*SCN4A*	missense	American Quarter Horse
Idiopathic hypocalcaemia/Hypoparathyroidism	002458-9796	*RAPGEF5*	nonsense	Thoroughbred
Immune-mediated myositis (IMM)	002141-9796	*MYH1*	missense	American Quarter Horse
Junctional epidermolysis bullosa (JEB1), LAMC2-related	001678-9796	*LAMC2*	insertion (1 bp)	Belgian, Italian draft horse, Trait Breton, Trait Comtois
Junctional epidermolysis bullosa (JEB2), LAMA3-related	001677-9796	*LAMA3*	deletion (6589 bp)	American Saddlebred
Malignant hyperthermia (MH)	000621-9796	*RYR1*	missense	American Quarter Horse
Myotonia	000698-9796	*CLCN1*	missense	New Forest pony
Naked foal syndrome (NFS)	002096-9796	*ST14*	nonsense	Akhal-Teke
Occipitoatlantoaxial malformation (OAAM)	000081-9796	*HOXD3*	deletion (2700 bp)	Arabian
Ocular squamous cell carcinoma (OSCC)	000735-9796	*DDB2*	missense	Belgian, Haflinger, Percheron

Disease	OMIA ID	Gene	Type of variant	Breeds
Polysaccharide storage myopathy type 1 (PSSM1)	001158-9796	*GYS1*	missense	American Paint Horse, Appaloosa, Draft, Quarter Horse, Warmblood
Severe combined immunodeficiency disease (SCID)	000220-9796	*PRKDC*	deletion (5 bp)	Arabian
Skeletal atavism (SA)	002013-9796	*SHOX*	2x large deletions	Shetland Pony
Warmblood fragile foal syndrome (WFFS)	001982-9796	*PLOD1*	missense	Warmblood, Thoroughbred, Haflinger, Knabstrupper, American Sport Pony*
related to coat color				
Congenital stationary night blindness (CSNB)/ Leopard Complex Spotting (LP)	001341-9796	*TRPM1*	insertion (1378 bp)	American Miniature Horse, Appaloosa, Australian Spotted Pony, British Spotted Pony, Knabstrupper, Noriker, Pony of the Americas, Thoroughbred
Lavender foal syndrome (LFS)	001501-9796	*MYO5A*	deletion (1 bp)	Arabian
Melanoma/ Gray	001356-9796	*STX17*	duplication	Many
Multiple Congenital Ocular Anomalies (MCOA)/ Silver	001438-9796	*PMEL*	missense	American Miniature Horse, Icelandic, Rocky Mountain
Overo lethal white syndrome (OLWS)	000629-9796	*EDNRB*	substitution (2 bp)	American Paint Horse, Miniature Horse, Pinto Horse, Quarter Horse, Thoroughbred

*breeds added from Bellone et al. (2020) and Reiter et al. (2020)

6.1.1. Warmblood fragile foal syndrome (WFFS)

WFFS is an autosomal recessive disorder that was first detected in warmblood horses. Affected foals are aborted during a late stage of gestation or are born with skin abnormalities and extreme flexibility in their digital joints. The phenotype of an affected foal is shown in Figure 3. In 2011, the causative mutation, a nonsynonymous SNV in the *procollagen-lysine-2-oxoglutarate-5-dioxygenase 1* gene (PLOD1:c.2032G>A, p.Gly678Arg), was identified (Winand, 2011). A base substitution from guanine to adenine results in an altered amino acid sequence of the polypeptide chain, with a severe impact. Homozygosity for this *PLOD1* variant seems to be incompatible with extra-uterine life.

Figure 3: Phenotype of a WFFS-affected foal. The foal had poor muscular development and protruding bony prominences (A), severe skin ruptures beneath the right eye (B), at the neck (C) and ventral abdomen (D). Figure taken from Metzger et al. (2020).

Foals born alive must be immediately euthanized (Aurich *et al.*, 2019). WFFS carriers were identified in a broad range of warmblood breeds with an average carrier frequency of 11% (Winand, 2011; Dias *et al.*, 2019). Apart from warmbloods, the WFFS allele was also detected in Thoroughbreds with a carrier frequency of 2.4% (Bellone et al., 2020). Possible origins of this disease have been discussed, including the Thoroughbred and the Arabian horse. Pedigree analysis of around 2000 horses showed that the Thoroughbred stallion Dark Ronald (1905–1928) was the only common ancestor of all carrier animals (Wobbe *et al.*, 2019). However,

Dark Ronald was refuted as the founder of WFFS. The WFFS allele was not detected in the skin of Dark Ronald, preserved at a museum, after sequencing the DNA isolated from this sample (Zhang *et al.*, 2020). In addition to Dark Ronald, the Arabian stallion Bairactar Or. Ar. (1813) was also hypothesized as a founder of WFFS, as symptoms similar to WFFS were recognised in offspring, inbred to Bairactar in the 19th century (Rueff, 1855).

6.2. Coat color genetics

Although an old wise saying is 'a good horse has no color,' horses were selected for their colors immediately after domestication (Ludwig *et al.*, 2009). Today, special-colored horses are extremely popular. Understanding the genetic background of colors in the horse is of great economic importance in breeds that are focused on rare coat colors. Commercial coat color tests are already available for a broad range of phenotypes (Sponenberg and Bellone, 2017; Bellone and Avila, 2020).

Coat color phenotypes of horses can roughly be divided into three categories, which are base coat colors, dilutions, and white patterns.

The base coat colors black, bay, and chestnut are regulated by the two genes melanocortin 1 receptor (*MC1R*) and agouti signalling protein (*ASIP*). While *MC1R*, also known as the Extension locus, specifies if the black pigment eumelanin or the red pigment pheomelanin is produced, *ASIP*, also known as Agouti locus, defines where the pigment is produced, on the whole body or restricted to the points (Marklund *et al.*, 1996; Wagner and Reissmann, 2000; Rieder *et al.*, 2001).

Dilutions, as the term implies, lighten the base colors bay, chestnut, or black and manifest as pale colors. The genes and alleles determining these special colors have already been described, and six 'dilution genes' have been reported in horses. The most common dilutions are line backed duns (*TBX3* locus) and cream-related (*SLC45A2* locus) colors (Mariat et al., 2003; Imsland *et al.*, 2016). Other dilutions, such as champagne (*SLC36A1* locus), silver dapple (*PMEL* locus), pearl (*SLC45A2* locus), and mushroom (*MFSD12* locus), are less frequent (Brunberg *et al.*, 2006; Cook *et al.*, 2008; Holl *et al.*, 2019; Tanaka *et al.*, 2019).

White patterns emerge through white hairs that are superimposed over the base coat color. The expression of white patterns is variable and ranges from small white markings on the head and legs to completely white horses. Any combination of white patterns with any base colors and dilutions is possible. Thus far, more than 40 mutations in 7 different genes have been described that contribute to white patterns in horses. It is also important to mention, that some white patterns are accompanied by pathogenic phenotypes (see chapter genetic diseases above). An

overview of the most common white patterns and the underlying genes is given in Figure 4.

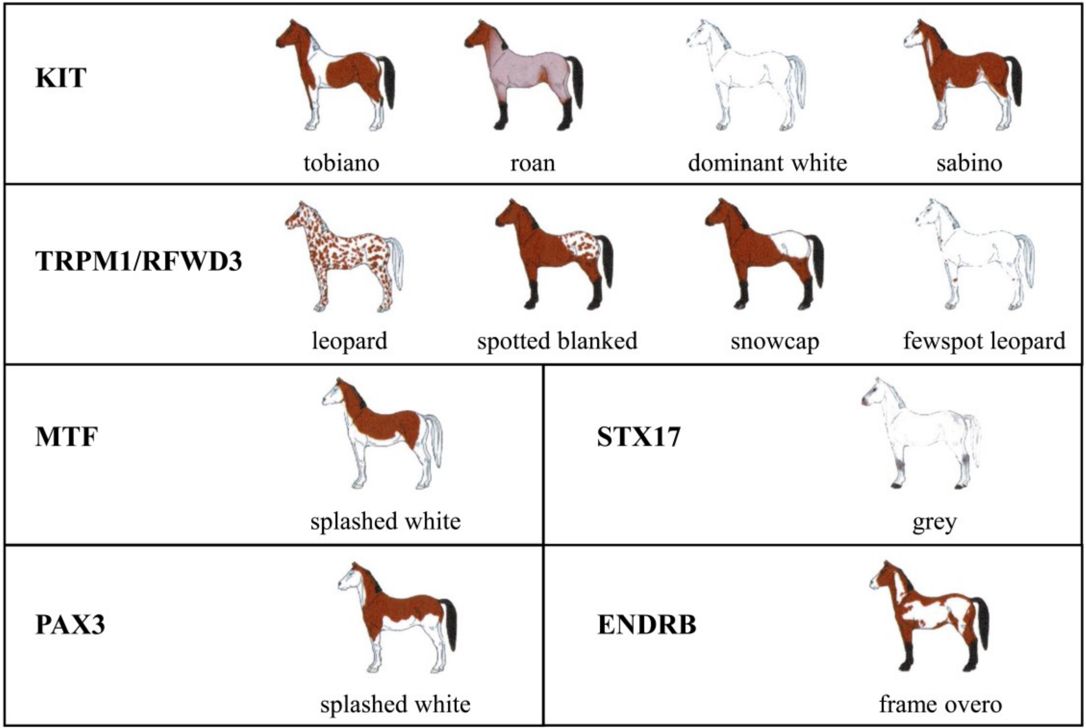

Figure 4: Overview of common white spotting patterns in the horse grouped according to the causal genes. Adapted from Figure 9.2 in Sponenberg and Bellone 2017.

6.2.1. Roan

The roan phenotype is a white spotting pattern characterized by intermixed white and colored hairs in the body, while the head, lower legs, mane, and tail remain colored (Figure 5). Roan can appear on any base color and is termed accordingly as blue roan (black), red roan (bay), and strawberry roan (chestnut). The roan phenotype is widespread and occurs in different pony, warmblood, and coldblood breeds (Sponenberg and Bellone, 2017; Grilz-Seger and Druml, 2018). In contrast to the phenotype 'gray,' where white hairs are equally distributed across the body and increase from year to year, roan is not progressive. A seasonal color change, however, is clearly pronounced. Therefore, in some breeds, including the Icelandic horse, roan is also designated as color-changing roan. Blue roans are common in the Austrian Noriker breed and are named 'Blauschimmel' or 'Mohrenkopf.' The mode of inheritance of roan is autosomal dominant. The genetic basis of roan was investigated in several studies, and the causal region has been mapped to equine chromosome 3, close to the KIT gene (Marklund et al., 1996; Grilz-Seger et al., 2019; Voß et al., 2020). Although several variants were detected to be associated with roan in diverse breeds, the causal one remains unknown. A marker test for this trait, informative in Quarter Horses, is commercially available (*Roan Zygosity Test/UC Davis*, 2021).

Figure 5: Blue roan Noriker mare with intermixed white hair on the body. Picture: Simone Reiter

III. STUDY OBJECTIVES

Mendelian traits are phenotypes that are determined by a single variant in the genome. As the understanding of Mendelian traits forms an important basis of genetic selection in animal breeding, effort has been made to resolve their causal variants. Depending on the type of variant, its detection and the proof of causality can be a challenging process, but once the genetic background of a trait is resolved, genetic testing is feasible. Genetic tests identify carrier animals of special traits and therefore enable straightforward selection. For example, certain coat colors can be promoted, or genetic disorders avoided. The reduction of hereditary defects in a population is an important point in terms of animal welfare and can be managed by suitable breeding programs.

In this work the significance of monogenic traits in horse breeding is illuminated based on two concrete publications, dealing with certain questions of this field. Within the first publication the genetic background of the roan coat color in Austrian Noriker horses was investigated. The second publication deals with the spread of hereditary defects and the potential of extended screening studies for horse breeding.

Publication I (co-first Authorship):

A Genome-Wide Association Analysis in Noriker Horses Identifies a SNP Associated With Roan Coat Color

In this article the associated region for roan coat color was confirmed near the *KIT* gene using GWAS. Based on NGS data from a homozygous roan Noriker horse, an attempt was made to identify the causal variant. Potential variants in the haplotype block associated with roan in the Noriker horse were checked by genotyping them in a random sample set of roan and non-roan animals in Noriker and other horse breeds.

Publication II (first Authorship):

Distribution of the Warmblood Fragile Foal Syndrome Type 1 Mutation (PLOD1.2032G>A) in Different Horse Breeds from Europe and the United States

In this study, the distribution of the WFFS allele was screened in a range of horse breeds from different countries. Breeds where WFFS was not suspected, and a historic DNA sample were included. The genotyping results provide data for breeders and breeding associations to determine whether genetic testing for WFFS in breeding animals is sensible. The DNA investigation from a historical museum sample enables conclusions about the origin of WFFS.

IV. Results

Two publications included in this thesis are presented within the results section. Both publications will not be part of any other doctoral thesis. References of the individual manuscripts are formatted in the style of the journal and are not included in the general reference section at the end of this thesis. Numeration of figures and tables refers to the respective journal article.

1. Publication I

A Genome-Wide Association Analysis in Noriker Horses Identifies a SNP Associated With Roan Coat Color

Gertrud Grilz-Seger, Simone Reiter, Markus Neuditschko, Barbara Wallner, Stefan Rieder, Tosso Leeb, Vidhya Jagannathan, Matjaz Mesarič, Markus Cotman, Hubert Pausch, Gabriella Lindgren, Brandon Velie, Michaela Horna, Gottfried Brem, Thomas Druml

Journal of Equine Veterinary Science, February 2020

PMID: 32303326

DOI: 10.1016/j.jevs.2020.102950

Journal of Equine Veterinary Science 88 (2020) 102950

Contents lists available at ScienceDirect

Journal of Equine Veterinary Science

journal homepage: www.j-evs.com

Original Research

A Genome-Wide Association Analysis in Noriker Horses Identifies a SNP Associated With Roan Coat Color

Gertrud Grilz-Seger [a, 1, *], Simone Reiter [a, 1], Markus Neuditschko [b], Barbara Wallner [a], Stefan Rieder [c], Tosso Leeb [d], Vidhya Jagannathan [d], Matjaz Mesarič [e], Markus Cotman [f], Hubert Pausch [g], Gabriella Lindgren [h, i], Brandon Velie [j], Michaela Horna [k], Gottfried Brem [a], Thomas Druml [l]

[a] Department of Animal Breeding and Genetics, University of Veterinary Sciences Vienna, Vienna, Austria
[b] Agroscope, Avenches, Switzerland
[c] Identitas AG, Bern, Switzerland
[d] Department of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland
[e] Clinic for Reproduction and Large Animals, University of Ljubljana, Veterinary Faculty, Ljubljana, Slovenia
[f] Department for Preclinical Sciences, University of Ljubljana, Veterinary Faculty, Ljubljana, Slovenia
[g] Animal Genomics, ETH Zürich, Lindau, Switzerland
[h] Department of Animal Breeding & Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden
[i] Livestock Genetics, Department of Biosystems, KU Leuven, Leuven, Belgium
[j] School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales, Australia
[k] Department of Animal Husbandry, Slovak University of Agriculture in Nitra, Nitra, Slovakia
[l] Lipizzan stud farm of Piber, Köflach, Austria

ARTICLE INFO

Article history:
Received 21 December 2019
Received in revised form
28 January 2020
Accepted 28 January 2020
Available online 5 February 2020

Keywords:
Roan
Horse
Coat color
KIT
GWAS
Genetic marker

ABSTRACT

The roan coat color in horses is characterized by dispersed white hair and dark points. This phenotype segregates in a broad range of horse breeds, while the underlying genetic background is still unknown. Previous studies mapped the roan locus to the *KIT* gene on equine chromosome 3 (ECA3). However, this association could not be validated across different horse breeds. Performing a genome-wide association analysis (GWAS) in Noriker horses, we identified a single nucleotide polymorphism (SNP) (ECA3:g.79,543.439 A > G) in the intron 17 of the *KIT* gene. The *G* -allele of the top associated SNP was present in other roan horses, namely Quarter Horse, Murgese, Slovenian, and Belgian draught horse, while it was absent in a panel of 15 breeds, including 657 non-roan horses. In further 379 gray Lipizzan horses, eight animals exhibited a heterozygous genotype (*A/G*). Comparative whole-genome sequence analysis of the *KIT* region revealed two deletions in the downstream region (ECA3:79,533,217_79,533,224delTCGTCTTC; ECA3:79,533,282_79,533,285delTTCT) and a 3 bp deletion combined with 17 bp insertion in intron 20 of *KIT* (ECA3:79,588,128_79,588,130delinsTTATCTCTA-TAGTAGTT). Within the Noriker sample, these loci were in complete linkage disequilibrium (LD) with the identified top SNP. Based upon pedigree information and historical records, we were able to trace back the genetic origin of roan coat color to a baroque gene pool. Furthermore, our data suggest allelic heterogeneity and the existence of additional roan alleles in ponies and breeds related to the English

Author's contribution: Grilz-Seger, Gertrud: conceptualization, writing - original draft, project administration; Reiter Simone: comparative sequence analysis, genotyping, fine-scale mapping; Neuditschko, Markus: formal analysis, visualization; Wallner, Barbara: comparative sequence analysis, fine-scale mapping; Leeb, Tosso: sequencing NO180, Jagannathan, Vidhya.: sequencing NO180, Mesaric, Matjaz: sampling, Cotman, Marko: sampling, DNA extraction; Pausch, Hubert: formal analysis, Lindgren, Gabriella: providing SNP data; Velie, Brandon: providing SNP data; Brem, Gottfried: funding acquisition; Druml, Thomas: formal analysis, Visualization, pedigree analysis, funding acquisition; all authors: reading, commenting and reviewing of the final manuscript draft.

