CSIRO PUBLISHING

Animal Production Science
http://dx.doi.org/10.1071/AN14605

SNP included in candidate genes involved in muscle, lipid and energy metabolism behave like neutral markers

Natalia Sevane[A], *Javier Cañon*[A], *John L. Williams*[B], *Hubert Levéziel*[C,D], *Alessio Valentini*[E], *Susana Dunner*[A,F] *and the GemQual Consortium*

[A]Dpto. de Producción Animal, Facultad de Veterinaria, Universidad Complutense, 28040 Madrid, Spain.
[B]Parco Tecnologico Padano, Via Einstein, Polo Universitario, 26900 Lodi, Italy.
[C]INRA, UMR 1061, F-87000 Limoges, France.
[D]Université de Limoges, UMR 1061, F-87000 Limoges, France.
[E]Department for Innovation in Biological, Agro-Food and Forest Systems, Università della Tuscia, via De Lellis, 01100 Viterbo, Italy.
[F]Corresponding author. Email: dunner@ucm.es

Abstract. Studies of population structure and diversity in cattle have provided insights into the origins of breeds, their history and evolution, and allow the identification of global livestock diversity hotspots, which is important for conservation of diversity. Genetic diversity, genetic relationship, population structure, and the presence of hotspots of genetic diversity among 15 European bovine breeds from five countries were assessed using 435 single nucleotide polymorphisms (SNP) markers identified in candidate genes for muscle, lipid and energy metabolism, thus providing the opportunity to compare the breed relationships obtained using putatively functional markers with previous data using neutral loci. Individuals belonging to 11 breeds tended to be clearly assigned to a single cluster when the number of pre-defined populations reached a maximum in the likelihood of the data at K = 12, whereas Asturiana de los Valles, Danish Red, Simmental, and Avileña-Negra Ibérica displayed a greater degree of admixture, which may be explained by their diverse ancestry. Although overall results were in agreement with those reported by studies based on neutral genetic variations, some additional breed relationship information emerged using markers in candidate functional loci, including the relationship between the Asturiana de los Valles and Piedmontese, and Danish Red and Charolais breeds. This study indicates that the analysed loci have not been main targets for selection or for adaptation processes, but also that SNP within candidate genes related with beef characteristics and performance may provide a slightly new perspective on past breeding and may also help in the development of strategies for the rational conservation of livestock diversity.

Additional keywords: adaptive variance, admixture, *Bos taurus*, genetic diversity, hotspots.

Received 21 November 2013, accepted 16 July 2014, published online 30 September 2014

Introduction

Cattle were introduced into northern Europe ~5000 years ago, and have been continuously selected for domestication traits over thousands of years. With the increasing demand for meat and milk in the last century, focussed selection has resulted in the creation of highly specialised breeds (Felius *et al*. 2011). Unimproved indigenous breeds are typically less productive and hence the widespread exploitation of a few highly selected breeds has caused a significant erosion of genetic diversity of domesticated cattle.

There is a growing awareness that local breeds may possess alleles associated with adaptation to local conditions and should be protected from extinction to preserve cattle diversity (Felius *et al*. 2011). In order to make rational decisions on conservation strategies, molecular characterisation of bovine genetic diversity is essential, and requires protocols that assess among-breed and within-breed genetic variation (Cañón *et al*. 2011). Most studies to date have focussed on neutral genetic variations, on the basis that these will not be biased by selection for functional loci or those controlling performance traits (Toro 2006). Adaptive variation, based on functional rather than neutral differences between populations, may provide new and complementary criteria to back up conservation decisions (Toro 2006), avoiding placing excessive dependence on genetic distances measurements based on neutral genetic variation.

Until recently, studies of population structure and diversity in cattle have used variations in mitochondrial DNA, Y-chromosomal variation, and simple sequence repeat markers (Wiener *et al*. 2004; Cortés *et al*. 2008, 2011; Martín-Burriel *et al*. 2011). After the sequencing of the bovine genome and with the availability of genotyping panels with many thousands of single nucleotide polymorphisms (SNP), this marker type has taken precedence for diversity studies and has been used to sample bovine populations from different parts of the world

Journal compilation © CSIRO 2014

(e.g. Decker *et al*. 2009; Gautier *et al*. 2010; Lin *et al*. 2010); however, these studies have invariably used random SNP, without differentiating between those linked to functional variation and those in neutral regions of chromosomes. In the present study, a set of 435 SNP markers was selected to be within functional candidate genes involved in muscle, lipid and energy metabolism (Williams *et al*. 2009; Sevane *et al*. 2011). This was used to examine the genetic diversity within, and the genetic relationships among, 15 European beef and dairy cattle from Spain, the United Kingdom, France, Denmark, and Italy, and to explore ancestry and the influence of selection on diversity at these loci for identifying the presence of diversity hotspots.

Material and methods

Animals

A total of 397 unrelated purebred bulls belonging to 12 European beef breeds including local and highly selected breeds, and three dairy breeds were used. The whole sample included: 28 Asturiana de los Valles, 26 Asturiana de la Montaña, 23 Avileña-Negra Ibérica, and 27 Pirenaica from Spain; 28 Jersey, 26 South Devon, 26 Aberdeen Angus, and 29 Highland from the United Kingdom; 31 Limousin, and 30 Charolais from France; 30 Piedmontese, and 27 Marchigiana from Italy; and 24 Holstein, 27 Danish Red Cattle, and 15 Simmental, from Denmark (Fig. 1).

