Journal of Heredity 2012:103(3):380–390
doi:10.1093/jhered/ess003

Maternal Lineages in Native Canadian Equine Populations and Their Relationship to the Nordic and Mountain and Moorland Pony Breeds

Jaclyn Mercedes Prystupa, Pamela Hind, E. Gus Cothran, and Yves Plante

From the Department of Animal and Poultry Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 5A8 (Prystupa and Plante); the Agriculture and Agri-Food Canada, Saskatoon, Saskatchewan, Canada S7N 0X2 (Prystupa, Hind, and Plante); and the Department of Veterinary Integrative Biosciences, College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, TX 77845 (Cothran).

Address correspondence to Jaclyn M. Prystupa at the address above, or e-mail: jmp479@mail.usask.ca.

Abstract

A 378-bp section of the mitochondrial displacement loop was used to estimate genetic diversity in the native Canadian equine populations. The inclusion of 10 Mountain and Moorland, 3 Nordic pony breeds, 2 feral populations, and 5 horse breeds were also investigated as they may have influenced the development (or rejuvenation) of the native Canadian populations. A total of 281 samples were sequenced, which produced 75 haplotypes derived from 54 informative sites. On further investigation, 36 of these 75 haplotypes were found to be previously unreported. Overall, total diversity was lowest in the feral Sable Island population with a haplotype diversity (0.27 ± 0.12), nucleotide diversity (0.0007 ± 0.0004), and pairwise difference of 0.286 ± 0.317. This is not surprising due to the geographic isolation of this population. Haplotype diversity was highest (1.00 ± 0.13) in the New Forest population, pairwise difference was highest (8.061 ± 4.028) in the Icelandic breed, whereas nucleotide diversity was highest in the Exmoor breed (0.0209 ± 0.0025). Within the Canadian populations, haplotype diversity was highest in the Newfoundland pony (0.96 ± 0.08), whereas pairwise difference and nucleotide diversity was highest in the Canadian horse (7.090 ± 3.581 and 0.0188 ± 0.0042, respectively). Three different estimates of genetic distances were used to examine the phylogenetic relationships amongst these populations. All 3 estimates produced similar topologies. In general, the native Canadian populations were highly represented in the D clade, with particular emphasis in the D1 and D2 clades. This is an important factor when considering the phylogenetic conservation of these Canadian equine populations.

Key words: *conservation, diversity, genetics, horse, mtDNA*

Viable conservation strategies are important for both domestic and wild species alike. One of the first, and crucial, steps in the development of any conservation strategy is to properly assess and characterize populations of interest at the phenotypic and genotypic levels (Glowatzki-Mullis et al. 2006; Plante et al. 2007; Gizaw et al. 2008). Different molecular tools and techniques have been utilized to assess the origin and genetic diversity present within a species, including mitochondrial DNA (mtDNA) sequence data (Vilà et al. 2001; Jansen et al. 2002; Cothran et al. 2005; Pedrosa et al. 2005; Luís et al. 2006; Pérez-Gutiérrez et al. 2008; Jia et al. 2010). In particular, the hypervariable region 1 of the displacement loop (D-loop) has been found to be particularly useful when estimating genetic diversity, phylogenetic relationships, and maternal origins among closely related populations (Aranguren-Méndez et al. 2001; Beja-Pereira et al. 2004; Cozzi et al. 2004; Cothran et al. 2005; Pedrosa et al. 2005; McGahern et al. 2006; Naderi et al. 2007; Jia et al. 2010).

Previous studies examining the origin of *Equus caballus* using mtDNA sequence data have reported a large number of ancestral haplogroups (17–19) belonging to 6 or 7 main clades (Vilà et al. 2001; Jansen et al. 2002; Cieslak et al. 2010). This is significantly higher than in other agricultural species, including cattle (Jia et al. 2010), sheep (Pedrosa et al. 2005), goats (Naderi et al. 2007), and donkeys (Beja-Pereira et al. 2004). Possible explanations for the differences observed among the species is likely a result of the

domestication process for each species. *E. caballus* is thought to have multiple origins of domestication (both in geography and time) along with a strong bias for the recruitment of mares over stallions from the wild (Vilà et al. 2001; Jansen et al. 2002; Wallner et al. 2003; Cieslak et al. 2010). Cieslak et al. 2010 has concluded that 39 of the 87 ancestral haplotypes found in various ancient horse remains (dating from 12000 BC to 1000 AD) are still present in modern day horse and pony breeds.

Canada is home to 3 native equine breeds (and feral population) consisting of 2 pony breeds, the Lac La Croix and the Newfoundland, and 1 horse breed, the Canadian, along with 1 feral equine population located on Sable Island (off the East Coast of Nova Scotia). These populations are all maintained in situ and are listed as endangered by a few conservation groups (Lynghaug 2009; Rare Breeds Canada [RBC] 2009; Equus Survival Trust [EST] 2011).

The Lac La Croix pony originated from the Great Lakes area (between the Canada–United States of America international border) and is believed to have arisen from crossing Canadian horses (breed) with Mustangs, although previous work investigating the native Canadian populations using microsatellites did not support this (Prystupa et al. 2011). The Lac La Croix population was reduced to 4 females in 1977, and in order to save the population, these mares were bred with Spanish Mustangs and were left to run feral on private property until the late 1990s.

The Canadian horse developed from a shipment of horses sent to Canada in 1665 from the king of France. The Canadian horse played a very important role in not only Canada but also in the United States as it influenced the development of many of the American gaited breeds (Lynghaug 2009). Previous studies using microsatellite loci have found the Canadian horse to be related to draft breeds such as the Belgian, Clydesdale, and Percheron (Plante et al. 2007; Prystupa et al. 2011).

The Newfoundland pony originated from the province of Newfoundland and Labrador and herds of ponies were frequently left to roam in a feral state when not required as a source of labor. Due to this early management strategy, the exact breeds used to develop the Newfoundland pony are unknown but are generally thought to have been derived from a number of the Mountain and Moorland pony breeds (Lynghaug 2009). The Mountain and Moorland breeds are native of the British Isles and include 11 pony breeds (Connemara, Dale, Dartmoor, Eriskay, Exmoor, Fell, Highland, Kerry Bog, New Forest, Shetland, and Welsh). As a result of multiple population reductions and bottlenecks experienced throughout history, several of the Mountain and Moorland breeds are also listed as endangered by various conservation organizations (Lynghaug 2009; RBC 2009; EST 2011).

