Department of Biomedical Sciences
University of Veterinary Medicine Vienna
Institute of Animal Breeding and Genetics
(Head: O.Univ.-Prof. Dr. Mathias Müller)

Genetic Diversity and Population Structure of Curly Horses

Diploma thesis

University of Veterinary Medicine Vienna

Submitted by:
Sarah Hofmeister

Vienna, October 2017

Supervisors and Institutions

Dr[in]. med. vet. Barbara Wallner
Institute of Animal Breeding and Genetics, University of Veterinary Medicine Vienna

Dr. rer. nat. Priv.-Doz. Claus Vogl
Institute of Animal Breeding and Genetics, University of Veterinary Medicine Vienna

Table of contents

1 Introduction and Hypothesis

A horse breed with only a short and unfortunately inconsistent history of breeding is the American Bashkir Curly Horse, occurring originally in Nevada, USA. Thus, there are many myths and theories about where it comes from. Moreover, opinions of experts and different breeding societies diverge on how American Bashkir Curly Horses should be defined.

Unlike other horses with straight hair coats the American Bashkir Curly Horse has, as the name implies, curly coat (Figure 1). However, the curliness is not restricted to the coat as the hairs of the mane, fetlocks and tail are curly too (Figure 2). The curliness is particularly visible in the winter but sometimes also in summer. Thus, it is also referred to as Curly Horse or simply Curly. The curliness of the coat varies from extreme with very tight and extreme curls e.g., dreadlock manes to minimal with curls in the ears and kinky mane and tail. Thus, even horses with straight hair coat exist within the American Bashkir Curly Horse. In spring some Curly Horses shed entirely the dense curly coat as well as the mane and tail. The new grown summer coat is lesser curly (Lynghaug 2009).

Figure 1: American Bashkir Curly Horse (www.wetcanvas.com).

Figure 2: Curly mane and fur of the ears (2.bp.blogspot.com).

The horse's exterior is a strong, medium-sized shape with characteristic curly fur. A straight, wedged head with a wide forehead is typical, as well as a muscular and short neck with little withers. The average height of the horse is 140 to 150 cm. Further, the Curly Horse distinguishes oneself with a very compact, low torso, beefy, muscular shoulders and croup as well as a robust and short back. Its dark, hard hoofs and the well-developed joints are well known in farmer circles as positive qualities. Besides, these horses are extremely robust, agile and persistent. Therefore, they are perfect workhorses and qualify for each and every purpose (Haller 2009). The Curly Horse is quite hardy and has the ability to withstand colder temperatures more than many other breeds. Tenacity and frugality are qualities particularly well-suited for rough winters (Ranch Namaspamoos 2007).

In contrast to other horse breeds many Curly horses do not show the same flight reaction as they do not frighten easily, thus they are naturally gentle. They wait patiently for help even in bad situation. As a result, Curlies are valued for their gentle disposition as well as their intelligence and versatility. They succeed in all disciplines ranging from working cattle, trail riding, and endurance to showing. Further, due to their gentleness Curly horses are searched partners in therapeutic riding schools (Lynghaug 2009).

Furthermore, it is worth mentioning that the American Bashkir Curly Horses are very special, not only because of their outward appearance, but because of the nature of their coat. Their

curly hair, which closely resembles mohair, is also considered to be hypoallergenic. According to the pilot observational case study of Wolfgang Mitlehner, hyper allergenic patients did not react significantly to the Curly Horses hair (Mitlehner 2013).

Under the microscope the hair of Curly Horses appears flat instead of round, barbed or feathered (Lynghaug 2009). Considering human hair, it is only known that curls mainly depend on the shape and position of the follicle (Figure 3). The rounder the follicle, the straighter the hair and the more oval the follicle, the curlier the hair gets (Figure 4). However, the position, meaning the angle of the follicle referring to the skin is important: in the case the tube is bended the hair will grow curlier compared to an upright follicle tube (Figure 5).

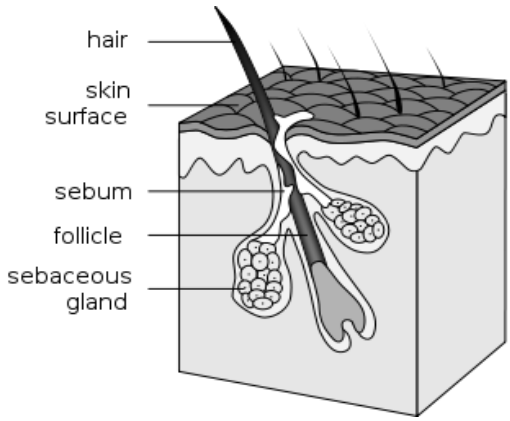

Figure 3: Sketch of a commonly grown human hair (helix.northwestern.edu)

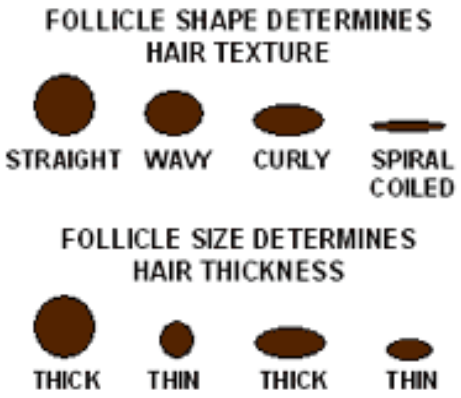

Figure 4: Influence of the follicle tube shape and size on the appearance of the hair (genetics.thetech.org)

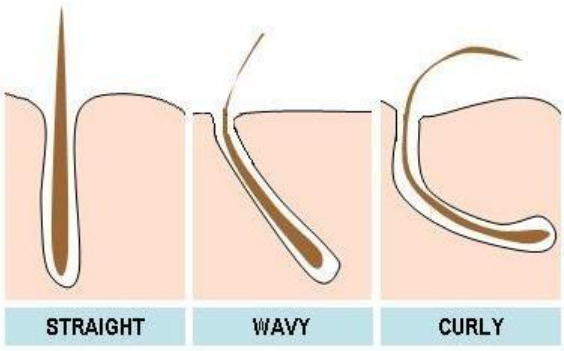

Figure 5: Grades of bender hair follicle tubes and its results in curliness (gurumagazine.org)

Another factor for the occurrence of curly hairs is the biochemical structure. The hair mainly consists of keratin, which is a protein fiber rich in the amino acid cysteine. Cysteine contains Sulphur which is necessary for the formation of interchain disulphide bridges. The number of these bonds influences the flexibility and strength of the hair. A higher amount of disulphide bonds can be found in curly hair (Tirado-Lee 2014).

The biochemical structure of the hair is determined by the genes and it has been demonstrated that the mode of inheritance of curliness in horses with curly coat is recessive. Thus, it can be assumed that these horses are homozygous. Nonetheless, it has been found that besides this recessive mode a dominant gene exists in Curly Horses. The dominant gene seems to be responsible for the annual shedding of the mane and tail (Sponenberg 1990).

At present, there are two major associations occupied with the conservation and breeding of the American Bashkir Curly Horse: one is the International Curly Horse Organisation and the other is the American Curly Horse Registry. Both aim to preserve, promote and protect this rare horse breed. However, there are some differences between these organisations.

The American Bashkir Curly Horse Registry (ABCR) was founded in 1971 by Benny Damele and other people sharing the interest of preserving the Curly horse. A major concern at that time was the inbreeding as they had just a small number of horses they could work with. Thus, they decided to cross their Curlies with four breeds to introduce new blood. The four breeds to fresh

up blood were Arabian, Morgan, Appaloosa and Missouri Fox Trotter. Arabians were selected for their endurance and the same short back as the Curlies, whereas the Morgans have a similar conformation. The shedding of mane and tail hair as well as endurance are characteristics belonging to Appaloosas, Native American horses. The reason for the selection of Missouri Fox Trotter was the smooth gait they have in common with the Curlies (Lynghaug 2009).

