RESEARCH ARTICLE

Memories, museum artefacts and excavations in resolving the history of maternal lineages in the Finnhorse

Laura Kvist[1] | Johanna Honka[1] | Daniela Salazar[1] | Tuija Kirkinen[2] | Karin Hemmann[3]

[1]Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland

[2]Department of Cultures, Archaeology, University of Helsinki, Helsinki, Finland

[3]Faculty of Agriculture and Forestry/ Animal Science, University of Helsinki, Helsinki, Finland

Correspondence
Laura Kvist, Ecology and Genetics Research Unit, PO Box 3000, 90014 University of Oulu, Oulu, Finland.
Email: laura.kvist@oulu.fi

Funding information
Alfred Kordelinin Säätiö; Erkki Rajakosken Rahasto; Oulun Yliopisto

Abstract

We used historical DNA samples to examine the history of a native horse breed, the Finnhorse. Samples were collected from private collections, museums, schools and excavations, representing the times prior to, during, and after the foundation of the breed; from the end of the 19th century and throughout the 20th century. We sequenced a fragment of mitochondrial DNA from these historical samples to study the history and evolution of maternal lineages of horses back to the early days of the breed, compared the mitochondrial DNA sequence diversity of different historical periods and modern day Finnhorses, estimated the effective population sizes, and searched for both temporal and geographic population genetic structure. We observed high maternal haplotype and nucleotide diversity at the time during the foundation of the breed, and a decrease in both measures during 1931–1970. In addition, we observed losses of some haplotypes present in the early stages of the breed. There was only slight evidence of geographical or temporal population structure. This study is, to our knowledge, the first to use such temporal sampling to reveal the history of a specific animal breed.

KEYWORDS
control region, Finnhorse, historical DNA, mitochondrial DNA, temporal sampling

INTRODUCTION

Domestic animals have been bred by humans for thousands of years, but breeds, as we know them today, have existed for much shorter periods. When a species is domesticated, only selected individuals, which carry desired traits, are allowed to reproduce. This kind of selective breeding forms the basis of breeds. This is true for horse breeds as well. The horse is commonly agreed to have been domesticated in the Pontic-Caspian Steppe, about 5300–5500 years ago (Anthony, 2007) with recent genomic studies pinpointing the lower Volga-Don as the domestication location (approximately 4000–4200 years

ago) of horses from which the modern horse lineage has descended (Librado et al., 2021). A few horse breeds are believed to have existed for a long time. One of them is the Caspian horse, a miniature horse that is believed to originate from the oldest known horse breed that existed already close to the time of horse domestication (Pluta et al., 2020), and another is the Icelandic horse that is assumed to originate from horses brought to Iceland by Norse settlers 1100 years ago (Aðalsteinsson, 1981). However, most horse breeds have more recent origins, some tracing back to the Middle Ages (e.g. the Arab, Kucera, 2013; and the Andalusian, Valera et al., 2005), but most have been officially founded within the last

couple of centuries (e.g. the Standardbred, Cothran et al., 1986; the Hanoverian, Hamann & Distl, 2008; the Friesian, Savelkouls, 2015).

In northern Europe, especially in Fennoscandia and Iceland, early horse populations evolved in geographically isolated areas, thereby retaining archaic features and reflecting adaptations to harsh local conditions (Saastamoinen & Mäenpää, 2005). As a result, the horses were tenacious, small in size, and able to cope with the scarcity of food. From the 18th to early 20th centuries, the breeding of native horses in northern Europe aimed to raise better quality horses for use in agriculture, forestry, military, and changing demands of the modern world. Desired traits were commonly added to landraces by crossing native horses with foreign breeds for gaining height, strength or certain stature. These early stages of breeds' histories have been studied almost entirely based on historical documents and, to some extent, also using DNA-analyses of modern horses (Castaneda et al., 2019; Hreidarsdóttir et al., 2014; Kvist et al., 2019).

The Finnish horse has remained essentially the same since the 18th century, but the size has increased and the colour spectrum has decreased during the 20th century. The Finnish horse has been characterised by its versatility since its early days. Although the Finnish horses were used mainly in agricultural work and strength was a desired trait, speed was also highly valued. According to historical documents, in the 19th century and at the beginning of the 20th century, the Finnish horses were to some extent crossed with foreign breeds, in order to achieve more strength, size and speed. The studbook of these Finnish horses was established in 1907, after which the breed was designated as a Finnhorse and no foreign breeds were allowed to be used for breeding any longer (Peltonen & Saastamoinen, 2007). Based on genetic studies, the Finnhorse is closely related to other Nordic or eastern horse breeds, e.g. the Norwegian Fjord, North Swedish horse, Gotland Russ, Icelandic horse, Estonian Native, Tuva horse, Yakutian horse and the Mongolian horse (Petersen et al., 2013; Sild et al., 2019). Many of the Nordic horse breeds have preserved a unique gene pool and are distinct from other native breeds around the Baltic Sea region and North-Eastern Europe (Sild et al., 2019).