The final native Canadian equine population consists of feral equines located on Sable Island (located off the East coast of Canada). The Sable Island horse population has been protected from human interference since 1960, but prior to that individual horse owners would leave stallions on the island to breed with the mares (Plante et al. 2007). The progeny would then be collected the following year and brought back to the mainland where they would be offered for sale (Nova Scotia Museum of Natural History [NSMNH] 2001). Previous work using microsatellite loci has found the Sable island population to be related to the Nordic pony breeds (Fjord, Icelandic, and Shetland; Plante et al. 2007; Prystupa et al. 2011).

The declaration of the convention on biological diversity serves as an important cornerstone regarding the importance of genetic diversity within livestock breeds (especially indigenous ones), as well as providing the first legal framework regarding estimating genetic diversity within farm animal species in various countries around the world (Food and Agriculture Organization [FAO] 2007). Furthermore, the Strategic Priority Area 1 of the FAO Global Plan of Action for Animal Genetic Resources (FAO 2007) clearly stresses the need to characterize farm animal genetic resources used for food and agriculture. Within this particular context, estimating molecular diversity in farm animal breeds and populations becomes a priority and an integral part of management strategies for the conservation of farm animal genetic resources.

This study examines the genetic diversity and relationships among the native Canadian equine populations and their relationship to the Mountain and Moorland (Connemara, Dale, Dartmoor, Eriskay, Exmoor, Fell, Highland, Kerry Bog, New Forest, Shetland, and Welsh) and Nordic (Fjord, Icelandic, and Shetland) pony breeds using a 378-bp section of the mtDNA D-loop. It should be noted that for the purpose of this study, the Shetland pony will be treated as a Nordic breed as previous studies using microsatellite loci have found it to be more closely related genetically to the Nordic breeds as opposed to European or Mountain and Moorland breeds (Plante et al. 2007; Leroy et al. 2009; and Prystupa et al. 2011). Results from this study will be combined with other microsatellite loci data previously collected and will aid in the development or possible amendment of conservation strategies for all populations included in this study (Prystupa et al. 2011).

Materials and Methods

Samples

In total, 281 (blood or hair) randomly selected samples were collected from 24 different populations with approximately 10 individuals sampled in each population. In addition to samples collected from the native Canadian equine populations (Canadian, Lac La Croix, Newfoundland, and Sable Island), samples were also collected from the Mountain and Moorland (Connemara, Dale, Dartmoor, Eriskay, Exmoor, Fell, Highland, Kerry Bog, New Forest, and Welsh), Nordic (Fjord, Icelandic, and Shetland), and historically important horse and feral populations. In total, 5 horse breeds (Caspian, Haflinger, Mongolian, Clydesdale, and Standardbred) along with 2 feral populations (Saint-Pierre et Miquelon and Grand Turk) were also included in this study. Samples from Saint-Pierre et Miquelon and

Grand Turk were included due to the close proximity of the French islands to Canada and historical records regarding trade flow between Turks and Caicos and North America. Samples were mainly from North American sources with the exception of the New Forest, Eriskay, Exmoor, and half of the Shetland samples that were collected in the United Kingdom. In addition, feral populations and Mongolian Domestic samples were collected from their respective countries of origin.

DNA Extraction

DNA was extracted from either whole blood or hair follicles using the same protocol as Prystupa et al. (2011). DNA samples of inferior quality for sequencing were reextracted from hair follicles (if available) using a DNeasy Blood & Tissue Kit (Qiagen Inc., Mississauga, Ontario) and a user-developed protocol available on the Qiagen website (http://www.qiagen.com/literature/render.aspx?id=519).

PCR and Sequencing

A 421-bp section of the mtDNA D-loop was amplified using the PCR and previously published primers (Cothran et al. 2005). PCR reactions were made using Platinum Taq polymerase (Invitrogen Canada Inc., Burling, Ontario) and the manufacturer's recommendations. PCR protocols involved an initial denaturation at 94 °C for 2 min followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min. PCR amplicons were purified using Exonuclease-I and Shrimp Alkaline Phosphatase (Fermentas Canada Inc., Burling, Ontario) according to the manufacturer's specifications.

Sequencing reactions were performed using BigDye Terminator Cycle Sequencing Kits v. 3.1 (Applied Biosystems, Foster City, CA) and the forward primer according to the manufacturer's recommendations. PCR protocols involved an initial denaturation at 96 °C for 1 min followed by 25 cycles of 96 °C for 10 s, 50 °C for 5 s, and 72 °C for 4 min. Sequencing reactions were purified using Agencourt CleanSEQ (Agencourt Bioscience Corporation/Beckman Coulter Company, Mississauga, Ontario) using the recommended protocol. Purified sequencing products were loaded on to a Genetic Analyzer 3130xl (Applied Biosystems) equipped with a 50 cm capillary array and filled with POP7 polymer. Sequences were examined using SEQMAN (Swindell and Plasterer 1997).

Sequences from the populations included in this study were entered into the National Center for Biotechnology Information (NCBI) GenBank database available at http://www.ncbi.nlm.nih.gov/ and were assigned accession numbers HQ592784–HQ593063 (see Supplementary Table 1).

Statistical Analysis

All sequences were aligned to the NCBI accession X79547 (Xu and Árnason 1994) and shortened to 378 bp to allow for maximum sample size, using CLUSTALW (Thompson et al. 1994), as implemented in MEGA4 (Tamura et al. 2007). A BLAST search using the NCBI database was used to determine any previously unreported haplotypes. Several genetic diversity parameters were estimated for the 24 populations of interest, with the consideration of gaps, including the number of variable sites, number of haplotypes. Using DNASP v. 5 (Librado and Rozas 2009), haplotype and nucleotide diversity were estimated. Pairwise comparisons were estimated using ARLEQUIN v. 3.5 (Excoffier and Lischer 2010).

A median-joining network showing the haplotypes proportionate to frequency was drawn, from all available sequences, using NETWORK 4.6 (Bandelt et al. 1999; Polzin and Daneshmand 2003) while removing 3 mutational hot spots and down weighing 3 sites in the mtDNA genome found at positions 15585, 15597, 15650 and 15604, 15659, and 15737, respectively, as recommended by Jansen et al. (2002) and Cieslak et al. (2010). Three additional statistical approaches (maximum parsimony, maximum likelihood, and Bayesian) were also used without removing mutation hot spots to also investigate the phylogenetic relationships among the equine populations (horse, pony, and feral) using programs available from the Cyberinfrastructure for Phylogenetic Research (CIPRES) Science Gateway available at http://www.phylo.org/ (Miller et al. 2010). Maximum parsimony was estimated using PAUPRAT (Sikes and Lewis 2001), with 200 iterations, 5 levels (5%, 10%, 15%, 20%, and 25%) of perturbed data, and 15 independent runs. Maximum likelihood was estimated using the general time reversible (GTR) model (Tavare 1986), with and without an estimate of the proportion of invariable sites (GTRGramma+1), as implemented in the RAXML-HPC black box (Stamatakis 2006). The Bayesian approach was estimated using MRBAYES (Huelsenbeck and Ronquist 2001) using the parameter settings of 2 runs, 4 chains, heating parameters of 0.2, sample frequency every 2000 generations, burn-in fraction of 0.25, and 1 million generations (the number of generations were increased until the "average standard deviation [SD] of split frequencies" fell below 0.05) using a 4 by 4 nuclear model and $N_{st} = 6$ setting in the CIPRES gateway. All phylogenetic consensus trees were depicted using MEGA4 (Tamura et al. 2007) and rooted to *E. ferus przewalski*, NCBI accession AF072994 (Lister et al. 1998).

