

## Assessment of genetic drift and migration in six cattle breeds

Kristína Lehocká\*, Radovan Kasarda, Barbora Olšanská, Nina Moravčíková

*Slovak University of Agriculture in Nitra, Faculty of Agrobiolgy and Food Resources, Department of Animal Genetics and Breeding Biology, Nitra, Slovakia*



Licensed under a Creative Commons Attribution 4.0 International License



Slovak Spotted cattle represents an endangered breed with cultural importance in Slovakia. The study was based on the panel of 34,604 SNPs that were used for genotyping of 451 individuals. We used a combination of two arrays Illumina BovineSNP50v2 BeadChip and ICBF International Dairy and Beef v3, for estimation of gene flow and genetic drift. Based on the admixture results, a gene flow network across the analysed breeds was created. Our result showed that the Jersey population was involved in the grading-up of the analysed breeds. Analysed breeds were not confirmed to influence genetic make-up of Jersey. In addition, the phylogenetic analysis of the six cattle breeds revealed that Jersey is separated from the others. In contrast, the other breeds showed a close relationship with each other according to the maximum-likelihood tree. Migration edges reached weight values below 0.2, apart the one observed among the Ayrshire/Swiss Simmental breeds into Jersey (0.4), reflecting that the donor population has made a significant contribution to the recipient population.

**Keywords:** Bayesian Population Structure Analysis, genetic drift, gene flow, Slovak Spotted, TreeMix

### 1 Introduction

The genetic diversity within and between breeds is derived from their origin, history and development (Troy et al., 2001; Kasprzak-Filipek et al., 2019). An essential task in biodiversity conservation is to protect native domestic breeds of cattle, which represent a reservoir of unique combinations of genes and alleles (Bulla et al., 2013). The present concept of the breeds was formulated by historical events such as fluctuations in population size and selection processes. In particular, intensive artificial selection performed in the last century has resulted in significant changes in the genetic composition of individuals and production systems (Kukučková et al., 2017). Genetic drift is an evolutionary process, which is defined as a change in the frequency of an already existing gene variant in a population due to random sampling of organisms (Masel, 2011). Genetic drift can act in a variety of ways, until the complete disappearance of some gene variants from a population, resulting in a reduction in overall genetic variation. Several authors defined genetic drift as a random change in allele frequencies over generations due to the final population size (Wright, 1929; Falconer and Mackay, 1996; Merilä, 2014).

In natural populations, gene flow and species distribution are key processes that affect demographics and the development of the structure of populations. The genetic structure, which represents the composition of animals in a population, can be strongly influenced by external factors such as genetic drift, gene flow and in some cases natural selection (Banks et al., 2013). The genetic structure of populations is affected by the limited flow of genes that occurs as geographical distances increase among them (Sexton et al., 2014). The genetic structure can be characterised by markers analysis, and in recent years, the use of single nucleotide polymorphism (SNP) has become prevalent in this research (Groeneveld et al., 2010). Studies indicate that SNP analyses explain bovine history more accurately than microsatellite sequence analysis and are therefore more commonly used in biodiversity conservation (Gautier et al., 2007; McKay et al., 2008; Socol et al., 2015). The aim of the

\* **Corresponding Author:** Kristína Lehocká, Slovak University of Agriculture in Nitra, Faculty of Agrobiolgy and Food Resources, Department of Animal Genetics and Breeding Biology, Tr. A. Hlinku 2, 94976 Nitra, Slovakia. E-mail: tina.lehocka@gmail.com

present study was to investigate genetic diversity and assess the genetic drift and gene flow within and among cattle breeds.

## 2 Material and methods

The SNP database was composed by combining new data and data from public repositories. The dataset was composed of 451 animals consisting of Slovak spotted cattle (SS = 85) along with historically-related breeds; Holstein (HOL = 99), Swiss Simmental (SIM = 78), Slovak Pinzgau (PIN = 151), Jersey (JER = 28) and Ayrshire (AYR = 10). Over the last 50 years, there has been a significant decrease in the number of autochthonous breeds in Slovakia (Chrenek et al., 2017). For this reason, protected herds of Slovak spotted and Slovak Pinzgau cattle were established. The main goal is to preserve these breeds by purebred breeding in situ conditions (FAO, 2003).