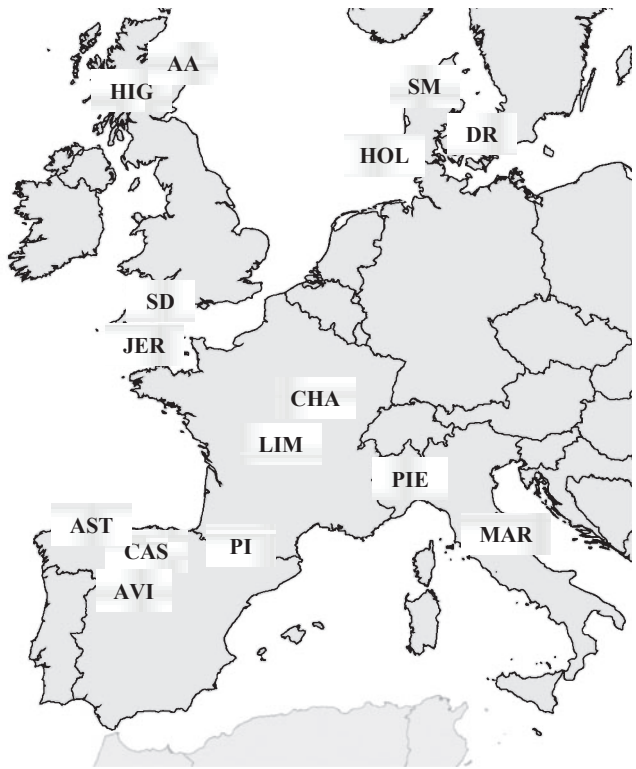

Fig. 1. Geographical distribution of the 15 bovine breeds studied. Complete breed names: Jersey (JER), South Devon (SD), Aberdeen Angus (AA), Highland (HIG), Holstein (HOL), Danish Red (DR), Simmental (SM), Limousin (LIM), Charolais (CHA), Piedmontese (PIE), Marchigiana (MAR), Asturiana de los Valles (AST), Asturiana de la Montaña (CAS), Avileña-Negra Ibérica (AVI), Pirenaica (PI).

Selection of SNP

The analysis of genetic diversity and relationships among breeds used the information provided by 435 SNP found in 192 candidate genes, some of which have been shown to be associated with muscle, lipid and energy metabolism. The whole dataset included: 389 SNP previously described by Williams *et al*. (2009) and genotyped by Kbioscience using the proprietary Kaspar methodology; 46 SNP previously described by Sevane *et al*. (2011) and genotyped by SNP multiplex and Primer Extension amplification. All SNP had a minor allele frequency above 5% in the breeds investigated. From an initial panel of 436 bulls, only those individuals with less than 20% of missing data were included in the analysis ($n = 397$) (Table 1).

Data analyses

Allele frequencies and observed and expected heterozygosities were obtained using the software GenePop version 4.0.7 (Raymond and Rousset 1995). Fisher's exact test for Hardy–Weinberg equilibrium (HWE) across loci and populations was performed using the Markov chain method, as implemented in GenePop version 4.0.7. Wright's index F_{IS} was calculated using GENETIX version 4.05 (Belkhir *et al.* 2004). Population pair-wise F_{ST} were calculated with ARLEQUIN version 3.11 (Excoffier *et al*. 2005).

Genetic relationships among populations was analysed using the standard genetic distance of Nei (1972), and a phenogram was constructed from the distance matrix using the unweighted pair group method with arithmetic mean (UPGMA) implemented in PHYLIP 3.69 (Felsenstein 1989). Bootstrap re-sampling ($n = 1000$) was performed to evaluate the robustness of the genetic clusters using the SEQBOOT, GENDIST, NEIGHBOUR, DRAWGRAM and CONSENSE programs in PHYLIP 3.69.

STRUCTURE 2.2 software (Pritchard *et al*. 2000) was used to infer population substructure in each cattle population with the admixture model and uncorrelated allele frequencies. This approach allowed subpopulations (K) with distinctive allele frequencies to be identified from the full dataset without using prior information on sampling groups. For each value of the putative number of K between 2 and 20, 10 independent runs were performed with a burn-in period of 50 000 followed by 100 000 Markov chain Monte Carlo (MCMC) repetitions. The methodology described by Evanno *et al*. (2005) was used to identify the optimal K value, and hence identify the most reliable result. DISTRUCT was used to graphically display results produced by the STRUCTURE program (Rosenberg 2004).

For each breed, the degree of admixture or ancestry diversity was calculated as $1 - \Sigma(q_k)^2$, where q_k is an average fraction of the genetic ancestry of the breed belonging to genetic clusters k, identified in STRUCTURE analysis (Tapio *et al*. 2010). The proportion of mixed ancestry was compared with the genetic diversity (expected heterozygosity) to analyse the relationship between the ancestry and the within-breed diversity.

The genetic structure was further analysed with a factorial correspondence analysis (FCA) to visualise the influence of SNP on discrimination patterns among breeds using the

Table 1. Descriptive statistics per breed

Complete breed names: Jersey (JER), South Devon (SD), Aberdeen Angus (AA), Highland (HIG), Holstein (HOL), Danish Red (DR), Simmental (SM), Limousin (LIM), Charolais (CHA), Piedmontese (PIE), Marchigiana (MAR), Asturiana de los Valles (AST), Asturiana de la Montaña (CAS), Avileña-Negra Ibérica (AVI), Pirenaica (PI)

