Date: 26-02-25

Final report

Genomic Characterization as a Tool Towards Sustainable Breeding of Nordic Native Horse Breeds
(Genomisk kartläggning som redskap för hållbart avelsarbete i inhemska nordiska hästraser)

Project number: H-21-47-653

Project period: 2022-01-01 to 2024-12-31

Main applicant:
Susanne Eriksson, Swedish University of Agricultural Sciences (SLU), *susanne.eriksson@slu.se*

Co-applicant(s):
Peer Berg, Norwegian University of Life Sciences, *peer.berg@nmbu.no* (main applicant from NMBU, Norway)
Sofia Mikko, SLU, *sofia.mikko@slu.se*
Iryna Shutava, SLU, *iryna.shutava@slu.se*
Maria Kjetså, NordGen, *maria.kjetså@nordgen.org*
Therese Selle, Norwegian Horse Association, *therese.selle@nhest.no*
Also active in the project: Nathalie Almaas Smogeli, NMBU, *nathalie.almaas.smogeli@nmbu.no*

Part 1: Detailed summary

Syftet med projektet var att studera genetisk variation, liksom brist på sådan inom och mellan inhemska nordiska hästraser på genomisk nivå med hjälp av data från helgenomsekvensering. Därtill var målet att jämföra tecken på avelsurval i genomet mellan de olika raserna och att studera i vilken utsträckning raserna bär på potentiellt skadliga genetiska varianter. Vi arbetade utifrån hypotesen att det finns distinkta genetiska rasskillnader med för raserna unika genetiska varianter i det studerade materialet, men även genetiska likheter mellan raser med liknande bakgrund och användning. Vidare förväntade vi oss att hitta genetisk variation inom och mellan raser både i form av SNP-markörer (variation i enskilda baspar) och strukturell variation i genomet.

DNA ur blodprov från 190 hästar av raserna svensk ardenner, nordsvensk brukshäst, gotlandsruss, kallblodstravare, dölehäst, norsk fjordhäst, nordlands/lyngshäst och färöisk häst sekvenserades med Illuminas NovaSeq 6000-system, med ett genomsnittligt läsdjup av drygt 18 gånger. Detta gav efter strikt kvalitetskontroll över 12 miljoner SNP-markörer och ca 35 tusen strukturella variationer för vidare analyser av variation, släktskap, inavel, tecken på selektion, samt riskbedömning av genetiska varianter.

Sekvenseringen av proverna tog längre tid och blev något dyrare än förväntat. Det gjorde att vi ännu inte hunnit jämföra hela vårt material med det från andra raser. Samarbetspartners är

The project has been financed by:
STIFTELSEN HÄSTFORSKNING

kontaktade och datainsamling påbörjad och analyserna är planerade att göras vid NordGen efter projektets slut. Sekvenseringen gav dock bättre kvalitet med högre läsdjup än vi hade räknat med och det material projektet tagit fram kommer vara till nytta i flera kommande forskningsprojekt. Vi fick även möjlighet att jämföra mot ett nytt referensgenom för finsk häst i tillägg till det från engelskt fullblod som är mindre närbesläktad med de studerade nordiska raserna.

Resultaten bekräftade dölehästens, nordsvenska hästens och den kallblodiga travhästens gemensamma ursprung, både genom låga fixationsindex (Fst) och principalkomponentanalys. De andra raserna visade starkare genetisk differentiering, speciellt den färöiska hästen som genomgått en kraftig genetisk flaskhals och för vilken inavelsnivån är mycket hög (genomiska F_{ROH}-värden över 30%). Den svenska ardennerhästen hade genomgående högst heterozygotigrad och lägst inavelsgrad (12,5%) vilket förmodligen beror den mycket stora population som funnits i Sverige fram till moderniseringen av jordbruket i mitten av 1900-talet.

Mönstret i släktskap mellan raserna var liknande då SNP-markörer och strukturella variationer analyserades. Generellt identifierades något mer genetisk variation när referensgenomet från finsk häst användes än det från engelskt fullblod, men skillnaderna var måttliga. Vi fann att alla raser bär på potentiellt riskabla genetiska varianter, med så kallat GERP (Genomic Evolutionary Rate Profiling) score högre än 4. Antalet sådana varianter skiljde sig dock åt mellan raserna där den mest inavlade, färöisk häst, hade totalt lägst antal, men fler som var vanliga, eller till och med fixerade i rasen än de andra raserna.

De preliminära resultaten för spår av avelsurval i genomet visade vissa förväntade skillnader, som att genomiska regioner med gener av betydelse för tillväxt och kroppsstorlek har selekterats för hos ardenner, medan sådana för liten kroppsstorlek verkar ha selekterats för hos färöisk ponny. Regioner med potentiell koppling till exempel fruktsamhet, beteende, anpassning till låg fodertillgång, etc. verkar skilja ut sig hos olika raser. Detta behöver studeras vidare mer i detalj då många av regionerna främst har studerats hos andra arter som människa tidigare och den kända kopplingen till egenskaper inte är direkt översättningsbar till hästars egenskaper.

