

# Analysis of Genetic Diversity in the Corolla Feral Horse herd of North Carolina

*E. Gus Cothran*

*Department of Veterinary Integrative Biosciences*

*College of Veterinary Medicine and Biomedical Science*

*Texas A & M University*

The purpose of this study was to analyze genetic variability in the Corolla feral horse herd from the Outer Banks of North Carolina in order to determine its conservation status and to understand its origins.

To start out I will give a little background. First, what is genetic diversity? Genetic diversity primarily refers to variation within the genetic make up of an individual. It also is variation in genetic make up among individuals within a population, genetic variation among different populations of a species, and genetic differences among species. For this work I am focusing on individual and populational variation.

Why is Genetic Diversity Important? It gives individuals flexibility to respond to different environmental conditions, it increased disease resistance, it gives populations the ability to adapt to environmental changes over time (for domestic animals this also can mean the ability to make selective improvements to a breed), and it buffers against deleterious recessive genes.

What factors act to change genetic diversity? The number one factor is population size, the smaller the size the lower the diversity. Also, inbreeding reduces individual variation. Out-crossing increases diversity when different groups are crossed. Selection, which will usually decrease variation, is another factor as is mutation which increases variation but is unimportant over short time scales.

How do we measure diversity? Individual diversity is measured as Heterozygosity (Observed  $H_o$ ) which is the proportion of the genome that varies within an individual. The Inbreeding coefficient also is a measure of basically the same thing but on a theoretical level. Both these measures are independent of sample size. For estimating genetic diversity of a whole population or breed we again use Heterozygosity but it is Expected Heterozygosity ( $H_e$ ) which is the average proportion of the genome that varies based upon population genetic theory. Another important measure is simply the number of variants observed in the population. These measures are correlated with sample size. We determine these different measures by typing of genetic markers. A gene marker is a variable spot in the genome that we can test. The different forms the marker can take are called alleles. Each marker can have two possible alleles in an individual but many across the population. For an individual, if the alleles are the same, the marker is homozygous. If different, the marker is heterozygous. Heterozygosity is the proportion of markers out of all tested that are heterozygous. This can be calculated as the average of those actually observed as heterozygous ( $H_o$ ) or by the predicted proportion based upon theory calculated from the frequency of each allele in the population ( $H_e$ ).

Table 1: Genetic variability estimates for the Corolla NC horse herd and values for selected domestic breeds.  $N$  = sample size,  $MNA$  = Mean Number of Alleles per locus,  $Ae$  = Effective Number of Alleles,  $Ho$  = Observed heterozygosity,  $He$  = Expected Heterozygosity,  $Fis$  = Estimated inbreeding based upon Heterozygosity ratio.

<b>Breed</b>	<b><math>N</math></b>	<b><math>MNA</math></b>	<b><math>Ae</math></b>	<b><math>Ho</math></b>	<b><math>He</math></b>	<b><math>Fis</math></b>
<b>Corolla</b>	<b>38</b>	<b>4.00</b>	<b>2.531</b>	<b>0.555</b>	<b>0.545</b>	<b>-0.018</b>
Andalusian	52	6.58	4.259	0.722	0.753	0.041
Arabian	47	7.17	3.814	0.660	0.727	0.092
Chilean Criollo	30	6.25	3.741	0.714	0.718	0.005
Cleveland Bay	47	4.92	2.934	0.609	0.627	0.027
Fell Pony	43	7.00	4.032	0.822	0.736	-0.116
Friesian	304	5.83	2.561	0.545	0.539	-0.011
Haflinger	386	7.42	3.874	0.717	0.729	0.017
Peruvian Paso	46	7.75	4.774	0.784	0.780	-0.005
Shire	32	5.58	3.181	0.677	0.656	-0.031
Thoroughbred	1195	5.75	3.918	0.734	0.726	-0.011
<b>Domestic Mean</b>	75	6.72	3.990	0.706	0.718	0.013
S.D.		1.42	0.737	0.078	0.072	0.088
<b>Feral Mean</b>	123	6.06	3.876	0.718	0.711	-0.013
S.D.		1.07	0.652	0.055	0.058	0.072

The results of the diversity analysis of the breed show that genetic diversity in the Corolla herd is low. Both measures of heterozygosity are among the lowest that have been found in horses. However, the allelic diversity measures  $MNA$  and  $Ae$  are especially low. The genetic variability measures confirm the effects of very small population size in this herd over the past few generations. The heterozygosity and  $Fis$  values are what would be expected from a recent population size bottleneck which certainly existed in the 1990s. The low allelic diversity indicates that much of the genetic diversity expected to be present in a horse population is gone and this can not be recovered. However, the future rate of loss can be minimized by maintaining a larger effective population size. In general terms, the minimum effective population size required to keep the rate of loss of genetic variation at about 1% per generation is 50. However, effective population size is a complex concept and the census number of horses needed to achieve an effective size of 50 is about 100 to 150 animals. More detailed analysis is underway and more discussion of conservation strategy will be done once all analyses are completed.

