

## **Genetic diversity and Population Structure of locally adapted horse breeds in Brazil.**

*P. Ianella<sup>1</sup>, M.S.A. Maués<sup>1</sup>, A.R. Caetano<sup>1</sup>, G.R. Biazio<sup>1</sup>, S. R. Paiva<sup>1</sup>, A.R. Nepomuceno<sup>2</sup>, L.F.R. Carvalho<sup>3</sup>, C.M. McManus<sup>4</sup>.*

<sup>1</sup> – *Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, Brazil.*

[patricia.ianella@embrapa.br](mailto:patricia.ianella@embrapa.br)

<sup>2</sup> - *Ministério da Agricultura, Pecuária e Abastecimento, Brasília, DF, Brazil.*

<sup>3</sup> – *Pós-Graduação em Biologia Animal, Universidade de Brasília, Campus Darcy Ribeiro, Brasília, DF, Brazil.*

<sup>4</sup> - *Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, Campus Darcy Ribeiro, Brasília, DF, Brazil.*

### **Summary**

Since the introduction of horses in Brazil a variety of local breeds with highly specific phenotypes has arisen. These breeds are important to local communities since they are strongly associated with subsistence farming activities. In order to subsidize management decisions this study investigated the genetic diversity and population structure from eight horse breeds locally adapted in Brazil (Campeira, Crioula, Lavradeira, Pantaneira, Mangalarga Marchador, Puruca pony and Baixadeiro genetic group). Data from 243 horses genotyped using the Illumina EquineSNP70 BeadChip were analyzed and results indicate low differentiation among most of the studied breeds and moderate differentiation among Crioula and the others. Population structure analyses revealed that breeds are clustered in genetic groups highly correlated with geographical proximity. This pattern suggests similar origin and/or crossbreeding among these breeds. The results provide a better definition than microsatellites markers and have showed new insights into the relationships among Brazilian horse breeds, highlighting the need for better management strategies for conservation of these populations.

*Keywords: creoule breeds; Equus caballus; SNP*

### **Introduction**

Since domestication horses have been widely used for transportation, agriculture and warfare, and therefore selected for physical performance and appearance traits, which resulted in a worldwide distribution of hundreds of breeds (Petersen et al. 2013). This species was introduced in Brazil after 1534 by Spanish and Portuguese settlers (Primo 2004). Five centuries of evolutionary processes shaped currently existing breeds, which have become adapted to local environmental and sanitary conditions, giving rise to a variety of local horse populations with highly specific phenotypes. Brazilian locally adapted breeds descend from Iberian Peninsula breeds and include breeds such as Campeira, Crioula, Lavradeira, Pantaneira, Mangalarga Marchador, as well as animals of smaller stature such as the Puruca pony and the Baixadeiro genetic group. These Brazilian horse breeds are extremely well adapted to local conditions and are widely dispersed throughout the country. Despite the local importance of these breeds, few studies have been carried out to investigate genetic diversity and population structure and determine relationships between them with SNP markers. The

present study aimed to better understand the genetic composition and relationships among these local breeds to provide support to development of better management and conservation strategies.

## Material and methods

A total of 243 DNA samples from eight locally adapted horse breeds from North (Baixadeira n = 20, Lavradeira n = 20, Marajoara n = 20, Puruca n = 20), South (Campeira n = 20, Crioula n = 11), Southeast (Mangalarga Marchador n = 20) and Centerwest regions (Pantaneiro n = 112), derived from Embrapa Genetic Resources and Biotechnology Genetic Bank were genotyped using EquineSNP70 BeadChip Illumina BeadChip (Illumina, Inc.). Initial quality control was performed in SNP & Variation Suite v7.x (Golden Helix) removing samples with call rates < 0.90, and markers with unknown map positions, call rates < 0.95, and minor allele frequencies < 0.05 across all breeds. In addition, pruning based on linkage disequilibrium was used to remove SNPs with  $r^2 > 0.5$  within a 50-SNP sliding window. Population structure and degree of admixture was evaluated using ADMIXTURE 1.3 (Alexander et al. 2009) and SNP & Variation Suite v7.x for principal component analysis (PCA). ADMIXTURE analysis were performed with a number of hypothetical pseudo-populations, K, which ranged from 2 to 11. To evaluate optimal partitioning, cross-validation (CV) error values were computed for each K using a 10-fold cross-validation procedure. To visualize clustering across runs of K, ADMIXTURE outputs were submitted to the CLUMPAK pipeline (Kopelman et al. 2015).  $F_{ST}$  estimates were obtained using Arlequin v 3.5.2.2 (Excoffier and Lischer 2010).