De olika analyserna visade minskad genetisk mångfald hos alla studerade raser. Resultaten understryker vikten av att upprätthålla genetisk mångfald och undvika stark inavel för rasernas överlevnad. Situationen i de flesta raser var inte alarmerande, men det är tydligt att ytterligare minskande betäckningssiffror i vissa av raserna utgör en risk för rasernas framtida överlevnad. Information och rådgivning har givits till rasföreningar och uppfödare under projektets gång, men det finns ytterligare behov av detta, liksom av att även de svenska rasföreningarna får tillgång till, och hjälp med, moderna metoder för att beräkna inavelsgrader och planera parningar. Sådant stöd finns redan genom Stiftelsen Norsk Hestesenter i Norge. Även stöd för kryokonservering av genetiskt material samt studier av hur väl det fungerar för de inhemska nordiska raserna vore önskvärt.

Part 2: Main report

Introduction

The Nordic countries are home to several unique national native horse breeds. All Nordic native horse breeds have been found to be at risk of extinction, except for the Finnhorse and the Icelandic Horse. The situation is more critical for some breeds than for others. There are broad knowledge gaps regarding our native breeds, and the true status of their genetic diversity and relatedness has not previously been extensively studied (Kierkegaard et al., 2020).

The main aim of the project was therefore to detect genetic diversity, and/or lack thereof, on a genomic level within and between Nordic native horse breeds using whole genome sequencing (WGS) data. Our main focus was to explore the Norwegian and Swedish native horse breeds, as several of these breeds share genetic background and history across the border. We also aimed to compare signatures of selection in the genome of breeds with different characteristics to learn more about potential genetic adaptations, and to study the extent of potentially harmful alleles segregating in the different breeds.

We got access to a new reference genome from another Nordic horse breed, a Finnhorse of working horse type (Pokharel et al., 2024), and we could thus also compare results from using this genome with the reference genome from a Thoroughbred horse (Kalbfleisch et al., 2018) that was, at the time, the most commonly used.

Material and methods

Sample collection and selection

The study included whole blood samples from 25 Fjord horses, 30 Nordland/Lyngen horses, 30 North Swedish horses, 28 Coldblooded Trotters, 15 Gotland Ponies, 10 Faroese horses, 30 Dola horses, and 22 Swedish Ardennes. Most horses were sampled by veterinarians with owner's consent at breeding shows and other gatherings in Sweden and Norway (with ethical permits Dnr 5.8.18-05055/2019 (Sweden) and FOTS ID 29635 (Norway)), but some Nordland/Lyngen horse samples were from a biobank and Faroese horse samples were provided from the breed association Felagið Føroysk Ross. We avoided including close relatives, but the average relatedness varied depending on breed. A final selection of samples to sequence was also done based on the DNA quality of the prepared samples.

DNA sequencing, data pre-processing, and quality control

Genomic DNA was prepared from whole blood samples. TruSeq PCRfree DNA libraries were prepared and paired-end sequencing was performed on an Illumina NovaSeq 6000 sequencing instrument. FastQC tools were used to inspect the raw data, and the GATK best practices workflow for short variant discovery was used for the data preprocessing. The samples were first mapped on the reference horse genome EquCab3.0 (Kalbfleisch et al., 2018), and later also on the EquCab_Finn (Pokharel et al., 2024) reference genome. The average read depth of samples was observed at a level of more than 18X for the former and 13X for the latter genome, which was higher than originally expected. Sequences from three of the 190 samples showed a lower quality, however.

The final SNP genotypes in the data were obtained using GATK tools. For SNP data quality control was imposed in two rounds using PLINK 2.0 (Chang et al., 2015), one for all analyses (leaving over 21 million SNPs) and one for strict filtering needed for the analysis of runs of homozygosity and heterozygosity (reducing the number to above 12 million SNPs).

The detection of the structural variants was done with a pipeline designed by us based on Delly Germline structural variants calling workflow. About 35 thousand structural variants remained after quality control.

Analysis

The SNP data was analyzed with several different types of methods, and comparisons were made between the two reference genomes:

- Fixation Index (F_{ST}) values between populations were estimated using the Hudson method in PLINK 2.0 (Chang et al., 2015).
- Principal Component Analysis (PCA) was done in PLINK 2.0.
- Runs of Homozygosity (ROH) and inbreeding (F_{ROH}) based on ROH were estimated, using PLINK 1.9 (Purcell et al, 2007).
- Heterozygosity (HET) in the populations was estimated using PLINK 2.0.
- Historical effective population size (Ne) was estimated based on LD decay patterns per breed using SNeP 1.1 (Barbato et al., 2015).
- Identity by state (IBS) and identity by descent (IBD), and the fraction of pairwise comparisons (FPC) comparing the number of animals found to have IBD using different genome assemblies, were estimated in PLINK 2.0.
- Private Alleles were detected using bcftools 1.17 (Danecek et al., 2021). The SNPs were then used with Ensembl Variant Effect Prediction (VEP) (McLaren et al., 2016) at default options to detect nearby genes of interest, and pathway analysis was then done in Enrichr-KG (Evangelista et al., 2023).
- Cross-population haplotype scores (XP-EHH) were analyzed in Selscan 2.0.2 (Zachary, 2024) to detect selective sweeps differing between populations.
- Shared Haplotypes were analyzed using the SHAPEIT software (Delaneau et al., 2012), and the R-package "GHap" (Utsunomiya et al., 2016).
- Genetic load was evaluated utilizing the Ensembl VEP command-line tool to obtain Genomic Evolutionary Rate Profiling (GERP) scores from the EPO 35 mammalian alignment (Cooper et al., 2005).