Today, ABCR represents the oldest and largest registry for Curly Horses worldwide. There exist affiliated organisations in Canada and Europe. The ABCR has registered over 4,100 Full, 750 Straight and 750 Part Bred Curly Horses since 1971 (ABCR 2016). For the ABCR an eligible American Bashkir Curly Horse has to fulfil the requirement of a special registration consisting of six sections. At least one section has to fit the animal's heritage description precisely, otherwise the horse cannot be a member of the registry.

For example section one is called „ABC Full Curly Registry". In order to fulfil section one's guidelines a curly haired foal has to have both parents fully registered or at least one full curly parent if the other one is full straight. Straight in this case means that the horses' fur is not curly neither the body coat nor the mane or tail. Section two is called „ABC Curly Blood Percentage Registry". For this one a curly foal needs to have no less than 50% Curly blood percentage, but nevertheless less than full blood percentage. Further, one parent needs to be a full registered parent and the other one has to be a blood percentage registered parent.

In contrast, the International Curly Horse Organisation (ICHO) is a non-profit organisation, founded in 2000 and located in Ohio, USA. It was founded by a few Curly owners feeling the need to create an open registry where all Curly horses regardless of their background or breeding are welcomed. The ICHO had over 800 horses registered in its North American Curly Horse Registry (ICHO 2012).

While the ICHO breeders have the freedom to choose whatever type of Curly breeding program they want to pursue, the ABCR differentiates far more accurately between Curly Horses. This is probably the most distinguishing factor between the two big associations. The one where everything is admissible and the other one with very strict rules (ICHO and ABCR 2016).

As currently no one knows for sure the origin of the American Bashkir Curly Horse it is the aim of this thesis to explore the history of curly horses and to investigate the population structure registered by the ABCR, the oldest breeding association for Curly Horses. First, a literature search for historical data on curly horses was performed to gather information on probable sources of origin for the Curly Horse. Second, for population genetic analysis, microsatellite data from routine parentage testing conducted by ABCR for Curly Horses were analysed together with data from other horse breeds. This was done since there are no exact historical records from where the American Bashkir Curly Horses originate, on the one hand American and Russian breeds which served as sources for the American Bashkir Curlies were included. On the other hand, European horses with a long breeding history and closed studbook have been selected, because they were expected to be genetically separated from the American Bashkir Curly Horses. The assumption is to see an existing link between the American Bashkir Curly Horse and American breeds like Morgan Horses, Quarter Horses and Appaloosas, as these may be its source breeds (Heise 1989, Hendricks 1995). However, a relationship with Russian breeds could be expected as well, in matters of the historical findings from northern Europe and Russia. In total twelve different breeds were used, among them American, European, Arabian and Russian breeds.

Thus, after the introduction section 2 of this thesis reports on the different materials and methods used to elucidate the origin of the American Bashkir Curly Horse. Section 3 gives an overview on the results obtained across the literature search for historical data on curly horses and the population analysis. The discussion of findings is conducted in section 4, whereas the conclusion can be found in section 5 followed by the abstract in section 6.

2 Material and Methods

2.1 Literature search for historical data about curly horses

A systematic literature search to find historical data about curly horses was conducted on the basis of scientific databases, the search engine Google and physically by means of books and articles. Only studies in German or English were used for the composing of this thesis. Further, the author was provided with copies of historical texts in electronic form on a server of the ABCR. These sources are indicated in the section references as a personal communication.

2.2 Data for the population analysis

The microsatellite data of 3818 horses were collected by several people and institutes worldwide as part of routine diagnostic procedures. 18 different microsatellite loci from the ISAG-panel were written in the ISAG standardised letter code (Animal Genetics Blog 2013). ISAG stands for International Society for Animal Genetics. Only certified laboratories are empowered to design these codes used for parental analysis (Animal Genetics Blog 2013).

The sample set of 3818 horses encompasses twelve different breeds (Table 1): American Bashkir Curly Horse, Akhal-Teke, Tuvanian Horse, Bashkir, Transbaikalian Horse, Arabian Horse, Thoroughbred, Morgan Horse, Quarter Horse, Appaloosa, Noriker Horse and Lipizzan Horse

Information on the country of origin and their phenotype (BIT) was available for Curly horses. Microsatellite data from American Bashkir Curly Horses were provided by Caren Schuman (ABCR); Akhal-Teke, Tuvanian Horse, Bashkir and Transbaikalian Horse data by Lilija Kalinkova (Laboratory of Genetics, The All-Russian Institute for Horse breeding); Arabian Horse, Thoroughbred, Morgan Horse, Quarter Horse, Appaloosa by Dr. E. Gus Cothran (Ph.D. Texas A&M University, College Station) and Lipizzan as well as Noriker data by O. Univ.-Prof. Dr. med. vet. Gottfried Brem (TZU, Institute of Animal Breeding and Genetics, University of Veterinary Medicine Vienna). Furthermore, there was information about the origin of individual horses for the Lipizzans (Table 2).

Table 1: Information on the breeds used for analysis

Horse breed	general Abbreviation for the breeds	Countries	Number of individuals in raw data	Number of individuals cleaned data	Number of loci in raw data
Tuvanian Horse	Tuva	Russia	40	40	14
Transbaikal	Trans	Russia	51	50	14
Bashkir	Bash	Russia	32	32	14
Curly Horse	Curly	Worldwide	2699	2688	15
Akhal - Teke	AT	Russia	30	30	16
Noriker	Nor	Europe	110	103	12
Lipizzaner	Lipi	Europe	615	536	12
Arabian Horse	Arab	Asia	50	50	16
Thoroughbred	Thor	Europe	70	70	16
Morgan Horse	Morg	America	34	34	16
Quarter Horse	Qua	America	36	36	16
Appaloosa	App	America	51	51	16

In order to work with this genetic data, two specialised population analysis programs were used, STRUCTURE and Arlequin. STRUCTURE, on the one hand, detects subsets of a given data sample by recognising allele frequencies. Therefore, based on analysis of likelihoods it can assign individuals to subpopulations (Porras-Hurtado et al. 2013). Arlequin, on the other hand, can perform both, interpopulational and intrapopulational methods. It is able to compare populations to other populations or examine each on its own (Excoffier et al. 2005).

Table 2: Origin of the Lipizzaner subpopulations

Lipizzaner Horse studs	Number of individuals
Beclean	24
Djakovo	44
Fagaras	84
Kladrub	46
Lipica	27
Monterotondo	58
Piber	75
Szilvasvarad	75
Topolcianky	39
Wien	64

2.3 Data processing

The first step was to write a sample list containing all the microsatellite data together with additional information on the samples (breed, country of origin, curly phenotype). Therefore, it was necessary to eliminate nine of the 18 microsatellite loci, in order to have standardised raw data. These loci were uniquely genotyped for particular breeds. In addition, some individuals with missing data were sorted out to avoid error messages from the analysing programs. Finally, a total number of 3720 individual horses all genotyped for the same nine microsatellite loci (Table 3) remained for the data analysis. The individuals belong to the twelve breeds listed in Table 1.