Comparison of allele frequencies can give measures of genetic resemblance. These are not true measures of ancestry but are approximations. Genetic similarity measure can be summarized by tree diagrams that give two dimensional representations of estimated relationships. The Corolla horse does not show close similarity to any domestic breed, mainly due to the low diversity. Figure 1 shows the Corolla herd compared to a set of domestic breeds. The Corolla herd pairs with the Brazilian Criollo, a breed with Iberian ancestry. However, the position of this couplet is between two larger branches that contain primarily American breeds, other Iberian breeds and some oriental breeds. The fit is not good as shown by the number 16 at the branch point. This is the percentage of times this branching occurred out of 100 times. Figure 2 shows the same set of breeds

but with the Shackelford Banks horse included as well. As had been shown in an earlier study of blood typing results, the Outer Banks feral horses show closer relationships to each other than to any domestic breeds which is the result seen here. The position of the branch with these two herds is not close to any of the domestic breeds. We are now doing DNA test for other Outer Banks herds and some other horses from the Southeastern US. Once these test are completed we will include them in the analysis.

Figure 1. Dendrogram of genetic similarity of the Corolla herd to domestic horse breeds. Numbers at the branching points show how many times out of 100 the cluster to the right occurred.

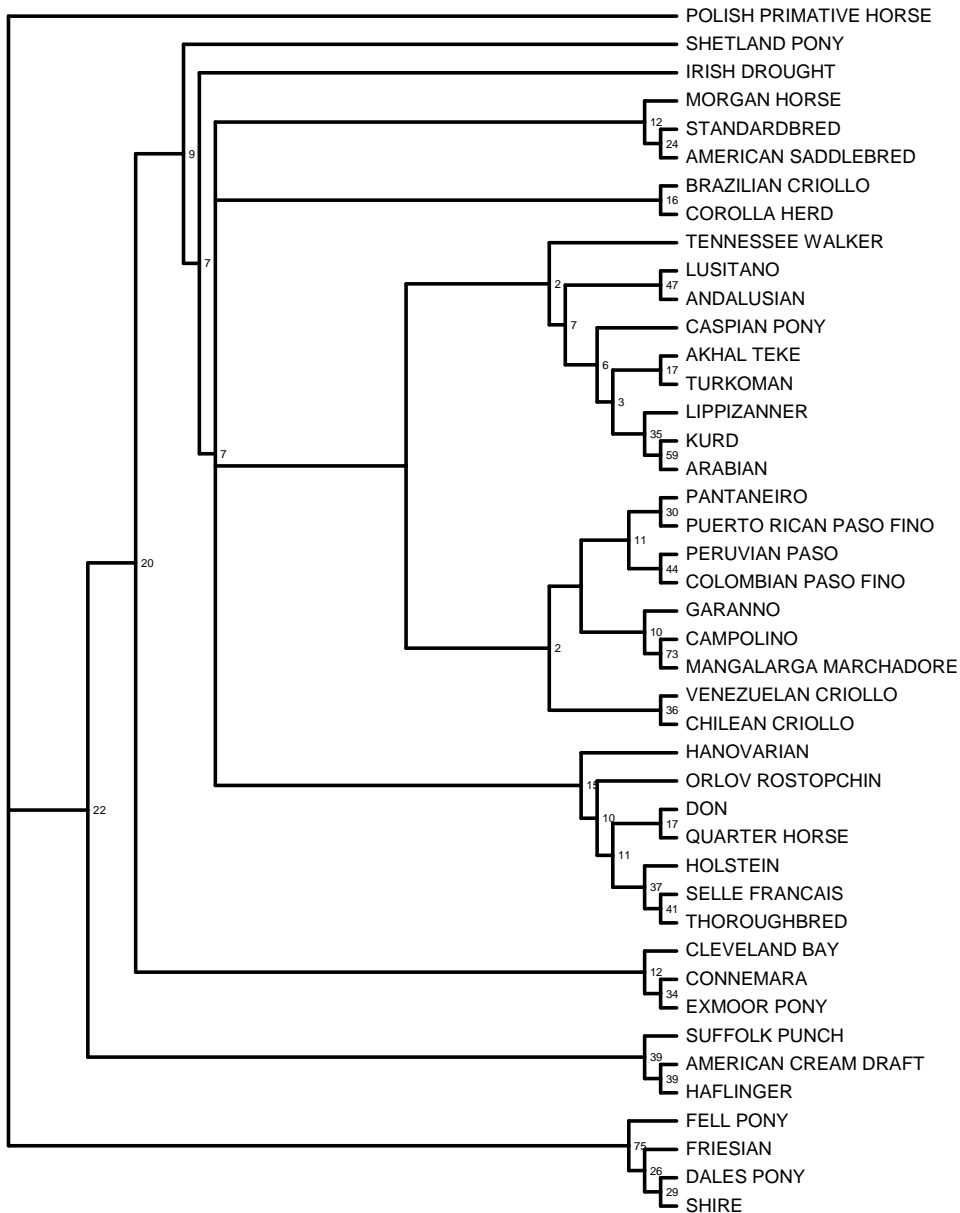


Figure 2. Dendrogram of genetic similarity of the Corolla herd to domestic horse breeds and the Shackleford Banks herd. Numbers at the branching points show how many times out of 100 the cluster to the right occurred.