For the analysis of private alleles and cross-population haplotype scores, the three breeds with a shared background, Coldblooded trotter, Dola horse and North Swedish horse were treated as one breed.

The proportion of detected structural variants by sub-types of variants was done with designed R-scripts. Fst, PCA, and private alleles in whole population and within and between breeds were also analyzed for structural variants, with PLINK1.9, bcftools, vcftools (Danecek et al, 2011) and R-scripts designed by us.

Results and discussion

Fixation Index (F_{ST})

The F_{ST} estimates confirmed a close relationship between the North Swedish and Dola horses, for which the lowest (0.04) value was estimated. This was expected based on both breed history and more recent exchange of genetic material between these breeds. Low to moderate values (0.07-0.10) were also found between these breeds and the Coldblooded trotter which originates from the Dola and North Swedish horse breeds, though with some known introgression of genetic material from Standardbred trotters (Jäderkvist Fegraeus et al., 2018). Moderate distances (0.11-0.12) were seen between the Norwegian Fjord horse, Swedish Ardennes and North Swedish horse, that are all of working horse type, though the Fjord horse is lighter built than the other breeds. The most inbred and isolated breed, the Faroese pony, consistently showed, as expected, the highest F_{ST} values regardless of breed combination. This is to be expected, as the breed has most likely lost genetic variation that may have been shared by other Nordic breeds previously.

Principal Component Analysis (PCA)

The PCA showed a pattern of relatedness between the breeds that largely agreed with their known breed history (Fig. 1). Again, the Coldblooded trotter was close to its founding breeds North Swedish horse and Dola horse, and this group of breeds were separated from the rest. In the EquCab_Finn mapped data, the first principal component separated the other working horse breeds Swedish Ardennes and Fjord horse from the pony-sized breeds Gotland pony, Faroese horse and Nordland/Lyngen horse, but the second principal component showed clear divisions between all of these breeds. The proportion of variation explained by the first (17.2%) and second (15.5%) principal components was relatively high, illustrating the different genetic backgrounds and selection for different usage of the breeds.

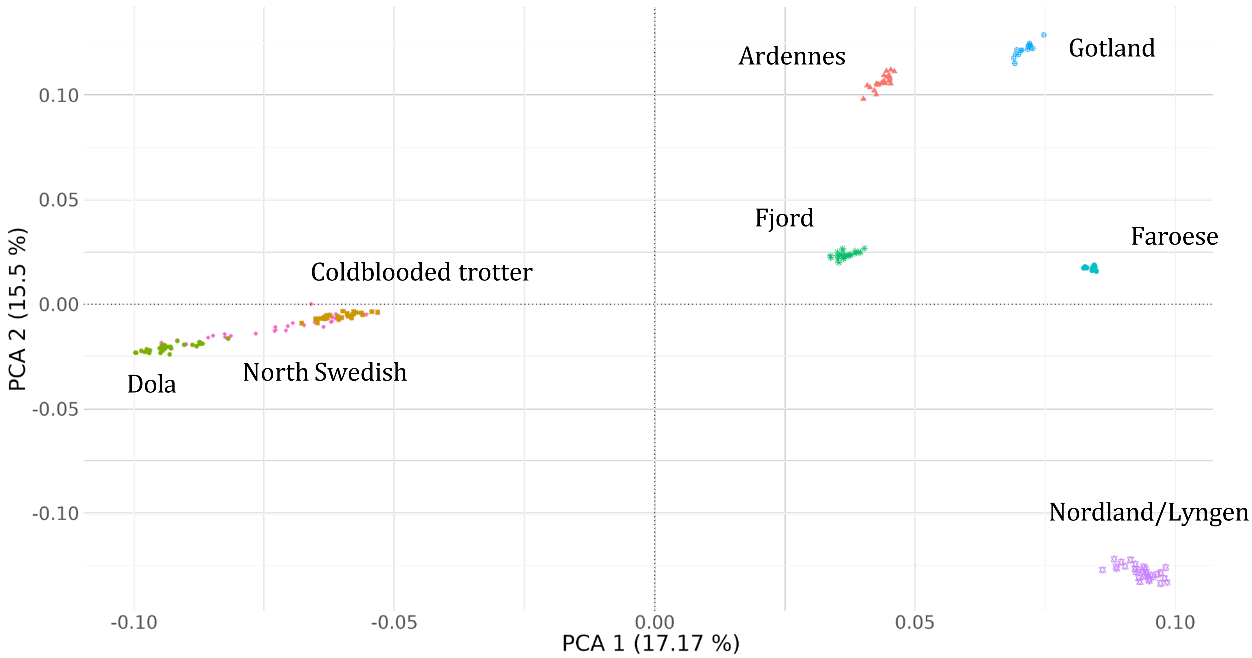

Figure 1. Principal component analysis (PCA), based on SNPs from the data mapped on the EquCab_Finn genome.