Traditional Slovak breeds (SS and PIN) were genotyped by two platforms - Illumina BovineSNP50v2 BeadChip and ICBF International Dairy and Beef v3. The Jersey were genotyped by GeneSeek GGP 150K genotyping platform. While genotypic data of other breeds (HOL, SIM and AYR) were obtained using web-based data archive (McTavish et al., 2013).

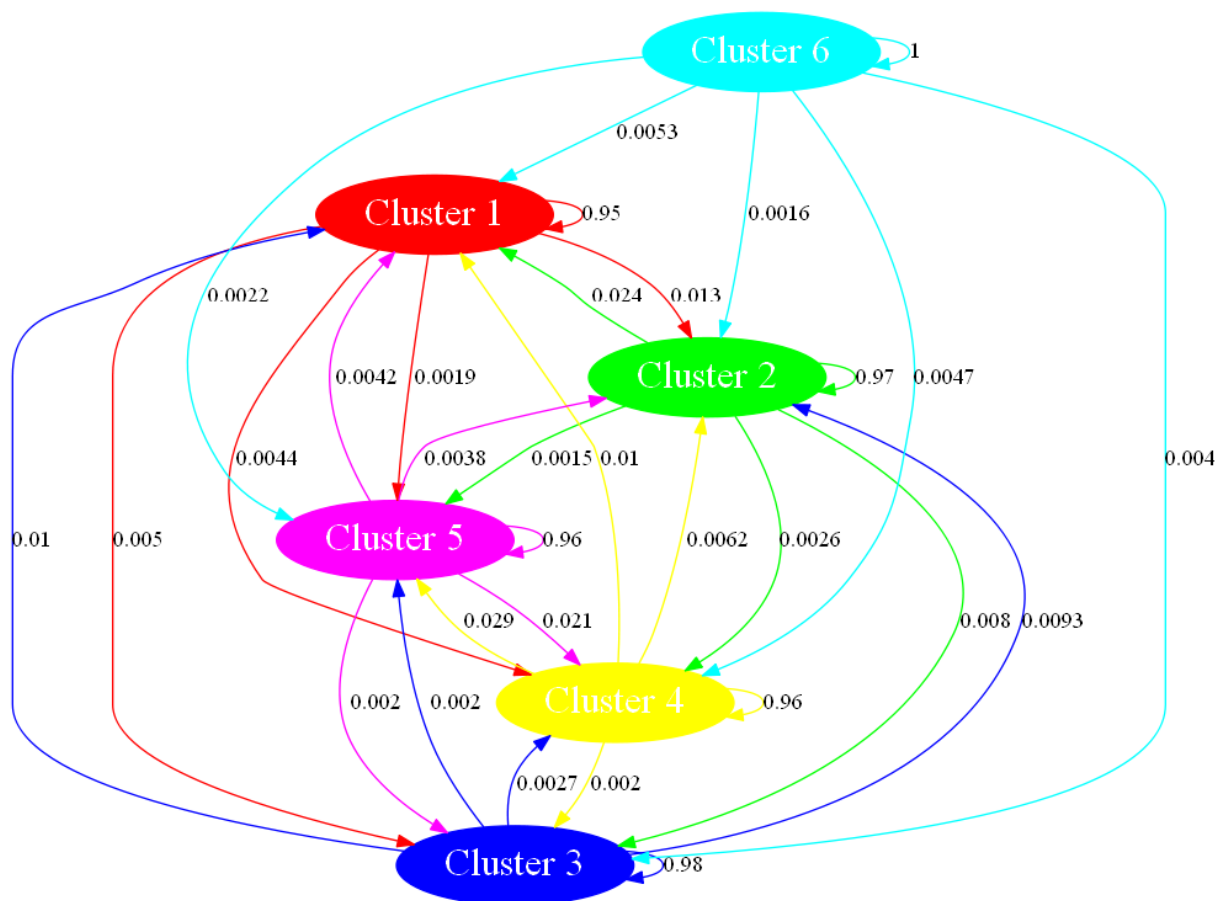
Standard quality control of SNPs was performed using PLINK v1.9 (Chang et al., 2015), according to Moravčíková et al. (2018). The dataset consisted of 34,604 SNP markers. Some SNPs were pruned due to high linkage disequilibrium (LD) in software PLINK v1.9. Kijas et al. (2009) reported that pruning of SNPs with high LD acts against the effect of bias in the detection and then expresses a meaningful comparison of breeds. A total of 5,584 SNPs remained for further analysis. The population structure of the analysed breeds was evaluated using Bayesian Population Structure Analysis (BAPS) version 6.0 (Corander and Tang, 2007). Our results were based on the number of cluster K 10-50 and 1,000 simulations from the posterior allele frequencies, according to Kukučková et al. (2018). The TreeMix software (Pickrell and Pritchard, 2012) was used to recognise genetic drift and migration between the analysed populations. Subsequently, the consistency of migration events was evaluated (10x separate runs) and then migration edges (-m set to max 8) were added to the generated graph (Upadhyay et al., 2019). The results were visualised using R software (R Core Team, 2014).

