COBISS: 1.08
Agris category code: L10

ESTIMATION OF INBREEDING IN SLOVENIAN BROWN-SWISS POPULATION

Jana OBŠTETER [1], Betka LOGAR [2]

ABSTRACT

Breeding programs require control of the level (F) and rate (ΔF) of inbreeding in order to avoid inbreeding depression. Increasing availability of genomic information has enabled a more accurate estimation of F and ΔF. This study aimed to investigate classical (Fped) and genomic inbreeding coefficients (FROH) in 214 genotyped Slovenian Brown-Swiss animals. Fped was obtained from pedigree analysis using PEDIG and FROH was estimated based on runs of homozygosity (ROHs) analysis using PLINK. The results show that year-averaged FROH exceeds year-averaged Fped for all the studied years. ΔFROH>1MB (0.00918) was ~3 times higher than ΔFped (0.00334) and was close to the suggested limit that still allows sustainable population management. While detected ROHs reveal more ancient as well as recent inbreeding, the majority reflects inbreeding dating back ~12 to ~3 generations. When stratified by ROH lengths, FROH reveals some differences in most highly inbred chromosomes according to shorter and longer ROHs suggesting some changes in selection goals during breed's history.

Key words: cattle, breeds, Brown-Swiss, inbreeding, pedigree, genotypes, runs of homozygosity, Slovenia

1 INTRODUCTION

Inbreeding is defined as the probability of the two alleles at a locus in an individual being identical by descend (IBD). Inbreeding coefficient is computed in respect to the base population which is assumed to be non-inbred and have inbreeding of zero (Falconer and Mackay, 1996). When analysing population's pedigree this condition is usually difficult to satisfy since the pedigree records are not always complete and include a limited definite number of generation. Especially with cattle, it is erroneous to assume that the first generation in the pedigree is unrelated since genetic material of the favourable bulls is widely and abundantly distributed (van der Werf, 1999). Consequently, the results could be spurious and inbreeding coefficient underestimated. Selection programs require a control of the level and the rate of inbreeding in the population, particularly in small populations, in order to maintain genetic diversity and prevent inbreeding depression that could lead to the reduction in mean fitness and production (Falconer and Mackay, 1996).

The increased use of genotyping arrays has allowed the development of new and more accurate methods to estimate inbreeding. Identification of long stretches of homozygous genotypes, i.e. runs of homozygosity (ROHs), enables detection of genome-wide autozygosity and IBD regions (McQuillan *et al.*, 2008). Therefore, longer homozygous tracts might provide evidence for genomic regions undergoing selection. It has also been denoted that genomic selection requires genomic control of inbreeding since they are both based on the same level (Sonesson *et al.*, 2012). Genomic selection for Brown-Swiss in Slovenia has started in 2009 by reason of participation in Interbull project InterGenomic (Santus, 2011). Since then 191 Slovenian bulls were added to the international reference population (Rigler *et al.*, 2016). The first genomic breeding values were predicted in 2013

1 Agricultural Institute of Slovenia, Hacquetova ulica 17, 1000 Ljubljana, Slovenia, e-mail: jana.obsteter@kis.si
2 Same address as 1, e-mail: betka.logar@kis.si

(Potočnik *et al.*, 2016). Now, all new breeding bulls are selected on the basis of genomic breeding values.

2 MATERIAL AND METHODS

A total of 214 Slovenian Brown-Swiss animals, 115 males and 99 females, born between 2003 and 2016, were genotyped on six different genotyping chips (Table 1). Genotyped animals served as a reference population for the construction of the pedigree. Final pedigree included 2653 animals, 880 males and 1773 females. Pedigree records were obtained from database in Cattle Information System (Logar *et al.*, 2005) which caters for most information requirements in the cattle breeding scheme in Slovenia. The quality of the pedigree was assessed with complete generation equivalent computed as a sum of $(1/2)^n$ terms over all known individual's ancestors, where n is the number of generations separating the individuals from the ancestor (Maignel *et al.*, 1996). Classical inbreeding coefficient (F) was computed based on pedigree information as defined by Meuwissen and Luo (1992) using PEDIG (Boichard, 2002) software:

