Analysis of Genome Regions Showing Strong Inbreeding in Brown Swiss and Fleckvieh Cattle

H. Schwarzenbacher
ZuchtData EDV Dienstleistungen GmbH, Vienna, Austria
Email: schwarzenbacher@zuchtdata.at

Abstract

Two breeds with diverse breeding history were analysed for footprints of recent inbreeding. For this purpose, SNP chip data from 4,049 Brown Swiss and 8,307 Fleckvieh bulls was analysed for long homozygous segments, so called *'runs of homozygosity'*. Presence and length of homozygous blocks in the genome was substantially higher in Brown Swiss when compared to Fleckvieh. Across chromosomes, BTA 5 and 6 showed the strongest inbreeding, particularly in Brown Swiss. The genome-wide search for regions with extensive homozygosity revealed large runs of homozygosity around 35 and particularly 91 MB on BTA 6. At a single location on BTA 6 a large homozygous block of more than 5 MB was present in 50% of all Brown Swiss bulls.

Introduction

Rise in levels of inbreeding is an inevitable consequence of intense artificial selection. QTL alleles with favorable effects on economically important traits tend to increase in frequency. Alleles, at linked, non-selected sites also increase in frequency, a phenomenon that was coined "Hitchhiking" by Maynard Smith and Haigh (1974). The higher the selection coefficient on the selected site, the more pronounced the hitchhiking effect. One typical signature of hitchhiking are regions of long haplotypes in a population as they increase in frequency faster than recombination can break up linkage between selected and non selected sites. This leads to loss of variability not only at the selected site but also in the surrounding genome region. Selected haplotypes might also harbor unfavorable alleles on health and fitness which hereby increase in frequency and eventually get fixed.

Recent availability of genome-wide SNP data allows for new possibilities to analyze and control inbreeding and genetic diversity in livestock populations. Genetic diversity, a key parameter that limits long term gain, can now be analysed throughout the genome at very high resolution. 'Runs of homozygosity' (ROH) are one measure, to assess especially recent inbreeding. This idea was first suggested by Broman and Weber (1999), who proposed identifying autozygous segments from runs of consecutive homozygous markers. ROH are thus defined as long chromosome segments where all loci are in homozygous status. Long ROH blocks indicate recent inbreeding, since recombination events tend to break up long range haplotypes in the long term.

In this study two breeds with a diverse breeding history were included in the analysis: Braunvieh cattle from Switzerland were exported to the US in the 17th and 18th century. After strong selection for milk production, gene flow turned back to Europe where the autochthonous Braunvieh cattle was more or less replaced by Brown Swiss (BS). The BS breed exhibited intense selection at periods where the number of selection candidates was very limited. The Fleckvieh breed (FV, dual purpose Simmental) dates back to the early 19th century where Swiss Simmental cattle were exported to Bavaria and crossed with local breeds. Pure breeding started from 1900 onwards with increasing selection intensity. Nevertheless, the active breeding population in FV has always been much larger than in BS.

The focus of this paper was the analysis of presence and length of long homozygous segments as a measure of selection induced inbreeding of two cattle breeds with a quite different breed history: BS and FV.

Materials and Methods

Genotype data: Illumina BovineSNP50 genotype data of 4,049 Brown Swiss- and 8,307 Fleckvieh bulls from the joint genetic evaluation of Austria and Germany was included in the analysis. Genotypes are validated by the Bavarian State Research Center for Agriculture (ITZ). Validation includes standard marker based criteria (call rate, MAF, HWE) and checks on accordance of genotypes of related animals based on pedigree information using direct comparisons (parent-offspring) and comparisons of marker-based with pedigree-based IBD-coefficients (Wang, 2002). Genotypes that show obvious conflicts are either deleted or set to unknown following a defined protocol. After these quality checks 38,312 markers in Brown Swiss and 41,082 in Fleckvieh were included in further analysis.

Measures on ROH: ROH blocks are defined as a sequence of at least two neighboring SNPs, all in homozygous state. Physical position of starts and ends of ROH blocks were assessed using UMD 3.1 map (ftp://ftp.cbcb.umd.edu/pub/data/assembly/Bos_taurus/Bos_taurus_UMD_3.1).

