ANIMAL GENETICS
Immunogenetics, Molecular Genetics and Functional Genomics

Identification of genomic regions involved in the phenotypic differences between original and modern Brown populations

Journal:	*Animal Genetics*
Manuscript ID	AnGen-20-06-0160.R1
Manuscript Type:	Full Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Moscarelli, Angelo; University of Palermo, Scienze Agrarie, Alimentari e Forestali Sardina, Maria Teresa; University of Palermo, Scienze Agrarie, Alimentari e Forestali Cassandro, Martino; University of Padova, Department of Animal Science Ciani, Elena; University of Bari, Biosciences, Biotechnologies and Pharmacological Sciences Pilla, Fabio; University of Molise, Agricoltura Ambiente Alimenti Senczuk, Gabriele; University of Molise, Dipartimento Agricoltura, Ambiente e Alimenti Portolano, Baldassare; University of Palermo, Scienze Agrarie, Alimentari e Forestali Mastrangelo, Salvatore; University of Palermo, Scienze Agrarie, Alimentari e Forestali
Keywords:	Brown cattle, FST, runs of homozygosity, SNPs, candidate genes

Identification of genomic regions involved in the phenotypic differences between original and modern Brown populations

A. Moscarelli[1], M. T. Sardina[1], M. Cassandro[2], E. Ciani[3], F. Pilla[4], G. Senczuck[4], B. Portolano[1] and S. Mastrangelo[1]*

[1] Dipartimento Scienze Agrarie, Alimentari e Forestali, University of Palermo, 90128 Palermo, Italy

[2] Dipartimento di Agronomia Animali Alimenti Risorse naturali e Ambiente, University of Padova, 35020 Legnaro, Italy.

[3] Dipartimento di Bioscienze, Biotecnologie e Biofarmaceutica, University of Bari, 70124 Bari, Italy

[4] Dipartimento Agricoltura, Ambiente e Alimenti, University of Molise, 86100 Campobasso, Italy

*Corresponding author

Salvatore Mastrangelo

salvatore.mastrangelo@unipa.it

Summary

Identifying genomic regions involved in the differences between breeds can provide information on genes that are under the influence of both artificial and natural selection. The aim of this study was to identify genomic regions involved in phenotypic differentiation among four different Brown populations (original *vs* modern) and to characterize the distribution of runs of homozygosity (ROHs) islands using the Illumina Bovine SNP50 BeadChip genotyping data. After quality control, 34,735 SNPs and 106 animals were retained for the analyses. Larger heterogeneity was highlighted for the original breeds. Patterns of genetic differentiation, multidimensional scaling, and the neighboring joining tree distinguished the modern breeds from the original populations. The F_{ST} - outlier identified several genes involved in many phenotypic differences between the two groups, such as stature and growth, behavior and adaptability to local environments. The ROH islands within both the original and the modern groups overlapped with QTL associated with relevant traits. In modern Brown, ROH islands harbored candidate genes associated with milk production traits, in evident agreement with the artificial selection conducted to improve this trait in these populations. In original Brown, we identified candidate genes related with fat deposition, confirming that breeding strategies for the original Brown populations aimed to produce dual-purpose animals. Our study highlighted the presence of several genomic regions which vary between Brown populations, in line with their different breeding histories.

Keywords Brown cattle, genetic diversity, F_{ST}, runs of homozygosity, candidate genes

Introduction

In cattle, natural and artificial selection has resulted in divergent breeds. In fact, the continuously increasing demand for work, milk and meat has enhanced between-population differences over the centuries (Zhao *et al*., 2015). An interesting situation regarding this divergence among populations is represented by the Brown cattle. The Brown cattle is derived from populations used in valleys

and mountain slopes of Switzerland since before historic records began. One theory says that this breed goes back to oriental origins, having been introduced into Central Europe from the steppes and valleys of Western Asia. Moreover, cattle bones found in the ruins of the Swiss Lake Dwellers indicate that a type of cattle, apparently closely related to actual Brown cattle, existed during the Bronze Age in Switzerland (Del Bo *et al.,* 2002). Nowadays, the Brown cattle reared in Europe can be grouped into three different populations: the Original Braunvieh, the Braunvieh, and the Brown Swiss (Hagger, 2005). The Original Braunvieh is a dual-purpose cattle breed reared in Switzerland and it represents the original population of Brown cattle (Bhati *et al*., 2020). This population is ancestral to the Brown Swiss; in fact, few individuals of the Original Braunvieh (about 170) originating from the mountain tops of Northeast Switzerland had been imported in the USA during the end of the 19th century giving rise to the current Brown Swiss, which is today one of the most widespread dairy breed in the world (Yoder & Lush 1937). Between 1967 and 1998, a crossbreeding between Original Braunvieh and the American Brown Swiss was conducted through artificial insemination, leading to the current Braunvieh population; this is mainly spread over the Alpine regions of Austria, Germany, Italy, and Switzerland.

Genetic drift, geographical barriers, different environments, and local management practices have affected the divergences among the original and modern Brown populations. Moreover, natural selection and modern breeding approaches, while acting on adaptation, morphological and production traits, also contributed to shape the genetic structure of these populations.

The availability of single nucleotide polymorphism (SNP) panels, consisting of thousands of markers, has greatly improved the power of genome-wide studies for a deep investigation of genetic diversity between and within population(s), allowing the identification of highly differentiated genomic regions between cattle breeds (Flori *et al*., 2009; Qanbari *et al*., 2011; Signer-Hasler *et al.*, 2017; Mastrangelo *et al.*, 2018a). The analytical approach, using the allele frequency-based inter-population genetic differentiation (F_{ST}) and intra-population runs of homozygosity (ROHs), has been recently applied in several livestock species for the identification of genomic regions involved

in phenotypic differences (Cesarani *et al.*, 2018; Onzima *et al.*, 2018; Elbeltagy *et al.*, 2019; Hulsegge *et al.*, 2019; Szmatoła *et al.*, 2019). Finding links between phenotypic and genotypic differences is of great importance in order to gain a better understanding of genetic mechanisms underpinning traits of interest, and it offers the opportunity to improve the efficiency of animal breeding through directed selection on favorable alleles (Rothammer *et al.*, 2013). Therefore, the objective of the present study was to identify differentiated genomic regions among the Brown populations (original *vs* modern) and to characterize the distribution of runs of homozygosity islands using the Illumina Bovine SNP50 BeadChip genotyping data, which may provide insights into the mechanisms underlying their genomic differences. We also checked if the genomic regions that were most associated with ROHs overlapped with reported quantitative traits loci (QTL).

Materials and methods

Sampling, genotyping and quality control

Overall, 106 animals belonging to four different Brown cattle populations were used: Original Braunvieh (ORBR), Braunvieh (BRHV), Brown Swiss (BRSW) and Italian Brown (ITBR). These breeds present differences in both phenotypic and production traits. For example, in the ORBR, selection breeding schemes are almost absent, whereas the Brown Swiss is the second largest producer of milk after the Holstein breed. Nevertheless, the original populations, such as the BRHV, are adapted to the harshness of mountain areas because of their good grazing characteristics, and are resistant to environmental conditions. These populations shows also a smaller body size compared to selected breeds.

Detailed information about the breeds, the samples and the grouping into original and modern cattle are reported in Table 1. All individuals were genotyped using the Bovine SNP50K BeadChip. SNPs were mapped using the *Bos taurus* ARS-UCD1.2 genome assembly. The markers were filtered using PLINK 1.07 (Purcell *et al.*, 2007) and considering only SNPs mapped in autosomal

chromosomes. SNPs with a minor allele frequency (MAF) lower than 0.01 and call rate lower than 95%, as well as individuals with missing genotyping rate lower than 90%, were removed.

Comparison of breeds

Pairwise genetic relationships were estimated to evaluate population substructure using identity-by-state (IBS) genetic distances calculated by PLINK 1.07 (Purcell *et al.*, 2007) and graphically represented by multidimensional scaling (MDS). PLINK was also used to estimate observed (Ho) and expected (He) heterozygosity, and the genomic inbreeding, which is based on the difference between the observed and expected numbers of homozygous genotypes (F_{HOM}). The ARLEQUIN version 3.5.2.2 (Excoffier & Lischer, 2010) was used to estimate population relatedness using pairwise estimates of F_{ST} among all four breeds. A neighbor-joining tree was constructed based on individual allele-sharing distances (--distance 1-IBS in PLINK) and visualized using SPLITSTREE (Huson & Bryant, 2006).

F_{ST}-outlier analysis

The F_{ST}-outlier approach implemented in the BayeScan software (Foll, 2012) was adopted to identify loci involved in the differentiation between the considered groups. BayeScan analyses comprised 20 pilot runs of 5,000 iterations, a burn-in of 50,000 iterations, a thinning interval of 10 (5,000 iterations were used for the estimation of posterior odds) with a resulting total number of 100,000 iterations. To control the number of false positives, significant markers were defined by applying a *q*-value threshold of 0.05 and using the 0.999 SNPs of F_{ST} percentile distribution.

Runs of homozygosity islands

Runs of homozygosity (ROHs) were estimated for each sample using PLINK 1.07 (Purcell *et al.*, 2007). The minimum length that constituted the ROH was set to 1 Mb. The following criteria were used to define the ROHs: (i) one missing SNP was allowed in the ROH and up to one possible heterozygous genotype; (ii) the minimum number of consecutive SNPs that constituted the ROH was set to 30; (iii) minimum density of 1 SNP every 100 kb; (iv) maximum gap between

consecutive SNPs of 1 Mb. The average length of ROH (L_{ROH}) was estimated. ROH were grouped into five classes of length (1 to ≤2 Mb, 2 to ≤4 Mb, 4 to ≤8 Mb, 8 to ≤16 Mb and >16 Mb) (Marras *et al*., 2014). The number and frequency of ROH within each ROH length category for the two groups were also determined. The percentage of SNPs residing within the ROH was estimated by counting the number of times that each SNP appeared in a ROH and by dividing that number by the number of animals in each group. To identify the genomic regions of "high homozygosity", also called ROHs islands, the top 0.999 SNPs of the percentile distribution of the locus homozygosity range within each group were selected.

