1Title

2

**TRACEABILITY OF FOUR EUROPEAN PROTECTED GEOGRAPHIC INDICATION (PGI)
4BEEF PRODUCTS USING SINGLE NUCLEOTIDE POLYMORPHISMS (SNP) AND
5BAYESIAN STATISTICS**

6

7Authors

8R Negrini¹, L Nicoloso², P Crepaldi², E Milanese^{1,2}, R. Marino¹, D. Perini¹., L
9Pariset³, S Dunner⁴, H Leveziel⁵, JL Williams⁶, P Ajmone Marsan¹

10

11¹Istituto di Zootechnica, Università Cattolica del S. Cuore, Piacenza, Italy

12²Sezione di Zootechnica Agraria, Dipartimento di Scienze Animali, Università degli
13Studi di Milano, Italy

14³Dipartimento Produzioni Animali, Università della Tuscia, Viterbo, Italy.

15⁴Dpto de Producción Animal, Facultad de Veterinaria, Madrid, Spain

16⁵Unité de Génétique Moléculaire Animale, INRA/Université de Limoges, Faculté des
17Sciences et Techniques, Limoges Cedex, France.

18⁶Parco Tecnologico Padano, Polo Universitario, Lodi 26900, Italy

19

20**Corresponding author:** Riccardo Negrini, Istituto di Zootechnica, Università
21Cattolica del S. Cuore, Via E. Parmense, 84, 29100

1Abstract

2The use of SNPs in combination with Bayesian statistics for the geographic traceability of cattle were
3evaluated using a dataset comprising 24 breeds from Italy, France, Spain, Denmark, the Netherlands,
4Switzerland and UK genotyped with 90 polymorphic markers.

5The percentage of correct assignment of the individuals to their Country of origin was 90%, with an
6average assignment probability of 93% and an average specificity of 92%. The higher value was
7observed for UK breeds (97% of correct assignment) while Swiss animals were the most difficult to
8allocate (77% of correct assignment).

9Tracing of Protected Geographic Indication (PGI) products, the approach correctly assigned 100% of
10Guaranteed Pure Highland Beef; 97% of “Vitellone dell’Appennino Centrale” breeds; 84% of Ternera
11de Navarra, and 80% of Boeuf de Chalusse.

12Methods to verify Products of Designated Origin (PDO) and Protected Geographic Indication (PGI)
13products will help to protect regional foods and promote the economic growth of marginal rural areas
14by encouraging the production of high quality niche market foods.

15Keywords

16SNPs, traceability, breed, cattle

17Introduction

18By the late 1990s, a new geographical diversity of agro-food emerged. While the globalization of
19trade in foodstuffs continues apace, Europe has experienced an increasing interest in foods with local
20and regional identities. Local agri-food production systems have indeed been characterized by various
21strategies to promote local/regional food products (Murdoch et al., 2000; Marsden et al., 2002;
22Goodman, 2004; Ilbery et al., 2005). This trend has led to legislation in Europe to provide legal
23protection to regional foods, through the ‘PGI’ (Protected Geographic Indication) and ‘PDO’
24(Protected Designation of Origin) labels (European Union Regulation (EEU) 2081/92). The former
25are products produced, processed and prepared in a specific geographic area using defined materials
26and methodology. The latter are products in which at least one stage of production is covered by the

1geographical protection. In both cases the geographic component is the key aspect of the protection.

2The aims of this legislation were: to support diversity in agricultural production, to protect consumers

3by giving them information on the specific characteristic of the product and to protect product names

4against fraud and imitation (Ilbery and Kneafsey, 2000; Parrott et al., 2006). Legislation appropriate

5methods to ensure “traceability” are indeed essential and play a key role in any modern food safety

6control and verification system for products. An effective traceability system contributes to prevent

7frauds, provides an effective method for the assessment and management of food risk, facilitates

8disease control procedures and contributes to increase the consumer confidence in product safety.

9DNA-based methods offer the possibility to identify animals and animal derived foodstuffs at

10different taxonomic levels, from single individuals, to breeds or population, species and higher taxon,

11along the food chain from the farm until consumption. Therefore, they can provide a way to verify the

12accuracy of traditional identification methods such as ear tagging animals and product labelling.