Runs of Homozygosity (ROH) and inbreeding (F_{ROH})

The Faroese horse exhibited very high genomic inbreeding coefficients, with an average of more than 30% in the EquCab_Finn mapped data (Table 1). This was not surprising as this critically endangered breed has been through a very narrow bottleneck and expanded from 5 animals to a small current population size (Kettunen et al., 2022). The Swedish Ardennes breed that has decreased dramatically in numbers from a very large population around the mid-20th century (Carlström et al, 1946), had the lowest mean inbreeding of close to 13%. The North Swedish horse showed on average a lower level of inbreeding compared to the Dola horse, and included one individual with close to zero F_{ROH}, that was found to be a crossbred between North Swedish and Dola horse, which is allowed in the studbook. The other studied breeds had relatively high inbreeding coefficients, ranging from 14% to 21% in the EquCab_Finn mapped data. In the EquCab3.0 mapped data, somewhat less genetic variation was found, causing the inbreeding coefficients to increase by 1-2 percentage units per breed. This could be explained by the closer relatedness of the native Nordic Finnhorse with the Nordic breeds studied, compared with the Thoroughbred horse that the EquCab3.0 is based on.

Heterozygosity (HET)

The mean observed heterozygosity was generally rather low and varied between 12% and 16% for the different breeds (Table 1). The highest level was seen for the Swedish Ardennes and the lowest for the Faroese horse. The Dole horse had the second lowest mean observed heterozygosity. For all breeds, the mean observed heterozygosity was higher than the expected heterozygosity, indicating active breeding measures to avoid matings of close relatives.

Table 1. Estimated mean F_{ROH} and observed and expected heterozygosity per breed in EquCab3.0 and EquCab_Finn mapped data.

	EquCab3.0	**EquCab_Finn**		
Breed	**Mean (and range) of F_{ROH} (%)**	**Mean (and range) of F_{ROH} (%)**	**Exp. Het. (%)**	**Mean Obs. Het. (%)**
Swedish Ardennes	13.8 (10.0-18.2)	12.5 (8.9-16.8)	14.9	16.0
Dola horse	23.0 (7.7-30.4)	21.2 (6.4-27.8)	12.7	13.4
North Swedish horse	15.6 (1.2-24.9)	14.3 (0.6-23.1)	14.3	14.8
Coldblooded Trotter	15.5 (2.6-20.3)	14.0 (2.1-19.0)	14.3	15.3
Gotland Pony	22.5 (17.1-25.9)	20.5 (15.6-24.4)	13.6	14.6
Fjord horse	14.7 (3.8-18.2)	13.3 (2.9-17.0)	14.4	15.2
Nordland/Lyngen horse	20.5 (14.7-25.3)	18.6 (12.6-23.5)	13.8	14.7
Faroese horse	33.4 (28.7-38.1)	31.4 (27.2-36.6)	10.6	11.8

Historical effective population size

The historical effective population size (Ne) 13 generations ago was highest for the North Swedish horse followed by the Swedish Ardennes, and lowest for the Faroese horse and second lowest for the Gotland pony. The two latter breeds are known to have been through severe genetic bottlenecks. These measures were only slightly higher using the EquCab_Finn mapped data instead of the EquCab3.0 data. It should be noted that estimates of historical Ne can be influenced by drastic changes in population size, however (Adepoju et al., 2024).

Identity by state and identity by descent, and the fraction of pairwise comparisons

The identity by state (IBS) values were similar between the breeds, whereas the estimates for identity by descent (IBD) and the fraction of pairwise comparisons (FPC) varied more. The IBD followed the general pattern of lower values for Swedish Ardennes (15%) and highest for the Faroese pony (40%) in the EquCab_Finn mapped data, and with only slightly higher values using the EquCab3.0 data. The highest FPC was found for the Dola horse and the Coldblooded trotters, indicating a lack of horses unrelated to the other individuals.

Private Alleles

Breed specific private alleles were observed throughout the genome but more commonly in certain chromosome regions, most pronounced in ECA20:31-32Mb, ECA20:33-34Mb, ECA20:61-62Mb, ECA25:5-6Mb, and ECA26:41-42Mb. Genetic annotations analysis showed a large number of biological pathways and gene involvements, including for example metabolic processes, immune regulation, neurological functions, and cardiovascular development. A few subnetworks showed association with inherited disorders detected in other species, and those would be of interest to study further.

Cross-population haplotype scores (XP-EHH) and shared haplotypes

Signatures of selection based on XP-EHH were detected for each breed, and 481 regions were found among the top 0.001% haplotypes across the eight breeds. Such regions include for example genes with known associations with muscle growth and body size for the Ardennes, but also for example genes with potential associations with postnatal lethality in Gotland ponies, and other regions which need further inspection. Recently, some closely related Gotland pony foals were born with suspected genetic defects (Picchi, 2024), highlighting the need to compare their genetic variants with suspected regions from the present study.

No observed haplotypes larger than one Mbp were found to be shared among the breeds, indicating a lack of recent common ancestry.

Genetic load

Relatively small breed differences were found in genetic load, except for the Faroese horse. However, breeds with lower average F_{ROH} tended to have a higher number of potentially deleterious alleles (GERP score > 4) segregating in the population (Swedish Ardennes, North Swedish horse, Coldblooded trotter and Norwegian Fjord horse) than those with higher F_{ROH} (Dola horse, Faroese horse, Gotland pony and Nordland/Lyngen horse). For most breeds, potentially harmful variants with high GERP scores occurred in low allele frequency, however. The Faroese horse tended to have more homozygotes for potentially harmful variants at high frequency, which is likely due to its special situation with very high inbreeding levels and small population size after a severe genetic bottleneck.

Structural variation

In total, 35,426 structural variants (SV) were found in all breeds combined before filtering out variants with rare (<1% or <5%) minor alleles. Most detected SVs were translocations or deletions (Table 2).