## 3 Results and discussion

Figure 1 shows the degree of gene flow identified across the six evaluated populations. The evaluation of the admixture of a particular population is shown by an arrow pointing to that population. A typical population contains its primary sources (the population itself), which are indicated by a looping arrow and consequently contains small proportions of gene flow from other populations (Tang et al., 2009). The observed level of admixture was expressed in the graph as arrows between breeds, with the corresponding numbers representing migration rate for the analysed population (Kukučková et al., 2018). Four modes of gene flow between SIM, HOL, AYR, PIN and SS populations (cluster 1 – 5) were found. Gene flow between clusters shows that clusters from 1 to 5 represent sources of migrants but also receive gene flow.

The exception was the JER population (cluster 6), which contributed to each population analysed. But cluster 6 was not affected by other breeds, because did not show four modes of gene flow. The cluster 6 seems to participate in the development history of the analysed breeds, but not vice versa. In the case of cluster 5 (SS), it reached a high proportion of genotypes that have same sequence signature s from other clusters, indicating extensive migration events between these populations. The SS has 97% of its genetic make-up, but only 3% of the DNA introduced was obtained through gene flow from other populations. The two major sources of gene flow for cluster 5 come from the clusters 3 (AYR) and 6 (JER), with a contribution of 3% and 0.44%, respectively. The remaining 0.45% of genes come from the other clusters, but neither of them participated with more than 0.4%. The results agree with the historical formation of the Slovak spotted breed in Slovakia.

Slovak spotted and PIN cattle come from the common already extinct ancestors of the Carpathian red and grey cattle. In the 18th century, the SIM participated in the formation of the breed, as stated by Kasarda et al. (2015). After 1972, other breeds such as HOL, JER and AYR participated in the formation of the breed to increase milk production (Association of Slovak Spotted cattle breeders - Cooperative, 2020).