13DNA-based approaches have also several further advantages over systems based on paper audits: i)

14DNA is a relatively stable molecule even when treated at high temperature (up to 120 C°); ii) the

15results of a DNA assay is independent of age and sex; iii) testing can be carried out starting from a

16wide variety of biological materials; iv) is the results are highly repeatable and can be easily

17standardized and automated (Lenstra,2003).

18Recent advances in high-throughput DNA sequencing and bioinformatics have facilitated the

19identification and validation of large collections of Single Nucleotide Polymorphisms (SNPs) in a

20number of species (Hall, 2007). SNPs are the fundamental unit of genetic variation and are very

21useful molecular markers due to their abundance (Heaton et al., 2005), genetic stability (Markotsova

22et al., 2000) and suitability to automated analysis (Lindblad-Toh et al., 2000). SNPs have been

23successfully employed in association analysis as markers for a wide range of traits (Chen et al., 2007;

24Wallstein et al., 2007), to estimate linkage disequilibrium (Angius et al., 2008) and to identify genes

25by linkage studies (Hamada et al., 2008).

1Recently, computational methods have been developed specifically to assign individuals to
2populations (Falush et al., 2007). The first tests were based on frequency statistics and calculated the
3probability of drawing a specific single multilocus genotype from each potential source group (e.g. a
4number of populations) by matching genotypes with the observed allele frequencies at each locus in
5each group (Paetkau et al., 1995). Subsequently, Rannala and Mountain (1997) pioneered the use of
6Bayesian statistics developing a partially Bayesian assignment approach to estimate population allele
7frequencies, and a frequency approach to compute the statistical significance of individual
8assignments. Now, several different methods and software are available for clustering populations and
9assigning individuals to likely populations of origin (Manel et al., 2005).

10The aims of this study were to evaluate SNP markers for their ability to allocate cattle individuals to
11their Country of origin and to trace individuals belonging to breeds with special beef certification
12(PGI) namely: “Vitellone dell’Appennino Centrale”, Italy, “Ternera de Navarra”, Spain, “Boeuf de
13Chalosse”, France and the “Guaranteed Pure Highland Beef”, UK. To address these issues we applied
14Rannala & Mountain Bayesian allocation algorithm to a SNP dataset comprising twenty-four beef,
15dairy and double-purpose breeds from Italy, France, Spain, Denmark and UK genotyped with 90 SNP
16markers.

17**Materials and methods**

18*Collection of samples*

19A large dataset of 1047 minimally related animals belonging to 24 European breeds (Table 1) was
20used. All animals were genotyped with 90 independent SNP markers (GeneBank Accession Number
21from ss77831721 to ss77831810) located in 72 genes that had the potential to affect meat quality and
22production traits. The animals came from two sources: 249 animals belonging to 13 cattle breeds were
23sampled by the authors and genotyped by KBiosciences using their proprietary KASPUR system
24(www.Kbioscience.com), while the remaining animals were sampled and genotyped within the EU
25project GemQual (QLRT-1999-30147). The average percentage of missing data in the dataset was
260.03%.

1The breeds sampled included those used in the production of four European PGI products,
2specifically:

3i) **Vitellone Bianco dell'Appennino Centrale**: an Italian PGI product for which only Chianina,
4Marchigiana and Romagnola breeds are allowed.

5ii) **Tenera de Navarra**: a Spanish PGI. Pirenaica, Blonde d'Aquitaine, Pardo Alpina (Brown Swiss
6cattle), Charolais and their crosses can be commercialised under this certification.

7iii) **Boeuf de Chalosse**: only Limousin, Blonde d'Aquitaine, Bazadaise (not sampled) are admitted to
8this brand.

9iv) **Guaranteed Pure Highland Beef**: only fresh beef products from Highland are backed up by a
10Certificate from the Highland Cattle Society approved by the Trade Mark Examiner.

11For our purpose we use the animals of the correct breeds already in our dataset as “illustrations” of
12animals that could be part of the PGI scheme.