Gene ontology enrichment

Genomic regions detected by the two statistical approaches were interrogated for genes annotated to the *Bos taurus* genome assembly ARS-UCD1.2 using Genome Data Viewer (https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_002263795.1) provided by NCBI.

The genes within ROH islands were further analyzed with the PANTHER Classification System (Mi *et al.*, 2013) to identify significant (P ≤0.05) gene ontology (GO) terms. Moreover, to investigate the biological function of the annotated genes, an accurate literature search was also conducted.

Finally, using "JBrowse" (Buels *et al.*, 2016), an open source JavaScript-based genome browser, available on National Animal Genome Research Program (https://www.animalgenome.org/jbrowse/), we checked if the genomic regions of high homozygosity overlapped with reported quantitative traits loci (QTL).

Results

Comparison of breeds

After quality control, 34,735 SNPs and 106 animals were retained for the analyses. Genetic diversity indices (Ho and He) and inbreeding coefficient (F_{HOM}), which are key parameters in the genetic management of populations, were used to determine the levels of genetic variability in the four Brown breeds. The original populations (ORBR and BRHV) displayed the highest genetic diversity, whereas the lowest value was found in modern populations (BRSW and ITBR) (Table 2). To examine and visualize the genetic relationships among the four Brown cattle populations, we used an MDS plot of the pairwise identity-by-state distance (Figure 1). The results showed that the populations formed two different clusters. The first dimension (C1) separated BRSW and ITBR from ORBR and BRHV. In agreement with MDS results, the neighbor joining (NJ) tree based on allele sharing distance (ASD) separated individuals according to their population of origin (Fig. S1). Genetic differentiation between all pairs of populations estimated by F_{ST} statistics are also reported in Table S1, and substantially confirms results from MDS and NJ tree. In particular, F_{ST} was low between BRSW and ITBR (0.006) or between ORBR and BRHV (0.016).

All these results were used to categorize the four Brown populations into two contrasting groups for comparative analysis: original (ORBR and BRHV) *versus* modern (BRSW and ITBR) cattle.

F_{ST} analysis

Results from the Bayesian population differentiation approach identified a total of 35 outlier SNPs between ORBR-BRHV and BRSW-ITBR (Table 3). These SNPs were identified on 15 different chromosomes (BTA). Most of these markers were located far apart from each other. Manhattan plots of F_{ST} values are reported in Figure 2. The locus with the highest value (0.305) was ARS-BFGL-NGS-40251 on BTA02. Only three outlier genomic regions showed adjacent SNPs on BTA02, BTA06, BTA17, in which *NYAP2*, *GRID2* and *TMEM132D* are mapped, respectively (Table 3). Among the outlier SNPs, 11 markers (31%) were located on BTA6. A total of 17 markers that exceeded the significance threshold were mapped within 14 protein-coding genes.

Runs of homozygosity

The total number of detected ROHs exhibited variation between groups, with the modern group having the highest value (1,725). Moreover, the modern Brown group had also the highest average length of ROHs (8.54 Mb *vs* 7.17 Mb). The frequency and the distribution of the number of ROHs in the different classes was similar (Table S2 and Fig. S2). The most represented ROH classes in both groups was that of length ranging from 4 to ≤8 Mb, representing 48% and 44% in original and modern groups, respectively. The largest number of ROH in the class of highest length (>16 Mb) was observed in modern (BRSW and ITBR) cattle (Fig. S2). To identify the genomic regions that were most associated with ROHs, the top 0.999 SNPs of the percentile distribution of the locus homozygosity range were chosen as an indicator of a possible ROH hotspot in the genome. The analysis was carried out within the same two groups of the pairwise comparison reported above for the F_{ST}-outlier approach. Table 4 provides the chromosome position, the start and the end of ROH regions and the relative genes list. Although the distribution of the ROHs was relatively balanced and the signals were moderate in height, we found a few outstanding peaks with a high occurrence of ROHs (Figure 3). In total, five genomic regions that frequently appeared in the ROHs were identified in the two Brown groups (Table 4). The ROHs islands were found on BTA11 and BTA26 in the ORBR/BRHV group (Figure 3A), and on BTA05 and BTA06 in the BRSW/ITBR group (Figure 3B), ranging in length from 0.90 Mb (BRSW/ITBR; 9 consecutive SNPs on BTA05) to 5.58 Mb (ORBR/BRHV; 68 consecutive SNPs on BTA11).

Within the ROHs islands reported above, we identified several known genes, together with uncharacterized genes (LOC) (Table 4). Some of the detected islands, such as on BTA5 in the modern Brown breeds, contained only three annotated genes. Moreover, some of these genomic regions overlapped with ROHs islands found in other studies (Table S3).

Gene ontology enrichment

The results of the GO enrichment analysis (Table S4) showed multiple categories that were statistically significant ($P \leqslant 0.05$). The genes within ROHs islands encompass a wide spectrum of molecular functions, biological processes, and cellular components. A PANTHER gene list analysis

revealed a high percentage of genes involved in the following categories: cellular process (GO:0009987), catalytic activity (GO:0003824), binding (GO:0005488), metabolic process (GO:0008152), biological regulation (GO:0065007), localization (GO:0051179), response to stimulus (GO:0050896). Finally, we also checked if the ROHs overlapped with previously reported quantitative traits loci (QTL). Overall, the ROHs islands overlap with known QTL regions associated with several traits (Table S5). We identified QTL associated with milk composition and milk fat content, life history traits, energy efficiency association and intramuscular fat. For example, the ROH island in BTA06 overlap with QTL associated with somatic cell score, clinical mastitis, length of productive life and calf size.

Discussion

Identifying genomic regions involved in the differences between breeds can provide information on genes that are under the influence of both artificial and natural selection, and thus, can help the identification of beneficial mutations and underlying biological pathways for economically important traits (Zhao *et al.*, 2015). Notably, the inter-population F_{ST}-outlier test compares the variation of allele frequencies within and between groups, and the locus that shows the largest differences in allele frequencies between populations is assumed to be a signal of selection (Qanbari *et al.*, 2011). Intra-population ROHs analysis has been used to identify regions that have an unfavorable effect on a phenotype when they are in the homozygous state but also to detect associations between traits of economic interest and genes present in these regions (e.g. Zhang *et al.*, 2015; Mastrangelo *et al.*, 2018b). In this study, we applied the two above-mentioned genome-scan analytical approaches to identify differentiated genomic regions between original and modern Brown cattle populations and to characterize the distribution of runs of homozygosity islands within these populations.

Several studies investigated the genomic diversity and population structure of the Brown populations using either pedigree (Hagger, 2005) or microarray data (Signer-Hasler *et al.*, 2017;

Senczuk *et al.*, 2020), but to the best of our knowledge, this is the first study on the identification of differentiated genomic regions between original and modern Brown cattle populations.

Patterns of genetic differentiation, multidimensional scaling, and the neighboring joining tree showed a marked divergence between the Brown populations investigated here, and clearly distinguishing the modern breeds from the original populations. Similar patterns were also observed in previous studies (e.g. Del Bo *et al.*, 2011; Signer-Hasler *et al.*, 2017; Senczuk *et al.*, 2020). The observed genetic differentiation among breeds was in accordance with their geographical and historical origins. Moreover, genetic drift and the different selective schemes to which these populations have been subjected could have further contributed to the observed scenario. We also found that modern Brown cattle present lower values of genetic diversity, related with the high selection pressure on these breeds for dairy traits. Moreover, this group showed the highest frequency of ROH > 16 Mb (Fig. S2) as also reported by Cesarani *et al*. (2018), and related to a more recent inbreeding (Mastrangelo *et al*., 2016). On the other hand, the original populations consists of traditional pure-bred animals and has maintained substantial genetic diversity, due to the use of a relatively high number of natural service sires and relatively weak genetic selection for milk yield (Hagger, 2005).

The genome-wide distribution of pair-wise F_{ST} values was used to identify highly differentiated SNPs between the two groups of Brown cattle considered in this study. In general, the F_{ST}-outlier approach confirmed what already observed based on phenotypic differences between original and modern Brown cattle: the main genetic differences between the two groups lie in morphological traits and adaptability to local environments. Indeed, the results highlighted growth-related candidate genes such as *NYAP2* (Meng *et al.*, 2017), *CACNA2D1* (Hou *et al.*, 2010), *GRID2* (Lee *et al.*, 2012), *STIM2* and *CPEB4* (Mudadu *et al.*, 2016), while the outlier SNP in the BTA28 mapped on *PTPN20*, a candidate gene involved in the genetic determinism of udder conformation in cattle, an important trait under selection in the specialized dairy cattle breeds (Marete *et al.*, 2018). These results are consistent with the knowledge that breeding strategies, for the original Brown

populations, aimed to produce dual-purpose animals. Moreover, this comparison identified a gene related with a local adaptation trait, i.e. exposure to UV radiation (*DKK2*) (Pausch *et al.* 2012). In transhumant systems, the animals belonging to the original Brown populations graze at alpine pastures (between 1000 and 2400 m above sea level) during the summer months and return to the stables for the winter months (Bhati *et al*., 2020). Therefore, genes related with solar radiation are important for these breeds, that are exposed to increased summer solar ultraviolet radiation. In addition, on BTA17, *TMEM132D* was retrieved, a gene involved in behavioral traits such as reduced fear in humans and aggressiveness, consequently to human-driven selection (Qanbari *et al.*, 2014; Vallée *et al.*, 2016). All the above genes may very well explain some of the phenotypic difference between the two Brown groups. In our study, the genetic differentiation analysis between groups did not detect SNPs located in genomic regions known to contain genes associated with milk production traits, probably because under selection within both groups, even if with different selection pressure. Consistently with our results, Signer-Hasler *et al.* (2017), in a study on selection signatures with the same SNP array, which also involved original and modern Brown populations, reported that potential candidate genes based on phenotypic evidence, such as *DGAT1* or casein clusters, have not left any recognizable signal in these populations. The authors interpreted they results as the consequence of a low SNP density of the adopted SNP array in these regions.