Breed	Country	Sample size	% of polymorphic SNP	No. of SNP in HWd[A]	Availability[B]	H_e[C]	H_o[D]	F_{IS}[E]
JER	UK	28	84	4	0.93	0.25	0.26	–0.0423
SD	UK	26	86	3	0.90	0.29	0.30	–0.0271
AA	UK	26	86	2	0.91	0.30	0.30	–0.0274
HIG	UK	29	80	18	0.91	0.24	0.24	0.0338
HOL	DK	24	90	5	0.94	0.29	0.29	–0.0083
DR	DK	27	90	3	0.93	0.31	0.32	–0.0495
SM	DK	15	86	2	0.93	0.30	0.30	–0.0323
LIM	FR	31	95	5	0.97	0.30	0.31	–0.0381
CHA	FR	30	95	5	0.95	0.30	0.31	–0.0231
PIE	IT	30	98	5	0.97	0.30	0.30	–0.0117
MAR	IT	27	94	7	0.95	0.30	0.31	–0.0318
AST	ES	28	96	5	0.91	0.31	0.32	–0.0220
CAS	ES	26	89	8	0.93	0.28	0.29	–0.0336
AVI	ES	23	95	5	0.90	0.32	0.31	0.0171
PI	ES	27	94	2	0.92	0.29	0.29	0.0014

[A]Number of SNP in Hardy–Weinberg disequilibrium within the breed ($P < 0.05$).
[B]Availability = *1 – number of observations/n.*
[C]Expected heterozygosity.
[D]Observed heterorozygosity.
[E]Not significant.

GENETIX version 4.05 program (Belkhir *et al.* 2004). FCA is a multivariate method similar to principal component analysis, but which is appropriate for categorical variables and allows a simultaneous representation of breeds and alleles as a cloud of points in a metric space. Allele frequencies were used as the variables to cluster the breeds based on Chi-square distances.

Results

Genetic diversity

Heterozygosities (H_e) ranged from 0.24 in the Highland breed to 0.32 in the Avileña-Negra Ibérica breed. The breeds having H_e values above 0.3 (Avileña-Negra Ibérica, Asturiana de los Valles and Danish Red) were considered as 'diversity hotspots'. Values for the observed heterozygosity (H_o) and expected H_e and the measure of the possible discrepancy between this value and the observed heterozygosity (F_{IS}) per population are shown in Table 1. The percentage of polymorphic SNP per breed ranged from 80 in Highland to 98 in Piedmontese (Table 1). About 52% of the 435 SNP were found to be polymorphic in all 15 breeds analysed and ~72% were polymorphic at least in 14 breeds. Only 10 markers were polymorphic in fewer than eight breeds. No significant differences were detected between expected and observed heterozygosity, showing that populations were in HE equilibrium. Although the results of the Fisher's exact test show no significant deviations from HWE for all populations, deviations from HWE across loci were significant in all populations ($P < 0.005$). The number of SNP in HW disequilibrium within breeds ranged from 2 in Aberdeen Angus, Simmental and Pirenaica to 18 in Highland (Table 1). Twenty-four SNP were not in HWE in one population while 8 SNP were not in HWE for more than two breeds, 5 of which had previously been associated with a production trait in cattle.

Genetic distances and clustering

Values for pair-wise comparisons of Nei's standard genetic distance (Nei 1972) between the 15 bovine breeds are shown in Table 1. Highland and Jersey were the most distant breeds on average (0.12 and 0.11, respectively), whereas Asturiana de los Valles and Charolais had the lowest average genetic distance (0.06). The global average value of F_{ST} was 0.12, which indicates that 88% of the total variability can be ascribed to the within-breed variation.

UPGMA clustering of the standard genetic distance (Nei 1972) separates the United Kingdom breeds, the Holstein dairy breed and the Italian breed Marchigiana into distinct groups, while the remaining breeds can be divided into two main clusters: one containing the Avileña-Negra Ibérica and Asturiana de la Montaña; and the second with the Asturiana de los Valles, Piedmontese, Pirenaica, Limousin, Simmental, Charolais and Danish Red (Fig. 2*a*). The phylogeny based on the information obtained through the Bayesian model-based procedure and assuming different K values is shown in Fig. 2*b*. The groups identified by UPGMA were also revealed by the STRUCTURE program when $K = 7$ except for the Marchigiana breed (Fig. 2*b*, Table 2). The four United Kingdom breeds are clearly assigned to a distinct single cluster, indicating a high degree of genetic differentiation. Holstein and Danish Red breeds split in two distinct clusters but, possibly as a result of introgression of Holstein into the Danish Red, the latter shared one of its main clusters with