13Statistical analysis

14Summary statistics were calculated using PowerMarker V3.0 (URL: <http://www.powermarker.net>)
15with the default setting, these included Reynolds, Weir, and Cockerham's (1983) genetic distance. Not
16considering the terms related to the population size, it can be simplified as:

$$\theta_w = \frac{\sum_{j=1}^m \sum_{i=1}^{a_j} (p_{ij} - q_{ij})^2}{2 \sum_{j=1}^m \left(1 - \sum_{i=1}^{a_j} p_{ij} q_{ij} \right)}$$

17
18where p_{ij} and q_{ij} are the frequencies of i_{th} allele at the j_{th} locus in populations X and Y respectively,
19while a_j is the number of alleles at the j_{th} locus, and m is the number of loci examined. Not
20considering mutation, this distance measure is expected to rise in a linear way depending on amount
21of genetic drift and therefore is appropriated to study cattle breeds that have diverged recently.

22The distance matrix was graphically visualised by Multi-Dimensional Scaling techniques [MDS]
23using the software Statistica ver. 7.0 (StatSoft Inc. <http://statsoft.com>).

1The relationships at individual level were assessed by the Correspondence Analysis using the “AFC
2sur populations” option available in the program GENETIX (Belkhir et al., 2004).

3Finally the variance associated to the differentiation among clusters of populations (here Country or
4PGI) was estimated by AMOVA using the freely available Arlequin software (Excoffier et al., 1992;
5<http://lgb.unige.ch/arlequin/>).

6Allocation test:

7The allocation of an anonymous animal to a predefined category (Country or PGI) within a set was
8performed by the Bayesian method developed by Rannala and Mountain (1997).

9All the allocation tests were performed using GeneClass2
10(<http://www.montpellier.inra.fr/URLB/geneClass/geneClass.html>).

11The probability of assignment was performed by simulating 1000 individuals by MC re-sampling
12procedure and setting the “Type I” error to 0.05 (Assignment threshold of score=0.05; Piry et al.,
132004). Five independent runs were compared.

14To evaluate the results of the allocation tests three indexes were used: *i*) sensitivity, or Rate of Correct
15Assignment, calculated as number of correct allocation to a category “j” / number of animals sampled
16from category “j”; *ii*) overall average assignment probability, as the average of the probability of any
17correct assignment calculated per category; *iii*) specificity, calculated as the number of correct
18assignment to category “j” / total (correct + incorrect) assignment to category “j”.

19Results and Discussion

20DNA-based techniques can be directly used for both of the two key requirements of the traceability:
21*i*) tracking, or the ability to follow a product through the supply chain from the farm to the consumer
22and *ii*) tracing, or the ability to identify the origins of an item upstream in the supply chain.

23In this paper we focused on bovine meat and meat products and used data from a 90 SNP panel
24together with a Bayesian statistics framework to test the assignment of individual cattle to their
25Country of origin: Italy, France, Spain, and UK. Using the same approach, we also investigated the
26accuracy of identification of individuals belonging to breeds covered by the PGI labels Ternera de

1Navarra, Beouf de Chalosse and Vitellone dell'Appennino Centrale; and to discriminate beef product
2from Highland breed protected by a special certified of origin and tracking system.

3As the power of the assignment depends closely on the level of genetic heterogeneity between breeds,
4the principal population genetics parameters per countries and per PGI (Table 2) were estimated and
5used. Only in the UK breeds was a statistical difference between Observed and Expected
6heterozygosity and a F_{is} value statistically different from 0: this occurs when significant deficit of
7heterozygotes is expected.

8Pair-wise Reynolds genetic distances were calculated between Countries and between PGIs and the
9distance matrices were graphically represented by Multidimensional Scaling multivariate techniques
10(Figure 1a, 1b). At the Country level, breeds from Italy, France and Spain clustered closed together,
11while breeds from UK and Switzerland were separated from all others. Therefore when the PGI
12brands were compared the Highland most clearly separated from Italian, French and Spanish PGI
13brands.

14Analysis of Molecular Variance a method of estimating population differentiation directly from
15molecular data and testing hypotheses about such differentiation, showed that the 10.7% of the total
16variation is explained by the between breeds component and that less than 3% accounts for the
17between countries component or betweenPGI.