Genetic studies of the domestication history of the horse, with the use of ancient or modern DNA, have revealed a previously unknown complexity in their early domestication. A good example of this is the finding that the Przewalski's horse is likely to be a feral domesticated horse of the Botai culture and not a true wild horse (Gaunitz et al., 2018). Further, genetic studies of tens of horse breeds have been published to gain insight into genetic variation within the breeds, relationships among breeds, and to reveal events, such as population size changes, and selection in the histories of the breeds (e.g. Kvist et al., 2019; Petersen et al., 2013; Poyato-Bonilla et al., 2022). These studies have used samples of recent DNA collections and state-of-the-art analysis methods, resulting in an increasing amount of knowledge on the genetic properties of breeds, which can be utilised in the improvement of endangered breeds (e.g. Janova et al., 2013; Luís et al., 2007).

The early stages of the formation of breeds have, to our knowledge, never been studied using DNA samples originating from those times. Here, we have collected historical samples of horses from the end of the 19th century and throughout the 20th century, originating from private collections, museums, schools, and excavations, in order to examine the history of the Finnhorse. Using mitochondrial DNA (mtDNA) sequences obtained from these historical samples and samples of modern day Finnhorses, we (1) studied how genetic diversity of the Finnhorse has changed since the foundation of the breed; and (2) elucidated the history and evolution of the maternal lineages of horses back to the early days of the breed. Further, (3) we estimated effective population size changes spanning the time before, during, and after the foundation of the breed.

MATERIALS AND METHODS

Sampling

The samples were either hair, tooth, bone, or hoof samples obtained from excavations or collected from museums, schools, and private collections (Figure 1 and Table S1). The samples were collected between 2017 and 2021 in Finland and included horses born between 1850 and 1990. A total of 416 samples were processed for DNA extraction and amplification of mitochondrial control region, of which 325 were hair samples and 91 were tooth, bone or hoof samples. For detailed description of sampling, DNA extraction, PCR, and sequencing, see Appendix S2.

Sequence analyses

CodonCodeAligner v.9.0.1 (https://www.codoncode. com/aligner/) and BioEdit v.7.0.0. (Hall, 1999) were used for editing and aligning sequences. After removing identical sequences originating from samples collected from the same place, we grouped the sequences in the following two ways: (1) Sequences of all the historical samples were divided into groups according to the period they represented: before 1900 ($N = 53$), 1901–1930 ($N = 31$), 1931–1950 ($N = 29$), 1951–1970 ($N = 25$), and 1971–1990 ($N = 5$). Henceforth, this dataset is designated in the text as 'Historical horses'. (2) Based on the biography of the samples, we omitted those historical samples that we considered might not represent Finnhorses (e.g. samples from agricultural schools that were used for educational purposes, most representing

FIGURE 1 Examples of sampled material for this study. (a) Horsehair (sample number mEca337, Table S1), private collection. Photo H. Aaltonen. (b) A hoof (mEca245) from a teaching collection of Mustiala agricultural school, Finland. (c) A horsehair filled pillow (mEca 331), private collection. Photo P. Hämäläinen-Eerola. (d) The taxidermied head of a Finnhorse named Eri-Aaroni (mEca2), in Laihia museum exhibition, Finland. (e) A tooth model (mEca269) from a teaching collection of the Otava agricultural school, Finland. (f) The remains of a Finnhorse named Valokas (mEca125), a famous trotter, in an archaeological excavation in Koria, Finland, in 2020. Photos T. Kirkinen unless otherwise mentioned

TABLE 1 Diversity indices of historical horses' and Finnhorses' datasets divided by periods

	N	H/H_{indels}	$\hat{h}$	π	θ
Historical horses	143	55/70	0.970 (0.005)	0.0194 (0.0006)	0.0207 (0.0059)
Before 1900	53	34	0.978 (0.009)	0.0225 (0.0009)	0.0226 (0.0074)
1901–1930	31	24	0.978 (0.015)	0.0219 (0.0015)	0.0219 (0.0079)
1931–1950	29	20	0.968 (0.018)	0.0182 (0.0012)	0.0196 (0.0073)
1951–1970	25	19	0.973 (0.019)	0.0198 (0.0016)	0.0184 (0.0071)
1971–1990[a]	5	5	1.000 (0.126)	0.0264 (0.0041)	0.0299 (0.0161)
Finnhorses	810	104/116	0.951 (0.003)	0.0200 (n.d.)	0.0269 (0.0059)
Before 1900[a]	2	2	1.000 (0.250)	0.0292 (0.0146)	0.0292 (0.0219)
1901–1930	23	20	0.988 (0.016)	0.0231 (0.0019)	0.0227 (0.0086)
1931–1950	22	15	0.952 (0.029)	0.0186 (0.0012)	0.0190 (0.0075)
1951–1970	16	14	0.983 (0.028)	0.0192 (0.0018)	0.0154 (0.0067)
1971–2000	145	50	0.959 (0.008)	0.0224 (0.0007)	0.0232 (0.0065)
2001–2017	602	87	0.950 (0.004)	0.0201 (n.d.)	0.02645 (0.0060)