Then blocks are sorted by length in decreasing order for each bull and within each of the 29 autosomes. For each bull and chromosome the average size of sorted ROH blocks was calculated that span together a physical distance of 5,10, ...,30 Mb. Chromosome size differences were accounted for by multiplying for each chromosome the physical distance (5,10,…30 Mb) for which mean ROH block lengths were analysed with a scaling factor *S*=(BTAx)/(BTA25), where BTAx is the chromosome of interest and BTA25 is the shortest autosome in cattle. Genome-wide results are presented aggregated across chromosomes within bulls, while chromosome-wide results are aggregated across bulls within each chromosome. The genomic relationship matrix, used to obtain genome-based inbreeding coefficients along the diagonals of the matrix was set up as described by VanRaden, (2008) ('type 1') using base allele frequencies following the approach of Gengler *et al.* (2007).

For the analysis of specific regions that show strong inbreeding, each SNP was tested on the proportion of bulls that exhibited a ROH block of at least 5 Mb containing the position of interest. Haplotypes were inferred using Beagle software v3.3.1 (Browning and Browning, 2007) for chromosomes, namely 5, 6 and 16 that show high rates of inbreeding. Pedigree information was not accounted for in haplotype inference. For each position of interest, a centered segment spanning 20 Mb was plotted. To assist interpretation, haplotypes were ordered by core SNP alleles and with widening the window in both directions of the chromosome, by decreasing haplotype frequency, respectively.

All statistical analyses were carried out using R (cran.r-project.org) statistical package.

Results and Discussion

The focus of the analysis of ROH block in this paper was to compare presence and length of homozygous blocks in the genome between BS and FV. A comparison of the average size of sorted ROH blocks that span together a distance of 5 to 30 Mb in Brown Swiss and Fleckvieh is presented in Table 1. Across all segment lengths analyzed, BS exhibit blocks that are approximately twice as long compared to FV. As expected, BS exhibits much stronger recent inbreeding than FV.

The comparison of the mean ROH block size spanning together a distance of 15 Mb within each chromosome is presented in Table 2.

Table 1. Genome-wide distribution of average length (in Mb) of sorted ROH blocks, spanning together a distance of 5 to 30 Mb in Brown Swiss (BS) and Fleckvieh (FV) cattle.

Segment-	BS		FV	
length	mean	95% CI	mean	95% CI
5	5.54	2.44-9.92	2.65	1.45-5.11
10	3.63	1.45-7.33	1.67	0.98-3.79
15	2.25	0.96-5.15	1.13	0.75-2.71
20	1.39	0.72-3.47	0.81	0.61-1.95
25	0.90	0.57-2.36	0.61	0.50-1.15
30	0.62	0.45-1.39	0.47	0.41-0.63

Relative to genome wide average, FV but especially BS show much longer ROH blocks for BTA 5 and 6 while BTA 23 and 25 show

shorter than average ROH blocks. This means that BTA 5 and 6, particularly in BS, show higher levels of recent inbreeding than average, most likely as a consequence of artificial selection.

Table 2. Chromosome-specific distribution of average length (in Mb) of sorted ROH blocks, spanning together a distance of 15 Mb in Brown Swiss (BS) and Fleckvieh (FV) cattle.

