Genetic diversity of local cattle*

Veronika Kukučková[1]✉, Nina Moravčíková[1], Ino Curik[2], Mojca Simčič[3], Gábor Mészáros[4] and Radovan Kasarda[1]

[1]Slovak University of Agriculture in Nitra, Faculty of Agrobiology and Food Resources, Department of Animal Genetics and Breeding Biology, Nitra, Slovak Republic; [2]University of Zagreb, Department of Animal Science, Zagreb, Croatia; [3]University of Ljubljana, Department of Animal Science, Ljubljana, Slovenia; [4]University of Natural Resources and Life Sciences, Division of Livestock Sciences, Vienna, Austria

The Slovak Pinzgau breed faces the bottleneck effect and the loss of diversity due to unequal use of founders and a significant population decline. Further population size reduction can lead to serious problems. Information obtained here and in other studies from high-throughput genotyping of 179 individuals was used to characterise genetic diversity and differentiation of Slovak Pinzgau, Austrian Pinzgau, Cika and Piedmontese cattle by Bayesian clustering algorithm. A gene flow network for the clusters estimated from admixture results was produced. The low estimate of genetic differentiation (F_{ST}) in Pinzgau cattle populations indicated that differentiation among these populations is low, particularly owing to a common historical origin and high gene flow. Changes in the log marginal likelihood indicated Austrian Pinzgau as the most similar breed to Slovak Pinzgau. All populations except the Piedmontese one displayed two ways of gene flow among populations, indicating that Piedmontese cattle was involved in producing of the analysed breeds while these breeds were not involved in creation of Piedmontese. Genetic evaluation represents an important tool in breeding and cattle selection. It is more strategically important than ever to preserve as much of the livestock diversity as possible, to ensure a prompt and proper response to the needs of future generations. Information provided by the fine-scale genetic characterization of this study clearly shows that there is a difference in genetic composition of Slovak and Austrian populations, as well as the Cika and Piedmontese cattle. Despite its population size, the Slovak Pinzgau cattle have a potential to serve as a basic gene reserve of this breed, with European and world-wide importance.

Key words: admixture, clustering, gene flow, genotyping array, Pinzgau cattle

Received: 07 November, 2017; revised: 01 March, 2018; accepted: 09 June, 2018; available on-line: 27 August, 2018

INTRODUCTION

Threats to biodiversity in terms of extinction rate, destruction of ecosystems and habitat, or loss of genetic diversity are increasing within livestock. The current concept of breed was formulated in the mid-1800s, followed up by intense artificial selection leading to striking and large-scale changes in the genetic makeup of individuals and the production systems. In this process, the farmers progressively substituted the less productive, locally adapted, native breeds with highly productive cosmopolitan breeds, abandoning marginal production areas. As a result of this process, a significant number of cattle, sheep, and goat breeds has already disappeared and many became endangered (Taberlet *et al.*, 2008).

The genetic structure of natural populations is influenced by a limited gene flow occurring when geographical distances increase between them. Non-random mating is a result of the preferential mating between geographically close individuals; this means that the genetic distance between individuals or populations is proportional to their spatial distance ("isolation by distance"). The isolation by distance models are used to study demographic, migratory reproductive aspects of populations (Loiseau *et al.*, 2009). Gene flow in natural populations depends on ecological characteristics and global or local environment, while in the case of livestock it is mainly influenced by human activities (Allendorf *et al.*, 2010).

In Central Europe, most breeds were of triple purpose. In remote Alpine valleys, productivity was improved only in the late 19th century by better feeding and management. Red-pied breeds such as the Dutch Meuse-Rhine-Yssel were more developed as dual purpose cattle, but in the late 20th century were influenced by red-pied Holstein sires. Successful breeds had spread beyond their region of origin and were even exported to other countries or continents. In contrast, locally adapted but less productive breeds declined in number or disappeared. Genetic isolation of breeds decreased the diversity at the molecular level, which can be monitored via an increase in homozygosity (Felius *et al.*, 2014).

The Slovak Pinzgau breed is endangered cattle in which selection with intense and highly unbalanced use of superior bulls magnified the loss of genetic diversity. The genetic gain increases along with the inbreeding coefficient, which may cause problems due to its small population size. Change of breeding goal to preserve the dual purpose character of Slovak Pinzgau is proposed, with an expected positive impact on the population structure (Kasarda *et al.*, 2014). Polymorphism of genes encoding leptin, leptin receptor (Trakovická *et al.*, 2013) and five others candidate genes associated with production parameters (Miluchová *et al.*, 2014) was evaluated in the Slovak Pinzgau cattle, with the determination of genotypic population structure using molecular genetic methods. The use of most recent results of quantitative and molecular genetics should help to sustainably manage population. With application of genomic selection,

a routine collection of genomic information provides an invaluable tool for effective management of breeding programs in small, endangered populations (Mészáros *et al.*, 2015).

