



# A microsatellite analysis of five Colonial Spanish horse populations of the southeastern United States

E. K. Conant, R. Juras and E. G. Cothran

Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843-4458, USA

## Summary

The domestic horse (*Equus caballus*) was re-introduced to the Americas by Spanish explorers. Although horses from other parts of Europe were subsequently introduced, some New World populations maintain characteristics ascribed to their Spanish heritage. The southeastern United States has a history of Spanish invasion and settlement, and this influence on local feral horse populations includes two feral-recaptured breeds: the Florida Cracker and the Marsh Tacky, both of which are classified as Colonial Spanish horses. The feral Banker horses found on islands off the coast of North Carolina, which include, among others, the Shackleford Banks, the Corolla and the Ocracoke, are also Colonial Spanish horses. Herein we analyse 15 microsatellite loci from 532 feral and 2583 domestic horses in order to compare the genetic variation of these five Colonial Spanish Horse populations to 40 modern horse breeds. We find that the Corolla horse has very low heterozygosity and that both the Corolla and Ocracoke populations have a low mean number of alleles. We also find that the Florida Cracker population has a heterozygosity deficit. In addition, we find evidence of similarity of the Shackleford Banks, Marsh Tacky and Florida Cracker populations to New World Iberian horse breeds, while the origins of the other two populations are less clear.

**Keywords** feral horse, genetic distance, microsatellite, phylogeny.

## Introduction

The modern horse was re-introduced to the Americas beginning with Columbus' second trip to the New World in 1493 (Loch 1986). The earliest horses to reach North America were of Spanish origin and of a type known as the jennet (Loch 1986; Bennett 1998). The Spanish explorers and Conquistadores brought horses to the West Indies, where they were bred and used as stock for expeditions to the mainland (Chard 1940). It is generally assumed that some of the horses from such expeditions escaped and formed feral populations. However, it appears unlikely that the feral horse of the southeastern United States was produced by individuals from these early expeditions (i.e. prior to 1580; Chard 1940; Denhardt 1975; Bennett 1998). It was the later Spanish settlers, rather than these explorers,

who left behind and traded horses (Denhardt 1975). By the 18th century, horses of Spanish extraction were found from Florida to California (Bennett 1998), although by this time there were also horses of Northern and Central European origin in North America.

In the southeastern United States, a few feral populations have existed in close contact with civilization from the time of European settlement. Herein we examine five of these populations: three feral island populations, and two feral-recaptured breeds. All are considered a critical conservation priority by the American Livestock Breeds Conservancy (<http://www.albc-usa.org>).

## Outer Banks horses

The Outer Banks islands off the eastern coast of North Carolina are home to several feral horse populations, including the Shackleford Banks, Corolla and Ocracoke herds (Fig. 1). Feral horses are known to have existed on these islands once the mainland was settled, primarily by Englishmen, from 1650 until the present. The founding individuals of these populations may have arrived earlier (Quinn *et al.* 1982), but there has been a population of

Address for correspondence

E. G. Cothran, Department of Veterinary Integrative Biosciences,  
Texas A&M University, CVM, TAMU 4458, College Station,  
TX 77843-4458, USA.

E-mail: GCothran@cvm.tamu.edu

Accepted for publication 2 February 2011



**Figure 1** Map of the Outer Banks islands off the coast of North Carolina. Horses from the Corolla, Ocracoke and Shackleford Banks populations were included in this study.

horses on the islands for at least 350 years (Prioli & Taylor 2007). The Banker populations are all managed for population size, and current numbers are between 20 and 25, 110 and 130, and 88 and 119 for the Ocracoke, Shackleford Banks and Corolla, respectively.

It is expected that the Shackleford Banks, Ocracoke and Corolla herds have some genetic affinity to each other due to

their geographic proximity. One study was carried out on the Ocracoke and Shackleford populations, but the sample size for the Shackleford Banks ( $n = 4$ ) was small (Goodloe *et al.* 1991). The Ocracoke population was found to have similar levels of genetic diversity to domestic horse breeds, but the Shackleford Banks was found to have reduced gene diversity (heterozygosity), although this result was most likely related to sample size.

#### Other Colonial Spanish horses

Other originally feral horse populations in the southern United States are the Cracker horse, which originated in Florida, and the Marsh Tacky of South Carolina. Both breeds are derived from feral horses that have been re-domesticated, and they are thought to be of Spanish origin. In 1989 when the Florida Cracker Horse Association (FCHA) was founded, 31 horses were registered. Currently there are over 800 horses registered by the FCHA.

Unlike the situation with Florida Crackers, the American Livestock Breeds Conservancy (<http://www.albc-usa.org>) estimates that there are fewer than 150 true Marsh Tackies left. The Marsh Tacky is thought to be closely related to the Florida Cracker, and both are said to have physical characteristics common to breeds of Iberian origin.

All these feral or formerly feral horses are considered to be Colonial Spanish Horses based on conformation, and as such are of interest both for genetic preservation as herds that are unique in their continual survival in harsh habitats, as well as for their historic value. They perhaps represent some of the few remaining relics in North America of the earliest modern horses introduced to the New World. Disturbingly, all these populations are today quite small or have undergone significant reductions in population size in the recent past. Small populations are often highly inbred, and such inbreeding can confound inferences regarding population origins. The additional effects of intermixture of these populations with other European breeds further complicate this picture.

Population analyses of genetic variation are an important conservation tool. Although previous studies of Old World breeds and feral populations (Canon *et al.* 2000; Juras *et al.* 2003; Morais *et al.* 2003; Aberle *et al.* 2004; Azor *et al.* 2007; Luis *et al.* 2007) and South American breeds (Kelly *et al.* 2002; Vega-Pla *et al.* 2006; Sereno *et al.* 2008; DeAssis *et al.* 2009) have been carried out, few studies have looked at North American feral horses. A recent paper on the Sable Island horses found them to be unique compared to other breeds (Plante *et al.* 2007). Some studies have looked at Western American feral horses (Bowling 1994; Cothran & Singer 2000; Ashley 2004), but to date no in depth study of southeastern American populations has been carried out.