Note: Standard deviations are indicated within parenthesis.

Abbreviations: N, number of samples, H, number of haplotypes, H_{indels}, number of haplotypes considering also indels, $\hat{h}$, haplotype diversity, π, nucleotide diversity, and θ, mutation parameter.

[a]Omitted from statistical analyses due to low sample size.

the period before 1900). We then combined these samples with modern Finnhorse sequences from Kvist et al. (2019). These were subsequently grouped according to periods: before 1900 (N = 2), 1901–1930 (N = 23), 1931–1950 (N = 22), 1951–1970 (N = 16), 1971–2000 (N = 145), and 2001–2017 (N = 602). Henceforth, this dataset is designated in the text as 'Finnhorses'.

These data were further divided into western and eastern geographic groups (Figure S3). DnaSp v.6.12 (Librado & Rozas, 2009) was used to calculate the basic genetic diversity estimates for all the temporal groups. The significance of nucleotide and haplotype diversities among periods was tested by permutation with Genetic_diversity_diffs v.1.0.6. (Alexander, 2017)

in R v.4.0.5 (R Core Team, 2021). In order to further assess the temporal genetic variation between the modern and the historical populations, temporal statistical parsimony networks were constructed using the program TempNet (Prost & Anderson, 2011) in R v.4.0.5 separately for the Historical horses' and the Finnhorses' data. In addition, pairwise φ_{ST} values were calculated and the exact test of population differentiation performed using Arlequin v.3.5.2.2 (Excoffier & Lischer, 2010). The same program was applied to test between the geographical groups. The best supported substitution model and gamma value were inferred from Akaike and Bayesian information criterion values obtained in MEGA X program (Tamura et al., 2013).

We searched for changes in female effective population size (N_e) through time using the Bayesian skyline analysis with the program BEAST v.2.6.6 (Bouckaert et al., 2019). The input was created using BEAUTI v.2.5 (Bouckaert et al., 2019), and TRACER v.1.7.2 (Rambaut et al., 2018), was used to inspect the posterior distributions and effective sample sizes, and to analyse the skyline plot. Details of all methods are found in Appendix S2.

RESULTS

Genetic variation

The haplotype diversity ($\hat{h}$) of Finnhorses (0.951) was slightly lower than for Historical horses (0.970), but nucleotide diversity (π) and mutation parameter (θ) were higher for Finnhorses (0.0200 and 0.0269) than for Historical horses (0.0194 and 0.02063), respectively (see Table 1). With the Historical horses' dataset, nucleotide diversity comparisons showed low, although not quite significant p-values only between the period before 1901–1930 and 1951–1970 ($p = 0.055$; all other p-values >0.130). No significant differences were detected in haplotype diversity with this dataset (all p-values >0.394). Nucleotide diversities between periods of the Finnhorses' dataset were significantly different between periods 1931–1950 and 1971–2000 ($p = 0.024$), and 1951–1970 and 1971–2000 ($p = 0.050$). Haplotype diversities did not differ between the periods (all p-values >0.181). However, with both datasets, decrease of haplotype and nucleotide diversities are seen in the period 1931–1950 compared to other periods (Figure 2).

