Western North American Naturalist 79(1), © 2019, pp. 85–98

Genetic structure of the Big Summit herd and neighboring wild horse populations inhabiting herd management areas of Oregon

Ketaki Deshpande[1,2,3], Evelyn Perez[1,2], Natalie Leyva[1,2], Merly Suarez[3], and DeEtta K. Mills[1,3,*]

[1]*International Forensic Research Institute, Florida International University, Miami, FL 33199*
[2]*Department of Chemistry and Biochemistry, Florida International University, Miami, FL 33199*
[3]*Department of Biological Sciences, Florida International University, Miami, FL 33199*

Abstract.—Feral horses roaming North America are distinct not only because of their acclimatization to the wild but also because of their diversified ancestry. The wild horses inhabiting the Ochoco National Forest near Prineville, Oregon, are one of the remaining wild mustang herds and are protected by the Wild Horse and Burro Acts of 1971 and 1978. With deterioration of their rangeland, drought, vegetation conditions, human encroachment, fencing off of grazing lands, competition with wildlife, and other livestock grazing on public lands, a critical balance between conserving the natural ecology of rangelands and protecting and managing the wild horses needs to be monitored. Supplementing traditional visual censuses with genetic analyses can enhance management and conservation efforts because DNA analyses provide insight into the genetic fitness and inbreeding status of the population. The objective of this study was to provide a genetic analysis using noninvasive sample collection methods to assess the genetic health of this small herd and to determine their genetic fitness. A total of 52 individuals from the Big Summit wild horse herd and adjacent herd management areas (HMAs) were genotyped for 17 short tandem repeat (STR) loci, mtDNA haplogroups, and major histocompatibility complex STR loci. Cluster analysis exhibited an admixed population with discernable contribution from Iberian ancestry, the major influence coming from Andalusian and Lusitano breeds. The deficiency of heterozygosity, a deviation from Hardy–Weinberg equilibrium, together with a positive inbreeding coefficient for the neutral STRs in the Big Summit population suggested a parallel to an "island population" phenomenon leading to loss of genetic diversity within the herd. These findings improve understanding of the genetic structure of feral herds from different HMAs, which in turn will enable enhanced conservation and management strategies.

Resumen.—Los caballos salvajes que habitan América del Norte se distinguen no sólo por su aclimatación a la vida silvestre sino también por su diversificada ascendencia. Los caballos salvajes que habitan el Bosque Nacional Ochoco, cerca de Prineville (Oregon), son una de las pocas manadas de caballos mesteños salvajes que aún persisten y que están protegidos por las Leyes de burros y caballos salvajes de 1971 y 1978. Dado el deterioro de los pastizales, la sequía, las condiciones de la vegetación, la ocupación humana, el cercamiento de las tierras aptas para el pastoreo, la constante competencia con otros animales silvestres y con el ganado que pasta en tierras publicas, es necesario monitorear el equilibrio entre la conservación de los pastizales, la protección y el control de los caballos salvajes y la ecología natural de sus hábitats. Complementar el censo visual tradicional con análisis genéticos puede promover los esfuerzos de manejo, ya que, los análisis de ADN proporcionan información sobre la adecuación genética y el estado de endogamia de la población. El objetivo de este estudio fue proporcionar un análisis genético utilizando métodos no invasivos de recolección de muestras para evaluar la salud y la adecuación genética de esta pequeña manada. En total, 52 individuos de la manada de caballos salvajes de Big Summit y de las áreas de manejo de manadas adyacentes (HMAs, por sus siglas en inglés) fueron genotipados 17 secuencias cortas repetidas en tándem (STR, por sus siglas en inglés), haplogrupos de ADN mitocondrial (mtDNA por sus siglas en inglés) y loci STR del complejo mayor de histocompatibilidad. El análisis de agrupamiento exhibió una población mezclada con una contribución discernible de ascendencia ibérica, la mayor influencia proviene de las razas andaluza y lusitana. La deficiencia en la heterocigosidad, la desviación del equilibrio Hardy–Weinberg, junto con un coeficiente de endogamia positivo de los STR neutrales de la población del Big Summit, sugieren un fenómeno análogo al de la "población de isla" llevando a la pérdida de la diversidad genética dentro de la manada. Estos resultados ayudan a comprender la estructura genética de las manadas salvajes de diferentes HMA que, a su vez, permiten promover las estrategias de conservación y manejo.

America's wild horses are descendants of European horses introduced during early exploration and colonization of the Americas (Luis et al. 2006, McGahern et al. 2006, Prystupa et al. 2012). However, these feral horses now represent a genetic mixture of more recent

*Corresponding author: millsd@fiu.edu

KD orcid.org/0000-0002-3131-5748 DKM orcid.org/0000-0002-9977-7834

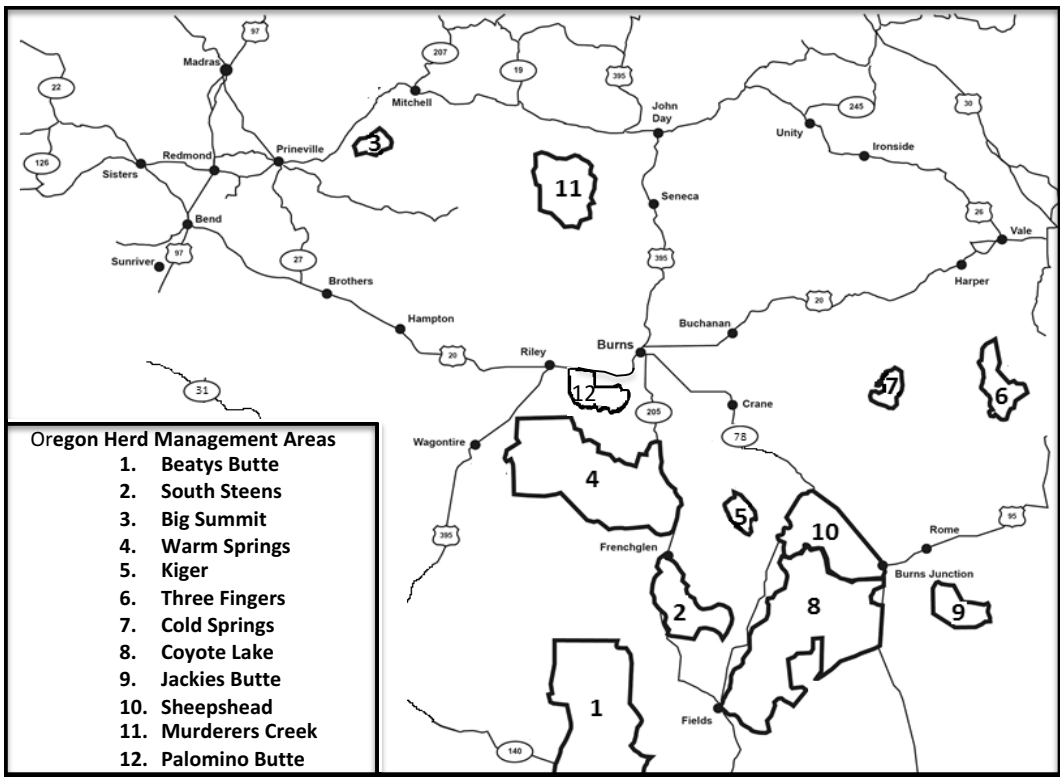

Fig. 1. Geographical locations of the Oregon herd management areas (HMAs) used in this study. Adapted from https://www.blm.gov/or/resources/whb/files/or_state_map_HMA.pdf. Sites included in the study were Beatys Butte (1), South Steens (2), Big Summit (3), Warm Springs (4), Kiger (5), Jackies Butte (9), and Murderers Creek (11). The numbers in the figure represent the HMA identifiers in Oregon.