As mentioned above, another method to identify candidate genomic regions, based on intra-population analysis, is the identification of runs of homozygosity islands. Since ROHs are normally abundant in regions under positive selection, their accumulation at specific loci has been used to identify genomic regions that reflect directional selection in livestock species (Mastrangelo *et al.*, 2018b; Purfield *et al.*, 2017; Metzger *et al.*, 2015). Moreover, several studies reported the occurrence of ROHs islands within QTLs for important production traits in cattle (Purfield *et al.* 2012; Szmatoła *et al.*, 2019). In this study, the ROH islands detected in the Brown populations overlapped with reported QTLs associated with relevant traits, such as milk composition and milk fat content, growth, body size, and length of productive life. Since ROH islands spanned many

candidate genes displaying several functions, in our analysis of the published literature, we mainly focused on genes related to livestock breeding traits.

In the original Brown group (ORBR/BRHV), at the beginning of the island on BTA11, we found the *ETAA1* gene, also reported within a ROH island in Sardo-Bruna cattle (Cesarani *et al.*, 2018), a local Italian Brown-derived population, thus confirming previously reported data. In fact, the overlapping ROHs islands among studies (Table S3) provided good evidence that they are not artifacts, but potential genomic regions affected by selection (Mastrangelo *et al.*, 2018). The island on BTA11 also included some known candidate genes (*C1D*, *WDR92, PNO1, PPP3R1, ARHGAP25*) related with fat deposition (Zhang *et al.*, 2019; Sorbolini *et al.*, 2015; Pant *et al.*, 2014), associated with high-altitude adaptation (*PPP3R1*) (Wei *et al.*, 2016), involved in immune and inflammatory response (*PLEK*) (Cremonesi *et al.*, 2012) and influencing the expression of flight speed (*ANTXR1*), used as a measure of temperament (Valente *et al.*, 2016). Based on these functions, some of these genes could be related with adaptation to cold climate, a characteristic of the original Brown populations, that commonly display better ability to adapt to harsh environments compared with the specialized modern breeds. These results confirmed the importance of the local populations as reservoirs of biodiversity and as models for studying the genetic basis of adaptability (Cesarani *et al*., 2018). Moreover, this island on BTA11 identified several candidates genes associated with fertility traits (*EHD3* and *PCBP1*). This may explain, at least partly, why local populations are generally characterized by high longevity. Finally, this island showed genes related with meat quality (*CAPN14*, *CAPN13*, *LBH* and *LCLAT1*), as previously reported within a significant selection signature in Original Braunvieh (Bhati *et al*., 2020; Rothammer *et al.*, 2013; Signer-Hasler *et al.*, 2017) and in line with the dual-purpose characteristics of these animals. Similarly, the ROH island on BTA26, also identified in the Italian Holstein cattle (Gaspa *et al.*, 2014), showed genes involved in bovine leg conformation (*BTRC* and *LBX1*) (Van den Berg *et at.*, 2014), related with lipid metabolism pathway (*FGF8* and *GBF1*) (Marques *et al.*, 2009; Iso-Touru *et al.*, 2016) and with milk protein and fat traits in cattle (*MGEA5*) (Lin *et al.*, 2019). Moreover, a

strong signature of selection (21-23 Mb) was reported in literature on this genomic region in Original Braunvieh (Signer-Hasler *et al.*, 2017). On the other side, in the ROH island on BTA5 of modern Brown, several candidate genes associated with milk production traits in cattle (*CSF2RB*, *NCF4* and *MPST*) (Lopdell *et al.*, 2019; Raven *et al.*, 2015) and in dairy buffaloes (*TMPRSS6*, *C1QTNF6*, *SSTR3*, *RAC2*, *CYTH4*) (De Camargo *et al.*, 2015) are mapped, in evident agreement with the artificial selection conducted to improve this traits in these populations. In the smallest ROH island in BTA5 we detected two genes, *TMTC1* and *ERIC2* associated with the *Mycobacterium avium* spp. *paratuberculosis* infection status (Pant *et al.*, 2010). Finally, of interest is the ROH island on BTA6 between about 86 and 87 Mb, a genomic region known to be functionally associated with traits related with mastitis, the disease that causes the greatest economic losses to the dairy industry worldwide. In fact, within this region, several candidate genes related with this disease, such as *SLC4A4* (Wu *et al.*, 2015), *GC* and *NPFFR2* (Sahana *et al.*, 2014) were retrieved.

Conclusion

Through the present study, we described some genetic differences between original and modern Brown populations. The different approaches used gave a picture of genetic relationships between original and modern breeds. As expected, larger heterogeneity was highlighted for the original breeds. Our study highlighted the presence of several genomic regions which vary between Brown populations, in line with their different breeding histories, and demonstrating that selection, other than genetic drift, contributed to the genetic differentiation among the original and modern breeds. The modern Brown has undergone the influences of intense artificial selection for milk production, losing some typical characteristics of the original Brown, like the dual propose attitude, establishing itself as a cosmopolitan breed for milk production. The original Brown maintains dual-purpose attitude, and have genetic characteristics linked to the marginal environments in which they are

reared and where continues to represent an important genetic resource for breeders. Most of the genes that were detected in our study were consistent with the phenotypic traits of these populations. Several genes and genomic regions here identified corroborate with previously reported studies carried out in other cattle breeds. However, further studies using the high-density array data, an increase in the number of genotyped animals and the collection of phenotypic records would be particularly relevant to refine and validate these results.

Conflict of interest

There are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

Bhati M., Kadri N.K., Crysnanto D. & Pausch H. (2020) Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. *BMC Genomics* 21(1), 1-14.

Buels R., Yao E., Diesh C.M. et *al*. (2016) JBrowse: a dynamic web platform for genome visualization and analysis. *Genome biology* 17(1), 66.

Cesarani A., Sorbolini S., Criscione A., Bordonaro S., Pulina G., Battacone G., Marletta D., Gaspa G. & Macciotta N.P.P. (2018) Genome-wide variability and selection signatures in Italian island cattle breeds. *Animal Genetics* 49(5), 371-383.

Cremonesi P., Capoferri R., Pisoni G. *et al*. (2012). Response of the goat mammary gland to infection with Staphylococcus aureus revealed by gene expression profiling in milk somatic and white blood cells. *BMC genomics*, 13(1), 540.

De Camargo G.M.F., Aspilcueta-Borquis R.R., Fortes M.R.S. *et al.* (2015). Prospecting major genes in dairy buffaloes. *BMC genomics* 16(1), 872.

Decker J.E., McKay S.D., Rolf M.M. *et al*. (2014) Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS genetics* 10(3).

Del Bo L., Polli M., Longeri M., Ceriotti G., Looft C., Barre-Dirie A., Dolf M. & Zanotti M. (2001) Genetic diversity among some cattle breeds in the Alpine area. *Journal of Animal Breeding and Genetics* 118(5), 317-325.

Elbeltagy A.R., Bertolini F., Fleming D.S., Van Goor A., Ashwell C.M., Schmidt C.J., Kugonza D.R., Lamont S.J. & Rothschild M.F. (2019) Natural Selection Footprints Among African Chicken Breeds and Village Ecotypes. *Frontiers in Genetics* 10, 376.

Excoffier L., Lischer H.E.L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10, 564-7.

Flori L., Fritz S., Jaffrézic F., Boussaha M., Gut I., Heath S., Foulley J.L. & Gautier M. (2009) The genome response to artificial selection: a case study in dairy cattle. *PloS one* 4(8).

Foll M. (2012) BayeScan v2. 1 user manual. *Ecology* 20, 1450-1462.

Gaspa G., Marras G., Sorbolini S., Ajmone-Marsan P., Williams J.L., Valentini A., Dimauro C. & Macciotta N.N.P. (2014) Genome-wide homozygosity in Italian Holstein cattle using HD panel. In

proceedings of the 10th World Congress of Genetics Applied to Livestock Production, 17 to 22 August, Vancouver, BC, Canada.

Gautier M., Laloë D. & Moazami-Goudarzi K. (2010) Insights into the genetic history of French cattle from dense SNP data on 47 worldwide breeds. *PloS one* 5(9).

Hagger C. (2005). Estimates of genetic diversity in the brown cattle population of Switzerland obtained from pedigree information. *Journal of Animal Breeding and Genetics* 122(6), 405-413.

Hou G.Y., Yuan Z.R., Gao X., Li J.Y., Gao H.J., Chen J.B. & Xu S.Z. (2010) Genetic polymorphisms of the CACNA2D1 gene and their association with carcass and meat quality traits in cattle. *Biochemical genetics* 48(9-10), 751-759.

Hulsegge I., Calus M., Hoving-Bolink R., Lopes M., Megens H.J. & Oldenbroek K. (2019) Impact of merging commercial breeding lines on the genetic diversity of Landrace pigs. *Genetics Selection Evolution* 51(1), 60.

Huson D.H., Bryant D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology Evolution* 23, 254–67.

Iso-Touru T., Sahana G., Guldbrandtsen B., Lund M.S. & Vilkki J. (2016) Genome-wide association analysis of milk yield traits in Nordic Red Cattle using imputed whole genome sequence variants. *BMC genetics* 17(1), 55.

Lee J.H., Li Y. & Kim J.J. (2012) Detection of QTL for carcass quality on chromosome 6 by exploiting linkage and linkage disequilibrium in Hanwoo. *Asian-Australasian journal of animal sciences* 25(1), 17.

Lin S., Zhang H., Hou Y., Liu L., Li W., Jiang J., Han B., Zhang S. & Sun D. (2019) SNV discovery and functional candidate gene identification for milk composition based on whole genome resequencing of Holstein bulls with extremely high and low breeding values. *PloS one* 14(8).