Results

In total, 75 haplotypes from 54 polymorphic sites were found in the 24 populations examined in this study. Unfortunately, due to poor DNA template quality, only 5 New Forest samples could be sequenced. A BLAST search revealed that of the 75 haplotypes, 36 were not previously reported. The exact nuclear base substitutions and positions for the 36 unreported haplotypes can be found in Table 1. The unreported haplotypes included 15 breed-specific haplotypes; one in the Icelandic and 1 in the Highland, along with one haplotype found only in pony breeds (Dale and Newfoundland). Two breed-specific

lineages were also confirmed during the BLAST search in the Shetland and Caspian breeds (Hill et al. 2002; Flannery and Cothran 2003).

A summary of the results from the genetic diversity parameters estimated and haplotypes found for each breed/population is presented in Table 2. Variable sites within the defined breeds and populations varied from only 3 in the Dartmoor, Eriskay, Fjord, and Grand Turk populations to as high as 25 in the Canadian horse. The New Forest pony had the highest haplotype diversity (1.00 ± 0.13) and pairwise difference (8.000 ± 4.480), whereas the Icelandic pony breeds had the highest nucleotide diversity (0.0213 ± 0.0018).

A total of a 75 haplotypes (see Figure 1) were found throughout all 24 equine populations included in this study. All statistical approaches (Bayesian, maximum likelihood, and median-joining) examined in this study produced similar topologies revealing a strong admixture amongst the horse, pony, and feral populations (Figures 1 and 2). Interestingly, the relationship of the haplotypes also remained consistent, despite the possibility that the various levels of perturbed data would cause a dampening of the data in the parsimony phylogenetic trees (data not shown). When the mutational hot spots were taken into account (for median-joining network only), the number of haplotypes decreased from 75 to 61 (Figure 2).

The native Canadian populations appeared most frequently in the D clade, specifically the D1 and D2 haplogroups. Within the D clade, there were found to be 3 main haplogroups, which correlated to the sequence motifs of the mtDNA clusters as reported by Jansen et al. (2002). When taking into account the additional mutational hot spots reported by Cieslak et al. (2010), it was found that some of the haplotypes that were no longer separately represented within their own distinctive haplogroup, as depicted in the MJ network as represented in Figure 1. These haplotypes were, however, combined within other larger haplogroups, and as a result, are not visibly represented. D1 haplogroup incorporated all of the feral Sable Island population along with a few Canadian, Fjord, and Mongolian horses. The D2 haplogroup consisted largely of the Newfoundland and Lac La Croix ponies. Many other horse and pony breeds including the Canadian, Dartmoor, Dale, Caspian, Highland, Connemara, Icelandic, Mongolian, and St-Pierre et Miquelon populations were also included within this haplogroup. Haplogroup D3 is represented primarily by horse breeds as well as the Mountain and Moorland and Nordic ponies. A small percentage of Newfoundland ponies (5%) and St-Pierre et Miquelon (9%) are represented within this clade. Within the E haplogroup, a number of Lac La Croix ponies (30%) were also found. All remaining native Canadian population samples were spread throughout the A, B, and C clades.

The Mountain and Moorland breeds were found to be distributed across the entire network. Interestingly, the D3 haplogroup was found more commonly in the Mountain and Moorland breeds than the D1 and D2 haplogroups, which as previously stated, are composed primarily of the native Canadian populations. In addition, a substantial proportion of the Dale, Kerry Bog, and Shetland ponies were found in the E haplogroup. The A and C clades were also prominent within the Mountain and Moorland breeds with the majority of the Fell ponies contained within A3 haplogroup. The A5 haplogroup consisted of the Dartmoor, Connemara, Welsh, and Exmoor breeds. A number of different breeds were also found in the C1 haplogroup including: Exmoor, Kerry Bog, and Welsh. Additional individuals of the New Forest and Exmoor breeds were also found frequently in the C2 haplogroup.

The Nordic pony breeds were distributed across all clades with the exception of the clades B and A. The majority of the Nordic samples fell into the D clade, specifically Fjord and Icelandic ponies which fell into the D1 haplogroup. In addition, 23% of the Shetland samples fell in the D3 haplogroup. A large proportion of the Nordic breeds (33% of the Icelandic and 23% of the Shetland) were found to belong to the F1 haplogroup. The E haplogroup was also prominent within this group and included 35% of the Shetland, along with a small number of individuals from the Icelandic and Fjord breeds. A very small percentage of the Icelandic and Shetland were found in the C clade.

Several of the other horse, pony and feral populations included in this study were also found throughout the network. A large number of the Canadian horse breeds fell into the D clade. Within the Grand Turk population, 90% of the population was found in the A3 haplogroup and shared one haplotype with the Lac La Croix pony population. A large portion of the B1 haplogroup consisted of Clydesdale, Standardbred, Caspian, and Saint-Pierre et Miquelon horses. Several purebred and feral horses were also found within the C clade and these include the Clydesdale, Standardbred, Saint-Pierre et Miquelon found specifically in the C1 haplogroup. An additional 41% of Standardbred and 30% of the Clydesdale samples along with others were found within the C2 haplogroup. A large number of ponies were found within the E group, with the exception of one horse from the Grand Turk population.