Western breeds. Because of the wild horse's iconic value as a reminder of America's West, the Wild Free-Roaming Horses and Burros Act of 1971 and 1978 mandated the Bureau of Land Management (BLM) to protect, manage, and control wild horses and burro populations on federal rangeland. Considered an integral part of the existing natural ecosystems on the public lands under the act, these feral horses are protected from indiscriminate capture, branding, harassment, and killing. The BLM management goals aim to achieve the appropriate management level (AML), which is the number of feral horses and burros that the BLM determines can exist in balance with other rangeland species, resources, and uses in a given area, by employing methods of random capture and removal of horses (http://www.blm.gov/or/resources/whb/herd-manage.php). These approaches often overlook the close relationships of the active breeding gene pool within herds and may be putting existing sequestered herds at risk of a local population crash. Decreased genetic diversity can subsequently lead to inbreeding, lower fitness, and ultimately to herd extinction. It is therefore vital that biologists choosing management strategies for feral horses sequestered on public rangelands consider employing genetic analyses so as to maintain a healthy, diverse gene pool (Beauclerc et al. 2010) at an optimal carrying capacity for the rangelands. Within the State of Oregon, 19 herd management areas (HMAs) have been established (Fig. 1), one of which, the Big Summit HMA, is under the direct management of the United State Department of Agriculture–Forest Service (USDA–FS). The others are managed by the BLM (or comanaged by the USDA–FS and BLM, such as Murderers Creek feral horse territory; https://www.blm.gov/or/resources/whb/files/MurderersCreekWHT.pdf. The herd

numbers in Oregon are increasing yearly by 20% on average.

The 2011 census indicated the Big Summit herd size to be approximately 85 individuals, which is significantly lower than the recommended minimum of 150–200 breeding individuals to maintain random mating (Cothran et al. 1998, Luís et al. 2007) within a natural, free-roaming population. In response to the 2011 census results, the Food and Agriculture Organization (FAO) classified the Big Summit HMA's feral horse population as "high–risk" for extinction (FAO 1998, Luís et al. 2007). The objective of the present study was to provide a genetic analysis using noninvasive collection methods (hair) to assess the genetic health of this small herd compared to horses within other HMAs. This study aimed to understand the patterns of major histocompatibility complex (MHC) gene-based adaptive variation in relation to the patterns of neutral genetic markers within feral equine populations of Oregon. The analyses included determination of genetic structure, distribution of observed genetic variation, phylogenetic analysis of HMA populations, and ancestral breed assignment from microsatellite alleles. Additionally, mtDNA analysis was carried out to reflect a species' population history. However, mtDNA provides information on maternal history and does not reflect the demographic events caused by biparental inheritance. Therefore, this study examined nuclear (short tandem repeats = STRs) and adaptive (MHC) markers in parallel with mtDNA sequences.

METHODS

Ethics Statement

The protocol was approved under the Institutional Animal Care and Use Committee (IACUC #15-001) of Florida International University. No specific permissions were required for sample collection. Hair samples were donated by horse owners and the USDA–FS, Prineville, Oregon, and were mailed to Florida International University.

Ochoco National Forest

The Big Summit HMA, within the Ochoco National Forest located in central Oregon, was the primary area of study. The rangeland covers approximately 11,050 ha of high desert valleys surrounded by conifer-covered mountains. The elevation ranges from 1200 to 2200 m, and the area is subject to cold, snowy winters and hot summers. The overall forest terrain with steep ravines and cliffs hinders traditional aerial and ground census surveys. Severe winters with low survivorship (Central Oregon Wild Horse Coalition personal communication), the encroachment of human development, and the fencing off of adjacent lands have further impacted this herd and eliminated any natural migration corridors between neighboring HMAs for several generations.

Sample Collection

A total of 73 individual (mane) hair samples were obtained. "Known" samples were furnished by owners of captive Big Summit horses (24) and by owners of other captive feral horses (Jackies Butte [6], South Steens [7], Warm Springs [5], Beatys Butte [4], Kiger [6], and Murderers Creek [9]). Other samples were collected from hairs (with follicles and intact roots) that were left behind on the bark of pine trees or on fences within the Ochoco Forest. These "unknown" samples (12) were representative of the current "feral" gene pool of the Big Summit herd. Because samples were collected over a period of time, the mtDNA analysis does not include the unknown samples. The published neutral STR genotypic data of 19 domestic equine breeds (Lusitano [43], Konik [50], Irish Cob [50], Welsh [50], Dartmore [22], Shire [29], Tennessee Walker [23], Andalusian [50], Fell [50], Connemara [40], Standardbred [50], Appaloosa [50], Fjord [50], Hackney [50], Thoroughbred [50], Icelandic [50], Dutchdraft [50], Haflinger [50], and Kaspian [17]) documented in Van de Goor et al. (2011) were used for STRUCTURE analysis and the phylogenetic tree.

DNA Extraction

Genomic DNA was extracted from hair samples using the hair protocol provided by the QIAamp® DNA Mini Kit (Qiagen, Germantown, MD). Five to 10 hair strands were used for samples originating from known horses. One hair root was used for unknown sources (those obtained from trees and fences) to ensure that the DNA isolated was from a single horse. DNA was quantified using the Qubit® dsDNA HS assay kit on the Qubit® 2.0 Fluorometer (Life Technologies, Grand Island, NY). The average DNA yield was 5–20 ng, depending on the number and quality of hair roots extracted.

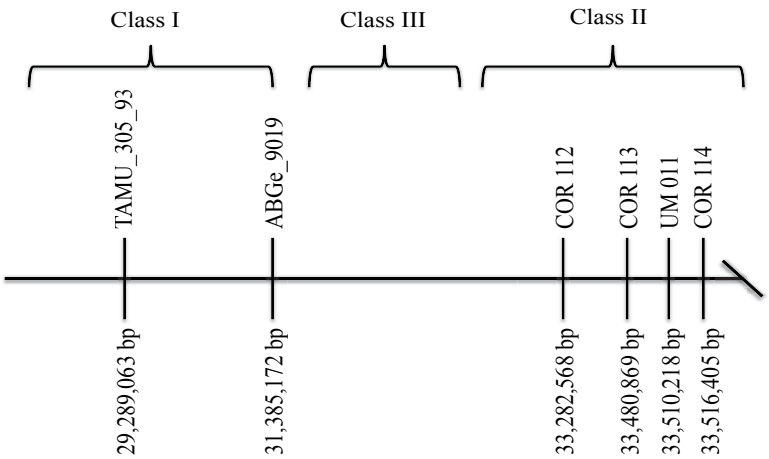

Fig. 2. Schematic representation of horse major histocompatibility complex (MHC) microsatellite loci located on ECA20. The slash (\) indicates the centromere end. The 6 microsatellite loci included in the study are shown. Positions of loci are indicated in base pairs (bp).