$$A_{ii} = \sum_{j=1}^{i} L^2 D_{jj}$$

where A_{ii} = i^{th} diagonal element of the pedigree relationship matrix, L = lower triangel of A matrix, D = diagonal matrix. Generation intervals and pedigree completeness were determined using PopRep software (Groeneveld *et al.*, 2009).

Individuals genotyped on GeneSeek chips (n = 86) (Table 1) were imputed onto 50K Illumina chip utilising FIMPUTE (Stachowicz *et al.*, 2011) via ZANARDI (Nicolazzi and Marras, 2015) software. SNPs exclusive to each of the GeneSeek chips and sex chromosome SNPs were excluded prior to the imputation. The accuracy of imputation was assessed with 10x cross validation and allelic concordance levels were reported (Table 1). No genotype quality control was applied prior to SNP imputation since this was shown to be the best strategy (Roshyara *et al.*, 2014).

ROH analysis was performed on the 214 imputed genotypes. Genotype quality control was applied prior to the analysis with the following parameters: call rate per SNP > 90 %, MAF > 0.01 and deviation from Hardy-Weinberg equilibrium $p > 0.0001$, resulting in 214 genotyped animals and 42,302 remaining SNPs. The analysis was carried out using PLINK 1.9 (Purcell *et al.*, 2007) with adjusted parameters for the sliding window of 20 SNPs, allowing one heterozygous and two missing SNPs within the window, minimum SNP density one SNP every 120 kB and setting the minimum number of SNPs in a segment to be called a ROH to 15 and minimum length to 1 MB (adopted from Purfield *et al.*, 2012 and Ferenčaković *et al.*, 2013a). Individual F_{ROH} was computed as described in

$$F_{ROH} = \sum L_{ROH} \Big/ \sum L_{AUTO}$$

where ΣL_{ROH} is the cumulative length of individual's ROH and ΣL_{AUTO} is the total length of the genome SNP coverage (i.e. 2.51 GB). Pedigree (ΔF_{ped}) and genomic (ΔF_{ROH}) rate of inbreeding was computed by regressing the natural logarithm of F_{ped} and F_{ROH} onto the year of birth:

$$\ln(1 - F_Y) = \ln(1 - F_0) + \beta Y + e$$

$$\Delta F = (1 - e^{\beta}) L$$

where F_Y is year-averaged F | F_{ROH}, $\beta = \ln(1 - \Delta F_Y)$ and L is the average generation interval. Effective population size (N_e) was subsequently estimated as $N_e = 1/(2 * \Delta F)$.

The identified ROHs were classified into the following length classes as proposed by other studies (Ferenčaković *et al.*, 2013a): [1–2], (2–4], (4–8], (8–16] and > 16 Mb. F_{ROH} was computed at five different cut-

Table 1: *The number of animals genotyped*

Genotyping chip	Number of SNPs	Number of animals	Accuracy of imputation* onto Illumina 50Kv02
Illumina 50Kv02	54,609	128	-
GGP v02	19,720	6	94.0 %
GGP v03	26,151	43	96.9 %
GGP v04	30,105	22	95.6 %
GGP HD	76,883	4	98.7 %
GGP HDv02	138,892	11	97.0 %
			Average = 96.4 %

GGP = GeneSeek Genomic Profiler, v = version. *Accuracy of imputation is reported as allelic concordance.

offs for ROH length: >1 MB, >2 MB, >4 MB, >8 MB and >16 MB. $F_{ROH>1MB}$ and $F_{ROH>8MB}$ were used for comparison with other studies, since they represent F_{ROH} consisting of all identified ROHs and ROHs that are most likely to reflect true identity by descent, respectively.