Lopdell T. J., Tiplady K., Couldrey C. et *al.* (2019). Multiple QTL underlie milk phenotypes at the CSF2RB locus. *Genetics Selection Evolution* 51(1), 3.

Marete A., Lund M.S., Boichard D. & Ramayo-Caldas Y. (2018) A system-based analysis of the genetic determinism of udder conformation and health phenotypes across three French dairy cattle breeds. *PLoS One* 13(7).

Marras G., Gaspa G., Sorbolini S., Dimauro C., Ajmone-Marsam P., Valentini A., Williams J.L. & Macciotta N.P.P. (2014). Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Animal Genetics* 46, 110–121.

Marques E., Nkrumah J.D., Sherman E.L. & Moore S.S. (2009) Polymorphisms in positional candidate genes on BTA14 and BTA26 affect carcass quality in beef cattle. *Journal of animal science* 87(8), 2475-2484.

Mastrangelo S., Tolone M., Gerlando R.D., Fontanesi L., Sardina M.T. & Portolano B. (2016) Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. *Animal* 10, 746–54.

Mastrangelo S., Ciani E., Ajmone-Marsan P.A.,et al. (2018a) Conservation status and historical relatedness of Italian cattle breeds. *Genetics Selection Evolution* 50:35.

Mastrangelo S., Sardina M.T., Tolone M., Di Gerlando R., Sutera A.M., Fontanesi L. & Portolano B. (2018b) Genome-wide identification of runs of homozygosity islands and associated genes in local dairy cattle breeds. *Animal* 12(12), 2480-2488.

Meng Q., Wang K., Liu X., Zhou H., Xu L., Wang Z. & Fang M. (2017) Identification of growth trait related genes in a Yorkshire purebred pig population by genome-wide association studies. *Asian-Australasian journal of animal sciences* 30(4), 462.

Metzger J., Karwath M., Tonda R., Beltran S., Águeda L., Gut M., Gut I.G. & Distl O. (2015) Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. BMC genomics 16(1), 764.

Mi H., Muruganujan A. & Thomas P.D. (2013) PANTHER in: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research* 41, D377–D386.

Mudadu M.A., Porto-Neto L.R., Mokry F.B. *et al.* (2016) Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. *Bmc Genomics* 17(1), 235.

Onzima R.B., Upadhyay M.R., Doekes H.P., Brito L., Bosse M., Kanis E., Groenen M.A.M. & Crooijmans R.P. (2018) Genome-wide characterization of selection signatures and runs of homozygosity in Ugandan goat breeds. *Frontiers in genetics* 9, 318.

Pant S.D., Karlskov-Mortensen P., Salicio S.C. *et al.* (2014) Genome-wide linkage disequilibrium linkage analysis (LDLA) of body fat traits in an F2 porcine model for human obesity. In 10th World Congress of Genetics Applied to Livestock Production.

Pant S.D., Schenkel F.S., Verschoor C.P., You Q., Kelton D.F., Moore S.S. & Karrow N.A. (2010) A principal component regression based genome wide analysis approach reveals the presence of a novel QTL on BTA7 for MAP resistance in holstein cattle. *Genomics* 95(3), 176-182.

Pausch H., Wang X., Jung S., Krogmeier D., Edel C., Emmerling R., Götz K.U. & Fries R. (2012) Identification of QTL for UV-protective eye area pigmentation in cattle by progeny phenotyping and genome-wide association analysis. *PloS one* 7(5).

Purcell S., Neale B., Todd-Brown K. *et al*. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81, 559–75.

Purfield D.C., Berry D.P., McParland S. & Bradley D.G. (2012) Runs of homozygosity and population history in cattle. *BMC genetics* 13(1), 70.

Purfield D.C., McParland S., Wall E. & Berry D.P. (2017) The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. PLoS ONE 12, e0176780.

Qanbari S., Gianola D., Hayes B., Schenkel F., Miller S., Moore S., Thaller G. & Simianer H. (2011) Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC genomics* 12(1), 318.

Qanbari S., Pausch H., Jansen S., Somel M., Strom T.M., Fries R., Nielsen R. & Simianer H. (2014) Classic selective sweeps revealed by massive sequencing in cattle. *PLoS genetics* 10(2).

Ramljak J., Bunevski G., Bytyqi H. *et al.*, (2018) Conservation of a domestic metapopulation structured into related and partly admixed strains. *Molecular ecology* 27(7), 1633-1650.

Raven L.A., Cocks B.G., Kemper K.E., Chamberlain A.J., Vander Jagt C.J., Goddard M.E. & Hayes B.J. (2016) Targeted imputation of sequence variants and gene expression profiling identifies twelve candidate genes associated with lactation volume, composition and calving interval in dairy cattle. *Mammalian genome* 27(1-2), 81-97.

Rothammer S., Seichter D., Förster M. & Medugorac I. (2013) A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC genomics* 14(1), 908.

Sahana G., Guldbrandtsen B., Thomsen B., Holm L.E., Panitz F., Brøndum R. F. & Lund M. S. (2014) Genome-wide association study using high-density single nucleotide polymorphism arrays

and whole-genome sequences for clinical mastitis traits in dairy cattle. *Journal of dairy science* 97(11), 7258-7275.

Senczuk G., Mastrangelo S., Ciani E. *et al*. (2020) The genetic heritage of Alpine local cattle breeds using genomic SNPs data. *Genetics Selection Evolution* Under review.

Signer-Hasler H., Burren A., Neuditschko M., Frischknecht M., Garrick D., Stricker C., Gredler B., Bapst B. & Flury C. (2017) Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genetics Selection Evolution* 49(1), 83.

Sorbolini S., Marras G., Gaspa G., Dimauro C., Cellesi M., Valentini A. & Macciotta N.P.P. (2015) Detection of selection signatures in Piemontese and Marchigiana cattle, two breeds with similar production aptitudes but different selection histories. *Genetics Selection Evolution* 47(1), 52.

Szmatoła T., Gurgul A., Jasielczuk I., Ząbek T., Ropka-Molik K., Litwińczuk Z. & Bugno-Poniewierska M. (2019) A Comprehensive Analysis of Runs of Homozygosity of Eleven Cattle Breeds Representing Different Production Types. *Animals* 9(12), 1024.

Valente T.S., Baldi F., Sant'Anna A.C., Albuquerque L.G. & da Costa M.J.R.P. (2016) Genome-Wide Association Study between single nucleotide polymorphisms and flight speed in Nellore cattle. *PloS one* 11(6).

Vallée A., Daures J., Van Arendonk J. A.M. & Bovenhuis H. (2016) Genome-wide association study for behavior, type traits, and muscular development in Charolais beef cattle. *Journal of animal science* 94(6), 2307-2316.

Van den Berg I., Fritz S., Rodriguez S., Rocha D., Boussaha M., Lund M.S. & Boichard D. (2014) Concordance analysis for QTL detection in dairy cattle: a case study of leg morphology. *Genetics Selection Evolution* 46(1), 31.

Wei C., Wang H., Liu G. *et al.* (2016) Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. *Scientific reports* 6, 26770.

Wu X., Lund M.S., Sahana G., Guldbrandtsen B., Sun D., Zhang Q. & Su G. (2015) Association analysis for udder health based on SNP-panel and sequence data in Danish Holsteins. *Genetics Selection Evolution* 47(1), 50.

Zhang Q., Guldbrandtsen B., Bosse M., Lund M.S. & Sahana G. (2015) Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genomics* 16(1), 542.

Zhang T., Gao H., Sahana G. *et al.* (2019) Genome-wide association studies revealed candidate genes for tail fat deposition and body size in the Hulun Buir sheep. *Journal of Animal Breeding and Genetics* 136(5), 362-370.

Zhao F., McParland S., Kearney F., Du L. & Berry D.P. (2015) Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genetics Selection Evolution* 47(1), 49.

Yoder D.M. & Lush J.L. (1937) A genetic history of the Brown Swiss cattle in the United States. *Journal of Heredity* 28(4), 154-160.

Table 1. Population type, name and code, number of animals (N), country of origin and source for the genotypic data of the Brown breeds used in the contrasted groups.

Group	Breed	Breed code	N	Country	Data source
1	Original Braunvieh	ORBR	20	Switzerland	Decker *et al.*, 2014
	Braunvieh	BRHV	35	Switzerland/Germany	Ramljak *et al.*, 2018
2	Brown Swiss	BRSW	19	Americas	Gautier *et al.*, 2010
	Italian Brown	ITBR	32	Italy	Mastrangelo *et al.*, 2018

Table 2. Genetic diversity indices in the four Brown populations: observed (Ho) and expected (He) heterozygosity, the genomic inbreeding (F_{HOM}) and standard deviation (SD).

Breed	Ho±SD	He±SD	F_{HOM}±SD
ORBR	0.360 ± 0.151	0.353 ± 0.134	0.014 ± 0.018
BRHV	0.359 ± 0.165	0.351 ± 0.137	0.020 ± 0.048
BRWS	0.343 ± 0.179	0.328 ± 0.153	0.055 ± 0.040
ITBR	0.332 ± 0.173	0.323 ± 0.155	0.085 ± 0.035

Table 3. Results of the BayeScan analysis showing outlier SNPs identified between original (ORBR and BRHV) and modern (BRSW and ITBR) Brown cattle populations.