Neutral microsatellite amplification and fragment analysis.—The DNA extracts were amplified using StockMarks® for Horse Equine Genotyping Kit (Applied Biosystems, Foster City, CA) per the manufacturer's protocol, and STRs were separated using an ABI Prism 310 Capillary Genetic Analyzer (Applied Biosystems, Foster City, CA) and analyzed by GeneMapper® Software Version 4.0 (Applied Biosystems, Foster City, CA). Allele sizes were assigned a letter score following the International Society of Animal Genetics (ISAG) guidelines (Gill et al. 1994, 1997, Bär et al. 1997, Schneider et al. 1998) and were later converted to allele repeat numbers based on the updated equine typing standards (Van de Goor et al. 2010). The ISAG guidelines allow the comparison of genotypes across various laboratories and eliminate binning errors. Domestic horses that had been DNA-typed at an external laboratory according to the American Quarter Horse Association (AQHA) registration requirements were used as positive controls, and these external DNA-typed profiles, as well as the internal DNA standard within the kit, were used to verify all allele calls for this study.

MHC microsatellite amplification.—DNA was amplified using the MHC microsatellites, depicted in Fig. 2, following published methods (Tseng et al. 2010, Brinkmeyer-Langford et al. 2013). These included TAMU_305_93 and ABGe_9019, belonging to Class I, and COR113, COR112, COR114, and UM011, belonging to Class II (Tseng et al. 2010, Brinkmeyer-Langford et al. 2013). Fragment analysis was achieved using capillary electrophoresis on an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA) and GeneMapper® Software Version 4.0 (Applied Biosystems, Foster City, CA).

mtDNA amplification and sequencing.—The DNA was amplified using mtDNA D-loop HVR1 region-specific equine primers (forward: 5′-CTA GCT CCA CCA TCA ACA CC-3′, reverse: 5′-ATG GCC CTG AAG AAA GAA CC-3′), which amplified a 410-bp region. Each 20-μL PCR reaction contained 10–50 ng DNA, 0.5 μM each primer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, and 0.6 U GoTaq® Hot Start Polymerase (Promega, Madison, WI). The PCR cycles consisted of a denaturation step of 95 °C for 2 min, followed by 30 cycles of 95 °C for 1 min, 60 °C for 45 s, and 72 °C for 1 min, with a final extension of 72 °C for 5 min on the C1000 Touch™ Thermal Cycler (Bio-Rad, Hercules, CA). The PCR products were purified using ExoSAP-IT® (USB, Cleveland, OH) and then sequenced using the Big Dye® Terminator version 3.1 sequencing kit (Applied Biosystems, Carlsbad, CA). Products were purified with 75% isopropanol and dried down at 80 °C for 1 min. Hi-Di™ formamide (Applied Biosystems, Carlsbad, CA) (12.5 μL) was added and the products were subsequently denatured at 95 °C for 2 min, loaded onto an ABI Prism® 310 Genetic Analyzer (Applied Biosystems,

Carlsbad, CA), separated, and analyzed. DNA sequences were aligned using the software ClustalW (European Bioinformatics Institute, Cambridge, United Kingdom).

Statistical Analysis

NEUTRAL AND MAJOR HISTOCOMPATIBILITY COMPLEX (MHC) MICROSATELLITE DATA.—Statistical analyses were performed on 14 neutral loci (Canon et al. 2000, Luís et al. 2006, Monies et al. 2011) and 6 MHC loci. Allele frequencies, number of alleles, inbreeding coefficient (F_{IS}), and observed (H_{Obs}) and expected (H_{Exp}) heterozygosity were calculated using GenAlex (Peakall and Smouse 2006, 2012) and GenePop (Raymond and Rousset 1995). GenAlex was used to evaluate pairwise genetic differentiation between populations and departure from Hardy–Weinberg equilibrium (HWE), using chi-square analysis and sequential Bonferonni correction on loci. Analysis of molecular variance (AMOVA) was performed to determine the amount of genetic variation attributable to within- and between-populations. Based on the 14 neutral STRs, individual relationships within the herd were estimated through ML-Relate (Kalinowski et al. 2006), which predicts the degrees of relatedness as parent–offspring (PO), full sibling (FS), half sibling (HS) or unrelated (U). Kinship likelihoods were calculated with 95% confidence (Kalinowski et al. 2006b, Wagner et al. 2006).

MTDNA SEQUENCES.—The genetic variability among the sequences was measured through haplotype and nucleotide diversity using the program Arlequin version 3.5.2.2 (Excoffier and Lischer 2010) (http://cmpg.unibe.ch/software/arlequin35/). The probability of identity (P_{ID}) was calculated as the sum of the square of the frequency of each haplotype in that population ($P_{ID} = \sum q_k^2$, where q_k is the frequency of the kth haplotype) (Hill et al. 2002, McGahern et al. 2006, Álvarez et al. 2012) and where P_{ID} is the likelihood that 2 individuals in the population share the same haplotype. In order to investigate population differentiation among the different HMAs studied, the population pairwise F_{ST} values were computed under the analysis of molecular variance (AMOVA) also using Arlequin version 3.5.2.2 (Excoffier and Lischer 2010). The statistical significance of the values was estimated through 1000 permutations.

Phylogenetic Analysis

NEUTRAL MICROSATELLITE DATA.—To construct phylogenetic relationships, an unweighted pair group method with arithmetic mean (UPGMA) tree was constructed with program POPTREE2 (Takezaki et al. 2010) based on F_{ST} genetic distance measures from allele frequency and 1000 bootstrap iterations.

MTDNA SEQUENCES.—The sequences were truncated to 388 bp, between the positions of 15459 and 15846 np, in accordance with the published nucleotide position numbering that follows Xu and Árnason (1994) (GenBank X79547), the mtDNA reference sequence of the horse. A maximum parsimony tree was constructed based on the comparisons of different sequence haplotypes that were identified using the software Network 4.611 (Fluxus Technology Ltd, Clare, UK; http://www.fluxus-engineering.com/sharenet.htm) (Bandelt et al. 1999).

POPULATION SUBSTRUCTURE.—Model-based clustering analysis using STRUCTURE 2.3.4 (Pritchard et al. 2000) was used (1) to determine the admixture from ancestral breeds of the feral horses in the HMAs, (2) to assign unknown hair samples to the most probable HMA, and (3) to assess the population structure of the HMAs based on the neutral microsatellites. An admixture model with correlated allele frequencies was adopted without prior delineation of populations to account for recent gene flow between breeds that share major alleles. The parameter for individual admixture alpha was set to be the same for all simulations with uniform prior probability distribution. Distinct populations (K) were estimated using 150,000 burn-ins, 350,000 Markov chain Monte Carlo repetitions, and 3 independent runs for each K (1–30) (Evanno et al. 2005). The optimal K was then calculated using an ad hoc quantity (ΔK) and estimated as $K = 7$ for 19 domestic breeds and the 7 individual HMA populations.

RESULTS

The unknown samples were verified to be from the Big Summit HMA based on F_{ST} genetic distance and known collection site within the Big Summit HMA. Subsequently, further microsatellite analysis assigned a total of 33 individuals (24 Big Summit horses and 9 unknown horses) to the Big Summit population.