Discussion

This study presents the first crucial look at mitochondrial diversity within the native Canadian equine populations. Although the data gathered here only reflects maternal gene flow among the populations, it complements our previous work investigating these populations using 38 microsatellite loci (Prystupa et al. 2011). The majority of the 24 populations of interest were found to have high haplotype and nucleotide diversity despite several of these populations undergoing recent herd reductions or bottlenecks. Similar results have been found in the breeds such as the Zemaitukai and German draft breeds which also had relatively high nucleotide diversity despite suffering from severe bottlenecks (Cothran et al. 2005; Aberle et al. 2007). Of the Canadian equines populations, the Canadian and Newfoundland were found to be the most diverse, whereas

Table 1 Summary of the 20 unreported haplotypes present in the 24 populations of interest

Haplotype	15494	15495	15496	15519	15521	15526	15533	15534	15538	15540	15542	15546	15585 *	15595	15596	15597 *	15600	15601	15602	15603	15604*	15615	15617	15635	15649	15650 *	15657	15659 *	15666	15683	15684	15703	15709	15718	15720	15726	15728	15737 *	15740	15770	15771	15775	15776	15777	15806	15807	15808
(X79547)	T	T	A	C	G	T	A	C	A	A	C	C	G	A	A	A	G	T	C	T	G	A	T	C	A	A	T	T	G	C	C	T	C	C	G	G	T	T	A	C	C	C	T	A	C	C	A
Hap_1	C	C	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	G	.	C	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_4	.	C	.	.	.	.	.	.	.	.	T	.	A	.	.	G	A	.	T	.	.	G	.	T	.	G	.	.	A	.	.	C	.	.	A	.	.	.	.	T	.	.	.	.	.	.	.
Hap_12	C	C	G	.	.	.	.	T	.	.	.	.	A	.	.	G	.	.	C	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_14	.	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	T	.	.	.	.	.	.	G	.	.	A	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_24	.	C	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	T	.	A	.	.	T	.	.	.	.	.	.	.	C	.	.	A	.	.	.	G	.	T	.	.	G	.	C	A
Hap_25	.	C	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	T	.	.	.	.	.	.	G	.	.	A	.	.	.	.	.	A	.	.	.	.	.	T	.	C	.	.	C	A
Hap_26	C	C	G	.	.	.	.	T	.	.	.	T	.	.	.	.	.	.	T	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_27	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	T	C	A
Hap_28	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_29	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	C	.
Hap_31	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	T	T	C	.
Hap_33	.	C	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_37	.	C	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	T	.	.	.	.	.	.	.
Hap_38	.	C	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	C	.	T	.	.	.	.	.	.	.
Hap_40	.	C	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	.	.	G	.	T	.	.	G	.	C	A
Hap_45	.	C	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	G	.	C	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_46	.	C	G	.	.	.	.	T	.	.	.	.	A	.	.	.	.	.	T	C	A	.	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_47	C	C	C	.	.	.	.	T	.	.	.	.	A	.	.	.	.	.	T	C	A	.	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_49	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	G	.	.	.	.	.	C	.	.	A	A	.	.	G	.	T	.	.	.	.	.	.
Hap_51	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	T	.	.	.	.	.
Hap_53	.	C	G	.	.	.	.	T	.	.	.	.	A	.	.	.	.	.	T	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_56	C	C	G	.	.	.	.	T	.	.	.	.	A	.	.	.	.	.	T	C	A	.	.	.	G	.	.	.	.	.	.	.	.	.	A	.	C	.	.	.	T	.	.	.	.	.	.
Hap_58	.	C	.	T	.	.	.	.	G	.	.	.	.	.	.	G	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	.	T	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_59	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	.	T	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	T	A
Hap_61	.	C	.	.	.	.	.	.	.	.	T	.	A	.	.	G	A	.	T	.	.	G	.	T	.	G	.	.	A	.	.	C	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.
Hap_62	.	C	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	T	.	.	.	C	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	T		A
Hap_63	.	C	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	.	T	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_64	.	C	.	.	.	.	G	.	.	.	.	.	A	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	T	.	A
Hap_65	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	T	.	.	.	.	A	.	.	.	.	.	T	.	.	.	T	.	A
Hap_66	.	C	.	.	.	.	.	.	G	.	.	.	A	.	G	.	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	.	T	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_69	.	C	.	.	.	C	.	.	.	G	.	.	A	.	.	.	.	T	T	.	.	.	.	.	G	.	.	.	.	.	.	.	T	A	.	.	.	.	.	T	.	C	.	.	.	A
Hap_70	C	C	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	G	.	.	.	.	.	T	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_73	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	C	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	T	.	A
Hap_74	.	C	.	G	C	.	.	.	G	.	.	.	.	G	.	.	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	.	T	.	A	.	.	.	.	.	T	.	.	.	.	.	.
Hap_75	C	C	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.

The summary of the 20 unreported haplotypes present in the 24 populations of interest for this study. Specific mutations in sites are listed. * denotes sites that are considered "mutational hotspots" by previous authors (Jansen et al. 2002; Cieslak et al. 2010) and were reweighted or omitted according to recommendations by the previous authors for the median-joining network.

the Sable Island population was the least diverse population in the entire data set. These results are not unexpected as the Newfoundland and Sable Island populations were found to also have high and low autosomal diversity, respectively, when estimated using microsatellite loci (Prystupa et al. 2011).

All 4 statistical approaches used (Bayesian, maximum likelihood, maximum parsimony, and median-joining) consistently showed the same relationships among the various haplotypes (and therefore populations) found within this study. In general, the majority of individuals from the 4 native Canadian equine populations were most prominent in the D clade. Previous studies have reported that individuals belonging to this clade are of British and Iberian descent, therefore, our findings suggest that the majority of our

Table 2 Summary of the genetic diversity estimates and haplotypes found within the 24 populations of interest

