Genetic Characterization Of The Pinzgauer Breed Reared In Italy

E. D'Agaro, M. Cimolino, A. Zof, and F. Costa*

Introduction

The Pinzgauer horse, also known as the Noriker horse, is a moderately heavy Austrian draught breed. The Pinzgauer horse is an indigenous breed of the central Alpine region of Europe and is believed to have originated in the Salzburg region since 1897. In early history, this region became known as the Roman province Noricum. At the end of 19th century, the original name "Pinzgauer" was changed into "Noriker", due to the romanophile attitude at that time. The Pinzgauer breed has played an important role for several centuries in the transportation of goods through the Alps. Nevertheless, in the last decades, heavy draught horses has become increasingly unimportant because of the intense mechanization of agriculture and transport systems, and, therefore, the populations decreased to alarmingly low numbers in the 1970s. Fortunately, a governmental support and the work of several horse breeders helped to save this culturally important breed from extinction. Pinzgauer horses started to be imported from Austria to Friuli Venezia Giulia (FVG) region in recent years and the current population (15-20 animals) derives mainly from local animals and crosses with Austrian horses. In recent years, many studies have used the mitochondrial D-loop region (Cozzi *et al*., 2004), the most variable part of mtDNA, as a genetic marker in phylogenetic horse studies. Here, we report the preliminary results of the genetic analysis of the Pinzgauer horse populations in the FVG and South Tirol regions.

Material and methods

Buccal spit samples were collected from Pinzaguer horses reared in the FVG region (n=15) and South Tirol (n=20) using FTA swabs. Polymerase chain reaction (PCR) primers for the mitochondrial hyper variable region were designed according to the published horse sequence (X79547; Xu and Arnason, 1994). The D-loop region of the mtDNA was amplified using the primers (forward primer) 5'-AAC GTT TCC TCC CAA GGA CT-3' and reverse primer 5'-GCA TTT TCA GTG CCT TG CTT-3', resulting in a 478-bp fragment. PCR was performed in a 25-μl reaction containing approximately 50 ng of horse DNA, 10x PCR buffer, 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1U of Taq polymerase (Sigma, Milano) and 0.5 μM of each primer. The amplification was carried out in a Sprint Thermal Cycler system (THERMO HYBAID, Ashford, UK) under the following conditions: an initial denaturation step at 94°C for 3 min, followed by 34 cycles at 94°C for 45s, annealing at 62°C for 45s and a final extension of 80 s at 72°C. Visualization was performed on a 1.5% ethidium bromide

* University of Udine Via S. Mauro 2, 33010 Pagnacco (UD), Italy.

stained agarose gel. Successful PCR products were purified by the Quiagen purification Kit (Quiagen, Milan), and then sequenced using the Big Dye Terminator method (Applied Biosystem) on an ABI 3770 automated sequencer. For most samples, the forward and reverse sequences were obtained. Multiple sequence alignments were analyzed using the DNA sequence polymorphism DNASP software version 4.0 (Rozas *et al*., 2003). Haplotype diversity (Hd) was computed using equations suggested by Nei (1987). Nucleotide diversity was estimated from the number of polymorphic segregating (S) sites (Watterson 1975; Nei, 1987). Calculations were based on all (silent and non synonymous) segregating sites. Neutrality test statistics D, D* (Tajima, 1989), and F* (Fu and Li, 1993) were calculated using 10,000 simulations to test the hypothesis that mutations in the gene are selectively neutral (Kimura, 1983). Phylogenetic inference was performed using the Neighbour-Joining (NJ), the maximum likelihood (ML) (PhyML and RaxML) and the Mrbayes (MB) analysis using TOPALI V2.5 software. The optimal model of nucleotide evolution for NJ and ML analyses was determined using the model test which is implemented in the program package. Model selection is done running ML on each model (in parallel on clusters or multi-core machines). The results are sorted in ascending order according to the BIC score (BIC=-2lnl (log likehood)+ df (degrees of freedom)+ log (n) (sample size)). Model selection was also tested for rate of heterogeneity (proportion of invariant sites and gamma distribution). Model selection for MrBayes was tested as a subset of the above models.