TABLE 1. Statistical results and indices for the Big Summit HMA based on 14 microsatellite loci and 6 MHC loci. Ne = The number of effective alleles = $1/(\sum p_i^2)$, where p_i is the frequency of the ith allele at a locus; HWE = Hardy–Weinberg equilibrium; F_{IS} = fixation index (inbreeding coefficient); H_{Obs} = observed heterozygosity; H_{Exp} = expected heterozygosity

	Alleles	Ne	H_{Obs}	H_{Exp}	HWE P value	Significance	F_{IS}	Private alleles	Frequency
Neutral loci									
AHT4	8	4	0.64	0.76	0.52	NS	0.16	30	0.16
HMS7	10	3	0.55	0.70	0.00	$P < 0.001$	0.22	15	0.02
								24	0.02
HTG4	6	3	0.39	0.70	0.00	$P < 0.001$	0.44	—	—
VHL20	6	5	0.55	0.78	0.10	NS	0.30	—	—
AHT5	9	4	0.67	0.75	0.19	NS	0.11	22	0.03
								25	0.02
HMS6	8	3	0.42	0.71	0.00	$P < 0.001$	0.40	13	0.48
HMS3	10	5	0.58	0.81	0.00	$P < 0.001$	0.29	22	0.03
								23	0.02
HTG10	10	4	0.58	0.77	0.00	$P < 0.001$	0.25	25	0.12
								28	0.06
								29	0.02
ASB17	9	6	0.64	0.84	0.00	$P < 0.001$	0.24	28	0.06
CA425	7	3	0.52	0.67	0.00	$P < 0.001$	0.23	26	0.03
								29	0.02
ASB23	8	5	0.42	0.79	0.00	$P < 0.001$	0.46	24	0.02
								31	0.05
HMS2	6	4	0.67	0.72	0.00	$P < 0.001$	0.08	25	0.03
ASB2	6	4	0.67	0.78	0.03	$P < 0.05$	0.14	16	0.24
								19	0.02
HTG7	6	3	0.45	0.63	0.00	$P < 0.001$	0.28	21	0.03
AVERAGE			0.55	0.74			0.26		
TOTAL	109								
MHC loci									
MHC I 305	5	3	0.67	0.62	0.65	NS	−0.07		
MHC I 168	9	5	0.92	0.79	0.00	$P < 0.01$	−0.16		
MHC II COR113	9	7	0.88	0.86	0.00	$P < 0.001$	−0.02		
MHC II COR112	7	4	0.63	0.76	0.03	$P < 0.05$	0.17		
MHC II COR114	6	4	0.75	0.76	0.70	NS	0.01		
MHC II UM011	7	4	0.75	0.76	0.28	NS	0.01		
AVERAGE			0.76	0.76			−0.01		
TOTAL	43								

TABLE 2. Mean statistical results and indices for 7 herd management area populations based on 14 microsatellite loci. A high F_{IS} is representative of inbreeding. H_{Obs} = observed heterozygosity, H_{Exp} = expected heterozygosity, F_{IS} = inbreeding coefficient.

Population	*N* (individuals)	H_{Obs}	H_{Exp}	F_{IS}
Big Summit	33	0.55	0.74	0.26
Jackies Butte	6	0.69	0.75	0.07
Murderers Creek	9	0.69	0.73	0.04
Beatys Butte	4	0.63	0.68	0.08
Warm Springs	5	0.64	0.69	0.06
South Steens	7	0.77	0.73	−0.04
Kiger	6	0.55	0.61	0.09

TABLE 3. Results of *F* statistics over Big Summit and 6 HMA populations for each locus. A negative F_{IS} value indicates excess of heterozygosity, while a positive value indicates a trend toward loss of heterozygosity. F_{ST} values closer to 0 indicate little genetic differentiation, while values 0.25 to 1 indicate greater genetic differentiation.

F_{IS} = (mean H_{Exp} − mean H_{Obs})/mean H_{Exp}
F_{IT} = (H_t − mean H_{Obs})/H_t
F_{ST} = (H_t − mean H_{Obs})/H_t

Mean H_{Exp} = average H_e across the populations, Mean H_{Obs} = average H_o across the populations; H_t = total expected heterozygosity

Locus	F_{IT}	F_{IS}	F_{ST}
AHT4	0.074	−0.050	0.118
HMS7	0.031	−0.080	0.103
HTG4	0.243	0.167	0.092
VHL20	0.153	0.009	0.146
AHT5	0.070	−0.036	0.102
HMS6	0.333	0.203	0.163
HMS3	0.202	0.138	0.075
HTG10	0.108	−0.019	0.124
ASB17	0.403	0.306	0.140
CA425	0.342	0.224	0.153
ASB23	0.278	0.195	0.103
HMS2	0.308	0.041	0.279
ASB2	0.283	0.127	0.179
HTG7	0.077	−0.033	0.106

Polymorphisms

NEUTRAL MICROSATELLITES.—A total of 109 alleles were detected across 14 microsatellite markers in the Big Summit herd. The mean number of alleles per locus ranged from 6 to 10. Statistical results for H_{Obs}/H_{Exp} and deviation from Hardy–Weinberg equilibrium (HWE) at each locus for Big Summit and other HMA populations show reduced heterozygosity in the Big Summit populations (Tables 1, 2). The expected heterozygosity for the Big Summit population was higher than the observed heterozygosity for all the markers tested. Except for AHT4, VHL20, and AHT5, 11 other markers deviated significantly from HWE ($P < 0.001$). The inbreeding coefficient F_{IS} for each marker in the Big Summit herd ranged between 0.08 for HMS2 and 0.46 for ASB23 (Table 1). The overall Big Summit inbreeding coefficient was 0.26, and the Big Summit population demonstrated higher homozygosity compared to other HMA populations (Table 2).

Allele frequencies for the Big Summit HMA (Supplementary Material 1) demonstrated a broad range of alleles for all 14 loci compared to the domestic equine breeds (Van de Goor et al. 2011). The HMA populations had a greater number of alleles that were found only in the feral and not in the domestic populations, with the isolated Big Summit HMA harboring the majority of the private alleles (Table 1).

MAJOR HISTOCOMPATIBILITY COMPLEX (MHC) MICROSATELLITES.—A total of 43 alleles were detected across 6 MHC loci within the Big Summit Herd. Contrary to the statistical results of the neutral markers for H_{Obs}/H_{Exp} and deviation from HWE, the Big Summit population showed optimal heterozygosity compared to the expected heterozygosity at the MHC loci studied (Table 1). Additionally, the inbreeding coefficient of −0.01 displays a trend of higher polymorphism being maintained at the MHC loci compared to the neutral markers.

POPULATION DIFFERENTIATION.—The values for F_{IT}, F_{ST}, and F_{IS}, which denote the inbreeding coefficient (I) relative to the total (T) population, the effect of subpopulations (S) compared to T, and I relative to S, respectively, are shown in Table 3. Considering how F_{ST} is calculated, positive values indicate some degree of variation in the allele frequencies across subpopulations. The F_{ST} values (Table 3) exhibit moderate genetic variation in the allele frequencies across populations. The variability at the group level between the 7 populations assessed by population pairwise F_{ST} values for the mtDNA is summarized in Table 4.

Table 4. Population pairwise F_{ST} values (below diagonal) and the P value significance of F_{ST} (above diagonal) between 7 populations for the mtDNA. A high value is indicative of a lesser degree of gene flow with complete differentiation among the populations. A plus symbol (+) indicates $P < 0.05$, and NS indicates no statistical significance.