Breed	*n*	Variable sites	Number of haplotypes	Haplotype diversity ± SD	Nucleotide diversity ± SD	Pairwise difference ± SD	Haplotypes (*n*)
Canadian equine populations							
Canadian	12	25	6	0.85 ± 0.07	0.0188 ± 0.0042	7.090 ± 3.581	Hap2 (3); **Hap4 (1)**; Hap5 (1); Hap9 (2); Hap10 (1); Hap11 (4)
Lac La Croix	10	14	4	0.64 ± 0.15	0.0155 ± 0.0033	5.867 ± 3.062	Hap2 (6); Hap20 (2); Hap30 (1); Hap71 (1)
Newfoundland	18	24	11	0.96 ± 0.08	0.0148 ± 0.0028	4.392 ± 2.275	**Hap1 (1)**; Hap2 (4); Hap3 (5); **Hap4 (1)**; Hap6 (1); Hap13 (1); Hap18 (1); Hap19 (1); **Hap45 (1)**; Hap68 (1); **Hap75 (1)**
Sable Island	21	3	4	0.27 ± 0.12	0.0008 ± 0.0004	0.286 ± 0.317	Hap11 (18); **Hap46 (1)**; **Hap47 (1)**; **Hap53 (1)**
Mountain and Moorland breeds							
Connemara	12	22	8	0.89 ± 0.08	0.0200 ± 0.0027	7.879 ± 3.944	Hap2 (1); Hap5 (4); Hap30 (1); Hap32 (1); **Hap33 (1)**; Hap34 (2); Hap35 (1); **Hap40 (1)**
Dale	12	18	5	0.58 ± 0.16	0.0146 ± 0.0040	5.515 ± 2.852	**Hap (1)**; Hap2 (1); Hap7 (8); Hap8 (1); Hap9 (1)
Dartmoor	12	10	3	0.44 ± 0.16	0.0107 ± 0.0040	4.045 ± 2.171	Hap2 (1); Hap5 (9); Hap6 (2)
Eriskay	12	16	3	0.59 ± 0.11	0.0170 ± 0.0030	6.440 ± 3.280	Hap20 (7); **Hap69 (4)**; **Hap70 (1)**
Exmoor	11	21	6	0.84 ± 0.09	0.0209 ± 0.0025	7.890 ± 3.978	Hap8 (4); Hap10 (1); **Hap58 (1)**; **Hap59 (1)**; **Hap60 (1)**; **Hap61 (3)**
Fell	9	8	4	0.58 ± 0.18	0.0056 ± 0.0025	2.111 ± 1.300	**Hap28 (1)**; **Hap29 (1)**; Hap30 (6); **Hap31 (1)**
Highland	11	16	6	0.84 ± 0.09	0.0191 ± 0.0026	7.236 ± 3.674	Hap2 (1); Hap3 (1); Hap9 (1); Hap16 (3); **Hap17 (4)**; Hap18 (1)
Kerry Bog	10	15	6	0.84 ± 0.10	0.0155 ± 0.0040	5.867 ± 3.062	Hap7 (4); Hap19 (2); Hap20 (1); Hap21 (1); Hap22 (1); Hap30 (1)
New Forest	5	15	5	1.00 ± 0.13	0.0202 ± 0.0033	8.000 ± 4.480	Hap18 (1); Hap44 (1); Hap57 (1); **Hap64 (1)**; **Hap65 (1)**
Welsh	10	17	7	0.87 ± 0.11	0.0116 ± 0.0030	6.267 ± 3.250	Hap5 (1); Hap18 (4); Hap19 (1); Hap41 (1); Hap42 (1); Hap43 (1); Hap44 (1)
Nordic breeds							
Fjord	11	18	3	0.62 ± 0.10	0.0207 ± 0.0032	7.818 ± 3.944	Hap11 (6); Hap36 (4); Hap72 (1)
Icelandic	12	20	7	0.88 ± 0.08	0.0213 ± 0.0018	8.061 ± 4.028	Hap2 (1); Hap3 (1); Hap11 (2); Hap19 (1); **Hap37 (1)**; **Hap38 (2)**; Hap39 (4)
Shetland	17	22	6	0.80 ± 0.06	0.0206 ± 0.0013	7.790 ± 3.818	Hap10 (1); Hap16 (4); Hap20 (6); Hap39 (4); **Hap62 (1)**; **Hap63 (1)**
Feral populations							
Grand Turk	11	8	3	0.56 ± 0.02	0.0054 ± 0.0024	2.036 ± 1.236	Hap20 (1); Hap30(7); Hap57 (3)
Saint-Pierre and Miquelon	11	19	6	0.84 ± 0.09	0.0172 ± 0.0024	6.510 ± 3.340	Hap2 (3); Hap13 (1); Hap18 (1); Hap19 (1); Hap48 (4); **Hap49 (1)**
Horse breeds							
Caspian	10	19	8	0.96 ± 0.06	0.0189 ± 0.0021	7.133 ± 3.657	Hap2 (1); **Hap12 (2)**; Hap13 (2); **Hap14 (1)**; Hap15 (1); Hap50 (1); **Hap51 (1)**; Hap52(1)
Clydesdale	10	17	7	0.93 ± 0.06	0.0161 ± 0.0024	6.270 ± 3.250	Hap10 (2); Hap13 (2); Hap18 (1); Hap19 (2); Hap34 (1); **Hap73 (1)**; **Hap74 (1)**

Table 2 Continued

Breed	*n*	Variable sites	Number of haplotypes	Haplotype diversity ± SD	Nucleotide diversity ± SD	Pairwise difference ± SD	Haplotypes (*n*)
Haflinger	11	24	6	0.87 ± 0.07	0.0199 ± 0.0042	7.530 ± 3.801	Hap18 (3); Hap23 (3); **Hap24 (1)**; **Hap25 (1)**; **Hap26 (1)**; **Hap27 (2)**
Mongolian	11	17	8	0.95 ± 0.05	0.0174 ± 0.0020	6.581 ± 3.369	Hap2 (2); Hap11 (1); Hap19 (1); Hap32 (2); Hap54 (1); Hap55 (2); **Hap56 (1)**; Hap57 (1)
Standardbred	12	15	6	0.82 ± 0.10	0.0132 ± 0.0023	5.000 ± 2.614	Hap11 (1); Hap13 (1); Hap19 (2); Hap55 (1); **Hap56 (2)**; Hap57 (5); **Hap66 (1)**

The summary of the genetic diversity estimates (using a 378 section of the D-Loop) and haplotypes found within the 24 populations of interest including the sample size (*n*), number of variable sites, number of haplotypes, haplotype diversity (± standard deviation [SD]), nucleotide diversity (± SD), and pairwise difference (± SD). Bolded haplotypes are previously unreported. Underlined haplotypes are breed specific.

native Canadian equine populations likely share a common maternal ancestry with equines of British and/or Iberian stocks (McGahern et al. 2006; Lei et al. 2009). The Canadian horse shared haplotypes with all the remaining native Canadian populations, along with draft, Nordic, and Mountain and Moorland breeds. This is suggestive of past gene flow among all native Canadian equine populations. This information supports our previous work that found a strong relationship between the Canadian horse and the draft breeds (Plante et al. 2007; Prystupa et al. 2011).

Contrary to the information gathered using microsatellite loci, the Newfoundland did not show a strong relationship with the Sable Island herd as these 2 populations did not share overlapping haplotypes (Prystupa et al. 2011). This may be due to the fact the Newfoundland population underwent a severe herd reduction and bottleneck in the 1980s, and any haplotypes that were previously shared between these populations were simply lost (Lynghaug 2009). These potential maternal lineages from the Sable Island population have never been reintroduced into the Newfoundland breed as the Sable Island population has been protected from human interference since 1960 (Plante et al. 2007). Despite the severe herd reduction and bottleneck, the Newfoundland pony was found to have the highest number of haplotypes observed within the data set and shared several of these with horse and pony breeds (Mountain and Moorland, and Nordic) as well as the feral population of Saint-Pierre et Miquelon. This may be a result of early management with a combination of several types of origin and/or ancestral variability (Vilà et al. 2001; Cieslak et al. 2010). There is speculation that some Newfoundland ponies were crossed with some horse breeds (in particular the Clydesdale and Standardbred) in order to improve certain desired traits. This study showed that some of the Newfoundland ponies did share haplotypes dominated by both these breeds, suggesting that at one time at least some Newfoundland ponies were crossed with Standardbreds and Clydesdales, however; this was not observed by Prystupa et al. (2011) while investigating genomic diversity using microsatellite markers. In contrast, the strong relationship observed between the feral Saint-Pierre et Miquelon population and the Newfoundland pony supports previous work (Prystupa et al. 2011).