Herein we use microsatellite loci to examine the genetic diversity of five horse populations with regard to internal

variation, similarity to each other and similarity to other modern horse populations. Information on the genetic structure can give insight into the history of the populations as well as in assisting in management improvements and conservation planning.

## Materials and methods

### Sample collection

A total of 532 feral horses (287 Shackleford Banks, 38 Corolla, 37 Ocracoke, 46 Florida Crackers, 124 Marsh Tackies) were compared to 2583 horses from 40 other breeds (see Table 1 for a list of breeds). In addition, Miniature Donkeys ( $n = 134$ ) were included as an outgroup.

### Microsatellite analysis

A total of 15 microsatellite loci (*AHT4*, *AHT5*, *ASB2*, *ASB17*, *ASB23*, *HMS2*, *HMS3*, *HMS6*, *HMS7*, *HTG4*, *HTG6*, *HTG7*, *HTG10*, *LEX33* and *VHL20*) was examined using the methods described previously (Juras *et al.* 2003; Vega-Pla *et al.* 2006). This specific set of highly polymorphic loci maximizes the number of breeds available to be used in this study. All 15 loci are included in the Food and Agriculture Organization/International Society of Animal Genetics (FAO/ISAG) Measurement of Domestic Animal Diversity (MoDAD) panel (Hoffmann *et al.* 2004).

### Statistical analysis

Gene diversity was calculated using ARLEQUIN 3.1 (Excoffier *et al.* 2005). An analysis of molecular variance (AMOVA) was also calculated with ARLEQUIN 3.1.  $F$ -statistics, deviation from Hardy–Weinberg equilibrium and allele frequencies were calculated with FSTAT 2.9.3.2 (Goudet 1995). There was no evidence of linkage disequilibrium, so loci were treated individually. The presence of genetic bottlenecks in the Colonial populations was tested for under the Luikart model using BOTTLENECK 1.2.02 (Cornuet & Luikart 1996). Restricted maximum likelihood (RML) trees based on chord distance were generated from 1000 bootstrapped allele frequency datasets using PHYLIP 3.69 (Felsenstein 2005). Restricted maximum likelihood trees were chosen because they have been shown to generate clades consistent with historical knowledge of horse breeds (Cothran & Luis 2005). Majority-rule consensus trees were created with PHYLIP 3.69. Trees were visualized with ARCHAEOPTERYX version 0.957 beta (Han & Zmasek 2009).

The data were also examined using STRUCTURE 2.1 (Pritchard *et al.* 2000), which assigns samples to clusters using a Bayesian method. Burn-in was set to 20 000, and 100 000 MCMC repetitions were used, as recommended by the software developer (Falush *et al.* 2007). Runs for each different selected value of  $K$  (the number of clusters) were repeated

ten times. STRUCTURE works best with a small number of discrete populations (Pritchard *et al.* 2000); therefore the five southeastern populations were analysed along with the Donkey outgroup. CLUMPP (Jakobsson & Rosenberg 2007) was used to concatenate the data from the multiple runs for each  $K$ , and DISTRICT (Rosenberg 2004) was used to graphically display the results.

A factorial correspondence analysis was carried out using GENETIX version 4.05 in order to further examine the relationships of the five populations studied here to other breeds with which they may share ancestry (Belkhir *et al.* 1996–2004). GENECLASS 2.0 (Piry *et al.* 2004) was used for individual assignment to predefined populations using a Bayesian algorithm developed previously (Rannala & Mountain 1997). This model primarily examines the effects of admixture in a population due to immigration rather than genetic drift or mutation.

## Results

Standard measures of diversity are shown in Table 1. With the exception of the Corolla, all of the test populations examined had observed and expected heterozygosity ( $H_o$  and  $H_e$ ) values within 1.5 standard deviations of the mean for all other breeds. The Corolla population was nearly four standard deviations away from the mean for  $H_e$  and nearly 3.5 standard deviations away from the mean for  $H_o$ . Only the Friesian breed (of the populations used in this analysis) had lower values for  $H_o$  and  $H_e$  than the Corolla. The Corolla and the Ocracoke populations had the lowest two values for the mean number of alleles of all the populations studied. Of the five populations examined, only the Florida Cracker population had a significant ( $P = 0.005$ ) heterozygosity deficit ( $F_{IS} = 0.058$ ), which is uncommon in horse breeds (Table 1). No statistically significant evidence of genetic bottlenecks was found in any of the five Colonial Spanish populations.

The AMOVA analysis based on all the listed horse breeds (excluding Miniature Donkeys) found 90.5% of the variance was within individuals, and 9.2% was between populations. This result is similar to what is seen in other studies (Bjornstad *et al.* 2000; Canon *et al.* 2000; Aberle *et al.* 2004; Leroy *et al.* 2009).