Genetic structure

No significant differentiation was seen between the periods in pairwise φ_{ST} (Historical horses: φ_{ST}-values between −0.026 and 0.007, p-values 0.234–0.991, for Finnhorses: φ_{ST}-values between −0.019 and 0.002, p-values 0.315–0.838; Table S4). For Finnhorses, the exact
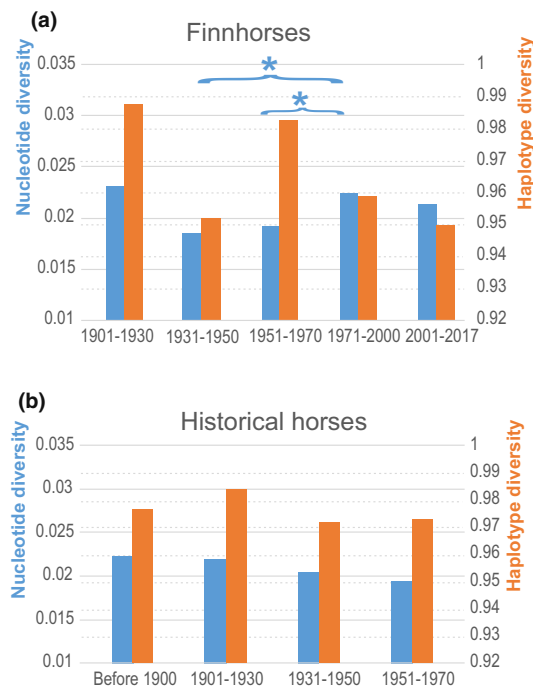

FIGURE 2 Nucleotide (blue) and haplotype (orange) diversities divided by periods for (a) Finnhorses' and (b) historical horses' datasets. Statistically significant differences for nucleotide diversities between periods are shown with asterisks, haplotype diversities did not differ significantly between any periods

test of differentiation suggested differentiation between the period 1901–1930 from periods 1971–2000 and 2001–2017, and period 1951–1970 from 2001–2017, whereas for Historical horses' data, no differentiation was detected. Eastern and western groups of Finnhorses did not differ significantly, with respect to φ_{ST} (0.0146, $p = 0.189$), but again, the differentiation was significant with the exact test of differentiation ($p = 0.000$).

Temporal statistical parsimony networks (Figure 3a,b) showed that only a few haplotypes were detected throughout the 20th century. Among the Historical horses, altogether 41 of the 70 haplotypes were private (each detected in only one individual). Only two of the haplotypes were present throughout all periods (individual IDs of the sequences included in Figure 3 can be found in Table S5). The Finnhorses data revealed 54 private haplotypes out of the 115 detected, and altogether 68 haplotypes were found in one period only, largely from the most recent periods. Only two were found throughout the 20th century. Of the haplotypes present in the two earliest periods, 17 haplotypes of the 22 were detected also in the two latest periods. Altogether 10 haplotypes out of 26 present in 1931–1970 were no longer found in the two latest periods. Of the 18 horse haplogroups (defined by Achilli et al., 2012) only B, D, G, H, I, L, M, and Q were detected in the two earliest periods, but all, except groups K and O, are found in the latest period. A more comprehensive report of these results is described in Appendix S6.
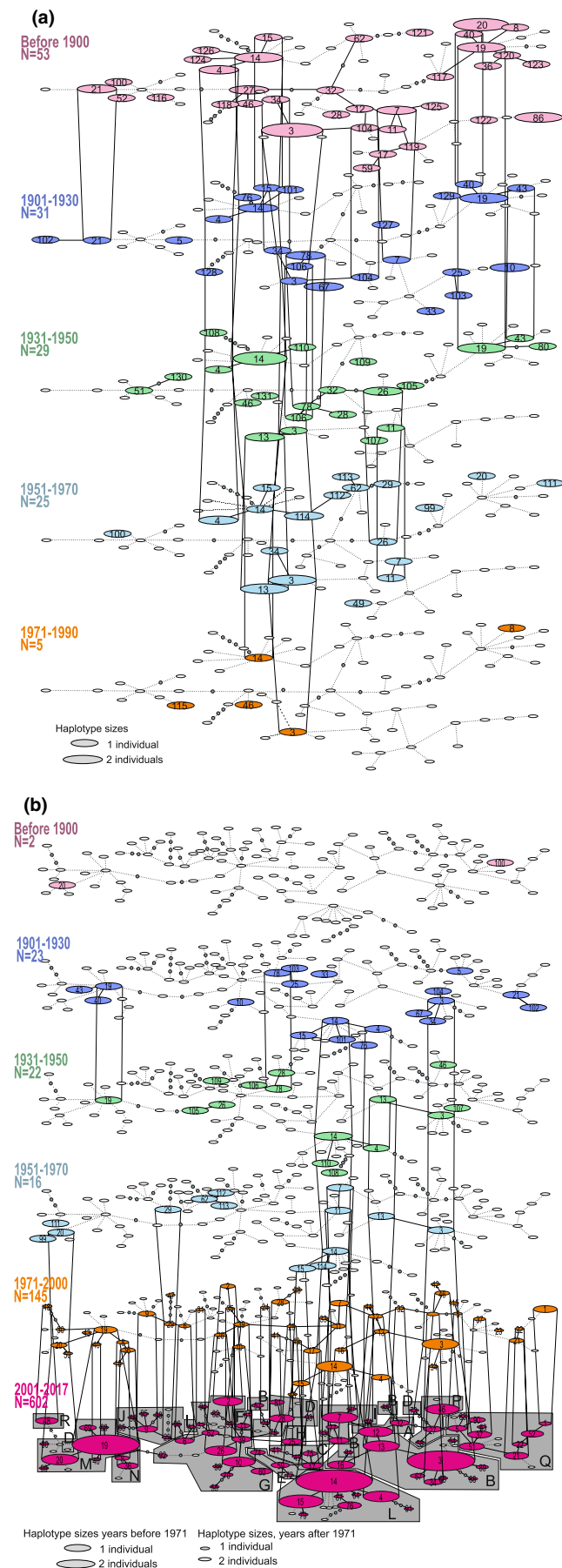
FIGURE 3 Temporal statistical parsimony network of mitochondrial DNA control region (277 bp) showing (a) historical horses' dataset divided into five periods: before 1900 (purple), 1901–1930 (blue), 1931–1950 (green), 1951–1970 (turquoise), and 1971–1990 (orange) and (b) Finnhorses' dataset divided into six periods: before 1900 (purple), 1901–1930 (blue), 1931–1950 (green), 1951–1970 (turquoise), 1971–2000 (orange), and 2001–2017 (magenta). Coloured circles indicate haplotypes present in that period and their sizes are proportional to the number of identical haplotypes found. Numbers within the circles depict haplotype IDs and correspond in panels (a) and (b) (see also Table S5). Small black dots connecting haplotypes within each period illustrate hypothetical haplotypes not present in the sample, and bars depict substitutions between the haplotypes. Small white circles illustrate haplotypes found in other periods, but which are absent from that particular period. Haplotypes of the latest period are grouped into haplogroups (shaded in grey) and represented with letters A–R after Achilli et al. (2012)