The most common haplotype within the Sable Island population was also found in individuals from a number of different breeds including the Canadian, Connemara, Fjord, Icelandic, Mongolian, and Standardbred, perhaps indicating that this population shares a common maternal ancestry with several different types of equines and supports historical data regarding the types of horses released to breed with the population on the island in the past (NSMNH 2001). This study reemphasizes the previous studies involving microsatellite loci and the close relationship among the Nordic breeds and their relationship to the Sable Island population (Plante et al. 2007; Leroy et al. 2009; Prystupa et al. 2011). The low diversity in the Sable Island population is most likely a result of demographic aggregation. Subsequently, the 4 observed haplotypes may reflect the resource distribution on Sable Island, maternal lines

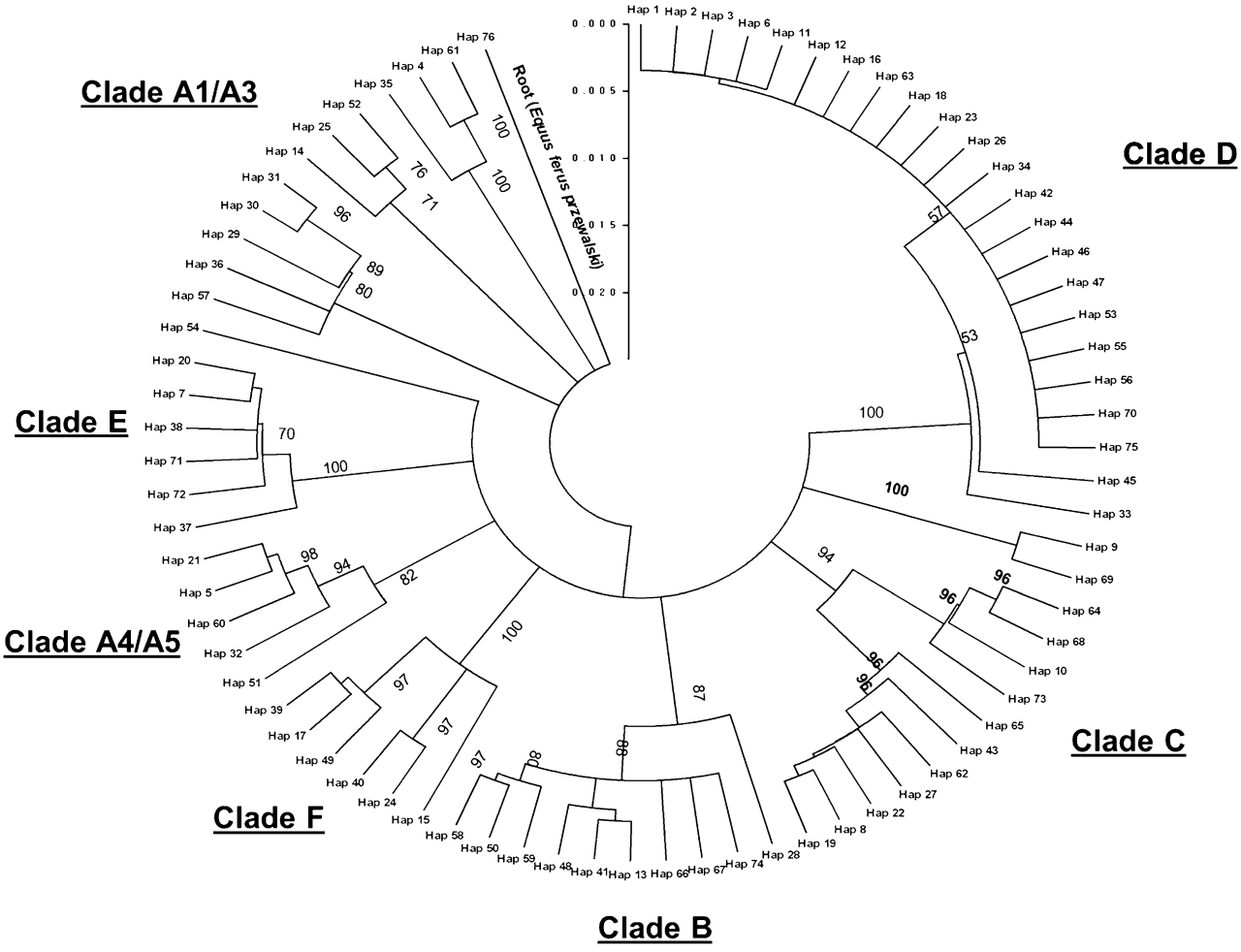

Figure 1. The phylogenetic tree created using a Bayesian approach of the 75 haplotypes observed in this study rooted to the Prezwalski haplotype (Hap_76). Gray numbers on individual branches indicate the posterior probabilities associated with them.

associated with certain areas or territories, or the fluctuation in population size from year to year (Lucas et al. 2009).

The haplogroups observed in the Lac La Croix were those found frequently in other populations (horse, pony, and feral). The Lac La Croix shared one haplotype in common with the Canadian horse, as well as Mountain and Moorland and Nordic populations. Interestingly, the previous study examining the relationship between the Lac La Croix and the Canadian horse using microsatellite loci was unable to find a relationship between these breeds, and this also likely is due to the more recent infusion of Mustangs into the Lac La Croix population (Prystupa et al. 2011). In addition, the Lac La Croix also shared 2 haplotypes (Hap_20 and Hap_30) with the Grand Turk population and gives support to the opinion that some indigenous Canadian equine populations could have been influenced from equines shipped to Canada from Grand Turk in the past. Alternatively, both these populations have arisen from Spanish stock (Mustangs and horse breeds), and this might also explain the overlapping of haplotypes.

One of the most interesting findings in this study was the large proportion of pony breeds (Canadian, Mountain and Moorland, and Nordic) belonging to the rare E haplogroup. The majority of the individuals assigned to this haplogroup were ponies with the inclusion of one feral horse from Grand Turk. Previous studies have reported this haplogroup to be not only rare worldwide, appearing in only 3% of living equines (although very predominant within the Kerry Bog ponies [40%]); but also to be old, dating back as far as the Bronze Age (3200 BC to 1200 BC; Kakoi et al. 2007; Cieslak et al. 2010). This study also found a number of samples from the Dale, Eriskay, Shetland, Lac La Croix, Icelandic, and Fjord breeds belonging to this group. Interestingly, in the Prystupa et al. (2011) study, the Eriskay and Lac La Croix were among the top breeds recommended for the development of a conservation strategy, indicating that these breeds are not only distinct using autosomal but also mitochondrial information. Both of these breeds had haplotypes found within the E clade, as well as breed-specific haplotypes found in the Eriskay.