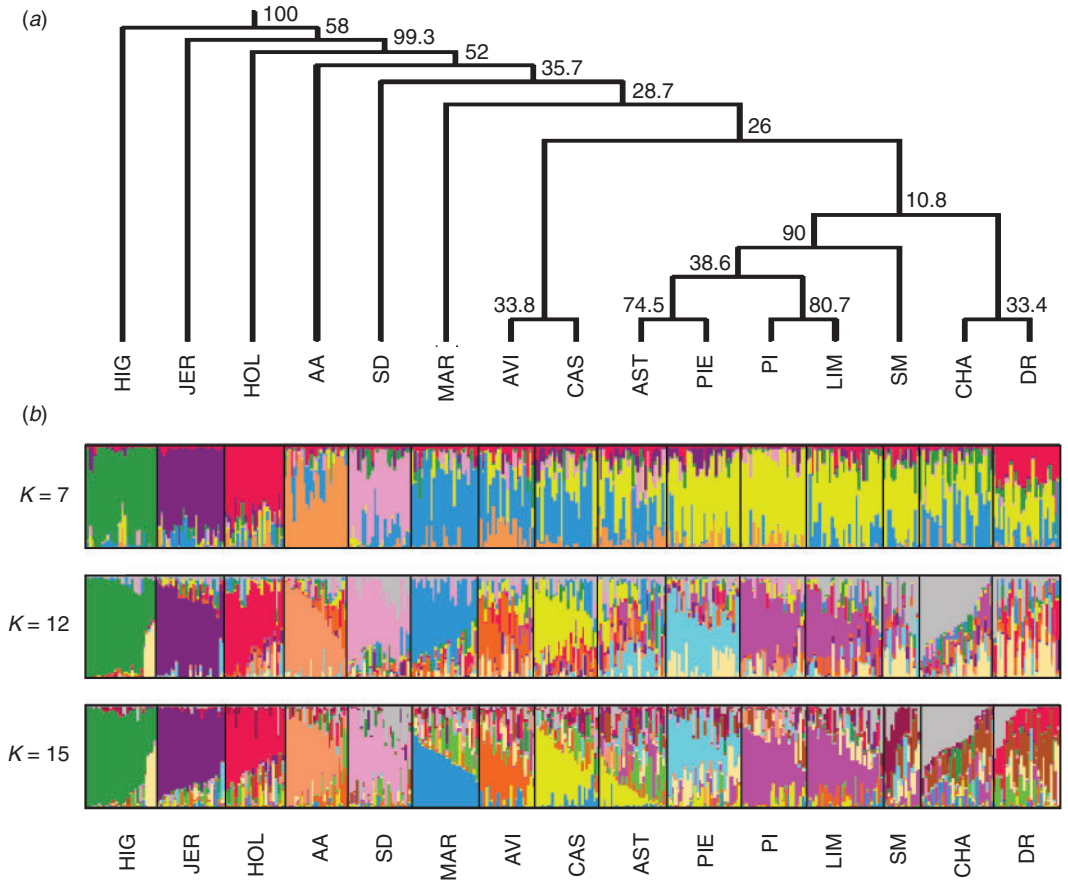

Fig. 2. Genetic distances and clustering of 15 European bovine breeds: (*a*) unweighted pair group method with arithmetic mean tree using Nei's (1972) original distance constructed using the SEQBOOT, GENDIST, NEIGHBOUR, and CONSENSE programs in the PHYLIP version 3.69 software package (bootstrap support values in percentage – 1000 replicates – are indicated at the nodes); (*b*) estimated membership fractions of each bovine breed assuming K = 7, 12 and 15 inferred using STRUCTURE 2.2 (each individual is represented by a single vertical line divided into K colours, where K is the number of clusters assumed and the length of the coloured segment represents the individual's estimated proportion of membership to a particular cluster). Complete breed names: Jersey (JER), South Devon (SD), Aberdeen Angus (AA), Highland (HIG), Holstein (HOL), Danish Red (DR), Simmental (SM), Limousin (LIM), Charolais (CHA), Piedmontese (PIE), Marchigiana (MAR), Asturiana de los Valles (AST), Asturiana de la Montaña (CAS), Avileña-Negra Ibérica (AVI), Pirenaica (PI).

Holstein. The remaining breeds grouped in two clusters which were the same as obtained using the UPGMA tree.

The STRUCTURE analysis reached a maximum likelihood for the data with K = 12. The proportional contribution of the inferred ancestral clusters per breed when K = 12 is shown in Table 3. In 10 breeds the proportional contribution of the inferred clusters was greater than 50%, whereas Danish Red, Simmental, Asturiana de los Valles, Avileña-Negra Ibérica, and, to a lesser extent, Charolais, displayed a greater degree of admixture. Apart from clusters including Holstein and Danish Red (Fig. 2*b*, red when K = 12, 15), and Asturiana de los Valles and Asturiana de la Montaña breeds (Fig. 2*b*, yellow when K = 12, 15), the clusters obtained were generally consistent with the UPGMA tree.

A high significant positive correlation ($r = 0.86$, $P < 0.0001$) was obtained between the expected heterozygosity and the level of admixture based on global STRUCTURE results (K = 12) for the 15 bovine breeds studied, suggesting that admixture may explain the presence of diversity hotspots.

The genetic structure was further analysed with FCA clustering methods. The first two axes contributed 16.2% and 13.9% of the total inertia (Fig. 3). Highland and Jersey were the only breeds clearly separated from the main group, representing 48% of total inertia in axis 1 and 66% in axis 2, respectively. The most important alleles were: the mutated A allele of *GDF8*_F94L SNP previously associated with increased muscularity (Esmailizadeh *et al.* 2008), the T allele of *VCL*_a1_160, the A allele of *MYLK2*_b1_203, the T allele of *CALM3*_a1_149, the A allele of *ATP1B2*_a1_307, the C allele of *SPARC*_b1_268, and the G allele of *RORA*.

Discussion

Genetic diversity

Genotype data from 435 SNP markers in 397 animals belonging to European beef and dairy bovine breeds were used to assess the genetic structure and the genetic relationships among breeds. The levels of within-breed diversity in the 15 cattle breeds

Table 2. Genetic distance of Nei (1972) between 15 bovine breeds

Complete breed names: Jersey (JER), South Devon (SD), Aberdeen Angus (AA), Highland (HIG), Holstein (HOL), Danish Red (DR), Simmental (SM), Limousin (LIM), Charolais (CHA), Piedmontese (PIE), Marchigiana (MAR), Asturiana de los Valles (AST), Asturiana de la Montaña (CAS), Avileña-Negra Ibérica (AVI), Pirenaica (PI)

