doi:10.1111/j.1365-2052.2006.01441.x

Characterization of single nucleotide polymorphisms in sheep and their variation as evidence of selection

L. Pariset*, I. Cappuccio*, S. Joost†, M. D'Andrea‡, D. Marletta§, P. Ajmone Marsan¶, A. Valentini* and The ECONOGENE Consortium**

*Dipartimento di Produzioni Animali, Università della Tuscia, Viterbo, Italy. †Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. ‡Dipartimento di Scienze Animali Vegetali e dell'Ambiente, Università degli Studi del Molise, Campobasso, Italy. §Dipartimento di Scienze Agronomiche, Agrochimiche e delle Produzioni Animali, Università di Catania, Catania, Italy. ¶Istituto di Zootecnica, Università Cattolica del Sacro Cuore, Piacenza, Italy. **http://lasig.epfl.ch/projets/econogene/

Summary

The discovery of SNPs was performed using animals from eight European sheep breeds. Eleven SNPs were further characterized using about 1700 sheep belonging to 57 breeds. A method for the identification of loci that were likely subject to selection was applied; three of the 11 SNPs lying outside the 95% confidence region of the conditional joint distribution of F_{ST} and mean heterozygosity were identified as outliers.

Keywords outliers, selection, sheep, SNPs.

Signatures of natural selection can be detected by quantifying the variation of allele frequencies between populations using the F_{ST} statistic (Lewontin & Krakauer 1973; Beaumont & Nichols 1996). Under selective neutrality, F_{ST} is influenced by genetic drift, which will affect all loci across the genome in a similar fashion. Conversely, selection is a locus-specific force that will change the F_{ST} for a selected gene and its closely linked genetic markers.

Genes involved in key metabolic pathways influencing production, disease resistance and morphological traits have been selected for SNP discovery under an EU ECONOGENE project (http://lasig.epfl.ch/projets/econogene/). These genes may have high adaptive value and may therefore be under natural or artificial selection. In this study, we identified loci behaving differently from neutral models; therefore, we suggest that these loci are under natural or artificial selection.

Genomic DNA was extracted from blood samples of 1748 sheep, of which about one-third were males, belonging to 57 breeds, of which 52 were autochthonous (Table 1). Whenever possible, three samples per farm and 11 farms per breed were collected. A panel of 16 unrelated sheep belonging to eight European breeds was used for SNP screening and characterization. Exon and intron sequences were screened for polymorphisms by direct sequencing and DHPLC analyses. Eleven of 16 genes had at least one SNP. PCR primers were designed from genomic sequences of *Ovis aries* or from the consensus sequence of a related species available in GenBank (Table S1).

Heterozygosity and F_{ST} were calculated using the FDIST software (http://www.rubic.rdg.ac.uk/~mab/software.html). Population datasets were built by bootstrapping 200 000 replications on real data using a coalescent model. Upper and lower 95% confidence intervals were assumed for conditional joint distribution of F_{ST} vs. mean heterozygosity. Differences in allele frequencies were assessed using a student *t*-test.

The characterization of sequences from animals of eight breeds revealed 12 SNPs in 6686 bp sequenced, giving an average density of one SNP every 557 bases. Genotyping was performed on 11 SNPs, including six transitions (four G>A and two T>G) and five T>G transversions. Allele frequencies are reported in Table 2, with frequencies <0.5% considered monomorphic.

Heterozygosity of the loci determined from SNP frequencies ranged from 0.063 to 0.497, with a mean of 0.28 (Table 2). The F_{ST} among sheep breeds determined from SNP frequencies (Table 2) indicates a low level of differentiation, although the mean heterozygosity of the loci was 0.28, a value not particularly low for this type of marker (Brouillette & Venta 2002; Matsuzaki *et al.* 2004; Morin *et al.* 2004). Our sample included a large number of populations spanning a considerable geographical area. A possible explanation for the low level of differentiation observed is that the ease of transportation, even in ancient times, shuffled the genetic pool and did not permit stratification through genetic drift.

Address for correspondence

Lorraine Pariset, Dipartimento di Produzioni Animali, Facoltà di Agraria, Università della Tuscia, via S. Camillo de Lellis, 01100 Viterbo, Italy.
E-mail: pariset@unitus.it

Accepted for publication 16 February 2006

Table 1 Breeds, countries of origin and numbers of genotyped sheep.

