Genetic Diversity and Relationships among Spanish Beef Breeds Assessed by a Bovine High-density Chip

J.J. Cañas-Álvarez[1], A. González-Rodríguez[2], S. Munilla[2], L. Varona[2], C. Díaz[3], J. A. Baro[4], C. Moreno[2], A. Molina[5] and J. Piedrafita[1]

[1]Universitat Autònoma de Barcelona, Spain, [2]Universidad de Zaragoza, Spain, [3]INIA, Spain, [4]Universidad de Valladolid, Spain, [5]Universidad de Córdoba, Spain

ABSTRACT: High-density SNP genotyping was used to describe the genetic variability and relationships among five Spanish beef cattle breeds (Asturiana de los Valles, Avileña – Negra Ibérica, Bruna dels Pirineus, Pirenaica and Retinta). A total of 239 samples were genotyped with the BovineHD (778 K) Beadchip, which proved its suitability for the study of these breeds. Genetic diversity, measured as the expected heterozygosity, was around 0.31, similar to other European cattle breeds. The degree of genetic differentiation was little to moderate, as the F_{ST} estimates ranged from 0.024 to 0.065. PCA analysis was able to distinguish five separated clusters corresponding to each breed. The clusters were confirmed through an analysis for ancestries which also showed admixture among some breeds. This contributes background information to explore the feasibility of genomic selection in Spanish beef breeds.
Keywords: SNP variability, beef cattle

Introduction

The genetic histories of the populations have involved processes of isolation, drift and selection or adaptation to their particular habitat. The availability of SNP chips for massive genotyping has been useful to characterize genetically populations of domestic cattle and to establish their degree of divergence (Gautier et al. (2007); de Roos et al. (2008)). Several studies have demonstrated that the predictive ability of SNP effects estimated in one population to estimate the GEBV of candidates of selection from another population depends on the map density and the genetic distance between populations (see review of Garrick (2011)). Using microsatellite markers it has been demonstrated that the Spanish cattle breeds are related (Beja-Pereira et al. (2003); Martin-Burriel et al. (2011)), but the relationships among the breeds assessed from genomic data has not been investigated so far. Thus, using a high- density SNP chip, in this study we aim at characterizing the diversity of five Spanish beef breeds and evaluating their degree of relationship as a first step towards exploring the feasibility of genomic selection in these breeds.

Materials and Methods

Animals and sampling. A total of 239 blood samples were collected from the caudal vein of animals belonging to five Spanish beef cattle populations, including Asturiana de los Valles (AV, 50), Avileña – Negra Ibérica (ANI, 48), Bruna dels Pirineus (BP, 50), Pirenaica (Pi, 48) and Retinta (Re, 43) breeds. The animals were the parents of 116 trios and were chosen as less unrelated as possible. Genomic DNA was extracted by standard protocols.

SNP genotyping and diversity statistics. High-density SNP genotyping was performed by using the BovineHD Genotyping Beadchip (IlluminaInc, USA) designed to genotype 777,962 SNPs, according to the protocol of the manufacturer at a commercial laboratory (Xenética Fontao, Lugo, Spain). For this study, SNPs mapped to chromosomes X and Y, mitochondria, or which were in repeated positions, as well as with Mendel errors > 0.05 were removed. Quality controls were performed using PLINK v1.07 software (Purcell et al. (2007)). Descriptive statistics for each breed included the percentage of markers genotyped in more than 5% of the samples, markers with a MAF > 0.05, expected heterozygosity, mean number of alleles (MNA) and percentage of loci in H-W equilibrium ($p>0.01$).

Relationships among breeds. To analyze the divergence among breeds, an individual call rate of ≥ 0.95 and SNP call rate ≥ 0.95 were required. As background LD can affect both PCA and clustering analysis, we thinned the marker set by excluding SNPs in strong LD (i.e., with a pairwise $r^2 > 0.1$). The final result was 54,693 SNPs left for the analysis. Pairwise F_{ST} statistics (Wright (1951)) and standard D distances (Nei (1972)) were computed by ARLEQUIN v3.5 (Excofier et al. (2005)) and PHYLIP v3.695 (Felsenstein (1989)), respectively. Neighbor-Joining (N-J) tree estimation (Saitou and Nei (1987)) from Nei distances was implemented in APE R-package (Paradis et al. (2004)). Clustering analysis for ancestries was performed using ADMIXTURE (Alexander et al. (2009)). PCA analysis and another intermediate calculus and figures were performed in an R environment (http://www.r-project.org).