	JER	SD	AA	HIG	HOL	DR	SM	LIM	CHA	PIE	MAR	AST	CAS	AVI	PI
JER	–	0.12	0.14	0.16	0.11	0.12	0.12	0.10	0.10	0.09	0.12	0.08	0.10	0.12	0.11
SD	–	–	0.09	0.11	0.11	0.08	0.10	0.08	0.06	0.07	0.08	0.06	0.07	0.08	0.08
AA	–	–	–	0.11	0.09	0.08	0.10	0.09	0.07	0.08	0.09	0.06	0.08	0.07	0.09
HIG	–	–	–	–	0.12	0.10	0.12	0.11	0.10	0.12	0.11	0.11	0.11	0.12	0.14
HOL	–	–	–	–	–	0.06	0.10	0.09	0.09	0.09	0.08	0.07	0.08	0.07	0.12
DR	–	–	–	–	–	–	0.07	0.06	0.05	0.06	0.07	0.05	0.07	0.06	0.08
SM	–	–	–	–	–	–	–	0.05	0.05	0.05	0.09	0.06	0.09	0.08	0.06
LIM	–	–	–	–	–	–	–	–	0.04	0.05	0.07	0.04	0.06	0.06	0.04
CHA	–	–	–	–	–	–	–	–	–	0.05	0.06	0.04	0.06	0.06	0.06
PIE	–	–	–	–	–	–	–	–	–	–	0.07	0.04	0.06	0.06	0.05
MAR	–	–	–	–	–	–	–	–	–	–	–	0.05	0.07	0.06	0.10
AST	–	–	–	–	–	–	–	–	–	–	–	–	0.04	0.05	0.05
CAS	–	–	–	–	–	–	–	–	–	–	–	–	–	0.06	0.08
AVI	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.08
Average	0.11	0.08	0.09	0.12	0.09	0.07	0.08	0.07	0.06	0.07	0.08	0.06	0.07	0.07	0.08

Table 3. Proportional contribution of the inferred ancestral clusters (K = 12) in each breed studied

Complete breed names: Jersey (JER), South Devon (SD), Aberdeen Angus (AA), Highland (HIG), Holstein (HOL), Danish Red (DR), Simmental (SM), Limousin (LIM), Charolais (CHA), Piedmontese (PIE), Marchigiana (MAR), Asturiana de los Valles (AST), Asturiana de la Montaña (CAS), Avileña-Negra Ibérica (AVI), Pirenaica (PI). Bold denotes main ancestral clusters in each breed

	Clusters												Ancestry diversity	
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	(0.917)[A]	H_e
JER	0.015	0.043	0.009	**0.762**	0.024	0.013	0.020	0.030	0.029	0.023	0.015	0.017	0.414	0.25
SD	0.029	0.025	0.010	0.022	0.026	0.013	0.007	0.018	0.043	0.052	**0.627**	0.128	0.583	0.29
AA	0.023	0.026	**0.648**	0.008	0.087	0.015	0.043	0.038	0.025	0.016	0.029	0.043	0.564	0.30
HIG	0.084	0.005	0.004	0.004	0.008	0.003	0.009	0.005	**0.835**	0.013	0.014	0.017	0.295	0.24
HOL	0.035	0.038	0.049	0.030	0.035	0.030	**0.649**	0.031	0.024	0.038	0.011	0.029	0.567	0.29
DR	**0.268**	0.063	0.011	0.008	0.095	0.040	**0.211**	0.031	0.023	0.085	0.008	0.158	0.835	0.31
SM	0.196	0.170	0.008	0.024	0.039	**0.242**	0.023	0.016	0.033	0.008	0.005	**0.235**	0.814	0.30
LIM	0.062	0.043	0.007	0.041	0.062	**0.500**	0.050	0.039	0.030	0.033	0.006	0.127	0.716	0.30
CHA	0.086	0.051	0.032	0.031	0.037	0.151	0.018	0.038	0.046	0.037	0.024	**0.450**	0.756	0.30
PIE	0.130	**0.533**	0.024	0.032	0.028	0.082	0.020	0.052	0.013	0.023	0.018	0.044	0.684	0.30
MAR	0.034	0.032	0.022	0.015	0.056	0.025	0.038	0.033	0.028	**0.627**	0.053	0.036	0.593	0.30
AST	0.039	**0.222**	0.074	0.046	0.064	0.104	0.071	**0.167**	0.037	0.087	0.037	0.053	0.881	0.31
CAS	0.028	0.038	0.008	0.039	0.072	0.030	0.071	**0.509**	0.040	0.081	0.025	0.061	0.714	0.28
AVI	0.088	0.025	0.064	0.031	**0.382**	0.049	0.063	**0.119**	0.029	0.058	0.049	0.042	0.812	0.32
PI	0.063	0.053	0.060	0.018	0.023	**0.602**	0.012	0.051	0.008	0.014	0.064	0.032	0.619	0.29

[A]Maximum value for the ancestry diversity $[1 - \Sigma(q_k)^2]$, where q_k is an average fraction of the genetic ancestry of the breed belonging to genetic clusters k, identified in STRUCTURE analysis.

($H_e = 0.29$; $H_o = 0.30$) were similar to those found in other diversity studies of cattle breeds using SNP loci (e.g. Gautier *et al*. 2010; Lin *et al*. 2010). In contrast to these earlier studies, the SNP analysed here were within candidate genes involved in muscle, lipid and energy metabolism that may affect production traits, including growth, carcass, physical variables, lipid profiles and organoleptic properties of meat, and thus have higher probabilities of having been selected in past generations. However, despite these premises, only 18 of the SNP used here have been shown to have a significant effect, ranging from 1.5% to 19%, on meat quality-related traits in the breeds studied (Dunner *et al*. 2013*a*, 2013*b*; Sevane *et al*. 2013). Genetic selection is invariably focussed on easily measurable production traits such as daily gain or final weight, whereas meat quality traits, which are more difficult and expensive to measure, are usually not included in selection programs (Simm *et al*. 2009). Thus, the lack of selection pressure on the allele frequencies at the loci used may explain the lack of significant deviations from HWE in the 15 breeds.

