

Genetic variation in nine European cattle breeds as determined on the basis of microsatellite markers. II. Gene migration and genetic distance

Grzegorz Grzybowski, Beata Prusak

Polish Academy of Sciences Institute of Genetics and Animal Breeding,
Jastrzębiec, 05-552 Wólka Kosowska, Poland

(Received January 5, 2004; accepted February 27, 2004)

Basing on the polymorphism at 26 DNA microsatellite *loci* an estimation was made of the mean number of genes migrating in one generation ($N_e m$) between populations of nine cattle breeds maintained in Germany, Switzerland and Poland: Angler, German Simmental, Brown Swiss, Swiss Simmental, Holstein, Eringer, Polish Red (PR – included in the National Rare Livestock Breeds Preservation Programme), Polish Black-and-White (PBW) and Polish Red-and-White (PRW). The $N_e m$ value for the whole population amounted to 2.048. When the breeds were grouped according to the country of origin the greatest flow of genes was observed within the Polish (2.023), while the lowest within the German population (1.214). The $N_e m$ values estimated when comparing each breed with each show that the greatest exchange of genetic material took place between breeds from the same geographic regions. On the phylogenetic N-J tree, constructed on the basis of genetic distance D_A , the PR were classified together with PBW and PRW cattle. However, when evaluating the relations between breeds on the basis of values estimated for the D_{sw} genetic distance, the PR cattle were classified together with the Angler breed. Bootstrap values were higher when the tree was constructed on the basis of the D_A and not the D_{sw} distance, what indicates a higher reliability of the genealogical structure determined on the basis of the former. In turn, the D_{sw} distance gave a more precise information about the common background of the German and Swiss Simmentals, as well as about the share of the Angler breed used in the past for upgrading the PR.

KEY WORDS: cattle / gene migration / genetic distance / microsatellite markers

Parallel to the FAO programme aiming at preservation of farm animal genetic resources throughout the world [Anonymous 1996a], a programme was initiated for the analysis of genetic diversity in cattle (European Concerted Action AIRE 2006 for the Analysis of Genetic Diversity to Preserve Future Breeding Options). The programme

was based on an analysis of the polymorphism of microsatellite DNA markers [Anonymous 1996b]. The results obtained are collected in an internet database and designated for the preserve breeding programmes in individual countries. The unification of the principles for selecting material, as well as of the panel of the microsatellite *loci* analysed, renders it possible to test different models and thus elaborate such formulas, which result in the most precise description of a population genetic structure on the basis of differentiated microsatellite markers and reflects to the greatest degree the existence of inter-breed genetic links.

Part I of this study [Lubieniecka *et al.* 2001] was conducted according to the criteria of the European programme mentioned. Determined was the differentiation within and between Angler, German Simental, Brown Swiss, Swiss Simental, Holstein, and Eringer breeds. On this background the genetic structure was presented of Polish cattle breeds: Red (PR – included in the National Rare Livestock Breeds Preservation Programme – NRLBPP), Black-and-White (PBW) and Red-and-White (PRW).

One may accept that from the time of introduction of pedigree books the population of European cattle was divided into several groups, isolated as regards reproduction. Before that, an exchange of the genetic material did probably take place, through random or planned matings. Moreover, native breeds, such as PR, were upgraded to improve their production traits and this also led to the incorporation of genes foreign to them.

The present work aimed at estimating potential gene migration between the populations mentioned as well as determining two genetic distances: D_A , considering the genetic drift as the principal factor genetically differentiating populations, and D_{sw} , considering the stepwise model for generating mutations that reflects DNA polymerase slippage.

Material and methods

The animal material, criteria for the selection of microsatellite *loci* used for analyses as well as the molecular procedure of automated DNA sizing technology for allele identification were described in Part I of this study [Lubieniecka *et al.* 2001].

In the present Part II the mean number of genes migrating between populations in one generation was determined on the basis of parameter $N_e m$ [Slatkin 1985] where N_e is the effective population size and m the per cent of migrating genes. The $N_e m$ value was estimated using the GENEPOP software [Raymond and Rousset 1995]. Genetic relations between populations were determined on the basis of genetic distances D_A and D_{sw} .