Animal welfare/ethical statement: This study was discussed and approved by the institutional ethics and animal welfare committee in accordance with GSP guidelines and national legislation in protocols by the institutional Commission for Ethics and Animal Welfare, University of Veterinary Medicine, Vienna (ETK-194/12/2019, ETK-06/05/2015) and by the Ethics Committee for Animal Experiments in Uppsala, Sweden (Number: C 121/14). Hair samples for Slovenian Coldblooded Horse, Bosnian Mountain Horse, Posavina, Quarter Horse were collected in the context of routine procedures during the studbook registration of horses by the Institute for Breeding and Health Care of Horses of the Veterinary Faculty, Ljubljana.

Conflict of interest statement: None of the authors has conflicts of interest.

* Corresponding author at: Gertrud Grilz-Seger, Institute of Animal Breeding and Genetics, Department for Biomedical Sciences; University of Veterinary Medicine Vienna, Veterinärplatz 1, A-1210 Vienna, Austria.
E-mail address: gertrud.grilz@vetmeduni.ac.at (G. Grilz-Seger).
[1] These two authors equally contributed to this work.

https://doi.org/10.1016/j.jevs.2020.102950
0737-0806/© 2020 Elsevier Inc. All rights reserved.

Thoroughbred. In order to study the roan phenotype segregating in those breeds, further association and verification studies are required.

© 2020 Elsevier Inc. All rights reserved.

1. Introduction

The knowledge of coat color genetics in horses rapidly progressed within the last decades. One milestone was the detection of a single missense variant in the equine melanocortin-1-receptor gene (*MC1R*), causative for chestnut coat color by Marklund et al. [1]. With the identification of an 11-bp deletion in exon 2 of the agouti-signaling-protein *(ASIP)* gene associated with black coat color by Rieder et al. [2], the basic colors of horses could be described on a molecular level. Our knowledge regarding diluted colors was deepened by a series of publications by Mariat et al. (cream, *SLC45A2/MATP*) [3], Imsland et al. (dun, *TBX3*) [4], Cook et al. (champagne, *SLC36A1*) [5], and Brunberg et al. (silver, *PMEL*) [6]. In 2008 Pielberg et al. [7] successfully identified the causative variant for gray coat color, whereas for the high phenotypic variability of leopard spotting (*LP*) patterns up to date two genetic factors were identified: the causative variant for *LP* [8] and the modifier pattern1 (*PATN1*) [9]. The increasing scientific interest in the white patterns in horses started with the assignment of the Overo spotting pattern to *ENDRB* [10]. In on-going research, the *KIT* gene encoding the KIT proto-oncogene, receptor tyrosine kinase, was found to harbor numerous genetic variants, which are responsible for a wide range of white coat color patterns. To date, the patterns Sabino1 [11], Tobiano [12], Dominant White (up to 28 different alleles within *KIT*) were described by [13–23]. Splashed White patterns were found to be caused by variants in *MITF* or *PAX3* [17,24].

The causative genetic background for roan coat color, which segregates in a wide range of horse breeds, is still unknown. Marklund et al. [25] proposed *KIT* as a candidate gene for roan coat color and identified an association between several *KIT* variants and the roan coat color in Belgian draught horses. The proposed variants could not be verified in other horse breeds (Welsh, Shetland, and Gotland Pony), exhibiting a roan phenotype, which led the authors to conclude that the genetic background of roan coat color may be heterogeneous across breeds. Furthermore, Marklund et al. [25] confirmed the dominant mode of inheritance of the roan allele *Rn* and claimed the lethality of homozygous roan horses, which was already proposed by Hinz and VanVleck [26]. The hypothesis of the lethality in homozygous roan horses currently has been disputed by Sponenberg and Bellone [27], who postulate the existence of homozygous roan horses.

Performing a genome-wide association analysis (GWAS) in Noriker horses, we identified an SNP (AX-103594067, ECA3:g.79,543,439A > G) in intron 17 of the *KIT* gene with the *G*-Allele associated with roan coat color [28]. Pedigree analysis and offspring ratios supported the homozygous state of *G/G* roan Noriker horses. Based on these results, a homozygous G/G roan Noriker stallion was identified as a reference animal for further whole-genome sequence analysis [29]. The aim of this study was to

Table 1
Description of the samples used for this study, including the data source for the top-SNP AX-103594067(ECA3:g.79,543,439A > G) and the three genetic variants A2 (3:79533214), A3 (3:79533281), B1 (3:79588127–29).

Non-roan horses	coat color	N	data source top-SNP	data source genetic variants
Akhal Teke	non-roan	36	SNP chip	
Appaloosa	non-roan	2		genotyping all 4 variants
Belgian draught horse	roan	1		genotyping all 4 variants
Bosnian Mountain Horse	non-roan	23	SNP chip	
Exmoor Pony	non-roan	256	SNP chip	
Franches Montagnes	non-roan	80	SNP chip	
German Sport horse	non-roan	1		genotyping all 4 variants
German Sport horse	roan	3		genotyping all 4 variants
Gidran	non-roan	20	SNP chip	
Lipizzan black/bay	non-roan	10	SNP chip	
Lipizzan, gray*	non-roan	379	SNP chip	
Murgese	non-roan	2		genotyping all 4 variants
Murgese	roan	2		genotyping all 4 variants
Noriker	non-roan	31		genotyping all 4 variants
Noriker	non-roan	7	SNP chip	genotyping all 4 variants
Noriker	non-roan	119	SNP chip	
Noriker	roan	26		genotyping all 4 variants
Noriker	roan	11	genotyping top-SNP	
Noriker	roan	14	SNP chip	
Posavina	non-roan	28	SNP chip	
Quarter Horse	non-roan	3		genotyping all 4 variants
Quarter Horse	non-roan	2		genotyping all 4 variants
Quarter Horse	roan	5		genotyping all 4 variants
Quarter Horse	roan	2	SNP chip	genotyping all 4 variants
Shagya Arabian	non-roan	33	SNP chip	
Shetland Pony	non-roan	2		genotyping all 4 variants
Shetland Pony	roan	1		genotyping all 4 variants
Shetland Pony	roan	2	SNP chip	genotyping all 4 variants
Slovenian draught horse	roan	3	SNP chip	genotyping all 4 variants
Trakehner	roan	1		genotyping all 4 variants
Trotter	non-roan	1		genotyping all 4 variants
All		1106		

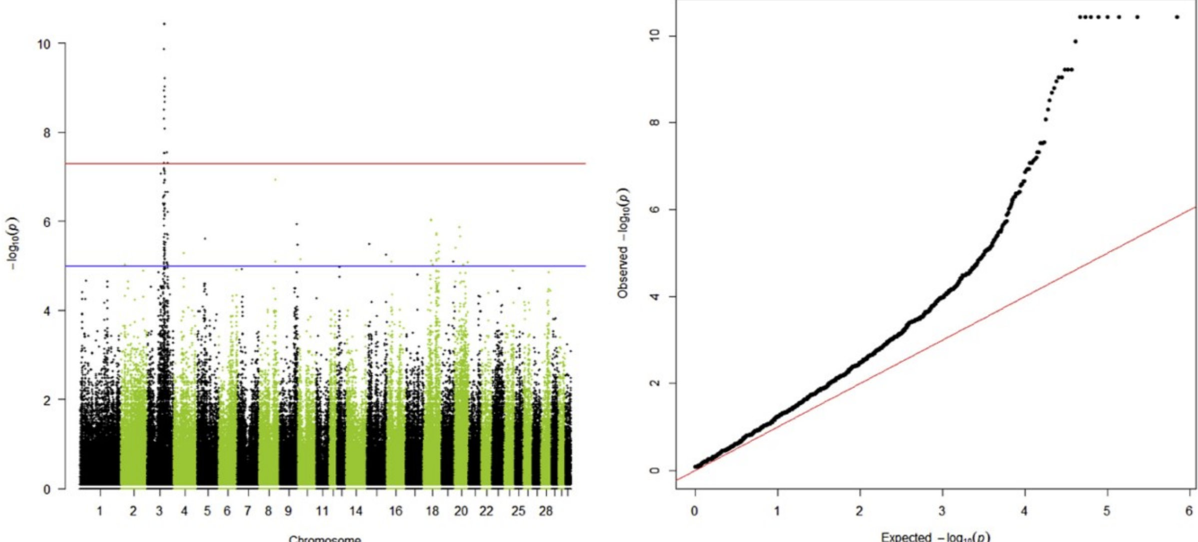

Fig. 1. Manhattan plot and Quantile-Quantile plot for the GWAS of roan versus black Noriker horses.

Table 2
Genotypes for the top-SNP AX-103594067(ECA3:g.79,543,439A > G) for 1035 non-roan/71 roan horses, derived from 670k SNP data and genotyping with KASP technology.

Non-roan horses	N	Genotype *A/A*	Genotype *A/G*	Genotype *G/G*
Appaloosa	2	2	0	0
Akhal Teke	36	36	0	0
Shagya Arabian	33	33	0	0
Franches Montagnes	80	80	0	0
Bosnian Mountain Horse	23	23	0	0
Posavina	28	28	0	0
Gidran	20	20	0	0
Lipizzan black/bay	10	10	0	0
Lipizzan, gray[a]	379	371	8	0
Quarter Horse	3	3	0	0
Quarter Horse	2	2	0	0
Noriker	31	31	0	0
Noriker	7	7	0	0
Noriker	119	119	0	0
Exmoor Pony	256	256	0	0
Shetland Pony	2	2	0	0
German Sport horse	1	1	0	0
Murgese	2	2	0	0
Trotter	1	1	0	0
All[b]	1035	1027	8	0
Roan horses				
Noriker	51	0	36	15
Quarter Horse	7	0	7	0
Slovenian draught horse	3	0	3	0
Murgese	2	0	2	0
Belgian draught horse	1	0	0	1
Shetland Pony	3	3	0	0
Trakehner	1	1	0	0
German Sport horse	3	3	0	0
All[b]	71	7	48	16

[a] for the 379 gray Lipizzan horses the underlying coat color was not known, but roan color is known to segregate in a specific mare family.
[b] Fisher's exact test for genotype association (A/A and G/-) with roan/non roan coat color for the entire sample revealed a *P*-value $P < .001$.

further explore the genome region around the associated top-SNP AX-103594067 and to better understand the underlying genetic background of roan coat color in horses.

2. Material and Methods

2.1. Genome-wide SNP (Single Nucleotide Polymorphism) Data and Genome-wide Association Analysis

From previous studies [30–33] 670k SNP data for a total of 1012 horses was available, including the following breeds: 140 Noriker, 36 Akhal Teke, 33 Shagya Arabian, 80 Franches Montagnes, 23 Bosnian Mountain Horse, 28 Posavina, 20 Gidran, 389 Lipizzan, two Quarter horses, two Shetland Ponies, three Slovenian draught horses, and 256 Exmoor Ponies (Table 1).

The GWAS was performed, including 40 Noriker horses using a case-control design (case-group: 14 blue roan Noriker horses, control-group: 26 black Noriker horses) and associations were corrected for multiple testing using the Fisher's exact test. SNP extraction and GWAS were performed using the software package PLINK v.1.7 [34]. Quality control (QC) and SNP filtering were applied on the Noriker data set using a minor allelic frequency (MAF) threshold of <0.01. After QC, a total of genome-wide 464,880 SNPs were included in the GWAS.

For the verification of genotype and haplotype distribution of roan/non-roan horses of the entire multibreed data set, we applied Fisher's exact test. Statistical analyses and graphical representations were performed using the R-platform (www.r-project.com).

2.2. Genotyping

For further verification of our top SNP (AX-103594067) and further putative variants, we genotyped 97 horses (roan/non-roan), including the following breeds: Noriker (26/38), Quarter horse (7/5), Appaloosa (0/2), Murgese (2/2), Belgian draught horse (1/0), Slovenian draught horse (3/0), Shetland Pony (3/2), Trakehner (1/0), Trotter (0/1), Warmblood (3/1) (Table 1). In addition, we genotyped 11 archived roan Noriker horses only for the top SNP (AX-

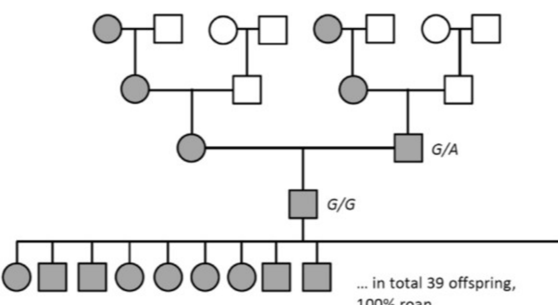

Fig. 2. Pedigree information of the roan stallion NO180 given as an example for the segregation of roan coat color in Noriker horses. Stallions are illustrated by squares, mares by circles; roan phenotype is represented by gray color, non-roan phenotype by white color. On the bottom NO180, homozygous *G/G* at the top SNP AX-103594067 (ECA3:g.79,543,439A > G), and his 100% roan offspring, including 39 foals are shown.

103594067). Genomic DNA was isolated from hair roots using nexttecTM Tissue & Cells Kit, following the manufacturer's protocol. For genotyping, competitive allele-specific PCR SNP genotyping assays (KASP) were used. LGC KASP assays were designed to genotype five variants (Table 3) using the LGC service (http://www.lgcgroup.com). KASP screening was performed as described by the supplier on a CFX96 Touch Real-Time PCR Detection System.

2.3. Pedigree- and ROH Analysis

The genotype plausibility (checking sire and dam coat color, offspring ratios) of 51 roan Noriker horses were examined, including the pedigree information of 55.567 horses from the Noriker studbook [35]. Furthermore, we performed a ROH analysis for eight roan Noriker horses homozygous *G* for the top-SNP AX-103594067 with an overlapping window approach as implemented in PLINK v1.7 [34] based on the following settings: minimum SNP density of ROH segments was set to one SNP per 50 kb with a maximum gap length of 50 kb; final segments were called runs of homozygosity (ROH) if the minimum length of the homozygous segment was greater than 125 kb and comprised more than 20 homozygous SNPs; one heterozygote and one missing genotype were permitted within each segment.

The whole-genome sequence of one blue roan (black base color) Noriker stallion (NO180), verified as homozygous for the roan factor [28], was available from [29]. The variant callings from NO180 (published in [29]) in the ROH region of homozygous Noriker horses, spanning the *KIT* gene (ECA3:79,472,667–79,694,802), were comparatively analyzed with the sequences published for 87 horses from different breeds [29] (4 Akhal-Teke, 4 American Paint Horse, 1 American Standardbred, 2 Arabian, 1 Polish Warmblood, 3 German Warmblood, 29 Franches-Montagnes, 4 Haflinger, 2 Hannoveraner, 8 Holsteiner, 2 German Riding Pony, 1 Morgan Horse, 4 Icelandic, 1 Dutch Warmblood, 3 Oldenburger, 3 Quarter Horse, 4 Swiss Warmblood, 3 Shetland Pony, 1 Thoroughbred, 3 Trakehner, 2 UK Warmblood, 2 Welsh Pony). For the majority of these horses (84), information on coat color phenotypes was provided for the present study (Suppl. 1). Among the 84 horses with color phenotypes and whole-genome information [29], three Quarter horses were roan. However, although the roan coat color is known in a variety of horse breeds, the overall incidence of roan horses in the sample set of the 88 horses with whole-genome information [29] remained low. Accordingly, we filtered the variant calls published in [29] for positions on ECA3:79,472,667–79,694,802 that are homozygous in NO180 and called at maximum in five additional horses.

Finally, a comparison analysis between the variants identified by Marklund et al. [25] in the Belgian draught horse and the variants inferred in this study was conducted.

3. Results

The GWAS in 40 Noriker horses revealed eight SNPs on ECA3 that were significantly associated with roan coat color. These variants were located in the *KIT* genic region (Fig. 1). The top-SNP (AX-103594067, ECA3:g.79,543,439A > G,) was either heterozygous or homozygous for the alternate allele in 64 roan horses (51 Noriker, three Slovenian draught horses, seven Quarter horses, two Murgese horses, and one Belgian draught horse), whereas all non-roan horses were homozygous for the wildtype-allele (*A/A*) (Table 2). One-third of the roan Noriker horses (n = 15) and one Belgian draught horse were homozygous for the roan associated *G* allele (Table 2). The *G* allele was not present in three roan Shetland Ponies, one roan Trakehner, and three roan German Sport horses. In Table 2, the genotype distribution for the top-SNP AX-103594067 of 1036 non-roan horses and 71 roan horses is illustrated. Among 379 gray Lipizzan horses, eight heterozygous *A/G* carriers could be detected (Table 2).