	Big Summit	Kiger	Murderers Creek	Warm Spring	Jackies Butte	Beatys Butte	South Steens
Big Summit	0	+	+	+	+	+	+
Kiger	0.424	0	+	NS	NS	+	NS
Murderers Creek	0.237	0.190	0	NS	+	+	+
Warm Spring	0.341	−0.018	0.028	0	NS	NS	NS
Jackies Butte	0.463	0.208	0.151	−0.062	0	NS	NS
Beatys Butte	0.609	0.325	0.288	0.176	0.231	0	NS
South Steens	0.550	0.150	0.264	0.124	0.231	0.097	0

Genetic diversity and relationships.—Genetic relationship data for the captive Ochoco herd members ($n = 33$), examined using ML-Relate, revealed 4 parent–offspring relations, 14 full siblings, and 49 half siblings. The mtDNA genetic diversity estimates found within the 7 populations studied, including haplotype diversity (h), nucleotide diversity (π), and the probability index (Table 5), indicated that Warm Springs and Beatys Butte had the most diverse populations ($h = 1.00 \pm 0.13$ and $h = 1.00 \pm 0.18$, respectively), even with only 4 horses sampled. These HMAs exhibited one haplotype per horse sampled in each population. Big Summit horses had the most conserved mtDNA haplotypes ($h = 0.51 \pm 0.04$) with 24 horses sampled. The analysis of the nucleotide diversity revealed that Warm Springs and Murderers Creek exhibited the most variation ($\pi = 0.026 \pm 0.017$ and $\pi = 0.024 \pm 0.014$, respectively). On the other hand, Beatys Butte and South Steens ($\pi = 0.008 \pm 0.006$ and $\pi = 0.009 \pm 0.006$, respectively) demonstrated very limited nucleotide diversity, though fewer horses from these HMAs were sampled. The estimated P_{ID} of the haplotypes identified ranged from 0.20 to 0.51 for the HMAs in Oregon and was 0.51 for the Big Summit HMA.

Phylogenetic analysis.—The phylogenetic trees were constructed using 14 loci on Nei et al. (1983) genetic distances and an unrooted neighbor joining (NJ) to compare the 7 HMAs in Oregon to 19 domestic breeds (Nei et al. 1983). The phylogenetic tree indicated 5 distinct clusters of related breeds (Fig. 3). The cold-blooded breeds Dutch Draft and Haflinger grouped together, the warm-blooded breeds like the Thoroughbred grouped closer to the Iberian breeds, while the Konik, Fjord, Appaloosa, Tennessee Walker, and Icelandic breeds formed a separate branch. Although the feral horses branched separately from the domestic breeds, the Big Summit HMA formed a separate cluster from the 6 HMA and domestic breeds. However, the feral horses from Oregon appear to cluster closer to Spanish breeds than to other domestic breeds.

A median-joining network tree was constructed for mtDNA according to the identified haplotypes and significant clusters (Fig. 4). There was a clear differentiation of all haplogroups assigned, evidenced by their apparent clustering. Haplogroup I was exclusively composed of horses from the Big Summit HMA. Almost all other haplogroups had no clear geographic affiliation and were composed of a mixture of horses from different HMAs, with the exception of haplogroups Q1 and Q3 (Fig. 4).

Population substructure.—Cluster analysis performed on the 7 HMAs (Fig. 5) showed distinct groupings that support the phylogenetic tree (Fig. 3). Dissimilar populations are divided by a black line and are numbered using the representative domestic breeds and HMAs (Fig. 5). Each individual is a column divided into K colors, each color representing a breed cluster. In concurrence with the phylogenetic tree, the Icelandic, Fjord, and Konik group together; Fell, Irish Cob, and Shire group as one cluster; the Iberian breeds, Andalusian and Lusitano, group together; and the HMA populations group distinctively from the domestic breeds. Within the HMAs, 2 clusters are seen in the Big Summit HMA and the other 6 HMAs as subpopulations. The HMAs as a whole showed significant ancestral contributions from over 5 domestic breeds, as observed with the presence of various colors (Fig. 5). The major color (light blue), corresponding to the Andalusian and Lusitano breeds, was seen in Big Summit HMA, followed by orange, dark blue, and others corresponding

TABLE 5. Genetic diversity indices of mtDNA lineages. h = haplotype diversity, π = nucleotide diversity, P_{ID} = probability of identity. Standard errors are in parentheses.

HMA subpopulations	Individuals	Haplotypes	Haplotype diversity (h)	Nucleotide diversity (π)	Probability of identity (P_{ID})
Big Summit	24	2	0.51 (0.04)	0.012 (0.007)	0.51
Warm Springs	5	5	1.00 (0.13)	0.026 (0.017)	0.20
Jackies Butte	6	3	0.73 (0.16)	0.017 (0.011)	0.39
Murderers Creek	9	6	0.89 (0.09)	0.024 (0.014)	0.21
Beatys Butte	4	4	1.00 (0.18)	0.008 (0.006)	0.25
Kiger	6	2	0.60 (0.13)	0.019 (0.012)	0.50
South Steens	7	3	0.67 (0.16)	0.009 (0.006)	0.43

to Standardbred, Thoroughbred, Appaloosa, Icelandic and other breeds. The pink color in the group of 6 HMAs is similar to that of the Konik breed.

DISCUSSION

The horses inhabiting the Ochoco National Forest are considered feral but are managed as "wildlife" by the USDA–FS. They are no doubt recent descendants of domesticated horses that helped settle the western USA and either escaped captivity or were emancipated by their owners during the settling of America. However, these horses are unlike most domestic breeds, which are often bred under heavily managed animal husbandry practices, with many breeds specifically linebred for certain characteristics. Previous genetic analysis (Cothran 2011) of 12 horses from Big Summit HMA using 12 microsatellites indicated inbreeding within the herd as well as a strong ancestral lineage to the Spanish Andalusian breed, the North American gaited breeds, and Arabian breeds (www.blm.gov/or/districts/burns/wildhorse/). The results of the present study also showed a strong ancestral contribution to Big Summit horses from Iberian breeds, specifically Andalusian and Lusitano breeds, supporting the Cothran (2011) study. Although contributions from other modern breeds were seen across all the Oregon HMAs, it was interesting to see a strong correlation to the semi-feral Polish Konik breed within the 6 other HMA populations—a population of horses not associated with the settling of the American West. This diverse admixture association may be a better indicator of the lack of human intervention and the ability to randomly mate than it is of the Konik contribution via bloodlines. In other words, if a population has a more diverse breeding pool plus an opportunity for mate selection via random mating, that could be reflected in the ancestral diversity seen in some HMA populations. In theory, given enough geographic area to mate randomly, the wild horses should be more genetically diverse as they will inherently select mates based on dissimilarities—nature's method of preventing inbreeding. However, when herds are sequestered like the Big Summit herd, the scenario is set up for nonrandom mate selection and unmanaged (without human intervention) breeding strategies for the herd.