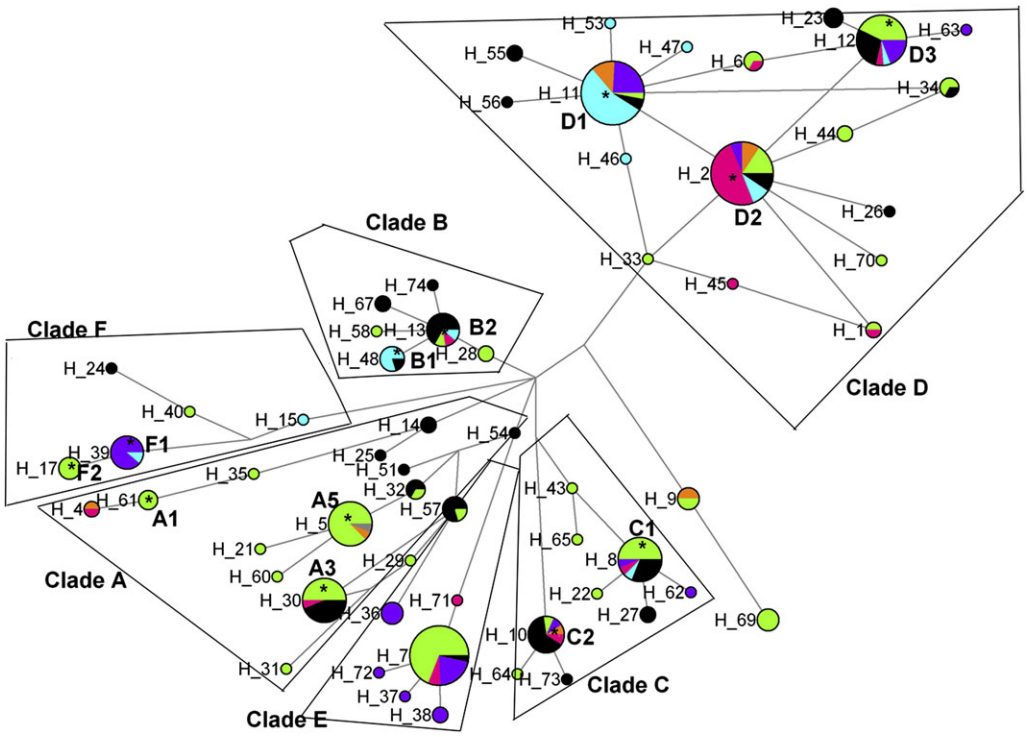

Figure 2. The median-joining network representing all 24 equine populations and taking into account the "mutational hot spots" with the recommendations of previous research (Jansen et al. 2002; Cieslak et al. 2010). The major equine clades originally presented by Vilà et al. (2001) along with the common haplogroups identified by Jansen et al. (2002) are labeled with circles proportionate to the frequency of haplotypes. Colors represent different equine groups: horse breeds = black; Mountain and Moorland breeds = green; Nordic breeds = purple; Canadian pony breeds = pink; Canadian feral populations = blue; Canadian horse populations = orange; reference sample X79547 (Hap_5) populations = gray.

A few other observations were also found within this study regarding the relationship among some of the breeds and are worth noting. The first is that although the Fell and Dales ponies have very similar phenotypic traits and are very similar based on protein and microsatellite variation, they did not share overlapping haplotypes (Luís et al. 2007; Prystupa et al. 2011). The majority of the Fell ponies were grouped within the A clade; whereas the majority of the Dale ponies grouped into the E clade. The second is that this study also found that the Mountain and Moorland breeds shared 2 haplotypes with the Grand Turk population. One possible explanation for this may be that ponies from Europe were likely sent to the island at one time where they either remained or were traded to other countries including Canada.

Lastly, our work found that some horse breeds, specifically the Standardbred and Clydesdale, were just as prominent in the C1 haplogroup as some of the Mountain and Moorland pony breeds. This is surprising because the C1 haplogroup has been previously reported to contain predominately Northern European pony breeds (Jansen et al. 2002).

Conclusions

This study provided the first real insight into the maternal gene flow and mitochondrial diversity within the native Canadian equine populations. In general, the information gathered here supports previous information published regarding the relationship among these breeds while using microsatellite loci (Prystupa et al 2011). The Sable Island population was shown to share one haplotype with the Nordic breeds along with the Canadian horse. The Newfoundland pony appears to have the highest number of maternal lines, indicating that perhaps mares of many different backgrounds were used to repopulate the breed or that the breed contains a high level of ancestral variability. Interestingly, the native Canadian pony breeds did not share haplotypes with the Sable Island herd; instead the Newfoundland pony shared haplotypes with the Saint-Pierre and Miquelon population (as previously found). The Lac La Croix shared haplotypes with the Grand Turk population (Prystupa et al. 2011), and the Canadian horse shares haplotypes with all the remaining native Canadian equine populations along with other draft, Mountain and Moorland, and Nordic breeds. The information gathered here along with previous microsatellite data will be combined to produce (or amend) conservation strategies for all breeds examined.

Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding

This work was supported by the Canadian Animal Genetic Resources program of Agriculture and Agri-Food Canada.

Acknowledgments

The authors would like to acknowledge the many breed associations and societies, Maxxam Analytics (Ontario, Canada), Zoe Lucas from the Green Horse Society, and Patricia Morris for providing valuable support and samples.

References

Aberle KS, Hamann H, Drögemüller C, Distl O. 2007. Phylogenetic relationships of German heavy draught horse breeds inferred from mitochondrial DNA D-loop variation. J Anim Breed Genet. 124:94–100.

Aranguren-Méndez J, Jordana J, Gomez M. 2001. Genetic diversity in Spanish donkey breeds using microsatellite DNA markers. Genet Sel Evol. 33:433–442.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 16:37–48.

Beja-Pereira A, England PR, Ferrand N, Jordan S, Bakhiet AO, Abdalla MA, Mashkour M, Jordana J, Taberlet P, Luikart G. 2004. African origins of the domestic donkey. Science. 304:1781.

Cieslak M, Pruvost M, Benecke N, Hofreiter M, Morales A, Reissmann M, Ludwig A. 2010. Origin and history of mitochondrial DNA lineages in domestic horses. PLoS One. 5:e15311. doi:10.1371/journal.pone.0015311

Cothran EG, Juras R, Macijauskiene V. 2005. Mitochondrial DNA D-loop sequence variation among 5 maternal lines of the Zemaitukai horse breed. Genet Mol Biol. 28:677–681.