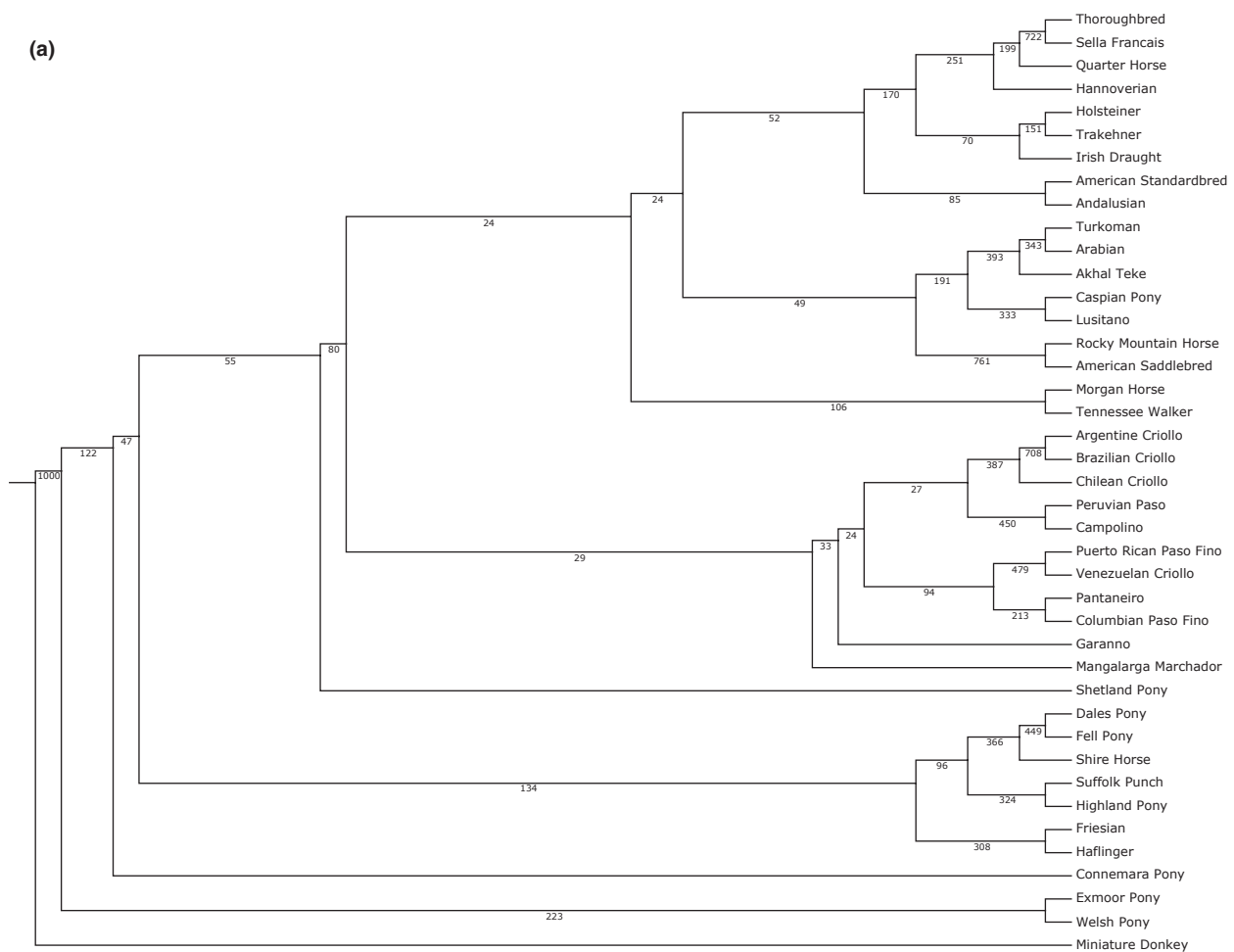
PHYLIP 3.69 was used to generate majority-rule consensus trees from 1000 bootstrapped RML trees of the 40 horse breeds along with the five test populations and the Miniature Donkey (Fig. 2). As is common with trees of horse breeds, bootstrap values are low, but the clades stay consistent (Cothran & Luis 2005). A tree of all 41 breeds with none of the feral populations was generated first to establish a reference (Fig. 2a). In this tree, the main clades are pony and draft horses, South American breeds of Iberian origin, and Old World and North American breeds. Next, the five test populations were examined together with the 41 other breeds (Fig. 2b). The Ocracoke population did not fall into a

**Table 1** Observed and expected heterozygosity ( $H_o$  and  $H_e$ ),  $F_{IS}$ , and mean number of alleles for 40 horse breeds as well as the three outer banks populations, the Florida Cracker and the Marsh Tacky.