Genetic distances and clustering

Many factors complicate the classification of cattle, including the scarce documentation of the history and diversity of cattle

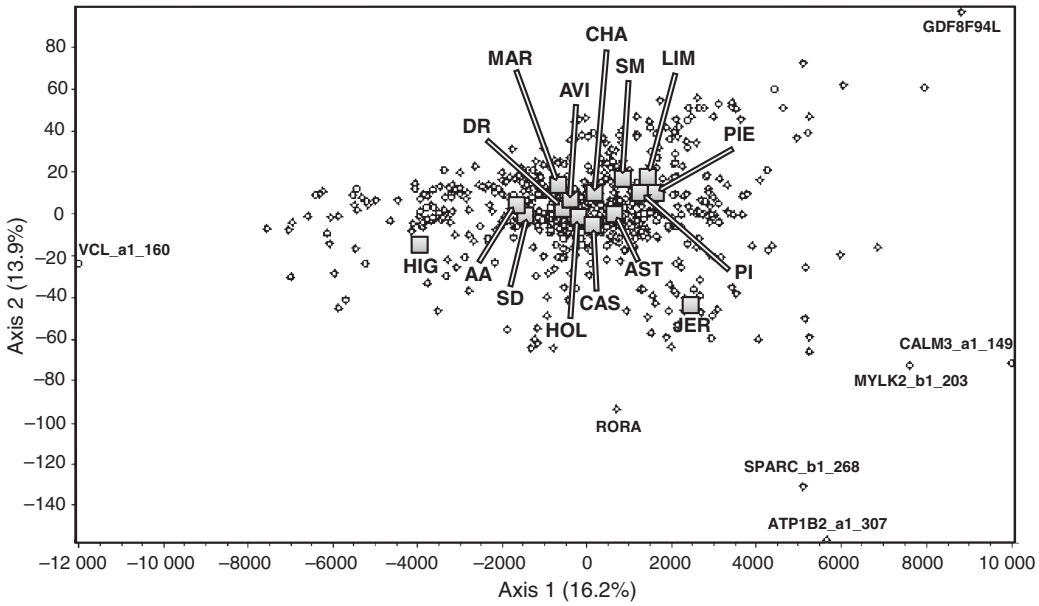

Fig. 3. Spatial representation of the factorial correspondence analysis (FCA) of 15 bovine breeds (squares) and 435 SNP markers (points) obtained with FCA 3D by population (GENETIX version 4.05).

populations before breed formation, the influence of migration and the unknown genetic roots of the current breeds. Since the formation of breeds in Europe, selection has produced phenotypic differences between breeds, which may, or may not, have originated relatively recently from a common gene pool. The global average F_{ST} found in several studies is in the range 0.07–0.19 – F_{ST} 0.07 (Cañón *et al.* 2011) for Iberian and French breeds using microsatellites; F_{ST} 0.09 (Liron *et al.* 2006) for Creole breeds using microsatellites; F_{ST} 0.10 (McKay *et al.* 2008) for Angus, Charolais, Limousin, Dutch Black and White Dairy, and Holstein breeds using SNP; and F_{ST} 0.19 (Gautier *et al.* 2010) for different *Bos taurus* and *B. indicus* breeds using SNP – similar to the F_{ST} of 0.12 reported here.

The persistent signatures of admixture present in many breeds confuses the apparent and observed genetic relationships of these breeds (Decker *et al.* 2009). The different procedures used by the UPGMA clustering, based on genetic distance matrices which lose information by collapsing all genotype data for pairs of breeds into a single number, and the STRUCTURE clustering model-based method, which allows gene flow between breeds to be estimated, may explain the discrepancies obtained between both methods for Holstein and Danish Red, and Asturiana de los Valles and Asturiana de la Montaña breeds.

The main breed clusters obtained in this study were in accordance with previous studies, which showed that Angus, Holstein and South Devon breeds share close positions in cladograms (Decker *et al.* 2009). In addition, the isolation of the United Kingdom breeds seen here has also been reported by others (Wiener *et al.* 2004).

Among the Spanish breeds, Pirenaica clusters close to French breeds (Fig. 2, Table 2), as seen by Felius *et al.* (2011); this may be explained by genetic migration promoted by geographical proximity. Asturiana de la Montaña and Avileña-Negra Ibérica cluster together in the Iberian group (Fig. 2*a*). Asturiana de los Valles shows an intermediate position between Iberian and Central Brown groups (Table 2), which is in accordance with the genetic proportion shared with the Brown Swiss breed (Q = 7.5%) as reported by Martín-Burriel *et al.* (2011), and with previous studies that described the influence of neighbouring populations (Asturiana de la Montaña, Pasiega and Rubia Gallega) on this breed (Cañón *et al.* 2011; Martín-Burriel *et al.* 2011), as reflected by a common ancestral origin shared between Asturiana de los Valles and Asturiana de la Montaña when the number of assumed populations (K) is set to 7, 12 or 15 (Fig. 2*b*, Table 2).

The French Limousin and Charolais breeds have been shown to be closely related in some studies (Decker *et al.* 2009), whereas Felius *et al.* (2011) placed these breeds in two different categories, the French-Pyrenean and the North-West Intermediate, respectively, which is in agreement with the results obtained here (Fig. 2*a*).

Danish Red and Holstein breeds are not closely related historically (Felius *et al.* 2011), but here these breeds cluster together when the number of populations (K) was set to 7, 12 and 15 (Fig. 2*b*, Table 2). The likely explanation of this is the recent admixture between these two breeds, in particular the introgression of Holstein genetics into the Danish Red population.