3 RESULTS AND DISCUSSION

The constructed pedigree consisted of animals born between 1952 and 2016 with an average generation interval of 7.3 years. The mean of complete generation equivalents was 3.41 for the whole and 5.97 for the reference population. The pedigree completeness was above 90 % for up to four generation and dropped to 79.3 % for six generation pedigree depth. F_{ROH} was computed for animals born between 2003 and 2016 hence this period was used for F_{ped} and F_{ROH} comparison. There were 666 animals in the pedigree with unknown date of birth and were therefore used only for inbreeding computation but not for rate of inbreeding and effective population size calculation, since their F_{ped} could not be included in the calculation of the year averages.

Figure 1 shows the distribution of the F_{ped} (Fig. 1a) and F_{ROH} (Fig. 1b) for the reference animals. The mean values for inbreeding coefficients in this study were 0.0142 for $F_{ped,}$ 0.103 for $F_{ROH>1MB}$, 0.102 for $F_{ROH>2MB}$, 0.0898 for $F_{ROH>8MB}$ and 0.0607 for $F_{ROH>16MB}$ (Fig. 1c). Therefore the highest genomic inbreeding coefficient was observed at ROH lengths >1 MB. This was expected since $F_{ROH>1MB}$ and was computed based on all identified ROHs and therefore captures ancient as well as recent inbreeding. All year-averaged F_{ROH} exceeded year-averaged F_{ped} for all of the considered years (Fig. 1c), $F_{ROH>1MB}$ exceeded F_{ped} ~5 times. Both, year-averaged F_{ped} and all F_{ROH} were lowest in 2003 (F_{ped} = 0.00245, $F_{ROH>1MB}$ = 0.0680). Conversely, F_{ped} was highest in 2005 (0.0367) and all F_{ROH} in 2010 ($F_{ROH>1MB}$ = 0.130). Other studies investigating inbreeding in Brown-Swiss reported values of $F_{ROH>1MB}$ = 0.156, $F_{ROH>8MB}$ = 0.074 and F_{ped} = 0.048, which is slightly higher than in our study. The highest correlation of F_{ped} and F_{ROH} in this study was observed for $F_{ROH>16MB}$ (0.464) and it was significantly different from 0 (1.533e-11). Ferenčakovič *et al.* (2013a) observed the highest correlation of F_{ped} with $F_{ROH>1MB}$ for Brown-Swiss (0.660). Studies investigating other cattle breeds reported similar or even weaker correlation and larger discrepancies between F_{ped} and $F_{ROH.}$ Study from Hillestat *et al.* (2015) observed ~5 times larger F_{ROH} values compared to F_{ped} for Norwegian Red Cattle and Gurgul *et al.* (2016) reported as large as 10 times larger F_{ROH} values for Holstein. Studies investigating other cattle breeds reported values of $F_{ROH>1MB}$ and $F_{ROH>8MB}$ ~0.090 and ~0.020 for Simmental, 0.088 and 0.035 for Fleckvieh, ~0.080 and ~0.0250 for Pinzgau, 0.048 and 0.0140 for Nellore cattle. This illustrates that inbreeding levels differ between breeds and even within the same breed depending on the local population's demography.

Estimates of ΔF from pedigree and genotype analysis (ROH > 1MB) were 0.00334 (N_e = 149.5) and 0.00918 (N_e = 54.4). It has been pointed out that managing inbreeding rate is more important than managing inbreeding level in a population. While ΔF of 0.025 is considered to be high risk for a population (N_e = 20), ΔF of 0.01 (N_e = 50) was suggested to be sufficient for sustainable management of a population. However, sometimes lower rate are desirable (Woolliams *et al.*, 1998). In this study both ΔF_{ped} and ΔF_{ROH} are within the acceptable limits of ΔF. However, although ΔF_{ped} according to pedi-