The main objective of this study was to provide a detailed insight into genomic characterisation of the Pinzgau cattle (Slovak and Austrian Pinzgau) by utilizing high-throughput molecular information. Bayesian clustering algorithms were used to detect gene flow among the Pinzgau breeds, Cika and Piedmontese cattle.

MATERIALS AND METHODS

The analysed dataset consists of 179 cattle genotypes (19 Slovak Pinzgau, 105 Austrian Pinzgau, 26 Cika, and 29 Piedmontese). Slovak and Austrian Pinzgau cattle are dual purpose Alpine cattle breeds suitable for mountain regions, with good longevity, fertility and adaptability to poor production conditions. Recently, the Slovak Pinzgau cattle population size has declined due to the crossing and replacement with highly intensive breeds, e.g. Red Holstein-Friesian cattle.

The following sources of information were used: all active Slovak Pinzgau (SP) breeding bulls were genotyped in this study with the Illumina BovineSNP50 v2 BeadChip (Illumina Inc., San Diego, CA); genotyping information (BovineSNP50 v1 BeadChip) for Austrian Pinzgau (AP) bulls was obtained from Ferenčaković *et al.* (2013); (BovineSNP50 v2 BeadChip) data for Cika (CK) bulls was provided in Simčič *et al.* (2015); and (BovineSNP50 v1 BeadChip) information for Piedmontese (PI) breed was retrieved from the Dryad Digital Repository (http://dx.doi.org/10.5061/dryad.42tr0.2; McTavish *et al.*, 2013b).

The data sets were merged and quality controlled, SNPs with average heterozygosity >0.5, with call rate <0.8, with significant deviation from Hardy-Weinberg equilibrium at $P<10^{-5}$ were removed, following McTavish and coworkers (McTavish *et al.*, 2013a). All individuals with >10% missing SNPs placed at 29 autosomal chromosomes and all SNPs that were missing in >10% of individuals were removed. Minor allele frequency threshold of 0.05 was imposed in all calculations. The LD-based pruning was performed to remove SNPs that were in linkage disequilibrium (LD) with one another by using PLINK's (Purcell *et al.*, 2007) simple pairwise threshold model (command: --indep-pairwise 50 5 0.05). This removed 25,433 SNPs that exceeded LD threshold of 0.05, after which 9,433 SNPs remained for further analyses.

Bayesian approach using BAPS version 6.0 (Corander *et al.*, 2008) was used to find admixture and to produce gene flow network. The optimal number of clusters was directly inferred by the implemented algorithm. The maximum number of clusters was set to 5, as recommended in case of higher than expected number of populations. An admixture analysis conditional on the optimal genetic mixture estimate from the individual level analysis was performed. Results were based on 1000 simulations from the posterior allele frequencies. The number of clusters containing more than 19 individuals as a point estimate of K was used. Furthermore, to assess the significance of the admixture estimates, 200 individuals were generated from each identified ancestral source (initial breed) to provide an approximation to the distribution of the estimates under the hypothesis of no admixture. Number of iterations for the reference individuals was set to 10. A gene flow network for the clusters estimated from admixture results was produced (Tang *et al.*, 2009).

RESULTS AND DISCUSSION

As a starting point of this study, previous analyses, partly mentioned by Kukučková and coworkers (Kukučková *et al.*, 2017), were used. The degree of the genetic differentiation between analysed cattle populations (SP, AP) was initially investigated. The level of F_{ST} indicating differences between SP and AP was 0.01. The authors had found a low level of genetic differentiation between CK and PI as well (0.04), thus a more thorough study was needed. The similarity of 4 evaluated populations is shown in more detail in Fig. 1. Parameter estimation in metapopulations is more valuable and more accurate (than in isolated populations with no gene flow), and facilitate investigations of the relative importance as well as interactions among drift, selection and migration in populations of conservation concern (Allendorf *et al.*, 2010).

Figure 1B presents 15 exact clusters for the metapopulations of 15 European breeds (K=15) using unsupervised clustering without prior information about populations, based on 1000 simulations from the posterior allele frequencies. On the contrary, Fig. 1A visualised results based on a smaller number of simulations, with the resulting number of clusters K=13. The Pinzgau populations are in one cluster (coloured yellow), while the CK and PI are grouped in another cluster, coloured dark blue. The influence of the Pinzgau cattle in formation of the CK population is also apparent.

