Genetic diversity and genetic distances in threatened horse breeds from Western Pyrenees by microsatellites.

A. SOLIS, B. JUGO, F. RENDO, A. ESTONBA

Dpt. of Animal Biology and Genetics. Univ. of the Basque Country, Spain.

eman ta zabal zazu

universidad   euskal herriko
del país vasco   unibertsitatea

1. INTRODUCTION

Since times immemorial there have been feral populations of horses in the Cantabrian-Pyrenean mountains and some breeds have evolved under wild or semi-wild conditions. They are phenotipically primitive and robust and well adapted to their local environment.

As a part of a larger work on the diversity of Cantabrian-Pyrenean breeds, the genetic polymorphism of four autochthonous horse breeds has been investigated:

* **Pottoka** (Pot), also named the Basque poney, can be found in the 3 provinces of the Basque Autonomus Community as well as in the French Basque Country. Attending to the coat colour and morphological characteristics they can be classified in three different sections: A, B or C.

* **Jaca Navarra** (JN), named also the Navarre poney, is a native breed located mainly in Navarre. There is no Stud-book but a Conservation Plan has been started.

* **Euskal Herriko Mendiko Zaldia** (EHMZ) also known with other different names is located mainly in Araba where five different areas (C, G, L, V and M) can be distinguised.

* **Burguete** (BU) is located in the NorthWestern part of Navarre. It is perhaps the least well defined breed. Its origin lies in the intercross of Jaca Navarra dams with sires of French breeds.

The aims of this work are twofold: Firstly, the genetic characterization of these horse populations, to quantify the genetic diversity and their phylogenetic relationships for reference for the interbreeding. The second aim of the work is the individual identification and the parentage control. Both are needed for correct management of these horse populations.

To achieve these aims we choose microsatellite markers because of their high degree of polymorphism and the ease which they can be analyzed by PCR. The usefulness of microsatellites for the estimation of genetic distances among closely related populations has been documented previously by numerous studies. Lastly they are very useful in parentage control also.

2. MATERIAL AND METHODS

BREEDS

A total of 339 animals representing four native breeds have been analysed (See Table 1). These animals were sampled trying to avoid closely related animals. In adition 18 Thoroughbred horses were included for comparative purposes.

MICROSATELLITE TYPING

DNA was extracted from the blood using standard protocols. The analysed microsatellites and the PCR conditions were those described in Stock Marks for Horses. Equine Paternity PCR typing kit (see Table 2). Samples of the 1999 Comparison Test were genotyped in order to standardize our nomenclature with other research groups.

STATISTICAL ANALYSIS

* **Expected heterozigosity**, polymorphic information content **(PIC)**, and **average exclusion probabilities** were estimated using the CERVUS program (I. Marshall, 1998)

*The test for **deviation from HWE** was performed with GENEPOP package version 1.2 (Raymond and Rausset, 1995a).This method uses an exact test of HWE with the addition that, for loci/population combinations with five or more alleles, a Markov chain algorithm is used to obtain an unbiased estimate of the exact probability of being wrong in rejecting HWE.

* Genetic differentiation between populations was analyzed across all loci using **Fst values** (Weir and Cockerham, 1984) using FSTAT program (Goudet, 1995) and their statistical significance tested using permutation test.

* A **tree diagram** was built up using the UPGMA method based on the Nei distance (Nei, 1972), also called the Ds distance. The robustness of the tree was evaluated by carrying out 1000 **bootstrap** iterations with the PHYLIP statistical package (Felsestein, 1989)
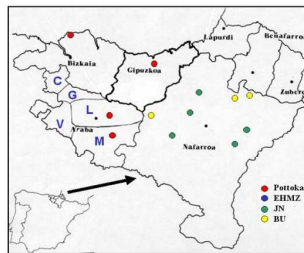
3.- RESULTS

Some parameters indicating variability at each population are presented in Table 3: MNA, observed and expected heterozigosity and PIC. There are not very big differences between populations. The greatest variability was found in the EHMZ and the lowest values were found in the populations with the lowest number of individuals analysed, indicating that more individuals should be analysed for comparison. The mean PIC can be considered as highly informative (PIC>0.5).

Six populations were in equilibrium for all markers tested, but the other five were not (see Table 4). HMS3 was not in equilibrium for 3 populations. HMS7 and HTG10 in two and HTG4 and HTG6 in one each.

As fas as the distribution within and between breeds of the observed genetic variation is concerned, 2% of the variability could be atributed to differences among breeds (mean $F_{ST} = 0.02$, p<0.001) (see Table 5).

The robustness of the tree diagram (Figure 2) was gauged by performing 1000 bootstrap iterations, and the numbers show the percentage of times a particular node was found in the 1000 bootstraped trees. As expected, Pottoka-A and Pottoka-B samples came together in 848 of the 1000 trees. Most of the EHMZ samples are grouped together (except one, EHMZC, which is a very small sample) and with Burguete breed. Further studies with a bigger sample size are needed to clarify this cluster. The more distant population corresponds to Thoroughbred.

4. CONCLUSIONS

* In accordance with other investigations our results indicate that , in general, variability at microsatellite loci in these populations is as high as that of other populations, being a useful tool in populations studies and parentage testing.

* These data can not be taken as definitive given that few animals were analysed in Jaca Navarra and Burguete breeds.