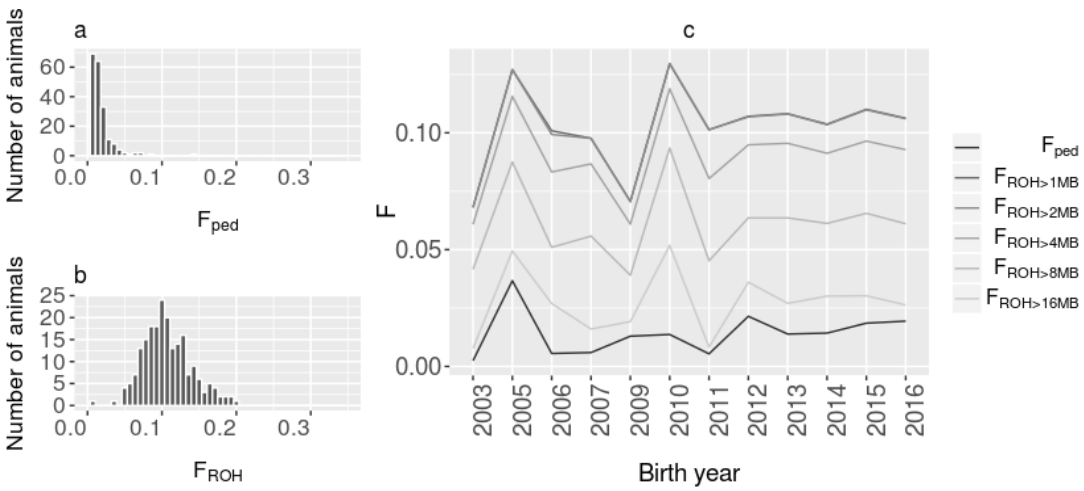

Figure 1: *a) Distribution of F_{ped} in the reference population; b) Distribution of $F_{ROH>1MB}$ in the reference population; c) Year-averaged F_{ped} and F_{ROH} in the reference population. F_{ped} = Meuwissen inbreeding coefficient, F_{ROH} = genomic inbreeding coefficient.*