Breed	Origin	No. animals
Akkaraman*	Turkey	17
Altamurana	Italy	31
Anogeiano	Greece	31
Bardhoka	Albania	31
Bergamasca*	Italy	31
Churra	Portugal	30
Cikta	Hungary	31
Colmenareña	Spain	31
Cypriot fat-tailed sheep	Cyprus	32
Dagliç	Turkey	31
Delle Langhe	Italy	31
Exmoor Horn	Great Britain	31
Gentile di Puglia	Italy	31
German Grey Heath	Germany	31
German Merino	Germany	31
Heri	Saudi-Arabia	29
Hungarian Merino	Hungary	31
Hungarian Tsiagia	Hungary	29
Kalarritiko	Greece	31
Kamieniec	Poland	31
Karagouniko*	Greece	31
Karakul	Romania	31
Karayaka	Turkey	31
Kefallinias	Greece	31
Kymi	Greece	31
Laticauda	Italy	31
Lesvos	Greece	31
Manchega	Spain	31
Morkaraman	Turkey	31
Naemi	Saudi-Arabia	29
Najdi	Saudi-Arabia	31
Orino	Greece	31
Ossimi	Egypt	31
Piliorritiko	Greece	31
Polish Merino	Poland	31
Polish Mountain (Owza gorska)	Poland	31
Pomorska	Poland	31
Racka	Hungary	31
Rhönsheep*	Germany	31
Romanian Merino	Romania	31
Romanian Tsigaia	Romania	31
Rubia del Molar*	Spain	31
Ruda	Albania	31
Scottish Blackface	Great Britain	31
Segureña	Spain	31
Sfakia	Greece	31
Shkodrane	Albania	31
Skopelos	Greece	31
Spanish Merino I	Spain	31
Spanish Merino II	Spain	32
Swaledale	Great Britain	31
Thône et Martod	France	31
Turcana*	Romania	31
Welsh Mountain*	Great Britain	31
White and Brown Mountain Sheep	Germany	31
Wrzosowka	Poland	31
Zelazna*	Poland	31

*Breeds included in the SNPs detection panels.

Table 2 Locus, gene name, allele frequencies, heterozygosity and F_{ST} for each genotyped SNP.

Locus	Gene name	Allele 1	Allele 2	Homozygotes allele 1 (%)	Heterozygotes (%)	Homozygotes allele 2 (%)	Sample heterozygosity	Sample F_{ST}	Test statistic	P simulated F_{ST} < sample F_{ST}
SERPINA3	*Alpha-1 antiproteinase, antitrypsin*	A	G	1.1	16.5	82.4	0.170	0.039	−1.407	0.094
MEG3	*Maternally expressed 3*	T	G	13.3	42.3	44.4	0.451	0.030	−2.555	0.007
CTSB	*Cathepsin B*	A	G	1.3	14.5	84.1	0.169	0.022	−2.647	0.005
GHRHR	*Growth hormone releasing hormone receptor*	G	T	5.5	31.8	62.7	0.338	0.054	−0.736	0.256
GHR	*Growth hormone receptor*	G	A	30.2	46.0	23.8	0.497	0.085	0.864	0.218
IGF1	*Insulin like growth factor 1*	T	C	38.8	44.1	17.1	0.466	0.046	−1.303	0.113
ITGB1	*Integrin beta 1*	T	C	1.4	15.1	83.5	0.166	0.068	0.115	0.463
GDF8	*Growth differentiation factor 8*	T	C	0.5	5.3	94.2	0.063	0.055	−0.181	0.441
TYRP1	*Tyrosinase related protein 1*	C	T	2.8	23.4	73.9	0.254	0.057	−0.480	0.339
ZP2	*Zona pellucida glycoprotein 2*	T	C	4.8	27.7	67.5	0.308	0.087	0.873	0.215
MYH1	*Myosin, heavy polypeptide 1*	G	A	84.4	15.0	0.6	0.169	0.196	3.353	0.001