We inted to extend our analyses with more animals in these breeds and with more microsatellite markers in order to build up a complete picture of the genetic structure and relationships of these horse populations.
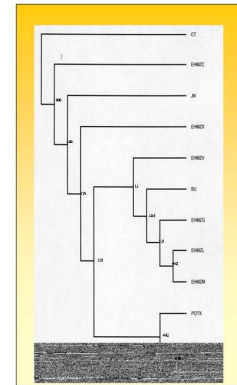
Table 3 Mean number of alleles per locus (MNA), observed and expected heterozigosities, Polymorphic Information Content and Exclusion probability.

BREED	N	MNA	H_O	H_E	PIC	EXCL.
EHMZ	144	10.50	0.739	0.755	0.722	0.999954
EHMZmix	22	5.92	0.643	0.670	0.610	0.999020
Burguete	13	4.67	0.723	0.670	0.601	0.998929
Pottoka A	71	8.58	0.739	0.762	0.728	0.9999959
Pottoka B	35	7.83	0.655	0.751	0.707	0.999999
Pottokamix	16	7.67	0.748	0.764	0.709	0.99958
Jaca Navarra	14	5.83	0.750	0.735	0.671	0.992173
Thoroughbred	18	4.25	0.625	0.651	0.575	0.998010

Table 4. Markov chain probability test for HWE proportions. A light gray cell indicates a significant deviation at the P = 0.05 level and a dark gray cell indicates a significant result at the P<0.01 level

	MZC	MZG	MZL	MZM	MZV	POTA	POTB	MZX	POTX	BU	JN
AHT4	9572	6254	1606	8772	7574	1893	1320	4001	5188	2707	7041
AHT5	7715	3618	3579	9701	3236	4628	5427	1186	0529	3664	3745
ASB2	8182	5889	9018	4444	6945	5418	3243	5973	1797	7213	2937
HMS2	2386	6563	8069	2307	7461	5216	8412	2814	5902	6288	3141
HMS3	4341	3028	5453	8243	5468	0000	0392	0322	9663	7657	3207
HMS6	5273	0756	1630	5600	8684	8815	0724	5230	4287	8041	6809
HMS7	6364	4440	2212	9131	0454	1978	0817	0290	2833	9612	6106
HTG10	8377	4556	3442	7891	2331	0227	0000	0762	4237	7811	5913
HTG7	4108	7996	5779	5648	3445	1251	3685	7674	4200	1	6171
HTG4	7514	7380	3818	0218	2716	4488	9178	1	9616	3055	1575
HTG6	1602	4143	1513	8460	0000	1956	6723	2963	2995	1	4807
VHL20	7514	6335	8881	3141	2960	2078	1308	5127	5441	6333	9206

Table 5. F statistics estimates and their significance by locus in the breeds analysed. Significance: * 0.05>p>0.01; ** 0.01>p>0.001;*** p<0.001

	Fis	Fit	Fst
AHT4	0.043 ±0.043	0.024 ± 0.044	0.020 ± 0.003
AHT5	0.018 ± 0.023	0.013 ± 0.022	0.005 ± 0.003
ASB2	0.077 ± 0.034	0.089 ± 0.026***	0.013 ± 0.002
HMS2	0.011 ± 0.031	-0.013 ±0.039	0.004 ± 0.003
HMS3	0.097 ± 0.047***	0.052 ± 0.036***	0.036 ± 0.013 ***
HMS6	0.046 ± 0.039	0.018 ± 0.037*	0.028 ± 0.013 ***
HMS7	0.017 ± 0.037	0.010 ± 0.042	0.016 ± 0.002**
HTG10	0.057 ± 0.034	0.039 ± 0.037	0.011 ± 0.007 **
HTG7	0.028 ± 0.026	-0.030 ± 0.034	0.034 ± 0.013 ***
HTG4	0.039 ± 0.024	0.028 ± 0.024	0.012 ± 0.007
HTG6	0.059 ± 0.039***	0.072 ± 0.037***	0.018 ± 0.003 ***
VHL20	0.039 ± 0.026	0.039 ± 0.032***	0.012 ± 0.003
All loci	0.052 ± 0.011***	0.032 ± 0.011***	0.021 ± 0.002 ***

Table 1 Breeds analysed, number of animals sampled and estimated population size.

Breed	Location	Number of animals sampled	Population size*	Studbook
Euskal Herriko mendiko zaldia		166	4.500	Yes (1999)
	-Llanada Alavesa	58		
	-Valles Alaveses	32		
	-Montaña Alavesa	31		
	-Cantábrica Alavesa	9		
	-Gorbea	24		
	- Mixed	22		
Burguete		13	5.200	No
Pottoka		146	1.141-2.000#	Yes (1995)
	-A section	89		
	-B section	41		
	-C section	16		
Jaca Navarra		14	250	No
Thoroughbred		18		
TOTAL		337		

* Naturalis (1999) 14-157-165
# Bizkaiko Foru Aldundia (1997), Carreras (1962)

Table 2. Analysed microsatellites and its conditions

| Microsatellite	Cromosome	Referece
VHL20	30	Van Haeringen (1994)
HTG4	9	Ellegren et al. (1992)
ASB2	24q4	Greima et al. (1994)
HMS7	1	Greima et al. (1994)
HTG6	Eq23q2?	Ellegren et al. (1992)
HMS6	4	Greima et al. (1994)
HTG7	4	Marklund et al. (1994)
HMS3	9	Greima et al. (1994)
AHT5	6	Binns et al. (1995)
HTG10	21	Marklund et al. (1994)
HMS2	10	Greima et al. (1994)

Figure 1. The Basque Autonomous Community with its 3 provinces: Bizkaia, Gipuzkoa and Araba, and Navarre. Sampling points are indicated.
Figure 2. (on the right). UPGMA tree between Pyrenean-Cantabrian horse populations.Thoroughbred is included.