Pedigree analyses of roan Noriker horses were in concordance with the genotyping results and supported a homozygous state and the dominant mode of inheritance. Eight of the investigated roan Noriker horses with homozygous state *G/G* of the top SNP AX-103594067 had 72 documented offspring, whereas 70 horses were roan, one colt was leopard-spotted, and one foal was tobiano. In Fig. 2 the pedigree information of the sequenced Noriker stallion (NO180) is presented. Based on 670K SNP chip data from eight roan Noriker horses homozygous *G* for the top-SNP AX-103594067, we identified a 222 kb homozygous region from ECA3:79,472.667–79,694.802 containing the entire *KIT* gene. The comparative sequence analysis of the region ECA3:79,472,667–79,694,802 in 88 horses [29], resulted in a total of 1552 variants. Assuming that the roan factor should be present in a homozygous state in animal NO180, 314 variants remained, from which 88 variants were private for this animal. Applying a filter that allows five animals to share these variants, 115 candidate variants were retained. Finally, we selected two regions (A and B) harboring structural variants as putative candidate loci for roan coat color. In the 3′-flanking region of *KIT* (region A), three deletions were identified and in intron 20 of the *KIT* gene (Region B), a 3 bp

Table 3
Location, reference allele (REF) and alternate allele (ALT) for the three deletions found in region A (A1, A2, A3), the top-SNP AX-103594067 from the GWAS (SNV) and the insertion in region B (B1) associated with roan coat color.

Variant	Position on EquCab3	REF	ALT (NO180)
A1	ECA3:79,531,997_79,532,011	TTCCATGATTAATTA	—
A2	ECA3:79,533,217_79,533,224	CTCGTCTT	—
A3	ECA3:79,533,282_79,533,285	TTCT	—
SNV	ECA3:79,543,439	A	G
B1	ECA3:79,588.127–29	ACA	TTATCTCTATAGTAGTT

Table 4
Genotyping results of the four genotyped variants and respective haplotypes (homozygous for reference allele (Hom.REF), homozygous for alternate allele (hom.ALT) or heterozygous (Het.), roan horses without the roan haplotype, are marked with*.

Non-roan horses	n	Haplotype	A2 (3:79533214)	A3 (3:79533281)	SNP (3:79543439)	B1 (3:79588127–29)
Noriker	38	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
Quarter Horse	5	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
Appaloosa	2	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
Murgese	2	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
Trotter	1	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
German Sport horse	1	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
Shetland Pony	2	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
All[a]	51					
Roan horses						
Noriker	18	11-1-1/00-0-0	Het.	Het.	Het.	Het.
Noriker	8	11-1-1/11-1-1	Hom.ALT	Hom.ALT	Hom.ALT	Hom.ALT
Quarter Horse	7	11-11/00-0-0	Het.	Het.	Het.	Het.
Slov.Coldblood	3	11-1-1/00-0-0	Het.	Het.	Het.	Het.
Belgian Draught	1	11-1-1/11-1-1	Hom.ALT	Hom.ALT	Hom.ALT	Hom.ALT
Murgese	2	11-1-1/00-0-0	Het.	Het.	Het.	Het.
Shetland Pony*	3	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
Trakehner*	1	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
German Sport horse*	3	00-0-0/00-0-0	Hom.REF	Hom.REF	Hom.REF	Hom.REF
All[a]	46					

[a] Fisher's exact test for haplotype association (Hom.REF and Hom.ALT/Het.) with roan/non-roan coat color for the entire sample revealed a P-value $P < .001$.

deletion combined with 17 bp insertion was detected (Table 3). All five variants were homozygous for the alternate allele (ALT) in the Noriker stallion NO180. LGC KASP assays were designed for the genotyping of the associated top SNP AX-103594067 from the 670K SNP Chip, the three deletions located downstream of the *KIT* gene (A1-A3) and the 3 bp deletion combined with 17 bp insertion in intron 20 of *KIT* (B1). The genotyping results revealed that the deletion A1 was not roan specific, and was, therefore, excluded from further analyses.

The remaining two deletions (A2 and A3) and the 3 bp deletion combined with 17 bp insertion (B1) were solely detected in roan horses for the breeds Noriker, Quarter Horse, Murgese, Slovenian and Belgian draught horse. These three loci were in complete LD with the top SNP AX-103594067 at ECA3g:79,543,439A > G and the resulting roan-associated haplotype was assigned 11-1-1 (variants phased, REF = 0, ALT = 1) (Table 4., Fig. 3). This roan haplotype was not present in three roan German Sport Horses and one roan Trakehner and three roan Shetland Ponies. All these animals exhibited the same haplotype (00-0-0) as non-roan horses.

The comparison of the identified variants with the regions associated with roan coat color by Marklund et al. [25] showed a close proximity, especially for the SNP AX-103594067 (Table 5), and two roan associated sites of Marklund et al. [25] (ECA3:79,545,912G > A in site SSCP, ECA3:79,540,501G > A in site TaqI) were confirmed in the homozygous stallion NO 180 (Table 5).

4. Discussion

For a long time, the homozygous state of a dominant roan coat color allele was supposed to be lethal in utero according to

Fig. 3. Roan color phenotype associated with different roan haplotypes: (1) Murgese stallion (haplotype: 11-1-1/00-0-0), (2) Lipizzan colt before greying (haplotype: 11-1-1/00-0-0), (3) Noriker stallion (haplotype: 11-1-1/11-1-1), (4) Quarter Horse (haplotype: 11-1-1/00-0-0) and (5) Slovenian draught horses (haplotype: 11-1-1/00-0-0). The Shetland Pony in (6) is not carrying the reported roan haplotype (haplotype: 00-0-0/00-0-0) (images by Grilz-Seger, Mesaric, archive Druml).

Table 5
Distance between the top SNP AX-103594067 and the sites associated with roan coat color by Marklund et al. (1999).

ID	Study	position	Distance to top-SNP	NO180	REF	ALT
CfoI	Marklund et al. (1999)	ECA3:79,538,738	4701 bp	Hom.REF	C	T
TaqI	Marklund et al. (1999)/this study	ECA3:79,540,501	2938 bp	Hom.ALT	G	A
SNV	this study	ECA3:79,543,439		Hom.ALT	A	G
SSCP	Marklund et al. (1999)/this study	ECA3:79,545,912	2473 bp	Hom.ALT	G	A

offspring ratios [25,26]. Sponenberg and Bellone [27] proposed the existence of living homozygous Rn/Rn roan horses, which was confirmed in our study.

Marklund et al. (1999) [25] reported an incomplete association between roan phenotype and genotype information in 33 roan and 92 nonroan horses, mapped to the *KIT* gene. With the investigation of additional breeds (Welsh, Shetland, and Gotland Pony), the level of association declined, which led the authors to assume genetic heterogeneity. Our study recapitulates these findings as the genetic association for roan coat color was not complete. However, we found a complete association between a roan haplotype in Noriker and further draught horses (Belgian and Slovenian draught horse), as well as Murgese, and seven Quarter horses of our dataset. The identified roan haplotype was not present in three roan Shetland Ponies, three German Warmblood horses, and one Trakehner. The three roan Quarter horses published in [29] did also not exhibit the roan haplotype. Our four sampled roan German Sport horses and Trakehner horses all descended from the same stallion and might carry a very rare roan allele. The existence of breed-specific segregation of coat color alleles is well known for Sabino (*SB1*) and chestnut (*MC1R*) [11,36]. The *SB1* allele associated with the Sabino phenotype was firstly detected in the Tennessee Walking horse [11], and it was further documented in numerous breeds, including American Miniature, Paint Horse, Azteca, Missouri Foxtrotter, Shetland Pony, and Spanish Mustang. Brooks et al. [11] further demonstrated that within breeds with the classical Sabino phenotype like Shire Horse or Clydesdale, the SB1 allele did not segregate, a result that was also confirmed by Reissmann et al. [37]. Further allelic heterogeneity at the KIT locus was reported for Dominant White, where several breeds or even family-specific mutations are responsible for depigmented phenotypes in horses [38].

The breed-specific segregation of roan associated alleles may be explained by the population history of the breeds. Roan coat color represented a major breeding objective in Old-Italian and Old-Spanish horse populations before 1800 [39,40]. The breeds Lipizzan, Murgese, Noriker, Slovenian draught horse (mainly derived from Noriker horses), Belgian draught horse, and Quarter Horse, are directly connected to this baroque gene pool or founder populations are proven by either pedigree or historical records.

For the draught horse breeds in this sample, a direct connection to the baroque gene pool can be found in imperial stud farms established in original breeding areas of respective breeds: for example, in the clerical stud farm Rif near Salzburg, which interacted with local breeding (horse in Noriker-like type), roan horses were already recorded in the year 1652 [41]. A similar gene-flow between a baroque gene pool and local working horse populations of later on Belgian draught horse type like horses and introgression of roan can be assumed from the imperial stud farm Alost, founded by the Austrian empire in Belgium in 1770 [42].

In our sample, Shetland Ponies and breeds derived from English Thoroughbred were characterized by the absence of the roan associated *G*-allele and haplotype, and thus, the roan phenotype in these breeds may be traced back to another founder population. Introgression of single dominant inherited traits was mainly conducted using single breeding animals, especially stallions. Y-chromosomal studies [43] revealed that Northern European Pony breeds and the Thoroughbred cluster show distinct population history and genetic distances to the core clusters, which mainly evolved from a unique group, including the descendant breeds from a prior baroque gene pool.

5. Conclusion

Our results suggest the SNP AX-103594067 (ECA3g:79,543,439A > G) in the *KIT* gene as a genetic marker for roan coat color in Noriker horses and as a putative marker in draught horse breeds, Lipizzans and Murgese. Our data supports the previously postulated allelic heterogeneity in roan horses, as the marker was not identified in roan Shetland ponies and roan Thoroughbred-related breeds. To validate the marker and the associated roan haplotype in other breeds than Noriker, we suggest further investigation within a bigger panel comprising additional multibreed samples.

Acknowledgments

We want to thank the Austrian, Slovenian horse breeding associations and Viktoria Remer for supporting data collection, Annik Gmel for DNA extraction. The Exmoor Pony Society in the UK is thanked for the sampling of horses.

Funding: This work was financially supported by the Austrian Research Promotion Agency (FFG), Contract number 843464; the Federal Ministry for Sustainability and Tourism (BMNT), Contract number 101332, and Slovenian Research Agency program P4-0053 to M. Cotman. Data collection and genotyping of the Exmoor Pony received funding from the European Union's Seventh Framework Program managed by REA-Research Executive Agency; http://ec.europa.eu/research/rea (FP7/2007–2013) under grant agreement no. 606142 and The Exmoor Pony Society in the UK, and grant 31003A_172964 of the Swiss National Science Foundation.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jevs.2020.102950.

References

[1] Marklund L, Moller MJ, Sandberg K, Andersson L. A missense mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with the chestnut coat color in horses. Mamm Genome 1996;7:895–9.

[2] Rieder S, Taourit S, Mariat D, Langlois B, Guerin G. Mutation in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (Equus caballus). Mamm Genome 2001;12:450–5.

[3] Mariat D, Taourit S, Guérin G. A mutation in the MATP gene causes the cream coat color in the horse. Genet Sel Evol 2003;35:119–33.

[4] Imsland F, McGowan K, Rubin CJ, Henegar C, Sundström E, et al. Regulatory mutations in TBX3 disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses. Nat Genet 2016;48:152–8.

[5] Cook D, Brooks S, Bellone R, Bailey E. Missense mutation in exon 2 of SLC36A1 responsible for champagne dilution in horses. PLoS Genet 2008;4:e1000195.

[6] Brunberg E, Anderson L, Cothran G, et al. A missense mutation in PMEL17 is associated with the Silver coat color in the horse. BMC Genet 2006;7:46.

[7] Pielberg RG, Golovko A, Sundström E, Curik I, Lennartsson J. A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. Nat Genet 2008;40:1004–9.
[8] Bellone R, Holl H, Setulari V, et al. Evidence for a retroviral insertion in TRPM1 as the cause of congenital stationary night blindness and leopard complex spotting in the horse. PLoS One 2013;8:e78280.
[9] Holl HM, Brooks SA, Archer S, Brown K, Malvick J, Penedo MC, Bellone RR. Variant in the RFWD3 gene associated with PATN1, a modifier of leopard complex spotting. Anim Genet 2016;47:91–101.
[10] Yang GC, Croaker D, Zhang A, Manglick P, Cartmill T, Cass D. A dinucleotide mutation in the endothelian-B receptor gene is associated with letal white foal syndrome (LWFS) – a horse variant of Hirschprung-disease (HSCR). Hum Mol Genet 1989;7:1047–52.
[11] Brooks S, Bailey E. Exon skipping in the KIT gene causes a Sabino spotting pattern in Horses. Mamm Genome 2005;16:893–902.
[12] Brooks S, Lear T, Adelson L, Bailey E. A chromosome inversion near the KIT gene and the Tobiano spotting pattern in horses. Cytogenet Genome Res 2007;119:225–30.
[13] Haase B, Brooks SA, Schlumbaum A, Azor PJ, Bailey E, Alaeddine F, Mevissen M, Burger D, Poncet PA, Rieder S, Leeb T. Allelic heterogeneity at the equine KIT locus in dominant white (W) horses. PLoS Genet 2007;3:e195.
[14] Haase B, Brooks SA, Tozaki T, Burger D, Poncet PA, Rieder S, Hasegawa T, Penedo C, Leeb T. Seven novel KIT mutations in horses with white coat color phenotypes. Anim Genet 2009;40:623–9.
[15] Holl H, Brooks S, Bailey E. De novo mutation of KIT discovered as a result of a non-hereditary white coat color pattern. Anim Genet 2010;41(Suppl 2):196–8.
[16] Haase B, Rieder S, Tozaki T, Hasegawa T, Penedo MC, Jude R, Leeb T. Five novel KIT mutations in horses with white coat color phenotypes. Anim Genet 2011;42:337–9.
[17] Hauswirth R, Jude R, Haase B, Bellone RR, Archer S, Holl H, Brooks SA, Tozaki T, Penedo MC, Rieder S, Leeb T. Novel variants in the KIT and PAX3 genes in horses with white-spotted coat color phenotypes. Anim Genet 2013;44:763–5.
[18] Haase B, Jagannathan V, Rieder S, Leeb T. A novel KIT variant in an Icelandic horse with white-spotted coat color. Anim Genet 2015;46:466.
[19] Holl HM, Brooks SA, Carpenter ML, Bustamante CD, Lafayette C. A novel splice mutation within equine KIT and the W15 allele in the homozygous state lead to all white coat color phenotypes. Anim Genet 2017;48:497–8.
[20] Dürig N, Jude R, Holl H, Brooks SA, Lafayette C, Jagannathan V, Leeb T. Whole genome sequencing reveals a novel deletion variant in the KIT gene in horses with white spotted coat color phenotypes. Anim Genet 2017;48:483–5.
[21] Capomaccio S, Milanesi M, Nocelli C, Giontella A, Verini-Supplizi A, Branca M, Silvestrelli M, Cappelli K. Splicing site disruption in the KIT gene as strong candidate for white dominant phenotype in an Italian Trotter. Anim Genet 2017;48:727–8.
[22] Hoban R, Castle K, Hamilton N, Haase B. Novel KIT variants for dominant white in the Australian horse population. Anim Genet 2018;49:99–100.
[23] Hug P, Jude R, Henkel J, Jagannathan V, Leeb T. A novel KIT deletion variant in a German Riding Pony with white-spotting coat color phenotype. Anim Genet 2019;50:761–3.
[24] Hauswirth R, Haase B, Blatter M, Brooks SA, Burger D, Drögemüller C, et al. Mutations in MITF and PAX3 cause "splashed white" and other white spotting phenotypes in horses. PLoS Genet 2012;8:e1002653.
[25] Marklund S, Moller M, Sandberg K, Andersson L. Close association between sequence polymorphism in the KIT gene and the roan coat color in horses. Mamm Genome 1999;10:283–8.
[26] Hintz HF, Van Vleck LD. Lethal dominant roan in horses. J Hered 1979;70:145–6.
[27] Sponenberg DP, Bellone R. Equine color genetics. 4th ed. Ames: Wiley-Blackwell; 2017.
[28] Grilz-Seger G, Neuditschko M, Wallner B, Horna M, Brem G, Druml T. Association of roan coat color with KIT in Noriker draught horses. Prag: Poster, Animal Science Days; 2019.
[29] Jagannathan V, Gerber V, Rieder S, Tetens J, et al. Comprehensive characterization of horse genome variation by whole-genome sequencing of 88 horses. Anim Genet 2018;50:74–7.
[30] Grilz-Seger G, Druml T, Neuditschko M, Mesarič M, Cotman M, Brem G. Analysis of ROH patterns in the Noriker horse breed reveal signatures of selection for coat color and body size. under review. Anim Genet 2019;50:334–46.
[31] Grilz-Seger G, Druml T, Neuditschko M, Dobretsberger M, Horna M, Brem G. High-resolution population structure and runs of homozygosity reveal the genetic architecture of complex traits in the Lipizzan horse. BMC Genomics 2019;20:174.
[32] Grilz-Seger G, Neuditschko M, Ricard A, Velie BD, Lindgren G, et al. Genome-wide homozygosity patterns and evidence for selection in a set of European and near Eastern horse breeds. Genes (Basel) 2019;10:491.
[33] Gmel A, Druml T, Niederhäusern R, Leeb T, Neuditschko M. Genome-wide association studies based on equine joint angle measurements reveal new QTL affecting the conformation of horses. Genes (Basel) 2019;10:370.
[34] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.
[35] ARGE Noriker. PDV Database export. Stadl-Paura, Austria: Arbeitsgemeinschaft (ARGE) Noriker; 2019.
[36] Wagner HJ, Reissmann M. New polymorphism detected in the horse MC1R gene. Anim Genet 2000;31:289–90.
[37] Reissmann M, Musa L, Zkizadech S, Ludwig A. Distribution of coat-color-associated alleles in the domestic horse population and Przewalski's horse. J Appl Genet 2016;57:519–25.
[38] Haase B, Brooks SA, Schlumbaum A, Azor PJ, Bailey E, Alaeddine F, Mevissen M, Burger D, Poncet PA, Rieder S, Leeb T. Allelic heterogeneity at the equine KIT locus in dominant white (W) horses. PLoS Genet 2007;3:e195.
[39] Buonavolontà G, Silvestrelli M, editors. Il murgese. Italy: Equestri, Milano; 1986.
[40] Grilz-Seger G, Druml T. Lipizzaner hengststämme. Graz, Austria: Vehling Verlag; 2011.
[41] Grilz-Seger G, Druml T. Das Noriker Pferd. Band II. Graz, Austria: Vehling Verlag; 2012.
[42] De Theulegoet H. Monographie des Belgischen Lastpferdes. Germany: Hannover; 1912.
[43] Wallner B, Palmieri N, Vogl C, Rigler D, Bozlak E, Druml T, Jagannathan V, et al. Y chromosome uncovers the recent oriental origin of modern stallions. Curr Biol 2017;27:2029–2035.e5.

