Analysis of the history and population structure of the Icelandic horse using pedigree data and DNA analyses

Gunnfríður Elín Hreiðarsdóttir[1], Thorvaldur Árnason[1], Vilhjálmur Svansson[2], and Jón Hallsteinn Hallsson[1*]

[1]*Department of Land and Animal Resources, Agricultural University of Iceland, Keldnaholt, 112 Reykjavik, Iceland.*

[2]*Institute for Experimental Pathology, University of Iceland, Keldur, 112 Reykjavik, Iceland.*
[*]*jonhal@lbhi.is* (*Corresponding Author*)

ABSTRACT

The Icelandic horse breed is believed to have been founded on the basis of horses brought to Iceland eleven hundred years ago. It has in past decades gone from close to obsolete to enjoying considerable popularity with about 250,000 horses in at least 19 countries. To characterize the population structure pedigree data for over 300,000 individuals was analysed, calculating parameters such as effective population size, effective number of founders and ancestors, as well as inbreeding. DNA sequence analysis was also used to assess population structure and genetic diversity. The Icelandic horse population is genetically uniform despite differences in the use of genetic material between areas, has high levels of heterozygosity, numerous D-loop haplotypes, and an effective population size above one hundred horses. Despite long-term isolation, fluctuations in population size, and a selective breeding program the Icelandic horse population is well-off in regards to genetic diversity.

Keywords: Pedigree data, genetic diversity, microsatellites, D-loop.

YFIRLIT

Greining á sögu og stofngerð íslenska hestsins með ætternisgögnum og DNA greiningu

Íslenski hesturinn er talinn afkomandi hrossa sem bárust til Íslands með landnámsmönnum fyrir 1100 árum. Íslenski hesturinn hefur á undanförnum árum notið aukinna vinsælda og telur nú um 250 000 hross í um 19 löndum. Til að lýsa byggingu stofnsins voru ætternisgögn 300 000 hrossa greind og reiknaðar breytur sem gefa mynd af stöðu stofnsins, til dæmis virk stofnstærð, virkur fjöldi forfeðra og innræktun. Hvatbera DNA var raðgreint og örtungl greind til að varpa ljósi á stöðu stofnsins og byggingu hans. Niðurstöður sýna að þrátt fyrir að nýting erfðaefnis milli svæða sé mismunandi þá er stofninn erfðafræðilega einsleitur, býr yfir töluverðri arfblendni, geymir umtalsverðan fjölda D-lykkju setraða og virk stofnstærð er vel yfir eitt hundrað einstaklingar. Þrátt fyrir langa einangrun, sveiflur í stofnstærð og skipulagt ræktunarstarf stendur stofninn því vel hvað varðar erfðafjölbreytni.

INTRODUCTION

The genetic diversity of the domestic horse (*Equus ferus caballus*) is diminishing, through both loss of breeds and within breed diversity. Breeding programs and the use of reproductive technology can increase the use of relatively few 'superior' individuals, usually select stallions, which in turn lowers the effective population size. For conservation of individual horse breeds and responsible use of genetic resources detailed analysis of genetic diversity is a necessary first step.

Little is known for certain about the origin of the Icelandic horse. It is generally believed that

Icelandic livestock descend from animals brought to the country with Norse settlers around 1100 years ago (Adalsteinsson 1981), a view substantiated through molecular analysis (Kantanen et al. 2000, Bjørnstad et al. 2003, Tapio et al. 2005). Historical records state that horses were brought to the country during the first centuries of settlement but the exact origin and number of animals is not recorded (Björnsson & Sveinsson 2004). It is commonly believed that the Icelandic domestic breeds have remained isolated ever since settlement, although this remains to be demonstrated conclusively (Adalsteinsson 1981, Björnsson & Sveinsson 2004).

Throughout its history the Icelandic horse has been used primarily for riding and its most important attributes were considered strength and endurance along with a stable temperament and soft gaits. The importance of the horse as a means of transportation over often difficult terrain combined with at times limited fodder led to a strong self-sufficient horse, showing remarkable endurance (Björnsson & Sveinsson 2004). Isolation and late industrialization led to a breed that kept its importance into the twentieth century long after machines had replaced horses in most other western countries.

All the Icelandic domestic breeds have fluctuated dramatically in size in the past with some becoming critically small in size. The Icelandic goat population went down to only 62 individuals in 1885 and currently counts fewer than a thousand individuals (Baldursdottir et al. 2012). The Icelandic cattle, however, remains the main production breed in traditional agriculture and retains a large population size (Asbjarnardottir et al. 2010). Records show that, contrast, the size of the Icelandic horse population has fluctuated quite dramatically. For example, the consequences of the Laki volcanic eruption in 1784-1785 reduced the number of horses to 8,000-9,000 horses. Sixty years later the population had grown to about 40,000 horses (Icelandic Historical Statistics 1997, Statistics Iceland 2013). Despite known bottlenecks recent results suggest that the effective population size of the Icelandic horse has remained stable over the last 150 years (Campana et al. 2012).

Although the oldest pedigree records for the Icelandic horse date back to the late 19th century when breeding societies were first established it was not until the middle of the 20th century that breeding according to breeding goals became common. The breed has now become a popular riding horse, both within Iceland and abroad, with over two hundred thousand horses worldwide and breeding societies in 19 countries. The number of horses outside Iceland now exceeds the number of horses in Iceland, with large numbers in Germany (34,000 horses), Denmark (33,000 horses), and Sweden (20,000 horses). To guard against the accidental importation of disease, reimporting of livestock to Iceland is not allowed under Icelandic law (Act No. 54/1990) and exported horses are therefore not allowed to be returned, resulting in one way traffic of genetic material. This limits the options of Icelandic breeders, but the long term effects of this on the Icelandic breeding effort are unknown.

Analysis of population structure and changes in genetic diversity can be achieved using either pedigree data or molecular markers. Although the available pedigree data for the Icelandic horse includes only a short period of the breed's long history it has been used to shed light on current levels of inbreeding (Kristjánsson 2005). Molecular analysis has also shed light on the Icelandic horse breed; a study investigating the relationship between Norwegian and Mongolian horses placed the Icelandic horse closer to the Mongolian horse than Norwegian horse breeds (Bjørnstad et al. 2003), while another study found evidence for two branches within the Icelandic population, one between the Shetland and Nordland breeds and another between the Fjord and Coldblooded trotter, which might be evidence for population structure within the Icelandic breed (Bjørnstad & Røed 2001).

The aim of the present study is to shed light on the current status of the Icelandic horse breed, population structure, and breeding history through the use of both pedigree data and molecular analysis. Such analysis is important for closed populations such as the Icelandic horse as it can aid in the management and maintenance of the population's genetic diversity.

MATERIALS AND METHODS

Pedigree data and analysis

Pedigree data for all Icelandic horses born in Iceland and registered before spring 2009 was obtained from the WorldFengur pedigree database (www.worldfengur.com) resulting in a pedigree file with a worldwide collection of more than 300,000 horses. ID numbers of individuals and their parents were obtained with each number containing necessary information for pedigree analysis: country of birth (two letters), year of birth (four digits), sex (1 = male and 2 = female), area code (two digits), and an individual identifier (three digits). Both horses born in Iceland (Icelandic) and abroad (non-Icelandic) were used. Iceland was divided into eight areas based on historical distinction and/or geography that could have given rise to distinct breeding lines: Reykjanes (REY, 4% of current population), Vesturland (VST, 14%), Húnavatnssýslur (HUN, 9%), Skagafjörður (SKG, 21%), Norðausturland (NAU, 10%), Austurland (AUS, 5%), Hornafjörður (HRN, 2%), and Suðurland (SDR, 35%). Area codes from WorldFengur were grouped as follows: 10-26 as REY, 30-38 as VST, 55-56 as HUN, 50, 51, 57, and 58 as SKG, 60-67 as NAU, 70-76 as AUS, 77 as HRN, and 80-88 as SDR. Area codes 45, 46, 47, and 49 (individuals from the Vestfirðir area (VFI)) were excluded due to the low number of horses registered.