Breeds	<i>n</i>	Heterozygosity				$F_{IS}$	<i>P</i> value	Mean number	
		Expected	SD	Observed	SD			of alleles	SD
Highland Pony	28	0.67	0.08	0.68	0.08	-0.013	0.702	5.333	1.193
Akhal Teke	29	0.65	0.18	0.73	0.23	-0.111	1.000	5.400	1.405
Argentine Criollo	25	0.73	0.07	0.74	0.11	-0.021	0.787	6.000	1.461
Welsh Pony	31	0.75	0.07	0.74	0.09	0.013	0.288	6.867	1.707
Shire	32	0.66	0.11	0.69	0.14	-0.032	0.892	5.400	1.083
Turkoman	113	0.77	0.08	0.75	0.08	0.034	0.003	8.733	2.081
Puerto Rican Paso Fino	60	0.68	0.08	0.70	0.07	-0.027	0.680	6.667	1.075
Andalusian	33	0.73	0.10	0.70	0.08	0.047	0.026	5.933	1.181
Arabian	89	0.70	0.11	0.64	0.14	0.091	0.000	6.867	2.029
Connemara Pony	69	0.75	0.08	0.77	0.08	-0.017	0.847	6.800	1.833
Caspian Pony	87	0.74	0.12	0.74	0.14	-0.001	0.517	7.400	1.818
Fell Pony	55	0.72	0.08	0.72	0.12	0.002	0.486	5.933	1.340
Friesian	134	0.52	0.20	0.50	0.19	0.041	0.015	5.133	2.093
Haflinger	46	0.67	0.18	0.66	0.19	0.015	0.247	5.667	1.491
Hannoverian	34	0.77	0.06	0.75	0.13	0.035	0.090	6.733	1.436
Holsteiner	36	0.74	0.07	0.72	0.16	0.027	0.158	5.800	1.424
Lusitano	62	0.72	0.13	0.68	0.14	0.058	0.009	6.067	1.569
Quarter Horse	40	0.76	0.08	0.74	0.09	0.033	0.057	7.067	1.436
Peruvian Paso	42	0.75	0.11	0.76	0.10	-0.010	0.708	7.333	1.619
American Standardbred	33	0.69	0.10	0.67	0.12	0.032	0.254	5.867	1.500
Thoroughbred	139	0.71	0.08	0.69	0.10	0.026	0.047	5.667	1.300
Columbian Paso Fino	88	0.75	0.07	0.75	0.11	0.006	0.356	7.800	2.104
Campolino	30	0.74	0.11	0.71	0.16	0.034	0.089	6.867	1.500
Chilean Criollo	30	0.72	0.08	0.69	0.12	0.038	0.067	5.933	1.181
Pantaneiro	25	0.75	0.09	0.77	0.12	-0.025	0.826	6.667	1.350
Irish Draught	47	0.78	0.06	0.78	0.07	0.003	0.431	7.533	1.746
Rocky Mountain Horse	45	0.74	0.07	0.72	0.07	0.035	0.039	7.067	1.289
Dales Pony	86	0.69	0.13	0.71	0.12	-0.031	0.978	6.200	1.327
Exmoor Pony	46	0.63	0.10	0.60	0.11	0.052	0.020	5.267	1.181
Garrano	37	0.79	0.06	0.75	0.13	0.042	0.021	7.533	1.857
Morgan Horse	59	0.76	0.08	0.71	0.10	0.060	0.002	7.533	1.962
Shetland Pony	97	0.75	0.10	0.72	0.08	0.038	0.012	7.800	2.688
American Saddlebred	174	0.71	0.14	0.70	0.16	0.019	0.095	7.133	1.996
Mangalarga Marchador	111	0.74	0.07	0.70	0.08	0.053	0.208	7.600	1.925
Sella Francais	41	0.74	0.09	0.73	0.13	0.010	0.342	6.267	1.526
Suffolk Punch	50	0.71	0.12	0.73	0.14	-0.033	0.956	5.467	1.543
Tennessee Walker	49	0.68	0.12	0.72	0.17	-0.052	0.956	6.400	1.818
Brazilian Criollo	50	0.74	0.08	0.76	0.10	-0.027	0.924	7.200	1.558
Trakehner	64	0.73	0.10	0.76	0.09	-0.042	0.986	6.067	2.144
Venezuelan Criollo	213	0.78	0.06	0.76	0.06	0.026	0.002	8.400	2.471
Florida Cracker	46	0.71	0.07	0.67	0.06	0.058	0.005	6.533	1.784
Shackleford Banks	287	0.65	0.15	0.67	0.15	-0.028	0.499	6.667	1.445
Corolla	38	0.53	0.18	0.54	0.18	-0.021	0.751	4.000	1.265
Ocracoke	37	0.67	0.08	0.72	0.10	-0.069	0.993	4.867	1.360
Marsh Tacky	124	0.73	0.08	0.74	0.10	-0.013	0.869	7.733	2.048

major clade and branches away first from the outgroup. The Florida Cracker pairs with the Tennessee Walker, which branches off of into a group containing the Venezuelan Criollo, Puerto Rican Paso Fino, the Shackleford Banks, and the Marsh Tacky. All these populations fall into a larger clade that contains primarily South American Iberian

breeds. Finally, the Corolla horse falls next to the Shetland Pony. Because adding in test populations slightly changed the overall topology of the tree, we further examined each southeastern American population individually against the 41 other breeds (data not shown). The placement of each test population was similar to that seen in Fig. 2b, although



**Figure 2** Consensus tree from 1000 bootstrapped RML trees (chord distance) based on majority-rule, in 40 horse breeds plus the Donkey outgroup (a) as generated by PHYLIP, and drawn with Archaeopteryx; (b) the consensus tree based on majority-rule in 40 horse breeds plus all five feral and feral-recaptured populations with Donkey outgroup, as generated by PHYLIP and drawn with the ARCHAEOPTERYX program. Bootstrap values are the number of times out of 1000 that the arrangement to the right occurred.

the Marsh Tacky did not cluster with the New World Iberian horses, but instead paired with the Tennessee Walker. The Corolla did not pair with the Shetland Pony, but fell instead between the Exmoor and Connemara Ponies.

A Bayesian model from the program STRUCTURE was used to examine the relationship of the five populations with respect to each other (Fig. 3 and Table 2). STRUCTURE divides the input data into a fixed number ( $K$ ) of distinct groups. By varying  $K$  we obtained a sense of the relative strengths of the genetic signals of population differentiation between these five feral populations. The best value of  $K$  to fit the data was calculated as described in Evanno *et al.* 2005 and was found to be six, i.e. equal to the number of populations included in the analysis. At  $K = 2$ , all the feral populations are found to be distinct from the Miniature Donkey. From  $K = 3$ , the model identifies the Shackleford Banks population as the most distinct group, with 98% of Shackleford Banks genotypes belonging to a single cluster along with 19% of the Ocracoke genotypes. At  $K = 4$ , 87% of the

Florida Cracker genotypes and 93% of the Marsh Tacky genotypes share a cluster, while 79% and 65% of the Corolla and Ocracoke fall into a second cluster and 90% of the Shackleford Banks genotypes fall into one cluster. At  $K = 5$ , Shackleford Banks, Corolla and Marsh Tackies predominantly fall into unique clusters (86%, 98% and 90%, respectively). At  $K = 6$ , the Shackleford Banks population begins to show internal differentiation. The Corolla, Ocracoke and Marsh Tackies again primarily fall into unique clusters, although the Florida Cracker does not.