2. Publication II

Distribution of the Warmblood Fragile Foal Syndrome Type 1 Mutation (PLOD1.2032G>A) in Different Horse Breeds from Europe and the United States

Simone Reiter, Barbara Wallner, Gottfried Brem, Elisabeth Haring, Ludwig Hoelzle, Monika Stefaniuk-Szmukier, Bogusława Długosz, Katarzyna Piórkowska, Katarzyna Ropka-Molik, Julia Malvick, Maria Cecilia T Penedo, Rebecca R Bellone

Genes, December 2020
PMID: 33353040
DOI: 10.3390/genes11121518

Article

Distribution of the Warmblood Fragile Foal Syndrome Type 1 Mutation (PLOD1 c.2032G>A) in Different Horse Breeds from Europe and the United States

Simone Reiter [1], Barbara Wallner [1], Gottfried Brem [1], Elisabeth Haring [2,3], Ludwig Hoelzle [4], Monika Stefaniuk-Szmukier [5], Bogusława Długosz [5], Katarzyna Piórkowska [6], Katarzyna Ropka-Molik [6], Julia Malvick [7], Maria Cecilia T. Penedo [7] and Rebecca R. Bellone [7,8,*]

1 Institute of Animal Breeding and Genetics, University of Veterinary Medicine Vienna, 1210 Vienna, Austria; simone.reiter1@gmail.com (S.R.); barbara.wallner@vetmeduni.ac.at (B.W.); gottfried.brem@bremkg.de (G.B.)
2 Central Research Laboratories, Museum of Natural History, 1010 Vienna, Austria; elisabeth.haring@nhm-wien.ac.at
3 Department of Evolutionary Biology, University of Vienna, 1090 Vienna, Austria
4 Institute of Animal Science, University of Hohenheim, 70599 Stuttgart, Germany; Ludwig.Hoelzle@uni-hohenheim.de
5 Department of Animal Reproduction, Anatomy and Genomics, University of Agriculture in Kraków, al. Mickiewicza 24/28, 30-059 Kraków, Poland; monika.stefaniuk-szmukier@urk.edu.pl (M.S.-S.); boguslawa.dlugosz@urk.edu.pl (B.D.)
6 Department of Animal Molecular Biology, National Research Institute of Animal Production, Krakowska 1, 32-083 Balice, Poland; katarzyna.piorkowska@izoo.krakow.pl (K.P.); katarzyna.ropka@izoo.krakow.pl (K.R.-M.)
7 Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, CA 95616, USA; jzcollins@ucdavis.edu (J.M.); mctorrespenedo@ucdavis.edu (M.C.T.P.)
8 Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA 95616, USA
* Correspondence: rbellone@ucdavis.edu; Tel.: +1-(530)-752-9299

Received: 9 November 2020; Accepted: 16 December 2020; Published: 18 December 2020

Abstract: Warmblood fragile foal syndrome (WFFS) is an autosomal recessive disorder caused by a single nucleotide variant in the *procollagen-lysine-2-oxoglutarate-5-dioxygenase 1* gene (PLOD1:c.2032G>A, p.Gly678Arg). Homozygosity for the PLOD1 variant causes an Ehler-Danlos-like syndrome, which has to date only been reported in warmblood breeds but the WFFS allele has been also detected in the Thoroughbred. To investigate the breed distribution of the WFFS allele, 4081 horses belonging to 38 different breeds were screened. In total, 4.9% of the horses representing 21 breeds carried the WFFS allele. The affected breeds were mainly warmbloods, with carrier frequency as high as 17% in the Hanoverian and Danish Warmblood. The WFFS allele was not detected in most non-warmblood breeds. Exceptions include WFFS carriers in the Thoroughbred (17/716), Haflinger (2/48), American Sport Pony (1/12), and Knabstrupper (3/46). The origin of the WFFS allele remains unknown. The Arabian breed and specifically the stallion Bairactar Or. Ar. (1813), whose offspring were reported to have a similar phenotype in the 19th century, were hypothesized as the origin. DNA from a museum sample of Bairactar Or. Ar. showed that he did not carry the mutated allele. This result, together with the genotypes of 302 Arabians, all homozygous for the reference allele, does not support an Arabian origin of the WFFS allele. Our extensive survey shows the WFFS allele to be of moderate frequency and concern in warmbloods and also in breeds where it may not be expected.

Keywords: PLOD 1; warmblood fragile foal syndrome; Arabians; museum sample

Genes **2020**, *11*, 1518

1. Introduction

Ehler-Danlos syndrome (EDS) is a heterogeneous group of connective tissue disorders in humans caused by mutations in at least 20 different genes [1]. Characteristic symptoms of EDS are fragility of the soft connective tissues with widespread manifestations in skin, ligaments and joints, blood vessels, and internal organs [2]. In the last several decades, EDS-like syndrome was described in a broad range of mammals [3] as well as in several horse breeds including heavy horses [4], warmblood horses [5–8], Arabians [9,10], Quarter horses [11–23] and Thoroughbreds [24].

Two single nucleotide variants (SNVs) are documented as causal for EDS-like syndromes in horses. The first, hereditary equine regional dermal asthenia (HERDA), is a degenerative skin disease occurring in the Quarter horse and Quarter Horse-related breeds caused by a missense mutation (PPIB: c.115 G>A) in the cyclophilin B (PPIB) gene [25].

The second EDS-like syndrome was named Warmblood fragile foal syndrome type 1 (WFFS) because it was initially identified in warmblood breeds. It is proposed to be an autosomal recessive disorder caused by a SNV in the procollagen-lysine-2-oxoglutarate-5-dioxygenase 1 gene (PLOD1:c.2032G>A, p.Gly678Arg) [26,27]. Homozygosity for this PLOD1 variant (WFFS allele) is thought to be incompatible with extra-uterine life with a total of 18 cases described to date in the literature [26–29]. The predominant manifestation is death during the later stages of gestation (n = 4) or live births with foals being non-viable or euthanized because of poor prognosis (n = 14) [28]. Affected foals show skin abnormalities, including hyperextensible and abnormally thin skin that result in open lesions, as well as abnormal flexibility of digital joints at the time of birth. However, the role of the WFFS allele in embryonic loss is yet unknown.

Screening of multiple horse breeds during the commercialization of this test at the Veterinary Genetics Laboratory, University of California, Davis, USA (VGL) identified the presence of the WFFS allele in a broad range of warmblood breeds from different countries and also in the Thoroughbred. The presence of WFFS carriers was further investigated in the Thoroughbred and was determined to be at low frequency (2.4% in the 716 horses evaluated). Further, the WFFS allele was not associated with catastrophic breakdown [30]. Most recently, the WFFS allele frequency was estimated to be 14% in Hanoverians [29]. Pedigree analyses of WFFS carriers in warmbloods suggested that the variant could trace to the English Thoroughbred stallion Dark Ronald, a formative ancestor in many German warmblood lines [31]. However, DNA isolated from the preserved skin of Dark Ronald showed that he was homozygous for the wild-type PLOD1 variant, thus rejecting this hypothesis [32]. To date, the exact origin of WFFS is unclear. The famous imported Arabian stallion Bairactar Or. Ar. (1813) has also been hypothesized to be the founder of the WFFS allele, as offspring that resulted from inbreeding to him were reported to have a disease, similar to WFFS [9].

The stallion Bairactar Or. Ar. was imported from the Middle East in 1817, to the German stud in Weil. Until the first half of the 20th century the sire line was maintained mainly at Weil, but recently Bairactar Or. Ar.'s descendants are found in many warmblood breeds and bloodlines. Among these is the stallion Amurath Sahib who was the most influential Arabian sire line used in Poland. To investigate whether Bairactar Or. Ar. was a potential founder of WFFS, we isolated DNA from his museum remains. We also isolated DNA from other Arabian horses to test for the presence of the WFFS allele among Arabians from diverse lines including those that trace to Amurath Sahib.

In this study, carrier frequency and distribution of the WFFS allele were examined in a diverse sample set of over four thousand individuals representing 38 different breeds, including several warmblood breeds. Warmbloods are a group of horses with medium size body types representing many breeds that primarily originated in Europe and that have known draft ("coldblood"), Thoroughbred, and Arabian ("hotblood") ancestry. They are registered with various breed registries that are characterized by open studbook policy, studbook selection, and are often bred for competitive equestrian sports. In addition to the warmbloods, Thoroughbreds, coldbloods, and ponies were also investigated. Horses were selected to cover a broad range of breeds and samples from both Europe and the United States to assess allele frequencies and distribution of the WFFS allele among breeds. Identification and

estimation of breed-specific allele frequencies can guide breed genetic testing recommendations related to WFFS.

2. Materials and Methods

2.1. Sample Description Population Screening

Hair and blood samples of 4081 horses from 38 different breeds were analyzed in this study. The collection includes Thoroughbreds, warmblood horse breeds, Quarter Horses, breeds derived from warmbloods, and breeds in which warmbloods are allowable crosses. The sample set was also supplemented by a large number of Arabians (302), a selection of baroque horses, such as Friesians and Lippizans, and draft and pony breeds (Table 1).

A total of 2127 DNA samples representing 32 different breeds, banked at the UC Davis Veterinary Genetics Laboratory (VGL), were available for investigation. Samples were filtered for relatedness within breeds, based on one degree of separation. A total of 1238 horses were previously sampled for other research projects and stored in the Department of Animal Reproduction, Anatomy and Genomics, under the University of Agriculture in Krakow requirements (Ethical Agreement no. 00665 and 1173/2015). The sampled Arabians include representatives of descendants of Bairactar Or. Ar. (16 mares and 3 stallions) used in the Polish breeding program. The experimental protocol for Arabians was approved by the Animal Care and Use Committee of the Institute of Pharmacology, Polish Academy of Sciences in Krakow (no. 1173/2015).

The 716 Thoroughbred horses included in this study were used in Bellone et al., 2019 [30] and were included here for completeness of the dataset.

2.2. DNA Isolation and Genotyping of WFFS

DNA isolation and genotyping were performed with two different methods. In method A, DNA isolation from hair follicles was done as previously described in Locke et al. 2002 [33]. DNA was amplified for the PLOD1 c.2032G>A variant using the commercially available assay at the UC Davis VGL (https://vgl.ucdavis.edu/test/wffs). To ensure accurate genotyping, assays were run with three positive controls (one for each genotype) and one negative control. Positive controls genotyped as expected and negative controls did not yield detectable PCR product.

In method B, DNA was isolated from whole blood or hair follicles using Sherlock AX (A&A Biotechnology) according to the manufacturer's protocol. The PLOD1 c.2032G>A variant was genotyped using a PCR-RFLP method. The primers for the PLOD1 gene were designed based on ENSECAG00000022842 (fwd:5′-CTCGTGGTAGTGCGTGAGTC-3′ and rev: 5′-AGGGCCCAG CTTCCTCTT-3′) reference using PrimerInput3 (0.4.0). The endonuclease FauI, which only cuts the unmutated G-Allele but not the mutated A-Allele, was selected using NebCutter V2.0. PCR was performed using AmpliTaq Gold™ 360 Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) according to the protocol with an annealing temperature of 57 °C. PCR fragments were digested with FauI (New England Biolabs, Ipswich, MA, USA) according to the protocol and visualized on a 4% agarose gel (G-Allele 102 and 64 bp; A-Allele 166 bp). Amplicons from all heterozygous and some homozygous samples were confirmed by Sanger sequencing using BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) and 3500×L Genetic Analyzers (Applied Biosystems, Foster, CA, USA; Thermo Fisher Scientific).

Allele and carrier frequencies were calculated for each breed and over all horses using Excel (Microsoft Office). Additionally, 95% confidence intervals were calculated for those breeds with more than 30 samples and for the overall sample set using the modified Wald Method in Graph Pad (https://www.graphpad.com/quickcalcs/confInterval2/) [34].

2.3. Historic DNA—Bairactar Or. Ar.

A molar tooth of the stallion Bairactar Or. Ar. (1813–1838) was taken from his skeleton at the Stud Museum Offenhausen in Germany (Figure 1). To avoid contamination, DNA extraction was performed in the clean room of the Central Research Laboratories at the Natural History Museum in Vienna.

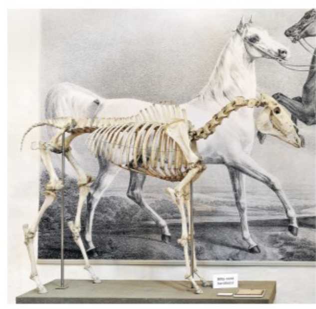

Figure 1. Skeleton of the Arabian stallion Bairactar Or. Ar. (1813–1838) in the Stud Museum Offenhausen in Germany (copyright: Stephan Kube). A tooth from this skeleton was used for DNA analysis.

DNA was extracted from 1000 mg tooth powder using the GEN-IAL® All-tissue DNA-Kit adjusting the manufacturer's instructions (find details in File S1). Finally, DNA was dissolved in 30 μL nuclease-free water (Invitrogen™, Carlsbad, CA, USA). A final DNA concentration of 2.5 ng/μL was determined with Qubit™ dsDNA HS Assay Kit. No-template controls were included for the whole process of DNA isolation.

A 235 bp PCR product flanking the WFFS allele was amplified in 4 independent 50 μL reactions containing 1.75 μL DNA extract, 1 μM each of the primers PLOD1_B_fwd: 5′-GTCAC TCCACAAGGCACAAG-3′ and PLOD_1_B_rev: 5′-GTGGTAGTGCGTGAGTCGTC-3′, 0.25 mM of each dNTP, 2 mM MgCl2, 1× AmpliTaq Gold® 360 Buffer, and 1.25 U AmpliTaq Gold® 360 DNA Polymerase (Applied Biosystems™). PCR conditions were 5 min at 95 °C, followed by 40 cycles of 30 s at 95 °C, 30 s at 59 °C, 30 s at 72 °C, and a final extension of 7 min at 72 °C. PCR products were checked on 2% agarose gels, no product was visible.

Subsequently, 8 nested PCRs using 0.3 μM each of the primers PLOD1_C_fwd: 5′-AAACT GACGCTTCCTGTTGG-3′ and PLOD1_B_rev: 5′-GTGGTAGTGCGTGAGTCGTC-3′, resulting in a 143 bp product were performed in 25 μL reactions. Nested PCR reactions contained 3 μL PCR product of the first round as template, 0.25 mM of each dNTP, 1.5 mM MgCl2, 1× Buffer, and 0.5 U Taq Polymerase (Biozym, Hessisch Oldendorf, Germany). PCR conditions were 5 min at 95 °C, followed by 35 cycles of 30 s at 95 °C, 30 s at 59 °C, 30 s at 72 °C, and a final extension of 7 min at 72 °C. PCR products were visualized on 2% ethidium bromide-stained agarose gels. Negative controls, including the no-template controls from the first PCR rounds as input, were included.

As positive controls, two PCRs of known heterozygous horses using the primers PLOD1_B_fwd and PLOD_1_B_rev were done on a separate day to avoid contamination, following the nested PCR protocol described above.

DNA bands of the nested PCRs and the amplicons of the two controls were purified with Qiagen Gel-extraction-kit on different days and sent for Sanger sequencing at LGC genomics in Berlin (Germany).

All 20 sequencing traces (8 products, fwd and rev each from Bairactar Or. Ar.; 2 products, fwd and rev each from heterozygous controls) were analyzed using the software CodonCode Aligner 3.0.1.

3. Results

3.1. Distribution of the WFFS Allele

This study includes WFFS genotyping results of 4081 horses representing 38 different breeds, as summarized in Table 1. In total, 200 horses (4.9%) in our dataset carried the WFFS allele. The WFFS

allele was detected in 21 breeds and carrier frequency ranged from 0% in most non-warmblood horse breeds (with the exception of Thoroughbred and Haflinger) to over 17% in the Hanoverian and Danish Warmblood. Among warmbloods, the WFFS allele was present in 17 of 19 breeds studied. It was not detected in the Swedish Warmblood (n = 16) and the Zangersheide Warmblood (n = 10), but the sample numbers of these breeds were small. The average carrier frequency across the 1610 warmblood horses tested in this study was 11%.