BTA	Pos	SNP name	F_{ST}	Gene
2	113,621,322	ARS-BFGL-NGS-40251	0.305	
2	113,650,536	Hapmap51816-BTA-20000	0.296	*NYAP2*
2	130,360,932	ARS-BFGL-NGS-43912	0.285	*ZBTB40*
4	38,606,134	BTB-00178575	0.266	*CACNA2D1*
4	75,566,919	ARS-BFGL-NGS-2431	0.279	-
5	15,943,853	ARS-BFGL-NGS-20187	0.251	*MGAT4C*
5	96,563,734	Hapmap28380-BTA-74649	0.245	-
6	18,195,932	ARS-BFGL-NGS-91848	0.262	*DKK2*
6	31,738,269	Hapmap33117-BTC-032493	0.253	*GRID2*
6	31,761,997	Hapmap30962-BTC-032558	0.248	*GRID2*
6	39,437,165	Hapmap33744-BTC-050901	0.245	-
6	39,573,948	BTB-00252917	0.275	-
6	39,667,998	BTB-00406718	0.273	-
6	46,364,510	ARS-BFGL-NGS-31387	0.271	*STIM2*
6	47,983,800	Hapmap52342-rs29016265	0.254	*STIM2*
6	70,387,975	Hapmap32219-BTC-042322	0.249	-
6	90,103,125	BTA-86242-no-rs	0.248	-
6	109,282,605	Hapmap32834-BTA-149103	0.255	-
8	111,318,888	ARS-BFGL-NGS-89567	0.258	-
11	27,312,691	Hapmap23233-BTA-88497	0.257	-
13	22,505,137	Hapmap35027-BES1_Contig330_454	0.284	-
14	4,534,257	Hapmap30091-BTC-005211	0.244	*FAM135B*
14	48,435,848	ARS-BFGL-NGS-70879	0.253	-
17	47,697,987	Hapmap38389-BTA-17091	0.284	*TMEM132D*
17	47,740,843	Hapmap42713-BTA-87646	0.283	*TMEM132D*
18	41,475,285	Hapmap42198-BTA-39980	0.246	-
18	54,783,476	ARS-BFGL-NGS-70161	0.246	-
20	5,665,364	BTB-01104222	0.295	*CPEB4*
20	70,694,677	ARS-BFGL-NGS-98971	0.255	-
22	11,818,533	ARS-BFGL-NGS-24627	0.252	*XYLB*
22	32,711,432	Hapmap51256-BTA-54100	0.282	*FAM19A4*
27	43,602,071	Hapmap47470-BTA-101224	0.268	*ZNF385D*
28	17,475,241	BTA-63612-no-rs	0.286	-
28	42,481,492	ARS-BFGL-NGS-81842	0.259	*PTPN20*
29	13,770,679	BTB-01423078	0.247	-

Table 4. List of genomic regions of extended homozygosity (ROHs islands) identified in original (ORBR and BRHV) (group 1) and modern (BRSW and ITBR) (group 2) Brown cattle populations.

Group	BTA	n of SNPs	n of Genes	Start bp	End bp	Length (Mb)	Genes
1	11	68	59	65,681,864	71,263,018	5.58	*ETAA1, LOC101904865, C1D, WDR92, PNO1, PPP3R1, CNRIP1, LOC104973409, PLEK, FBXO48, APLF, PROKR1, ARHGAP25, BMP10, LOC509961, LOC104973410, GKN2, GKN1, ANTXR1, GFPT1, NFU1, AAK1, LOC787229, LOC104973411, ANXA4, GMCL1, SNRNP27, MXD1, ASPRV1, LOC104973412, LOC101905499, PCBP1, LOC107132938, C11H2orf42, TIA1, PCYOX1, LOC107132939, SNRPG, EHD3, CAPN14, GALNT14, CAPN13, LCLAT1, LOC101905676, LBH, YPEL5, LOC101905873, LOC104968430, ALK, CLIP4, C11H2orf71, FAM179A, LOC107132940, WDR43, TRMT61B, SPDYA, PPP1CB, PLB1*
	26	17	36	21,539,987	22,954,453	1.41	*LOC104975965, PAX2, SLF2, SEMA4G, MRPL43, C26H10orf2, LZTS2, PDZD7, SFXN3, KAZALD1, TLX1, LOC107131879, LBX1, LOC100847491, BTRC, LOC104975969, LOC783067, POLL, DPCD, FBXW4, LOC101908075, FGF8 NPM3, MGEA5, KCNIP2, C26H10orf76, HPS6, LOC104975987, LDB1, PPRC1, NOLC1, LOC101902227, LOC785229, ELOVL3, PITX3, GBF1*
2	5	56	43	75,086,818	78,560,464	3.49	*CACNG2, IFT27, LOC101906363, LOC101906435, PVALB, LOC107132513, NCF4, CSF2RB, LOC788541, TEX33, TST, MPST, KCTD17, TMPRSS6, IL2RB, LOC510185, C1QTNF6, SSTR3, RAC2, MIR1835, CYTH4, LOC107132508, ELFN2, LOC107132514, MFNG, CARD10, USP18, ALG10, SYT10, LOC107132509, PKP2, YARS2, DNM1L, LOC101907810, FGD4, BICD1, LOC782092, KIAA1551, LOC100137780, LOC104972515, AMN1, ETFBKMT, DENND5B*
	5	9	3	79,334,574	80,235,852	0.90	*TMTC1, LOC518980, ERGIC2*
	6	21	7	86,399,795	87,818,793	1.42	*LOC112447099, SLC4A4, LOC782958, GC, NPFFR2, ADAMTS3, TRNAC-GCA*

Figures

Figure 1 Principal components analysis for the genetic differentiation among the four Brown cattle breeds. Original Braunvieh = ORBR; Braunvieh = BRHV; Brown Swiss = BRSW; Italian Brown = ITBR.

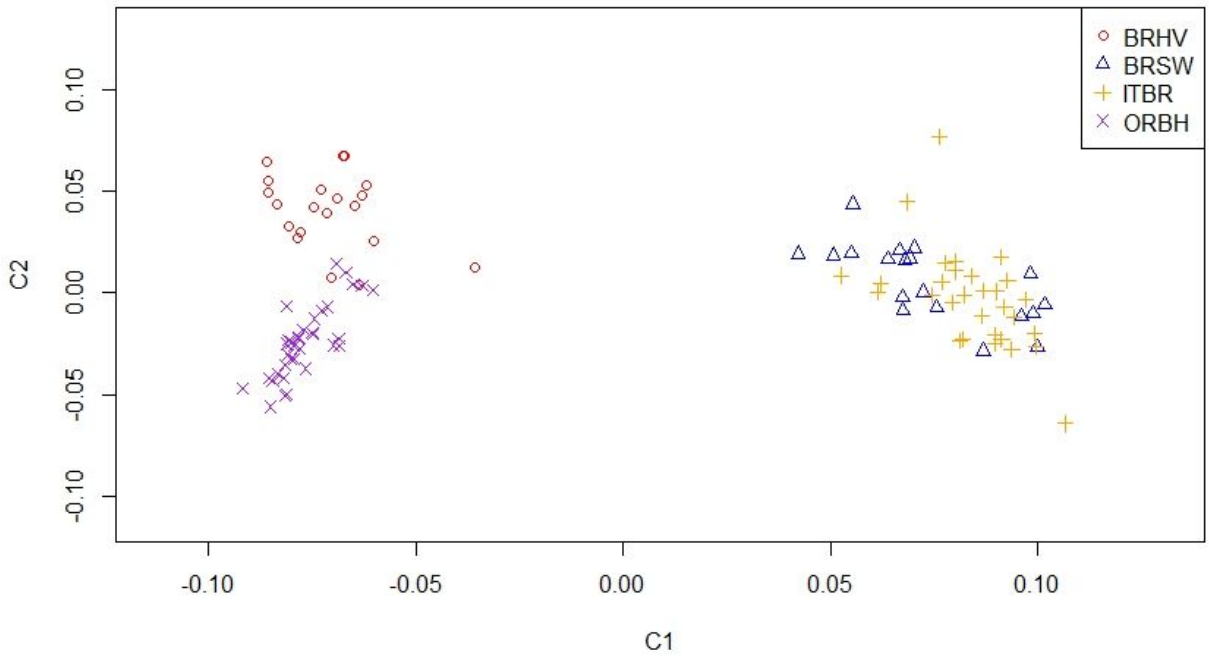


Figure 2 Manhattan plot of the pairwise genome-wide autosomal F_{ST} analysis generated by BayeScan between group 1 (ORBR/BRHV) and group 2 (BRSW/ITBR). The red line indicates the threshold of F_{ST} value (0.243) corresponding to 0.999 of the F_{ST} percentile distribution.