Table 3: List of microsatellite markers used, allelic size, number of identified alleles, mean number of alleles per locus

Consecutively numbered	Locus	Number of identified alleles	Mean number of alleles per locus	mean observed Heterozygosity	mean expected Heterozygosity
1	AHT4	10	7,000	0,72197	0,73485
2	AHT5	9	6,583	0,75884	0,76864
3	ASB2	17	8,917	0,76682	0,77783
4	HMS3	10	7,000	0,63612	0,73219
5	HMS6	11	6,083	0,70033	0,71967
6	HMS7	9	6,667	0,75281	0,75989
7	HTG10	13	8,750	0,77055	0,77879
8	HTG4	8	5,917	0,65067	0,64288
9	VHL20	12	8,167	0,79135	0,78516
Mean		11	7,231	0,72772	0,74444

Thereafter, the ISAG letter code was transformed into a numerical code since these are needed for the programs used. Further, nullalleles had to be imputed because of the analysis program Arlequin, whereas the program STRUCTURE is able to work with them. This discrepancy might generate problems when comparing the results. Accordingly, the allelic types of phenotypically monomorphic loci with either a nullallele or the homozygous alleles were replaced assuming the Hardy-Weingberg equilibrium. This step was performed using an in-house Perl script written by Claus Vogl a member of the institute.

Loci with nullalleles show a lower observed heterozygosity compared to that expected under Hardy-Weinberg equilibrium (Table 3).

2.4 Data analysis

2.4.1 Arlequin Parities FST model

The program Arlequin (Excoffier et al. 2005) was used for the general population genetic data analysis and genetic differentiation. First, the file with imputed null alleles was used as input for the calculation of the number of alleles and allelic size range over all samples for each locus. As a result, the observed heterozygosities were printed out. Second, we calculated expected heterozygosities under Hardy Weinberg equilibrium for each breed and locus combination and averaged the results over loci for an estimate per breed. Furthermore, we performed a Hardy Weinberg equilibrium test which works locus by locus with a number of 1.000.000 steps in Markov chain and a number of 100.000 dememorization steps (Guo and Thompson 1992) and also the pairwise FST for all pairs of populations.

2.4.2 STRUCTURE USEPOPINFO model

We used the program STUCTURE (Pritchard et al. 2000), a model-based clustering method, in order to infer population structure. The USEPOPINFO model was chosen for our data in combination with the „selection flag“, because of the prior breed information we had. It is noteworthy that this type of model usually assumes that the predefined populations are correct, therefore this model is more stable than a model without prior class structure (Pritchard et al. 2000).

We started by picking a parameter set with a burn-in period of 20 000 MCMC iterations and 20 000 iterations for inference. Following that we had to find the best K, where K represents the number of clusters of individuals in the dataset. Therefore, we used a tutorial of the STRUCTURE software (Sim and Merk 2013) to evaluate the best K, keeping in mind that we should always aim for the smallest value of K that captures the major structure in the data (Pritchard et al. 2000). In doing so, we set a maximum K of ten and ran STRUCTURE from K=1 to K=10 twenty times for each K using the whole dataset as input. After finishing that process, we took the log likelihood for each K, Ln P(D) = L(K) values of the simulation summary, which approximate the posterior probability of the Ks, and calculated the inference of the best K using the delta K method according to Evanno and co-worker (2005).

3 Results

The first part of this section presents the findings of the literature search on historical data to the origin of curly coated horses. The second part reports the results of the population structure analysis by means of two different softwares. Eventually, the population genetic structure of 3720 horses from twelve different breeds, namely Tuvanian Horses, Transbaikalian Horses, Bashkir Horses and Akhal-Teke as Russian breeds, Noriker, Lipizzaner Horses and Thoroughbreds as European breeds, Morgan Horses, Quarter Horses and Appaloosas as American breeds, Arabian Horses and the American Bashkir Curly Horses (Table 1) was studied. The Arlequin software was used to gather information on the genetic diversity of the horse samples whereas the STRUCTURE software was used to infer the population structure among the different horse breeds.

3.1 Historical references of curly coated horses

Horses in general belong to the order *Perissodactyla* in which the genus *Equidae* encompasses the taxonomic family of horses (*Equus caballus*) and related animals, including donkeys (*Equus asinus*), hemione or half-ass (*Equus hemionus*) and zebras (*Equus zebra*) as extant species, and many other species as fossils. All of these subgenera originate from one common ancestor the *Pliohippus*. The *Pliohippus* is the first soliped that existed 10 million years ago at the beginning of the Pliocene. This animal was adapted to a life in the grass steppe and had a dentition and limbs close to extant horses. However, the direct ancestor of the *Equus caballus* is the *Plesihippus* that appeared in the Pleistocene (about 2 million years to 10.000 BC). It was a big and robust animal (Haller 2009).

Although, the oldest equine residues were found in south-west Europe horses spread through Europe, and reached over North Atlantic land corridors the northern part of America. At the end of the Eocene (about 35 million years ago) horses died out in the Old World and their evolution continued in North America. From there they spread in several waves of immigration to Eurasia, either from west across North Atlantic land bridges or from east over the Bering bridge (Figure 6). The Panama bridge formed in the Pliocene enabled the migration to South America (Wehner and Gehring 1995). During the Pliocene horses died out in both Americas,

thus this animal was unfamiliar to Native Americans until they were imported by the Spanish conquistadors in the 16th century (Cox and Moore 1987).

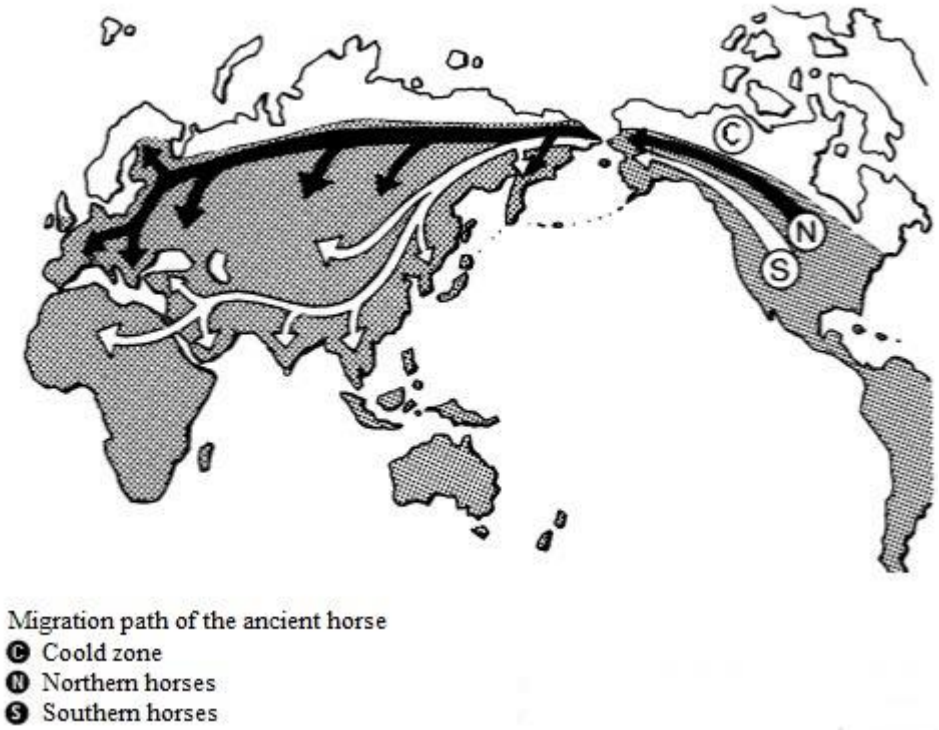

Figure 6: Migration path of the ancient horse (Haller 2009, modified).

About the further development of the horses exist different opinions. One group assumes that the two known wild horse formations Tarpan and Mongolian wild horse (*Equus przewalski*) developed all forms of the domestic horse known to us by human influence. The other group considers more likely a strong type differentiation before the immigration of *Equus* from America to Eurasia (Haller 2009).