PCA-plots based on structural variants showed a similar pattern as those based on SNPs, with the Coldblooded trotter and its two founding breeds, Dola horse and North Swedish horse, clustering together. The type of SV and the filtering applied influenced the pattern seen in similarities between the other breeds, but in most cases the Nordland/Lyngen horses were separated from the other breeds.

When keeping only rare SVs with minor allele frequency <0.05 (Figure 2), then the Swedish Ardennes, Gotland pony, and to an even higher degree, Faroese horse was more distinguished from the rest of the breeds, as well as showing more within-breed variation. Among such breed-specific rare SVs, there may be some potentially deleterious variants of interest to study further.

Table 2. Number of identified structural variants (SVs) in the whole data set by subtypes, and after filtration for minor allele frequency (MAF)

Structural variants subtypes	N variants	N after removing SVs with MAF<1%	N after removing SVs with MAF<5%
Translocations	16 244	12 649	9 746
Deletions	10 523	8 579	6 254
Insertions	764	764	755
Inversions	3 124	2 525	1 618
Duplications	4 771	3 687	2 443

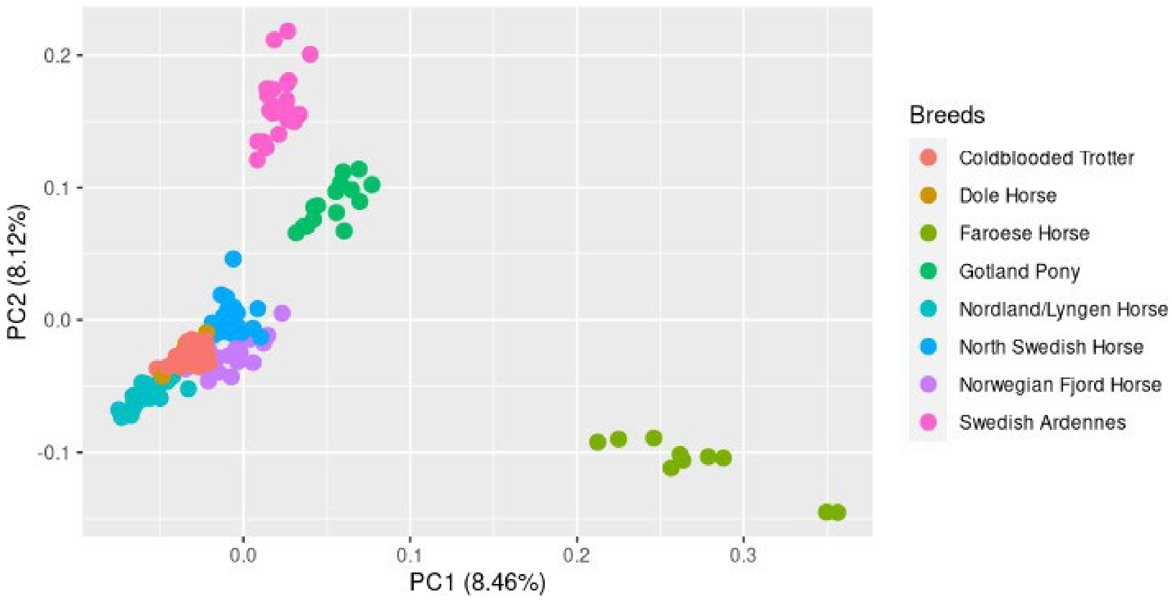

Figure 2. Principal component analysis (PCA) based on structural variants in the form of deletions in the data mapped on the EquCab_Finn genome filtered to keep only those with minor allele frequency <0.05 (i.e. rare structural variants).

Limitations of the study

The horses in this study were chosen to represent the current potential breeding animals in their breeds but are only a sample of all individuals in the populations. Thus, we cannot claim to have captured all available genetic variation. Whereas the sequencing gave a higher read depth than originally planned for, there are still regions in individuals that are less well covered, which may influence for example the number of potentially harmful variants found. The chosen settings are also known to be important for estimation of measures such as ROHs, and we therefore tested different settings for the analysis. Despite these potential limitations, the results do agree well with the known breed history, and the results from various analyses support each other, strengthening our confidence in the conclusions drawn.

Conclusions

Low genetic diversity was seen in all breeds, and confirmed in several different analyses, highlighting the importance of maintaining genetic health and diversity to ensure a sustainable future for the breeds. The standing genetic variation in breeds with a larger historical population size, especially the Swedish Ardennes, was larger than for breeds that have been through known severe genetic bottlenecks, such as the Faroese horse and the Gotland Pony. It also seems that the Dola horse has lost more genetic variation than the North Swedish Horse. A generally higher observed than expected heterozygosity shows that measures have been taken to avoid mating closely related individuals in the studied breeds.

The general pattern of relatedness within and across breeds was similar regardless of method and reference genome used, with clear distinctions between the different breeds except for a closer relatedness between the three breeds that are known to share a common background: Coldblooded trotter, North Swedish horse, and Dola horse. Somewhat more genetic diversity was revealed using the more closely related Finnhorse reference genome instead of that from a Thoroughbred horse, though in general the differences between the two genomes were small to moderate.

Several genomic regions in the breeds show signatures of selection, opening up new research questions for the future. All studied breeds carried genetic variants with a potential to be harmful, though most of these were in low frequency. The highest frequency of homozygotes for potentially harmful alleles was seen in the breed with the highest mean genomic inbreeding (F_{ROH}) coefficient.