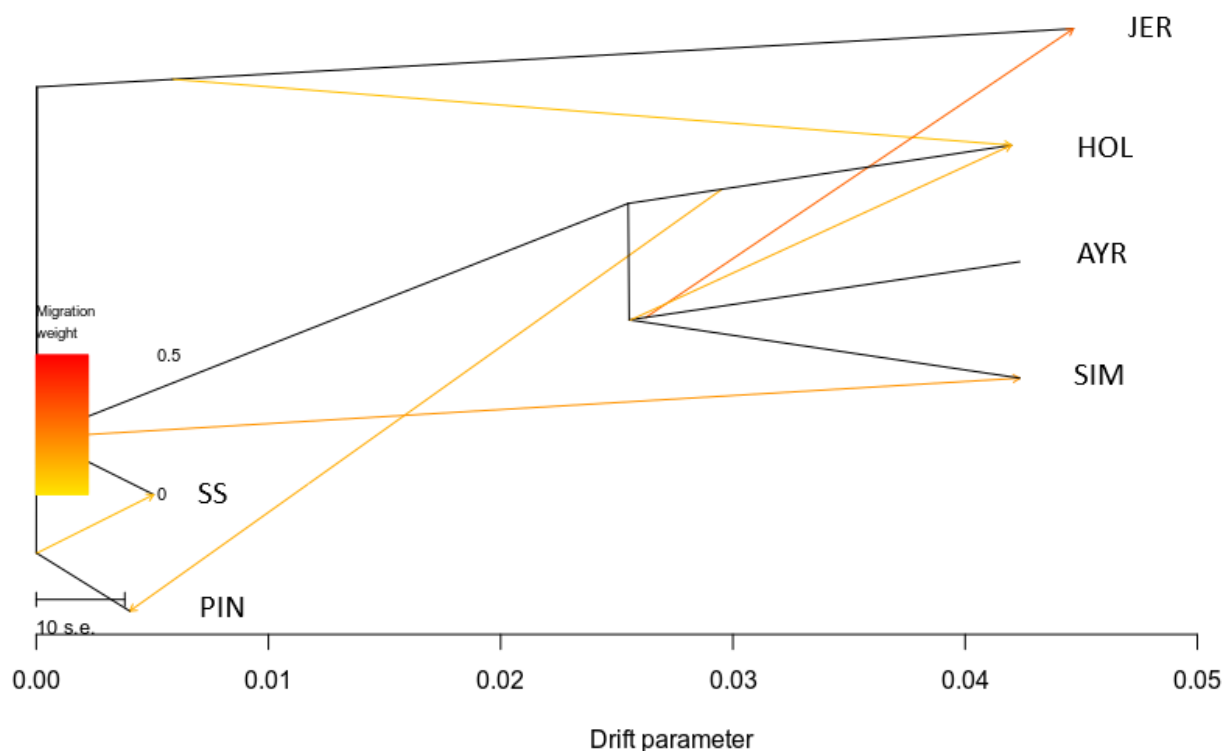


**Figure 1** Tentative gene flow graph in evaluated populations (Swiss Simmental – Cluster 1, Holstein – Cluster 2, Ayrshire – Cluster 3, Slovak Pinzgau – Cluster 4, Slovak Spotted – Cluster 5, Jersey – Cluster 6)

The maximum-likelihood (ML) phylogenetic ancestry graph was generated to examine genetic differentiation and gene flow among evaluated cattle breeds. For our dataset, the ML tree topology generated a graph on which populations are separated into two groups (JER and other breeds). Figure 2 shows the maximum-likelihood tree for the six evaluated cattle populations and expresses the migration arrows, which are coloured according to their weight and degree of admixture among breeds.

The amount of gene flow between the evaluated populations is proportionally expressed by a scale on the horizontal axis and by the length of the horizontal branches. The horizontal axis scale was expressed as 10x the average standard error of the covariance matrix between populations based on allele frequencies (Karimi et al., 2016).

All breeds except SS and PIN had long branches in the ML graph, which indicates that these breeds were carried away similarly, as reported by Rochus et al. (2020) in Swedish sheep. A medium-weight migration edge connecting to the SIM cattle branch comes from SS breed. It is also interesting that there is a high migration border between the roots of the AYR and SIM and JER populations. Compared with results of Browett et al. (2017), our findings point the hypothesis of historical gene flow from the SIM and AYR populations into the ancestral population of modern JER cattle. All migrations edges showed a lower weight below 0.2, the only exception being the migration from AYR/SIM to JER, which weighted 0.4, pointing the donor population made a significant genetic contribution to the consignee breed, similarly to that stated by Orozco-terWengel et al. (2015). It was found that between the evaluated populations, the Slovak breeds showed the lowest degree of divergence, similar to what reported by Jemaa et al. (2015) for Tunisian local cattle.



**Figure 2** The maximum-likelihood (ML) tree expressing the inferred relationship between analysed cattle populations (JER= Jersey, HOL= Holstein, AYR= Ayrshire, SIM = Swiss Simmental, SS= Slovak Spotted, PIN= Slovak Pinzgau)

#### 4 Conclusions

Many breeds of cattle are considered endangered, and the erosion of the existing genetic variability is leading to irretrievably losing these populations. For this reason, the assessment of diversity and genetic variability is crucial. Our results pointed out low to slightly genetic drift as well as high gene flow among analysed breeds. These findings reflected that JER was separated from all the other breeds. Ayrshire and SIM have resulted as highly connected breeds. The gene pool of SS was influenced by other breeds; this observation follows the historical formation of the SS breed. This study brought new information on the level of genetic drift and gene flow among evaluated breeds, which should be considered in conservation programs to preserve genetic diversity. Slovak Spotted represents a unique breed that would not be possible to restore to the form we know it today in case of extinction. A similar analysis will be applied in the future to beef breeds, in the formation of which SS participated.

#### Acknowledgments

The Slovak Research and Development Agency (APVV-14-0054 and APVV-17-0060) is gratefully acknowledged for supporting this study.