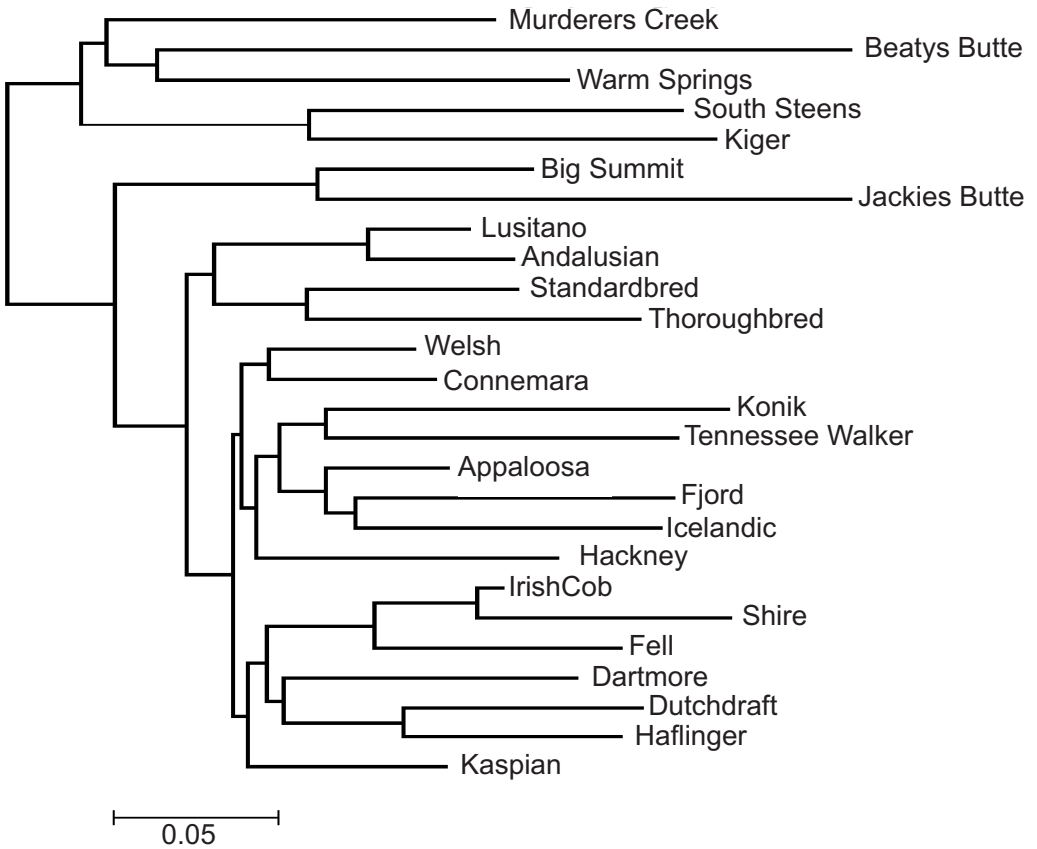

Fig. 3. Phylogenetic tree of HMA and domestic equine breeds. This neighbor-joining dendrogram was constructed in POPTREE2 from Nei's D_a distances (Nei et al. 1983) and shows genetic relationships across 7 Oregon herd management areas (Big Summit, Jackies Butte, South Steens, Warm Springs, Beatys Butte, Kiger, and Murderers Creek) in relation to 19 known equine breeds. Hair samples from Oregon HMAs were collected from 2008 to 2011 to increase the sample size for analysis.

When mate choice is limited to only close relatives, inbreeding is the genetic outcome. When domestic horses are line bred and undesirable traits appear, these traits can be controlled by outcrossing or removing those individuals, and thus the undesirable traits, from a herd. This control is not possible in a small, unmanaged wild horse herd like the Big Summit herd. Therefore, because of the contribution from diverse ancestral breeds and the lack of human intervention or selective breeding practices as compared to domestic breeds, one would expect the Oregon feral horses to be more genetically diverse.

However, reproductive isolation, reduced herd sizes, geographical and physical barriers to migration corridors, and human settlement have sequestered the Big Summit herd on an "island" of land. The small Big Summit population appears to have been trapped in a fragmented habitat, unable to migrate and breed outside of their local gene pool for several generations, mimicking the inbreeding seen in island populations. While the mating behavior of free-roaming horses should prevent them from breeding with close relatives (Berg 1986), within a small population horses have a higher probability of encountering close relatives. The Big Summit HMA analysis in this study displayed a large number of full and half siblings, indicative of close relatedness coupled with an apparent lack of recent gene flow between neighboring HMAs. However, given the small sample sizes, the horse's hierarchal social structure, and the serendipitous capture methods used, these results may be reflective of

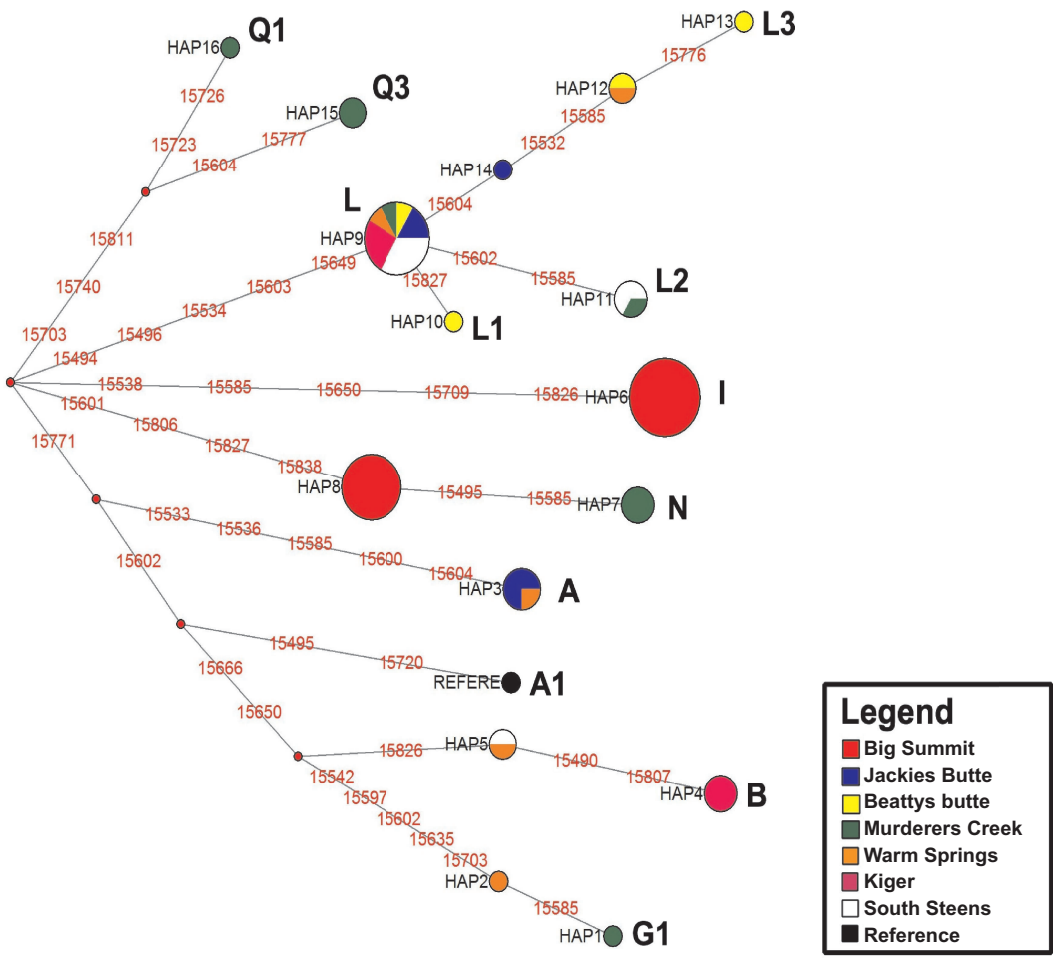

Fig. 4. Median-joining tree of wild horse mtDNA sequences. Haplogroups and haplotypes are designated accordingly. Circles are proportional to the number of horses they represent. Pie slices within circles indicate the herd management area from which the horses came.