18The relationships between individuals were also estimated by the Factorial Correspondence Analysis
19(AFC), a multivariate canonical analysis particularly suitable for the treatment of qualitative data. The
20results showed an overall weak genetic differentiation between breeds (data not shown). From the
21traceability point this reduces the possibility to unequivocally assign an individual correctly to the
22population of origin, as the power of the assignment is directly correlated with the genetic distance
23between the populations (Latch et al., 2006). Nevertheless, molecular approaches, in combination
24with specific statistical methods have been successfully used to identify the source breed of cattle
25individuals (Ciampolini et al., 2006; Dalvit et al., 2008). The recent availability of high-throughput

1molecular methods to genotype large SNP panels in combination with algorithms, using Bayesian
2statistics, are promising tools for the secure assignment of individuals to populations.

3To test these methods for identifying the Country of origin of an animal, and to trace PGI meat
4products a panel of 90 SNPs identified within genes potentially affecting meat quality were used. The
5SNP markers were tested in a 24 breeds including breeds that have undergone intensive genetic
6selection for milk production, beef production, or for dual purposes (dairy and beef). This wide range
7of breeds was examined considering that a significant source of meat in the European market is
8represented by young bulls belonging to dairy breeds – mainly Holstein Friesian - that are sold at
9significant low price respectto high quality meat from PGI.

10The choice of the SNPs in genes that were candidates for meat quality could potentially result in and
11ascertainment bias, - as the variability in these genes in dairy breeds may be higher than in beef
12breeds as the dairy breeds have not been under selection for meat quality. To exclude this potential
13bias, the samples were labelled according to their use (beef, dairy or dual purpose) and the efficiency
14of assigning individuals to breed to was tested and compared between these classifications. The
15average percentage of correct assignment was comparable between dairy and beef cattle and higher
16than 90% in both cases. The double purposes animals had a lower percentage of correct assignment
17(however still about 80%), possibly because selection have been less intense and hence these breeds
18are more genetically diverse. These results indicate that the SNP panel used was not bias between the
19genetic selection categories.

20The panel was therefore used to address test how well the Countries of origin could be assigned
21exploiting the geographic component of the genetic variance in cattle (Troy et al., 2001; Cymbron et
22al., 2005). The results are reported in Table 3. The overall rate of correct assignment was 90%, the
23average assignment probability 93% and the average specificity 88%.

24The highest value was observed for the UK breeds with 97% of individuals correctly allocated with
25an average assignment probability of 98.4%. Swiss animals were the most difficult to allocate in spite
26of their high genetic distances (Figure 1). However, the presence in the dataset of the Danish Red

1Cattle and Italian Red Pied, which are genetically closely related to Swiss Simmental, decreased both
2the rate of correct assignment and the specificity (78% of correct assignment and 72% of specificity).
3In addition to assigning individuals to breeds it is commercially useful to verify the origin of meat at
4Country level. Kapferer (1992), observed that marketing food products using a geographic indication
5of origin as part of the label permits consumers' associations with the region and provides the product
6with a positive image. Moreover such geographic traceability increases transparency, and consumer
7confidence in the face of recent food scares e.g. Bovine Spongiform Encephalopathy, the dioxin and
8the poultry avian influenza crisis (Lloyd et al., 2006; Puntaric, 2000; Peiris et al., 2007). The
9approaches described identify genetic variations, which are often linked to specific geographic
10distributions, such as indigenous and local breeds. However using genetic data to assign individuals to
11breeds is probabilistic, therefore does not give a definitive assignment and the interpretation of the
12data is dependent on the thresholds set. Nevertheless, the confidence level obtained in terms of
13specificity and sensitivity here is very high, for some breeds.

14The approach was also evaluated for the identification of particular PGI or PDO products linked to
15specific breeds. Like trademark, a PGI or PDO is a legally supported way to identify goods as
16originating from a particular region, and is associated with a given quality attributable to the place of
17origin. This labelling of typical local foods or brands is particularly effective for increasing their value
18and to enhance their market competitiveness.

19The present study focused on four PGI products linked to specific breeds, therefore are suitable for a
20DNA based traceability: Vitellone Bianco dell'Appennino Centrale (Chianina, Marchigiana and
21Romagnola breeds); Ternera de Navarra (Pirenaica, Blonde d'Aquitaine, Brown Swiss cattle,
22Charolais and their crosses); Boeuf de Chalosse (Limousin, Blonde d'Aquitaine) and Guaranteed Pure
23Highland Beef (beef products from Highland). The analyses were performed either considering each
24of these specific PGI products individually or testing the four PGI products simultaneously to
25evaluate if they could be distinguished. Although for some PGI products – e.g. Ternera de Navarra-

1cross-bred individuals are accepted; in this paper we consider only pure animals. The results are
2reported in Table 4.