#### References

- ASSOCIATION OF SLOVAK SPOTTED CATTLE BREEDERS – COOPERATIVE (2020). HISTORY OF BREED ORIGIN, Year 2020. ASSOCIATION OF SLOVAK SPOTTED CATTLE BREEDERS – COOPERATIVE, Retrieved Jun 11, 2020 from <https://www.simmental.sk/o-plemene/historia-vzniku-plemena.html>
- Banks, S.C., Cary, G.J., Smith, A.L., Davies, I.D., Driscoll, D.A. Gill, A.M., Lindenmayer, D.B., Peakall, R. (2013). How does ecological disturbance influence genetic diversity? *Trends in Ecology & Evolution*, 28(11), 670–679. <https://doi.org/10.1016/j.tree.2013.08.005>
- Bulla, J., Polák, P., Chrenek, P. (2013). Pinzgauer cattle in Slovakia. *Slovak J. Anim. Sci.*, 46, 151–154
- Chang, Ch.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., Lee J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>

Chrenek, P., Kubovičová, E., Makarevich, A. (2017). CURRENT SITUATION IN THE GENE BANK OF ANIMAL GENETIC RESOURCES IN SLOVAKIA: A review. *Slovak J. Anim. Sci.*, 50, 2017 (4): 135–138.

Corander, J. and Tang, J. (2007). Bayesian analysis of population structure based on linked molecular information. *Math Biosci*, 205,19 –31. <http://dx.doi.org/10.1016/j.mbs.2006.09.015>.

Falconer, D. S. and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*, Ed 4. Longmans Green, Harlow, Essex, UK

FAO (2003). The State of Farm Animal Genetic Resources in the Slovak Republic. *National Consultative Committee for the Use and Conservation of Farm Animal Genetic Resources*. 29. Nitra. FAO. <http://www.fao.org/tempref/docrep/fao/010/a1250e/annexes/CountryReports/SlovakRepublic.pdf>

Gautier, M., Faraut, T., Moazami-Goudarzi, K., Navratil, V., Foglio, M., Grohs, C., Boland, C., Garnier J-G., Boichard, D., Lathrop, G.M., Gut, I.G., Eggen, A. (2007). Genetic and haplotypic structure in 14 European and African cattle breeds, *Genetics*, 177,2, 1059–1070. <https://doi.org/10.1534/genetics.107.075804>

Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R., Finlay, E.K., Jianlin, H., Groeneveld, E., Weigend, S. (2010). Genetic diversity in farm animals – a review. *Anim. Genet.*, 41 (Suppl.1),6–31. <https://doi.org/10.1111/j.1365-2052.2010.02038.x>

Jemaa, S.B., Boussaha, M., Mehdi, M.B., Lee, J.H., Lee, S.H. (2015). Genome-wide insights into population structure and genetic history of Tunisian local cattle using the Illumina bovinesnp50 bead-chip. *BMC genomics*, 16(1), 677.

Karimi, K., Strucken, E.M., Moghaddar, N., Ferdosi, M.H., Esmailzadeh, A., Gondro, C. (2016). Local and global patterns of admixture and population structure in Iranian native cattle. *BMC Genet* 17, 108. <https://doi.org/10.1186/s12863-016-0416-z>

Kasarda, R., Moravčíková, N., Trakovická, A., Mészáros, G., Kadlečík, O. (2015). GENOME-WIDE SELECTION SIGNATURES IN PINZGAU CATTLE. *Potravinárstvo*, 9 (1), 268–274. <https://doi.org/10.5219/478>

Kasprzak-Filipek, K. Sawicka-Zugaj, W., Litwinczuk, Z., Chabuz, W., Šveistienė, R., Bulla, J. (2019). Assessment of the genetic structure of Central European cattle breeds based on functional gene polymorphism, *Global Ecology and Conservation*, 17, e00525, ISSN 2351-9894

Kijas, J.W., Townley, D., Dalrymple, B.P., Heaton, M.P., Maddox, J.F., McGrath, A., Wilson, P., Ingersoll, R.G., McCulloch, R., McWilliam, S., Tang, D., McEwan, J., Cockett, N., Oddy, V.H., Nicholas, F.W., Raadsma, H. (2009). A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. *PLOS ONE*, 4. <https://doi.org/10.1371/journal.pone.0004668>