According to Felius and coworkers (Felius *et al.*, 2014), the Piedmontese beef breed belongs to the Central European cattle but does not belong to a breed cluster of the spotted dairy cattle from the eastern Alps (Cika and Pinzgau). By using a Bayesian approach, McTavish and coworkers (McTavish *et al.*, 2013) have confirmed the hypothesis that the New World cattle (as well as modern breeds from Southern Europe, e.g. Piedmontese cattle) are influenced by an ancestral gene flow from northern Africa.

To determine the degree of genetic differentiation among the four studied local populations, the SNP chip data were analysed here using a genetic mixture model for unlinked markers. Using BAPS, changes in the log marginal likelihood based on posterior mode estimates of the admixture coefficients (set as significant at the 5% level) are shown in Fig. 2. There are 4 estimated populations (clusters) and cluster 3 was chosen as the 'source'. According to Tang and coworkers (Tang *et al.*, 2009), changes in the log marginal likelihood close to zero indicate that the mixture model judges both assignments (source cluster and target cluster) to be statistically reasonable for an individual. In contrast, values further away from zero indicate that the relative genetic affinity of the clusters has decreased. In Fig. 2, it is apparent that AP is genetically much closer to SP, than CK and PI. The density curve for AP is shifted more towards zero than the other populations, hence it implies a closer relationship to SP. The low estimate of genetic differentiation between SP and AP indicated that differentiation among these populations is low, particularly due to a common historical origin and high gene flow. Analyses of Allendorf (Allendorf, 2010) assume that genetic distance provides an estimation of genetic connectivity, by reflecting the degree to which gene flow and drift affect genetic divergence among populations.

The assignment of individuals from high-throughput genotyping data was achieved using Bayesian inference, confirming that it is possible to distinguish among the
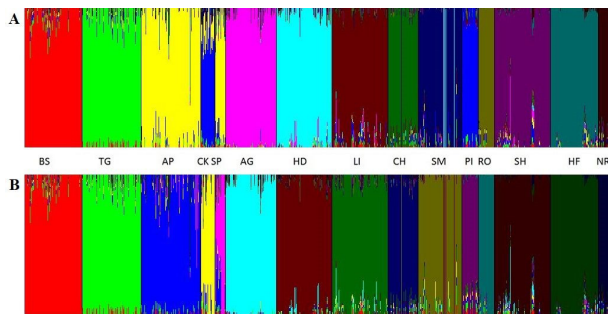
Figure 1. Unsupervised hierarchical clustering of the 1040 individuals belonging to 15 breeds.
The optimal number of clusters was directly inferred by the implemented algorithm. BS, Brown Swiss; TG, Tyrol Grey; AP, Austrian Pinzgau; CK, Cika; SP, Slovak Pinzgau; AG, Angus; HD, Hereford; LI, Limousine; CH, Charolais; SM, Simmental; PI, Piedmontese; RO, Romagnola; SH, Shorthorn; HF, Holstein; NR, Norwegian Red. (**A**) Results were based on 100 simulations from the posterior allele frequencies. (**B**) Results were based on 1000 simulations from the posterior allele frequencies.
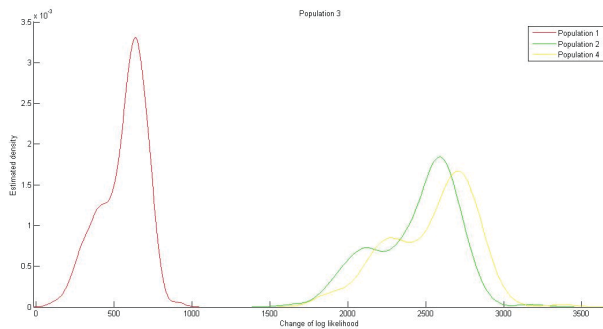