For calculating the D_A value the procedure of the modified Cavalli-Sforza (f_0) distance according to Nei *et al.* [1983] was applied. The value of distance D_{sw} (stepwise weighted genetic distance), which takes into consideration the stepwise mutation model, was estimated according to Shriver *et al.* [1995].

The data on genetic distances between breeds were used for constructing phylogenetic trees. With this aim the neighbour-joining (N-J) method was applied as described by Saitou and Nei [1987]. The values of the D_A distance were calculated using the DISPAN [Ota 1993], while those of the D_{sw} – using the PHYLIP softwares [Minch

1998]. The reliability of the trees obtained (and thus of the adequacy of the calculation model applied) was estimated by the bootstrap test [Felsenstein 1985] and expressed by the bootstrap value – a per cent share of the frequency of a given branch in the phylogenetic tree per 1000 simulated repeats.

Results and discussion

In order to determine the degree of the genetic material migration between populations of European cattle the mean number of genes exchanged in one generation ($N_e m$) was estimated basing on an analysis of the frequency of alleles rarely occurring in individual breeds. Sladkin [1985] demonstrated, that the existence of a direct relation between the mean frequency of alleles occurring in only one population and the $N_e m$ value is important for the applicability of this procedure in population analyses. This renders it possible to elaborate a regression curve, which may be used for estimating the mean per cent of genes exchanged between populations ($N_e m$) at a defined mean frequency of rare alleles – $p(1)$. The determination of $N_e m$ seems to be considerably informative when estimating relations between breeds and populations on the basis of polymorphism at DNA microsatellite *loci*. Decisive in this respect is the high polymorphism at these *loci* and thus the high probability of rare alleles to occur in the populations analysed (and this is precisely the key element of the analysis). The principal condition ensuring an objective comparison of results obtained by different research groups is a unification of procedures for the determination of allele size. This is especially important in the case of domestic cattle, as in those animals the DNA microsatellite polymorphism arises from the presence of di-nucleotide tandem repeats. The lack of a unified allele identification procedure may easily lead to an incorrect evaluation of individual alleles of similar sizes, what in turn renders difficult a fully objective comparison of the polymorphism of DNA microsatellites from different cattle breeds. The strict criteria, adopted by the European Concerted Action AIRE 2006 for the Analysis of Genetic Diversity to Preserve Future Breeding Options [Anonymous 1996b], applied in the analyses presented, aim at eliminating such divergences. Barton and Slatkin [1986] developed a method, which takes into consideration uneven numbers of animals in populations by introducing an interpolation to the nearest regression curve. This ad-

Table 1. Mean values for genes exchanged per generation ($N_e m$) as estimated for German, Swiss and Polish cattle populations

Population	Breed	$N_e m$	$p(1)$
German	Angler, German Simmental	1.214	0.0348
Swiss	Swiss Brown, Eminger, Swiss Simmental, Holstein	1.486	0.0484
Polish	Polish Red, Polish Black-and-White, Polish Red-and-White	2.023	0.0295
Total		2.049	0.0353

$p(1)$ – mean frequency of unique ("private") alleles.

ditionally increases the efficiency of the comparative polymorphism analysis. The $N_e m$ values obtained and the mean frequency for alleles occurring in only one population – $p(1)$ – are presented in Tables 1 and 2. The $N_e m$ value for a joint cattle population, composed of nine breeds, amounted to 2.049 (Tab. 1). In turn, when the breeds were grouped depending on the country of origin (*i.e.* Germany, Switzerland and Poland) the greatest gene migration was observed in the Polish (2.023), while the lowest in the German (1.214) population – Table 1.