Kukučková, V., Moravčíková, N., Curik, I., Simčíč, M., Mészáros, G., Kasarda, R. (2018). Genetic diversity of local cattle. *Acta biochimica Polonica*,65(3). [https://doi.org/10.18388/abp.2017\\_2347](https://doi.org/10.18388/abp.2017_2347)

Kukučková, V., Moravčíková, N., Ferenčakovič, M., Simčíč, M., Mészáros, G., Sölkner, J., Trakovická, A., Kadlečík, O., Curik, I., Kasarda, R. (2017). Genomic characterisation of Pinzgau cattle: genetic conservation and breeding perspectives. *Conserv. Genet.*18, 893–910. <https://doi.org/10.1007/s10592-017-0935-9>

Masel, J. (2011). *Genetic drift*. Current Biology. Cell Press. 21 (20): R837-8. <https://doi.org/10.1016/j.cub.2011.08.007>

McKay, S.D., Schnabel, R.D., Murdoch, B.M., Matukumalli, L.K., Aerts, J., Coppieters, W., Crews, D., Neto, E.D., Gill, C.A., Gao, Ch., Mannen, Ch., Wang, Z., Van Tassell, C.P., Williams, J.L., Taylor, J.F., Moore, S.S. (2008). An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. *BMC Genet.*, 9, 37. <https://doi.org/10.1186/1471-2156-9-37>

McTavish, E.J., Decker, J.E., Schnabel, R.D., Taylor, J.F., Hillis, D.M. (2013). New world cattle show ancestry from multiple independent domestication events, *Proc Natl Acad Sci U S A*, 110(15):1398–1406. <https://doi.org/10.1073/pnas.1303367110>

Merilä, J. (2014). Lakes and ponds as model systems to study parallelel evolution. *J. Limnol.*,73,33–45.

Moravčíková, N., Kadlečík, O., Trakovická, A., Kasarda, R. (2018) Autozygosity island resulting from artificial selection in Slovak spotted cattle. *Agriculture & Forestry*, 64(4): 21-28.

Orozco-terWengel, P., Barbato, M., Nicolazzi, E., Biscarini, F., Milanese, M., Davies, W., Williams, D., Stella, A., Ajmone-Marsan, P., Bruford, M.W. (2015). Revisiting demographic processes in cattle with genome-wide population genetic analysis. *Front Genet.* 6,191. <https://doi.org/10.3389/fgene.2015.00191>

Pickrell, J.K. and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet*, 8: e1002967. <https://doi.org/10.1371/journal.pgen.1002967>

R CORE TEAM. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing. [Internet]. Vienna, Austria; 2013. [cited 2020 May 8] Available from: <http://www.R-project.org/>.

Rochus, C. M. and Johansson, A. M. (2017). Estimation of genetic diversity in Gute sheep: pedigree and microsatellite analyses of an ancient Swedish breed. *Hereditas*, 154,4. <https://doi.org/10.1186/s41065-017-0026-4>

Sexton, J.P., Hangartner, S.B., Hoffmann, A.A. (2014). Genetic isolation by environment or distance: which pattern of gene flow is most common? *Evolution* 68, 1–15. <https://doi.org/10.1111/evo.12258>

Socol, C.T., Iacob, L., Mihalca, I., Criste, F.L. (2015). Molecular and population genetics tools for farm animal genetic resources conservation: brief overview. *Anim. Sci. Biotechnol.*, 48, 95–102.

Troy, C. S., MacHugh, D.E., Bailey, J.F., Magee, D.A., Loftus, R.T., Cunningham, P., Chamberlain, A.T., Sykes, B.C., Bradley, D.G. et al (2001). Genetic evidence for Near-Eastern origins of European cattle. *Nature*, 410,1088–1109

Upadhyay M., Bortoluzzi C., Barbato M., Marsan P.A., Colli L., Ginja C., Sonstegard, T.S., Bosse, M., Lenstra, J.A., Groenen, M.A.M., Crooijmans, R.P.M.A. et al. (2019). Deciphering the patterns of genetic admixture and diversity in southern European cattle using Genome-wide SNPs. *Evol Appl.* John Wiley & Sons, Ltd (10.1111); 2019.

Wright, S. (1929). The evolution of dominance. *Am. Nat.* 63, 556–561.