In addition to warmbloods, the WFFS allele was identified in the American Sport Pony, one of twelve horses examined was a carrier, and in the Knabstrupper with a carrier frequency of 6.5%. Two of 48 Haflingers were also identified as carriers (carrier frequency 4.2%). In reviewing pedigrees, a common ancestor among these two carriers was not identified in as far back as six generations.

The occurrence of WFFS carriers in Thoroughbreds has been shown previously to be 2.4% in Bellone et al., 2019 [30], but all 146 Thoroughbreds from Poland genotyped in this study were homozygous for the reference allele (PLOD1:c.2032G).

The WFFS allele was not detected in pony breeds such as the Shetland Pony, Hucul, Polish Konik, and the Norwegian Fjord. It was also absent in the tested draft breeds (Polish Heavy Draft and Shire), Quarter Horse, Appaloosa, and Baroque-type breeds like the Friesian, Friesian crosses, and Lipizzan. Akhal Teke, Arabian, and the tested gaited breeds (Tennessee Walker and Rocky Mountain Horse) also exclusively showed the reference allele.

Table 1. Warmblood fragile foal syndrome (WFFS) genotyping summary for all investigated breeds with the calculated carrier frequency, allele frequency, and 95% confidence interval for allele frequency. A short breed information is given for all breeds where carriers were detected.

Breed	Total	Carriers	Carrier Frequency %	Allele Frequency %	95% CI of WFFS Allele Frequency	Breed Information
Akhal Teke	35	0	0	0		
American Sport Pony	*12*	*1*	*8.33*	*4.17*		Warmblood-derived, studbook not closed
American Warmblood	57	8	14.04	7.02	3.41 to 13.43	Warmblood
Appaloosa	43	0	0	0		
Arabian	302	0	0	0		
Baden-Württemberger	*3*	*1*	*33.33*	*16.67*		Warmblood
Belgian Sport Horse	*10*	*1*	*10.00*	*5.00*		Warmblood
Belgian Warmblood	44	5	11.36	5.68	2.14 to 12.93	Warmblood
Canadian Warmblood	*29*	*3*	*10.34*	*5.17*		Warmblood
Danish Warmblood	127	22	17.32	8.66	5.74 to 12.82	Warmblood
Dutch Warmblood	249	19	7.63	3.82	2.42 to 5.92	Warmblood
Friesian	197	0	0	0		
Friesian Cross	72	0	0	0		
Haflinger	48	2	4.17	2.08	0.12 to 7.74	Origin in Europe with influence from several breeds including Arabians, studbook not closed until 1946
Hanoverian	283	49	17.31	9.01	6.90 to 11.67	Warmblood
Polish Heavy Draft	209	0	0	0		
Hessen	*2*	*1*	*50.00*	*25.00*		Warmblood
Holsteiner	132	11	8.33	4.17	2.26 to 7.39	Warmblood
Hucul	146	0	0	0		
Knabstrupper	46	3	6.52	3.26	0.72 to 9.55	Baroque type; Warmblood crosses allowed, studbook not closed

Table 1. *Cont.*

Breed	Total	Carriers	Carrier Frequency %	Allele Frequency %	95% CI of WFFS Allele Frequency	Breed Information
Lesser Poland Warmblood	157	3	1.91	0.96	0.19 to 2.91	Warmblood
Lippizan	42	0	0	0		
Norwegian Fjord	42	0	0	0		
Oldenburg	219	34	15.53	7.76	5.58 to 10.68	Warmblood
Polish Konik	96	0	0	0		
Quarter Horse	112	0	0	0		
Rheinland	12	2	16.67	8.33		Warmblood
Rocky Mountain Horse	89	0	0	0		
Selle Français	52	3	5.77	2.88	0.62 to 8.50	Warmblood
Shetland Pony	40	0	0	0		
Shire	39	0	0	0		
Silesian Horse	96	12	12.50	6.25	3.51 to 10.71	Warmblood
Swedish Warmblood	*16*	*0*	*0*	*0*		
Tennessee Walker	39	0	0	0		
Thoroughbred *	146	0	0	0		
Thoroughbred **	716	17	2.37	1.19	0.73 to 1.91	Arabian, Barb, and Turkoman ancestry
Trakehner	64	1	1.56	0.78	0.01 to 4.73	Warmblood
Westfalen	47	2	4.26	2.13	0.12 to 7.89	Warmblood
Zangersheide	*11*	*0*	*0*	*0*		
total	**4081**	**200**	**4.90**	**2.47**	**2.16 to 2.84**	

* Thoroughbreds sampled in Poland and firstly published in this article. ** Thoroughbreds already published in Bellone et al., 2019 [30]. Data in italic font denote breeds with small sample sizes ($n < 30$). Values for the total data set are given in bold.

3.2. Historic DNA–Bairactar Or. Ar.

The WFFS allele was not detected in any of the 16 sequences (8 fwd and 8 rev) generated out of DNA extracted from a tooth of Bairactar Or. Ar. (Figure 2). For two control horses, we confirmed that they were heterozygous carriers. Based on these results, we assume that the DNA extracted from the tooth is homozygous for the G-Allele (reference allele) at position PLOD1 c.2032. All 20 sequences are shown in File S1.

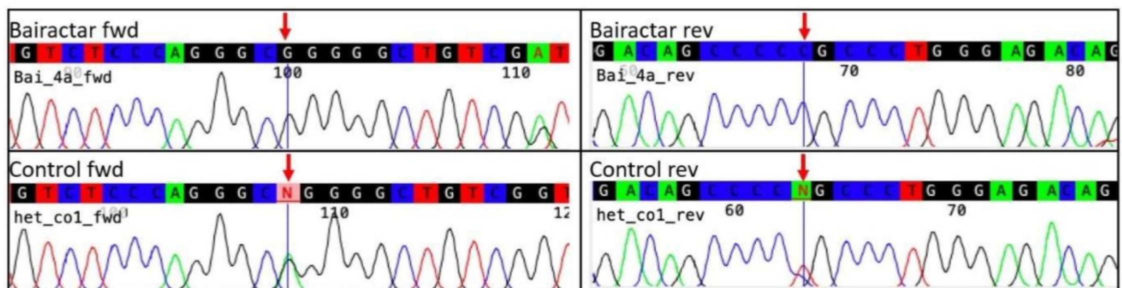

Figure 2. Sanger sequencing traces of Bairactar Or. Ar. (**upper panels**) and one heterozygous control horse (**lower panels**). The electropherograms correspond to sequences obtained with forward (**left side**) and reverse (**right side**) primers. The position of the WFFS single nucleotide variant (SNV) (procollagen-lysine-2-oxoglutarate-5-dioxygenase 1 gene (PLOD1) c.2032G>A) is marked with a red arrow.

4. Discussion

The objective of this study was to determine the presence and frequency of the WFFS allele in diverse horse breeds. Based on our large dataset, which consisted of 4081 horses from 38 different breeds, we confirmed that, with exception of the Thoroughbred and the Haflinger, WFFS mainly occurs in warmblood breeds or those deriving from or allowing outcrosses to warmbloods.

It has been reported that the variant within PLOD1 that converts the amino acid in position 678 in the protein chain from Gly to Arg was carried by 11% of a small warmblood test population. Further pedigree analyses showed that WFFS might segregate among Hanoverian, Selle Français, Dutch Warmblood, Oldenburg, and Westphalian breeds [27]. The same carrier frequency of 11% was initially determined for 374 warmblood samples from Brazil in Dias et al., 2019 [35] and further confirmed in the present study for 1610 warmbloods from Europe and the United States. We detected the WFFS allele in 17 out of 19 tested warmblood breeds, with the highest estimated carrier frequency of 17.3% in the Hanoverian (283 horses) and the Danish Warmblood (127 horses). In a study of 849 Hanoverians, Metzger et al. 2020 [29] reported an allele frequency of 14%, making the expected carrier frequency 24%, which is notably higher than any of the other breeds investigated here. Pedigree analyses of Hanoverian WFFS carriers identified one stallion of the traditional sire line F/W as a most recent common ancestor. This stallion, born in 1861, maybe one of the major contributors to the spread of the WFFS allele in warmblood populations [29]. In our study, similarly high numbers of carriers as in the Hanoverian were detected in the Oldenburger (15.5%, 219 horses tested) and the American Warmblood (14.0%, 57 horses tested). The Silesian horse, a heavy Polish Warmblood breed developed from old Oldenburger bloodlines and recently influenced by the Thoroughbred, showed a slightly lower carrier frequency (12.5%, 96 horses tested). It should be noted that the sample size of the two warmblood breeds with no carriers detected was low (Swedish Warmblood $n = 16$, Zangersheide Warmblood $n = 10$) and, therefore, additional horses from these breeds should be evaluated before any conclusions on the presence of the WFFS allele can be drawn.

In addition to the traditional warmblood breeds, the WFFS allele was also detected in one American Sport Pony and three Knabstruppers. Given the warmblood influence in these breeds, this finding was not surprising. Specifically, crossing to warmblood horses is allowable for Knabstrupper. The American Sport Pony was recently derived in the United States from several breeds including warmblood horses. We also identified two Haflingers with a single copy of the WFFS allele, which was unexpected. Inspection of the pedigree records did not identify a common ancestor of these two individuals. However, one of the carriers had six unknown ancestors listed in the six-generation pedigree. The Haflinger breed was developed in Austria and Northern Italy with the foundation sire foaled in 1874, an offspring of a half Arabian stallion. The stud book was closed in 1946, whereby it is possible that prior to this, an infusion of European warmblood or Thoroughbred bloodlines could account for the WFFS allele in the Haflinger breed. Yet, so far the source of the WFFS allele in the Haflinger remains unknown, and further evaluation of the genetic contribution by other breeds to the Haflinger remains to be investigated.

Thoroughbred was the only non-warmblood breed in which the WFFS allele was previously described. In Bellone et al. 2019 [30] a carrier frequency of 2.4% in 716 Thoroughbreds was reported and it was shown that there was no association between the WFFS allele and catastrophic breakdown in this breed. In contrast to this result, we could not identify any carriers among the 146 Thoroughbreds from Poland genotyped for this study. This is likely best explained by the relatively low frequency overall in the Thoroughbred and the smaller sample size utilized for the Polish samples here compared to the previous work. Additionally, regionally restricted selection of samples may also explain why no carriers were detected in the Polish Thoroughbred sample set, as these may represent distinct subpopulations. Nevertheless, there is evidence for the origin of WFFS in the Thoroughbred because of the occurrence of the WFFS allele in this breed and the large influence of Thoroughbreds in warmblood breeds where the allele mainly occurs. However, the suggestion of Wobbe et al., 2019 [31] that Bay Ronald XX or his son Dark Ronald XX represent the common ancestor of all carriers could not be

confirmed in Metzger et al., 2019 [29]. Indeed, Dark Ronald XX has been excluded as the founder in Zhang et al., 2020 [32] through DNA testing of skin remains.

We investigated the possibility that the Arabian breed was the origin of WFFS, which is based on writing from 1855 where symptoms similar to WFFS were described in foals inbred to Bairactar Or. Ar. [9]. Given that both the Thoroughbred and the Haflinger, two breeds with identified WFFS carriers, were influenced by Arabian horses, we sought to investigate this breed as the origin of the WFFS allele. Through DNA testing of historic remains, we found no evidence that the influential Arabian stallion Bairactar Or. Ar. (1813) was a carrier of the WFFS allele. Therefore, the hypothesis that Bairactar Or. Ar. spread the WFFS allele in Europe can be rejected. Nevertheless, it is still possible that a de novo germline mutation happened in Bairactar Or. Ar.'s son Amurath (*1829), from whom no remains were available for genotyping in this study. Further support of a non-Arabian origin of WFFS are the genotyping results of this study, wherein no carriers were identified among 302 Arabians tested. Thus, the likelihood of an Arabian origin of the WFFS allele is low, although given the presence of the WFFS allele in Haflinger, thought to have descended from a half Arabian stallion, this hypothesis cannot be definitively ruled out at this time.

The distribution of the WFFS allele appears to be restricted to warmblood horse breeds, the Thoroughbred, and a few other breeds that are likely derived from these. We did not detect carriers in Quarter Horses, Appaloosa, and gaited horses like the Tennessee Walker, which are breeds with a known Thoroughbred influence, suggesting these breeds are likely clear of WFFS caused by the PLOD 1 variant. In a broad range of pony and coldblooded breeds included in this study, the WFFS allele was not present in any of them, which might reflect their different origins. For example, Huculs and Polish Koniks represent primitive breeds described as feral. While Koniks are considered as Tarpan relics [36], Huculs originate from Carpathian Mountains horses [37]. The Polish Heavy Drafts arose after the Second World War and are based on local mares of coldblooded type crossed with imported breeding stock, mainly Ardennais and Bretons [38]. Given the results of this study, we propose that it is unlikely that Shetland Ponies, Polish Koniks, Norwegian Fjords and Huculs, Polish Heavy Drafts, and Shire horses will have WFFS caused by the PLOD1 variant. Nevertheless, in some instances, a limited number of samples were available. Six breeds had fewer than 30 samples available for testing (Table 1) and thus in these cases, further investigation is warranted to get a more accurate estimate of allele frequency [39], especially in instances where WFFS is suspected in an affected foal.

Our data further support the use of genetic testing for WFFS in warmblood breeds to prevent the production of affected foals. DNA test results can be utilized to avoid mating between carriers while still retaining carriers in the breeding program to maximize diversity and the production of other desirable traits. Given the presence of the WFFS allele in the Knabstrupper and American Sport Pony, it is also advisable to test breeding animals from warmblood-derived breeds or from breeds where crossing to warmbloods is allowed. DNA screening should even be performed in those breeds that may have genetic influence from warmblood breeds, as exemplified by the occurrence of the WFFS allele in the Haflinger breed. Screening for genetic diseases such as WFFS is not only of significant economic importance but it is also essential to avoid animal suffering and thus proper use of genetic testing in marker-assisted selection to prevent disease traits in horses cannot be over- emphasized.

5. Conclusions

Our comprehensive dataset reflects that WFFS is a concern in warmbloods and even occurs in breeds where it was not expected, like the Haflinger. While the origin of WFFS remains unclear, it is recommended to perform genetic tests even in non-warmblood related breeds, at least until there is more clarity about the occurrence of the WFFS allele.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/11/12/1518/s1, File S1: Detailed description of historic DNA analysis from Bairactar Or. Ar.

Author Contributions: Conceptualization of the project was done by M.C.T.P., R.R.B., B.W., and M.S.-S. Funding acquisition was performed by B.W., M.C.T.P., G.B., and R.R.B. Methodology was determined by R.R.B., S.R., E.H., K.P., K.R.-M., B.D., and M.S.-S. Data collection was performed by M.C.T.P., J.M., M.S.-S., and B.W. Resources were provided by L.H. and G.B. Data analysis was performed by R.R.B., M.C.T.P., B.W., S.R., K.P., K.R.-M., and M.S.-S. Supervision of the project was performed by B.W. and R.R.B. Manuscript was prepared by S.R., B.W., and R.R.B., with all authors contributed to the final editing and review of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Innovation Fund of the Austrian Academy of Sciences (ÖAW), grant number IF_2015_17 and was supported by the UC Davis Veterinary Genetics Laboratory.

Acknowledgments: The authors would like to thank Minh Le, Natalia Ocampo, and Barbara Tautscher for their technical assistance. The authors also gratefully acknowledge the horse owners who provided samples that made this study possible.

Conflicts of Interest: J.M., M.C.T.P., and R.R.B. are affiliated with the UC Davis Veterinary Genetics Laboratory, a genetic testing laboratory offering genetic diagnostic testing in horses and other species.