The results from this study demonstrate the value of using genomic data when monitoring the genetic diversity and genetic load in local breeds. The data will be useful in future studies to learn more about specific genetic regions, variants, and structural variation in horses.

Relevance for the practical horse sector incl. recommendations

The various analyses showed reduced genetic diversity in all the studied breeds, even though the situation does not appear critical at present for most of them. Potentially harmful genetic variants were found in all breeds, though in most cases in low frequency. The results highlight the importance of preserving genetic variation and avoiding a rapid inbreeding rate to ensure the future genetic health of these breeds. The ongoing decline in the annual number of covered mares in several of the breeds is clearly worrying, and it poses a real threat to the future survival of the breeds, considering the already reduced genetic diversity.

Information and advice have been given to breed associations and breeders during the project, but there is a further need for this. It is also clear that especially some of the Swedish breed associations need access to, and help to utilize, modern tools to routinely estimate and monitor relatedness and inbreeding and to plan suitable matings. In Norway, such resources are available through Stiftelsen Norsk Hestesenter, and resources are also available through the trotting organizations for Coldblooded trotters. Support for the use of cryopreservation of genetic material from the native horse breeds, as well as further studies of how well such methods work for these breeds, would also be desirable.

The whole genome sequence data produced in our project is more comprehensive than could be fully investigated within this three-year project. It will be possible to use it in future research. We especially see possibilities to look further into regions found in other projects to be of interest for traits such as fertility or health, or suspected of harboring potential genetic defects in native Nordic horse breeds, and to some extent, such investigations have already begun. The data can also be used for further studies of genetic similarities with other horse breeds, increasing our knowledge of their historical background as well as providing information about suitable breeds for crossbreeding, should such need arise to keep the breeds healthy in the future.

References

Special resources utilized in the project

Breed associations and several veterinarians contributed to the sample collection. Samples were also provided by the AS Biobank in Hamar, Norway, and the Faroese horse Association. The Swedish Trotting Association, Svenska Hästavelsförbundet, and the Norwegian Equine Association provided pedigree information.

The Animal Genetics Laboratory at SLU prepared the DNA samples. SNP&SEQ Technology Platform in Uppsala that performed the sequencing is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory and supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation.

The computations and data handling were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725.

The Orion High-Performance Computing Center (OHPCC) at the Norwegian University of Life Sciences (NMBU) also provided computational resources.

The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB89544 (https://www.ebi.ac.uk/ena/browser/view/PRJEB89544), and is currently under processing and quality control there. The code scripts for the analysis are published on the GitHub page: https://github.com/NattyandMinnie.

Scientific references

Adepoju, D., Ohlsson, I.J., Klingström, T., Rius-Vilarrasa, E., Johansson, A.M., Johnsson, M. Population history of Swedish cattle breeds: estimates and model checking. 2024. https://doi.org/10.1101/2024.10.03.616479.

Barbato, M., Orozco-terWengel, P-, Tapio, M., et al. 2015. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. Front. Genet. 6, https://doi.org/10.3389/fgene.2015.00109

Carlström O, Aaby-Ericsson A, Wilhelmsson F. 1946. Betänkande med förslag till åtgärder för främjande av ridhästaveln m.m. Statens offentliga utredningar 1946:45. Stockholm: Kungliga boktryckeriet P.A. Nordstedt & Söner. https://lagen.nu/sou/1946:45 (accessed June, 2025)

Chang C.C., Chow C.C., Tellier, et al. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets, GigaScience 4, s13742–015–0047–8. https://doi.org/10.1186/s13742-015-0047-8

Cooper, G.M., Stone E.A., Asimenos, G., et al., 2005. Distribution and intensity of constraint in mammalian genomic sequence, Genome Res. 15:901-913 https://doi.org/10.1101/gr.3577405

Danecek, P., Auton, A., Abecasis, G., et al. 2011. The variant call format and VCFtools, Bioinformatics 27:2156–2158, https://doi.org/10.1093/bioinformatics/btr330

Danecek P, Bonfield JK, Liddle J, et al. 2021. Twelve years of SAMtools and BCFtools. GigaScience 2021;10. https://doi.org/10.1093/gigascience/giab008

Delaneau, O., Marchini, J. & Zagury, JF. 2012. A linear complexity phasing method for thousands of genomes. Nat Methods 9, 179–181. https://doi.org/10.1038/nmeth.1785

Evangelista, J.E., Xie, Z., Marino, G.B., et al. 2023. Enrichr-KG: bridging enrichment analysis across multiple libraries, Nucleic Acids Res. 51, W168–W179, https://doi.org/10.1093/nar/gkad393

Jäderkvist Fegraeus K, Velie BD, Axelsson J, et al. 2018. A potential regulatory region near the EDN3 gene may control both harness racing performance and coat color variation in horses. Physiol Rep. 6:e13700. https://doi.org/10.14814/phy2.13700

Kalbfleisch, T. S., Rice, E. S., DePriest, M. S., et al. 2018. EquCab3, an updated reference genome for the domestic horse. bioRxiv. Preprint posted April 25, 2018. https://doi.org/10.1101/306928

Kierkegaard, L. S., Groeneveld, L. F., Kettunen, A., et al. (2020). The status and need for characterization of Nordic animal genetic resources. Acta Agric. Scand. A — Anim. Sci., 69:2–24. https://doi.org/10.1080/09064702.2020.1722216