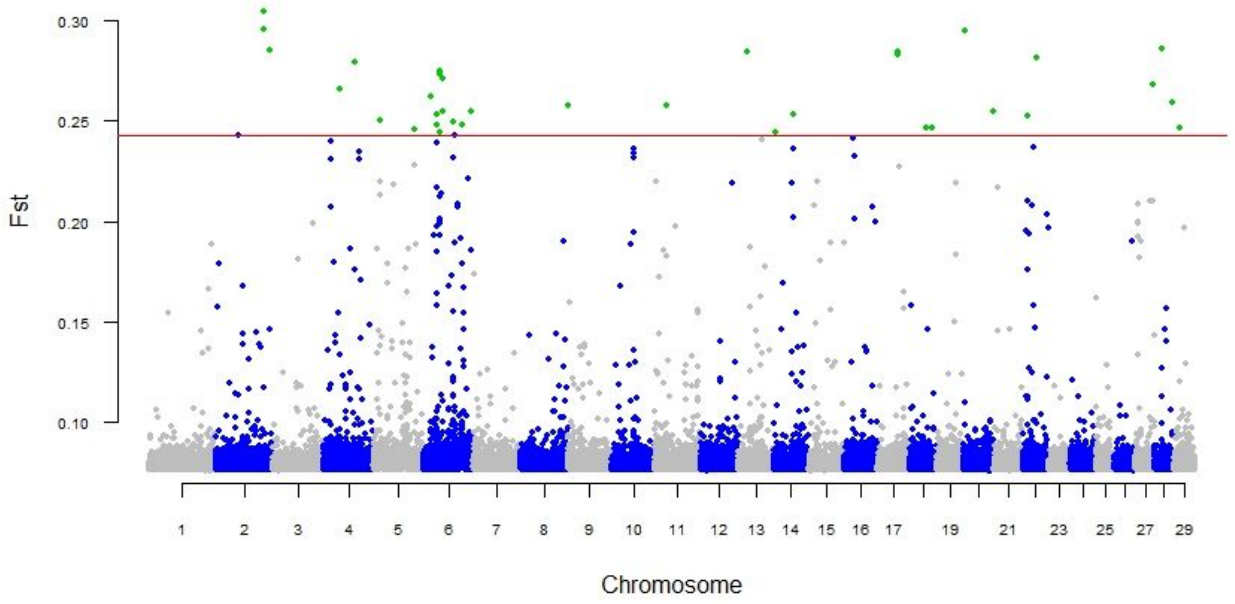
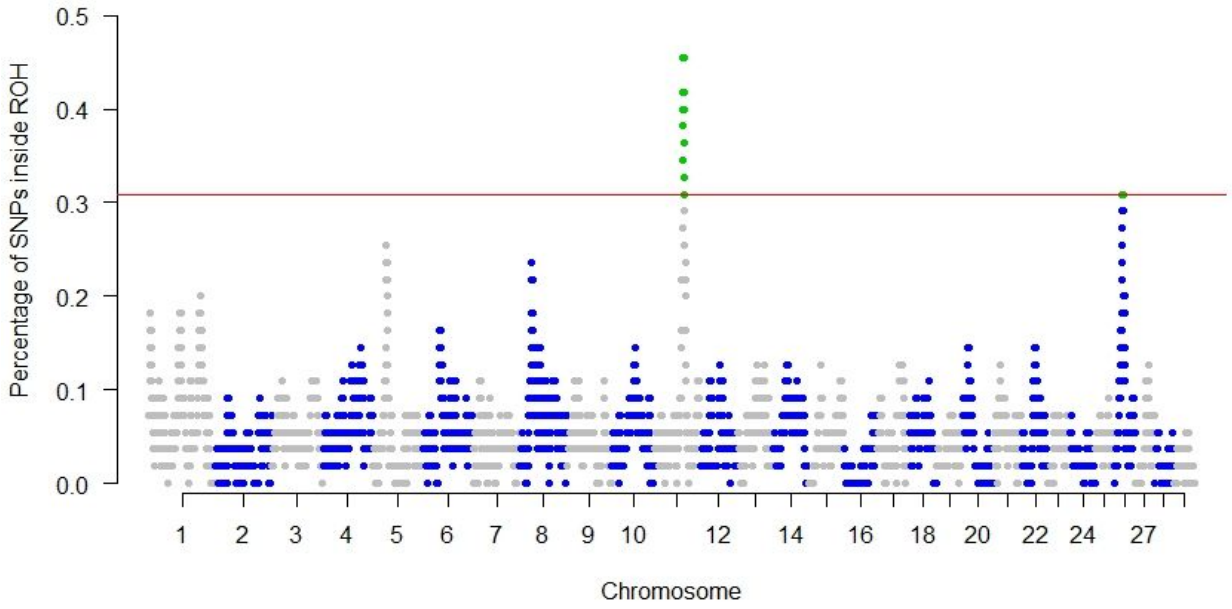
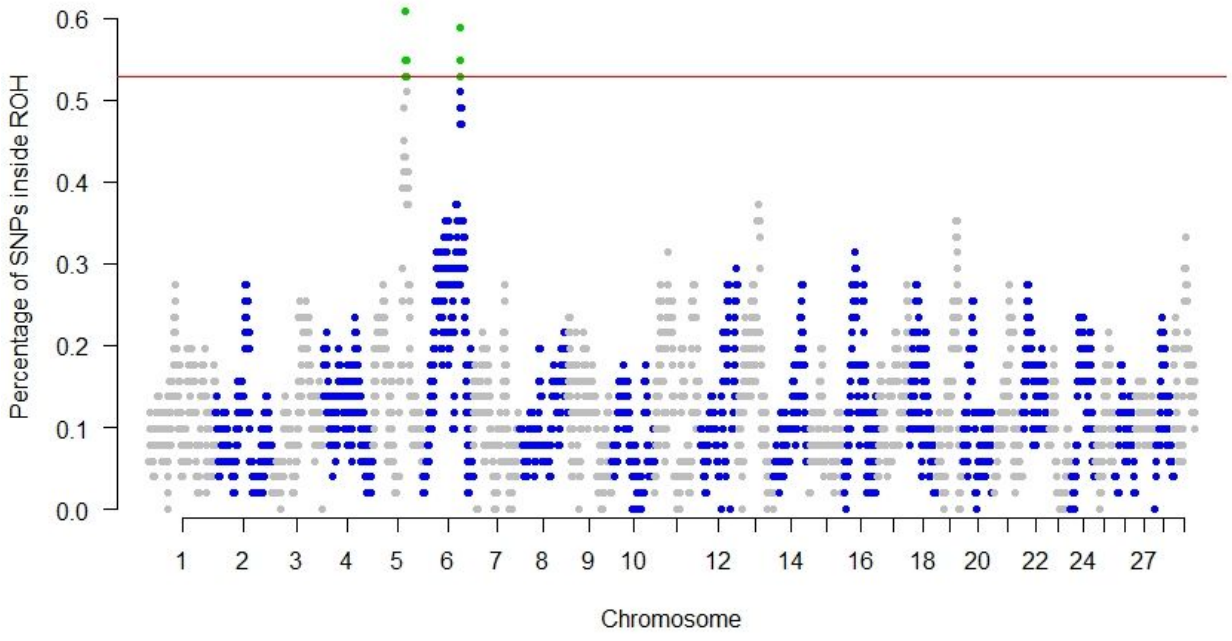


Figure 3 Manhattan plots of regions of homozygosity for (A) the group 1 (ORBR/BRHV) and (B) the group 2 (BRSW/ITBR). Thresholds used to detect high-homozygosity regions are indicated with a red line.

A)



B)



Supporting information

Figure S1 Neighbor-Joining tree relating all the individuals. The tree was constructed using allele-sharing distances.

Figure S2 Box plots of the distribution of ROH into five classes of length (1 to ≤2 Mb, 2 to ≤4 Mb, 4 to ≤8 Mb, 8 to ≤16 Mb and >16 Mb) for the two Brown cattle groups.

Table S1. Genetic differentiation among the four Brown cattle breeds measured using F_{ST}

Table S2. Descriptive statistics of the number and the frequency distribution of runs of homozygosity (ROH) in five different size classes (Mb) for the two Brown cattle groups.

Table S3. Comparison among overlapped runs of homozygosity (ROHs) islands here detected and those reported in previous studies.

Table S4. Gene ontology (GO) terms enriched (P<0.05) based on runs of homozygosity islands and number of involved genes (n) in the two Brown cattle groups.

Table S5. Bovine QTL mapped within the ROHs islands identified within the two Brown cattle groups.

Identification of genomic regions involved in the phenotypic differences between original and modern Brown populations

A. Moscarelli[1], M. T. Sardina[1], M. Cassandro[2], E. Ciani[3], F. Pilla[4], G. Senczuck[4], B. Portolano[1] and S. Mastrangelo[1]*

[1] Dipartimento Scienze Agrarie, Alimentari e Forestali, University of Palermo, 90128 Palermo, Italy

[2] Dipartimento di Agronomia Animali Alimenti Risorse naturali e Ambiente, University of Padova, 35020 Legnaro, Italy.

[3] Dipartimento di Bioscienze, Biotecnologie e Biofarmaceutica, University of Bari, 70124 Bari, Italy

[4] Dipartimento Agricoltura, Ambiente e Alimenti, University of Molise, 86100 Campobasso, Italy

*Corresponding author

Salvatore Mastrangelo

salvatore.mastrangelo@unipa.it

Summary

Identifying genomic regions involved in the differences between breeds can provide information on genes that are under the influence of both artificial and natural selection. The aim of this study was to identify genomic regions involved in phenotypic differentiation among four different Brown populations (original *vs* modern) and to characterize the distribution of runs of homozygosity (ROHs) islands using the Illumina Bovine SNP50 BeadChip genotyping data. After quality control, 34,735 SNPs and 106 animals were retained for the analyses. Larger heterogeneity was highlighted for the original breeds. Patterns of genetic differentiation, multidimensional scaling, and the neighboring joining tree distinguished the modern breeds from the original populations. The F_{ST} - outlier identified several genes involved in many phenotypic differences between the two groups, such as stature and growth, behavior and adaptability to local environments. The ROH islands within both the original and the modern groups overlapped with QTL associated with relevant traits. In modern Brown, ROH islands harbored candidate genes associated with milk production traits, in evident agreement with the artificial selection conducted to improve this trait in these populations. In original Brown, we identified candidate genes related with fat deposition, confirming that breeding strategies for the original Brown populations aimed to produce dual-purpose animals. Our study highlighted the presence of several genomic regions which vary between Brown populations, in line with their different breeding histories.

Keywords Brown cattle, genetic diversity, F_{ST}, runs of homozygosity, candidate genes

Introduction

In cattle, natural and artificial selection has resulted in divergent breeds. In fact, the continuously increasing demand for work, milk and meat has enhanced between-population differences over the centuries (Zhao *et al*., 2015). An interesting situation regarding this divergence among populations is represented by the Brown cattle. The Brown cattle is derived from populations used in valleys

and mountain slopes of Switzerland since before historic records began. One theory says that this breed goes back to oriental origins, having been introduced into Central Europe from the steppes and valleys of Western Asia. Moreover, cattle bones found in the ruins of the Swiss Lake Dwellers indicate that a type of cattle, apparently closely related to actual Brown cattle, existed during the Bronze Age in Switzerland (Del Bo *et al.,* 2002). Nowadays, the Brown cattle reared in Europe can be grouped into three different populations: the Original Braunvieh, the Braunvieh, and the Brown Swiss (Hagger, 2005). The Original Braunvieh is a dual-purpose cattle breed reared in Switzerland and it represents the original population of Brown cattle (Bhati *et al*., 2020). This population is ancestral to the Brown Swiss; in fact, few individuals of the Original Braunvieh (about 170) originating from the mountain tops of Northeast Switzerland had been imported in the USA during the end of the 19th century giving rise to the current Brown Swiss, which is today one of the most widespread dairy breed in the world (Yoder & Lush 1937). Between 1967 and 1998, a crossbreeding between Original Braunvieh and the American Brown Swiss was conducted through artificial insemination, leading to the current Braunvieh population; this is mainly spread over the Alpine regions of Austria, Germany, Italy, and Switzerland.