Table 2 Mean values for genes exchanged per generation ($N_e m$) as summed between European cattle breeds

Population	$N_e m$	$p(1)$
Polish Red – Polish Red-and-White	1.073	0.039
Polish Red – Polish Black-and-White	1.395	0.032
Polish Red – Holstein	1.038	0.040
Polish Red – Anglo	0.483	0.064
Polish Red – Charolais Simmental	0.902	0.043
Polish Red – Swiss Simmental	0.603	0.056
Polish Red – Brown Swiss	0.668	0.052
Polish Red – Etzger	0.685	0.051
Polish Black-and-White – Polish Red-and-White	2.385	0.062
Polish Black-and-White – Holstein	1.185	0.056
Polish Black-and-White – Charolais Simmental	1.081	0.059
Polish Black-and-White – Anglo	0.609	0.084
Polish Black-and-White – Swiss Simmental	0.606	0.084
Polish Black-and-White – Brown Swiss	0.830	0.069
Polish Black-and-White – Etzger	0.585	0.086
Polish Red-and-White – Holstein	1.255	0.054
Polish Red-and-White – Anglo	0.564	0.088
Polish Red-and-White – Charolais Swiss	1.106	0.058
Polish Red-and-White – Swiss Simmental	0.647	0.081
Polish Red-and-White – Brown Swiss	0.813	0.070
Polish Red-and-White – Etzger	0.560	0.088
Holstein – Anglo	0.805	0.070
Holstein – Charolais Simmental	1.196	0.055
Holstein – Brown Swiss	1.050	0.060
Holstein – Swiss Simmental	1.088	0.059
Holstein – Etzger	1.279	0.053
Anglo – Charolais Simmental	1.214	0.055
Anglo – Swiss Simmental	0.774	0.072
Anglo – Brown Swiss	0.560	0.088
Anglo – Etzger	0.691	0.077
Charolais Simmental – Swiss Simmental	1.444	0.049
Charolais Simmental – Brown Swiss	1.105	0.058
Charolais Simmental – Etzger	1.078	0.059
Swiss Simmental – Brown Swiss	1.221	0.055
Swiss Simmental – Etzger	1.279	0.051
Etzger – Brown Swiss	1.427	0.050

$p(1)$ – mean frequency of unique (private) alleles

The $N_e m$ values estimated for individual population pairs (Tab. 2) indicate the greatest exchange of the genetic material between breeds originating from the same geographical region. Within the breeds analysed, only the Holstein cattle showed a comparatively wide gene exchange with all the populations analysed ($N_e m = 1$) – the Angler cattle constituting the only exception ($N_e m = 0.805$). A considerable gene migration between Holstein cattle and the remaining breeds may be interpreted to be the result of a wide use of Holstein breed throughout the world for improving the dairy performance in numerous populations (including PBW and PRW).

A comparison of the $N_e m$ values indicates a greater gene migration between the PR and two other Polish breeds, Holsteins ($N_e m > 1$) and German Simmentals ($N_e m = 0.902$) than between the PR and Angler cattle ($N_e m = 0.483$) – Table 2. As in the past the PR cattle have been crossed with the Angler breed [Czaja 1991] one could expect this to be reflected by a high $N_e m$. However, the PR cattle, analysed in the present study, cannot be treated as a random sample from the population. The material was selected for the National Rare Livestock Breeds Preservation Programme according to strictly defined criteria. The low level of gene migration between the PR and Angler cattle indicates that the material selected retained the principal traits of former pure-bred PR cattle. On this basis one may state that the selection of animals for preserve breeding, although based on exterior and the pedigrees (not verified by blood or DNA tests) proved to be highly efficient. This is also an indirect proof that, as regards exterior traits, PR cattle is highly characteristic, what renders it possible to differentiate this breed from single coat-colour red cattle bred in other countries. The greatest gene migration was observed between the PBW and PBR cattle ($N_e m = 2.385$) – Table 2, what seems to reflect the existence of common genetic roots for both those populations.

In order to determine the degree of genetic relations between the cattle breeds analysed an estimation was made of the D_A and D_{sw} genetic distances. The results are presented in Figures 1 and 2 as phylogenetic neighbour-joining (N-J) trees. Estimating the topology of the tree constructed on the basis of the D_A distance one may state that the nine cattle breeds analysed create three separate territorial groups, originating from the same countries (Fig. 1). Thus, the grouping of German, Swiss and Polish breeds reflects the effect of the geographic origin on the genetic structure of the European cattle. This could arise from a lower gene migration between breeds originating from different countries and a greater probability of there existing common ancestors for breeds maintained in a geographical proximity. However, the fact that Holstein cattle groups with Polish breeds (though at a low bootstrap level – 37%) indicates a certain degree of gene migration between these breeds. The highest bootstrap values (> 80%) were obtained for a group consisting of PBW and PRW breeds and for the group of Swiss breeds.