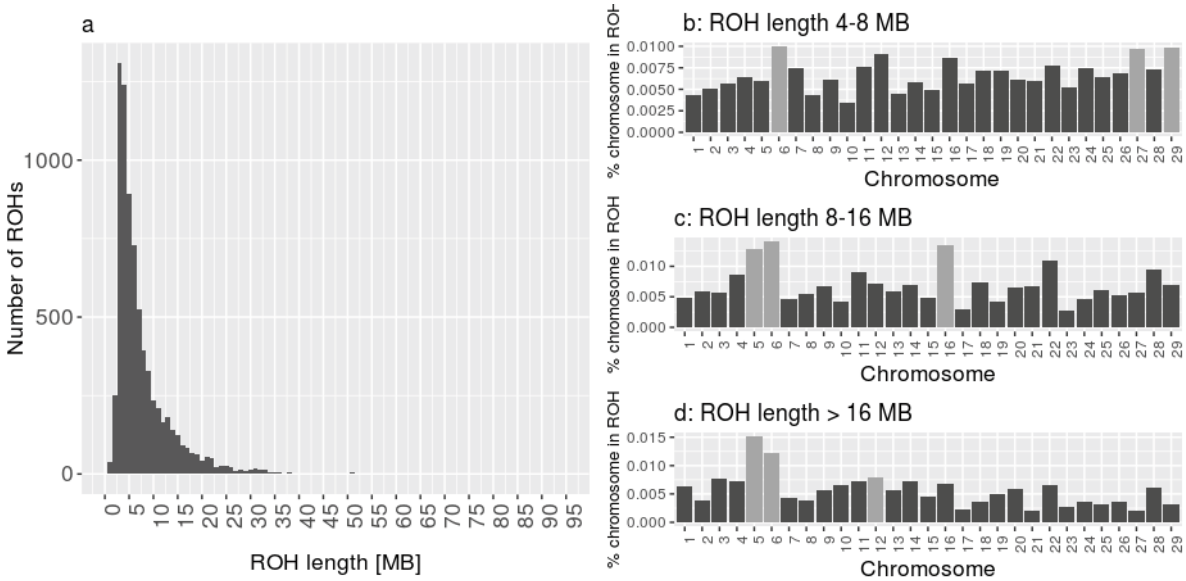

Figure 2: *a) Distribution of identified ROH lengths in MB; b) Percentage of chromosomes in ROHs 4–8 MB; c) Percentage of chromosomes in ROHs 8–16 MB; d) Percentage of chromosomes in ROHs >16 MB. ROH = runs of homozygosity.*

gree records does not imply a special concern should be dedicated to the management of inbreeding in Slovenian Brown-Swiss population, ΔF_{ROH} is close to the proposed rate and illustrates the need for an accurate control of the inbreeding in the breeding scheme. Results show that retrieved pedigree records are not adequate to accurately estimate the level and rate of inbreeding in the population. F_{ped} can capture only the inbreeding since the beginning of the pedigree records. The pedigree records for Slovenian Brown-Swiss population date back to 1950, therefore animals from this generation are assumed to be unrelated and non-inbred, an assumption which is most likely violated. F_{ROH} does not depend on pedigree information and is therefore not limited with the period and accuracy of record keeping.

A total of 7,470 ROHs larger than 1 MB were detected in the analysis (Fig. 2a). Lengths of homozygous segments follow exponential distribution with a mean of ½ g Morgan, where g in the number of generation since the common ancestor (Howrigan *et al.*, 2011). Therefore ROHs >1 MB date back ~50 generation, >2 MB ~25 generations, >4 MB ~12.5 generations >8 MB ~6 generations, and >16 MB ~3 generations. Shorter ROHs reflecting ancient inbreeding are more difficult to detect with pedigree analysis and require higher chip density or sequence information for detection (Zhang *et al.*, 2015). The distribution of identified ROH lengths suggests that the local population of Brown-Swiss experienced ancient as well as some recent inbreeding The majority of identified ROHs in this study fall into 2–4 MB class (29.0 %) dating back ~25–12.5 generations and 4–8 MB class (39.3 %) dating back ~12.5–6 generations. More ancient inbreeding, i.e. dating back 12.5 generations (~75 years), could be due to a bottle neck in Slovenian Brown population's history caused by the second world war and inbreeding afterwards.

Figure 2b, 2c and 2d illustrate the percentage of genome covered in ROH of different lengths, i.e. 4–8 MB, 8–16 MB and >16 MB. Stratifying by ROHs lengths and chromosomal location revealed some differences in highest inbred chromosomes according to ROHs of different length classes. Since ROHs reveal selection signatures and further more, ROHs of different lengths direct to a different point in population's history, this could provide insight into the history of selection decision and goals (Kim *et al.*, 2013). Bovine chromosome (BTA) 6 was among most highly inbred chromosomes according to all classes of ROHs (Fig. 2b, 2c, 2d). This could be explained with BTA6 having been associated with milk and mastitis traits and bearing most of these QTLs among all chromosomes (Ogorevc *et al.*, 2009). This is in concordance with other studies observing a ROH hotspot on BTA6 (Ferenčaković *et al.*, 2013a). Chromosomes shown as most highly inbred according to longer ROHs (Fig. 2b, 2c), indicating recent inbreeding (< 6 generations back) have been associated with dairy traits, i.e. BTA5 has been associated with milk production in two breeds (Raven *et al.*, 2014) and BTA12 has been associated with fertility in different cattle breeds (Olsen *et al.*, 2011; Minozzi *et al.*, 2013). However, more detailed investigation should be conducted in order to determine the cause of inbreeding of the specific chromosomes. Further on, a

comparison study between breeds should be performed once genotypes of other Slovenian breeds become available.