## Results and Discussion

Five samples (1 EBA, 2 ECA, 1 ECR and 1EMA) and 22.749 SNPs were excluded after quality control analysis. LD pruning resulted in a final set of 33,681 SNPs from 238 individuals for downstream analysis. Estimated genetic differentiation indices ( $F_{ST}$ ) between breeds are shown in Table 1 ( $P < 0.00001$ ). The lowest and highest observed genetic differences were between Puruca and Marajoara ( $F_{ST} = 0.012$ ) and between Crioula and Lavradeira ( $F_{ST} = 0.089$ ), respectively. Crioula breed showed moderate differentiation from all other breeds.

Table 1: Estimated  $F_{ST}$  indexes among locally adapted horse breeds from Brazil.

Breed	EBA	ECA	ECR	EL	EMA	EM	EPA
EBA	0.0						
ECA	0.049	0.0					
ECR	0.082	0.076	0.0				
ELV	0.034	0.057	0.089	0.0			
EMA	0.015	0.046	0.078	0.03	0.0		
EMM	0.037	0.040	0.080	0.049	0.035	0.0	
EPA	0.027	0.037	0.068	0.04	0.027	0.032	0.0
EPU	0.015	0.049	0.082	0.034	0.012	0.039	0.03

EBA – Baixadeira; ECA – Campeira; ECR – Crioula; ELV – Lavradeira; EMA – Marajoara; EMM – Mangalarga; EPA – Pantaneira; EPU – Puruca. All pairwise analysis were significant ( $P < 0.00001$ )

$\Delta K$  estimates obtained with CLUMPAK revealed a best K value of 4. Breeds can be separated according to geographic proximity. ECR was the less admixed breed and EPA was

the most. EMM and ECA have major contributions from one specific population, while EBA, ELA, EMA and EPU from yet another. Estimated pairwise  $F_{ST}$  values corroborates PCA (Figure 1) and ADMIXTURE (Figure 2) results. Altogether, Lavradeiro, Puruca, Marajoara and Baixadeiro breeds are close to each other when compared to the rest of studied breeds. According to Reis *et al.* (2008) Marajoara horses were derived from Portuguese horses from Cabo Verde, while the Puruca pony was derived from nine Shetland ponies imported from France. Albeit Marajoara horses are geographically isolated on Marajo Island, these results can suggest that a certain degree of crossbreeding may have occurred with northern breeds. Alternatively, observed results may be a consequence of eventual crossbreeding with specialized breeds to increase physical stature and improve performance. It can be also observed that ECR and ECA seem to be the most genetically differentiated from other breeds. Crioula horses have been subject to long term breeding strategies with a closed herd book maintained for almost 100 years (Maciel *et al.*, 2014), and were the most distant genetically in relation to the other breeds.

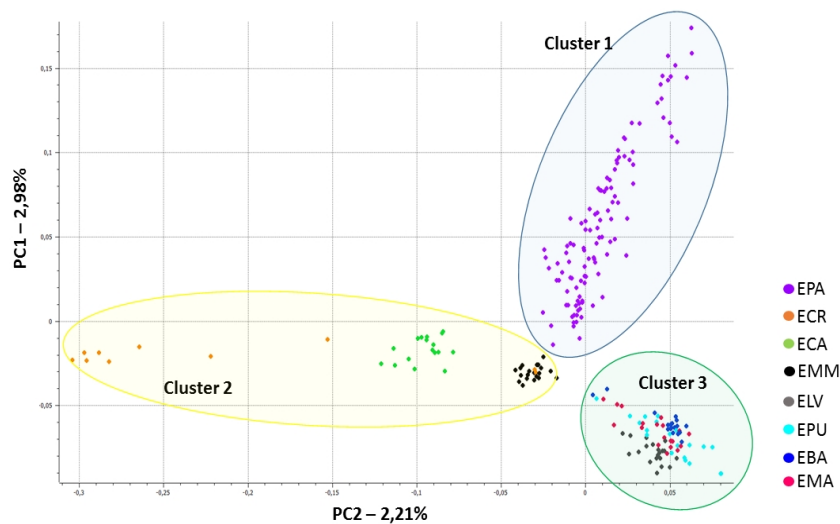


Figure 1 – Principal component analysis plot for Brazilian locally adapted horse breeds. PC1 and PC2 explain 2.98 and 2.21% of the total variance observed, respectively (EBA – Baixadeira; ECA – Campeira; ECR – Crioula; ELV – Lavradeira; EMA – Marajoara; EMM – Mangalarga; EPA – Pantaneira; EPU – Puruca).