3The panel of SNPs allocated correctly the 97% of animals belonging to Vitellone dell'Appennino
4Centrale with an average assignment probability of 97%. Only 22% (11 out of 93) of correct
5allocations had a probability of being correctly assigned lower than 90%. For Boeuf de Chalosse 80%
6of the samples were correctly assigned with an average probability of 88% and sensitivity of 81%.
7Samples from the Ternera de Navarra PGI were correctly assigned in the 84% of cases with an
8average probability of 93% and a specificity of 91%. Finally, Highland samples were assigned
9completely.

10Of course decreasing the assignment threshold will result in more animals correctly assigned but also
11in a loss of specificity, although with a slow rate (Figure 2).

12Alternatively to use all the available data, carrying out this type of verification analysis in a
13commercial setting the interest can be limited to the identification of the most likely contamination or
14source of fraud. In this case the cluster can be carried out against potential fraud breeds only
15increasing the percentage of correct allocation.

16The correct assignment was highest Vitellone del Appennino Centrale and Highland. This is for a
17number of reasons, firstly these products use only pure breeds and these breeds are unique to the
18respective PGI. The Ternera de Navarra and Boeuf de Chalosse products have a breed in common
19(Blonde d'Aquitaine) making an overlap in the definition of these products inevitable and hence
20lowering both assignment and specificity of the tests. The efficiency to identify cross-bred individuals
21using this marker set and method for analyses remains to be investigated.

22Conclusion

23Recent market research has shown that consumers are concerned with the provenance of the food they
24buy and therefore it is important to develop simple tests to control the origin of a product. In this
25view, DNA-based trace-back systems are able to identify the source of meat products and trace them

1through the supply chain. Availability of such test data is of potential benefit for consumers, while it
2also enables processors and retailers to ensure the identity of products.

3Here we showed that SNP and Bayesian statistics has the potential to be implementation for
4geographical traceability, and to test the authenticity of DOP and PGI products linked to a specific
5breed. Considering the unequal geographical distribution of DOP and PGI registered regional
6foodstuffs, with more than 75% of the products registered in five southern EU states (France, Italy,
7Greece, Portugal and Spain) and considering also that most these products come from rural areas, a
8DNA-based effective methods of brands protection - through the verification of authenticity of the
9products - may also provide tools for sustaining the viability of small farming and rural communities
10improving the Economy of marginal areas.

11The main barrier to implementation of these control methods is the cost of DNA based tests, which is
12still high. In addition the choice of markers could be better optimised to increase the assignment rate,
13which is mainly correlated with the allele frequencies of markers used. In a near future both these
14constraints may be removed with the availability of a larger number of SNP coming from the Bovine
15whole genome sequencing and HapMap projects, and by the implementation of novel high-throughput
16typing technologies that will significantly reduce test costs.

17**Acknowledgments**

18The authors acknowledge the Eu-TRACE project (contract n° 006942; www.trace.eu.org) and the
19Fondazione CARIPLO (n° 2003.0721/11.8094) for financial support and the Eu-GemQual project
20QLRT-1999-30147 for the provision of part of data.

21The authors acknowledge also the Italian Breeder Association for the provision of samples and
22pedigree information. The content of this paper does not necessarily reflect the view of the European
23Commission or its services.

24**Bibliography**

25Murdoch, J. (2000). Networks – a New Paradigm of Rural Development?. *Journal of Rural Studies*,
2616, 407–419.

1 Marsden, T., Banks, J. & Bristow, G. (2002). The Social Management of Rural Nature:
2 Understanding Agrarian-based Rural Development. *Environment and Planning A*, 34, 809–825.

3 Goodman, D. (2004). Rural Europe Redux? Reflections on Alternative Agro-food Networks and
4 Paradigm Change. *Sociologia Ruralis*, 44, 3–16.