Effective population size

The posterior distribution of the Markov chain Monte Carlo runs in BEAST was unimodal but, even though many Markov chain Monte Carlo runs were used, the effective sample size remained a bit low (119). The Bayesian skyline showed an increasing female effective population size until the end of the 19th century, reaching an N_e of 16003 (median; 95% CI 3096–5.82 × 10^5) by the 1880s. Thereafter, it started to slowly decline until the 1970s when the N_e was the smallest, 13209 (median; 95% CI 3283–5.63 × 10^5). The present N_e was estimated to be 13806 (median; 95% CI 3078–5.83 × 10^5; Figure 4).

DISCUSSION

This study is, to our knowledge, the first to use samples originating from the times of breed foundation to reveal the history of a specific animal breed. We observed high maternal haplotype and nucleotide diversities at the time during the foundation of the Finnhorse breed, as well as a decrease in both during 1931–1970. In addition, some haplotypes present in the early stages of the breed were seemingly lost as these were not detected in modern Finnhorses. A drop in genetic diversity was accompanied by a reduction of the female effective population size. We found some evidence of geographical and temporal population structure.

DNA extraction and PCR success of historical samples

About 25% of the hair samples yielded mtDNA control region sequences, while the success rate of bone, tooth, and hoof samples was about 85%. While demonstrating bone, tooth, and hoof samples are especially good sources of historical DNA, horse hairs are a much more accessible sample type, as many horse hair bundles have been stored by private persons. In hairs, the poor PCR
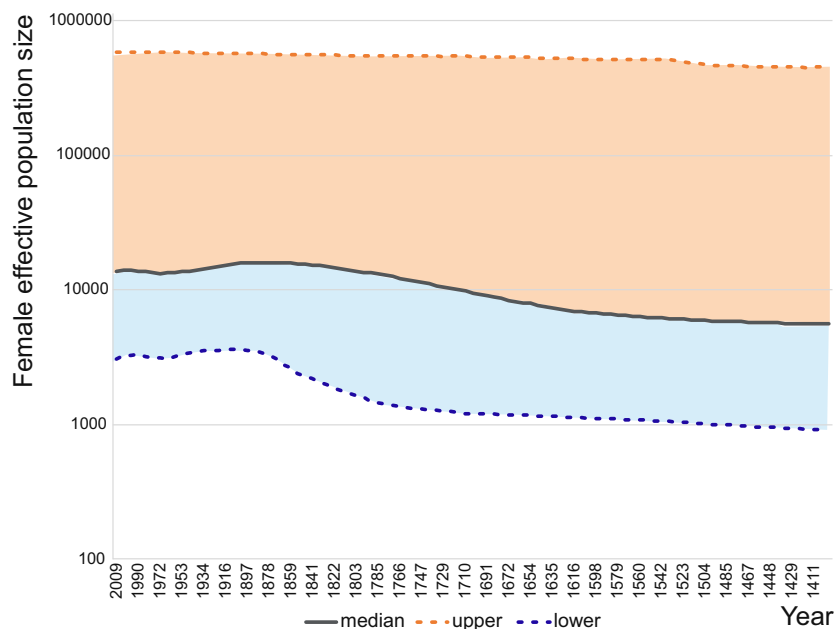
FIGURE 4 Bayesian skyline plot of female effective population size of historical (1850–1990) and modern Finnhorses. The effective population size is on a logarithmic scale. Median value is shown with a solid line and 95% upper and lower credibility intervals with dashed lines