To further understand the relationships of these five populations to other horse breeds, we carried out a factorial correspondence analysis (Fig. 4). Based on the location of the feral populations within the consensus tree, 32 of the original 40 breeds were used in this analysis. All the North and South American breeds, along with the pony and heavy horse cluster, were compared to the five southeastern US horse populations. GENETIX was used to generate a factorial correspondence analysis using the AFC 3D by populations

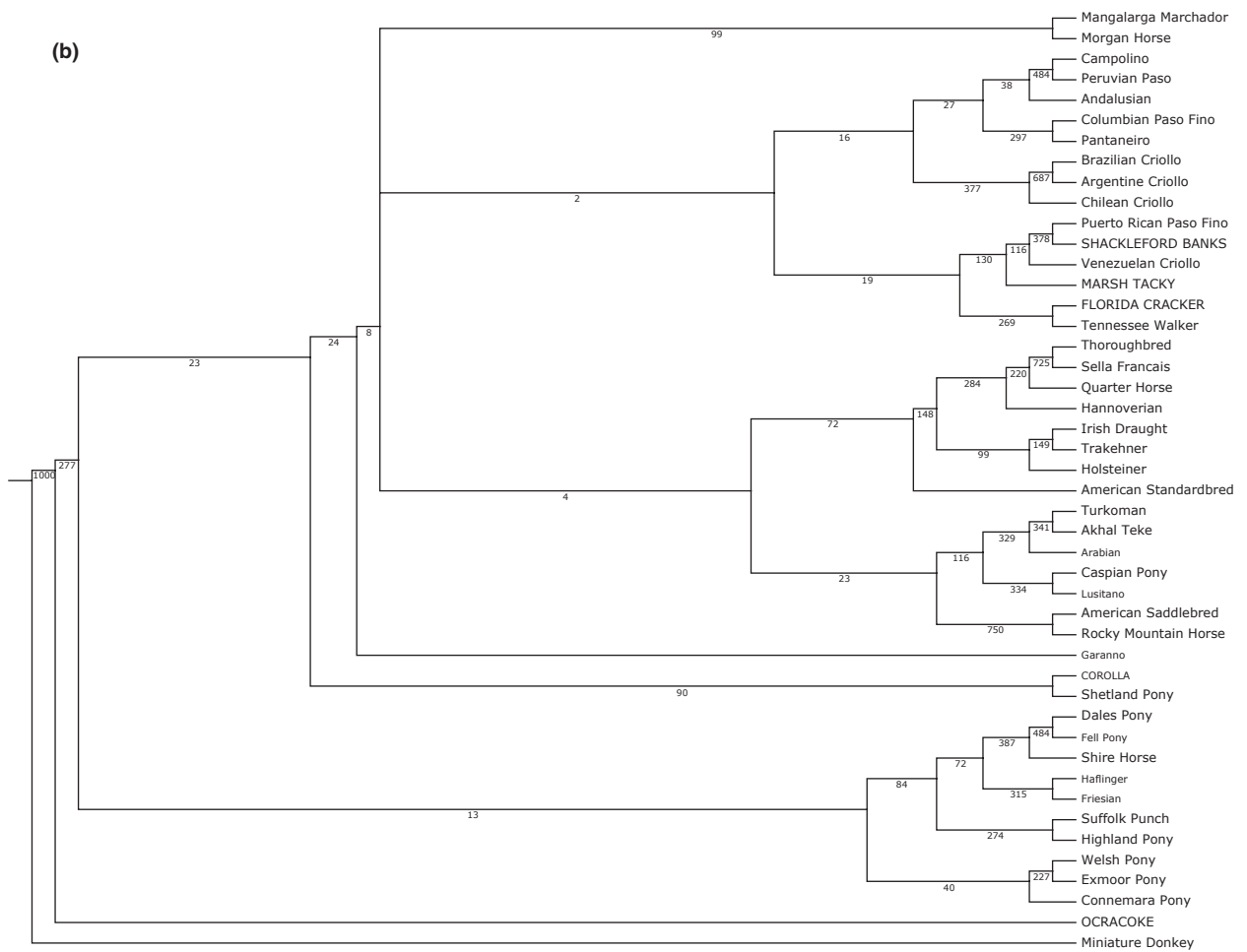


Figure 2 Continued.

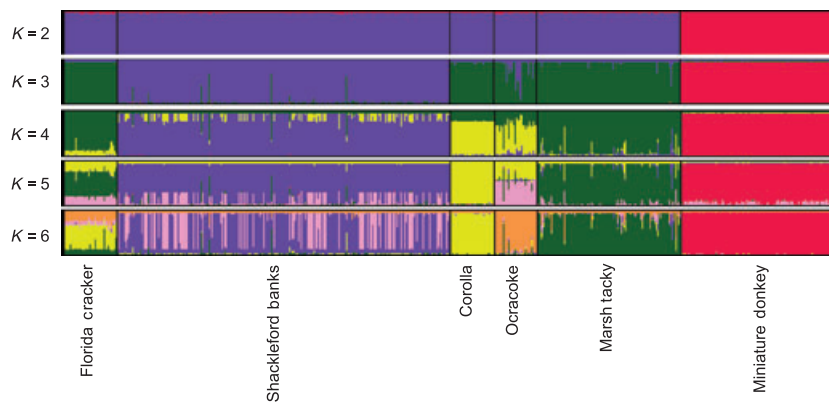


Figure 3 Proportion of membership of 656 individuals from five populations: the Florida Cracker, Shackleford Banks, Corolla, Ocracoke, and Marsh Tacky, with the Miniature Donkey used as an out group for  $K = 2-6$ , as calculated by STRUCTURE software. Multiple runs for each  $K$  were concatenated using CLUMPP, and DISTRUCT was used to generate images.

option. This option calculates a centre of gravity point for each population, which was then plotted in 2D. The comparison of the first vs. second, and second vs. third axes are shown. In the top plot (Fig. 4a), the Shackleford Banks

population is found to be differentiated from all other breeds on the first axis. The three island populations all have positive  $x$  values, unlike all other breeds except the Shire, Haflinger and Highland Pony. In the bottom plot (Fig. 4b),

**Table 2** Genotype assignment based on Structure analysis at  $K = 6$  and individual assignment as calculated by GeneClass for the Colonial Spanish populations