5 Ilbery, B. & Maye, D. (2005). Food Supply Chains and Sustainability: Evidence from Specialist Food
6 Producers in the Scottish/English Borders. *Land Use Policy*, 22, 331–344.

7 Ilbery, B. & Kneafsey, M. (2000). Producer Constructions of Quality in Regional Speciality Food
8 Production: a Case Study from South West England. *Journal of Rural Studies*, 16, 217–230.

9 Parrott, N., Wilson, N. & Murdoch, J. (2002). Spatializing Quality: Regional Protection and the
10 Alternative Geography of Food. *European Urban and Regional Studies*, 9, 241–261.

11 Heaton, M.P., Keen, J.E., Clawson, M.L., Harhay, G.P., Bauer, N., Shultz, C., Green, B.T., Durso, L.,
12 Chitko-McKown, C.G. & Laegreid, W.W. (2005). Use of bovine single nucleotide polymorphism
13 markers to verify sample tracking in beef processing. *Journal of the American Veterinary Medical*
14 *Association*, 226, 1311-1314.

15 Lindblad-Toh, K., Winchester, E., Daly, M.J., Wang, D.G., Hirschhorn, J.N., Laviolette, J.P., Ardlie,
16 K., Reich, D.E., Robinson, E., Sklar, P., Shah, N., Thomas, D., Fan, J.B., Gingeras, T., Warrington, J.,
17 Patil, N., Hudson, T.J. & Lander, E.S. (2000). Large-scale discovery and genotyping of single-
18 nucleotide polymorphisms in the mouse. *Nature Genetics*, 24, 381-386.

19 Lloyd, T. A., McCorriston, S., Morgan, C. W. & Rayner, A. J. (2006). Food scares, market power and
20 price transmission: the UK BSE crisis. *European Review of Agricultural Economics*, 33, 119-147.

21 [Markovtsova, L., Marjoram, P. & Tavaré, S. \(2000\).](#) The effects of rate variation on ancestral
22 inference in the coalescent. *Genetics*, 156, 1427-1436.

23 McKean, J.D. (2001). The importance of traceability for public health and consumer protection.
24 [Revue scientifique et technique \(International Office of Epizootics\)](#), 20, 363-371.

25 Lenstra, J.A. (2003). DNA methods for identifying plant and animal species in food. In Food
26 Authenticity and Traceability. Whoodhead Publishing Ltd- Cambridge UK.

1Chen, W.M. & Abecasis, G.R. (2007). Family-based association tests for genomewide association.
2*American journal of human genetics*, 81, 913-926.

3Angius, A. , Hyland, F.C., Persico, I., Pirastu, N., Woodage, T. & De la Vega, F.M. (2008). Patterns
4of linkage disequilibrium between SNPs in a sardinian population isolate and the selection of markers
5for association studies. *Human heredity*, 65, 9-22.

6Hamada, D., Takata, Y., Osabe, D., Nomura, K., Shinohara, S., Egawa, H., Nakano, S., Shinomiya,
7F., Scafe, C.R., Reeve, V.M., Miyamoto, T., Moritani, M., Kunika, K., Inoue, H., Yasui, N. & Itakura,
8M. (2005). Association between single-nucleotide polymorphisms in the SEC8L1 gene, which
9encodes a subunit of the exocyst complex, and rheumatoid arthritis in a Japanese population *Arthritis*
10*and rheumatism* 52, 1371-1380.

11Wollstein, A., Herrmann, A., Witting, M., Nothnagel, M., Franke, A., Nürnberg, P., Schreiber, S.,
12Kreawczak, M. & Hampe, J. (2007). Efficacy assessment of SNP sets for genome-wide disease
13association studies. *Nucleic Acids Research*, 35,17, e113.

14Falush, D., Stephens, M. & Pritchard, J.K. (2007). Inference of population structure using multilocus
15genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, 7, 574-578.

16Paetkau, D., Calvert, W., Stirling, I. & Strobeck, C. (1995). Microsatellite analysis of population
17structure in Canadian polar bears. *Molecular Ecology*, 4, 347-354.

18Rannala, B. & Mountain, J.L. (1997). Detecting immigration by using multilocus genotypes.
19*Proceedings of the National Academy of Science of the USA*, 94, 9197-9221.