Pedigree analysis revealed the total number of records, total number of horses born in Iceland, and number of horses recorded each year. To estimate the proportion of horses registered in WorldFengur each year the data was compared to the total number of horses alive in the autumn of each year, available from official records where track is kept of all livestock in Iceland. Programs from PEDIG (Boichard 2002), a Fortran package for pedigree analysis, were used for analysis of pedigree data using males and females born in 2007 as reference populations with the number of most important ancestors for the run set as 1000. Similar results were observed for both sexes; female foals were consequently chosen as a reference population for subsequent analysis. The average number of founders and ancestors as well as number of equivalent complete generations traced were calculated using the program *ngen*. The program *vanrad* was used to estimate levels of inbreeding with runs tracing three, five, and ten generations back. The number of founders (f), the effective number of founders (f_e), that is, founders that with an equal genetic contribution, would produce the same genetic diversity in the reference population, and the effective number of ancestors (f_a) were calculated using the program *prob_orig*, and the effective number of founder genomes (f_g) calculated using *segreg*.

Gene flow analysis was carried out using Fortran 95 and Fortran 77 programs relying on the method of Kennedy and Trus (1993). The same pedigree file as before was used with females born in 2007 as a reference population and the pedigree traced 15 generations back. Founders are defined as individuals with both parents unknown and for horses with one known parent a dummy parent was created using the birth location of the offspring and setting the birth year as the birth year of the offspring minus 4. For horses with unknown year of birth, birth year was set as 1900. Different runs were made with various maximum birth years of founders in order to detect possible differences in gene flow over time.

Mitochondrial D-loop sequencing and data analysis

For mitochondrial DNA D-loop sequencing fifty samples were selected from Icelandic

horses with no pedigree information to represent horses that are not a part of the main breeding population and 393 samples from registered horses from the following areas: 15 from Reykjanes (REY), 45 from Vesturland (VST), 49 from Húnavatnssýslur (HUN), 72 from Skagafjörður (SKG), 36 from Norðausturland (NAU), 23 from Austurland (AUS), 10 from Hornafjörður (HRN), 143 from Suðurland (SDR). WorldFengur was used to obtain pedigree information about sampled individuals, horses were assigned to an area according to the origin of their mother, the proportion of samples from each area was kept as close as possible to the total number of horses in a given area, and only a single offspring was sampled from each mare.

DNA was isolated from either whole blood or buffy coat using MasterPure DNA isolation kit (www.epicenter.com) or Puregene Blood Core Kit B (www.qiagen.com) according to the manufacturer's instructions and samples diluted to a concentration of 20 ng/µL. A 30 µL polymerase chain reaction (PCR) was performed using primers Forw1 (5'-ACCATCAACACCCAAAGCTG-3', nucleotide positions (nps) 15 425-15 444) and Rev1 (5'-GCATTTTCAGTGCCTTGCTT-3', nps 23-42) creating a 1 278 bp fragment covering the D-loop (nps as in reference sequence X79547 (Xu & Arnason 1994)). The PCR conditions were as follows: pre-denaturation at 94°C for 2 min; 39 cycles of 95°C for 30 sec, 65°C for 30 sec, 72°C for 1 min followed by 72°C for 8 min and a 4°C hold. Amplification was confirmed using gel electrophoreses on a 1.6% agarose gel. Primers Forw1 and Seq1-Rev (5′-ATGGCCCTGAAGAAAGAACC-3′, nps 15 867-15 848) were used for sequencing of the upstream region.

Consensus sequences were aligned and truncated to a length of 247 base pairs (nps 15 494-15 740, as in Jansen et al. (2002)) in order to maximize the number of sequences available for analysis. A total of 443 sequences from the Icelandic breed sequenced here and eight sequences previously deposited in GenBank with accession numbers AJ413717-413723 (Jansen et al. 2002) and AF072988 (unpublished results) were analysed. The sequences are available in GenBank (KJ741404-KJ741846, www.ncbi.nlm.nih.gov/genbank). The nomenclature of Jansen et al. (2002) was used to define haplotypes and the differentiation of haplotype composition between areas was carried out by way of an exact test of population differentiation (Rousset 2008). For analysis of changes in haplotype frequencies over time pedigrees were traced back for individuals with known D-loop haplotypes and the frequencies compared for the years 1947 and 2000.

Only horse populations with 20 or more sequenced individuals were used for comparison, resulting in a total of 493 previously published sequences: AF014413-AF014415 (Kim et al., 1999), AF132568-AF132594 (Bowling et al. 2000), AF354425-AF354441 (Yang et al. 2002), AF465996-AF466005, AF466007-AF466008 (Mirol et al. 2002), AJ413615-AJ413648, AJ413650-AJ413657, AJ413682-AJ413693 (Jansen et al. 2002), AF516502-AF5160504 (Luis et al. 2006), AY293975-AY293991, AY805645-AY805648, AY525091-AY525096 (Lopes et al. 2005), AY462441-AY462445 (Cozzi et al. 2004), AY519907-AY519913 (Royo et al. 2005), DQ327856-DQ327967, DQ327986-DQ328001, DQ328038-DQ328057 (McGahern et al. 2006), D14991, D23665, D23666 (Ishida et al. 1994), AF481305-AF481323 (Hill et al. 2002), AY246174-AY246185, AY246201-AY246208, AY246219-AY246224, AY246242-AY246247, AY246266-AY246271 (unpubl. results), AF072989-AF072993 and AF072996 (unpubl. results), and AF056071 (Kim et al. 1999).

Arlequin 3.1 (Excoffier et al. 2005) was used for analysis of mtDNA D-loop sequences, including calculations of gene diversity (GD) and three different theta (θ) indices; theta(k) based on the expected number of alleles (Ewens 1972), theta(S) based on number of segregating sites (Watterson 1975), and

theta(π) based on the mean number of pairwise differences (Tajima 1983).

DNA isolation, microsatellite markers and data analysis

DNA was extracted from 4822 nasal swabs or buffy coat blood samples from Icelandic horses and analysed by MATIS (www.matis.is) using eleven microsatellite markers: ASB2 and ASB17 (Breen et al. 1997), ASB23 (Lear et al. 1999), HMS1, HMS2, HMS3 and HMS6 (Guérin et al. 1994), UM011 (Swinburne et al. 2000), AHT4 (Binns et al. 1995), CA425 (Eggleston-Stott et al. 1997), and VHL20 (van Haeringen et al. 1994). The samples were grouped as previously with the addition of unregistered horses used for blood collection (BLD) and Icelandic horses born in Norway (NOR).

Standard diversity indices for microsatellite loci were calculated using GenAlEx 6.5 (Peakall & Smouse 2012) including number of alleles for individual markers (N_A), major allele frequencies (MAF), effective number of alleles (EN_A), mean number of alleles (MNA), number of genotypes (N_g), observed (H_O) and expected (HE) heterozygosity, fixation index (F), and Chi-Square tests for deviations from Hardy-Weinberg equilibrium (HWE) for each locus per area.

Inbreeding coefficients were estimated using PowerMarker 3.25 (Liu & Muse 2005) using two different approaches, the expectation-maximization (EM) algorithm to find the maximum likelihood estimation (MLE) of F_{IS} and method of moments (mom) (Weir & Hill 2002). In both cases 95% confidence intervals (CI) were calculated based on bootstrapping using 1000 random permutations. ONeSAMP 1.0, which relies on linkage disequilibrium, (Tallmon et al. 2008) and LDNe, which uses approximate Bayesian computation to estimate N_e, (Waples & Do 2008) were used to estimate the effective population size (N_e) based on the microsatellite data. Using BOTTLENECK 1.2.02 (Piry et al. 1999) two methods were employed to test for a recent reduction in effective population size (N_e). Firstly, a test for significant excess or deficiency in gene diversity with two statistical tests (sign test and Wilcoxon test) applied under three models of microsatellite evolution, the infinite-allele model (IAM), the stepwise mutation model (SMM), and the two-phase model of mutation (TPM; with 70% SMM and 30% variation). The expected distribution was found using 1000 iterations. Secondly, a graphical method which assumes that in non-bottlenecked populations alleles at low frequency (0.0-0.1) are more abundant than alleles at more intermediate frequencies (Luikart et al. 1998). A mode-shift distribution of allele frequencies is expected in recently bottlenecked populations.

Multilocus 1.3 (Agapow & Burt 2001) was used to test the significance of multilocus linkage disequilibrium among alleles (Brown et al. 1980, Haubold et al. 1998) by calculating the Index of Association (I_A) and rbarD, a version of I_A modified to remove the dependency on the number of loci (Maynard Smith et al. 1993). These two parameters give information on whether two individuals sharing an allele at one locus are more likely than not to share an allele at another locus. The I_A and rbarD values are equal to zero when there is no linkage and increase as the linkage disequilibrium increases. The null hypothesis of no linkage was tested by comparing observed values to 1000 randomized datasets.