Different opinions exist too on the origin of the American Bashkir Horse, whose name suggests that they are descendants of an ancient Russian breed called Bashkir. However, another Russian breed showing quite often a curly coat is the Lokai. Thus, it seems that Bashkir horses having occasionally a curly coat are the result of crossing with the Lokai. The Lokai is a breed occurring in the central and southern regions of Tajikistan (Hendricks 1995, Lynghaug 2009). Nonetheless, curly haired horses were first mentioned on paper in 1247, where the ambassador

of pope Innozenz IV wrote down the insights of their journey (Ramusio due 1247). After that, there is a long period without documentation about curly haired horses in Europe.

Later on, beginning in 1567 the term Bachmatt appears from time to time for horses with curly coat (Von Herberstein 1567). With time more historical writings dealing with curly horse became available. According to one of these descriptions, these horses are very frugal and look for their own food by digging for roots. They are strong, weather resistant and faithful (Von Herberstein 1567, Anonymous 1701). The Bachmatts were also said to be the Tartar's horses (Von Herberstein 1567). In 1742 Bachmatts were said to be Russian horses with „kraushårigt" fur. Moreover, they were supposed to be wild and have hard hoofs, which was an advantage, because they did not need shoes. Furthermore, it was mentioned that among those the „gesprendelten" or tiger-dappled horses were rare. In 1749 the Bachmattes were called „Podolische Pferde" and described as steady horses with long manes, curved forehead and again, it was said, that they have hard hoofs. In the winter their fur is long and frizzly like a poodle's fur and they run days and night without needing to eat more than a mouth full of grass every once in a while. Therefore, they are very frugal (Anonymous 1749).

The Bachmatts were first scientifically named and classified as *Equus hirsutus* (Ridinger 1768). Sometimes, however, it is written about Bachmatts, then Podolian horses or Russian horses. But put into context it is always meant the same, considering Podolia is another name for West Russia which describes the lowlands.

Gradually, the documentations become more and more accurate, so in 1805 the first „Pudelpferde" could be seen in Austria. There they were kept in Europeans oldest zoo, the zoo of Schönbrunn in Vienna, as attraction for the people (Richter 1805). Curly haired horses were not only in Vienna an attraction, but also in Edinburgh and Paris where they were thought to be something special as well. The Parisians even had a frizzled horse in a museum, after it was brought to the city from Cossack troops in 1815 (Blackwood 1831).

While the curly haired horses became more frequent in Europe, the documentations were no longer about the characteristics of these animals. Instead people became more interested in the

horses' origin and were much more curious what resulted in more details in their writings. Thus, we can find in writings from 1831 where Tartar Horses were mentioned and described with all the characteristics of the Bachmatt again, but information on the inhabitation of the deserts between Persia and the Caspian Sea was added as well. These Russian horses were all wandering herds inhabiting the plains of central Asia (Brown 1831). There is also a paper where Russian curly horses have been found alongside the River Don (Blackwood 1831). Furthermore, the frizzled fur gets called a meaningless freak of nature. These „Pudelpferde" can be found in Poland, Russia and the Krim (Schwab 1836), which stands in contrast to a writing from 1821, where the influence of climate is taken into account for being the reason for the curly hair. However, this would mean, that if the cold weather was responsible for the development of this phenotype, the influence must have been long and profound, since these horses can also be found in warmer climate areas (Histoire Naturell 1821).

The first reference of curly haired horses in America were drawings on a Sioux winter count during the winter of 1801-1802 (Thomas et al. 1989). Typically, Indian tribes recorded important events by means of drawings on animal hides, called "winter counts". So, the winter count of 1801 shows the Sioux stealing Curly horses from Crow Indians (Lynghaug 2009). Further, nearly everything we know about the Curly Horse is because of one family, the family Damele. John (former Giovanni) Damele arrived in Eureka, Nevada from Genoa, Italy, in 1879. For 11 years he worked as a woodcutter and provided wood to the charcoal ovens in the region. The charcoal was used as fuel for the smelter furnaces for local silver and lead mines. John could save enough money in order to bring his wife and three children from Italy to join him in Eureka. In 1931 the family came to the curly horses by chance when they discovered a few horses throughout the Hanson, Roberts Mountains and Red Hill Canyon in Nevada. So, the Dameles caught a Curly Horse from a mustang herd. They brought the horse back to the ranch and broke it to ride. Later on, the horse was sold. In the region the winter of 1932 was extremely harsh and all straight hair horses of the family had died but the curly ones had survived. This happened again in the winter 1951 and 1952 (Laughlin 2004). Due to economic problems the family started to capture some of these horses, broke them, trained them to work cattle and eventually sold them. But regardless of the horse's characteristic fur, the family bred and crossed them with other stock horses to produce what they needed as working horses (Figure 7).

Therefore, they bred those curly horses with Quarter Horses or Mustangs in their possession (Thomas et al. 1989).

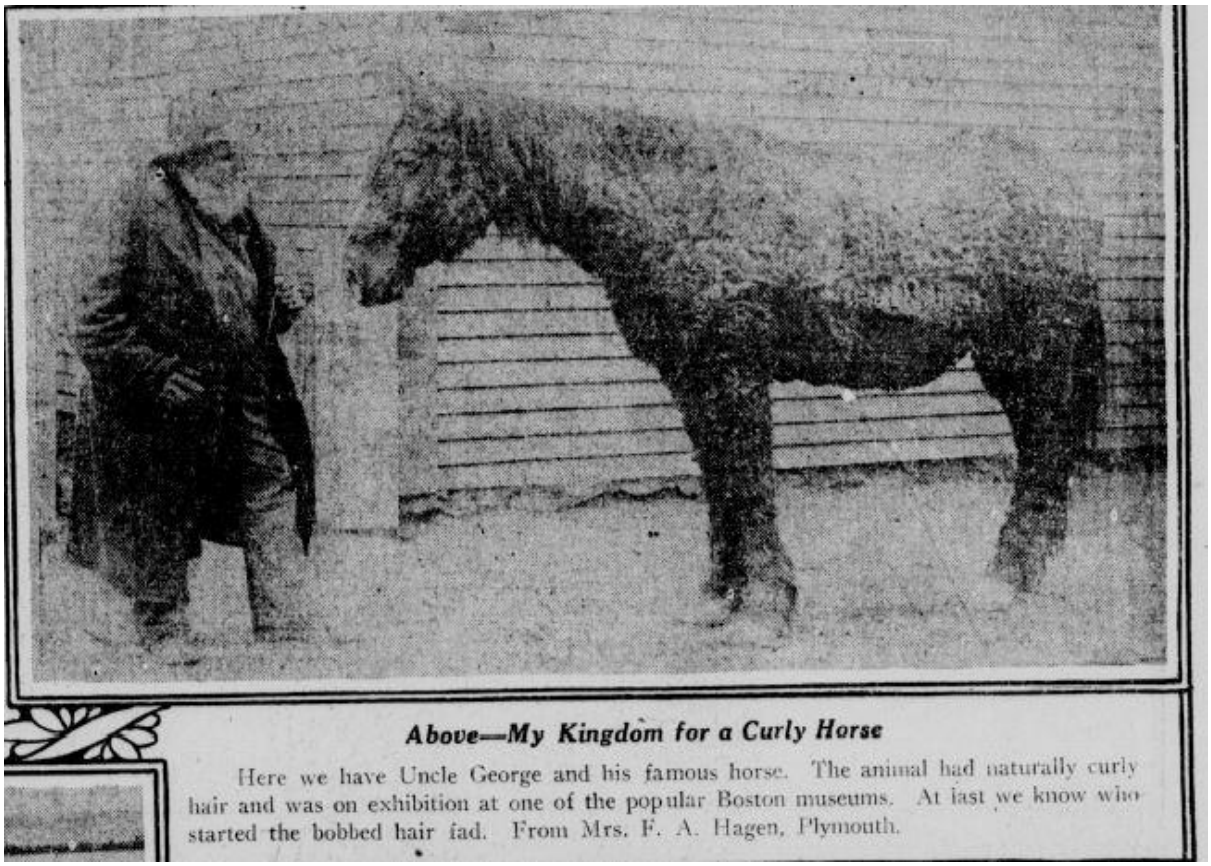

Figure 7: Historical picture from 1922, showing a horse with naturally curly hair in Boston. (4.bp.blogspot.com)

Today, curly haired horses are often seen in North America, but it is still unclear how these animals found their way to America. The historical traditions are sometimes contradictory, if existent at all. Although after reading the paper on „The Curly Horse in America“, it seems, that the American Bashkir Curly Horse is wrongfully named Bashkir. The Bashkirs have been retained in their original cold blood type, however the American Bashkir Curly Horses nowadays are warm-blooded (Thomas et al. 1989).