Kettunen, A., Joensen, S.K., Berg, P. 2022). Optimum contribution selection (OCS) analyses prompted successful conservation actions for Faroese horse population. Genet. Resources, 3:59–67. https://www.dx.doi.org/10.46265/genresj.KKXV5870

McLaren, W., Gil, L., Hunt, S.E. et al. 2016. The Ensembl Variant Effect Predictor. Genome Biol. 17, 122. https://doi.org/10.1186/s13059-016-0974-4

Picchi, F. 2024. Genomic Analysis of a Hypermobility Syndrome in Gotland Ponies. MSc thesis, Swedish University of Agricultural Sciences, Dept. of Animal Biosciences, https://stud.epsilon.slu.se/20663/

Pokharel, K., Weldenegodguad, M., Reilas, T., et al. 2024. EquCab_Finn: A new reference genome assembly for the domestic horse, Finnhorse. Anim. Genet., 55:766–771. https://doi.org/10.1111/age.13463

EquCab_Finn: A new reference genome assembly for the domestic horse, Finnhorse. Animal Genetics, 55(5), 766–771. https://doi.org/10.1111/age.13463

Purcell S, Neale B, Todd-Brown K, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559-75. https://doi.org/10.1086/519795

Utsunomiya, Y.T., Milanesi, M., Utsunomiya, A.T.H., et al. 2016. GHap: an R package for genome-wide haplotyping, Bioinformatics 32:2861–2862, https://doi.org/10.1093/bioinformatics/btw356

Zachary A Szpiech, selscan 2.0: scanning for sweeps in unphased data, Bioinformatics 40, btae006, https://doi.org/10.1093/bioinformatics/btae006

Part 3: Result dissemination

Scientific publications, *published*	Smogeli, N.A., Shutava, I., Joensen, S.K., Kjetså, M., Kantanen, J., Pokharel, K., Selle, T., Mikko, S., Eriksson, S., Berg, P. 2026. Characterizing Genetic Diversity Within and Between Native Nordic Horse Breeds Utilizing and Comparing the EquCab3 and EquCab_Finn Reference Genomes. *Accepted for publication in Genetic Resources February 2026.* *https://doi.org/10.46265/genresj.TXWX7641*
Scientific publications, *submitted*	
Scientific publications, *manuscript*	Smogeli, N.A., Shutava, I., Joensen, S.K., Kjetså, M., Kantanen, J., Pokharel, K., Selle, T., Mikko, S., Eriksson, S., Berg, P. Genomic Characterisation as a Tool Towards Sustainable Breeding of Native Nordic Horse Breeds. *The manuscript is almost complete, but we have waited for the first paper to be accepted so that we can refer to that for breed descriptions.*
	Shutava, I. Mikko, S., Smogeli, N.A., Joensen, S.K., M. Kjetså, Kantanen, J., Pokharel, K., Selle, T., Berg, P., Eriksson, S. Structural variations in Native Nordic Horse breeds: Insights from PCA and Variant Profiling *(Very early stage, results are in place but the full manuscript is not yet drafted, it is possible that it will only be a short communication or similar).*
Conference publications/ presentations	S. Eriksson, S. Mikko, M. Kjetså, T. Selle, I. Shutava, C. Brekke, Kallsoy Joensen, S. & P. Berg. 2022. Genomic characterization of native Nordic horse breeds. Report 2021/2022 of Dept. Animal Breeding and Genetics, Swedish University of Agricultural Sciences. Report no. 154, 22.
	Presentation by S. Eriksson at the Research days for Dept. of Animal Breeding and Genetics 10 November 2022.
	Almaas Smogeli, N. Shutava, I. Joensen, S.K., Mikko, S., Eriksson, S. Berg, P. 2024. Genomic Dive into Nordic Horse Breeds. The 75th Ann. Meeting European Federation of Animal Science, 1-5 September, Florence, Italy. p. 582. Oral pres.
	Joensen, S.K., Heller, R., Pečnerová, P., Mikko, S., Eriksson, S. 2024. Whole genome investigation of the Faroese horse. NordGen Farm Animals Conference 2024, February 7-8, Uppsala, Sweden. Poster (prize winning).
Other publications, *media etc(.*	Written information (1 page): "Nytt forskningsprojekt om genetisk diversitet hos nordiska hästraser" about the project provided to all Swedish breed associations in January/February 2022 for their own use in informing horse breeders.
	Eriksson, S. 2022. Hur genetiskt lika är svenska ardennerhästar? ArdennerNytt Nr 1 2022.
	Information on the SLU web page 2022: https://old.slu.se/fakulteter/vh/forskning/forskningsprojekt/hast/genomisk-kartlaggning-som-redskap-for-hallbart-avelsarbete-i-inhemska-nordiska-hastraser2/ (SLU web is under reconstruction)
	Online information on the SLU web page 2022: https://old.slu.se/en/faculties/vh/research/forskningsprojekt/hast/genomic-characterization-as-a-tool-towards-sustainable-breeding-of-nordic-native-horse-breeds2/ (SLU web is under reconstruction)
	Online information on NordGen webpage September 2022: https://www.nordgen.org/sv/nyheter/stor-kartlaggning-av-nordens-hastraser-ska-bidra-till-ett-battre-bevarandearbete/
	Popular scientific information from NordGen on Facebook 2022: Stor kartläggning av Nordens hästraser ska bidra till ett bättre bevarandearbete, including a pressrelease.
	T.Selle. 2022. Forskning; in Nøkkeltal om dei nasjonale hesterasane - rapport for 2022, page 28. https://www.nhest.no/forskning-nasjonale-hesteraser.480576.no.html
	NordGen newsletter 06.11.2024: Information about project and link to webinar (https://www.nordgen.org/sv/projekt/nanohorse/om-projektet/). The newsletter has at least 2000 subscribers.