Principal co-ordinate analysis (PCoA) was carried out using GenAlEx 6.5 based on Nei's genetic distance. PowerMarker 3.25 was used to calculate Nei et al.'s D_A distance (Nei et al. 1983) and Cavalli-Sforza chord distance (Cavalli-Sforza & Edwards 1967) with bootstrap analysis (Felsenstein 1985) performed with 1000 random permutations. The corresponding Neighbor-Joining trees were displayed using MEGA 5 (Tamura et al. 2011). BAPS 5.2 (Corander et al. 2004, 2008) was used to study the clustering of pre-defined groups of individuals (areas defined as previously described) using fixed K-clustering for K = 2-10.

RESULTS

Pedigree analysis

At the beginning of the 20th century the number of horses in Iceland was about 40,000 having risen from 8,683 horses in 1784 (Figure 1A). The earliest pedigree records date back to the late 19th century with twelve horses in the pedigree born before 1900, the oldest being the mare Gráskjóna from Gullberastöðum born in 1860 (IS1860235700). Before 1950 on average 1% of foals born each year were recorded in the database. From the mid-20th century an increasing number of foals were registered, with almost complete registration since 2004 (Figure 1B). A clear difference is seen in registration of males and females with males lagging until 1981 when a steady rise in male registration is seen (Figure 1C). The estimated equivalent number of known generations (data not shown) and the average number of ancestors grows with time as the number of registered foals increases (Figure 1D).

At the end of 2008 approximately 77,000 horses were registered in Iceland with a non-uniform distribution between areas, with the largest number of horses in Suðurland (25% of population) and Skagafjörður (21% of population). Gene flow analysis using three different years as a birth year of

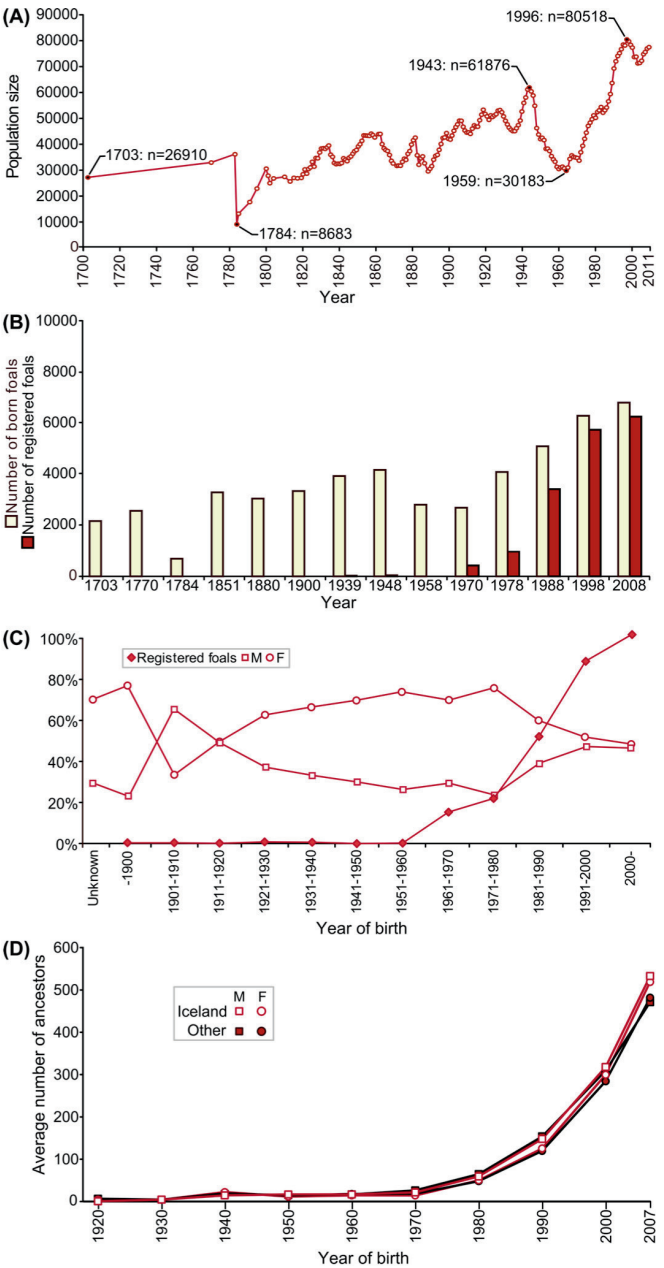

Figure 1. Changes in the Icelandic horse population. (A) Size of the Icelandic horse population in 1703 to 2008, (B) number of foals born and number of foals registered, (C) percentage of registered foals and ratio between males and females, and (D) the average number of ancestors for horses born in Iceland and abroad (other) for males (M) and females (F).

founders shows a considerable difference in the contribution of different areas to the total population (Table 1). Skagafjörður and Hornafjörður stand out at all time points with much use of their own genetic material; for example in the Hornafjörður area in 1920 and 1950 more than 90% of the genetic material used was from within the area (underlined in Table 1). This trend remains relatively constant in Skagafjörður but a change is seen in Hornafjörður in 2000 when the use of their own material was down to 84.8% from 90.4%. Areas such as Reykjanes and Austurland show a different trend where more than half of the genetic material originated from outside the region. Looking at the contribution of genetic material to other areas it is clear that Skagafjörður has been the most influential area, contributing, for example, 24.9% of the genetic input in Reykjanes and 21.6% in Austurland. The contribution from Skagafjörður into most other areas remained constant throughout the century (Table 1).

Table 1. Gene flow between areas for the Icelandic horse population at three time points based on pedigree data. Results shown as percentage (%) with number of horses in each region in parentheses and the use of genetic material within regions underlined.

		DONOR REGION							
	YEAR 1920	REY	VST	HUN	SKG	NAU	AUS	HRN	SDR
ACCEPTOR REGION	REY (427)	36.5	7.4	6.0	28.6	1.3	0.5	7.6	12.2
	VST (2004)	1.1	61.7	5.6	15.5	1.5	0.1	9.5	4.9
	HUN (1972)	0.9	3.6	74.1	12.7	1.8	0.2	2.8	3.9
	SKG (3434)	0.5	2.6	1.8	87.2	2.3	0.1	3.2	2.2
	NAU (1008)	0.4	5.8	5.1	21.7	54.1	0.8	7.6	4.5
	AUS (345)	1.3	8.5	5.8	16.1	2.6	45.2	15.7	4.9
	HRN (256)	0.2	1.7	0.7	4.8	0.51	0.7	90.1	1.3
	SDR (3284)	1.5	5.6	5.0	20.4	1.61	0.2	6.0	59.7
	YEAR 1950	REY	VST	HUN	SKG	NAU	AUS	HRN	SDR
ACCEPTOR REGION	REY (461)	37.6	7.4	6.2	28.3	1.5	0.6	7.0	11.5
	VST (2124)	1.0	62.4	5.6	15.4	1.6	0.2	9.0	4.7
	HUN (2068)	0.91	3.5	73.5	13.5	2.0	0.2	2.7	3.8
	SKG (3598)	0.54	2.5	2.4	86.8	2.4	0.1	3.1	2.2
	NAU (1084)	0.4	5.5	5.2	22.5	54.0	0.7	7.1	4.6
	AUS (376)	2.0	7.8	6.0	17.7	2.5	44.2	15.1	4.8
	HRN (282)	0.2	1.5	0.7	4.8	0.5	0.7	90.4	1.2
	SDR (3499)	1.5	5.3	5.5	20.5	2.0	0.2	5.7	59.4
	YEAR 2000	REY	VST	HUN	SKG	NAU	AUS	HRN	SDR
ACCEPTOR REGION	REY (945)	42.4	5.4	3.8	24.9	2.7	0.5	4.1	16.2
	VST (3419)	2.2	58.3	3.9	17.5	2.0	0.9	5.8	9.8
	HUN (3713)	1.6	3.3	64.5	16.8	2.5	0.5	1.8	9.0
	SKG (5980)	0.6	2.3	1.8	86.0	2.6	0.1	2.0	4.6
	NAU (2275)	1.2	3.3	3.1	23.9	56.3	0.6	3.8	7.7
	AUS (641)	1.6	5.7	3.7	21.6	3.5	46.2	9.4	8.5
	HRN (422)	0.4	1.3	0.5	7.1	1.2	1.1	83.8	4.5
	SDR (6466)	2.6	3.8	3.3	21.7	2.5	0.5	3.4	62.2

Founders, ancestors, and inbreeding

Pedigree analysis reveals more diversity in the pedigree of the non-Icelandic reference population (mares born abroad) than the Icelandic reference population, with the effective number of founders (f_e) estimated as 151.0 individuals for the Icelandic population compared to 159.1 for the non-Icelandic population. The effective number of ancestors (f_a) and the effective number of founder genomes (f_g) are estimated as 42.5 and 26.7 for the Icelandic population compared to 50.2 and 31.9 for the non-Icelandic population, respectively (Table 2).