It is likely that during the Russian occupation of Alaska and parts of the North American coast curly coated horses were imported and eventually sold to miners and farmers from Nevada.

However, horses with curly coat have been found in South America too (Hendricks 1995), at least Charles Darwin noticed horses with curly coat there in the early 1800s (Lynghaug 2009).

Recent genetic analyses found a lot of rare variants in Curly Horses, what indicates that their origin is outside from modern breeds. It seems that they are related to Quarter Horses and other American breeds (Morgans, Saddlebreds and Standardbreds) and to the Great Basin feral horses (Bowling 1990).

3.2 FST analysis with the program Arlequin

Nine polymorphic loci with a minimum number of eight alleles per locus and a maximum number of 17 alleles (Table 3) per locus fulfilled the recommendation from the FAO for genetic distance studies (FAO 1998). According to this recommendation, there should not be less than four alleles per locus. The mean number of alleles per locus ranges between 5,917 and 8,917 (Table 3).

As displayed in Table 4 the mean number of alleles per breed range from 9,889 in American Bashkir Curly Horses to 5,444 in English Thoroughbreds.
If we have a look at the expected heterozygosity calculated for every breed, we see that the mean expected and observed heterozygosity for all breeds are all above 0,500 (Table 4).

For American Bashkir Curly Horses, the mean observed heterozygosity is 0,76244 and the mean expected heterozygosity is 0,77642. In comparison the Lipizzaner Horses have values from 0,70502 observed and 0,74032 expected heterozygosity. Further, American Bashkir Curly Horses and Lipizzaner Horses show a significant deviation from the Hardy Weinberg Equilibrium at multiple loci.

Table 4: Genetic diversity for all 12 breeds.

breed	number of individuals	mean number of alleles	mean observed heterozygosity	mean expected heterozygosity	deviation from HWE at locus no,
Tuvanian Horse	40	7,556	0,73889	0,74522	4,7
Transbaikal	50	8,111	0,72444	0,80357	4,7
Bashkir	32	7,667	0,78125	0,76576	4
Curly Horse	2688	9,889	0,76244	0,77642	2,3,4,5,6,7,9
Akhal - Teke	30	6,222	0,71111	0,71268	-
Noriker	103	7,444	0,66127	0,69237	2,6,7
Lipizzaner	536	7,889	0,70502	0,74032	1,3,4,7,8,9
Arabian Horse	50	5,667	0,68667	0,69692	-
Thoroughbred	70	5,444	0,71429	0,70710	-
Morgan Horse	34	7,000	0,75817	0,75973	-
Quarter Horse	36	6,222	0,74074	0,74309	-
Appaloosa	51	7,667	0,75381	0,79003	-
Mean	310	7,2315	0,728175	0,74443	

Moreover, we calculated the pairwise FST values for all 12 breeds as shown in Table 5. FST values are in general comparatively low between the breeds in our dataset. Every FST measuring 0,05 and below means that there is no significant differentiation (Hartl and Clark 2007).

Therefore, the closer the breeds are to one another, the lower the FST will be. Conversely, the highest FST values imply a relevant population differentiation. For example, the pairwise FST of Arabian Horses and Noriker is 0,152. Furthermore, the Noriker reach values with 0,146 compared against Akhal-Teke, but also with 0,128 against Quarter Horses and Tuvanian Horses the genetic distance is high. In addition, Thoroughbreds compared to Noriker have a FST value of 0,157 and compared to Akhal-Theke 0,142.

In contrast, the American Bashkir Curly Horse has low values for all pairwise comparisons with the lowest FST=0,008 versus Appaloosa horses and the highest FST=0,069 versus Noriker. These results are displayed in Table 5.

Table 5: Results of the pairwise Fst for all 12 breeds (abbreviations according to table 1)

	Tuva	Morg	Qua	App	Trans	Bash	Curly	AT	Nor	Lipi	Arab	Thor
Tuva	0,000											
Morg	0,060	0,000										
Qua	0,059	0,062	0,000									
App	0,047	0,037	0,013	0,000								
Trans	0,040	0,043	0,027	0,014	0,000							
Bash	0,067	0,064	0,052	0,027	0,024	0,000						
Curly	0,061	0,051	0,038	0,008	0,026	0,026	0,000					
AT	0,090	0,112	0,062	0,061	0,051	0,067	0,065	0,000				
Nor	0,128	0,116	0,128	0,074	0,093	0,082	0,069	0,146	0,000			
Lipi	0,088	0,074	0,056	0,033	0,061	0,064	0,035	0,088	0,085	0,000		
Arab	0,113	0,120	0,069	0,068	0,089	0,113	0,067	0,117	0,152	0,069	0,000	
Thor	0,115	0,087	0,061	0,056	0,076	0,093	0,063	0,142	0,156	0,099	0,129	0,000

3.3 STRUCTURE analysis

For the number of populations, the inference of true K was determined by the approach using an ad hoc quantity Delta K (ΔK). In a first step the estimated probability of Ks (Ln P(D)) were calculated by the program. In a further step the second order rate of change of the likelihood (ΔK) divided by the corresponding standard deviation was determined and plotted against the different K values (Figure 8).

L(K) = an average of 20 values of Ln P(D)

