Runs of Homozygosity Reveal Genome-wide Autozygosity in the Austrian Fleckvieh Cattle

Maja FERENCAKOVIC [1] (✉)
Edin HAMZIC [2]
Birgit GREDLER [2]
Ino CURIK [1]
Johann SÖLKNER [2]

Summary

Runs of homozygosity (ROH) are recognized as potential inbreeding measure in studies on humans. Inbreeding coefficients derived from ROH (F_{ROH}) measure proportion of the genome arranged in long homozygous segments and highly correlate with those derived from pedigree (F_{ped}). From that we assumed that ROH represent an alternative to pedigree inbreeding levels in studies on animals too, because pedigree can be incorrect, incomplete and can not fully explain what happened in meiosis. To confirm our premise we used pedigree and genotype data from 500 Austrian dual purpose Simmental bulls to determine correlation between F_{ROH} and F_{ped}. ROH were obtained using Fortran 90 software created by the authors. Proportions of genome in ROH were calculated for lengths of ROH of >1, >2, >4, >8 and >16 Mb. Pedigree data were analyzed and inbreeding coefficients for complete pedigree (F_{pedT}) and five generations (F_{ped5}) were calculated using ENDOG software. We found low F_{pedT} and F_{ped5} (means of 1.5% and 0.9%) while F_{ROH} for segments >1Mb suggested much higher values (9.0%) indicating old inbreeding that can not be traced using pedigree. The highest correlations were found between F_{ROH} calculated from ROH of length >4Mb and F_{pedT} (0.68) that is consistent with studies on humans. We conclude that inbreeding coefficients derived from ROH are useful for measuring levels of inbreeding in cattle, because ROH are not subject to mistakes as pedigrees and calculations made from those.

Key words

inbreeding, runs of homozygosity, genome-wide autozygosity, pedigree, cattle

[1] University of Zagreb, Faculty of Agriculture, Department of Animal Science, Svetošimunska 25, 10000 Zagreb, Croatia
✉ e-mail: mferencakovic@agr.hr
[2] University of Natural Resources and Life Sciences Vienna, Department of Sustainable Agricultural Systems, Division of Livestock Sciences, Gregor Mendel Str. 33, A-1180 Vienna, Austria

Received: June 1, 2011 | Accepted: July 14, 2011

ACKNOWLEDGEMENTS
This research was partially funded by Scientific project of Ministry of Science Education and Sports no. 178-1780460-0546 (Strategies for elimination of genetic defects in selected populations). We gladly acknowledge support in part by an "Ernst Mach grant – worldwide" financed by the Austrian Federal Ministry of Education and Science. We are very grateful to Zuchtdata EDV-Dienstleistungen GmbH and the Association of Austrian Simmental breeders (AGÖF) for provision of genotype and pedigree data.

Introduction

Mating of related individuals results with inbred offspring that are generally less viable, less fertile or/and smaller than the population mean. The phenomenon is also known as inbreeding depression and occurs regularly in animal and plant breeding, in small natural populations (Pirchner, 1985; Charlesworth and Charlesworth, 1987) and in humans (Schull and Neel, 1965). The inbreeding coefficient is a measure of inbreeding defined as the proportion of an individual's genome that is autozygous, relative to that of a poorly characterized founder generation. Since it was developed (Wright, 1922) inbreeding coefficient has mainly been estimated from the pedigree information, here denoted as F_{ped}. The advent of high throughput methods enabled genotyping of individual (animals) for a large number of molecular markers spread all over the genome, and further stimulated development of molecular measures that estimate autozygosity of an individual (Leutenegger et al., 2003; Carothers et al., 2006; Polasek et al., 2010). Runs of homozygosity (ROH) were recently proposed as a measure quantifying individual autozygosity (McQuillan et al., 2008; Nalls et al., 2009). A ROH is a continuous or uninterrupted part of genome without heterozygosity in the diploid state. As recombination interrupts long chromosome segments over the time, it is expected that long identical segments come, through the parents, from the same haplotype of their common ancestor. Furthermore, the number of segregations to the common ancestors is lower for long segments in comparison to the shorter homozygous segments.