success of many samples can be explained by several factors such as damage caused by microorganisms and insects, temperature variations and humidity during storage (Kirkinen et al., 2022). In addition, most of the DNA in hairs is found in the follicles, but most of our samples did not have any. Both sample types contained fragmented DNA, seen as a low PCR success with the primer pair amplifying the longest fragment. Ancient and historical DNA is highly fragmented, as DNA degradation starts immediately after death (Hofreiter et al., 2001; Pääbo et al., 2004), and in keratinized hairs, DNA is already fragmented during the lifetime of the animal (Bengtsson et al., 2012).

Genetic diversity and structure, and the history of the Finnhorse

Improvement of the Finnish horse population began already in 1835 by placing state-owned breeding stallions (so-called 'crown stallions') in the care of private people. By the turn of the 20th century, the horse population was affected by well over 100 state-owned crown stallions, as well as some crossings with foreign breeds (Solala, 2021). Alfthan (1911) described the 19th century horse population as mixed, and distinguished several local landraces with different characteristics. Finnhorses in the east and west of the country were claimed to differ in their morphology (Alfthan, 1911; Vilkuna, 1967; broader information of the history of the Finnhorse in Appendix S7). In our Historical horses' dataset, nucleotide diversity was the highest in the period before 1900 that predates the establishment of the Finnhorse studbook (in 1907) and could

thus include more foreign-breed background than later periods.

Nucleotide diversity in Finnhorses was the highest in 1901–1930, but dropped slightly from previous periods during 1931–1970 (also in the Historical horses' dataset). This was most probably caused by several factors. The studbook, with strict regulations for the traits of accepted horses, severely limited the number of breeding individuals. The effect of this may have been somewhat delayed, not least because the owners of the horses did not always follow all the instructions given. As an example, they used horses that were successful in horse races for breeding, even though they were not approved in the studbook. World Wars I and II, and changes in agricultural and forestry practices during the 1950s, together with the movement of people from rural to urban areas, decreased horse numbers (Kumpulainen, 2007) and probably also affected the genetic diversity.

Changes in breeding practices have, likewise, also probably had an effect on genetic diversity. Within the breed, a variable number of breeding sections have been formed for different types of horses. Since 1971, the studbook has been divided into four sections: trotters, riding horses, pony-sized horses, and draught horses (Ojala et al., 2007). These divisions could have further limited genetic diversity, as only certain horses met the criteria and were used for breeding. By contrast, many registered Finnhorses were (and still are) not in the studbook and were (and are) used for breeding, leading to the creation of a common gene pool for the breeding sections. In line with estimates of genetic diversity, maternal effective population size showed an increase until the end of the 19th century and decreased slightly after the turn of the 20th century, fitting well with the Finnhorse's history

(Appendix S7) and the effect of the establishment of the studbook and breeding sections.

The exact test for population differentiation considered genetic differentiation to be significant among western and eastern Finnhorses. However, we found no geographical differentiation based on φ_{ST}. Similarly, some of the periods differed being based on the exact test, but not based on φ_{ST}. The reason for the disparate results of these tests probably lies in the way φ_{ST} also takes into account the genetic distances between haplotypes, but the exact test relies solely on haplotype numbers. In horses, mitochondrial differentiation is rare, and even between breeds there does not tend to be differentiation; instead, most breeds comprise of most haplogroups having large and overlapping distributions over vast geographical areas (Lu et al., 2019). This is most likely to be due to frequent dispersal events (both naturally in the wild and by people when domesticated) in the history of horses, and the large maternal input to the breeding pool (Lippold et al., 2011), which is a result of the mating structure in the wild and the use of a few stallions but many mares in breeding.

Haplotypes and haplogroups

Of the 18 horse haplogroups defined by Achilli et al. (2012) only eight (haplogroups B, D, G, H, I, L, M, and Q) were detected in the two earliest periods (before 1930s) in the Finnhorses' dataset, whereas 16 (haplogroups K and O are missing) are found in the modern Finnhorse. Evidently, sample sizes have an effect, but it is still interesting to note that all six of the presently most common haplogroups (B, G, I, M, L, and Q; Kvist et al., 2019) were detected already in the earliest periods. Haplotype diversity of modern samples is the lowest from all periods, although the sample size is the largest. It seems that some maternal lineages were lost during the 20th century, as some of the haplotypes present in the earliest periods are not found in modern Finnhorses. Whether these were just not detected in modern Finnhorses, lost randomly, or selected off due to these lineages carrying foreign or otherwise undesired traits, can only be speculated.