L'(K) = L(K)n – L(K)n-1

L''(K) = L'(K)n – L'(K)n-1

ΔK = [L''(K)]/Standard deviation

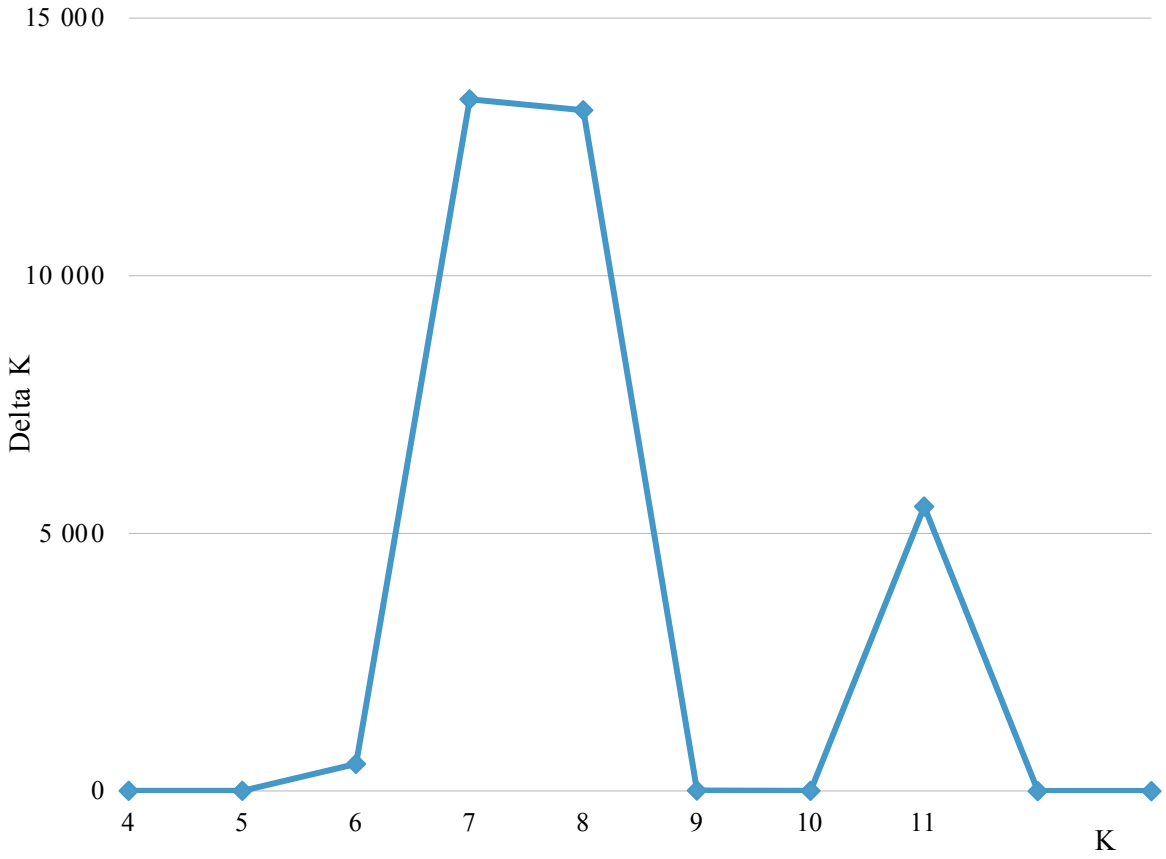

Figure 8: Determination of the best K possible for further analysis with the software STRUCTURE

At the true value of K, the ΔK shows typically a peak (Sim and Merk 2013). Thus, we estimated the optimal value of inferred clusters to be K=7 for the whole dataset. “We may not always be able to know the TRUE value of K, but we should aim for the smallest value of K that captures the major structure in the data” Pritchard et al. (2000).

After that all individual samples got assigned to the 7 clusters showing the proportion of membership of each predefined population in each of these 7 clusters (Figure 9).

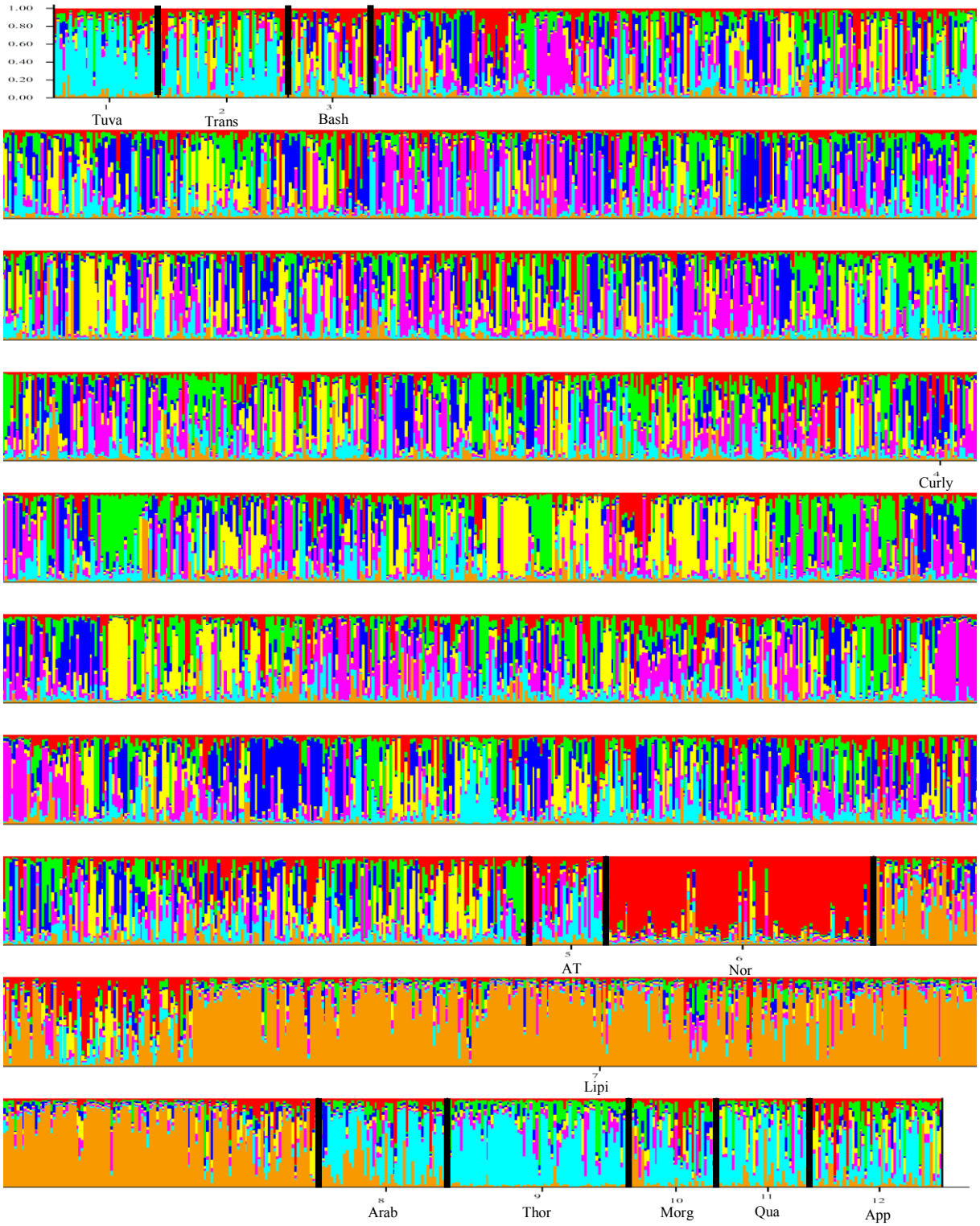

Figure 9: Structure plot from the whole dataset assuming 7 clusters. Individuals are sorted by breeds as indicated. Clusters are marked with different colours and the colours visualise the assignment into respective clusters for each individual.

Lipizzaner and Noriker Horses are assigned most strongly to one particular cluster in comparison to the other ten breeds. In Noriker Horses 0,776 got assigned to cluster 1, whereas assignment probabilities to other clusters are much lower, only 0,029 and 0,052. The Lipizzaner Horses got most often assigned to cluster 7 with a frequency of 0,629, with assignment to the other clusters again evenly low. Thoroughbreds are grouped into Cluster 6 with a value of 0,597. In contrast, the American Bashkir Curly Horse is evenly assigned to all seven clusters (Table 6). Many other breeds, e.g., the Bashkir Horse consistently are represented in all of the seven clusters however, not as evenly as the American Bashkir Curly Horse.

Table 6: Proportion of membership of the pre-defined population in each of the 7 clusters

Breed*	Inferred Clusters							Number of Individuals
	1	2	3	4	5	6	7	
Tuva:	0,123	0,089	0,046	0,054	0,048	0,592	0,048	40
Trans:	0,175	0,104	0,037	0,127	0,069	0,437	0,051	50
Bash:	0,214	0,040	0,124	0,230	0,104	0,203	0,085	32
Curly:	0,114	0,169	0,184	0,166	0,181	0,123	0,062	2688
AT:	0,250	0,031	0,064	0,073	0,266	0,248	0,067	30
Nor:	0,776	0,029	0,030	0,049	0,029	0,036	0,052	103
Lipi:	0,096	0,050	0,052	0,055	0,062	0,055	0,629	536
Arab:	0,108	0,071	0,069	0,036	0,054	0,456	0,206	50
Thor:	0,030	0,146	0,054	0,050	0,093	0,597	0,030	70
Morg:	0,149	0,171	0,072	0,072	0,117	0,319	0,100	34
Quat:	0,052	0,124	0,042	0,113	0,084	0,499	0,086	36
App:	0,141	0,133	0,062	0,151	0,087	0,325	0,100	51

*) For the abbreviation of the breed see table 1

4 Discussion

The opening question was if, it would be possible to locate the genetic origin of the American Bashkir Curly Horse by means of the present microsatellite analysis. This interpretation and discussion of the results should help to answer the initial research question.