Population	$n$	Inferred Clusters						Assignment	
		I	II	III	IV	V	VI	To self ( $n$ )	To other ( $n$ )
Florida Cracker	46	<b>0.146</b>	<b>0.545</b>	0.009	0.098	0.001	<b>0.201</b>	40	5
Shackleford Banks	287	0.014	0.008	<b>0.650</b>	<b>0.316</b>	0.002	0.010	282	5
Corolla	38	0.005	<b>0.969</b>	0.009	0.011	0.001	0.005	38	0
Ocracoke	37	0.045	0.014	0.016	0.023	0.002	<b>0.901</b>	33	4
Marsh Tacky	124	<b>0.869</b>	0.027	0.024	0.033	0.001	0.046	101	23
Miniature Donkey	134	0.001	0.002	0.001	0.001	<b>0.994</b>	0.001	134	0

\*Numbers in bold highlight the cluster(s) representing the greatest contribution from each population.

the ponies and heavy horses are primarily found in +x, +y quadrant and the New World Iberian breeds fall into the -y quadrants. The North American breeds are found in the -x, +y quadrant. As seen in Fig. 2, the Marsh Tacky, Florida Cracker, and Shackleford Banks populations are near the Puerto Rican Paso Fino and the Venezuelan Criollo. The Ocracoke and Corolla are in the same quadrant as the ponies and heavy horses.

Finally, we tested the ability of these genetic data to recover the given breed assignments by using GENECLASS. Of the 3048 individuals considered, 28% were correctly assigned to their population of origin. Within the test populations, the percentage of self-assignment was higher (see Table 2). Forty of 45 Florida Crackers were correctly assigned, and 101 of 124 Marsh Tackies were correctly assigned. For the island populations, 282 of 287 Shackleford Banks, 33 of 37 Ocracoke and all of the Corolla horses were correctly assigned. Of the feral and feral-recaptured horses assigned to different populations, no clear pattern was apparent, with samples falling into North and South American breeds, breeds with strong thoroughbred influence, as well as pony breeds.

## Discussion

With the exception of the Corolla, the feral and feral-recaptured populations examined herein all show levels of microsatellite diversity within the range of other horse populations, despite their small population sizes. The low observed heterozygosity and mean number of alleles seen in the Corolla is likely due to a historically small population size.

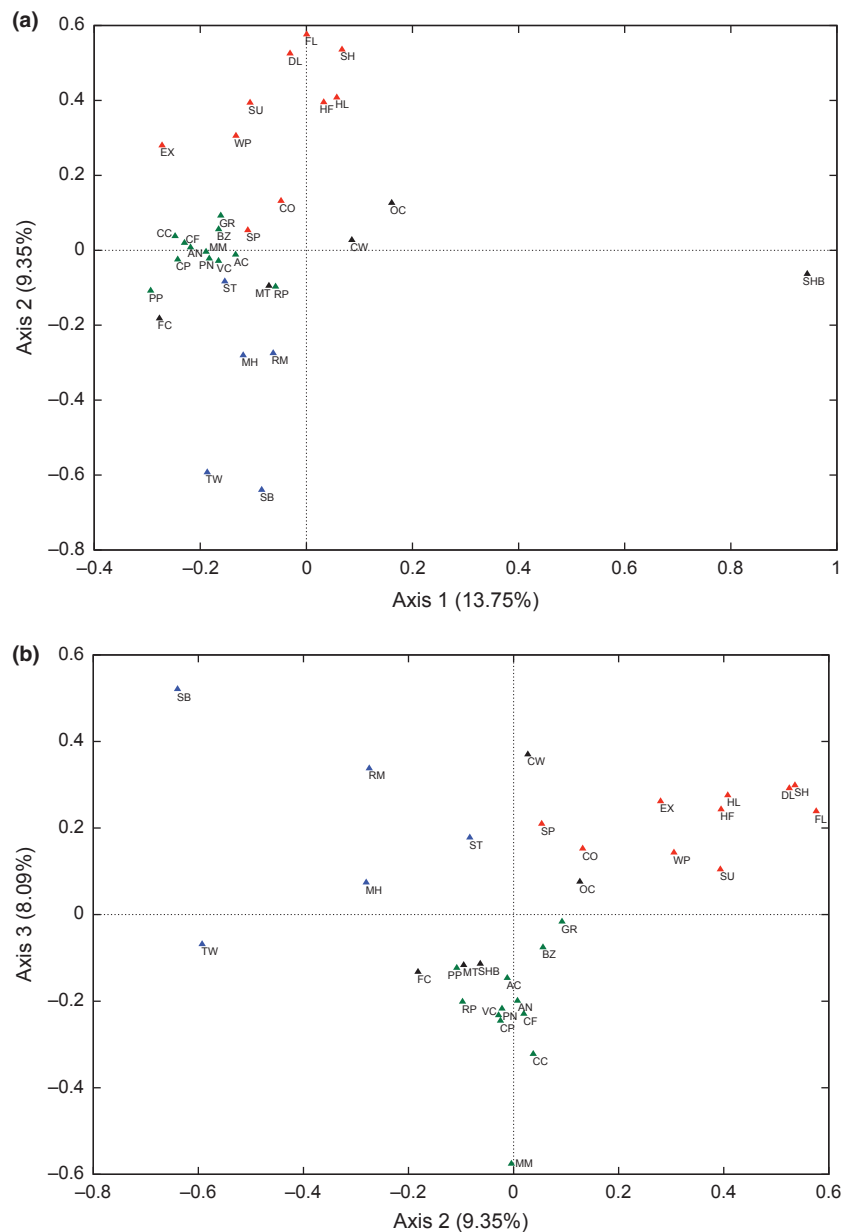
Of the five populations we studied, the Florida Cracker population was the only population to have a significant heterozygosity deficit. The cause of this deficit is not apparent, but may be due to population subdivision based upon isolation of animals on different farms leading to nonrandom mating within the breed.