In addition, in an inbred population we expect to find more and longer homozygous segments than in outbreed populations (Gibson et al., 2006). Human genome studies have also shown that individuals born in consanguineous unions (marriages between close relatives), have levels of homozygosity that are even higher than were expected (Woods et al., 2006; Broman and Weber, 1999) form pedigree information. Precise estimating inbreeding coefficients from pedigrees do not cover ancient relatedness and correct pedigrees. Even if the pedigree is well known and correct the estimates of inbreeding for single individuals can differ from expectation due to the stochastic pattern of inheritance. The mean inbreeding coefficient of the offspring of the first cousins is 0.0625 with a standard deviation of 0.0243 (Carothers et al., 2006). This variance increases with each meiosis, so it is possible for offspring of the third cousins to be more autozygous than offspring of second cousins (McQuillan et al., 2008). The availability of genome scan technology, which can genotype individual at large number of markers, provides us with the opportunity to "observe" levels of true inbreeding. Thus, distribution and size of ROH can provide information for calculating true individual level of autozygosity.

Aim of this study was to compare pedigree inbreeding coefficients with measures derived from ROH information for a 500 Austrian dual purpose Simmental bulls. We will also compare our results with the similar studies obtained in human populations and provide information for utility of ROH as a tool for measuring inbreeding coefficients from molecular data in cattle.

Materials and methods

Overall 1837 Austrian dual purpose Simmental, 447 Brown Swiss and 217 Tyrol Grey bulls were genotyped using the Illumina 50K bovine SNP chip (San Diego, CA, USA). All markers with unknown position and/or chromosome assignment as well as with GC-score lower than 0.2 were removed before preparing input files for PLINK software (Purcell et al., 2007). After the application of PLINK software (Purcell et al., 2007), by applying parameters -- mind 0.05, --maf 0.001 and --geno 0.25, 42262 markers were left for analyses. Additionally we excluded 529 SNP assigned to X chromosome. Final data set was including 41733 SNP on 29 chromosomes and they cover 2557.47 Mb of genome. For this pilot study we only used the 500 youngest Austrian Simmental bulls (born in 2001 to 2004) available in the data set.

For the analyses of Austrian Simmental population the pedigree included 41090 animals. From the pedigree data we calculated the equivalent complete generations and pedigree inbreeding coefficients referred to all (F_{pedT}) and five generation long pedigree (F_{ped5}) using ENDOG v4.8 (Gutiérrez and Goyache, 2005).

ROH were determined using Fortran 90 software developed by authors. The software simply counts homozygous SNP along chromosome and by their bp-positions providing information on length of ROH within given parameters. Depending on the minimum length of ROH in which no heterozygote SNP were allowed, we calculated ROH1, ROH2, ROH4, ROH8 and ROH16 according to the size of ROH being 1, 2, 4, 8 and 16 Mb long, respectively. Every ROH was required to have a minimum of 15 SNP. We also calculated molecular inbreeding coefficients based on ROH. Depending on the ROH size molecular inbreeding coefficients F_{ROH1}, F_{ROH2}, F_{ROH4}, F_{ROH8} and F_{ROH16} were calculated by dividing the sum distances covered by the ROH per individual by length of genome covered by SNP as described in Leutenegger et al. (2003). All statistical analyses and figures were done with SAS software v9.2 (SAS, 2009)

Results and discussion

On the population of 500 genotyped Austrian Simmental we observed average complete generation equivalent of 7.30 (±0.41; range of 5.91 to 8.32). The maximum number of generations tracked in a pedigree was 17. Descriptive statistics of the pedigree inbreeding coefficient estimations is given in Table 2. All animals (except one) were inbred for all generations period while 74.6% were inbred for five generations period. Both average pedigree inbreeding coefficients were low (up to 1.50 % and 0.9%) that was consistent with previously reported levels of inbreeding in this population (Maximini et al., 2011). Descriptive statistics of total length and number of determinate ROH in 500 Austrian Simmental bulls is given in Table 1, while descriptive statistics of molecular inbreeding coefficients calculated from ROH are given in Table 2.