Figure 2. Changes in the log marginal likelihood for the four evaluated populations (1=Austrian Pinzgau, 2=Cika, 3=Slovak Pinzgau, 4=Piedmontese) with Slovak Pinzgau as a source cluster.
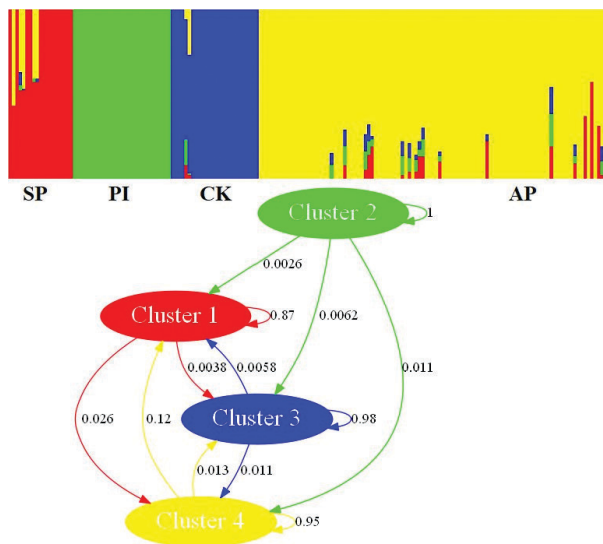

Figure 3. Admixture analysis in the population of Slovak Pinzgau (SP – Cluster 1), Piedmontese (PI – Cluster 2), Cika (CK – Cluster 3) and Austrian Pinzgau (AP – Cluster 4).
(**Top**) Each column represents a single individual and is coloured according to the proportion of genetic information assigned to each population. The final cluster assignment is shown by the colour underneath. (**Bottom**) A network of clusters indicating gene flow by weighted posterior admixture estimates for evaluated breeds. Direction of the gene flow is represented by arrows.

breeds by using SNPs and that the breeds are genetically distinct (Fig. 3 Top). To investigate the ancestral admixture of a certain population, one can look at the arrows pointing at this population. A typical population contains major sources of its own gene pool, denoted as a self-looping arrow, and small proportions of gene flow from other populations (Tang *et al.*, 2009). The actual rates of admixture represented by the arrows and corresponding numbers were randomly generated from a uniform distribution (Fig. 3 Bottom). Two ways of gene flow between SP and CK, SP and AP, CK and AP were found. The PI population contributed to each of the analysed populations. This was rather surprising as all populations except PI indicated two ways of gene flow, which means that PI was involved in formation of the analysed breeds, while these breeds were not involved in creation of PI. In numbers, 87% of SP is unique for the breed, while 12% of its genetic information comes from AP, 0.58% from Cika and 0.26% from Piedmontese cattle.

Gene flow in our study is presented as a proportion of genetic information identical for a pair of breeds, while Achmann and coworkers (Achmann *et al.*, 2004) estimated gene flow as a number of migrant horses per generation from a subpopulation and to a subpopulation. Furthermore, the pool of genotypes for the entire breed enables to continuously scan the population and allow a swift reaction in identifying carriers of lethal or potentially harmful haplotypes. The new information can be used to eliminate undesirable alleles through the mating process. Similarly, the breed proportions due to admixture could be estimated with the goal to fix a desirable ratio or to preserve the purity of the breed (Mészáros *et al.*, 2015).

Although the diversity of the current cosmopolitan cattle is still large enough to belie the claim that cattle becomes an endangered species, the loss of local breeds does erode genetic resources that are difficult to replace. Furthermore, these breeds belong to our cultural heritage and are of local cultural importance, even if most of the breeds are only one or two centuries old (Felius *et al.*, 2014).

Nielsen and coworkers (Nielsen *et al.*, 2009) found that adaptive population divergence may be prevalent despite seemingly high levels of gene flow, as found in most marine fishes. Their results have important implications for understanding of the interplay of evolutionary forces in general, and for the conservation of marine biodiversity under rapidly increasing evolutionary pressure from climate and fisheries induced changes in the local environments.

CONCLUSION

We demonstrated the power of using molecular markers obtained by high-throughput genotyping in clustering cattle breeds, even in case of a very weak differentiation, as between AP and SP (F_{ST}=0.012). Changes in the log marginal likelihood also indicated AP as the most similar breed to SP. All populations except the Piedmontese breed, showed two ways of gene flow among populations, which means that PI was involved in formation of the analysed breeds while these breeds were not involved in creation of PI. The obtained genetic data shed light on some issues related to the local Pinzgau breeds' origin and structure. Our study has shown that the Slovak, as well as the Austrian cattle, are important and viable targets for conservation because they display special traits of phenotypic, cultural and historical na-

ture. It is more strategically important than ever to preserve as much of the livestock diversity as possible, to ensure a prompt and proper response to the needs of future generations. In this context, approaches based on the combination of genomics and spatial analyses are of great help.