4 CONCLUSION

To conclude, control of inbreeding level is crucial in order to prevent inbreeding depression, especially in populations undergoing selection. This preliminary study shows that inbreeding computed based on pedigree records might not be adequate to capture the true inbreeding of the population and might not enable efficient control of the inbreeding. This becomes even a bigger concern with the introduction of genomic selection since it risks a higher rate of inbreeding comparing to classical selection due to a shortened generation interval (Boichard *et al.*, 2015). The development of genomic technologies enabled a more accurate estimation of inbreeding levels by using genotypic data to detect ROHs. The latter could provide a more powerful method to estimate true inbreeding levels since the pedigree estimation relies on the pedigree depth and captures only the inbreeding since the beginning of pedigree recordings. This has been shown also in this study where F_{ROH} exceeded F_{ped} for all the studies years. Although the majority of detected ROHs in this study are >4 MB reflecting inbreeding dating back less than 12.5 generation (~75 years) ago and pedigree records date back ~65 years, the latter are not complete. Furthermore, while estimated ΔF did not raise any special concern, estimated ΔF_{ROH} was close to the suggested upper limit for ΔF implying special attention needs to be paid to the control of inbreeding in the Slovenian Brown-Swiss breeding program. Additionally, it has been stressed that genomic selection requires genomic control of inbreeding (Sonesson *et al.*, 2012). However, it has to be taken into mind that it has been shown that 50K chip density can overestimate $F_{ROH>1MB}$ due to false positives (Ferenčakovič *et al.*, 2013b). Addition of more animal genotype on higher chip densities and expansion of animals' birth year will improve the detection of short ROHs reflecting more ancient inbreeding and enable a more accurate estimation of the inbreeding rate in the population.

4 REFERENCES

Boichard, D. (2002). Pedig: a fortran package for pedigree analysis suited for large populations. In: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, 19–23 Aug. 2002, CD-ROM communication č. 28–13.

Boichard, D., Ducrocq, V., Fritz, S. (2015). Sustainable dairy cattle selection in the genomic era. *Journal of Animal Breeding and Genetics*, *132*(2), 135–143.

Falconer, D. S., Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Harlow, UK: Longman.

Ferenčaković, M., Hamzić, E., Gredler, B., Solberg, T. R., Klemetsdal, G., Curik, I., Sölkner, J. (2013a). Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *Journal of Animal Breeding and Genetics, 130*(4), 286–293.

Ferenčaković, M., Sölkner, J., Curik, I. (2013b). Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genetics, Selection, Evolution : GSE, 45*(1), 42.

Groeneveld, E., Westhuizen, B. v D., Maiwashe, A., Voordewind, F., Ferraz, J. B. S. (2009). POPREP: a generic report for population management. *Genetics and Molecular Research: GMR*, *8*(3), 1158–1178.

Gurgul, A., Szmatoła, T., Topolski, P., Jasielczuk, I., Żukowski, K., Bugno-Poniewierska, M. (2016). The use of runs of homozygosity for estimation of recent inbreeding in Holstein cattle. *Journal of Applied Genetics*, 1–4.

Howrigan, D. P., Simonson, M. A., Keller, M. C. (2011). Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics, 12*, 460.

Kim, E.-S., Cole, J. B., Huson, H., Wiggans, G. R., Van Tassell, C. P., Crooker, B. A., Liu, G. et al. (2013). Effect of Artificial Selection on Runs of Homozygosity in U.S. Holstein Cattle. *PLoS ONE*, *8*(11). Retrieved from http://doi.org/10.1371/journal.pone.0080813

Logar, B., Podgoršek, P., Jeretina, J., Ivanoviè, B., Perpar, T. (2005). Online-available milk-recording data for efficient support of farm management. V: Knowledge transfer in cattle husbandry : new management practices, attitudes and adaptation. Wageningen, Wageningen Academic Publisher, 227–230.