existing family groups being captured together. Nevertheless, the overall fixation index of 0.26 and deviation from Hardy–Weinberg equilibrium at neutral loci supported the conclusion that inbreeding is rising because of nonrandom mating and limited mate selection in the Big Summit HMA. However, the overall fixation index of −0.01 at the nonneutral loci may be affected by the highly variable and polymorphic nature of the MHC genes. Population pairwise F_{ST} values based on mtDNA data provided additional evidence that the Big Summit horses are experiencing restricted gene flow and show little contribution from other HMA herds. Similar findings have been seen in studies on Spanish Celtic horse breeds (Canon et al. 2000), Portuguese horse breeds (Luís et al. 2007), and Iberian breeds (Luís et al. 2007) where distinct genetic differentiation and partition of the genetic variability and structure was observed within breeds (Kavar and Dovc 2008). Conversely, genetic data cannot fully resolve geographic dispersal or whether or not physical migration corridors between herds prior to the initiation of the Act of 1971 and human encroachment were present. It is crucial to understand that analysis cannot definitively resolve whether the shared alleles are common by descent or common by location. A number of private alleles were found in the Big Summit HMA, although this may be due to a larger sample size from the Big Summit population compared to other HMA populations. Notably, a similar pattern has been

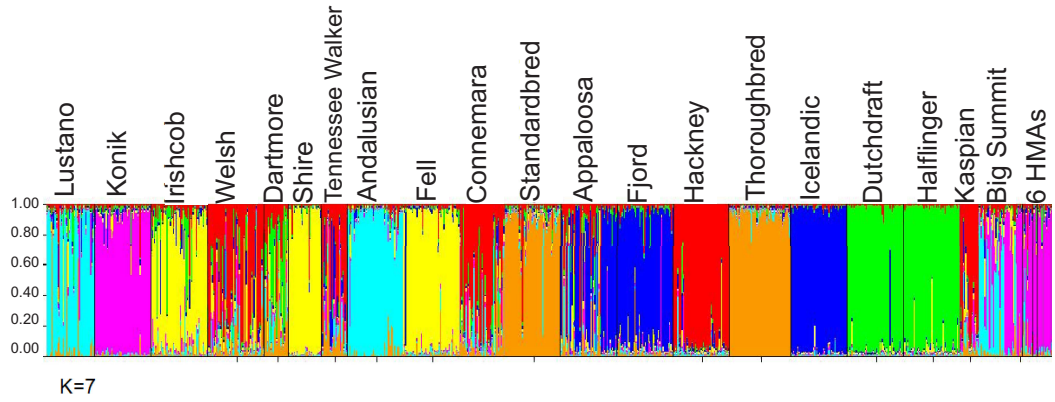

Fig. 5. Cluster analysis for herd management areas (HMA) and 19 domestic equine breeds. Clustering for 7 Oregon HMAs and 19 domestic breeds with K = 7 inferred clusters (colors) was estimated by STRUCTURE 2.3.4 (Pritchard et al. 2000). A contribution of Spanish breeds (Andalusian and Lusitano domestic breeds) was seen in Big Summit, as well as Konik horses. Similarities to the Konik semiferal European (Poland), Standardbred, and Thoroughbred horses were seen in other HMAs.

seen in feral horses isolated and protected on Sable Island, Canada, where a large number of breed-specific alleles were noted (Plante et al. 2007).

Evidence is mounting in conservation genetics that small insular populations, analogous to some of the Oregon HMAs, are greatly impacted by inbreeding and could be candidates for genetic restoration (Nei et al. 1975, Ingvarsson 2001, Tallmon et al. 2004). Research shows that bottlenecks, founder effects, and genetic drift increase the risk of a decline in genetic variation and result in a decrease in heterozygosity at numerous loci. A cumulative loss of genetic variation can lead to the expression of deleterious alleles, which may become fixed in small populations such as the Big Summit herd (Nei et al. 1975, Westemeier et al. 1998). The loss of genetic variation, in turn, decreases fitness and increases the probability of extinction of small isolated populations. Inbreeding seems to already be impacting these horses, as they appear more susceptible to diseases and have more physical deformities. Additionally, the herd suffers from high infant morbidity and mortality (USDA–FS personnel personal communication).

Other studies have demonstrated that destruction of habitat, sequestration, and fragmentation of populations subsequently led to a bottleneck resulting in inbreeding and limited gene diversity (Frantzen et al. 2001, Ouborg et al. 2010). In the Losina breed from Burgos, the loss of heterozygosity and increased homozygosity displayed a trend away from Hardy–Weinberg equilibrium. The authors concluded that the significant deviation from HWE for 3 loci (HTG10, HMS3, and AHT4; $P < 0.01$) could be a direct consequence of a small breeding population and differentiation of the breed over time (Canon et al. 2000). Other studies have implied that a substantial loss of genetic diversity reduces fitness, limits gene flow into the populations, and in turn alters the adaptive capacity of a population (Bryant et al. 1999, Beauclerc et al. 2010).

Genetic diversity is an essential part of any functioning ecosystem (Croteau et al. 2012). A thorough understanding of what influences patterns of genetic diversity and population structure is essential for managing populations of wide-ranging species, such as the feral horses (Croteau et al. 2012). Providing genetic data on these Oregon herds will allow informed management decisions where inbreeding can be minimized and genetic diversity maximized. Integrating scientific analyses with current management strategies is critical for saving these small herds of feral horses. The ultimate goal for current and future management, whether on a large or small scale, should always place its focus on maintaining the health of the herds, gene flow, and the highest form of genetic diversity.

SUPPLEMENTARY MATERIAL

One online-only supplementary file accompanies this article (scholarsarchive.byu.edu/wnan/vol79/iss1/9).

SUPPLEMENTARY MATERIAL 1. Allele frequencies observed in the Big Summit Herd Management Area (33 individuals) for 14 microsatellite loci.

ACKNOWLEDGMENTS

The authors thank the USDA–FS Prineville, Oregon, and their personnel for collecting hair samples from the Big Summit range. Funding for this project was supported in part by the USDA–Forest Service: Challenge Cost-share Agreement #NFS 09-CS-11060701-008 to DM. We are grateful to members of the Central Oregon Wild Horse Coalition, and especially Gayle Hunt, for providing hair samples from the feral horses representing HMAs in Oregon. A special thanks goes to the International Forensic Research Institute at Florida International University for use of the Forensic DNA Profiling Facility. This work represents the point of view of the authors and may not be the official view of the USDA–FS. The authors have no competing interests.

LITERATURE CITED

ÁLVAREZ, I., I. FERNÁNDEZ, L. LORENZO, L. PAYERAS, M. CUERVO, AND F. GOYACHE. 2012. Founder and present maternal diversity in two endangered Spanish horse breeds assessed via pedigree and mitochondrial DNA information. Journal of Animal Breeding Genetics 129:271–279.

BANDELT, H.J., P. FORSTER, AND A. RÖHL. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16:37–48.

BÄR, W., B. BRINKMAN, B. BUDOWLE, A. CARRACEDO, P. GILL, P. LINCOLN, W. MAYR, AND B. OLAISEN. 1997. DNA recommendations—further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. Forensic Science International 87:179–184.

BEAUCLERC, K.B., B. JOHNSON, AND B.N. WHITE. 2010. Genetic rescue of an inbred captive population of the critically endangered Puerto Rican crested toad (*Peltophryne lemur*) by mixing lineages. Conservation Genetics 11:21–32.

BERG, W.J. 1986. Effective population size estimates and inbreeding in feral horses: a preliminary assessment. Journal of Equine Veterinary Science 6:240–245.