The number of inbred individuals and mean inbreeding were estimated by tracing the pedigree three, five, and ten generations back. Two reference populations were used, the Icelandic population with 30,203 individuals and the non-Icelandic population with 26,289 individuals. Inbreeding decreases as the pedigree is traced further back suggesting a time-dependent increase in population-wide inbreeding. Tracing three generations back the mean inbreeding was slightly lower for the Icelandic population than the non-Icelandic

Table 2. Summary statistics of pedigree analysis for the Icelandic horse population including effective population size (N_e), number of founders (f), effective number of founders (f_e), effective number of ancestors (f_a), and effective number of founder genomes (f_g).

	Iceland	Abroad
Size of reference population	3212	2215
Effective population size (N_e)	113.4	175.9
Number of founders (f)	4547	3875
Effective number of founders (f_e)	151.0	159.1
Effective number of ancestors (f_a)	42.5	50.2
Effective number of founder genomes (f_g)	26.7	31.9

Table 3. Differences in levels of inbreeding between mares born in 2007 in Iceland and abroad.

		Number of inbred individuals	Mean inbreeding	Max inbreeding
Iceland	3 generations	3206	7.0	43.75
	5 generations	10371	3.5	43.75
	10 generations	13484	3.0	43.75
	All	13484	3.0	43.75
Abroad	3 generations	2272	7.7	43.75
	5 generations	8360	3.3	43.75
	10 generations	11299	2.7	43.75
	All	11300	2.7	43.75

population, with a maximum inbreeding of 43.8% for both populations (Table 3).

Of the ten most influential genetic contributors eight appear in both the Icelandic and non-Icelandic populations (Table 4). The two most influential horses for both populations were the stallion Hrafn from Holts-múli with a marginal contribution of 0.106, followed by the only mare on the list, Síða from Sauðárkrókur, with a marginal contribution of 0.066. Five of the most influential ancestors in the Icelandic population and six in the non-Icelandic population originated from Skagafjörður (area codes 50, 51, 57, and 58), underlining the area's influence in Icelandic horse breeding (Table 4).

Mitochondrial D-loop variation in the Icelandic horse population

Analysis of mitochondrial D-loop sequences from 442 Icelandic horses revealed 24 polymorphic sites giving rise to 26 haplotypes (Figures 2 and 3) assigned to eight sub-haplogroups (Table 5). Only in Húnavatnssýslur and Hornafjörður was the haplotype composition significantly different from the other areas (exact test, $P < 0.05$). Although the differences in haplotype frequencies between areas were only slight, comparing haplotype frequencies within areas at two time points, namely for the years 1947 and 2000, showed considerable changes over time (data not shown). D-loop sequencing results estimate genetic diversity (GD) as 0.88, theta(k) based on the expected number of alleles as 5.85, theta(S) based on number of segregating sites as 3.59, and theta(π) based on the mean number of pair-wise difference as 5.71.

Variation in microsatellite markers and population structure

All microsatellite markers were polymorphic, with numbers of alleles ranging from 6 (*HMS1* and *HMS6*) to 17 (*ASB17*) and the mean number of alleles (MNA) of 10.2 (Table 6). Four markers (*ASB17*, *ASB2*, *HMS3*, and *UM011*) showed deviations from Hardy-Weinberg equilibrium (HWE) in the combined dataset ($P < 0.01$) but no marker showed deviations from HWE in all areas. The effective number of alleles (EN_A) ranged from 2.6 (*HMS1*) to 7.0 (*ASB2*), in all cases lower than the values for the observed number of alleles. The mean observed heterozygosity (H_O) was 0.73 and the mean expected heterozygosity (H_E) was 0.75 (Table 6). Analysis of allele frequencies based on geographical regions revealed relatively high levels of heterozygosity in all areas (Table 7). Fourteen private alleles were found: nine alleles in Suðurland, three in Skagafjörður, and one in both Norðausturland and Hornafjörður. This large difference was somewhat reduced when the samples were rarefied to lessen the effect of the unequal number of horses in different areas (data not shown).

The average inbreeding based on microsatellite data was 3.7% (95% CI=0.001-0.092)

and the effective population size (N_e) was estimated as 214.6 (95% CI=196.8-233.4). Observed heterozygosity was higher than expected under conditions of equilibrium ($H_{obs}>H_{eq}$) under both the two-phased mutation model (TPM) and the infinite alleles model (IAM) ($p < 0.001$), suggesting a recent population bottleneck. Results under the stepwise mutation model (SMM) were statistically non-significant and the modeshift analysis of allele frequencies revealed a typical L-mode shape indicative of a non-bottlenecked population (data not shown).

No evidence of population structure was seen suggesting that the Icelandic

Table 4. Total and marginal genetic contributions of the ten most influential ancestors for the reference populations of mares born in Iceland and abroad in 2007.

	Id		TC	MC	Progeny
Iceland	IS1968157460	Hrafn from Holtsmúli	0.106	0.106	348
	IS1952257001	Síða from Sauðárkrókur	0.066	0.066	10
	IS1986186055	Orri from Þúfa	0.066	0.039	276
	IS1974158602	Ófeigur from Flugumýri	0.039	0.039	197
	IS1943156690	Fengur from Eiríksstaðir	0.039	0.037	16
	IS1967186102	Þáttur from Kirkjubær	0.025	0.025	151
	IS1973135980	Gáski from Hofsstaðir	0.022	0.022	116
	IS1957158589	Hörður from Kolkuós	0.018	0.018	99
	IS1941177415	Nökkvi from Hólmur	0.034	0.016	59
	IS1961158588	Rauður from Kolkuós	0.015	0.015	67
Abroad	IS1968157460	Hrafn from Holtsmúli	0.088	0.088	239
	IS1952257001	Síða from Sauðárkrókur	0.065	0.065	10
	IS1974158602	Ófeigur from Flugumýri	0.039	0.039	141
	IS1943156690	Fengur from Eiríksstaðir	0.037	0.037	18
	IS1957158589	Hörður from Kolkuós	0.034	0.034	134
	IS1967186102	Þáttur from Kirkjubær	0.027	0.027	137
	IS1986186055	Orri from Þúfa	0.038	0.021	145
	IS1937177180	Skuggi from Bjarnanesi	0.027	0.018	35
	IS1961158588	Rauður from Kolkuós	0.017	0.017	62
	IS1947157008	Goði from Sauðárkrókur	0.018	0.017	12

The sex of reference population is set as female (2). Id: Individual identification number with the year of birth underlined. TC: Total genetic contribution. MC: Marginal genetic contribution.

Table 5. Mitochondrial D-loop haplotypes found in the Icelandic horse population compared to reference sequence.