Genetic drift, geographical barriers, different environments, and local management practices have affected the divergences among the original and modern Brown populations. Moreover, natural selection and modern breeding approaches, while acting on adaptation, morphological and production traits, also contributed to shape the genetic structure of these populations.

The availability of single nucleotide polymorphism (SNP) panels, consisting of thousands of markers, has greatly improved the power of genome-wide studies for a deep investigation of genetic diversity between and within population(s), allowing the identification of highly differentiated genomic regions between cattle breeds (Flori *et al*., 2009; Qanbari *et al*., 2011; Signer-Hasler *et al.*, 2017; Mastrangelo *et al.*, 2018a). The analytical approach, using the allele frequency-based inter-population genetic differentiation (F_{ST}) and intra-population runs of homozygosity (ROHs), has been recently applied in several livestock species for the identification of genomic regions involved

in phenotypic differences (Cesarani *et al.*, 2018; Onzima *et al.*, 2018; Elbeltagy *et al.*, 2019; Hulsegge *et al.*, 2019; Szmatoła *et al.*, 2019). Finding links between phenotypic and genotypic differences is of great importance in order to gain a better understanding of genetic mechanisms underpinning traits of interest, and it offers the opportunity to improve the efficiency of animal breeding through directed selection on favorable alleles (Rothammer *et al.*, 2013). Therefore, the objective of the present study was to identify differentiated genomic regions among the Brown populations (original *vs* modern) and to characterize the distribution of runs of homozygosity islands using the Illumina Bovine SNP50 BeadChip genotyping data, which may provide insights into the mechanisms underlying their genomic differences. We also checked if the genomic regions that were most associated with ROHs overlapped with reported quantitative traits loci (QTL).

Materials and methods

Sampling, genotyping and quality control

Overall, 106 animals belonging to four different Brown cattle populations were used: Original Braunvieh (ORBR), Braunvieh (BRHV), Brown Swiss (BRSW) and Italian Brown (ITBR). These breeds present differences in both phenotypic and production traits. For example, in the ORBR, selection breeding schemes are almost absent, whereas the Brown Swiss is the second largest producer of milk after the Holstein breed. Nevertheless, the original populations, such as the BRHV, are adapted to the harshness of mountain areas because of their good grazing characteristics, and are resistant to environmental conditions. These populations shows also a smaller body size compared to selected breeds.

Detailed information about the breeds, the samples and the grouping into original and modern cattle are reported in Table 1. All individuals were genotyped using the Bovine SNP50K BeadChip. SNPs were mapped using the *Bos taurus* ARS-UCD1.2 genome assembly. The markers were filtered using PLINK 1.07 (Purcell *et al.*, 2007) and considering only SNPs mapped in autosomal

chromosomes. SNPs with a minor allele frequency (MAF) lower than 0.01 and call rate lower than 95%, as well as individuals with missing genotyping rate lower than 90%, were removed.

Comparison of breeds

Pairwise genetic relationships were estimated to evaluate population substructure using identity-by-state (IBS) genetic distances calculated by PLINK 1.07 (Purcell *et al.*, 2007) and graphically represented by multidimensional scaling (MDS). PLINK was also used to estimate observed (Ho) and expected (He) heterozygosity, and the genomic inbreeding, which is based on the difference between the observed and expected numbers of homozygous genotypes (F_{HOM}). The ARLEQUIN version 3.5.2.2 (Excoffier & Lischer, 2010) was used to estimate population relatedness using pairwise estimates of F_{ST} among all four breeds. A neighbor-joining tree was constructed based on individual allele-sharing distances (--distance 1-IBS in PLINK) and visualized using SPLITSTREE (Huson & Bryant, 2006).

F_{ST}-outlier analysis

The F_{ST}-outlier approach implemented in the BayeScan software (Foll, 2012) was adopted to identify loci involved in the differentiation between the considered groups. BayeScan analyses comprised 20 pilot runs of 5,000 iterations, a burn-in of 50,000 iterations, a thinning interval of 10 (5,000 iterations were used for the estimation of posterior odds) with a resulting total number of 100,000 iterations. To control the number of false positives, significant markers were defined by applying a *q*-value threshold of 0.05 and using the 0.999 SNPs of F_{ST} percentile distribution.

Runs of homozygosity islands

Runs of homozygosity (ROHs) were estimated for each sample using PLINK 1.07 (Purcell *et al.*, 2007). The minimum length that constituted the ROH was set to 1 Mb. The following criteria were used to define the ROHs: (i) one missing SNP was allowed in the ROH and up to one possible heterozygous genotype; (ii) the minimum number of consecutive SNPs that constituted the ROH was set to 30; (iii) minimum density of 1 SNP every 100 kb; (iv) maximum gap between

consecutive SNPs of 1 Mb. The average length of ROH (L_{ROH}) was estimated. ROH were grouped into five classes of length (1 to ≤2 Mb, 2 to ≤4 Mb, 4 to ≤8 Mb, 8 to ≤16 Mb and >16 Mb) (Marras *et al*., 2014). The number and frequency of ROH within each ROH length category for the two groups were also determined. The percentage of SNPs residing within the ROH was estimated by counting the number of times that each SNP appeared in a ROH and by dividing that number by the number of animals in each group. To identify the genomic regions of "high homozygosity", also called ROHs islands, the top 0.999 SNPs of the percentile distribution of the locus homozygosity range within each group were selected.

Gene ontology enrichment

Genomic regions detected by the two statistical approaches were interrogated for genes annotated to the *Bos taurus* genome assembly ARS-UCD1.2 using Genome Data Viewer (https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_002263795.1) provided by NCBI.

The genes within ROH islands were further analyzed with the PANTHER Classification System (Mi *et al.*, 2013) to identify significant (P ≤0.05) gene ontology (GO) terms. Moreover, to investigate the biological function of the annotated genes, an accurate literature search was also conducted.

Finally, using "JBrowse" (Buels *et al.*, 2016), an open source JavaScript-based genome browser, available on National Animal Genome Research Program (https://www.animalgenome.org/jbrowse/), we checked if the genomic regions of high homozygosity overlapped with reported quantitative traits loci (QTL).

Results

Comparison of breeds

After quality control, 34,735 SNPs and 106 animals were retained for the analyses. Genetic diversity indices (Ho and He) and inbreeding coefficient (F_{HOM}), which are key parameters in the genetic management of populations, were used to determine the levels of genetic variability in the four Brown breeds. The original populations (ORBR and BRHV) displayed the highest genetic diversity, whereas the lowest value was found in modern populations (BRSW and ITBR) (Table 2). To examine and visualize the genetic relationships among the four Brown cattle populations, we used an MDS plot of the pairwise identity-by-state distance (Figure 1). The results showed that the populations formed two different clusters. The first dimension (C1) separated BRSW and ITBR from ORBR and BRHV. In agreement with MDS results, the neighbor joining (NJ) tree based on allele sharing distance (ASD) separated individuals according to their population of origin (Fig. S1). Genetic differentiation between all pairs of populations estimated by F_{ST} statistics are also reported in Table S1, and substantially confirms results from MDS and NJ tree. In particular, F_{ST} was low between BRSW and ITBR (0.006) or between ORBR and BRHV (0.016).

All these results were used to categorize the four Brown populations into two contrasting groups for comparative analysis: original (ORBR and BRHV) *versus* modern (BRSW and ITBR) cattle.

F_{ST} analysis

Results from the Bayesian population differentiation approach identified a total of 35 outlier SNPs between ORBR-BRHV and BRSW-ITBR (Table 3). These SNPs were identified on 15 different chromosomes (BTA). Most of these markers were located far apart from each other. Manhattan plots of F_{ST} values are reported in Figure 2. The locus with the highest value (0.305) was ARS-BFGL-NGS-40251 on BTA02. Only three outlier genomic regions showed adjacent SNPs on BTA02, BTA06, BTA17, in which *NYAP2*, *GRID2* and *TMEM132D* are mapped, respectively (Table 3). Among the outlier SNPs, 11 markers (31%) were located on BTA6. A total of 17 markers that exceeded the significance threshold were mapped within 14 protein-coding genes.

Runs of homozygosity

The total number of detected ROHs exhibited variation between groups, with the modern group having the highest value (1,725). Moreover, the modern Brown group had also the highest average length of ROHs (8.54 Mb *vs* 7.17 Mb). The frequency and the distribution of the number of ROHs in the different classes was similar (Table S2 and Fig. S2). The most represented ROH classes in both groups was that of length ranging from 4 to ≤8 Mb, representing 48% and 44% in original and modern groups, respectively. The largest number of ROH in the class of highest length (>16 Mb) was observed in modern (BRSW and ITBR) cattle (Fig. S2). To identify the genomic regions that were most associated with ROHs, the top 0.999 SNPs of the percentile distribution of the locus homozygosity range were chosen as an indicator of a possible ROH hotspot in the genome. The analysis was carried out within the same two groups of the pairwise comparison reported above for the F_{ST}-outlier approach. Table 4 provides the chromosome position, the start and the end of ROH regions and the relative genes list. Although the distribution of the ROHs was relatively balanced and the signals were moderate in height, we found a few outstanding peaks with a high occurrence of ROHs (Figure 3). In total, five genomic regions that frequently appeared in the ROHs were identified in the two Brown groups (Table 4). The ROHs islands were found on BTA11 and BTA26 in the ORBR/BRHV group (Figure 3A), and on BTA05 and BTA06 in the BRSW/ITBR group (Figure 3B), ranging in length from 0.90 Mb (BRSW/ITBR; 9 consecutive SNPs on BTA05) to 5.58 Mb (ORBR/BRHV; 68 consecutive SNPs on BTA11).