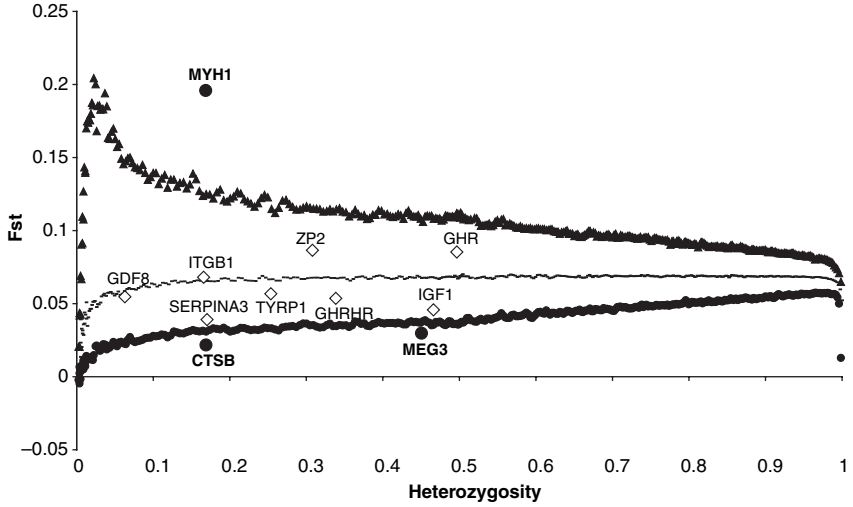

Figure 1 Plot of F_{ST} against heterozygosity for the 11 SNPs analysed. Outliers are indicated with closed circles; all other SNPs are shown with diamonds. ▲, upper and lower 95% confidence limits. ■, median of 200 000 replications of expected F_{ST} and heterozygosity using the coalescent model.

Three out of 11 loci (*MYH1*, *MEG3* and *CTSB*) were found to lie outside the 95% confidence region of the conditional joint distribution of F_{ST} and mean heterozygosity (Fig. 1), possibly because of selection. These genes are related to meat production, and differences in the genes could potentially result in morphological modifications. However, using a student *t*-test to compare breeds grouped according to their production emphasis (dairy or meat; Table 1), we did not find an association between allele frequencies of these three genes and the production emphases of the breeds. Allele frequencies for *MYH1* across Hungarian breeds ranged from 0.48 to 0.52, suggesting that there was no selection acting on the gene in this region.

Detecting evidence for selection is challenging because the effect of selection on the distribution of genetic variation can be mimicked by population demographic history (Akey *et al.* 2002). The use of SNPs as genetic markers provides the opportunity to identify a genome-wide signature of selection, as shown in humans (Sunyaev *et al.* 2000; Fay *et al.* 2001). Using F_{ST} and mean heterozygosity as measures of genetic differentiation for each locus, we have identified three genes in sheep that may have been targets of selection.

Acknowledgements

This work has been partially supported by the EU ECONOGENE contract QLK5-CT-2001-02461. The content of the publication does not represent necessarily the views of the Commission or its services.

References

Akey J.M., Zhang G., Zhang K., Jin L. & Shriver M.D. (2002) Interrogating a high-density SNP map for signatures of natural selection. *Genome Research* **12**, 1805–14.

Beaumont M.A. & Nichols R.A. (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London B*, **263**, 1619–26.

Brouillette J.A. & Venta P.J. (2002) Within-breed heterozygosity of canine single nucleotide polymorphisms identified by across-breed comparison. *Animal Genetics* **33**, 464–7.

Fay J.C., Wyckoff G.J. & Wu C.I. (2001) Positive and negative selection on the human genome. *Genetics* **158**, 1227–34.

Lewontin R.C. & Krakauer J. (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**, 175–95.

Matsuzaki H., Dong S., Loi H. *et al.* (2004) Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. *Nature Methods* **1**, 109–11.

Morin P.A., Luikart G., Wayne R.K. & the SNP workshop group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution* **19**, 208–16.

Sunyaev S.R., Lathe W.C. III, Ramensky V.E. & Bork P. (2000) SNP frequencies in human genes: An excess of rare alleles and differing modes of selection. *Trends in Genetics* **16**, 335–7.

Supplementary Material

The following supplementary material is available for this article online from http://www.blackwell-synergy.com:
Table S1 Characterization of SNPs.