Results and discussion

Pairwise alignment of Pinzgauer haplotypes with the reference sequence (GeneBank X79547, Xu and Arnason, 1994) showed a high identity pattern (97%) and 0 gaps. A detailed description of the results obtained on the haplotype and nucleotide diversity analysis for the aligned nucleotide sequences are presented in Table 1.The sequence alignment consisted of about 478 bp. We found 272 polymorphic sites (56%) and 93 parsimony informative sites. The nucleotide frequencies were: A, 28.97; C, 29.39; G, 15.48 and T/U, 26.15%; average pair wise distance: 0.8454; transition/transversion ratio: 0.52. Haplotype (Hd) and nucleotide diversity values (Pi , $\boldsymbol{\pi}$) were: 1 ± 0.177 (mean ± SD) and 0.6147 ± 0.1536, respectively. The nucleotide diversity (the average number of nucleotide differences per site between two sequences) calculated using the Jukes and Cantor correction was: 2.2267. Theta ( ) (per site) was calculated from Eta (total number of mutations) or from S (the number of segregating (polymorphic) sites) on base pair basis. Theta ( ) = 4Nm for an autosomal gene of a diploid organism (N and m are the effective population size and the mutation rate per nucleotide site per generation, respectively). The variance of this estimator depends on the recombination between sites. These variances were computed on a per nucleotide site basis: variance (per nucleotide site) = variance (per DNA sequence) / m*m where m is the total number of nucleotides studied. The level of polymorphism can be quantified by θ estimate of the neutral mutation (Watterson 1975). To allow the comparison between different regions, we estimated θ per site. In the present data, theta was 0.4975. This value is consistent with previous estimates (Fu and Li, 1993). This parameter summarizes the rate at which processes of mutation and random genetic drift generate and maintain variation within a gene, assuming that natural selection has not been operating. Although the number of segregating sites does not represent all the information in the sample, under the neutral infinite-sites model the frequency spectrum of sites is determined by , which in turn is estimated by S. Applying the likelihood ratio test procedure, the selected model of DNA substitution was the H81 model (Hasegawa *et al*,

1985). Molecular phylogeny tree of Pinzgauer populations reared in FVG and South Tirol inferred by ML method is shown in figure 1. The phylogenetic analysis resulted in a congruent tree topology, in agreement with the population geographical distribution, for the NJ, MB and ML (PhyML and RaxML). All trees revealed similar topology patterns with only minor differences, indicating that these methods are relatively insensitive to the initial assumptions. A notable separation from haplotypes imported from Austria, the out-group and the other groups was observed. The information obtained in this study should be helpful in making conservation decisions.

Table 1: Summary of the statistical analysis performed on the aligned sequences of Pinzgauer breed

Haplotype (gene) diversity, Hd: 1.000
Variance of haplotype diversity: 0.0312
Standard deviation of haplotype diversity: 0.039
Nucleotide diversity, Pi: 0.7404
Sampling variance of Pi: 0.0008
Standard deviation of Pi: 0.0299
Nucleotide diversity (Jukes and Cantor), Pi(JC): 2.6348
Theta (per site) from Eta: 0.9553
Theta (per site) from S, Theta-W: 0.3414
Variance of Theta (no recombination): 0.0717
Standard deviation of theta (no recombination): 0.1353
Variance of Theta (free recombination): 0,0005
Standard deviation of theta (free recombination): 0.0236

Figure 1: Molecular phylogeny of Pinzgauer populations inferred by ML method

Conclusion

This is a preliminary study on the genetic diversity of Pinzgauer populations reared in FVG and South Tirol regions. This research provides new morphological and genetic data concerning the horse populations living in the two Italian regions. Now, an efficient strategy for a conservation programme can be developed.

References

Cozzi, M.C., Strillacci, M.G., Valiati, P., Bighignoli, M., Cancedda, M, Zanotti, M (2004). *Genet. Sel. Evol*., 36: 663–672.

Fu, Y. X., Li, W. H. (1993). *Genetics,* 133: 693-709.

Hasegawa, M., Kishino, H. Yano, T. (1985). *J. Mol. Evol.* *22*: 160–174.

Kimura, M. (1983). *The neutral theory of molecular evolution. Cambridge* Univ. Press, Cambridge,UK.

Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia Univ. Press, New York.

Rozas, J.C., Sanchez-Delbarrio, X., Rozas, R., 2003. *Bioinformatics,* 19: 2496-2497.

Tajima, F., 1989. *Genetics,* 123: 585-595.

Watterson, G.A., 1975. *Theor. Pop. Biol.,* 7: 256-276.

Xu, X, Arnason, U. (1994). *Gene,* 148(2): 357-62.