The highest ancestral diversity was found in Asturiana de los Valles, Danish Red, Simmental, and Avileña-Negra Ibérica breeds, all of which had ancestry diversity values above 0.8. Low values were seen for Aberdeen Angus, Holstein, South Devon, and Marchigiana (ancestry diversity 0.5–0.6); however, the Highland and Jersey breeds had particularly low values (0.3 and 0.4, respectively). This trend is also observed from the genetic distance matrices. The high significant positive

correlation between H_e and ancestry diversity indicates that admixture may explain the higher diversity in some of the breeds.

Although breed relationships obtained here using SNP, located in candidate genes involved in muscle characteristics and beef production traits, were generally in agreement with those reported by studies based on neutral variations, some new relationships among breeds were identified including those between Asturiana de los Valles and Piedmontese, and between Danish Red and Charolais breeds. Selection for muscular hypertrophy and the genetic linkage (hitchhiking) around the myostatin gene (*GDF8*) may explain the relationship between the Asturiana and Piedmontese as both breeds are double muscled, although as a result of different mutations in *GDF8*. The apparent genetic similarity between the Danish Red and Charolais is, however, difficult to explain.

Finally, the FCA plot of multilocus genotypes and breeds showed limited differentiation among breeds, except for Highland and Jersey populations. Concerning the particular influence of SNP on discrimination patterns among breeds, it is worth highlighting the possible influence of *GDF8*_F94L on Limousin, *VCL*_a1_160 on Highland, and *MYLK2*_b1_203, *CALM3*_a1_149, *ATP1B2*_a1_307, *SPARC*_b1_268, and *RORA* on the Jersey breed.

In conclusion, the analysis data presented here for 435 SNP within candidate genes genotyped in 15 European breeds broadly confirmed the among-breed relationships obtained using neutral markers. Further analysis using a larger number of markers, not just located on candidate genes, but found to be associated to different production traits, may reveal different signatures of human selection pressure according to breed purpose. The study revealed that there is greater admixture found between continental breeds compared with United Kingdom breeds. This is consistent with the influence of migration of human populations and livestock over many generations, mainly through the Mediterranean route (Cortés *et al*. 2008), and more recently crossbreeding between European cattle (Martín-Burriel *et al*. 2011).

Acknowledgements

This work was supported in part by funding from the European Union grant QLK5 – CT2000–0147 for the 'GemQual' project.

GeMQual Consortium

Albertí P.[1], Amarger V.[2], Boitard S.[3], Cañón J.[4], Checa M.L.[4], Christensen M.[5], Crisá A.[6], Delourme D.[7], Dunner S.[4], Ertbjerg P.[5], Failla S.[8], García D.[4], Gigli S.[8], Hocquette J.F.[9], Levéziel H.[2], Mangin B.[3], Marchitelli C.[6], Miranda M.E.[4], Nute G.[10], Richardson I.[10], Olleta J.L.[11], Panea B.[11], Pérez R.[4], Razzaq N.[12], Renand G.[13], Sañudo C.[11], Valentini A.[6], Williams. J.L.[14]

[1]Centro de Investigación y Tecnología Agroalimentaria, Gobierno de Aragón, 50080, Zaragoza, Spain; [2]INRA, UMR 1061, 87000 Limoges, France and Université de Limoges, UMR 1061, 87000 Limoges; [3]INRA Chemin de Borde-Rouge-Auzeville, BP 52627, 31326 Castanet-Tolosan cedex, France; [4]Departamento de Producción Animal, Facultad de Veterinaria, 28040 Madrid, Spain; [5]Department of Food Science, University of Copenhagen, 1958 Frederiksberg C., Denmark; [6]Dipartimento di Produzioni Animali, Università della Tuscia, via De Lellis, 01100 Viterbo, Italy; [7]Université de Limoges, UMR 1061, 87000 Limoges, France; [8]CRA, Istituto Sperimentale per la Zootecnia, 00016 Monterotondo, Italy; [9]INRA, UR1213, Unité de Recherches sur les Herbivores, Centre de Clermont-Ferrand./Theix 63122, France; [10]Division of Farm Animal Science, University of Bristol, BS40 5DU, United Kingdom; [11]Dept de Producción Animal y Ciencia de los Alimentos, Universidad de Zaragoza, 50013, Zaragoza, Spain; [12]Roslin Institute, Roslin, Midlothian, Scotland.EH25 9PS, UK; [13]INRA, UR337, Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas cedex, France; [14]Parco Tecnologico Padano, Via Einstein, Polo Universitario, 26900 Lodi, Italy.

References

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) 'GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.' (Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II: Montpellier, France)

Cañón J, García D, Delgado JV, Dunner S, Gama LT, Landi V, Martín-Burriel I, Martínez A, Penedo C, Rodellar C, Zaragoza P, Ginja C (2011) Relative breed contributions to neutral genetic diversity of a comprehensive representation of Iberian native cattle. *Animal* **5**, 1323–1334. doi:10.1017/S1751731111000267

Cortés O, Tupac-Yupanqui I, Dunner S, García-Atance MA, García D, Fernández J, Cañón J (2008) Ancestral matrilineages and mitochondrial DNA diversity of the Lidia cattle breed. *Animal Genetics* **39**, 649–654. doi:10.1111/j.1365-2052.2008.01782.x

Cortés O, Tupac-Yupanqui I, Dunner S, Fernández J, Cañón J (2011) Y chromosome genetic diversity in the Lidia bovine breed: a highly fragmented population. *Journal of Animal Breeding and Genetics* **128**, 491–496. doi:10.1111/j.1439-0388.2011.00951.x