Oral communication, *to horse sector, students etc.*	Online presentation by S. Eriksson at a meeting about breeding Coldblooded trotters 20 January 2022.
	Presentation by M. Kjetså about project for entire NordGen staff 21 January 2022
	Online information meeting with representatives from the four Swedish breed associations and two breeding organizations (SH, ST) 4 February 2022. (S. Eriksson and S. Mikko)
	Online lecture by S. Eriksson about genetic diversity and inbreeding, including information about the project, for Swedish (and some Norwegian) breeders of Coldblooded trotters 10 February 2022.
	Presentation by M. Kjetså about project for the Nordic Animal Genetic Resources Council, 27.04.2022
	Pre-recorded online information by S. Eriksson for breeders of Coldblooded trotters streamed via Facebook on 17 September 2022.
	Online lecture by M. Kjetså on conservation and management of small populations, including information about the project for Finnish students in Animal Science at Helsinki University, 28 Sept. 2022
	Lectures by S. Mikko and S. Eriksson about genetic defects and inbreeding at "How to breed a champion" arranged by ASVT and SWB for breeders in Upplands Väsby 30 October 2022.
	Short seminar by S. Eriksson at the Research Day for Dept. of Animal Breeding and Genetics 10 November 2022.
	Presentation at the yearly meeting for the North Swedish Horse Association by C. Hamilton (using information and PPt from the project) 13 November 2022.
	Lecture by Peer Berg about inbreeding at Järvsöfaks uppfödarseminarium, Årjäng, 19 November 2022.
	Presentations by several project participants and discussions at NordGen Native Horse Network meeting in Alnarp, 18-19th November 2022.
	Presentation by S. Eriksson for Norwegian horse breeders at webinar arranged by the Norsk Hestesenter 30 November 2022.
	EuroHorse Paddock show "SLU: De nordiska hästraserna" by S. Mikko at Gothenburg Horse Show 25 February 2023.
	Project presentation by Peer Berg for the Norwegian committee for the breeding and conservation of native horses, an advisory committee to the Norwegian Ministry for Food and Agriculture. Gardermoen 22 March 2023.
	Lecture/Presentation by S. Eriksson at the yearly meeting of the Svenska Hästavelsförbundet in Skara on 17 June 2023.
	Information about the project to the Faroese horse Association by Signa Kallsoy Joensen during summer 2023.
	Visit sample collection, discussions, and breeding advice by S. Mikko at Lojsta Hed, Gotland 31 July 2023.
	Presentation by S. Eriksson at breeding conference by the Swedish Trotting Association, Örsundsbro, 9 September 2024
	NordGen webinar by most project participants for breeders and owners of Swedish and Norwegian native horse breeds 24 October 2024 with over 322 registrations and 200 participants.
	Presentations "Genetisk variation och avel i små populationer" by S. Eriksson for breeders of Swedish Ardennes 16 November in Vessigebro and 23 November in Askersund 2024.
	Presentation for Nordlands/Lyngshest breeders: Breeding in small populations. 30. Nov. 2024. Peer Berg
	Presentation "Storleken har betydelse – om populationsstorlek, genetisk variation och hållbart avelsarbete" by S. Eriksson at a webinar arranged by Svenska Hästavelsförbundet 18 May 2025.
	Presentation " Aktuell forskning om avel och genetik - varmblod och kallblod "by S. Eriksson and P. Berglund at seminar for breeders of trotters, at Menhammar, Ekerö, 1 Nov. 2026.
	Presentation"Genetisk variation och avel i små populationer" by S. Eriksson for breeders of North Swedish Horses in Askersund 15 Nov. 2025.

	Presentation"Genetisk variation och avel i små populationer" by S. Eriksson for breeders of Gotland Ponies, remotely, 22 Nov. 2025.
	Presentation"Genetisk variation och avel i små populationer" by S. Eriksson for breeders of North Swedish Horses in Norrbotten, online, 18 Dec. 2025.
	Presentation"Genetisk variation och avel i små populationer, med fokus på gotlandsruss" by S. Eriksson at webinar arranged by Svenska Hästavelsförbundet, online, 16 Febr. 2026.
Student theses	Signa Kallsoy Joensen, supervised by Rasmus Heller and Patrícia Pečnerová, with assistance from Susanne Eriksson and Sofia Mikko, 2024, The Genomic Diversity and Population Structure of the Faroese Horse – The first ever whole-genome study
	Fanny Margrethe Karlsen Vognild, supervised by Peer Berg and Nathalie Almaas Smogeli, 2025, Runs of Homozygosity and Heterozygosity to identify historical inbreeding and signatures of selection in eight Native Nordic horse breeds, MSc thesis (30 ECTS). https://nmbu.brage.unit.no/nmbu-xmlui/handle/11250/3198663
Other	