CONCLUSION

This study showed that historical material sourced from museums, private individuals, and archaeological excavations are valuable for examining the history of domestic animal breeds. Instead of examining modern DNA, this type of material provides a tool to follow the development of breeds as it happens, using a time series of historical samples. With historical DNA, we followed the evolution of maternal lineages and could detect, for example, loss of maternal lineages, and changes in female effective population size and genetic diversity.

CONFLICT OF INTEREST
The authors declare no conflicts of interest.

ORCID
Laura Kvist https://orcid.org/0000-0002-2108-0172

REFERENCES
Achilli, A., Olivieri, A., Soares, P., Lancioni, H., Hooshiar, K.B., Perego, U.A. et al. (2012) Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proceedings of the National Academy of Sciences, USA*, 109, 2449–24454. https://doi.org/10.1073/pnas.1111637109

Aðalsteinsson, S. (1981) Origin and conservation of farm animal populations in Iceland. *Journal of Animal Breeding and Genetics*, 98, 258–264. https://doi.org/10.1111/j.1439-0388.1981.tb00349.x

Alexander, A. (2017) genetic_diversity_diffs v1.0.6. https://github.com/laninsky/genetic_diversity_diffs

Alfthan, A. (1911) *Suomalainen hevonen*. Helsinki: Kustannusosakeyhtiö Otava. (In Finnish).

Anthony, D.W. (2007) *The horse, the wheel and language: how bronze age riders from the Eurasian steppes shaped the modern world*. Princeton: Princeton University Press.

Bengtsson, C.F., Olsen, M.E., Brandt, L.Ø., Bertelsen, M.F., Willerslev, E., Tobin, D.J. et al. (2012) DNA from keratinous tissue. Part 1: hair and nail. *Annals of Anatomy*, 194, 17–25. https://doi.org/10.1016/j.aanat.2011.03.013

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A. et al. (2019) BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 15, 4, e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Castaneda, C., Juras, R., Khanshour, A., Randlaht, I., Wallner, B., Rigler, D. et al. (2019) Population genetic analysis of the Estonian native horse suggests diverse and distinct genetics, ancient origin and contribution from unique patrilines. *Genes*, 10, 629. https://doi.org/10.3390/genes10080629

Cothran, E.G., MacCluer, J.W., Weitkamp, L.R. & Guttormsen, S.A. (1986) Genetic variability, inbreeding, and reproductive performance in standardbred. *Zoo Biology*, 5, 191–201. https://doi.org/10.1002/zoo.1430050213

Excoffier, L. & Lischer, H. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Molecular Ecology Resources*, 10, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Gaunitz, C., Fages, A., Hanghøj, K., Albrechtsen, A., Khan, N., Schubert, M. et al. (2018) Ancient genomes revisit the ancestry of domestic and Przewalski's horses. *Science*, 360, 111–114. https://doi.org/10.1126/science.aao3297

Hall, T. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.

Hamann, H. & Distl, O. (2008) Genetic variability in Hanoverian warmblood horses using pedigree analysis. *Journal of Animal Science*, 86, 1503–1513. https://doi.org/10.2527/jas.2007-0382

Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M. & Pääbo, S. (2001) Ancient DNA. *Nature Reviews Genetics*, 2, 353–359. https://doi.org/10.1038/35072071

Hreidarsdóttir, G., Arnason, T., Svansson, V. & Hallson, J. (2014) Analysis of the history and population structure of the Icelandic horse using pedigree data and DNA analyses. *Icelandic Agricultural Sciences*, 27, 63–79.

Janova, E., Futas, J., Klumplerova, M., Putnova, L., Vrtkova, I., Vyskocil, M. et al. (2013) Genetic diversity and conservation in a small endangered horse population. *Journal of Applied Genetics*, 54, 285–292. https://doi.org/10.1007/s13353-013-0151-3

Kirkinen, T., Honka, J., Salazar, D., Kvist, L., Saastamoinen, M. & Hemman, K. (2022) Determination of different predictors affecting DNA concentration isolated from historical hairs of the Finnhorse. *Journal of Archaeological Science: Reports*, 41, 103262. https://doi.org/10.1016/j.jasrep.2021.103262

Kucera, F. (2013) Arab horse in the culture of orient. *Sensus Historiae*, 12, 11–26.