	15 494	15 495	15 496	15 521	15 524	15 534	15 585	15 595	15 596	15 597	15 600	15 601	15 602	15 603	15 604	15 617	15 626	15 649	15 650	15 659	15 666	15 703	15 718	15 720	15 726	15 737	15 740	No. of individuals
Ref.	**T**	**T**	**A**	**G**	**T**	**C**	**G**	**A**	**A**	**A**	**G**	**T**	**C**	**T**	**G**	**T**	**A**	**A**	**A**	**T**	**G**	**T**	**C**	**G**	**G**	**T**	**A**	**0**
D108	C	C	G	.	.	T	A	.	.	.	.	.	T	C	.	.	.	G	.	.	.	.	.	A	.	.	.	29
D124	C	C	G	.	.	T	A	.	.	.	.	.	T	.	.	.	.	G	.	.	.	.	.	A	.	.	.	1
D201	C	C	.	.	N	.	A	.	.	.	.	.	.	C	.	.	.	G	.	.	.	.	.	A	.	.	.	1
D205	C	C	G	.	.	T	A	.	.	.	.	.	.	C	.	.	.	G	.	.	.	.	.	A	.	.	.	65
D208	C	C	G	.	.	T	A	.	.	.	.	.	.	C	.	.	G	G	.	.	.	.	.	A	.	.	.	9
D210	C	C	G	.	.	T	A	G	.	.	.	.	.	C	.	.	.	G	.	.	.	.	.	A	.	.	.	1
D211	C	C	G	.	.	T	.	.	.	.	.	.	.	C	.	.	.	G	.	.	.	.	.	A	.	.	.	1
E102	.	C	.	A	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	A	.	C	.	3
E105	.	C	.	A	.	.	.	.	.	.	A	.	T	.	.	.	.	.	.	.	.	.	.	A	.	C	.	10
E106	.	C	.	A	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.	.	A	.	C	.	11
E108	.	C	.	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	A	.	C	.	36
E109	.	C	.	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	T	A	.	C	.	11
E110	.	C	.	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.	G	.	.	.	.	A	.	C	.	2
E114	.	C	.	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	T	A	.	.	.	1
C204	.	C	.	.	.	.	A	.	.	.	.	C	T	.	.	.	.	.	.	.	.	C	.	A	.	.	.	4
C205	.	C	.	.	.	.	A	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.	.	A	.	.	.	88
C212	.	C	.	.	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.	.	A	.	.	.	6
C104	.	C	.	.	.	.	A	.	.	.	.	.	T	.	.	C	.	.	.	C	.	.	.	A	.	.	.	2
C106	.	C	.	.	.	.	.	.	.	.	.	.	T	.	.	C	.	.	.	C	.	.	.	A	.	.	.	41
C110	.	C	.	.	.	.	.	.	.	G	.	.	T	.	.	C	.	.	.	C	.	.	.	A	.	.	.	1
A306	.	C	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	A	.	.	A	.	.	.	17
A625	.	C	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	A	.	.	.	1
F202	.	C	.	.	.	T	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	C	.	A	A	.	G	6
F203	.	C	.	.	.	.	A	.	.	.	.	.	T	.	A	.	.	.	.	.	.	C	.	A	A	.	G	11
F204	.	C	.	.	.	.	.	.	G	.	.	.	T	.	A	.	.	.	.	.	.	C	.	A	A	.	G	1
F207	.	C	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	C	.	A	A	.	G	92
																												451

Haplotypes defined according to Jansen et al. (2002). The reference sequence is X79547 (Xu & Arnason 1994). Only nucleotides deviating from the reference sequence are shown.

horse population is a genetically homogenous group. Analysis aimed at assigning pre-defined 'sub-populations' (areas) to fixed numbers of clusters (K = 2 - 10) first separated the Vestfirðir area (VFI), next Hornafjörður (HRN) and then the blood collecting horses (BLD) (Figure 4A). Similar results were seen using principal co-ordinate analysis (PCoA) based on Nei's genetic distance (Figure 4B), where 67.8% of the variation can be explained on the first two axes and 85.6% by adding the third axis. A consensus Neighbour-Joining tree based on Nei et al.'s D_A distance had low bootstrapping values for all branches, underlining conflict between individual trees and highlighting the lack of population structure (Figure 4C). The Cavalli-Sforza chord distance yielded a topologically identical tree (isomorphic) (data not shown). I_A and rbarD values were statistically non-significant, the null hypothesis being that of no linkage (data not shown).

Table 6. Diversity indices for the Icelandic horse breed by microsatellite marker.

Marker	N_{obs} (avail)	N_A	MAF	EN_A (SE)	H_O (SE)	H_E (SE)	F (SE)
ASB17	4577 (95.7%)	16	0.318	4.53 (0.09)	0.79 (0.01)	0.78 (0.00)	-0.02 (0.01)
ASB23	4713 (98.5%)	8	0.250	4.81 (0.15)	0.81 (0.02)	0.79 (0.01)	-0.02 (0.01)
HMS6	4558 (95.3%)	6	0.435	3.49 (0.06)	0.69 (0.02)	0.71 (0.01)	0.03 (0.02)
HMS3	4477 (93.6%)	8	0.313	4.43 (0.20)	0.60 (0.04)	0.77 (0.02)	0.21 (0.07)
HMS2	4694 (98.1%)	10	0.480	3.27 (0.13)	0.69 (0.03)	0.69 (0.01)	0.00 (0.03)
ASB2	4194 (87.6%)	17	0.214	7.03 (0.13)	0.83 (0.02)	0.86 (0.00)	0.04 (0.02)
HMS1	4694 (98.1%)	6	0.462	2.60 (0.07)	0.58 (0.03)	0.61 (0.01)	0.05 (0.03)
UM011	4669 (97.6%)	15	0.280	5.00 (0.21)	0.81 (0.01)	0.80 (0.01)	-0.02 (0.02)
AHT4	4781 (99.9%)	8	0.417	4.40 (0.13)	0.74 (0.03)	0.77 (0.01)	0.04 (0.04)
CA425	4380 (91.5%)	8	0.449	3.94 (0.18)	0.75 (0.02)	0.74 (0.01)	-0.02 (0.02)
VHL20	4753 (99.3%)	10	0.281	4.43 (0.16)	0.76 (0.02)	0.77 (0.01)	0.01 (0.02)
Mean	4590 (95.9%)	10.2	0.355	4.36 (0.11)	0.73 (0.01)	0.75 (0.01)	0.03 (0.01)

N_{obs}: Number of observations with availability in parentheses. N_A: Number of alleles. MAF: Major allele frequency. EN_A = Effective number of alleles. H_O: Observed heterozygosity. H_E: Expected heterozygosity. F = Fixation index. Standard error in parentheses (SE).

DISCUSSION

Analysis of both pedigree data and genetic material can be used to estimate the status of genetic diversity and the future potential of a population, with microsatellites and mitochondrial haplotypes as commonly used markers. Although pedigree records are often the most appropriate and readily available source of information researchers often depend on molecular markers for analysis of genetic diversity, especially in populations where adequate pedigree data is lacking. Here both types of data were used to assess population structure and genetic diversity for the Icelandic horse population.

According to historical records the Icelandic horse breed has experienced dramatic fluctuations in population size. Although such fluctuations can be expected to affect genetic diversity, recent results suggest that the effective population size and "genetic profile" of the Icelandic horse have remained stable for the last 150 years (Campana et al. 2012). The results presented here on genetic bottlenecks are not clear, as the results differed based on the methods and models selected for analysis.

In the early days of active breeding the Icelandic horse breed consisted of about 40,000 horses with little or no tradition for recording pedigree data. By the end of the 19th century breeding according to official breeding goals and registration of horses had become a common practice. For the first half of the 20th century most of the horses recorded were of significance in breeding or ancestors of horses that later became influential. This is likely explained by the fact that co-ordinated registration did not start until the late 20th century with the construction of the WorldFengur pedigree database. Earlier records were less formal, often personal documentation of individual breeders, which might explain the emphasis on registering only pedigree data of select individuals. In 1950-2000 a rapid rise occurred in the level of registration, from around 1% to close to 100%. This level of pedigree recording gives breeders an excellent

tool to monitor future changes in the breed's status and can be used to minimize the negative effects of active breeding.

Gene flow analysis clearly underlines the influence of Skagafjörður in contributing breeding material to other parts of the country. Only Hornafjörður shares the extensive use of local horses seen in Skagafjörður, but this is most likely explained by the long-term isolation of the area. Areas such as Reykjanes and Austurland show a different trend with diverse use of genetic material and little influence in other regions. The great importance of Skagafjörður is evident also when considering the total and

Table 7. Diversity indices for microsatellite markers by geographical regions.

Region	N (avail)	N_g	H_O (SE)	H_E (SE)	F (SE)
REY	235 (96.9%)	26.7	0.73 (0.03)	0.75 (0.02)	0.03 (0.03)
VST	850 (94.9%)	32.0	0.75 (0.03)	0.77 (0.02)	0.03 (0.03)
VFI	13 (92.3%)	7.5	0.74 (0.04)	0.71 (0.02)	-0.04 (0.06)
HUN	370 (96.1%)	27.8	0.74 (0.02)	0.75 (0.02)	0.02 (0.02)
SKG	755 (96.5%)	30.3	0.73 (0.03)	0.76 (0.02)	0.03 (0.03)
NAU	434 (95.9%)	28.1	0.72 (0.03)	0.76 (0.02)	0.04 (0.03)
AUS	167 (98.3%)	22.8	0.75 (0.03)	0.76 (0.02)	0.02 (0.03)
HRN	35 (97.4%)	14.1	0.72 (0.04)	0.76 (0.02)	0.04 (0.05)
SDR	1874 (97.2%)	35.4	0.74 (0.03)	0.76 (0.02)	0.03 (0.03)
BLD	20 (96.8%)	10.9	0.67 (0.04)	0.75 (0.03)	0.12 (0.04)
NOR	69 (92.4%)	17.6	0.78 (0.03)	0.76 (0.02)	-0.02 (0.02)
ALL	4822 (96.4%)	40.4	0.73 (0.01)	0.75 (0.01)	0.03 (0.01)

N: Number of samples with availability in parentheses. N_g: Number of genotypes observed. H_O: Observed heterozygosity. H_E: Expected heterozygosity. F = Fixation index. Standard error in parentheses (SE).