According to the various myths and theories people have transported over the years about curly haired horses, we expected to determine genetically confirming findings to make a clear statement. Knowingly, the American Bashkir Curly Horse has a great genetic diversity due to its short breeding history, the different definitions associations have on what an American Bashkir Curly Horse is, and of course the open stud book for the last decades. But even tough, using the highly specialised genetic data it was expected to differentiate the American Bashkir Curly Horse at least from breeds with long closed stud books and an accurate, defined breeding goal, such as Lipizzaner Horses and Noriker Horses.

So, the analysis of microsatellite data from 12 different horse breeds, some central European, some Russian/Asian and also some central American along with the American Bashkir Curly Horses itself, should have allowed an overall view on the genetic situation and the relationship to one another. Especially, the central European breeds were expected to differentiate themselves perfectly from the others, whereas the English Thoroughbred is known from history as a refiner breed, as well as the Arabian Horse. This means they were used over centuries to improve the appearance or gait pattern of other horse breeds. As a result, these breeds appear genetically in several genetic population structures of other breeds.

Looking at the results of the genetic analysis, they revealed that the figures (Figure 9) generated with the software STUCTURE confirm that there is a distinction between the breeds mentioned in sections above and the remaining Russian and American breeds.

Additionally, the analysis showed that the Eurasian breeds do not differentiate themselves clearly from the American breeds. These breeds seemingly inherited genetic information from the Europeans. This picture is consistent with the American Bashkir Curly Horse

shown in figure 9, as the plot illustrates the demarcation of the different horse breeds. Other breeds like the Akhal-Theke or all of the American breeds used in this thesis have unclear population structure as well. Their differentiation level is hardly represented in the analysis. This may be due to the historical fact, that the American horses are more or less European horses, since Spanish conquistadores once brought them there (Tierische Geschichte von ihrer schönsten Seite 2013)

Having this information, we used another software called Arlequin (Excoffier et al. 2005) in order to circumstantiate the STRUCTURE results. After examination of the genetic data of the Lipizzaner horses it was just possible to separate them from the American Bashkir Curly Horses. This result might be attributed to the ten different studs (Table 2) the horses come from, as their breeding goals differ a lot. So, there is an existence of subpopulations within the Lipizzaner Horse breed, which was already confirmed by other studies (Achman et al. 2004, Kavar et al. 2002).

The population subdivision of the Lipizzaner is also evident in the calculations on the deviation from Hardy Weinberg Equilibrium (Table 3 and Table 4). Again, both breeds the Lipizzaner horse and the American Bashkir Curly Horse deviate from the Hardy Weinberg Equilibrium at 6 and 7 loci (Table 4). This could be explained by population substructure, also called the Wahlund effect, since both have low FST values and multiple clusters (Hartl and Clark 2007).

With none of the two programs used in the analysis it was possible to draw a definite relationship between the American Bashkir Curly Horse and one of the other breeds or at least to exclude some breeds from having a genetic relationship with them. It seems that the American Bashkir Curly Horses might be a mixture of all the breeds that were analysed in this research. Not even the American breeds can be described as closer to the American Bashkir Curly Horses, although these are said to be their source breeds.

One possible explanation could be that we used microsatellite data from American Bashkir Curly Horses mainly from America and Europe, also only from one association, the ABCR.

Therefore, this finding may be incorrect due to unilateral genetic information. Another explanation might be the fact that just 9 microsatellite loci were used for the analysis or the uneven sample size for the different breeds examined. Thus, a study of Puechmaille (2016) found that ”the program STRUCTURE does not reliably recover the correct population structure when sampling is uneven”. Further, it has to be considered that according to Evanno et al. (2005) results are sensitive to the type of genetic marker used, the number of loci scored, the number of populations sampled, and the number of individuals typed in each sample.

Moreover, there is the additional challenge of two main Curly horse associations among many smaller ones worldwide that disagree on the definition of the American Bashkir Curly Horse. The associations’ studbooks are based on different breeding goals which clearly affect the horses’ genetics.

5 Conclusion

In contrast to the initial assumption this research could not contribute to a better understanding of the American Curly Horses' origins and how the curly haired horses came to America. Interestingly, not one of the breeds analysed could be entirely differentiated from the American Bashkir Curly Horse.

However, the existence of subpopulations within the Lipizzaner Horse breed, as found in other studies, could be confirmed.

For further research in this field it is recommended to use high density SNP chip data to identify regions in the genome that might be responsible for the curly coat in order to prove, whether the curly coat is simply a phenotype or if this is truly the characteristics of a new breed.

Furthermore, it should be considered to use the same amount of data of each breed to get a more balanced result. Otherwise miscalculations by the analysis programs cannot be completely precluded.

6 Summary in English and German

A horse breed with only a short and unfortunately inconsistent history of breeding is the American Bashkir Curly Horse, occurring originally in Nevada, USA. Thus, there are many myths and theories about where it comes from. Moreover, opinions of experts and different breeding societies diverge on how American Bashkir Curly Horses should be defined. The aim of the current thesis was to determine an existing genetic link between the American Bashkir Curly Horse and 11 other breeds (Akhal-Teke, Tuvanian Horse, Bashkir Horse, Transbaikalian Horse, Arabian Horse, Thoroughbred, Morgan Horse, Quarter Horse, Appaloosa, Lipizzaner Horse and Noriker). Thus, at first a literature search was conducted to gather information on the origin of Curly horses. Second, a microsatellite data analysis of 3720 individuals and 9 loci was conducted through two genetic analysing softwares, Structure and Arlequin. Although historical data on curly coat horses can be traced back to the 13th century the origin of the Curly horses cannot be determined. Further, the data analysis with both softwares could not clearly separate the American Bashkir Curly Horse from the other breeds. However, the analyses revealed the formation of subpopulations within the breed of Lipizzaner horses. Thus, findings of earlier studies (Achman et al. 2004, Kavar et al. 2002) could be confirmed.

Deutsche Zusammenfassung:

Genetische Diversität und Populationsstruktur der Curly Horses

Eine Pferderasse mit nur einer kurzen und leider uneinheitlichen Zuchtgeschichte ist das amerikanische Bashkir Curly Horse, welches ursprünglich aus Nevada (USA) stammt. So gibt es viele Mythen und Theorien darüber, woher es kommt. Darüber hinaus unterscheiden sich die Meinungen von Fachleuten und verschiedenen Zuchtverbänden darin, wie die amerikanischen Bashkir Curly Horses definiert werden sollen. Ziel dieser Diplomarbeit war es, eine genetische Verbindung zwischen dem amerikanischen Bashkir Curly Horse und 11 weiteren Rassen (Achal-Tekkiner, Tuva Pferd, Baschkire, Transbaikalisches Pferd, Araber, Thoroughbred, Morgan, Quarter Horse, Appaloosa, Lipizzaner und Noriker) zu finden. So wurde zunächst eine Literatursuche durchgeführt, um Informationen über den Ursprung der Curly-Pferde zu sammeln. Weiter wurde eine Mikrosatelliten-Datenanalyse an 3720 Individuen und 9 Loci mit Hilfe von zwei genetischen Analyse-Softwaren, Structure und Arlequin, durchgeführt. Obwohl

historische Daten über lockige Pferde bis in das 13. Jahrhundert zurückverfolgt werden können, kann der Ursprung der Curly-Pferde aus der vorhandenen Literatur nicht ermittelt werden. Ferner konnte auch die Datenanalyse mit beiden Programmen das amerikanische Bashkir Curly Horse genetisch nicht von den anderen Rassen trennen. Allerdings zeigten die Analysen die Bildung von Subpopulationen innerhalb der Rasse der Lipizzaner. Damit ergab sich eine Bestätigung der Ergebnisse früherer Arbeiten (Achman et al. 2004, Kavar et al. 2002).