Decker JE, Pires JC, Conant GC, McKay SD, Heaton MP, Chen K, Cooper A, Vilkki J, Seabury CM, Caetano AR, Johnson GS, Brenneman RA, Hanotte O, Eggert LS, Wiener P, Kim JJ, Kim KS, Sonstegard TS, Van Tassell CP, Neibergs HL, McEwan JC, Brauning R, Coutinho LL, Babar ME, Wilson GA, McClure MC, Rolf MM, Kim J, Schnabel RD, Taylor JF (2009) Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18 644–18 649. doi:10.1073/pnas.0904691106

Dunner S, Sevane N, García D, Valentini A, Williams JL, Mangin B, Cañón J, Levéziel H GeMQual Consortium (2013*a*) Association of genes involved in carcass and meat quality traits in fifteen European bovine breeds. *Livestock Science* **154**, 34–44. doi:10.1016/j.livsci.2013.02.020

Dunner S, Sevane N, García D, Levéziel H, Williams JL, Mangin B, Valentini A GeMQual Consortium (2013*b*) Genes involved in muscle lipid composition in 15 European Bos taurus breeds. *Animal Genetics* **44**, 493–501. doi:10.1111/age.12044

Esmailizadeh AK, Bottema CD, Sellick GS, Verbyla AP, Morris CA, Cullen NG, Pitchford WS (2008) Effects of the myostatin F94L substitution on beef traits. *Journal of Animal Science* **86**, 1038–1046. doi:10.2527/jas.2007-0589

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47–50.

Felius M, Koolmees PA, Theunissen B, European Cattle Genetic Diversity Consortium Lenstra JA (2011) On the breeds of cattle-historic and current classifications. *Diversity* **3**, 660–692. doi:10.3390/d3040660

Felsenstein J (1989) phylip – phylogeny inference package (version 3.2). *Cladistics* **5**, 164–166.

Gautier M, Laloë D, Moazami-Goudarzi K (2010) Insights into the genetic history of French cattle from dense SNP data on 47 worldwide breeds. *PLoS ONE* **5**, e13038. doi:10.1371/journal.pone.0013038

Lin BZ, Sasazaki S, Mannen H (2010) Genetic diversity and structure in *Bos taurus* and *Bos indicus* populations analyzed by SNP markers. *Animal Science Journal* **81**, 281–289. doi:10.1111/j.1740-0929.2010.00744.x

Liron JP, Peral-Garcia P, Giovambattista G (2006) Genetic characterization of Argentine and Bolivian Creole cattle breeds assessed through microsatellites. *The Journal of Heredity* **97**, 331–339. doi:10.1093/jhered/esl003

Martín-Burriel I, Rodellar C, Cañón J, Cortés O, Dunner S, Landi V, Martínez-Martínez A, Gama LT, Ginja C, Penedo MC, Sanz A, Zaragoza P, Delgado JV (2011) Genetic diversity, structure, and breed relationships in Iberian cattle. *Journal of Animal Science* **89**, 893–906. doi:10.2527/jas.2010-3338

McKay SD, Schnabel RD, Murdoch BM, Matukumalli LK, Aerts J, Coppieters W, Crews D, Dias Neto E, Gill CA, Gao C, Mannen H, Wang Z, Van Tassell CP, Williams JL, Taylor JF, Moore SS (2008) An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. *BMC Genetics* **20**, 9–37.

Nei M (1972) Genetic distance between populations. *American Naturalist* **106**, 283–292. doi:10.1086/282771

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure from multilocus genotype data. *Genetics* **155**, 945–959.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *The Journal of Heredity* **86**, 248–249.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137–138. doi:10.1046/j.1471-8286.2003.00566.x

Sevane N, Crespo I, Cañón J, Dunner S (2011) A Primer-Extension Assay for simultaneous use in cattle Genotype Assisted Selection, parentage and traceability analysis. *Livestock Science* **137**, 141–150. doi:10.1016/j.livsci.2010.10.011

Sevane N, Armstrong E, Wiener P, Pong Wong R, Dunner S GeMQual Consortium (2013) Association of bovine meat quality traits with genes included in the PPARG and PPARGC1A networks. *Meat Science* **94**, 328–335. doi:10.1016/j.meatsci.2013.02.014

Simm G, Lambe N, Bünger L, Navajas E, Roehe R (2009) Use of meat quality information in breeding programmes. In 'Improving the sensory and nutritional quality of fresh meat'. (Eds JP Kerry, D Ledward) pp. 264–291. (Woodhead Publishing Ltd: Cambridge, UK)

Tapio M, Ozerov M, Tapio I, Toro MA, Marzanov N, Cinkulov M, Goncharenko G, Kiselyova T, Murawski M, Kantanen J (2010) Microsatellite-based genetic diversity and population structure of domestic sheep in northern Eurasia. *BMC Genetics* **10**, 11–76.

Toro MA (2006) Assessing genetic diversity between breeds for conservation. *Journal of Animal Breeding and Genetics* **123**, 289. doi:10.1111/j.1439-0388.2006.00610.x

Wiener P, Burton D, Williams JL (2004) Breed relationships and definition in British cattle: a genetic analysis. *Heredity* **93**, 597–602. doi:10.1038/sj.hdy.6800566

Williams JL, Dunner S, Valentini A, Mazza R, Amarger V, Checa ML, Crisa A, Razzaq N, Delourme D, Grandjean F, Marchitelli C, García D, Pérez Gómez R, Negrini R, Ajmone Marsan P, Levéziel H (2009) Discovery, characterization and validation of single nucleotide polymorphisms within 206 bovine genes that may be considered as candidate genes for beef production and quality. *Animal Genetics* **40**, 486–491. doi:10.1111/j.1365-2052.2009.01874.x