Results and Discussion

Genetic diversity analysis. The proportion of markers genotyped on 95% of the samples was around 97%-98% in all breeds (Table 1) which suggests the suitability of the chip to characterize the breeds studied. MAF > 0.05 values ranged between 86-89%, the frequencies for MAF > 0.01 being only slightly bigger. Our values are higher than those of Gautier et al. (2007) obtained across 696 SNPs. These authors found MAF > 0.05 in 73%-83% of the European breeds, while for African breeds the estimates were in the range 47%-71%. Our

bigger estimates could be related to the criteria used to select the markers included in the chip. The distribution of the less frequent allele was very similar among breeds at different ranges of allele frequencies (Figure 1). All this suggests a similar within breed variability. The mean number of alleles was similar among breeds, ranging from 1.86 to 1.89. As a consequence, the expected heterozygosity was also very similar among breeds (0.30 – 0.32) and also close to that observed by Gautier et al. (2007) in European breeds. The percentage of markers in H-W equilibrium was always over 99%.

Table 1. Genetic variability in five Spanish beef breeds.

Breed	Markers genotyped on >95% of the sample (%)	Markers with MAF > 0.05 (%)	Expected hetero-zygosity (SD)	MNA	Markers in HWE (P > 0.01) (%)
AV	97.87	89.24	0.319 (0.172)	1.89	99.38
ANI	98.03	87.13	0.306 (0.177)	1.88	99.17
BP	97.06	88.17	0.309 (0.176)	1.88	99.31
Pi	98.22	86.60	0.299 (0.180)	1.87	99.43
Re	98.03	86.39	0.305 (0.178)	1.87	99.27

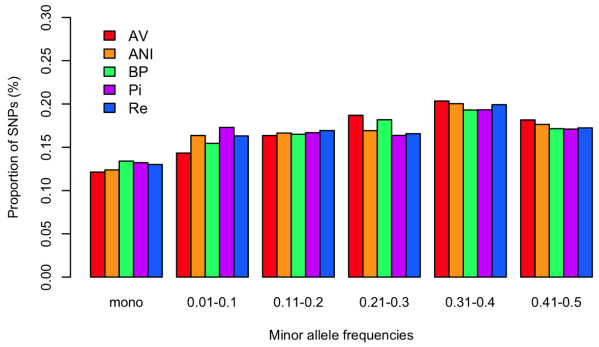

Figure 1: Minor allele frequency distribution

Relationship among breeds. Principal component analysis showed five clusters (Figure 2). The F_{ST} estimates show little (0.024) to moderate (0.065) differentiation among breeds (Table 2). The lowest F_{ST} estimated were observed between AV and the rest of the breeds, and the highest between Pi and both ANI and Re (> 0.06). The N-J tree computed from the Nei distance (Figure 3) shows a main group including ANI and Re, more closely related, and AV, as well as two more independent breeds. These results parallel with the ones found by Martín-Burriel et al. (2011) using 19 microsatellite markers and the Reynolds distance.

Table 2. F_{ST} statistics (below the diagonal) and Nei D distances (above the diagonal) among populations.

Breed	AV	ANI	BP	Pi	Re
AV		0.0127	0.0101	0.0148	0.0137
ANI	0.0329		0.0177	0.0207	0.0149
BP	0.0245	0.0505		0.0149	0.0190
Pi	0.0412	0.0620	0.0430		0.0218
Re	0.0358	0.0412	0.0543	0.0655	