ROH greater than 1Mb cover on average 9.0 % of the genome while pedigree inbreeding indicates a proportion of only 1.5 %. The similar level of autozygosity was also estimated by ROH greater than 16 Mb, thus, indicating recent inbreeding. Difference is due to "old" inbreeding that can not be traced using pedigree data but can be with short ROH. This is confirmed by the observation that correlations of F_{ROH} and inbreeding coefficients (F_{pedT} and F_{ped5}) are higher for ROH greater than 4Mb than those for F_{ROH1} and F_{ROH2} (Table 3). Studies on humans (McQuillan et al., 2008) give similar information. Low level of autozygosity from pedigree data (0.38%) was confirmed with low level of

Table 1. Descriptive statistics for the total length (in Mb) and number (in brackets) of runs of homozygosity (ROH) of different size (>1, 2, 4, 8 and 16 Mb) in 500 Austrian Simmental bulls

Total length of ROHs (Number of ROHs)	Mean	Standard deviation	Minimum	Maximum
>1 Mb	229.25 (96.79)	55.02 (13.37)	81.87 (48)	498.86 (135)
>2 Mb	139.17 (30.49)	52.20 (6.99)	24.15 (7)	419.65 (56)
>4 Mb	82.91 (9.65)	48.24 (3.99)	5.11 (1)	358.04 (29)
>8 Mb	52.86 (3.50)	42.79 (2.35)	8.09 (1)	290.70 (17)
>16 Mb	47.16 (1.89)	33.03 (1.23)	16.03 (1)	181.99 (7)

Table 2. Descriptive statistics of the pedigree and molecular, derived from runs of homozygosity (ROH) of different size (1, 2, 4, 8 and 16 Mb), inbreeding coefficients in 500 Austrian Simmental bulls

Inbreeding coefficient	Mean	Standard deviation	Minimum	Maximum
F_{pedT}	0.015	0.013	0.000	0.090
F_{ped5}	0.009	0.012	0.000	0.085
F_{ROH1}	0.090	0.022	0.032	0.195
F_{ROH2}	0.054	0.020	0.009	0.164
F_{ROH4}	0.032	0.019	0.002	0.140
F_{ROH8}	0.021	0.017	0.003	0.114
F_{ROH16}	0.018	0.013	0.006	0.071

Table 3. Correlations between pedigree and molecular, derived from runs of homozygosity (ROH) of different size (1, 2, 4, 8 and 16 Mb), inbreeding coefficients in 500 Austrian Simmental bulls

Inbreeding coefficient	F_{ROH1}	F_{ROH2}	F_{ROH4}	F_{ROH8}	F_{ROH16}	F_{pedT}
F_{ROH2}	0.969					
F_{ROH4}	0.920	0.965				
F_{ROH8}	0.866	0.914	0.950			
F_{ROH16}	0.755	0.800	0.839	0.890		
F_{pedT}	0.644	0.674	0.683	0.682	0.632	
F_{ped5}	0.613	0.648	0.663	0.671	0.625	0.973

autozygosity from ROH of minimum 5Mb (0.45%). Very small segments (minimum 0.5Mb), which could be observation of very old inbreeding give higher values (3.9%). The highest correlation between pedigree inbreeding coefficient and those derived from ROH in Orcadian is for length of ROH of minimum 1.5 Mb and it is 0.86. In our case of somewhat more inbred animals genome is covered with larger autozygous segments (4-8 Mb).

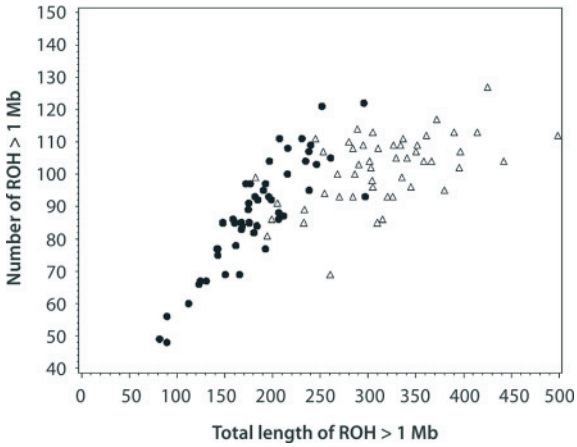

Figure 1. Relationship between number and total length of ROH, runs of homozigosity (ROH) greater than 1 Mb, in 50 animals with low (black dots) and in 50 animals with high (white triangles) total pedigree inbreeding coefficient

Number and length of ROH are getting higher as inbreeding coefficient grows. In animals with high inbreeding, the number of ROH is more constant while length of genome covered with ROH is different, suggesting very large segments. This can be observed from Figure 1 where we have chosen 50 animals with lowest F_{pedT} (0 - 0.005) and 50 animals with highest F_{pedT} (0.033 – 0.09).