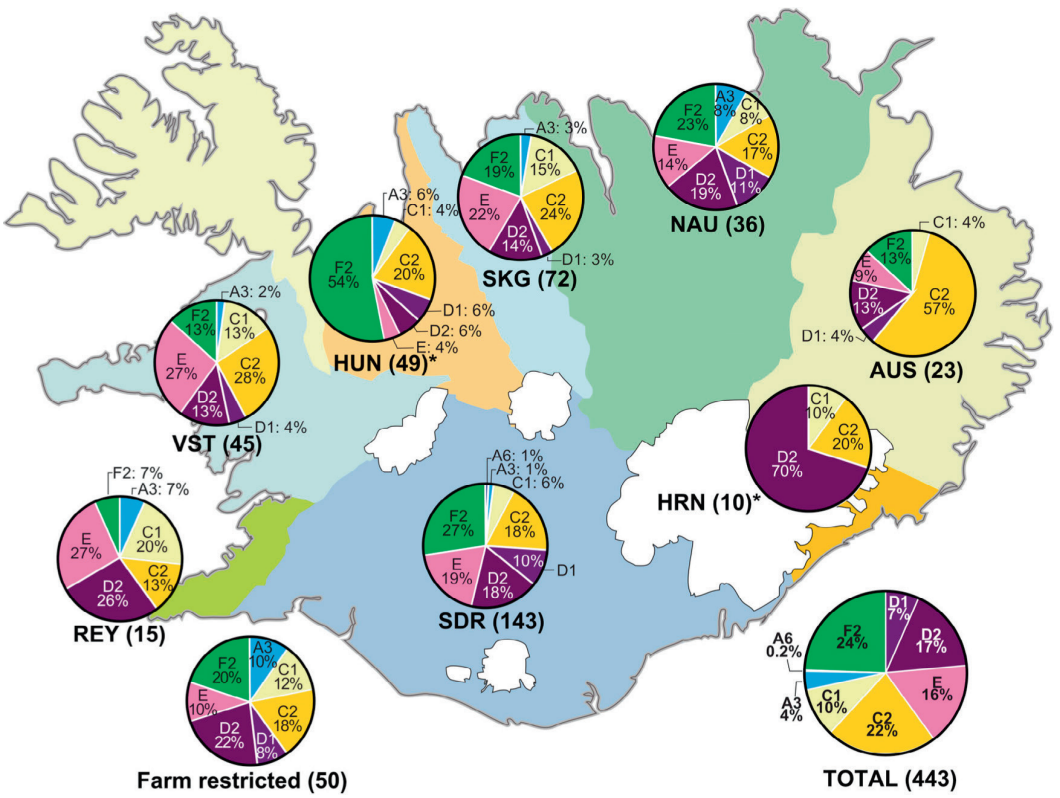

Figure 2. Mitochondrial D-loop haplotype frequencies in different geographical regions of Iceland. The number of horses analysed from each region is in parentheses. Only two regions (marked with an asterisk) had frequencies of haplotypes significantly different from all other ($P < 0.05$).

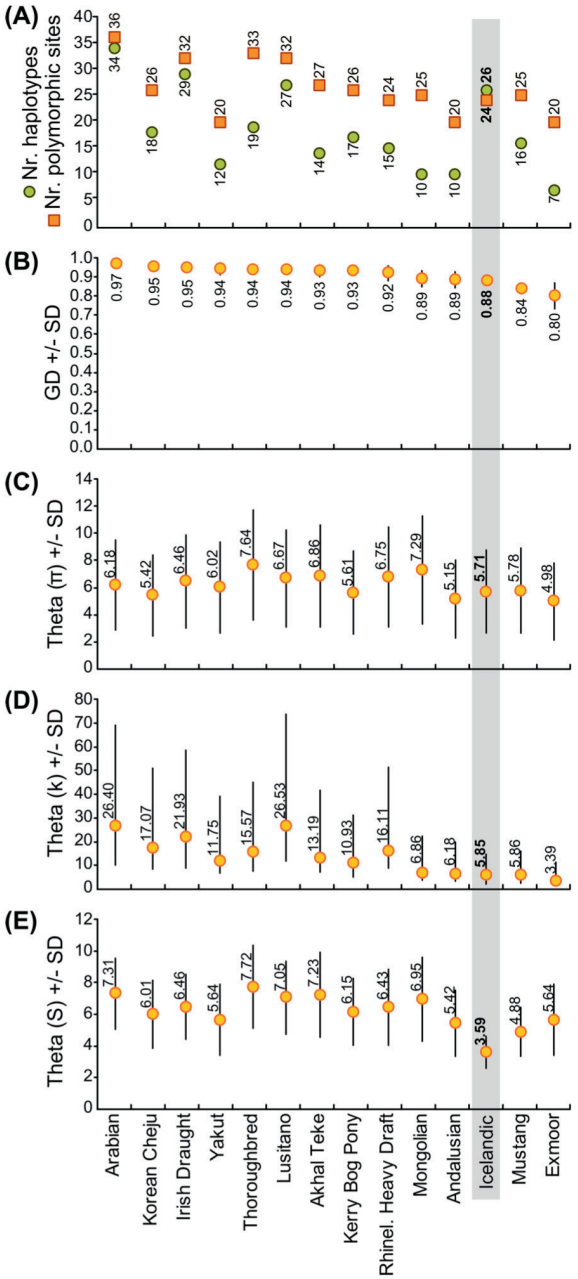

Figure 3. Diversity indices based on the results of mtDNA sequencing for the Icelandic horse breed compared to thirteen other breeds. (A) number of D-loop haplotypes and polymorphic sites, (B) genetic diversity, (C) theta (π) based on number of pairwise differences, (D) theta (k) based on the expected number of alleles, and (E) theta (S) based on number of segregating sites. Breeds are arranged from left to right according to levels of genetic diversity. All estimates shown ± standard deviation (SD).

marginal contributions of the most influential ancestors, where five (Icelandic population) and six (non-Icelandic population) horses in the top ten originated from Skagafjörður, although Skagafjörður constitutes only about 21% of the Icelandic horse population today.

Due to lack of pedigree information the number of founders is relatively large (4547). The effective population size (N_e) based on pedigree data for individuals born in Iceland is lower than for individuals born abroad (113 compared to 175), despite the higher number of founders (*f*). This is not unexpected as the total number of founders gives only limited information on diversity since it does not take into account the unequal contribution of individuals to subsequent generations. The effective number of founders (f_e) is therefore a more suitable index as it accounts for the unequal contribution of founders, but also here the observed values are higher for the non-Icelandic population, 151.0 compared to 159.1. For the Icelandic population the effective number of ancestors (f_a) is 42.5 individuals compared to 50.2 for the non-Icelandic population. The ratio f_e/f_a, which is 3.6 for the Icelandic population, can reveal the importance of bottlenecks on population