Maignel, L., Boichard, D., Verrier, E. (1996). Genetic variability of French dairy breeds estimated from pedigree information. *Interbull Bulletin*, *14*, 49.

McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., et al. (2008). Runs of homozygosity in European populations. *American Journal of Human Genetics, 83*(3), 359–372.

Meuwissen, T., Luo, Z. (1992). Computing inbreeding coefficients in large populations. *Genetics Selection Evolution, 24*, 305.

Minozzi, G., Nicolazzi, E. L., Stella, A., Biffani, S., Negrini, R., Lazzari, B., Ajmone-Marsan, P., Williams, J.L. (2013). Genome Wide Analysis of Fertility and Production Traits in Italian Holstein Cattle. *PLoS ONE, 8*, 11. Retrieved from 10.1371/journal.pone.0080219

Nicolazzi, E. L., Marras, G. (2015). Zanardi: an open-source pipeline for genomic analysis using SNP array data. Fondazione Parco Tecnologico Padano, Via Einstein, Loc. Cascina Codazza (26900) Lodi, Italy.

Ogorevc, J., Kunej, T., Razpet, A., Dovc, P. (2009). Database of cattle candidate genes and genetic markers for milk production and mastitis. *Animal Genetics, 40*(6), 832–851.

Olsen, H. G., Hayes, B. J., Kent, M. P., Nome, T., Svendsen, M.,

Larsgard, A.G., Lien, S. (2011). Genome-wide association mapping in Norwegian Red cattle identifies quantitative trait loci for fertility and milk production on BTA12. *Animal Genetics, 42*(5), 466–474.

Potočnik, K., Jenko, J., Gorjanc, G. (2016). Ali imamo možnosti za lastno genomsko selekcijo v slovenski populaciji rjavega goveda? *Rjavo govedo*, *11*, 7–10.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller J., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics, 81*(3), 559–575.

Purfield, D. C., Berry, D. P., McParland, S., Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC Genetics, 13*, 70.

Raven, L. A., Cocks, B. G., Hayes, B. J. (2014). Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics, 15*, 62.

Rigler, M., Sadar, M., Potočnik, K. (2016). Izvajanje genomske selekcije pri rjavi pasmi v Sloveniji. *Rjavo govedo, 11*, 12–13.

Roshyara, N. R., Kirsten, H., Horn, K., Ahnert, P., Scholz, M. (2014). Impact of pre-imputation SNP-filtering on genotype imputation results. *BMC Genetics, 15*, 88.

Santus, E. (2011). Intergenomics: business rules and transition into services. *Interbull Bulletin, 43*, 2.

Sonesson, A. K., Woolliams, J. A., & Meuwissen, T. H. E. (2012). Genomic selection requires genomic control of inbreeding. *Genetics, Selection, Evolution: GSE*, *44*, 27.

Stachowicz, K., Sargolzaei, M., Miglior, F., Schenkel, F. S. (2011). Rates of inbreeding and genetic diversity in Canadian Holstein and Jersey cattle. *Journal of Dairy Science, 94*(10), 5160–5175.

Van der Werf J. (1999). Inbreeding and the effects of increased prolificacy. V: Animal Breeding: Use of New Technologies. Sydney, Post Graduate Foundation in Veterinarian Science of the University of Sydney: 188–198.

Woolliams, J. A., Gwaze, D. P., Meuwissen, T. H., Planchenault, D., Renard, J. P., Thibier, M., Wagner, H. (1998). *Secondary guidelines for the development of national farm animal genetic resources management plans–management of small populations at risk*. Rome: Food and Agriculture Organization of the United Nations.

Zhang, Q., Calus, M. P., Guldbrandtsen, B., Lund, M. S., Sahana, G. (2015). Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. *BMC Genetics*, *16,* 88. Retrieved from http://doi.org/10.1186/s12863-015-0227-7