20Manel, S., Gaggiotti, O.E. & Waples, R.S. (2005). Assignment methods: matching biological
21questions with appropriate techniques. *Trends in Ecology & Evolution*, 20, 136-142.

22Reynolds, J., Weir, B.S., & Cockerham, C.C. (1983). Estimation of the Coancestry coefficient: basic
23for a short-term genetic distance. *Genetics*, 105, 767-779.

24Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (2004). GENETIX 4.05, logiciel
25sous Windows® pour la génétique des populations. Laboratoire Génome, Populations, Interactions,
26CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

1Excoffier, L, Smouse P.E., Quattro, J.M. (1992) Analysis of molecular variance inferred from metric
2distances among DNA haplotypes: application to human mitochondrial DNA restriction data.
3*Genetics*,131, 479-491.

4Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L. & Estoup, A. (2004). GeneClass 2: A
5Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity*, 95,
6536-539.

7Botstein, D., White, R.L., Skolnick, M. & Davis, R.W. (1980). Construction of a genetic linkage
8map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*,
932, 314-331.

10Weir, B.S. (1996). Genetic data analysis II, Sunderland, MA: Sinauer Associates, Inc.

11Latch, E.K., Dharmarajan, G., Glaubitz, J.C. & Rhodes, O.E. (2006). Relative performance of
12Bayesian clustering software for inferring population substructure and individual assignment at low
13levels of population differentiation. *Conservation Genetics*,7, 295–302.

14Ciampolini, R., Cetica, V., Ciani, E., Mazzanti, E., Fosella, X., Marroni, F., Biagetti, M., Sebastiani,
15C., Papa, P., Filippini, G., Cianci, D. & Presciuttini, S. (2006). Statistical analysis of individual
16assignment tests among four cattle breeds using fifteen STR loci. *Journal of Animal Science*, 84, 11–
1719.

18Dalvit, C., De Marchi, M., Dal Zotto, R., Gervaso, M., Meuwissen, T., Cassandro, M. (2008). Breed
19assignment test in four Italian cattle breeds. *Meat Science*doi: 10.1016/j.meatsci.2008.01.001

20Troy, C.S., MacHugh, D.E., Bailey, J.F., Magee, D.A., Loftus, R.T., Cunningham, P., Chamberlain,
21A.T., Sykes, B.C. & Bradley, D.G. (2001). Genetic evidence for Near-Eastern origins of European
22cattle. *Nature* 26, 410, 1088-1091.

23Cymbron, T., Freeman, A.R., Isabel Malheiro, M., Vigne, J.D. & Bradley, D.G. (2005). Microsatellite
24diversity suggests different histories for Mediterranean and Northern European cattle populations.
25Proceedings. *Biological Science* 272, 1837-1844.

1Kapferer, J.N. (1992). Strategic brand management: new approaches to creating and evaluating brand
2equity. London: Kogan Page.

3Lloyd, T. A., McCorriston, S., Morgan, C. W., & Rayner, A. J. (2006). Food scares, market power
4and price transmission: the UK BSE crisis. *European Review of Agricultural Economics*, 33, 119-147.

5Puntaric, D., Smit, Z., Bosnir, J., & Topolovec, J. (2000). Small countries and the dioxin scandal:
6how to control imported food? *Croatian Medical Journal*, 41, 150-153.

7Peiris, J.S.M., de Jong, M. D., & Guan, Y. (2007). Avian Influenza Virus (H5N1): a Threat to Human
8Health Clinica. *Microbiology Reviews*, 20, 243-267.

9Hall, N. (2007). Advanced sequencing technologies and their wider impact in microbiology. *Journal*
10*of Experimental Biology*, 210, 1518-1525.

1Figure Captions

2Figure 1a and 1b: Multi-dimensional scaling representation of the Reynolds genetic distances
3calculated between Countries (a)and PGI (b).

4Figure 2: Line plot of %sensitivity and %specificity against the % of the probability. As expected
5increasing the assignment threshold results in a higher sensitivity but also in a loss of specificity.

6Figures

7Figure 1a and 1b

8

1

2Figure 2

3

Table 1. Number of individuals sampled, country of origin and selection purpose of the breeds analysed.