BRINKMEYER-LANGFORD, C.L., J.J. CAI, C.A. GILL, AND L.C. SKOW. 2013. Microsatellite variation in the equine MHC. Animal Genetics 44:267–275.

BRYANT, E.H., V.L. BACKUS, M.E. CLARK, AND D.H. REED. 1999. Experimental tests of captive breeding for endangered species. Conservation Biology 13:1487–1496.

[BLM] BUREAU OF LAND MANAGEMENT. 2011. Oregon BLM wild horse and burro population data [online]. Report date: 6 April 2011. U.S. Department of the Interior, Bureau of Land Management; [accessed 20 March 2013]. http://www.blm.gov/or/resources/whb/files/2011foalcrop.pdf

[BLM] BUREAU OF LAND MANAGEMENT. 2014. Herd Management Areas Oregon/Washington [online]. U.S. Department of the Interior, Bureau of Land Management. [Accessed 12 March 2014]. http://www.blm.gov/or/resources/whb/herd-manage.php

CANON, J., M.L. CHECA, C. CARLEOS, J.L. VEGA-PLA, M. VALLEJO, AND S. DUNNER. 2000. The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. Animal Genetics 31:39–48.

COTHRAN, E.G. 2011. Report to USDA–FS: Genetic analysis of the Big Summit HMA, OR. 8 April 2011.

COTHRAN, E.G., S.A. SANTOS, M.C.M. MAZZA, T.L. LEAR, AND J.R.B. SERENO. 1998. Genetics of the Pantaneiro horse of the Pantanal region of Brazil. Genetics and Molecular Biology 21(3). https://dx.doi.org/10.1590/S1415-47571998000300009.

CROTEAU, E.K., E.J. HEIST, C.K. NIELSEN, J.R. HUTCHINSON, AND E.C. HELLGREN. 2012. Microsatellites and mitochondrial DNA reveal regional population structure in bobcats (*Lynx rufus*) of North America. Conservation Genetics 13:1637–1651.

EVANNO, G., S. REGNAUT, AND J. GOUDET. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611–2620.

EXCOFFIER, L., AND H.E.L. LISCHER. 2010. Arlequin Suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10:564–567.

[FAO] FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. 1998. Secondary guidelines for development of national farm animal genetic resources management plans. Management of Small Populations at Risk. Initiative for Domestic Animal Diversity. FAO, Rome, Italy.

FRANTZEN, M.A.J., J.W.H. FERGUSON, AND M.S. DE VILLIERS. 2001. The conservation role of captive African wild dogs (*Lycaon pictus*). Biological Conservation 100:253–260.

GILL, P., B. BRINKMANN, E. D'ALOJA, J. ANDERSEN, W. BAR, A. CARRACEDO, B. DUPUY, B. ERIKSEN, M. JANGBLAD, V. JOHNSSON, ET AL. 1997. Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Science International 87:185–192.

GILL, P., C. KIMPTON, E. DALOJA, J.F. ANDERSEN, W. BÄR, B. BRINKMANN, S. HOLGERSSON, V. JOHNSSON, A.D. KLOOSTERMAN, M.V. LAREU, ET AL. 1994. Report of the European DNA profiling group (EDNAP)—towards standardisation of short tandem repeat (STR) loci. Forensic Science International 65:51–59.

HILL, E.W., D.G. BRADLEY, M. AL-BARODY, O. ERTUGRUL, R.K. SPLAN, I. ZAKHAROV, AND E.P. CUNNINGHAM. 2002. History and integrity of thoroughbred dam lines revealed in equine mtDNA variation. Animal Genetics 33:287–294.

INGVARSSON, P.K. 2001. Restoration of genetic variation lost: the genetic rescue hypothesis. Trends in Ecology and Evolution 16:62–63.

KALINOWSKI, S.T., A.P. WAGNER, AND M.L. TAPER. 2006. ML-Relate: a computer program for maximum

likelihood estimation of relatedness and relationship. Molecular Ecology Notes 6:576–579.

Kavar, T., and P. Dovc. 2008. Domestication of the horse: genetic relationships between domestic and wild horses. Livestock Science 116:1–4.

Luís, C., C. Bastos-Silveira, E.G. Cothran, and M.M. Oom. 2006. Iberian origins of New World horse breeds. Journal of Heredity 97:107–113.

Luís, C., R. Juras, M.M. Oom, and E.G. Cothran. 2007. Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. Animal Genetics 38:20–27.

McGahern, A.M., C.J. Edwards, M.A. Bower, A. Heffernan, S.D.E. Park, P.O. Brophy, D.G. Bradley, D.E. MacHugh, and E.W. Hill. 2006. Mitochondrial DNA sequence diversity in extant Irish horse populations and in ancient horses. Animal Genetics 37:498–502.

Monies, D., N. Abu Al Saud, N. Sahar, and B.F. Meyer. 2011. Population studies and parentage testing for Arabian horses using 15 microsatellite markers. Animal Genetics 42:225–226.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29:1–10.

Nei, M., F. Tajima, and Y. Tateno. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. Journal of Molecular Evolution 19:153–170.

Ouborg, N.J., C. Pertoldi, V. Loeschcke, R.K. Bijlsma, and P.W. Hedrick. 2010. Conservation genetics in transition to conservation genomics. Trends in Genetics 26:177–187.

Peakall, R., and P.E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Resources 6:288–295.

Peakall, R., and P.E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537–2539.

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

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Prystupa, J.M., P. Hind, E.G. Cothran, and Y. Plante. 2012. Maternal lineages in native Canadian equine populations and their relationship to the Nordic and Mountain and Moorland pony breeds. Journal of Heredity 103:380–390.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248–249.

Schneider, H.R., S. Rand, H. Schmitter, and G. Weichhold. 1998. ACTBP2-nomenclature recommendations of GEDNAP. International Journal of Legal Medicine 111:97–100.

Takezaki, N., M. Nei, and K. Tamura. 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. Molecular Biology and Evolution 27:747–752.

Tallmon, D.A., G. Luikart, and R.S. Waples. 2004. The alluring simplicity and complex reality of genetic rescue. Trends in Ecology and Evolution 19:489–496.

Tseng, C.T., D. Miller, J. Cassano, E. Bailey, and D.F. Antczak. 2010. Identification of equine major histocompatibility complex haplotypes using polymorphic microsatellites. Animal Genetics 41:150–153.

Van de Goor, L.H.P., H. Panneman, and W.A. Van Haeringen. 2010. A proposal for standardization in forensic equine DNA typing: allele nomenclature of 17 equine-specific STR loci. Animal Genetics 41:122–127.

Van De Goor, L.H.P., W.A. Van Haeringen, and J.A. Lenstra. 2011. Population studies of 17 equine STR for forensic and phylogenetic analysis. Animal Genetics 42:627–633.

Wagner, A.P., S. Creel, and S.T. Kalinowski. 2006. Estimating relatedness and relationships using microsatellite loci with null alleles. Heredity 97:336–345.

Westemeier, R.L., J.D. Brawn, S.A. Simpson, T.L. Esker, R.W. Jansen, J.W. Walk, E.L. Kershner, J.L. Bouzat, and K.N. Paige. 1998. Tracking the long-term decline and recovery of an isolated population. Science 282:1695–1698.

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

Received 25 September 2017
Revised 4 March 2018
Accepted 31 August 2018
Published online 25 April 2019