References

1. Ehlers-Danlos Syndrome. Available online: https://ghr.nlm.nih.gov/condition/ehlers-danlos-syndrome (accessed on 17 September 2020).
2. Malfait, F.; De Paepe, A. The Ehlers-Danlos syndrome. *Adv. Exp. Med. Biol.* **2014**, *802*, 129–143. [CrossRef]
3. Halper, J. Connective tissue disorders in domestic animals. *Adv. Exp. Med. Biol.* **2014**, *802*, 231–240. [CrossRef]
4. Eßer, M.; Niederacher, V.; Pfeffer, K.; Scheuer, H. Über die selten auftretende Dermatosparaxie (Ehlers-Danlos-Syndrom) bei einem Fohlen—Ein Fallbericht. *Pferdeheilkunde* **1999**, *15*, 434–436. [CrossRef]
5. Witzig, P.; Suter, M.; Wild, P.; Rao, V.H.; Steinmann, B.; von Rotz, A. Dermatosparaxis in a foal and a cow—A rare disease? *Schweiz. Arch. Tierheilkd.* **1984**, *126*, 589–596.
6. Marshall, V.L.; Secombe, C.; Nicholls, P.K. Cutaneous asthenia in a Warmblood foal. *Aust. Vet. J.* **2011**, *89*, 77–81. [CrossRef]
7. Rüfenacht, S.; Straub, R.; Steinmann, B.; Winand, N.; Bidaut, A.; Stoffel, M.H.; Gerber, V.; Wyder, M.; Müller, E.; Roosje, P.J. Swiss warmblood horse with symptoms of hereditary equine regional dermal asthenia without mutation in the cyclophylin B gene (PPIB). *Schweiz. Arch. Tierheilkd.* **2010**, *152*, 188–192. [CrossRef]
8. Winter, T.; Börgel, C.; Aupperle, H.; Schoon, H. Eine dem Ehlers-Danlos Syndrom des Menschen ähnliche Bindegewebserkrankung der Haut bei einem Fohlen. *Pferdeheilkunde* **2004**, *20*, 19–22. [CrossRef]
9. Rueff, A.G. Eigenthümliche Krankheit neugeborener Fohlen (a peculiar disease of newborn foals). In *Jahrbuch für Pferdezucht*; Verlag Voigt: Weimar, Germany, 1855; pp. 267–277.
10. Gunson, D.E.; Halliwell, R.E.; Minor, R.R. Dermal collagen degradation and phagocytosis. Occurrence in a horse with hyperextensible fragile skin. *Arch. Dermatol.* **1984**, *120*, 599–604. [CrossRef]
11. Rashmir-Raven, A. Heritable equine regional dermal asthenia. *Vet. Clin. N. Am. Equine Pract.* **2013**, *29*, 689–702. [CrossRef]
12. White, S.D.; Affolter, V.K.; Bannasch, D.L.; Schultheiss, P.C.; Hamar, D.W.; Chapman, P.L.; Naydan, D.; Spier, S.J.; Rosychuk, R.A.W.; Rees, C.; et al. Hereditary equine regional dermal asthenia ("hyperelastosis cutis") in 50 horses: Clinical, histological, immunohistological and ultrastructural findings. *Vet. Dermatol.* **2004**, *15*, 207–217. [CrossRef]
13. Steelman, S.M.; Jackson, N.D.; Conant, E.; Juras, R.; Cothran, E.G.; Edwards, J.F.; Chowdhary, B.P. Ehlers-danlos syndrome in a quarter horse gelding: A case report of PPIB-independent hereditary equine regional dermal asthenia. *J. Equine Vet. Sci.* **2014**, *34*, 565–568. [CrossRef]
14. Borges, A.S.; Conceição, L.G.; Alves, A.L.G.; Fabris, V.E.; Pessoa, M.A. Hereditary equine regional dermal asthenia in three related Quarter horses in Brazil. *Vet. Dermatol.* **2005**, *16*, 125–130. [CrossRef] [PubMed]
15. Lerner, D.; McCracken, M. Hyperelastosis cutis in 2 Horses. *J. Equine Med. Surg.* **1978**, *2*, 350–352.
16. White, S.D.; Affolter, V.K.; Schultheiss, P.C.; Ball, B.A.; Wessel, M.T.; Kass, P.; Molinaro, A.M.; Bannasch, D.L.; Ihrke, P.J. Clinical and pathological findings in a HERDA-affected foal for 1.5 years of life. *Vet. Dermatol.* **2007**, *18*, 36–40. [CrossRef] [PubMed]
17. Hardy, M.H.; Fisher, K.R.; Vrablic, O.E.; Yager, J.A.; Nimmo-Wilkie, J.S.; Parker, W.; Keeley, F.W. An inherited connective tissue disease in the horse. *Lab. Investig.* **1988**, *59*, 253–262. [PubMed]

18. Brounts, S.H.; Rashmir-Raven, A.M.; Black, S.S. Zonal dermal separation: A distinctive histopathological lesion associated with hyperelastosis cutis in a Quarter Horse. *Vet. Dermatol.* **2001**, *12*, 219–224. [CrossRef] [PubMed]
19. Bowser, J.E.; Elder, S.H.; Pasquali, M.; Grady, J.G.; Rashmir-Raven, A.M.; Wills, R.; Swiderski, C.E. Tensile properties in collagen-rich tissues of Quarter Horses with hereditary equine regional dermal asthenia (HERDA). *Equine Vet. J.* **2014**, *46*, 216–222. [CrossRef]
20. Grady, J.G.; Elder, S.H.; Ryan, P.L.; Swiderski, C.E.; Rashmir-Raven, A.M. Biomechanical and molecular characteristics of hereditary equine regional dermal asthenia in Quarter Horses. *Vet. Dermatol.* **2009**, *20*, 591–599. [CrossRef]
21. Mochal, C.A.; Miller, W.W.; Cooley, A.J.; Linford, R.L.; Ryan, P.L.; Rashmir-Raven, A.M. Ocular findings in Quarter Horses with hereditary equine regional dermal asthenia. *J. Am. Vet. Med. Assoc.* **2010**, *237*, 304–310. [CrossRef]
22. Tryon, R.C.; Penedo, M.C.T.; McCue, M.E.; Valberg, S.J.; Mickelson, J.R.; Famula, T.R.; Wagner, M.L.; Jackson, M.; Hamilton, M.J.; Nooteboom, S.; et al. Evaluation of allele frequencies of inherited disease genes in subgroups of American Quarter Horses. *J. Am. Vet. Med. Assoc.* **2009**, *234*, 120–125. [CrossRef]
23. Finno, C.J.; Spier, S.J.; Valberg, S.J. Equine diseases caused by known genetic mutations. *Vet. J.* **2009**, *179*, 336–347. [CrossRef] [PubMed]
24. Solomons, B. Equine cutis hyperelastica. *Equine Vet. J.* **1984**, *16*, 541–542. [CrossRef] [PubMed]
25. Tryon, R.C.; White, S.D.; Bannasch, D.L. Homozygosity mapping approach identifies a missense mutation in equine cyclophilin B (PPIB) associated with HERDA in the American Quarter Horse. *Genomics* **2007**, *90*, 93–102. [CrossRef] [PubMed]
26. Monthoux, C.; de Brot, S.; Jackson, M.; Bleul, U.; Walter, J. Skin malformations in a neonatal foal tested homozygous positive for Warmblood Fragile Foal Syndrome. *BMC Vet. Res.* **2015**, *11*, 1–8. [CrossRef] [PubMed]
27. Winand, N.J. Identification of the Causative Mutation for Inherited Connective Tissue Disorders in Equines, United States Department of Commerce Application Number: 61/486,464; (Filing Date: May 16th, 2011). Available online: https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2012158711&recNum=1&maxRec=1&office=&prevFilter=&sortOption=Pub+Date+Desc&queryString=FP%3A%28WO2012158711%29&tab=PCT+Biblio (accessed on 20 September 2020).
28. Aurich, C.; Müller-Herbst, S.; Reineking, W.; Müller, E.; Wohlsein, P.; Gunreben, B.; Aurich, J. Characterization of abortion, stillbirth and non-viable foals homozygous for the Warmblood Fragile Foal Syndrome. *Anim. Reprod. Sci.* **2019**, *211*, 106202. [CrossRef] [PubMed]
29. Metzger, J.; Kreft, O.; Sieme, H.; Martinsson, G.; Reineking, W.; Hewicker-Trautwein, M.; Distl, O. Hanoverian F/W-line contributes to segregation of Warmblood fragile foal syndrome type 1 variant PLOD1:c.2032G>A in Warmblood horses. *Equine Vet. J.* **2020**, *22*, 1–9. [CrossRef]
30. Bellone, R.R.; Ocampo, N.R.; Hughes, S.S.; Le, V.; Arthur, R.; Finno, C.J.; Penedo, M.C.T. Warmblood fragile foal syndrome type 1 mutation (PLOD1 c.2032G>A) is not associated with catastrophic breakdown and has a low allele frequency in the Thoroughbred breed. *Equine Vet. J.* **2020**, *52*, 411–414. [CrossRef]
31. Wege zum Angemessenen Umgang mit WFFS und Anderen Genetischen Eigenschaften. Available online: https://www.vit.de/aktuelles/20191/wege-zum-angemessenen-umgang-mit-wffs-und-anderen-genetischen-eigenschaften/ (accessed on 17 December 2020).
32. Zhang, X.; Hirschfeld, M.; Schafberg, R.; Swalve, H.; Brenig, B. Skin exhibits of Dark Ronald XX are homozygous wild type at the Warmblood fragile foal syndrome causative missense variant position in lysyl hydroxylase gene PLOD1. *Anim. Genet.* **2020**, 1–2. [CrossRef]
33. Locke, M.M.; Penedo, M.C.T.; Bricker, S.J.; Millon, L.V.; Murray, J.D. Linkage of the grey coat colour locus to microsatellites on horse chromosome 25. *Anim. Genet.* **2002**, *33*, 329–337. [CrossRef]
34. Agrest, A.; Coull, B.S. Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions. *Am. Stat.* **1998**, *52*, 119–126. [CrossRef]
35. Dias, N.M.; de Andrade, D.G.A.; Teixeira-Neto, A.R.; Trinque, C.M.; de Oliveira-Filho, J.P.; Winand, N.J.; Araújo, J.P.; Borges, A.S. Warmblood Fragile Foal Syndrome causative single nucleotide polymorphism frequency in Warmblood horses in Brazil. *Vet. J.* **2019**, *248*, 101–102. [CrossRef]
36. Sandoval-Castellanos, E.; Wutke, S.; Gonzalez-Salazar, C.; Ludwig, A. Coat colour adaptation of post-glacial horses to increasing forest vegetation. *Nat. Ecol. Evol.* **2017**, *1*, 1816–1819. [CrossRef]

37. Georgescu, S.; Manea, M.; Dudu, A.; Costache, M.; Mol, G. Phylogenetic relationships of the Hucul horse from Romania inferred from mitochondrial D-loop variation. *Genet. Mol. Res.* **2011**, *10*, 4104–4113. [CrossRef]
38. Polak, G. Genetic variability of cold-blooded horses participating in genetic resources conservation programs, using pedigree analysis. *Ann. Anim. Sci.* **2019**, *19*, 49–60. [CrossRef]
39. Fung, T.; Kennan, K. General case of sampling form a finite diploid population of any size. *PLoS ONE* **2014**, *9*, e85925. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

V. DISCUSSION

Genome research has continued to develop since the molecular basis of a genetic disorder in horses was first determined 30 years ago. Today, many Mendelian traits are well characterized, and genetic testing is commercially available. Some tests are successfully implemented in breeding programs and thus are an important contribution to animal welfare.

This thesis deals with several questions on Mendelian traits in horse breeding, which were devised based on two publications. The findings of each publication will be discussed in detail followed by a more comprehensive general discussion.

Within the first publication an attempt was made to clarify the genetic background of a monogenetic trait, namely roan coat color.

Publication I:

A Genome-Wide Association Analysis in Noriker Horses Identifies a SNP Associated With Roan Coat Color

The coat color phenotype roan segregates in many different breeds around the world and is dominantly inherited. It has long been assumed that roan is lethal in a homozygous state (Hintz and van Vleck, 1979), but this assumption has been refuted multiple times (Sponenberg and Bellone, 2017; Grilz-Seger *et al.*, 2020; Voß *et al.*, 2020). Several studies published on this trait suggest a locus near the *KIT* gene, which was also confirmed as the causative region for roan in Noriker horses within this thesis.

In Grilz-Seger et al. (2020), we investigated the NGS sequence of a homozygous roan Noriker to detect potential causal variants. In close proximity of the associated top-SNP (ECA3:g.79,543.439 A > G), three variants were detected, a 3 bp deletion combined with 17 bp insertion in intron 20 and two deletions in the 3'-flanking region of the *KIT* gene, all in non-coding regions. The defined haplotype with a size of nearly 54.9 kb was completely associated with the roan phenotype in Noriker, Quarter Horses, Slovenian Coldblood, Belgian Drought, and Murgese. However, this haplotype was not detected in roan-colored Shetland Ponies, Trakehner, and German Sport horses (Grilz-Seger *et al.*, 2020). Based on the breeding history of those breeds that carry the 54.9 kb roan haplotype, a common origin can be presumed for this haplotype. This topic is discussed in more detail in the original article (Grilz-Seger *et al.*, 2020).

In Marklund et al. (1999), a similar observation, namely an association of the roan coat color

with *KIT* variants in some but not all breeds, was made. The screening of a TaqI polymorphism in exon 19 of the *KIT* gene in roan and non-roan horses from nine breeds revealed a complete association of this variant in the Belgian Drought, Connemara pony, American Quarter Horse, New Forest pony, Standardbred, and Swedish Halfbred but not in the Welsh, Shetland, and Gotland pony (Marklund et al., 1999). Two variants that were carried in roan individuals investigated in Marklund et al. (1999) were also detected in the whole genome sequenced homozygous roan Noriker (ECA3:79,540,501 G>A; ECA3:79,545,912 G>A). This leads to the assumption, that variants associated with roan in Marklund et al. 1999 correspond to the haplotype defined in the Noriker. A roan haplotype different to the one in the Noriker was recently described in the Icelandic horse. Five intronic and three exonic *KIT* variants were highly associated with roan in this breed (Voß *et al.*, 2020). Interestingly, although they were in close proximity to the variants reported by Marklund et al. (1999) and Grilz-Seger et al. (2020), those variants were not identical (Voß *et al.*, 2020).

The association of the roan coat color with *KIT* variants in some but not all breeds support the theory of allelic heterogeneity in different breeds. Despite multiple studies have been conducted to resolve the genetic basis of the roan coat color, the causal variant is still unknown. As the causal variant is not an exonic variant in the *KIT* gene, further investigations should be focused on intronic and structural variants.

Intronic variants can have an effect on mRNA and accordingly on the resulting protein. Mature mRNA contains only coding sequences, and it is the result of a complex process ('splicing') where intronic sequences are removed from the transcript. Intronic variants can alter the splicing process by disrupting existing splice sites or splicing regulatory sequences (intronic and exonic splicing silencers and enhancers), creating new ones or activating cryptic ones (Abramowicz and Gos, 2018).

About 9% of all variants reported in the Human Gene Mutation Database result in alternative splicing (*Human Gene Mutation Database*, 2022). Due to the challenging proof of their causality the volume and importance of splicing mutations is likely to be underestimated.

In Lim et al. (2011), the finding of a novel alternative splicing event in the porcine *KIT* gene was reported, which is speculated to be causative for the roan phenotype in the investigated pigs (Lim *et al.*, 2011). This study indicates that the intronic *KIT* mutations determined in this doctoral thesis should be further investigated, as they can be causative for the roan coat color in Noriker and other breeds with the related haplotype.

Determining the causality of a mutation in a non-model organism is challenging, especially for

intronic variants. One method for a proof of causality would be through studying the consequence of the variants in a genetically modified organism. Some studies have simulated genetic traits of horses in a mouse model. In Andersson et al. (2012), a mouse model was used to verify the so-called 'gait-keeper' mutation in the *DMRT3* gene of horses. This nonsense mutation is permissive for the ability to perform alternate gaits, such as ambling and passing (Andersson *et al.*, 2012). Another example is given in Liu et al. (2021), where a single SNP in the *TBX3* enhancer region was identified to drive size variation in Chinese horses. Mouse models confirmed that this SNP affects the level of *TBX3* transcription and consequently limb length (Liu *et al.*, 2021).

Large structural variations, which could also be a possible cause for roan, are difficult to detect with short-read NGS data. Long-read sequencing technologies can generate long continuous sequences, ranging from 10 kb to >1 Mb in length, which enable a more accurate detection of structural variants (Logsdon et al., 2020). One structural variant within the *KIT* region has already been described in horses. The tobiano white spotting pattern is caused by a large chromosomal inversion, which starts approximately 100 kb downstream of the *KIT* gene. Although this inversion does not interrupt any annotated genes, sequences regulating the *KIT* gene may be damaged and cause this pattern (Brooks et al. 2007). In addition to tobiano and roan, there are two further depigmentation phenotypes that are caused by variation in the KIT gene, namely dominant white and sabino.

'Special' coat colors in domestic animals are very popular. Like rare species in general, the attraction for special coat colors seems to be part of human nature (Angulo *et al.*, 2009). Thus, the trend of favoring unusual coat colors in horses started with the domestication of the horse around 5000 years ago, where a rapid and substantial increase in the number of coat color variations was recognized. While early wild horses were mainly bay or bay-dun, a variety of colors could be seen around the Bronze age. In addition to bay and black, the genotypes determining chestnut, tobiano, sabino, buckskin and black silver were detected in ancient bone samples (Ludwig *et al.*, 2009).
Today, unusually colored horses are in high demand and often achieve astonishingly high prices. Thus, the interest of the public in genetic testing to determine color characteristics is correspondingly high.

Although the causal variant for roan coat color remains unknown, genetic testing to determine the zygosity in Noriker (and in the other breeds that carry the Haplotype) is feasible through the associated variants. Because only an associated marker, instead of the causative variant can be monitored, it is an indirect genetic test (or marker test). However, there is always the

possibility of a recombination event, that could break up linkage and lead to a loss of the association between the marker and the causal mutation. Therefore, test results should always be viewed with caution and implausible results must be checked carefully. Nevertheless, results of the marker test enable selective mating with a more accurate prediction of coat color. Roan is a dominant inherited trait; thus, when mating a homozygous roan horse to a non-roan, all offspring will be roan.

The second publication of this work, on the other hand, deals with a completely opposite trait, namely a recessive, negatively selected disease in the horse. More specifically, we were interested here in the prevalence of a hereditary defect and the potential of an expanded screening study for horse breeding.

Publication II:

Distribution of the Warmblood Fragile Foal Syndrome Type 1 Mutation (PLOD1.2032G>A) in Different Horse Breeds from Europe and the United States

Warmblood fragile foal Syndrome is a recessive disorder that, as the name suggests, was first detected in warmbloods. Homozygosity for the WFFS allele is lethal, while heterozygous carriers are asymptomatic. The causative mutation for WFFS was first described in 2011(Winand, 2011). Since then, WFFS carrier animals have been detected in various warmblood breeds with an average frequency of 11% (Winand, 2011; Dias *et al.*, 2019). In 2019, WFFS was detected in a breed other than warmbloods. A carrier frequency of 2.4% was observed in the thoroughbred (Bellone et al., 2020). To gain clarity on the appearance and the origin of WFFS, in Reiter et al. (2020), we screened over 4000 horses of 38 different breeds for the WFFS allele (Reiter *et al.*, 2020). Within this study, a broad distribution of WFFS through warmbloods was confirmed. The highest carrier frequency was observed in the Hannoverian and Danish Warmblood with 17%, followed by the Oldenburger with 15%.