Country of origin	Country of sampling	Breed	Acronym	N° of individuals	Selection purpose
DK	DK	Red Cattle	RED	57	Dairy
CH	DK	Simmental	SIM	19	Beef
ES	ES	Asturiana de la montaña	CAS	55	Dairy
ES	ES	Asturiana de los Valles	RAV	56	Beef/Dairy
ES	ES	Avilena	AVI	53	Beef
ES	ES	Pirenaica	PIR	71	Beef
FR	FR	Blonde d'Aquitaine	BLO	19	Beef
FR	FR	Charolais	CHA	82	Beef
FR	FR	Limousin	LIM	96	Beef
FR	FR	Maine Anjou	MDA	19	Beef/Dairy
FR	FR	Parthenaise	PAR	14	Dairy
FR	FR	Salers	SAL	20	Beef/Dairy
IT	IT	Chianina	CHI	19	Beef
CH	IT	Brown	ITB	21	Dairy
NL	IT	Holstein Friesian	ITH	77	Dairy
IT	IT	Italian Red Pied	PRI	23	Beef/Dairy
IT	IT	Marchigiana	MCG	55	Beef
IT	IT	Maremmiana	MMA	22	Beef
IT	IT	Piemontese	PIM	85	Beef
IT	IT	Romagnola	ROM	19	Beef
UK	UK	Aberdeen Angus	ABA	38	Beef
UK	UK	Highland	HIG	46	Beef
UK	UK	Jersey	JER	46	Dairy
UK	UK	South Devon	SOD	35	Beef
		Total		1047	

Table 2. Summary statistics calculated per Country and per PGI. Av= Availability defined as $1 - 2\text{Obs}/n$, where Obs is the number of observations and n is the number of individuals sampled. Het. Exp. = Heterozygosity expected following Weir (1996); Het Obs. = Heterozygosity Observed; Fis= inbreeding-like effects within populations. * = t test significant thresholds $p < 0.05$; **= statistically different from 0.

	Major.Allele	No. Ind.	Av.	Het. Exp	Het Obs	Fis
Country						
France	0.7803	250	0.9772	0.3006	0.2869	0.0478
Italy	0.7822	224	0.9804	0.3026	0.2835	0.0744
Spain	0.7784	235	0.9667	0.3045	0.2884	0.0549
UK	0.7765	165	0.947	0.3121*	0.2679*	0.1451**
Swiss	0.7958	40	0.9808	0.2808	0.2763	0.0360
PGI						
PGI_SPAIN	0.7748	153	0.9651	0.302	0.2947	0.0312
PGI_FRANCE	0.7895	115	0.9832	0.2889	0.2863	0.0177
PGI_ITALY	0.7798	94	0.9774	0.3041	0.2927	0.0482
HIGHLAND	0.8366	46	0.9486	0.2233	0.2197	0.0282

1**Table 3.** Assignment statistics calculated grouping individuals according to the Country of origin. No.

2Ind.= Number of animal sampled per Country; Not Ass = number of animal not assigned; % Correct =

3Sens.= percentage of correct allocation with probability higher than 50%; Spec= Specificity; Av.

4Prob.= Average Probability of the correct assignments.

	No. Ind.	Not Ass.	Sens	Spec	Av Prob
Italy	200	22	0.89	0.89	0.92
France	250	30	0.88	0.89	0.9
UK	165	5	0.97	0.99	0.98
Spain	235	24	0.90	0.91	0.93
Swiss	40	9	0.78	0.72	0.92
Overall	890	90	0.90	0.88	0.93

5

6**Table 4.** Assignment statistics calculated grouping the breed allowed for the production of PGI. No.

7Sampled= Number of animal sampled per PGI. Not assigned= number of animal not assigned. %

8Correct = % correct = percentage of correct allocation with probability higher than 50%. Spec.=

9Specificity. Av. Prob.= Average Probability of the correct assignments.

	No. Ind.	Not Assigned	% correct	Spec.	Average Prob
Vitellone bianco	93	3	0.97	0.96	0.97
Boeuf de Chalosse	115	22	0.80	0.88	0.81
Ternera de Navarra	193	31	0.84	0.93	0.81
Hghland	96	0	1.00	1	1
Overall	497	56	0.90	0.94	0.90

10