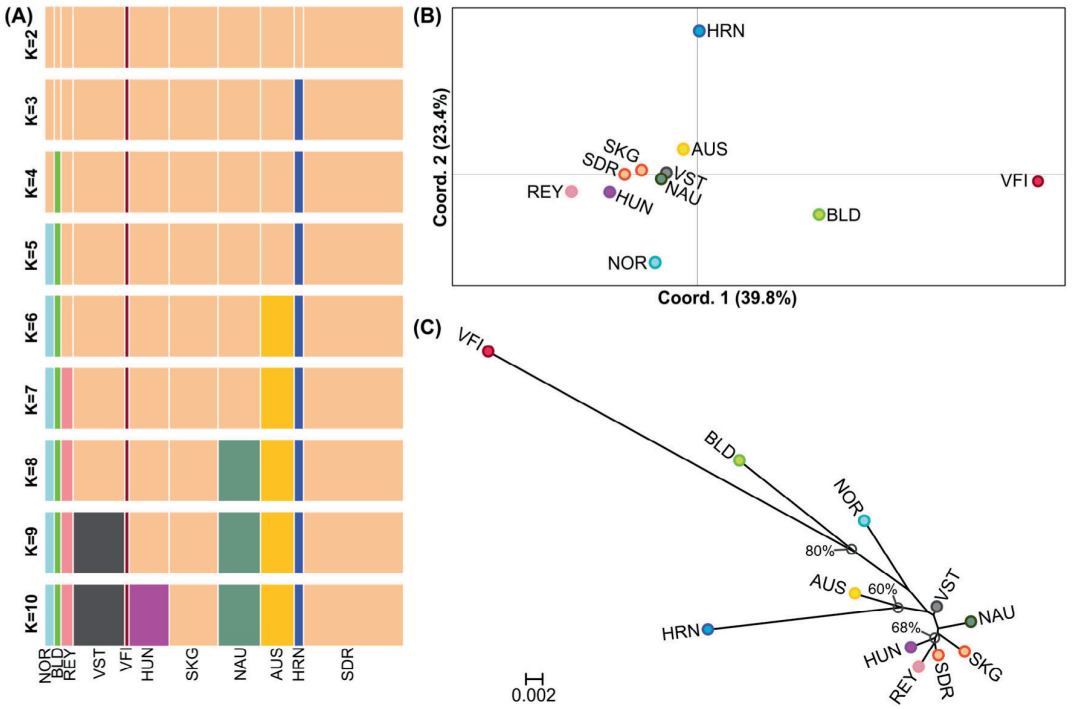

Figure 4. Results of microsatellite analysis showing (A) the results of Bayesian analysis of population structure where samples from different areas are grouped into a fixed number of clusters (K = 2-10), (B) the results of Principle Co-ordinates Analysis (PCoA) based on Nei's genetic distance, and (C) a Neighbour-Joining tree based on Nei's genetic distance (D_A). Bootstrap values below 60% are omitted from the tree.

diversity, when the ratio is close to one all individuals within the population are contributing equally to the breeding population (Boichard et al. 1997, Hamann and Distl 2008). Inbreeding is on the rise, although that might partly be explained by an increase in pedigree completeness and lack of pedigree information in the early part of the 19th century.

It is interesting to compare results for the effective number of founders, ancestors, and founder genomes for the Icelandic horse breed to results for the German Paint Horse which has almost the same number of founders as the Icelandic horse (f=4,216) but much higher values for f_e (560.8), f_a (207.8), and f_g (138.7) (Siderits et al. 2013). Despite unfavourable comparison to the German Paint Horse and the Hanoverian (f_e=244.9 and f_a=77.7) (Hamann and Distl 2008) the Icelandic horse fares better compared to breeds such as the Spanish Arab (f_e=38.6 and f_a=19.0)(Cervantes et al. 2008), Andalusian (f_e=39.6 and f_a=16.5)(Valera et al. 2005), Lipizzan (f_e=48.2 and f_a=26.2)(Zechner et al. 2002), and Austrian Noriker (f_e=117.2 and f_a =29.3)(Druml et al. 2009).

D-loop analysis shows that many haplogroups are absent from the Icelandic population, including haplogroups B and G and several of the A sub-haplogroups. Several factors might explain this observation, including founder effect, the long term isolation of the population, and fluctuations in population size. The Icelandic horse has lower levels of genetic diversity and theta values than many much smaller breeds such as the Irish Draught (est. population size of 1000 individuals), Lusitano

(6000 individuals), and Kerry Bog Pony (200 individuals) (Figure 3B-E).

Analysis of microsatellite data indicates that the Icelandic horse population is homogenous despite differences seen in the use of genetic material between areas. The results seen here for H_O (0.73) and H_E (0.75) are similar to what was previously reported for the Icelandic horse by Aberle et al. (2004). Comparing results for the Icelandic horse to results for other Nordic breeds reveals considerably higher levels of observed heterozygosity in the Icelandic horse (0.73) than in the Norwegian breeds Nordland/Lyngen (0.66), Fjord (0.69), and Døle (0.55) (Bjørnstad et al. 2000).

Over time both wild and domestic populations face environmental changes and require genetic diversity to adapt to these changes. Populations under selection become inbred at a faster rate and are therefore more vulnerable to environmental changes. The long term consequences of intense selection, seen for example in domestic populations, are increased inbreeding and loss of genetic diversity. Genetic diversity is the necessary fuel for a successful and sustainable breeding of any population and one of the goals of population management is therefore to maintain genetic diversity at a high level and inbreeding at a low level (Fernandez et al. 2005, Frankham et al. 2010). The large influence of certain individuals and areas is clearly demonstrated here for the Icelandic horse breed. It is clear that with ongoing intense breeding, inbreeding will increase in the population which increases the importance of monitoring the status of the breed but also raises the question of the use of optimum contribution selection (Meuwissen 1997) which might restrict the increase in inbreeding while maintaining a reasonably high rate of genetic progress in desirable traits.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Áslaug Helgadóttir and Emma Eyþórsdóttir for critically reading the manuscript, deCode genetics for use of facilities, Istexa for Icelandic samples, and Ingimar Sveinsson for good advice. Financial support for this study was provided by the Icelandic Research Council (RANNÍS), the Icelandic Horse Conservation Fund, and the Icelandic Council for Genetic Resources in Agriculture.

REFERENCES

Aðalsteinsson S 1981. Origin and conservation of farm animal populations in Iceland. J*ournal of Animal Breeding and Genetics* 98, 258–264.

Agapow PM & Burt A 2001. Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* 1, 101–102.

Asbjarnardottir MG, Kristjansson T, Jonsson MB & Hallsson JH 2010. Analysis of genetic diversity and population structure within the Icelandic cattle breed using molecular markers. *Acta Agriculturae Scandinavica Section A-Animal Science* 60, 203–210.

Baldursdottir BK, Kristjansson T & Hallsson JH 2012. Diversity of the Icelandic goat breed assessed using population data. *Acta Agriculturae Scandinavica, Section A - Animal Science* 62, 1–13.

Binns MM, Holmes NG, Holliman A & Scott AM 1995. The identification of polymorphic microsatellite loci in the horse and their use in thoroughbred parentage testing. *British Veterinary Journal* 151, 9–15.

Björnsson GB & Sveinsson HJ 2004. *Íslenski hesturinn.* [*The Icelandic horse*]. Mál og Menning, Reykjavík, 415 p. [In Icelandic].

Bjørnstad G & Røed KH 2001. Breed demarcation and potential for breed allocation of horses assessed by microsatellite markers. *Animal Genetics* 32, 59–65.

Bjørnstad G, Gunby E & Røed KH 2000. Genetic structure of Norwegian horse breeds. *Journal of Animal Breeding and Genetics* 117, 307–317.

Bjørnstad G, Nilsen NO & Røed KH 2003. Genetic relationship between Mongolian and Norwegian horses? *Animal Genetics* 34, 55–58.

Boichard D 2002. Pedig: A Fortran package for pedigree analysis suited for large populations. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, 19-23 august 2002, paper 28-13.

Boichard D, Maignel L, Verrier E 1997. The value of using probabilities of gene origin to measure

genetic variability in a population. *Genetics Selection Evolution* 29, 5–23.

Bowling AT, Del Valle A & Bowling M 2000. A pedigree-based study of mitochondrial D-loop DNA sequence variation among Arabian horses. *Animal Genetics* 31, 1–7.

Breen M, Lindgren G, Binns MM, Norman J, Irvin Z, Bell K, Sandberg K & Ellegren H 1997. Genetical and physical assignments of equine microsatellites - first integration of anchored markers in horse genome mapping. *Mammalian genome* 8, 267–273.

Brown AHD, Feldman MW & Nevo E 1980. Multilocus structure of natural populations of Hordeum spontaneum. *Genetics* 96, 523–536.

Campana MG, Stock F, Barrett E, Benecke N, Barker GWW, Seetah K & Bower MA 2012. Genetic stability in the Icelandic horse breed. *Animal genetics* 43, 447–449.

Cavalli-Sforza LL & Edwards AWF 1967. Phylogenetic analysis: Models and estimation procedures. *American Journal of Human Genetics* 19, 233–257.

Cervantes I, Molina A, Goyache F, Gutierrez JP & Valera M 2008. Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. *Livestock Science* 113, 24-33.

Corander J, Marttinen P, Siren J & Tang J 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9, 539.

Corander J, Waldmann P, Marttinen P & Sillanpaa MJ 2004. BAPS 2: Enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* 20, 2363–2369.

Cozzi MC, Strillacci MG, Valiati P, Bighignoli B, Cancedda M & Zanotti M 2004. Mitochondrial D-loop sequence variation among Italian horse breeds. *Genetics Selection Evolution* 36, 663–672.

Druml T, Baumung R & Sölkner J 2009. Pedigree analysis in the Austrian Noriker draught horse: Genetic diversity and the impact of breeding for coat colour on population structure. *Journal of Animal Breeding and Genetics* 126, 348–356.