Cozzi MC, Strillacci MG, Valiati P, Bighignoli B, Cancedda M, Zanotti M. 2004. Mitochondrial D-loop sequence variation among Italian horse breeds. Genet Sel Evol. 36:663–672.

Equus Survival Trust (EST). 2011. Conservation list [Internet]. Lowgap (NC); [cited 2011 Mar 11]. Available from: www.equus-survival-trust.org/documents/equineconservationlist.pdf

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res. 10:564–567.

Flannery AR, Cothran EG. 2003. Mitochondrial DNA sequence variation and the domestication pattern of the horse [Internet]. Lexington (KY); [cited 2011 Mar 11]. Available from: http://www.laclacroixindianpony.com/pdfs/cothranDNAreport.pdf.

Food and Agriculture Organization (FAO). 2007. The state of the world's animal genetic resources for food and agriculture. In: Rischkowsky B, Pilling D, editors. Rome (Italy).

Gizaw S, Komen H, Windig JJ, Hanotte O, Arendonk JAMV. 2008. Conservation priorities for Ethiopian sheep breeds combining threat status, breed merits and contributions to genetic diversity. Genet Sel Evol. 40:433–447.

Glowatzki-Mullis ML, Muntwyler J, Pfister W, Marti E, Rieder S, Poncet PA, Gaillard C. 2006. Genetic diversity among horse populations with a special focus on the Franches-Montagnes breed. Anim Genet. 37:33–39.

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. Anim Genet. 33:287–294.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 17:754–755.

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. Proc Natl Acad Sci U S A. 99:10905–10910.

Jia S, Zhou Y, Lei C, Yao R, Zhang Z, Fang X, Chen H. 2010. A new insight into cattle's maternal origin in six Asian countries. J Genet Genomics. 37:173–180.

Kakoi H, Tozaki T, Gawahara H. 2007. Molecular analysis using mitochondrial DNA and microsatellites to infer the formation process of Japanese native horse populations. Biochem Genet. 45:375–395.

Lei CZ, Su R, Bower MA, Edwards CJ, Wang XB, Weining S, Liu L, Xie WM, Li F, Liu RY, et al. 2009. Multiple maternal origins of native modern and ancient horse populations in China. Anim Genet. 40:933–944.

Leroy G, Callède L, Verrier E, Mériaux JC, Ricard A, Danchin-Burge C, Rognon X. 2009. Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. Genet Sel Evol. 41:5.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25:1451–1452.

Lister AM, Kadwell M, Kaagan LM, Stanley HF. 1998. Ancient and modern DNA in a study of horse domestication. London: Conservation Genetics, Institute of Zoology.

Lucas ZL, McLoughlin PD, Coltman DW, Barber C. 2009. Multiscale analysis reveals restricted gene flow and a linear gradient in heterozygosity for an island population of feral horses. Can J Zool. 87:310–316.

Luís C, Bastos-Silveira C, Cothran EG, Oom MM. 2006. Iberian origins of new world horse breeds. J Hered. 97:107–113.

Luís C, Juras R, Oom MM, Cothran EG. 2007. Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. Anim Genet. 38:20–27.

Lynghaug F. 2009. The official horse breeds standards guide: the complete guide to the standards of all North American equine breed associations. Minneapolis (MN): Voyageur Press.

McGahern AM, Edwards CJ, Bower MA, Heffernan A, Park SDE, Brophy PO, Bradley DG, MacHugh DE, Hill EW. 2006. Mitochondrial DNA sequence diversity in extant Irish horse populations and in ancient horses. Anim Genet. 37:498–502.

Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE); 2010 Nov 14; New Orleans (LA). p. 1–8.

Naderi S, Rezaei HR, Taberlet P, Zundel S, Rafat SA, Naghash HR, El-Barody MAA, Ertugrul O, Pompanon F, Abo-Shehada M, et al. 2007. Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. PLoS One. 2:e1012. doi:10.1371/journal.pone.0001012

Nova Scotia Museum of Natural History (NSMNH). 2001. Sable Island [Internet]. Halifax (Nova Scotia); [cited 2011 Mar 11]. Available from: http://museum.gov.ns.ca/mnh/nature/sableisland/index.htm

Pedrosa S, Uzun M, Arranz JJ, Gutiérrez-Gil B, San Primitivo F, Bayón Y. 2005. Evidence of three maternal lineages in near eastern sheep supporting multiple domestication events. Proc R Soc B Biol Sci. 272:2211–2217.

Pérez-Gutiérrez LM, De La Peña A, Arana P. 2008. Genetic analysis of the Hispano-Breton heavy horse. Anim Genet. 39:506–514.

Plante Y, Vega-Pla JL, Lucas Z, Colling D, De March B, Buchanan F. 2007. Genetic diversity in a feral horse population from Sable Island, Canada. J Hered. 98:594–602.

Polzin T, Daneshmand SV. 2003. On Steiner trees and minimum spanning trees in hypergraphs. Oper Res Lett. 31:12–20.

Prystupa JM, Juras R, Cothran EG, Buchanan FC, Plante Y. 2011. Genetic diversity and admixture among Canadian, Mountain and Moorland, and Nordic pony populations. Animal. 6:19–30.

Rare Breeds Canada (RBC). 2009. Canada's livestock and poultry conservation lists—2009 [Internet]. Notre-Dame-de-l'Île-Perrot (Quebec); [cited 2011 Mar 11]. Available from: http://www.rarebreedscanada.ca/2009conslist.pdf.

Sikes DS, Lewis PO. 2001. PAUPRat: PAUP* implementation of the parsimony ratchet. Storrs (CT): University of Connecticut.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.

Swindell SR, Plasterer TN. 1997. SEQMAN: Contig assembly. Methods Mol Biol. 70:75–89.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Tavare S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Am Math Soc Lect Math Life Sci. 17:57–86.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.

Vilà C, Leonard JA, Götherström A, Marklund S, Sandberg K, Lidén K, Wayne RK, Ellegren H. 2001. Widespread origins of domestic horse lineages. Science. 291:474–477.

Wallner B, Brem G, Müller M, Achmann R. 2003. Fixed nucleotide differences on the Y chromosome indicate clear divergence between *Equus przewalskii* and *Equus caballus*. Anim Genet. 34:453–456.

Xu X, Aárnason U. 1994. The complete mitochondrial DNA sequence of the horse, Equus caballus: extensive heteroplasmy of the control region. Gene. 148:357–362.

Received March 31, 2011; Revised December 15, 2011; Accepted January 22, 2012

Corresponding Editor: Ernest Bailey