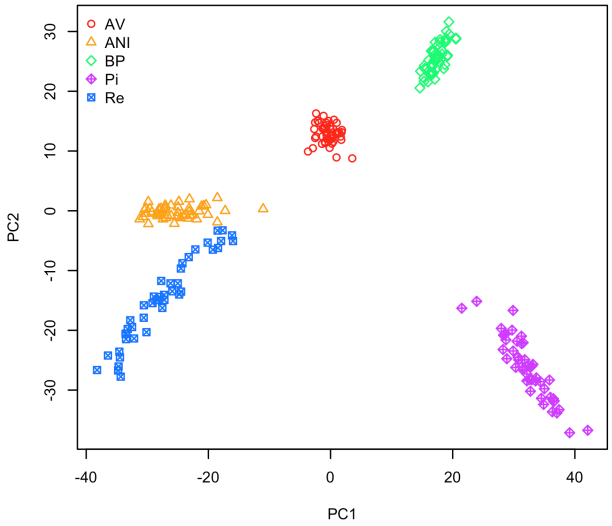

Figure 2: Principal component analysis

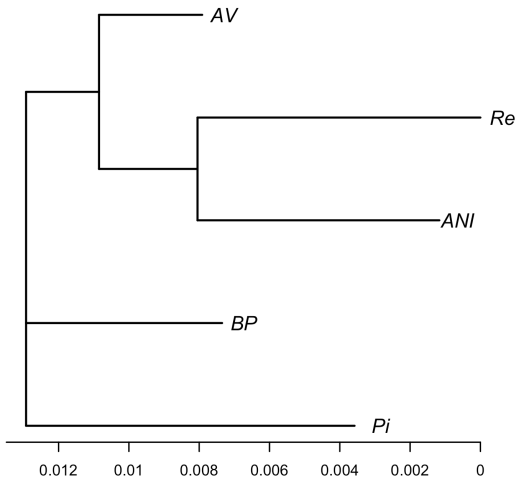

Figure 3: N-J representation of the pairwise Nei distances

The maximum likelihood estimation of ancestries assigned all individuals to clusters that coincide with the population of origin although also reveals some admixture among populations. A first clustering (K = 2) revealed the existence of two main groups, ANI-Re and BP-Pi, with AV sharing about half of these clusters. This is consistent with the N-J tree. The lowest cross validation errors (CVE) were 0.40201for K = 4 and 0.40210 for K = 5, indicating that 4 and 5 were the most parsimonious number of clusters. It is advised to choose the K with minor CVE, although in this case both values were very similar and then K = 5 (five clusters) was chosen as it can reflect better the relationships among breeds as depicted by PCA analysis. The clusters appear much more clearly defined that in the study of 40 Iberian cattle breeds (Martín-Burriel et al. (2011)), suggesting that high-density genotyping is a more powerful tool than microsatellites to unravel the relationship among breeds. AV contributed to all breeds, but mainly to Re and ANI. It is also important the contribution of ANI to Re, possibly related to transhumance (seasonal migrations) in search for pasture. BP contributed to AV, confirming a previous result of Martín-Burriel et al. (2011), who detected the contribution of the Brown Swiss, the origin of the BP breed, to the AV breed. BP also contributed to Pi breed, which could be explained by interchanges dues to their geographical proximity in the Pyrinean mountains.

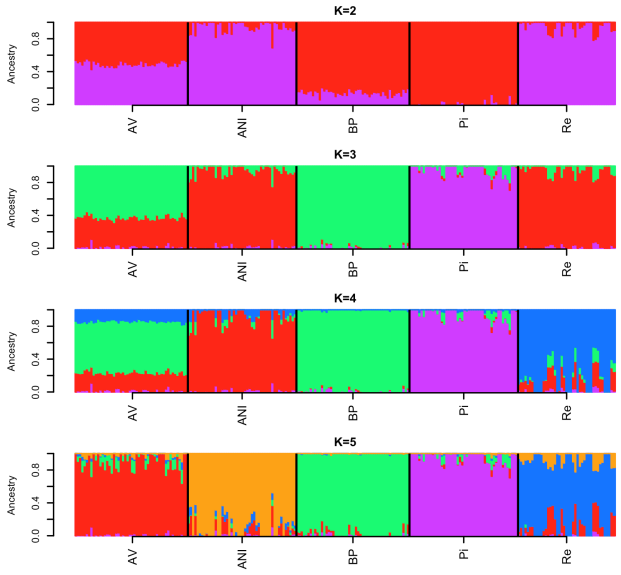

Figure 4: Estimated membership coefficients for each individual for K = 2-5

Conclusion

High-density SPN markers have proven to be useful to accurately describe the variability and relationships among Spanish beef breeds. Expected heterozygosity was in the range of other European cattle breeds. F_{ST} estimates indicated that the degree of differentiation among breeds ranked from little to moderate. It has been possible to differentiate clusters coincident with the breeds, notwithstanding the signals of admixture between some of them.

Acknowledgements

The financial support of Spanish AGL2010-15903 and UE FP7-289592-GENE2FARM grants made possible this resear
ch. The COLCIENCIAS Francisco José de Caldas 497/2009 and Spanish government BES-2011-045434 fellowships awarded to J. J. Cañas-Álvarez and to A. González-Rodríguez, respectively, are also recognized.

Literature Cited

Alexander, D. H., Novembre, J., and Lange, K. (2009). *Genom. Res*., 19, 1655.

Beja-Pereira, A., Alexandrino, P., Bessa, I., et al. (2003). *J. Hered.*, 94:243-250.

De Roos, A. P. W., Hayes, B. J., Spelman, R. J., et al. (2008). *Genetics*, 179:1503-1512.

Excoffier, L., Laval, G., and Schneider S. (2005). *Evol. Bioinf.*, 1:47-50.

Felsenstein, J. (1989). *Cladistics*, 5:164-166.

Garrick, D. J. (2011). *Genet. Sel. Evol*., 43:17.

Gautier, M., Faraut, T., Moazami-Goudarzi, K., et al. (2007). *Genetics*, 177:1059-1070.

Martín-Burriel, I., Rodellar, C., Cañón, J., et al. (2011). *J. Anim. Sci.*, 89:893-906.

Nei, M. (1972). *Am. Nat.*, 106(949):283-292.

Paradis, E., Claude, J., and Strimmer, K. (2004). *Bioinformatics*, 20:289-290.

Purcell, S., Neale, B., Todd-Brown, K., et al. (2007). *Am. J. Hum. Genet*. 81:559-575.

Saitou, N., and Nei, M. (1987). *Mol. Biol. Evol*., 4:406-425.

Wright, S. (1951). *Ann. Eugen*., 15:323-353.