7 Abbreviations

ABCR	American Bashkir Curly Horse Registry
App	Appaloosa breed
Arab	Arabian Horse breed
AT	Akhal-Theke breed
Bash	Bashkir breed
Curly	Curly Horse breed
BIT	Information on the country of origin and the phenotype
FST	Fixation index (statistic test for population genetics)
ICHO	International Curly Horse Organisation
ISAG	International Society for Animal Genetics
SNP	Single-Nucleotide Polymorphism
Lipi	Lippizzaner breed
Morg	Morgan Horse breed
Nor	Noriker breed
Qua	Quarter Horse breed
Thor	Thoroughbred breed
Trans	Transbaikla breed
Tuva	Tuvanian Horse breed
USA	United States of America

8 References

American Bashkir Curly Horse Registry. http://www.abcregistry.org (Accessed 27.7.2016).

Achmann R, Curik J, Dovc P, Kavar T, Bodo I, Habe F, Marti E, Sölkner J, Brem G. 2004. Microsatellite Diversity, population subdivision and gene flow in the Lipizzan horse. Animal Genetics, 35:285-292.

Animal Genetics Blog. 2013. The ISAG Profile for Canine and Equine DNA Testing. November 2013. https://animalgenetics.us/wordpress/2013/11/the-isag-profile-for-canine-and-equine-dna-testing/ (Accessed 26.10.2016).

Anonymous. 1701. An account of Livonia with a relation of the rise progress and decay of Karl Johann von Blomberg. Personal communication.

Anonymous. 1749. Allgemeines Haushaltungslexikon. Leipzig. Personal Communication.

Blackwood. 1831. The Journal of Agriculture. Band 2. Personal communication.

Bowling AT. 1990. Population genetics of curly horses. Genetic conservation of domestic livestock, 32:186-202.

Brown T. 1831. Biographical Sketches. Personal communication.

Cox CB, Moor PD. 1987. Einführung in die Biogeographie. Stuttgart, Gustav Fischer Verlag.

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

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: An Integrated Software Package for Population Genetics Data Analysis. Evolutionary Bioinformatics Online, 1:47–50.

FAO. 1998. Food and Agriculture Organization of the United Nations. Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans. Measurement of Domestic Animal Diversity (MoDAD): Original Working Group Report. http://dad.fao.org (Accessed 9.6.2016).

Guo SW, Thompson EA. 1992. Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. Biometrics, 48:361–372.

Haller M. 2009. Der neue Kosmos Pferdeführer. Vierte Aufl. Stuttgart, Franckh-Kosmos Verlag.

Hartl DL, Clark AG., 2007. In Principles of populations genetics. Sunderland, Sinauer Associates, Inc.

Heise L. 1989. Curly Horse Mystery Remains Unsolved. American Minor Breeds Conservancy News, 7(1). Personal communication.

Hendricks BL. 1995. International Encyclopedia of Horse Breeds. Norman, University of Oklahoma Press.

Histoire Naturell. 1821. Cheval a Poils Frises. Personal communication.

International Curly Horse Organization. http://www.ichocurlyhorses.com (Accessed 27.7.2016).

Kavar T, Brem G, Habe F, Sölkner J, Dovec P. 2002. History of Lipizzan horse maternal lines as revealed by mtDNA analysis. Genetic Selection Evolution, 34:635-648.

Laughlin M. 2004. History of the curly horse. http://curlyhorsecountry.com/history_curlyhorses.htm (Accessed 5.10.2017).

Lynghaug F. 2009. The Official Horse Breeds Standards Guide: The Complete Guide to the Standards of all North American Equine Breed Associations. London: Voyageur Press.

Mitlehner W. 2013. Allergy against horses: Are curly horses an alternative for horse-allergic riders? A pilot observational case study, Allergo Journal, 22(4):244–251.

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo Á., Lareu MV. 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. Frontiers in Genetics, 4:1-13. DOI 10.3389/fgene.2013.00098 (Accessed 22.9.2017).

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population Structure using multilocus genotype data. Genetics, 155:945–959.

Puechmaille S J. (2016), The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Molecular Ecology Resources, 16:608–627. DOI 10.1111/1755-0998.12512.

Ranch Namaspamoos. 2007. The Curly Horse (Breed characteristics and History). http://www.curlyquebec.com/breed.html (Accessed 11.6.2016).

Ramusio due. 1247. Personal communication.

Richter J. 1805. Briefe des jungen Eipeldauers an seinen Herrn Vettern in Kakran. Band 10. Personal communication.

Ridinger JE. 1768. Erste Ordnung Huftiere. Personal communication.

Schwab CL. 1836. Anleitung zur äußeren Pferdekenntnis. Personal communication.

Sim SC, Merk HL. 2013. Downstream Analysis of SNP Markers Using MSA, GGT, and STRUCTURE Software. http://articles.extension.org/pages/32370/downstream-analysis-of-snp-markers-using-msa-ggt-and-structure-software (Accessed 9.6.2016).

Sponenberg DP. 1990. Dominant Curly Coat in Horses. Genetics Selection Evolution, 22:257-260. DOI 10.1051/gse:19900208 (Accessed 22. 9.2017).

Thomas S, Geier D, Bowling A. 1989. The Curly Horse in America, Myth and Mystery. Personal communication.

Tierische Geschichte von ihrer schönsten Seite. 2013. Wie das Pferd nach Amerika kam. https://itigefo.wordpress.com/2013/07/14/wie-das-pferd-nach-amerika-kam/ (Accessed 12.7.2017).

Tirado-Lee L. 2014. The Science of Curls. https://helix.northwestern.edu/blog/2014/05/science-curls (Accessed 27.7.2016).

Von Herberstein S. 1567. Moscoviter wunderbare Historien. Personal communication.

Wehner R, Gehring W. 1995. Zoologie. 23. Aufl. Stuttgart, New York: Thieme.

9 List of figures

10 List of tables

11 Acknowledgement

First of all, I want to thank Dr[in]. Barbara Wallner very much for allowing me to pursue my diploma thesis under her supervision. Without her expertise and constant availability to discuss my work, it would probably not have been possible for me to finalise this thesis simultaneously with my clinical training.

Additionally, I am thankful to Dr. Claus Vogl who was always ready to give advice when needed.

Furthermore, Caren Schuhmann and Karen Sarul are the reason why this master thesis was even possible, so thank you for commissioning it and of course for providing lots of historical information as well.

Amongst others I want to mention my gratefulness especially to Wolfgang and also to Elisabeth and Franz for their technical support throughout this period. They were always there in times when I wished to have more knowledge about computer science.

Moreover, I have to thank Petra sincerely for stylistic advice.

Last but not least, I am most thankful to my parents for all their love, support and all the opportunities they have given me. Since my early childhood, when they recognised my interest and passion regarding horses and animals in general, they have encouraged me ever since to live my dreams. They taught me that with firmness, hard work and endurance everyone can reach their goals. For this and much more I am most thankful.