Kumpulainen, M. (2007) Suomenhevonen sotien jälkeen—jälleenrakennuksen aika. In: Saastamoinen, M. (Ed.) *Suomenhevonen*. Gummerus, Jyväskylä: Suomen Hippos Ry, pp. 36–50 (In Finnish).

Kvist, L., Niskanen, M., Mannermaa, K., Wutke, S. & Aspi, J. (2019) Genetic variability and history of a native Finnish horse breed. *Genetics Selection Evolution*, 51, 1–14. https://doi.org/10.1186/s12711-019-0480-8

Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Librado, P., Khan, N., Fages, A., Kusliy, M.A., Suchan, T., Tonasso-Calvière, L. et al. (2021) The origins and spread of domestic horses from the Western Eurasian steppes. *Nature*, 598, 634–640.

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

Lu, H., Bai, H., Wang, l. & Zhang, T. (2019) Abundant genetic diversity and maternal origins of modern horses. *Canadian Journal of Animal Science*, 99, 929–934. https://doi.org/10.1139/cjas-2018-0201

Luís, C., Cothra, E.G. & do Mar Oom, M. (2007) Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. *Journal of Heredity*, 98, 232–237. https://doi.org/10.1093/jhered/esm009

Ojala, M., Peltonen, T. & Saastamoinen, M. (2007) Suomenhevosen jalostus. In: Saastamoinen, M. (Ed.) *Suomenhevonen*. Gummerus, Jyväskylä: Suomen Hippos Ry, pp. 127–166 (In Finnish).

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N. et al. (2004) Genetic analyses from ancient DNA. *Annual Review of Genetics*, 38, 645–679. https://doi.org/10.1146/annurev.genet.37.110801.143214

Peltonen, T. & Saastamoinen, M. (2007) Suomenhevosen alkuperä ja historia—kehittyminen omaksi roduksi. In: Saastamoinen, M. (Ed.) *Suomenhevonen*. Gummerus, Jyväskylä: Suomen Hippos Ry, pp. 9–16 (In Finnish).

Petersen, J.L., Mickelson, J.R., Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E. et al. (2013) Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PLoS One*, 8, e54997. https://doi.org/10.1371/journal.pone.0054997

Pluta, M., Bańka, K., Cieśla, A. & Rogala, Ł. (2020) The state of breeding and use of Caspian horses in Europe and around the world. *Acta Scientiarum Polonorum Zootechnica*, 19, 79–84. https://doi.org/10.21005/asp.2020.19.3.10

Poyato-Bonilla, J., Laseca, N., Demyda-Peyrás, S., Molina, A. & Valera, M. (2022) 500 years of breeding in the *Carthusian strain* of Pura Raza Español horse: an evolutional analysis using genealogical and genomic data. *Journal of Animal Breeding and Genetics*, 139, 84–99. https://doi.org/10.1111/jbg.12641

Prost, S. & Anderson, C.N.K. (2011) TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. *Methods in Ecology and Evolution*, 2, 663–667. https://doi.org/10.1111/j.2041-210X.2011.00129.x

R Core Team. (2021) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67, 901–904. https://doi.org/10.1093/sysbio/syy032

Saastamoinen, M.T. & Mäenpää, M. (2005) Rare horse breeds in northern Europe. In: Bodó, I., Alderson, L. & Langlois, B. (Eds.) *Conservation genetics of rare horse breeds*. Wageningen: Wageningen Academic Publishers, pp. 129–136.

Savelkouls, J. (2015) The Friesian horse and the Frisian horse. The (re)invention and the historicity of an iconic breed. *De Vrije Fries*, 95, 9–42.

Sild, E., Värv, S., Kaart, T., Kantanen, J., Popov, R. & Viinalass, H. (2019) Maternal and paternal genetic variation in Estonian local horse breeds in the context of geographically adjacent and distant Eurasian breeds. *Animal Genetics*, 50, 757–760. https://doi.org/10.1111/age.12835

Solala, H. (2021) Suomalaisen hevosrodun synty, Maatiaishevonen ja kotieläinjalostuksen kansainvälinen murros 1893–1907. Doctoral dissertation. http://urn.fi/URN:ISBN:978-952-03-2031-7. (In Finnish).

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30, 2725–2729. https://doi.org/10.1093/molbev/mst197

Valera, M., Molina, A., Gutierrez, J.P., Gomez, J. & Goyach, F. (2005) Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. *Livestock Production Science*, 95, 57–66. https://doi.org/10.1016/j.livprodsci.2004.12.004

Vilkuna, K. (1967) Tamma ja "Kobylitsa Korela" (Tamma-Karjala) vuodelta 1338. *Virittäjä*, 71, 176.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.