Such high frequencies of a deleterious allele would be rather unusual in a population with natural selection, as the allele frequency of a recessive lethal allele is expected to slowly decline over generations. However, within managed populations, where mating is controlled by humans, the opposite–an increase of recessive disorders–can occur. This is often owed to influential breeding animals and is therefore referred to as the ‘popular sire effect.’ ‘Popular sires’ are stallions that are used disproportionately within breeding. Since the use of artificial insemination in horse breeding, the number of offspring per stallion has increased enormously, which has also accelerated the spread of undetected recessive diseases. In Metzger et al. (2019), pedigree analysis revealed a Hanoverian stallion, born in 1861, as the most recent common

ancestor of 76 genotyped WFFS carriers (Metzger *et al.*, 2020). The fact that this stallion is the ancestor of very successful dressage horses supports the theory of a heterozygous advantage of the WFFS allele, which can be another explanation for such high allele frequencies. Foals that are homozygous WFFS have defective connective tissue, which becomes apparent through skin and mucosa lacerations, fragile skin, and hyperextension of the articulations. Possibly, under heterozygous conditions, these symptoms are weakened and may represent a desired phenotype. Accordingly, carrying the WFFS allele could result in a selective advantage. Two independent studies on this theme identified a noticeable association between equine performance and the WFFS genotype, and a favorable effect of the WFFS allele for movement and dressage traits is suggested (Metzger *et al.*, 2020; Ablondi *et al.*, 2021).

Beneath the high frequencies of WFFS in warmbloods, in Reiter et al. (2020), the WFFS allele was detected for the first time in the Haflinger (2/48), American Sport Pony (1/12), and Knabstrupper (3/46). In another recently published work, one of six Haflingers was also identified as a WFFS carrier (Martin *et al.*, 2021). These data provide evidence, that the WFFS allele is more widespread than assumed. Further studies based on an even broader population-based sampling need to be carried out, to obtain a more meaningful allele frequency within the Haflinger, American Sport Pony, and Knabstrupper. Based on the recent findings, which provide evidence that WFFS is not limited to warmbloods, the term WFFS (Warmblood Fragile Foal Syndrome) was questioned, and renaming was considered. Therefore, the disease is termed Fragile Foal Syndrome (FFS), which is probably more appropriate, in the OMIA database.

Many studies have been carried out on WFFS, dealing with the pathology and distribution of the disease. However, the original source of the deleterious allele and its distribution mode is debated and not yet resolved. To determine the foundation individual, pedigree analysis could be an effective tool, as they enable tracing back segregating alleles within a breed. However, for an allele that is as widespread as WFFS, this analysis reached its limits. Furthermore, the length of the haplotype surrounding the mutated position can be investigated; the longer the associated haplotype, the more recent the shared founder. Another possibility is to trace the causative allele in historical or even ancient samples. For example, Bower et al. (2012) investigated the genetic origin and history of speed in the Thoroughbred racehorse by testing 12 historically important Thoroughbred stallions from Museums, for a variant in the myostatin gene (*MSTN*) (Bower *et al.*, 2012). Unfortunately, the identity of the Hanoverian stallion, mentioned in Metzger et al. (2020), has not been published; therefore, it is impossible to assess whether a museum sample of this horse is available. However, the likelihood of a museum specimen being available would be low, as only very view famous horses have been preserved.

Another study, based on Pedigree analysis pointed to the English Thoroughbred stallion Dark Ronald (1905) as the possible origin of WFFS (Wobbe *et al.*, 2019). Historic DNA analysis of Dark Ronalds' preserved skin provides evidence that Dark Ronald was not the founder of the WFFS causative variant (Zhang *et al.*, 2020). Rueff (1855) described symptoms similar to WFFS in horses inbred to the stallion Bairactar Or. Ar. (1813). Hence, Bairactar Or. Ar. was discussed as a possible origin of this disease. We obtained the privilege to isolate DNA from a tooth of the Arabian stallion Bairactar Or. Ar.. Sanger sequencing revealed that he was homozygous for the wild type allele at the SNV causative for WFFS in the *PLOD1* gene. Based on this finding, Bairactar Or. Ar. could be ruled out as a possible origin for the WFFS allele. This work again proves the potential of historical samples to elucidate the origin and dispersal of variants in the horse genome. However, this also underlines the risk that 'popular sires,' such as Bairactar Or. Ar. (Rueff, 1855), Dark Ronald (Zhang *et al.*, 2020), or the Hannoverian stallion, born in 1861 mentioned by Metzger et al., (2020), are automatically and unjustifiably accused of being the founder of some positively or negatively selected traits.

The increasing number of detected causal variants in horses has led to a wide range of available genetic tests. Large-scale screening studies are necessary to determine which alleles segregate in populations and whether they cause diseases. Such information is essential to select genetic tests to implement in breeding programs. In general, genetic testing is handled controversially through different breeding associations. For WFFS, the allelic status (carrier or free) is required for males registered as breeding stallions in most warmblood breeds. Since 2019, this is also the case for the Austrian Warmblood, where only stallions with a known WFFS status are listed in the main stallion book. WFFS carriers are not excluded from breeding, but rather get the note 'WFFS-carrier' in the registry. Consequently, WFFS-carriers are exclusively allowed to be mated with WFFS-free mares, as otherwise the risk of a non-viable foal is 25%. Animal welfare is an important issue in animal breeding. The breeding of animals with predictable pain, suffering, or harm is not only ethically questionable but also prohibited according to §5 of the Federal Act on the Protection of Animals in Austria (*RIS - Tierschutzgesetz*, 2022). Through the controlled mating of carrier animals, the WFFS allele should decrease slowly in the population, and the genetic diversity within a breed can be preserved.

In contrast to classic livestock animals that are selected for specific economic attributes, such as meat, milk or wool, the use of horses has clearly shifted in recent centuries. Horses have been selected for a broad range of characteristics and steadily developed away from livestock toward a companion and show animal. Accordingly, today's horse population is characterized by an extensive variation in color and other traits, leading to a range of phenotypic diversity

that can only be found in the dog (Hedrick and Andersson, 2011). Similar to dogs, some horse breeds are heavily inbred, which at the genome level, results in the accumulation of homozygous regions and a reduced heterogeneity. At the population level, this can lead to an extraordinary increase in allele frequency of some negative and even lethal variants, e.g. 17% of WFFS in Hanoverians (Reiter *et al.*, 2020). By studying overlapping homozygous regions (ROH islands) the impact of selection for different traits can be analyzed. In Grilz-Seger et al. (2019), ROH islands were used to study the impact of color selection in the Noriker breed. Although the Noriker breed is a closed population with a limited number of founders and ancestors, genetic fragmentation based on the color selection was shown (Grilz-Seger, Druml, *et al.*, 2019).

In summary, although there is enormous phenotypic variance in horses, this is not reflected in the genome. Breeding strategies that are applied to achieve rapid breeding success are at the same time responsible for the spread of deleterious alleles and act against genetic diversity (Orlando and Librado, 2019).

By studying Mendelian traits, their impact can be assessed, and their spread monitored. The development of suitable breeding programs can preserve desired phenotypes in a population, while reducing hereditary defects. Preservation of the still existing genetic diversity forms the basis of sustainable horse breeding, which also enables passion for horses in future generations.

VI. SUMMARY

Mendelian traits represent an important basis for genetic selection in modern animal breeding. Lots of effort has been made to resolve the underlying variants of Mendelian traits in horses. In this doctoral thesis, the various aspects of the research on Mendelian traits were processed based on two publications. The first publication deals with the challenges of detecting a causal variant, which are represented in the example of roan coat color. The roan phenotype is characterized by intermixed white and colored hairs in the body and occurs in different pony, warmblood, and coldblood breeds, including the Noriker. Roan is a Mendelian trait and follows a dominant mode of inheritance. To investigate the genetic background of roan 670,000 genome-wide SNP genotypes of 14 roan Noriker (case group) and 26 black Noriker (control group) were analyzed in a genome-wide association study (GWAS). A single nucleotide variant in intron 17 of the *KIT* gene was found to be completely linked with roan. Based on NGS data from a homozygous roan Noriker, a 54.9 kb haplotype, consisting of a single nucleotide variant, two deletions in the 3'-flanking region, and a 3 bp deletion combined with a 17 bp insertion in intron 20 of the *KIT* gene was determined. This haplotype, composed of four non-coding variants, was checked by genotyping a random sample set of roan and non-roan animals of different breeds. Complete association was observed in some breeds (Noriker, Quarter Horses, Slovenian Coldblood, Belgian Draft, Murgese) but not in others (Shetland Pony, Trakehner, German Sport horse). The results suggest allelic heterogeneity for the roan coat color through different breeds. Further work is needed to explicitly clarify the genetic basis of roan, but the associated variants identified within this work can be used for indirect genetic testing.

In the second publication, the importance of genetic testing for animal welfare was illustrated by a large scale screening study of the variant causing the Warmblood Fragile Foal Syndrome (WFFS). Most monogenetic disorders in the horse are recessively inherited, including WFFS. Homozygosity for the WFFS allele is always incompatible with extra-uterine life. The affected foals are aborted during a late stage of gestation or are born with skin abnormalities and extreme flexibility in their digital joints. The WFFS allele occurs at high frequencies in various warmblood breeds and has also been detected in the Thoroughbred. Within this work the distribution of the WFFS allele in the horse population was determined by screening 4081 horses belonging to 38 different breeds. The results confirmed high carrier frequencies in warmbloods, with the highest levels in Hannoverian and Danish Warmblood. For the first time, the WFFS allele was detected in breeds other than Warmbloods, namely in Haflinger, American Sport Pony, and Knabstrupper. Based on the inclusion of historical DNA, it has been shown

that an origin of WFFS in the Arabian breed, specifically in the stallion Bairactar Or. Ar. (1813), is unlikely.

This work shows that the clarification of Mendelian traits and the dense monitoring of the distribution of hereditary defects are of particular importance for applied horse breeding. Genetic testing enables the identification of animals carrying recessive deleterious alleles, which facilitates strategically planned mating and thus improves animal welfare.

VII. Zusammenfassung

Monogene Merkmale bilden eine wichtige Grundlage für die genetische Selektion in der modernen Tierzucht. An der Aufklärung der kausalen Varianten dieser Merkmale wird intensiv geforscht. In dieser Dissertation werden verschiedene Aspekte rund um monogene Merkmale beim Pferd, basierend auf zwei Publikationen, bearbeitet. Anhand der ersten Publikation werden die Herausforderungen der Detektion einer kausalen Variante am Beispiel roan aufgezeigt. Der roan-Phänotyp ist durch eine extreme Stichelhaarigkeit am ganzen Körper charakterisiert und kommt in verschieden Pony-, Warmblut-, und Kaltblutrassen vor, einschließlich dem Noriker. Bei roan handelt es sich um ein dominant vererbtes monogenes Merkmal. Um den genetischen Hintergrund von roan zu untersuchen, wurden 670.000 genomweite SNP Genotypen von 14 roan Norikern (Fallgruppe) und 26 Rappen (Kontrollgruppe) in einer genomweiten Assoziationsstudie (GWAS) analysiert. Dabei wurde ein mit roan assoziierter SNP in Intron 17 vom *KIT* Gen ermittelt. Basierend auf NGS Daten eines homozygoten roan Norikers wurde ein 54,9 kb großer Haplotyp, bestehend aus dem bereits erwähnten SNP, zwei Deletionen in der 3'-flankierenden Region und einer 3 bp Deletion kombiniert mit einer 17 bp Insertion in Intron 20 des *KIT* Gens, definiert. Dieser Haplotyp aus vier nicht-codierenden Varianten wurde durch Genotypisierung in einer zufälligen Stichprobe von roan und nicht-roan Pferden verschiedener Rassen überprüft. Einige Rassen (Noriker, Quarter Horse, slowenisches Kaltblut, belgisches Kaltblut, Murgese) zeigten eine vollständige Assoziation von diesem Haplotyp mit dem roan Phänotyp, andere Rassen (Shetlandpony, Trakehner, deutsches Sportpferd) jedoch nicht. Diese Ergebnisse sprechen für einen heterogenen Ursprung des roan Allels in den verschiedenen Rassen. Zur Bestimmung der kausalen Variante sind weitere Untersuchungen nötig, allerdings können die assoziierten Varianten aus dieser Arbeit für einen indirekten Gentest herangezogen werden.

In der zweiten Publikation dieser Arbeit wird die Bedeutung von Gentests für den Tierschutz anhand einer groß angelegte Screening-Studie der Variante, die das Warmblood Fragile Foal Syndrome (WFFS) verursacht, veranschaulicht. Ein Großteil aller monogenen Erbkrankheiten beim Pferd wird rezessiv vererbt, so auch WFFS. Homozygotie für das WFFS-Allel ist immer letal. Betroffene Fohlen werden gegen Ende der Trächtigkeit abortiert oder mit Hautveränderungen und einer extremen Flexibilität der Gelenke geboren. Das WFFS Allele kommt mit hohen Frequenzen in verschiedenen Warmblütern vor und wurde auch im Vollblut nachgewiesen. In dieser Arbeit wurde die Verbreitung des WFFS Allels in 4081 Pferden von 38 verschiedenen Rassen untersucht. Die Ergebnisse bestätigten hohe Trägerfrequenzen in den

Warmblütern, wobei diese im Hannoveraner und im Dänischen Warmblut am höchsten waren. Des Weiteren wurde das WFFS-Allel erstmals beim Haflinger, Amerikanischen Sportpony und Knabstrupper nachgewiesen. Durch Einbeziehung historischer DNA konnte zudem gezeigt werden, dass ein Ursprung von WFFS in den Arabern, speziell im Hengst Bairactar Or. Ar. (1813), unwahrscheinlich ist.

Die Abklärung monogener Merkmale und das Monitoring der Verbreitung von Erbfehlern sind für die angewandte Pferdezucht von großer Bedeutung. Durch die Identifikation von Trägern rezessiver Erbfehler ermöglichen Gentests eine strategische Anpaarung und leisten somit einen wichtigen Beitrag zum Tierschutz.

VIII. References

Ablondi, M. *et al.* (2021) *ISAG Abstract Book*.

Abramowicz, A. and Gos, M. (2018) 'Splicing mutations in human genetic disorders: examples, detection, and confirmation', *Journal of Applied Genetics*, 59(3), p. 253. doi: 10.1007/S13353-018-0444-7.

Andersson, L. S. *et al.* (2012) 'Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice', *Nature*, 488(7413), pp. 642–646. doi: 10.1038/nature11399.

Angulo, E. *et al.* (2009) 'Fatal attraction: rare species in the spotlight', *Proceedings. Biological sciences*, 276(1660), pp. 1331–1337. doi: 10.1098/RSPB.2008.1475.

Anthony, D. W. (2010) *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World*. Princeton: Princeton University Press.

Aurich, C. *et al.* (2019) 'Characterization of abortion, stillbirth and non-viable foals homozygous for the Warmblood Fragile Foal Syndrome', *Animal Reproduction Science*, 211(June), p. 106202. doi: 10.1016/j.anireprosci.2019.106202.

Behjati, S. and Tarpey, P. (2013) 'What is next generation sequencing?', *Archives of disease in childhood. Education and practice edition*, 98(6), pp. 236–238. doi: 10.1136/ARCHDISCHILD-2013-304340.

Bellone, R. *et al.* (2013) 'Evidence for a retroviral insertion in TRPM1 as the cause of congenital stationary night blindness and leopard complex spotting in the horse', *PloS one*, 8(10). doi: 10.1371/JOURNAL.PONE.0078280.

Bellone, R. R. *et al.* (2017) 'A missense mutation in damage-specific DNA binding protein 2 is a genetic risk factor for limbal squamous cell carcinoma in horses', *International Journal of Cancer*, 141(2), pp. 342–353. doi: 10.1002/ijc.30744.

Bellone, R. R. *et al.* (2020) 'Warmblood fragile foal syndrome type 1 mutation (PLOD1 c.2032G>A) is not associated with catastrophic breakdown and has a low allele frequency in the Thoroughbred breed', *Equine Veterinary Journal*, 52(3), pp. 411–414. doi: 10.1111/evj.13182.

Bellone, R. R. and Avila, F. (2020) 'Genetic Testing in the Horse', *Veterinary Clinics of North America - Equine Practice*, 36(2), pp. 211–234. doi: 10.1016/j.cveq.2020.03.003.

Bergamaschi, M. *et al.* (2020) 'Genome-wide association study for carcass quality traits and

growth in purebred and crossbred pigs', *Journal of Animal Science*, 98(1). doi: 10.1093/JAS/SKZ360.

Bower, A. *et al.* (2012) 'The genetic origin and history of speed in the Thoroughbred racehorse'. doi: 10.1038/ncomms1644.

Brooks, S. *et al.* (2007) 'A chromosome inversion near the KIT gene and the Tobiano spotting pattern in horses', *Cytogenetic and genome research*, 119(3–4), pp. 225–230. doi: 10.1159/000112065.