The relatively high heterozygosity of the Ocracoke, Marsh Tacky, Shackleford Banks and Florida Cracker may be due to recent outcrossing. At least one Andalusian stallion was introduced into the Ocracoke herd in recent times. The

Shackleford population has been isolated since the 1970s, but before that a low level of gene flow into the herd was possible. The Marsh Tacky did not exist as a closed breed until 2009 and it is certain that there was outbreeding with other horses present in the southeastern United States, which could have included the Tennessee Walker.

The three outer banks populations do not cluster together within the consensus tree, but do show similarity in the factorial correspondence analysis. The location of the Ocracoke and the Corolla within the tree may be best explained by the fact that these two populations are small and have low allelic diversity, as illustrated by their mean number of alleles. This low diversity makes placement within the consensus tree difficult. Although the Corolla is found with the pony breeds, it is worth noting that of the three island populations, the Corolla is said to be the least pony-like in size and conformity. Further examination of these populations in the STRUCTURE and factorial correspondence analysis yields evidence of shared ancestry among the island populations. It is likely that these populations have a distant shared ancestry, but recent reductions in population size and isolation from each other has led to the observed differentiation. No statistically significant evidence of genetic bottlenecks was found based on the Luikart model for any of the Colonial Spanish populations. However, the conditions under which these populations have existed may not show the expected variation normally associated with bottlenecks. These populations, especially in the case of the Banker horses, have remained small, rather than expanding following a single bottleneck event.

The Tackies and Crackers show similarity to each other in the STRUCTURE analysis at  $K = 2$  and 3. The Marsh Tackies, Florida Crackers and Shackleford Banks populations show similarity to each other in the RML tree and the factorial correspondence analysis. In both the consensus tree and the factorial correspondence analysis, the Shackleford, Marsh Tacky and Florida Cracker populations all show affinity to the Puerto Rican Paso Fino and other South American Iberian breeds, which is consistent with popular history and the Colonial Spanish designation for these populations. Given that early European explorers



**Figure 4** Factorial correspondence analysis carried out by GENETIX version 4.05 of 32 horse breeds of the four clades of interest, based on 15 microsatellites and summarized by population via a centre of gravity calculation. The first vs. second axes (13.75% and 9.35% of inertia explained, respectively) are shown at the top (a), and the second vs. third axes are shown at the bottom (b) (9.35% and 8.09% of inertia explained, respectively). Populations are coded by colour and abbreviations are as follows: Per clade seen in Fig. 2, heavy horse and pony breeds are shown in red: SP, Shetland Pony; CO, Connemara Pony; WP, Welsh Pony; SU, Suffolk Punch; HL, Highland Pony; EX, Exmoor Pony; HF, Haflinger; SH, Shire; DL, Dales Pony; FL, Fell Pony. South American breeds and Iberian breeds are in green: RP, Puerto Rican Paso Fino; CF, Columbian Paso Fino; CP, Campolino; PP, Peruvian Paso; VC, Venezuelan Criollo; MM, Mangalarga Marchador; PN, Pantaneiro; AC, Argentine Criollo; AN, Andalusian; CC, Chilean Criollo; BZ, Brazilian Criollo; GR, Garrano. North American breeds are in blue: SB, American Saddlebred; TW, Tennessee Walker; MH, Morgan Horse; RM, Rocky Mountain Horse; ST, American Standardbred. The populations examined in this paper are in black: MT, Marsh Tacky; FC, Florida Cracker; OC, Ocracoke; CW, Corolla; SHB, Shackleford Banks.

and settlers often obtained livestock from Puerto Rico and other Caribbean islands before heading to the eastern coast of the United States, this signature may truly show shared ancestry, which is rather remarkable after 400 years of separation. Colonial Spanish horses are

representative of the Iberian horses that were brought by early Spanish explorers and settlers to both North and South America. These horses were believed to be distinct from the Andalusian breed that was present in Spain at the time, and these North American Colonial Spanish



Horses show more similarity to other New World horses of Iberian origin than they do to the modern Lusitano and Andalusian.

Although these horse populations are known for being Colonial Spanish, no previous genetic analysis has examined their history in detail. Herein we find evidence of a shared ancestry amongst the Shackleford horses, the Florida Cracker and the Marsh Tacky, even though these populations have not interbred in recent times. We believe that these relic populations are worth preserving, both for their genetic as well as their historical heritage as descendants of the first modern horses in the Americas. We also feel that the evidence presented here should be considered in future decisions of population management of the Banker herds. Although we find that the diversity of most of these horses to be within the range of that seen in other breeds, the diversity will continue to decline, especially as population sizes are closely managed at small numbers. We hope that future monitoring will continue in order to preserve these breeds.

## Acknowledgements

We would like to thank two anonymous reviewers for their comments and helpful suggestions. We would like to thank Sue Stuska for her careful review of this manuscript as well as collecting samples. We also thank Jeannette Beranger (of the ALBC), Carolyn Mason (Foundation for Shackleford Horses), the Florida Cracker Horse Association, the Carolina Marsh Tacky Horse Association, the Corolla Wild Horse Fund and all the breeders and caretakers who provided samples for their critical assistance. Thanks also to Gavin Conant for his help with figures.