Eggleston-Stott ML, DelValle A, Bautista M, Dileanis S, Wictum E & Bowling AT 1997. Nine equine dinucleotide repeats at microsatellite loci UCDEQ136, UCDEQ405, UCDEQ412, UCDEQ425, UCDEQ437, UCDEQ467, UCDEQ487, UCDEQ502 and UCDEQ505. *Animal Genetics* 28, 370–371.

Ewens WJ 1972. Sampling theory of selectively neutral alleles. *Theoretical Population Biology* 3, 87–112.

Excoffier L, Laval G & Schneider S 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1, 47–50.

Felsenstein J 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* 39, 783–791.

Fernandez J, Villanueva B, Pong-Wong R & Toro MA 2005. Efficiency of the use of pedigree and molecular marker information in conservation programs. *Genetics* 170, 1313–1321.

Frankham R, Ballou JD & Briscoe DA 2010. *Introduction to Conservation Genetics*, 2nd ed. Cambridge University Press, Cambridge, UK, 644 p.

Guérin G, Bertaud M & Amigues Y 1994. Characterization of 7 New Horse Microsatellites - Hms1, Hms2, Hms3, Hms5, Hms6, Hms7 and Hms8. *Animal Genetics* 25, 62–62.

Hamann H & Distl O 2008. Genetic variability in Hanoverian warmblood horses using pedigree analysis. *Journal of Animal Science* 86, 1503–1513.

Haubold B, Travisano M, Rainey PB & Hudson RR 1998. Detecting linkage disequilibrium in bacterial populations. *Genetics* 150, 1341–1348.

Hill EW, Bradley DG, Al-Barody M, Ertugrul O, Splan RK, Zakharov I & Cunningham EP 2002. History and integrity of thoroughbred dam lines revealed in equine mtDNA variation. *Animal Genetics* 33, 287–294.

Icelandic Historical Statistics 1997. Number of livestock by type 1703-1852. Table 4.13a. CD-ROM. Statistics Iceland (Hagstofa Íslands), Reykjavik.

Ishida N, Hasegawa T, Takeda K, Sakagami M, Onishi A, Inumaru S, Komatsu M & Mukoyama H 1994. Polymorphic sequence in the D-loop region of equine mitochondrial-DNA. *Animal Genetics* 25, 215–221.

Jansen T, Forster P, Levine MA, Oelke H, Hurles M, Renfrew C, Weber J & Olek K 2002. Mitochondrial DNA and the origins of the domestic horse. *Proceedings of the National Academy of Sciences of the United States of America* 99, 10905–10910.

Kantanen J, Olsaker I, Holm LE, Lien S, Vilkki J, Brusgaard K, Eythorsdottir E, Danell B & Adalsteinsson S 2000. Genetic diversity and population structure of 20 North European cattle breeds. *Journal of Heredity* 91, 446–457

Kennedy BW & Trus D 1993. Considerations on genetic connectedness between management units under an animal model. *Journal of Animal Science* 71, 2341–2352.

Kim KI, Yang YH, Lee SS, Park C, Ma R, Bouzat JL & Lewin HA 1999. Phylogenetic relationships of Cheju horses to other horse breeds as determined by mtDNA D-loop sequence polymorphism. *Animal Genetics* 30, 102–108.

Kristjánsson Th 2005. Inbreeding depression and preservation of genetic variation in the Icelandic horse population. *MS thesis*. Hvanneyri, Iceland, 56 p.

Lear TL, Brandon R & Bell K 1999. Physical mapping of ten equine dinucleotide repeat microsatellites. *Animal Genetics* 30, 235–235.

Liu K & Muse SV 2005. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics Applications Note* 21, 2128–2129.

Lopes MS, Mendonca D, Cymbron T, Valera M, Da Costa-Ferreira J & Machado AD 2005. The Lusitano horse maternal lineage based on mitochondrial D-loop sequence variation. *Animal Genetics* 36, 196–202.

Luikart G, Allendorf FW, Cornuet JM & Sherwin WB 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* 89, 238–247.

Luis C, Bastos-Silveira C, Cothran EG & Oom MD 2006. Iberian origins of New World horse breeds. *Journal of Heredity* 97, 107–113.

Maynard Smith J, Smith NH, O'Rourke M & Spratt BG 1993. How clonal are bacteria? *Proceedings of the National Academy of Sciences of the United States of America* 90, 4384–4388.

McGahern A, Bower MAM, Edwards CJ, Brophy PO, Sulimova G, Zakharov I, Vizuete-Forster M, Levine M, Li S, MacHugh DE & Hill EW 2006. Evidence for biogeographic patterning of mitochondrial DNA sequences in Eastern horse populations. *Animal Genetics* 37, 494–497.

Meuwissen TH 1997. Maximizing the response of selection with a predefined rate of inbreeding. *Journal of Animal Science* 75, 934–940.

Mirol PM, Garcia PP, Vega-Pla JL & Dulout FN 2002. Phylogenetic relationships of Argentinean Creole horses and other South American and Spanish breeds inferred from mitochondrial DNA sequences. *Animal Genetics* 33, 356–363.

Nei M, Tajima F & Tateno Y 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* 19,153–170.

Peakall R & Smouse PE 2012. GenAlEx 6.5: Genetic analysis in Excel; Population genetic software for teaching and research--an update. *Bioinformatics* 28, 2537–2539.

Piry S, Luikart G & Cornuet JM 1999. Bottleneck: A computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity* 90, 502–503.

Rousset F 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8, 103–106.

Royo LJ, Alvarez I, Beja-Pereira A, Molina A, Fernandez I, Jordana J, Gomez E, Gutierrez JP & Goyache F 2005. The origins of Iberian horses assessed via mitochondrial DNA. *Journal of Heredity* 96, 663–669.

Siderits M, Baumung R, Fuerst-Waltl B 2013. Pedigree analysis in the German Paint Horse: Genetic variability and the influence of pedigree quality. *Livestock Science* 151, 152–157.

Statistics Iceland 2013. Database - Livestock from 1980. Accessed 04.04.2013 at the home-page of Statistics Iceland: http://www.statice.is/Statistics.

Swinburne J, Gerstenberg C, Breen M, Aldridge V, Lockhart L, Marti E, Antczak D, Eggleston-Stott M, Bailey E, Mickelson J, Røed K, Lindgren G, von Haeringen W, Guérin G, Bjarnason J, Allen T & Binns M 2000. First comprehensive low-density horse linkage map based on two 3-generation, full-sibling, cross-bred horse reference families. *Genomics* 66, 123–134.

Tajima F 1983. Evolutionary relationship of DNA-sequences in finite populations. *Genetics* 105, 437–460.

Tallmon DA, Koyuk A, Luikart G& Beaumont MA 2008. ONeSAMP: A program to estimate effective population size using approximate Bayesian computation. Molecular *Ecology Resources* 8, 299–301.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, & Kumar S 2011. MEGA5: Molecular evolutionary genetics analysis using maximum

likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28, 2731–2739.

Tapio M, Tapio I, Grislis Z, Holm LE, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H & Eythorsdottir E 2005. Native breeds demonstrate high contributions to the molecular variation in northern European sheep. *Molecular Ecology* 14, 3951–3963.

Valera M, Molina A, Gutierrez JP, Gomez J & Goyache F 2005. Pedigree analysis in the Andalusian horse: Population structure, genetic variability and influence of the Carthusian strain. *Livestock Production Science* 95, 57–66.

van Haeringen H, Bowling AT, Stott ML, Lenstra JA & Zwaagstra KA 1994. A highly polymorphic horse microsatellite locus: VHL20. *Animal Genetics* 25, 207–207.

Waples RS & Do C 2008. LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources* 8, 753–756.

Waples RS 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics* 7, 167–184.

Watterson GA 1975. Number of segregating sites in genetic models without recombination. *Theoretical Population Biology* 7, 256–276.

Weir BS & Hill WG 2002. Estimating F-statistics. *Annual Review of Genetics* 36, 721–750.

Xu XF & Arnason U 1994. The complete mitochondrial-DNA sequence of the horse, Equus-caballus - Extensive heteroplasmy of the control region. *Gene* 148, 357–362.

Yang YH, Kim KI, Cothran EG & Flannery AR 2002. Genetic diversity of Cheju horses (*Equus caballus*) determined by using mitochondrial DNA D-loop polymorphism. *Biochemical Genetics* 40, 175–186.

Zechner P, Solkner J, Bodo I, Druml T, Baumung R, Achmann R, Marti E, Habe F & Brem G 2002. Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information. *Livestock Production Science* 77, 137-146.

Manuscript received 15 April 2013
Accepted 15 April 2014