Conclusion

We conclude that inbreeding coefficients derived from ROH are a very useful tool for indicating levels of inbreeding, especially if pedigree data are missing or pedigrees are not correct. They are also giving a better picture about inbreeding in the ancestral population that we are usually not able to track. Information that we receive from ROH not only provides information on inbreeding level but also on its age. Using the observational approach rather than the probabilistic one applied in pedigree analysis, most likely provides more accurate information about the state of autozygosity of individuals.

References

Broman K.W. and Weber J.L. (1999). Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. Am J Hum Genet 65: 1493–1500.

Carothers A. D., Rudan I., Kolcic I., Polasek O., Hayward C., Wright A. F., Campbell H., Teague P., Hastie N. D. and Weber J. L. (2006). Estimating Human Inbreeding Coefficients: Comparison of Genealogical and Marker Heterozygosity Approaches. Ann Hum Genet 70: 666–676.

Charlesworth D. and Charlesworth B., (1987). Inbreeding depression and its evolutionary consequences. Annu Rev Ecol Syst 18: 237-268.

Gibson J., Morton N.E., Collins A.(2006). Extended tracts of homozygosity in outbred human populations. Hum Mol Genet 15: 789–795.

Gutiérrez J. P. and Goyache F. (2005). A note on ENDOG: a computer program for analysing pedigree information. J Anim Breed Genet 122: 172-176.

Leutenegger A.-L., Prum B., Genin E., Verny C., Lemainque A., Clerget-Darpoux F., Thompson E.A.(2003). Estimation of the inbreeding coefficient through use of genomic data. Am J Hum Genet 73: 516-523.

Maximini L., Fuerst-Waltl B., Gredler B. and Baumung R. (2011). Inbreeding Depression on Semen Quality in Austrian Dual-Purpose Simmental Bulls. Reprod Domest Anim 46: e102–e104.

McQuillan R., Leutenegger A.-L., Abdel-Rahman R., Franklin C.S., Pericic M., Barac-Lauc L., Smolej-Narancic N., Janicijevic B., Polasek O., Tenesa A., Macleod A.K., Farrington S.M., Rudan P., Hayward C., Vitart V., Rudan I., Wild S.H., Dunlop M.G., Wright A.F., Campbell H., Wilson J.F. (2008). Runs of homozygosity in European populations. Am J Hum Genet 83: 359–372.

Nalls M.A., Guerreiro R.J., Simon-Sanchez J., Bras J.T., Traynor B.J., Gibbs J.R., Launer L., Hardy J., Singleton A.B. (2009). Extended tracts of homozygosity identify novel candidate genes associated with late-onset Alzheimer's disease. Neurogenetics 10: 183-190.

Polasek O., Hayward C., Bellenguez C., Vitart V., Kolcic I., McQuillan R., Saftic V., Gyllensten U., Wilson J.F., Rudan I., Wright A.F., Campbell H., Leutenegger A.L. (2010). Comparative assessment of methods for estimating individual genome-wide homozygosity-by-descent from human genomic data. BMC Genom 11: 139.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. Am J Hum Genet 81: 559-575.

Pirchner F. (1985). Genetic structure of population. 1. Closed populations or matings among related individuals. In: General and Quantitative Genetics (A.B. Chapman, ed.), World Animal Science A4 Elsevier, Amsterdam. 227-250.

SAS (2009). SAS/STAT® User's Guide Version 9.2. SAS Institute Inc. Cary, NC, USA.

Schull W. J., Neel J. V. (1965). The Effect of Inbreeding on Japanese Children. Harper & Row, New York, USA.

Woods C.G., Cox J., Springell K., Hampshire D.J., Mohamed M.D., McKibbin M., Stern R., Raymond F.L., Sandford R., Malik Sharif S., Karbani G., Ahmed M., Bond J., Clayton D., Inglehearn C.F. (2006). Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. Am J Hum Genet 78: 889–896.

Wright S. (1922). Coefficients of inbreeding and relationship. Am Nat 56: 330–338.

---

acs76_63