Acknowledgements of Finnacial Support

REFERENCES

Achmann R, Curik I, Dovc P, Kavar T, Bodo I, Habe F, Marti E, Sölkner J, Brem G (2004) Microsatellite diversity, population subdivision and gene flow in the Lipizzan horse. *Anim Genet* **35**: 285–292. http://dx.doi.org/10.1111/j.1365-2052.2004.01157.x

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Genetics* **11**: 697–709. http://dx.doi.org/10.1038/nrg2844

Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9**: 539. http://dx.doi.org/10.1186/1471-2105-9-539

Felius M, Beerling M-L, Buchanan DS, Theunissen B, Koolmees PA, Lenstra JA (2014) On the history of cattle genetic resources. *Diversity* **6**: 705–750. http://dx.doi.org/10.3390/d6040705

Ferenčaković M, Sölkner J, Curik I (2013) Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genet Sel Evol* **45**: 42. http://dx.doi.org/10.1186/1297-9686-45-42

Kasarda R, Mészáros G, Kadleček O, Hazuchová E, Šidlová V, Pavlík I (2014) Influence of mating systems and selection intensity on the extent of inbreeding and genetic gain in the Slovak Pinzgau cattle. *Czech J Anim Sci* **59**: 219–226

Kukučková V, Moravčíková N, Ferenčaković M, Simčič M, Mészáros G, Sölkner J, Trakovická A, Kadleček O, Curik I, Kasarda R (2017) Genomic characterization of Pinzgau cattle: genetic conservation and breeding perspectives. *Conserv Genet* **18**: 893–910. http://dx.doi.org/10.1007/s10592-017-0935-9.

Loiseau L, Richard M, Garnier S, Chastel O, Julliard R, Zoorob R, Sorci G (2009) Diversifying selection on MHC class I in the house

sparrow (*Passer domesticus*). *Mol Ecol* **18**: 1331–1340. http://dx.doi.org/10.1111/j.1365-294X.2009.04105.x

McTavish EJ, Decker JE, Schnabel RD, Taylor JF, Hillis DM (2013a) New World cattle show ancestry from multiple independent domestication events. *Proc Natl Acad Sci U S A* **110**: E1398-E1406. http://dx.doi.org/10.1073/pnas.1303367110

McTavish EJ, Decker JE, Schnabel RD, Taylor JF, Hillis DM (2013b) New World cattle show ancestry from multiple independent domestication events. *Dryad Digital Repository*. http://dx.doi.org/10.5061/dryad.42tr0.2

Mészáros G, Boison SA, Pérez O'Brien AM, Ferenčaković M, Curik I, da Silva MVB, Utsunomiya YT, Garcia JF, Sölkner J (2015) Genomic analysis for managing small and endangered populations: A case study in Tyrol Grey cattle. *Front Genet* **6**: 173. http://dx.doi.org/10.3389/fgene.2015.00173

Miluchová M, Gábor M, Trakovická A (2014) Analysis of genetic structure in Slovak Pinzgau cattle using five candidate genes related to milk production traits. *Genetika* **46**: 865–875. http://dx.doi.org/10.2298/GENSR1403865M

Nielsen EE, Hemmer-Hansen J, Polsen NA, Loeschcke V, Moen T, Johansen T, Mittelholzer Ch, Taranger G-L, Ogden R, Carvalho GR (2009) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evol Biol* **9**: 276. http://dx.doi.org/10.1186/1471-2148-9-276

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ., Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet* **81**: 559–575. http://dx.doi.org/10.1086/519795

Simčič M, Smetko A, Sölkner J, Seichter D, Gorjanc G, Kompan D, Medugorac I (2015) Recovery of native genetic background in admixed populations using haplotypes, phenotypes, and pedigree information – using Cika cattle as a case breed. *PLoS ONE* **10**: e123253. http://dx.doi.org/10.1371/journal.pone.0123253

Taberlet P, Valentini A, Rezaei HR, Naderi S, Pompanon F, Negrini R, AjmoneMarsan P (2008) Are cattle, sheep, and goats endangered species? *Mol Ecol* **17**: 275–284. http://dx.doi.org/10.1111/j.1365-294X.2007.03475.x

Tang J, Hanage WP, Fraser Ch, Corander J (2009) Identifying currents in the gene pool for bacterial populations using an integrative approach. *PLoS Comput Biol* **5**: 8: e1000455. http://dx.doi.org/10.1371/journal.pcbi.1000455

Trakovická A, Moravčíková N, Kasarda R (2013) Genetic polymorphisms of leptin and leptin receptor genes in relation with production and reproduction traits in cattle. *Acta Biochim Pol* **60**: 783–787