BTA	BS		FV	
	mean	95% CI	mean	95% CI
1	1.64	0.65- 7.88	1.11	0.65-2.31
2	1.93	0.70-12.17	1.20	0.71-2.53
3	1.99	0.65-12.15	1.10	0.65-2.44
4	2.30	0.64-14.64	1.06	0.62-2.36
5	**4.55**	**0.86-25.92**	**1.60**	**0.78-6.12**
6	**4.54**	**0.79-22.48**	**1.50**	**0.69-3.88**
7	1.83	0.67- 9.63	1.13	0.66-2.15
8	2.02	0.59-14.02	0.97	0.60-1.66
9	1.98	0.69-12.64	1.26	0.69-3.36
10	1.61	0.65- 9.36	1.02	0.62-2.45
11	2.87	0.69-19.95	1.24	0.64-2.94
12	2.61	0.66-16.73	1.27	0.67-2.94
13	2.81	0.58-17.91	0.98	0.56-1.95
14	3.16	0.61-29.72	1.05	0.57-2.35
15	1.83	0.60-11.68	1.03	0.59-2.22
16	**3.10**	**0.65-18.35**	1.57	0.67-6.91
17	1.78	0.58- 9.93	1.10	0.56-2.64
18	2.28	0.58-14.52	1.13	0.58-2.88
19	1.85	0.54-11.84	0.84	0.51-2.05
20	2.04	0.55-13.80	1.12	0.57-3.32
21	2.07	0.61-13.60	1.05	0.57-2.36
22	2.55	0.56-22.77	1.06	0.56-2.66
23	**1.29**	**0.53- 6.21**	**1.05**	**0.53-3.04**
24	2.10	0.58-13.80	1.19	0.57-3.71
25	**1.39**	**0.43- 8.80**	**0.90**	**0.45-2.50**
26	1.74	0.52- 12.61	1.16	0.51-4.58
27	1.81	0.56- 9.90	1.01	0.56-2.41
28	1.46	0.49- 8.97	0.98	0.51-2.64
29	2.01	0.56- 14.58	1.05	0.55-3.03
Mean	**2.25**	**0.61-14.36**	**1.13**	**0.60-2.98**

Although the ROH block size statistic calculated in this paper has no direct biological meaning, Table 3 presents correlations with pedigree based (F_{NRM}) and genome-based (F_{IBS}) inbreeding coefficients. In both breeds correlations are highest for segments of 5 Mb length. While in FV correlations are higher to F_{IBS} as compared to F_{NRM}, it is vice versa in BS. However, correlation of F_{IBS} with F_{NRM} is 0.498 in FV while it is only 0.224 in BS. Results from internal, unpublished analyses indicate that distortion of allele frequencies due to population subdivision might cause the weak concordance of F_{IBS} with F_{NRM} in BS. The moderate correlations of F_{IBS} with our ROH measure should therefore not be over interpreted.

Table 3. Pearson correlation coefficient in Brown Swiss (BS) and Fleckvieh (FV) between genome-wide measures of average size of ROH blocks (in Mb), spanning together a distance of 5 to 30 Mb with pedigree-based and genome-based inbreeding coefficients, respectively.

Segment-length	BS		FV	
	r ROH-F_{NRM}	r ROH-F_{IBS}	r ROH-F_{NRM}	r ROH-F_{IBS}
5	0.33	0.12	0.57	0.64
10	0.31	0.15	0.54	0.61
15	0.29	0.17	0.49	0.55
20	0.27	0.18	0.43	0.47
25	0.23	0.17	0.36	0.39
30	0.20	0.15	0.28	0.32

Figure 1 shows the presence of ROH blocks with at least 5 Mb as proportion of all bulls under investigation. As already indicated in Table 2, BTAs 5 and 6 show strong inbreeding with up to 50% of all bulls being homozygous for 5 Mb segments at certain positions. Compared to BS the presence of large ROH blocks in FV cattle is much rarer.

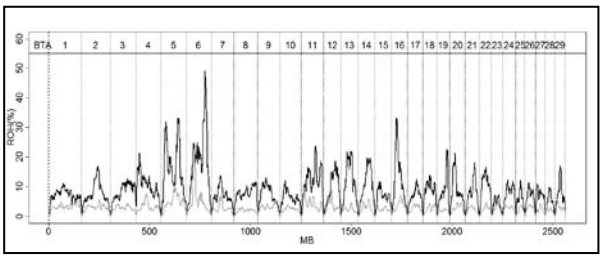

Figure 1. Percentage of bulls in Brown Swiss (black line) and Fleckvieh (grey line) that carry a ROH block of at least 5 Mb throughout the genome.

BTA 6 which is well known to harbor QTL for milk production traits is magnified in Figure 2. Regions with strong inbreeding can be found in BS around 91 Mb and with relatively good accordance in both breeds around 35 Mb.