Not all the elements of the tree structure constructed on the basis of distance D_A (Fig. 1) were confirmed in the topology of the tree constructed on the basis of distance D_{sw} (Fig. 2). The principal feature in the D_{sw} topology is the presence of PBW, PRW and Holstein cattle in one group (at a 40% bootstrap value). Moreover, the D_{sw} distance

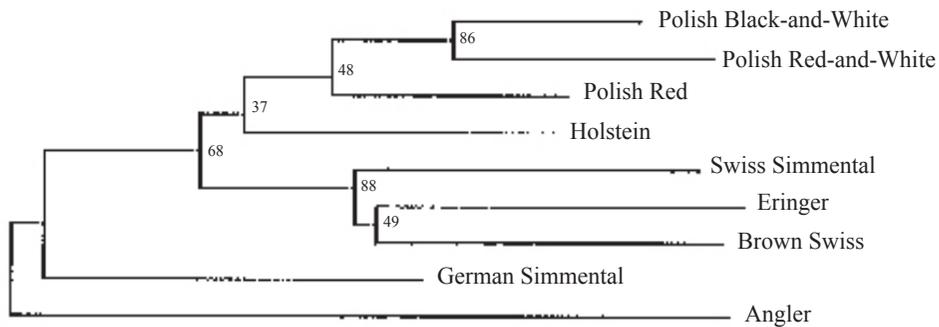


Fig. 1. Phylogenetic tree (neighbour-joining method) constructed on the basis of D_A distance between individual cattle breeds.



Fig. 2. Phylogenetic tree (neighbour-joining method) constructed on the basis of D_{SW} distance between individual cattle breeds.

indicates a closer relation between the PR and Angler breeds (a 64% bootstrap value) than did the D_A distance. Also, the D_{SW} values indicate a closer genetic relation between German and Swiss Simmentals than between Anglers and German Simmentals or between Swiss Simmentals and the remaining Swiss breeds.

The results obtained in this study indicate that the genetic structure of the European cattle analysed is not very stable, as the trees constructed show different structures depending on the calculation method used. This points to the necessity of a selective and critical choice of the method used for elaborating the data obtained from an analysis of microsatellite marker polymorphism. The divergences mentioned in the topology

of N-J phylogenetic trees constructed on the basis of D_A and D_{sw} , make it necessary to consider, which of them reflects the genetic relations between breeds more precisely, and which would be more valuable for future analyses based on DNA microsatellite polymorphism. Such information relating to cattle is incomplete, and in Poland such analyses have not yet been performed. Generally, the D_A distance seems more valuable for the analysis of populations related to each other, in which genetic differentiation is mainly dependent on genetic drift [Forbes *et al.* 1995, Goldstein *et al.* 1995]. However, the D_A model is not „sensitive” to the presence of rare alleles, what means that the f_0 value decreases with the increasing population size. The D_{sw} distance [Shriver *et al.* 1995] was elaborated specially for analysing microsatellite markers and other VNTR *loci* and takes into consideration the stepwise mutation model, specific for those markers. It is accepted that parameters estimated in this model are more adequate for determining the genetic structure of a population on the basis of an analysis of DNA microsatellite alleles, because they reflect more precisely the molecular mechanism generating the polymorphism known as DNA polymerase slippage. Thus, a stepwise mutation model would be particularly useful for the reconstruction of processes distant in time, differentiating breeds and populations, *i.e.* similar to those which in the past led to the separation of humped (zebu) and humpless cattle.

Summarizing the results obtained from an analysis of both genetic distances one may state, that the breed structure of the analysed populations of European cattle is principally a reflection of the geographic origin of individual breeds. As the phylogenetic tree constructed on the basis of distance D_A was characterized by higher bootstrap values, this distance seems to be more reliable. In turn, the fact that in the structure of the tree constructed on the basis of D_{sw} the Swiss Simmentals appear together with the German Simmentals while Polish Red cattle with Anglers, indicates that this distance more accurately reflects the history of individual breeds, and especially the existence of common ancestors for German and Swiss Simmentals, as well as crossing the PR cattle with Anglers.