Maintenance and management of local populations have often been neglected. Previous reports with microsatellites (Reis *et al.* 2008; Silva *et al.* 2012) and mtDNA (Ianella *et al.*, 2017) suggest variability loss in Brazilian horses breeds. In the present study, apparent crossbreeding or introgressions observed among northern breeds highlight the need for better management strategies for these populations. Genetic characterization is an essential tool for proper breed conservation as it contributes to increased knowledge about the level of genetic diversity and crossbreeding among breeds, while helping subsidize the development of breeding strategies and management plans. Since locally adapted breeds are usually found in highly challenging environments in association with subsistence farming activities, strategies and local policies for conservation and management strategies are of great importance for the food and economic security of local farming communities.

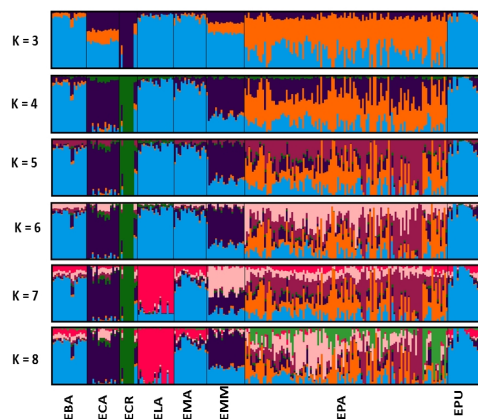


Figure 2 – Structure analysis of 235 individuals from eight populations. Population structure graphic representation. Each animal is represented by a vertical line divided into segments classified according to size and color, corresponding to genetic composition from inferred clusters. Distinct populations are separated by black lines (EBA – Baixadeira; ECA – Campeira; ECR – Crioula; ELV – Lavradeira; EMA – Marajoara; EMM – Mangalarga; EPA – Pantaneira; EPU – Puruca).

## References

- Alexander DH, Novembre J and Lange K (2009) Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Res* 19:1655–1664.
- De Assis JB, DeLaat DM, Peixoto MGCD, Bergmann JAG, Fonseca CG and Carvalho MRS (2009) Genetic diversity and population structure in Brazilian Mangalarga Marchador horses. *Genet Mol Res* 8:1519–24.
- Excoffier L and Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567.
- Ianella P, Albuquerque MSM, Paiva SR, Egito AA, Almeida LD, Sereno FTSP, Carvalho, LFR, Mariante AS, McManus CM (2017) D-Loop haplotype diversity in Brazilian horse breeds. *Genet Mol Biol*. Epub August 31.
- Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA and Mayrose I (2015) Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour* 15:1179–1191.
- Maciel FC, Bertoli CD, Braccini Neto J, Cobuci JA, Paiva SR and McManus CM (2014) Population structure and genealogical analysis of the Brazilian Crioula Horse. *Anim Genet Resour* 54:115–125.
- Petersen JL, Mickelson JR, Cothran EG, Andersson LS, Axelsson J, Bailey E, Bannasch D, Binns MM, Borges AS, Brama P et al. (2013) Genetic Diversity in the Modern Horse Illustrated from Genome-Wide SNP Data. *PLoS One*.
- Primo A (2004) *Conquista e Colonização: A fantástica História dos Conquistadores Híbericos e Seus Animais na Era dos Descobrimentos*. Editora movimento, Porto Alegre
- Reis SP, Gonçalves EC, Silva A and Schneider MPC (2008) Genetic variability and efficiency of DNA microsatellite markers for paternity testing in horse breeds from the Brazilian Marajó archipelago. *Genet Mol Biol* 31:68–72.
- Silva ACM, Paiva SR, Albuquerque MSM, Egito AA, Santos SA, Lima FC, Castro ST, Mariante AS, Correa PS and McManus CM (2012) Genetic variability in local Brazilian horse lines using microsatellite markers. *Genet Mol Res* 11:881–90.
- SNP & Variation Suite™ (Version 8.x) [Software]. Bozeman, MT: Golden Helix, Inc. Available from <http://www.goldenhelix.com>.