Brooks, S. A. *et al.* (2010) 'Morphological variation in the horse: Defining complex traits of body size and shape', *Animal Genetics*, 41(SUPPL. 2), pp. 159–165. doi: 10.1111/j.1365-2052.2010.02127.x.

Brunberg, E. *et al.* (2006) 'A missense mutation in PMEL17 is associated with the Silver coat color in the horse', *BMC genetics*, 7. doi: 10.1186/1471-2156-7-46.

Bush, W. S. and Moore, J. H. (2012) 'Chapter 11: Genome-Wide Association Studies', *PLoS Computational Biology*, 8(12). doi: 10.1371/journal.pcbi.1002822.

Ceballos, F. C. *et al.* (2018) 'Runs of homozygosity: windows into population history and trait architecture', *Nature reviews. Genetics*, 19(4), pp. 220–234. doi: 10.1038/NRG.2017.109.

Charlier, C. *et al.* (2008) 'Highly effective SNP-based association mapping and management of recessive defects in livestock', *Nature genetics*, 40(4), pp. 449–454. doi: 10.1038/NG.96.

Cook, D. *et al.* (2008) 'Missense Mutation in Exon 2 of SLC36A1 Responsible for Champagne Dilution in Horses', *PLoS Genetics*, 4(9), p. 1000195. doi: 10.1371/JOURNAL.PGEN.1000195.

Dias, N. M. *et al.* (2019) 'Warmblood Fragile Foal Syndrome causative single nucleotide polymorphism frequency in Warmblood horses in Brazil', *Veterinary Journal*, 248(226), pp. 101–102. doi: 10.1016/j.tvjl.2019.05.002.

van Dijk, E. *et al.* (2014) 'Ten years of next-generation sequencing technology', *Trends in genetics : TIG*, 30(9), pp. 418–426. doi: 10.1016/j.tig.2014.07.001.

Distl, O. (2013) 'The genetics of equine osteochondrosis', *Veterinary journal (London, England : 1997)*, 197(1), pp. 13–18. doi: 10.1016/J.TVJL.2013.03.036.

Drögemüller, C. *et al.* (2009) 'A Missense Mutation in the SERPINH1 Gene in Dachshunds with Osteogenesis Imperfecta', *PLOS Genetics*, 5(7), p. e1000579. doi: 10.1371/JOURNAL.PGEN.1000579.

Elsik, C. G. *et al.* (2009) 'The genome sequence of taurine cattle: A window to ruminant biology and evolution', *Science*, 324(5926), pp. 522–528. doi: 10.1126/science.1169588.

Ensembl genome browser (2021). Available at: https://m.ensembl.org/Equus_caballus/Info/Annotation#assembly (Accessed: 17 October 2021).

Fages, A. *et al.* (2019) 'Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series', *Cell*, 177(6), pp. 1419-1435.e31. doi: 10.1016/j.cell.2019.03.049.

Firshman, A. M. *et al.* (2003) 'Epidemiologic characteristics and management of polysaccharide storage myopathy in Quarter Horses', *American journal of veterinary research*, 64(10), pp. 1319–1327. doi: 10.2460/AJVR.2003.64.1319.

Firshman, A. M., Baird, J. D. and Valberg, S. J. (2005) 'Prevalences and clinical signs of polysaccharide storage myopathy and shivers in Belgian draft horses', *Journal of the American Veterinary Medical Association*, 227(12), pp. 1958–1964. doi: 10.2460/JAVMA.2005.227.1958.

Grilz-Seger, G., Druml, T., *et al.* (2019) 'Analysis of ROH patterns in the Noriker horse breed reveals signatures of selection for coat color and body size', *Animal Genetics*, 50(4), pp. 334–346. doi: 10.1111/age.12797.

Grilz-Seger, G., Neuditschko, M., *et al.* (2019) 'Association of roan coat color with KIT in Noriker draught horses', in. Prag.

Grilz-Seger, G. *et al.* (2020) 'A Genome-Wide Association Analysis in Noriker Horses Identifies a SNP Associated With Roan Coat Color', *Journal of Equine Veterinary Science*, 88, pp. 1–7. doi: 10.1016/j.jevs.2020.102950.

Grilz-Seger, G. and Druml, T. (2018) *Handbuch Pferdefarben*. 1st edn. Graz: Vehling Medienservice & Verlag GmbH.

Haase, B. *et al.* (2013) 'Accumulating Mutations in Series of Haplotypes at the KIT and MITF Loci Are Major Determinants of White Markings in Franches-Montagnes Horses', *PLoS ONE*, 8(9), p. 75071. doi: 10.1371/JOURNAL.PONE.0075071.

Hamilton, M. B. (2009) *Population Genetics*. 1st edn. Wiley-Blackwell.

Hauswirth, R. *et al.* (2012) 'Mutations in MITF and PAX3 cause "splashed white" and other white spotting phenotypes in horses', *PLoS genetics*, 8(4). doi: 10.1371/JOURNAL.PGEN.1002653.

Hedrick, P. W. and Andersson, L. (2011) ‘Are dogs genetically special?’, *Heredity*, 106(5), p. 712. doi: 10.1038/HDY.2010.98.

Hintz, F. and van Vleck, L. D. (1979) ‘Lethal dominant roan in horses’, *Journal of Heredity*, 70(2), pp. 145–146. doi: 10.1093/OXFORDJOURNALS.JHERED.A109213/2/70-2-145.PDF.GIF.

Holl, H. M. *et al.* (2019) ‘A candidate gene approach identifies variants in SLC45A2 that explain dilute phenotypes, pearl and sunshine, in compound heterozygote horses’, *Animal genetics*, 50(3), pp. 271–274. doi: 10.1111/AGE.12790.

Horse Genome Workshop (2021). Available at: https://horsegenomeworkshop.com/ (Accessed: 2 July 2021).

Human Gene Mutation Database (2022). Available at: http://www.hgmd.cf.ac.uk/ac/stats.php (Accessed: 25 February 2022).

Imsland, F. *et al.* (2016) ‘Regulatory mutations in TBX3 disrupt asymmetric hair pigmentation that underlies Dun camouflage color in horses’, *Natat. Genet.* doi: 10.1038/ng.3475.

ISAG (2022). Available at: https://www.isag.us/index.asp.

Jagannathan, V. *et al.* (2019) ‘Comprehensive characterization of horse genome variation by whole-genome sequencing of 88 horses’, *Animal genetics*, 50(1), pp. 74–77. doi: 10.1111/AGE.12753.

Kalbfleisch, T. S. *et al.* (2019) ‘Author Correction: Improved reference genome for the domestic horse increases assembly contiguity and composition (Communications Biology, (2018), 1, 1, (197), 10.1038/s42003-018-0199-z)’, *Communications Biology*. Nature Research. doi: 10.1038/s42003-019-0591-3.

Kelekna, P. (2009) *The Horse in Human History*. New York: Cambridge University Press.

Kwon, J. M. and Goate, A. M. (2000) ‘The candidate gene approach’, *Alcohol research & health : the journal of the National Institute on Alcohol Abuse and Alcoholism*, 24(3), pp. 164–8. Available at: https://pubmed.ncbi.nlm.nih.gov/11199286/ (Accessed: 4 January 2022).

Lander, E. S. *et al.* (2001) ‘Initial sequencing and analysis of the human genome’, *Nature*, 409(6822), pp. 860–921. doi: 10.1038/35057062.

Lander, E. S. and Botstein, D. (1987) ‘Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children’, *Science (New York, N.Y.)*, 236(4808), pp. 1567–1570. doi: 10.1126/SCIENCE.2884728.

Librado, P. *et al.* (2016) ‘The evolutionary origin and genetic makeup of domestic horses’, *Genetics*. Genetics Society of America, pp. 423–434. doi: 10.1534/genetics.116.194860.

Lim, H. T. *et al.* (2011) ‘Novel alternative splicing by exon skipping in KIT associated with whole-body roan in an intercrossed population of Landrace and Korean Native pigs’, *Animal genetics*, 42(4), pp. 451–455. doi: 10.1111/J.1365-2052.2011.02225.X.

Lindblad-Toh, K. *et al.* (2005) ‘Genome sequence, comparative analysis and haplotype structure of the domestic dog’, *Nature*, 438(7069), pp. 803–819. doi: 10.1038/nature04338.

Liu, X. *et al.* (2021) ‘A single-nucleotide mutation within the TBX3 enhancer increased body size in Chinese horses’, *Current Biology*, pp. 1–8. doi: 10.1016/j.cub.2021.11.052.

Logsdon, G. A., Vollger, M. R. and Eichler, E. E. (2020) ‘Long-read human genome sequencing and its applications’, *Nature reviews. Genetics*, 21(10), p. 597. doi: 10.1038/S41576-020-0236-X.

Ludwig, A. *et al.* (2009) ‘Coat color variation at the beginning of Horse domestication’, *Science*, 324(5926), p. 485. doi: 10.1126/science.1172750.

Makvandi-Nejad, S. *et al.* (2012) ‘Four loci explain 83% of size variation in the horse’, *PloS one*, 7(7). doi: 10.1371/JOURNAL.PONE.0039929.

Mariat, D., Taourit, S. and Guérin, G. (2003) ‘A mutation in the MATP gene causes the cream coat colour in the horse’, *Genetics, Selection, Evolution : GSE*, 35(1), p. 119. doi: 10.1186/1297-9686-35-1-119.

Marklund, L. *et al.* (1996) ‘A missense mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with the chestnut coat color in horses’, *Mammalian Genome*, 7(12), pp. 895–899. doi: 10.1007/s003359900264.

Marklund, S. *et al.* (1999) ‘Close association between sequence polymorphism in the KIT gene and the roan coat color in horses’, *Mammalian Genome*, pp. 283–288. doi: 10.1007/s003359900987.

Martin, K. *et al.* (2021) ‘Fragile Foal Syndrome (PLOD1 c.2032G>A) occurs across diverse horse populations’, *Animal Genetics*, 52(1), pp. 137–138. doi: 10.1111/age.13020.

McCue, M. E. *et al.* (2008) ‘Glycogen synthase (GYS1) mutation causes a novel skeletal muscle glycogenosis’, *Genomics*, 91(5), p. 458. doi: 10.1016/J.YGENO.2008.01.011.

Metallinos, D., Bowling, A. and Rine, J. (1998) ‘A missense mutation in the endothelin-B receptor gene is associated with Lethal White Foal Syndrome: an equine version of

Hirschsprung disease', *Mammalian genome : official journal of the International Mammalian Genome Society*, 9(6), pp. 426–431. doi: 10.1007/S003359900790.

Metzger, J. *et al.* (2020) 'Hanoverian F/W-line contributes to segregation of Warmblood fragile foal syndrome type 1 variant PLOD1:c.2032G>A in Warmblood horses', *Equine Veterinary Journal*, (April), pp. 1–9. doi: 10.1111/evj.13271.

Miglior, F. *et al.* (2017) 'A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle', *Journal of Dairy Science*, 100(12), pp. 10251–10271. doi: 10.3168/jds.2017-12968.

OMIA - Online Mendelian Inheritance in Animals (2022). Available at: https://omia.org/home/ (Accessed: 17 October 2021).

Orlando, L. and Librado, P. (2019) 'Origin and evolution of deleterious mutations in horses', *Genes*, 10(9). doi: 10.3390/genes10090649.

Outram, A. K. *et al.* (2009) 'The earliest horse harnessing and milking', *Science*, 323(5919), pp. 1332–1335. doi: 10.1126/science.1168594.

Reiter, S. *et al.* (2020) 'Distribution of the warmblood fragile foal syndrome type 1 mutation (Plod1 c.2032g>a) in different horse breeds from europe and the United States', *Genes*, 11(12), pp. 1–11. doi: 10.3390/genes11121518.

Ricard, A. *et al.* (2017) 'Endurance exercise ability in the horse: A trait with complex polygenic determinism', *Frontiers in Genetics*, 8(JUN), pp. 1–15. doi: 10.3389/fgene.2017.00089.

Rieder, S. *et al.* (2001) 'Mutations in the agouti (ASIP), the extension (MCIR), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (Equus caballus)', *Mammalian Genome*, 12(6), pp. 450–455. doi: 10.1007/s003350020017.

RIS - Tierschutzgesetz (2022). Available at: https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20003541 (Accessed: 13 January 2022).

Roan Zygosity Test/UC Davis (2021). Available at: https://vgl.ucdavis.edu/test/roan (Accessed: 2 July 2021).

Rosengren Pielberg, G. *et al.* (2008) 'A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse', *Nature genetics*, 40(8), pp. 1004–1009. doi: 10.1038/NG.185.

Rosseau, E. (2017) *Horses of the world*. Princeton University Press.

Rudolph, J. A. *et al.* (1992) 'Periodic paralysis in Quarter Horses: A sodium channel mutation disseminated by selective breeding', *Nature Genetics*, 2(2), pp. 144–147. doi: 10.1038/ng1092-144.

Rueff, A. G. (1855) 'Eigenthümliche Krankheit neugeborener Fohlen (a peculiar disease of newborn foals)', in *Jahrbuch für Pferdezucht 28*, pp. 267–277.

Schaefer, R. J. *et al.* (2017) 'Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds', *BMC Genomics*, 18(1). doi: 10.1186/s12864-017-3943-8.

Schaefer, R. J. and McCue, M. E. (2020) 'Equine Genotyping Arrays', *The Veterinary clinics of North America. Equine practice*, 36(2), pp. 183–193. doi: 10.1016/J.CVEQ.2020.03.001.

Signer-Hasler, H. *et al.* (2012) 'A genome-wide association study reveals loci influencing height and other conformation traits in horses', *PLoS ONE*, 7(5). doi: 10.1371/journal.pone.0037282.

Sponenberg, D. P. and Bellone, R. (2017) *Equine Color Genetics*. 4th edn. Wiley-Blackwell.

Tanaka, J. *et al.* (2019) 'Frameshift Variant in MFSD12 Explains the Mushroom Coat Color Dilution in Shetland Ponies', pp. 1–15.

Voß, K. *et al.* (2020) 'Coat Color Roan Shows Association with KIT Variants and No Evidence of Lethality in Icelandic Horses', *Genes*, 11(6), pp. 1–8. doi: 10.3390/GENES11060680.

Wade, C. M. *et al.* (2009) 'Genome sequence, comparative analysis, and population genetics of the domestic horse', *Science*, 326(5954), pp. 865–867. doi: 10.1126/science.1178158.

Wagner, H. J. and Reissmann, M. (2000) 'New polymorphism detected in the horse MC1R gene', *Animal Genetics*, 31(4), pp. 289–290. doi: 10.1046/j.1365-2052.2000.00655.x.

Willam, A. and Simianer, H. (2011) *Tierzucht*. 2nd edn. utb.

Winand, N. J. (2011) *Identification of the Causative Mutation for Inherited Connective Tissue Disorders in Equines*. Available at: https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2012158711&recNum=1&maxRec=1&office=&prevFilter=&sortOption=Pub+Date+Desc&queryString=FP%3A%28WO2012158711%29&tab=PCT+Biblio (Accessed: 20 September 2020).

Wobbe, M. *et al.* (2019) 'Wege zum angemessenen Umgang mit WFFS und anderen genetischen Eigenschaften', *Vereinigte Informationssysteme Tierhaltung w. V. (vit), Verden*.

Zhang, X. *et al.* (2020) 'Skin exhibits of Dark Ronald XX are homozygous wild type at the Warmblood fragile foal syndrome causative missense variant position in lysyl hydroxylase gene

PLOD1', *Animal Genetics*, pp. 1–2. doi: 10.1111/age.12972.

IX. DANKSAGUNG

An dieser Stelle möchte ich mich recht herzlich für die wertvolle Unterstützung und Motivation bedanken, die ich über die letzten Jahre hinweg von verschiedenen Seiten erhalten habe.

Als Erstes gilt mein Dank Herrn Priv.-Doz. Dr. Medugorac und Herrn O.Univ.-Prof. Dr. Brem, für die Betreuung und Korrektur meiner Arbeit.

Prof. Brem hat bereits im 6. Semester mit seiner Expertise im Rahmen der Tierzuchtvolesung meine Leidenschaft für dieses Fachgebiet geweckt. Eine Diplomarbeit und eine Dissertation später, blicke ich heute dankend auf mehrere Jahre bester Unterstützung, spannende Projekte und aufregende Reisen zurück, die mich nicht nur fachlich, sondern auch persönlich gewissermaßen geprägt haben.

Mein ganz besonderer Dank gilt Frau Priv. Doz. Dr. Wallner, die mir von Beginn an mit ihrem Knowhow unterstützend zur Seite stand und mich durch Höhen und Tiefen begleitete. Es war eine wahre Bereicherung für mich, Teil ihres konstruktiven und herzlichen Teams zu sein.

Recht herzlich bedanke ich mich auch bei Frau Mag. Grilz-Seger und Frau Prof. Bellone für die effektive und angenehme Zusammenarbeit, und bei Frau Mag. Riegler für den ausgezeichneten Support bei jeglichen Laborangelegenheiten.

Außerdem möchte ich mich bei Frau Priv. Doz. Dr. Haring und ihrem Team für die Bereitstellung der Räumlichkeiten und die fachliche Unterstützung zur Bearbeitung historischer Proben bedanken.

Zuletzt möchte ich meiner Familie, insbesondere meinen Eltern, großen Dank und Wertschätzung aussprechen. Sie haben es mir ermöglicht, meinen eigenen Weg zu gehen und mich stets bei meinen Vorhaben unterstützt.