## References

- Aberle K.S., Hamann H., Drogemuller C. & Distl O. (2004) Genetic diversity in German draught horse breeds compared with a group of primitive, riding and wild horses by means of microsatellite DNA markers. *Animal Genetics* **35**, 270–7.
- Ashley M.C. (2004) Population genetics of feral horses: implications of behavioral isolation. *Journal of Mammalogy* **85**, 611–7.
- Azor P.J., Valera M., Gomez M.D., Goyache F. & Molina A. (2007) Genetic characterization of the Spanish Trotter horse breed using microsatellite markers. *Genetics and Molecular Biology* **30**, 37–42.
- Belkhir K., Borsa P., Chikhi L., Raufaste N. & Bonhomme F. (1996–2004) GENETIX 4.05, Windows TM software for population genetics. In: *Laboratoire Genome, Populations, Interactions*. Universite Montpellier II, Montpellier, France.
- Bennett D. (1998) *Conquerors: The Roots of New World Horsemanship*. Amigo Publications, Solvang, CA.
- Bjornstad G., Gunby E. & Roed K.H. (2000) Genetic structure of Norwegian horse breeds. *Journal of Animal Breeding and Genetics-Zeitschrift Fur Tierzucht Und Zuchtungsbiologie* **117**, 307–17.
- Bowling A.T. (1994) Population-genetics of Great-Basin feral horses. *Animal Genetics* **25**, 67–74.
- Canon J., Checa M.L., Carleos C., Vega-Pla J.L., Vallejo M. & Dunner S. (2000) The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. *Animal Genetics* **31**, 39–48.
- Chard T. (1940) Did the first Spanish horses landed in Florida and Carolina leave Progeny? *American Anthropologist, New Series* **42**, 90–106.
- Cornuet J.M. & Luikart G. (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001–14.
- Cothran E.G. & Luis C. (2005) Genetic distance as a tool in conservation of rare horse breeds. In: *Conservation Genetics of Endangered Horse Breeds* (Ed. by I. Bodo, L. Alderson & B. Langolis), pp. 55–71. Wageningen, The Netherlands.
- Cothran E.G. & Singer F.J. (2000) *Analysis of Genetic Variation in the Pryor Mountain Wild Horse Herd*. U.S. Geological Survey, Midcontinent Ecological Science Center, Fort Collins, CO.
- DeAssis J.B., DeLaat D.M., Peixoto M.G., Bergmann J.A., Fonseca C.G. & Carvalho M.R. (2009) Genetic diversity and population structure in Brazilian Mangalarga Marchador horses. *Genetics and Molecular Research* **8**, 1519–24.
- Denhardt R.M. (1975) *The Horse of the Americas*. University of Oklahoma Press, Norman.
- Evanno G., Regnaut S. & Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611–20.
- Excoffier L., Laval G. & Schneider S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47–50.
- Falush D., Stephens M. & Pritchard J.K. (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7**, 574–8.
- Felsenstein J. (2005) *PHYLIP (Phylogeny Inference Package) Version 3.6*. Distributed by Felsenstein J. Department of Genome Sciences, University of Washington, Seattle.
- Goodloe R.B., Warren R.J., Cothran E.G., Bratton S.P. & Trembicki K.A. (1991) Genetic-variation and its management applications in eastern United States feral horses. *The Journal of Wildlife Management* **55**, 412–21.
- Goudet J. (1995) FSTAT (vers. 1.2): a computer program to calculate *F*-statistics. *Journal of Heredity* **86**, 485–6.
- Han M.V. & Zmasek C.M. (2009) phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics* **10**, 356.
- Hoffmann I., Ajmone Marsan P., Barker J.S.F., Cothran E.G., Hanotte O., Lenstra J.A., Milan D., Weigend S. & Simianer H. (2004) New MoDAD marker sets to be used in diversity studies for the major farm animal species: recommendations of a joint ISAG/FAO working group. In: *Proceedings of 29th International Conference on Animal Genetics*, p. 123, Tokyo, Japan.
- Jakobsson M. & Rosenberg N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–6.
- Juras R., Cothran E.G. & Klimas R. (2003) Genetic analysis of three Lithuanian native horse breeds. *Acta Agriculturae Scandinavica Section A, Animal Science* **53**, 180–5.
- Kelly L., Postiglioni A., De Andres D.F., Vega-Pla J.L., Gagliardi R., Biagetti R. & Franco J. (2002) Genetic characterisation of the Uruguayan Creole horse and analysis of relationships among horse breeds. *Research in Veterinary Science* **72**, 69–73.

- Leroy G., Caldele L., Verrier E., Meriaux J.C., Ricard A., Danchin-Burge C. & Rognon X. (2009) Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. *Genetics Selection Evolution* **41**, 5.
- Loch S. (1986) *The Royal Horse of Europe: The Story of the Andalusian and Lusitano*. J.A. Allen, London.
- Luis C., Juras R., Oom M.M. & Cothran E.G. (2007) Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. *Animal Genetics* **38**, 20–7.
- Morais J., Oom M.M., Malta-Vacas J. & Luis C. (2003) Genetic structure of an endangered Portuguese semiferal pony breed, the Garrano. *Biochemical Genetics* **43**, 347–64.
- Piry S., Alapetite A., Cornuet J.M., Paetkau D., Baudouin L. & Estoup A. (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536–9.
- Plante Y., Vega-Pla J.L., Lucas Z., Colling D., De March B. & Buchanan F. (2007) Genetic diversity in a feral horse population from Sable Island, Canada. *Journal of Heredity* **98**, 594–602.
- Prioli C.A. & Taylor S. (2007) *The Wild Horses of Shackleford Banks*. John F. Blair, Publisher, Winston-Salem, NC.
- Pritchard J.K., Stephens M. & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–59.
- Quinn D.B., Quinn A.M. & North Carolina, Division of Archives and History. (1982) *The First Colonists: Documents on the Planting of the First English Settlements in North America, 1584–1590*. Division of Archives and History, North Carolina Department of Cultural Resources, Raleigh.
- Rannala B. & Mountain J.L. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 9197–201.
- Rosenberg N.A. (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137–8.
- Sereno F.T.P.D., Sereno J.R.B., Vega-Pla J.L., Kelly L. & Bermejo J.V.D. (2008) Genetic diversity of Brazilian Pantaneiro horse and relationships among horse breeds. *Pesquisa Agropecuaria Brasileira* **43**, 595–604.
- Vega-Pla J.L., Calderon J., Rodriguez-Gallardo P.P., Martinez A.M. & Rico C. (2006) Saving feral horse populations: does it really matter? A case study of wild horses from Donana National Park in southern Spain. *Animal Genetics* **37**, 571–8.