REFERENCES

1. ANONYMOUS, 1996a – Domestic Animal Diversity Information System (DAD-IS) Database. Food and Agriculture Organization of the United Nations. <http://dad.fao.org/en/Home.htm>
2. ANONYMOUS, 1996b – CaDBase – Cattle Diversity DataBase, – The Analysis of Genetic Diversity in Cattle to Preserve Future Breeding Options – European Concerted Action Project under Food, Agriculture and Agro-Industry Research (FAIR) Programme. Roslin Institute, Edinburgh. http://www.ri.bbsrc.ac.uk/cdiv_www/
3. BARTON N.H., SLATKIN M., 1986 – A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56, 409-415.
4. CZAJA H., 1991 – Polskie bydło czerwone – wielowiekowa historia bez „happy endu” (Polish Red cattle – a history of many centuries but without a happy end). In Polish. *Przegląd Hodowlany* 10, 5-8.
5. FELSENSTEIN J., 1985 – Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 29, 783-791.

6. FORBES S.H., HOGG J.T., BUCHANAN F.C., CRAWFORD A.M., ALLENDORF F.W., 1995 – Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. *Molecular Biology and Evolution* 12 (6), 1106-1113.
7. GOLDSTEIN D.B., RUIZ LINARES A., CAVALLI-SFORZA L.L., FELDMAN M.W., 1995 – An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139, 463-471.
8. LUBIENIECKA J., GRZYBOWSKI G., LUBIENIECKI K., 2001 – Genetic variation in nine European cattle breeds as determined on the basis of microsatellite markers. I. Within-breed variation. *Animal Science Papers and Reports* 19 (4), 249-264.
9. MINCH E., 1998 – MICROSAT (version 1.5e). Stanford University Medical Center, Stanford, CA.
10. NEI M., TAJIMA F., TATENO Y., 1983 – Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* 19, 153-170.
11. OTA T., 1993 – DISPAN. Pennsylvania State University, Institute of Molecular Evolutionary Genetics.
12. RAYMOND M., ROUSSET F., 1995 – A population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86, 248-249.
13. SAITOU N., NEI M., 1987 – The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4, 406-425.
14. SHRIVER M.D., JIN L., BOERWINKLE E., DEKA R., FERRELL R.E., CHAKRABORTY R., 1995 – A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Molecular Biology and Evolution* 12(5), 914-920
15. SLATKIN M., 1985 – Rare alleles as indicators of gene flow. *Evolution* 39, 53-65.

Grzegorz Grzybowski, Beata Prusak

Zmienność genetyczna dziewięciu europejskich ras bydła określona na podstawie markerów mikrosatelitarnych.

II. Migracja genów i dystans genetyczny

Streszczenie

Na podstawie polimorfizmu w 26 *loci* mikrosatelitów DNA, oszacowano średnią liczbę migrantów wymienionych w jednym pokoleniu ($N_e m$) między populacjami bydła dziewięciu ras hodowanych w Niemczech, Szwajcarii i w Polsce: angler, simmental niemiecki, szwyc, simmental szwajcarski, holsztyn, eringer, polska czerwona (PR) z hodowli zachowawczej, cb (PBW) i czb (PRW).

Wartość $N_e m$ dla całej populacji wyniosła 2,048. Przy pogrupowaniu ras w zależności od kraju pochodzenia, największy przepływ genów zanotowano w populacji polskiej (2,023), a najmniejszy w niemieckiej (1,214). Wartości $N_e m$ oszacowane przy porównaniu każdej rasy z każdą, wskazują na największą wymianę materiału genetycznego między rasami pochodzącymi z tych samych regionów geograficznych. W drzewie filogenetycznym N-J zbudowanym na podstawie dystansu genetycznego D_A , wystąpiło grupowanie rasy pc z rasami cb i czb. Natomiast przy ocenie relacji między rasami na podstawie wartości oszacowanych dla dystansu genetycznego D_{sw} , stwierdzono grupowanie się rasy pc (PR) z rasą angler. Wyższe wartości bootstrapu w drzewie D_A , aniżeli w drzewie D_{sw} wskazują na większą wiarygodność struktury genealogicznej ras ustalonej na podstawie dystansu D_A . Natomiast dystans D_{sw}