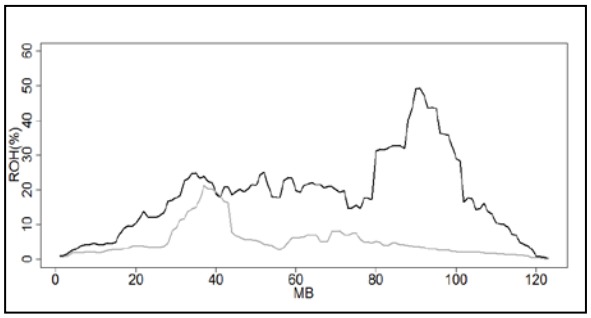

Figure 2. Percentage of Animals in Brown Swiss (black line) and Fleckvieh (grey line) that carry a ROH block of at least 5 Mb on BTA 6.

Selection on loci in the casein gene cluster, located around 88 Mb on BTA6 might be a possible cause for the obvious inbreeding in this region. Sodeland *et al.* (2011) have reported the presence of a long range haplotype in this region with favorable effects on milk production traits in Norwegian Red cattle. Moreover QTL on clinical mastitis were reported in an interval from 89-91MB (Nilsen *et al.*, 2006; Sodeland *et al.*, 2011).

Figure 3 presents for the same region on BTA 6 an ordered haplotype plot of a 20 Mb segment in BS. Extensive homozygosity due to selection- induced inbreeding extends to a large proportion of the chromosome. Such strong inbreeding is disadvantageous due to the loss of variability of hitchhiking loci and the risk that loci with unfavorable effects on fitness get fixed.

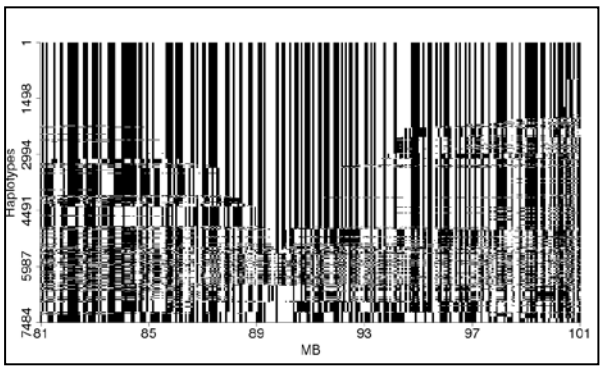

Figure 3. Ordered haplotype plot of BTA 6 in Brown Swiss showing a 20 Mb segment spanning the region around 91 Mb. Alleles coded with 0/1 are symbolized in white/black color, respectively. Barcode-like pattern is caused by massive numbers of identical haplotypes.

Conclusions

This analysis revealed regions under intense selection with extensive homozygous blocks especially in BS. The availability of massive, genome wide SNP data opens new possibilities to manage inbreeding in livestock which should be used for optimal allocation of genetic gain and genetic diversity in intensively selected cattle breeds.

References

Broman, K.W. & 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.

Browning, S.R. & Browning, B.L. 2007. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies using localized haplotype clustering. *Am. J. Hum. Genet. 81*, 1084-1097.

Gengler, N., Mayeres, P. & Szydlowski, M. 2007. A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual-purpose Belgian Blue cattle. *Animal 1*, 21-28.

Maynard Smith, J. & Haigh, J. 1974. The hitchhiking effect of a favorable gene. *Genet. Res. 23*, 23-35.

Nilsen, H., Olsen, H.G., Hayes, B., Nome, T., Sehested, E., Svendsen, M., Meuwissen, T.H.E., & Lien, S. 2009. Characterization of a QTL region affecting clinical mastitis and protein yield on BTA6. *Anim. Genet. 40:5*, 701-712.

Sodeland, M., Grove, H., Kent, M., Taylor, S., Svendsen, M., Hayes, B.J. & Lien, S. 2011. Molecular characterization of a long range haplotype affecting protein yield and mastitis susceptibility in Norwegian Red cattle. *BMC Genetics 12*, 70.

VanRaden, P.M. 2008. Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci. 91*, 4414-4423.

Wang, J. 2002. An Estimator for Pairwise Relatedness Using Molecular Markers. *Genetics 160*, 1203-1215.