Within the ROHs islands reported above, we identified several known genes, together with uncharacterized genes (LOC) (Table 4). Some of the detected islands, such as on BTA5 in the modern Brown breeds, contained only three annotated genes. Moreover, some of these genomic regions overlapped with ROHs islands found in other studies (Table S3).

Gene ontology enrichment

The results of the GO enrichment analysis (Table S4) showed multiple categories that were statistically significant ($P \leqslant 0.05$). The genes within ROHs islands encompass a wide spectrum of molecular functions, biological processes, and cellular components. A PANTHER gene list analysis

revealed a high percentage of genes involved in the following categories: cellular process (GO:0009987), catalytic activity (GO:0003824), binding (GO:0005488), metabolic process (GO:0008152), biological regulation (GO:0065007), localization (GO:0051179), response to stimulus (GO:0050896). Finally, we also checked if the ROHs overlapped with previously reported quantitative traits loci (QTL). Overall, the ROHs islands overlap with known QTL regions associated with several traits (Table S5). We identified QTL associated with milk composition and milk fat content, life history traits, energy efficiency association and intramuscular fat. For example, the ROH island in BTA06 overlap with QTL associated with somatic cell score, clinical mastitis, length of productive life and calf size.

Discussion

Identifying genomic regions involved in the differences between breeds can provide information on genes that are under the influence of both artificial and natural selection, and thus, can help the identification of beneficial mutations and underlying biological pathways for economically important traits (Zhao *et al.*, 2015). Notably, the inter-population F_{ST}-outlier test compares the variation of allele frequencies within and between groups, and the locus that shows the largest differences in allele frequencies between populations is assumed to be a signal of selection (Qanbari *et al.*, 2011). Intra-population ROHs analysis has been used to identify regions that have an unfavorable effect on a phenotype when they are in the homozygous state but also to detect associations between traits of economic interest and genes present in these regions (e.g. Zhang *et al.*, 2015; Mastrangelo *et al.*, 2018b). In this study, we applied the two above-mentioned genome-scan analytical approaches to identify differentiated genomic regions between original and modern Brown cattle populations and to characterize the distribution of runs of homozygosity islands within these populations.

Several studies investigated the genomic diversity and population structure of the Brown populations using either pedigree (Hagger, 2005) or microarray data (Signer-Hasler *et al.*, 2017;

Senczuk *et al.*, 2020), but to the best of our knowledge, this is the first study on the identification of differentiated genomic regions between original and modern Brown cattle populations.

Patterns of genetic differentiation, multidimensional scaling, and the neighboring joining tree showed a marked divergence between the Brown populations investigated here, and clearly distinguishing the modern breeds from the original populations. Similar patterns were also observed in previous studies (e.g. Del Bo *et al.*, 2011; Signer-Hasler *et al.*, 2017; Senczuk *et al.*, 2020). The observed genetic differentiation among breeds was in accordance with their geographical and historical origins. Moreover, genetic drift and the different selective schemes to which these populations have been subjected could have further contributed to the observed scenario. We also found that modern Brown cattle present lower values of genetic diversity, related with the high selection pressure on these breeds for dairy traits. Moreover, this group showed the highest frequency of ROH > 16 Mb (Fig. S2) as also reported by Cesarani *et al*. (2018), and related to a more recent inbreeding (Mastrangelo *et al*., 2016). On the other hand, the original populations consists of traditional pure-bred animals and has maintained substantial genetic diversity, due to the use of a relatively high number of natural service sires and relatively weak genetic selection for milk yield (Hagger, 2005).

The genome-wide distribution of pair-wise F_{ST} values was used to identify highly differentiated SNPs between the two groups of Brown cattle considered in this study. In general, the F_{ST}-outlier approach confirmed what already observed based on phenotypic differences between original and modern Brown cattle: the main genetic differences between the two groups lie in morphological traits and adaptability to local environments. Indeed, the results highlighted growth-related candidate genes such as *NYAP2* (Meng *et al.*, 2017), *CACNA2D1* (Hou *et al.*, 2010), *GRID2* (Lee *et al.*, 2012), *STIM2* and *CPEB4* (Mudadu *et al.*, 2016), while the outlier SNP in the BTA28 mapped on *PTPN20*, a candidate gene involved in the genetic determinism of udder conformation in cattle, an important trait under selection in the specialized dairy cattle breeds (Marete *et al.*, 2018). These results are consistent with the knowledge that breeding strategies, for the original Brown

populations, aimed to produce dual-purpose animals. Moreover, this comparison identified a gene related with a local adaptation trait, i.e. exposure to UV radiation (*DKK2*) (Pausch *et al.* 2012). In transhumant systems, the animals belonging to the original Brown populations graze at alpine pastures (between 1000 and 2400 m above sea level) during the summer months and return to the stables for the winter months (Bhati *et al*., 2020). Therefore, genes related with solar radiation are important for these breeds, that are exposed to increased summer solar ultraviolet radiation. In addition, on BTA17, *TMEM132D* was retrieved, a gene involved in behavioral traits such as reduced fear in humans and aggressiveness, consequently to human-driven selection (Qanbari *et al.*, 2014; Vallée *et al.*, 2016). All the above genes may very well explain some of the phenotypic difference between the two Brown groups. In our study, the genetic differentiation analysis between groups did not detect SNPs located in genomic regions known to contain genes associated with milk production traits, probably because under selection within both groups, even if with different selection pressure. Consistently with our results, Signer-Hasler *et al.* (2017), in a study on selection signatures with the same SNP array, which also involved original and modern Brown populations, reported that potential candidate genes based on phenotypic evidence, such as *DGAT1* or casein clusters, have not left any recognizable signal in these populations. The authors interpreted they results as the consequence of a low SNP density of the adopted SNP array in these regions.

As mentioned above, another method to identify candidate genomic regions, based on intra-population analysis, is the identification of runs of homozygosity islands. Since ROHs are normally abundant in regions under positive selection, their accumulation at specific loci has been used to identify genomic regions that reflect directional selection in livestock species (Mastrangelo *et al.*, 2018b; Purfield *et al.*, 2017; Metzger *et al.*, 2015). Moreover, several studies reported the occurrence of ROHs islands within QTLs for important production traits in cattle (Purfield *et al.* 2012; Szmatoła *et al.*, 2019). In this study, the ROH islands detected in the Brown populations overlapped with reported QTLs associated with relevant traits, such as milk composition and milk fat content, growth, body size, and length of productive life. Since ROH islands spanned many

candidate genes displaying several functions, in our analysis of the published literature, we mainly focused on genes related to livestock breeding traits.

In the original Brown group (ORBR/BRHV), at the beginning of the island on BTA11, we found the *ETAA1* gene, also reported within a ROH island in Sardo-Bruna cattle (Cesarani *et al.*, 2018), a local Italian Brown-derived population, thus confirming previously reported data. In fact, the overlapping ROHs islands among studies (Table S3) provided good evidence that they are not artifacts, but potential genomic regions affected by selection (Mastrangelo *et al.*, 2018). The island on BTA11 also included some known candidate genes (*C1D*, *WDR92, PNO1, PPP3R1, ARHGAP25*) related with fat deposition (Zhang *et al.*, 2019; Sorbolini *et al.*, 2015; Pant *et al.*, 2014), associated with high-altitude adaptation (*PPP3R1*) (Wei *et al.*, 2016), involved in immune and inflammatory response (*PLEK*) (Cremonesi *et al.*, 2012) and influencing the expression of flight speed (*ANTXR1*), used as a measure of temperament (Valente *et al.*, 2016). Based on these functions, some of these genes could be related with adaptation to cold climate, a characteristic of the original Brown populations, that commonly display better ability to adapt to harsh environments compared with the specialized modern breeds. These results confirmed the importance of the local populations as reservoirs of biodiversity and as models for studying the genetic basis of adaptability (Cesarani *et al*., 2018). Moreover, this island on BTA11 identified several candidates genes associated with fertility traits (*EHD3* and *PCBP1*). This may explain, at least partly, why local populations are generally characterized by high longevity. Finally, this island showed genes related with meat quality (*CAPN14*, *CAPN13*, *LBH* and *LCLAT1*), as previously reported within a significant selection signature in Original Braunvieh (Bhati *et al*., 2020; Rothammer *et al.*, 2013; Signer-Hasler *et al.*, 2017) and in line with the dual-purpose characteristics of these animals. Similarly, the ROH island on BTA26, also identified in the Italian Holstein cattle (Gaspa *et al.*, 2014), showed genes involved in bovine leg conformation (*BTRC* and *LBX1*) (Van den Berg *et at.*, 2014), related with lipid metabolism pathway (*FGF8* and *GBF1*) (Marques *et al.*, 2009; Iso-Touru *et al.*, 2016) and with milk protein and fat traits in cattle (*MGEA5*) (Lin *et al.*, 2019). Moreover, a

strong signature of selection (21-23 Mb) was reported in literature on this genomic region in Original Braunvieh (Signer-Hasler *et al.*, 2017). On the other side, in the ROH island on BTA5 of modern Brown, several candidate genes associated with milk production traits in cattle (*CSF2RB*, *NCF4* and *MPST*) (Lopdell *et al.*, 2019; Raven *et al.*, 2015) and in dairy buffaloes (*TMPRSS6*, *C1QTNF6*, *SSTR3*, *RAC2*, *CYTH4*) (De Camargo *et al.*, 2015) are mapped, in evident agreement with the artificial selection conducted to improve this traits in these populations. In the smallest ROH island in BTA5 we detected two genes, *TMTC1* and *ERIC2* associated with the *Mycobacterium avium* spp. *paratuberculosis* infection status (Pant *et al.*, 2010). Finally, of interest is the ROH island on BTA6 between about 86 and 87 Mb, a genomic region known to be functionally associated with traits related with mastitis, the disease that causes the greatest economic losses to the dairy industry worldwide. In fact, within this region, several candidate genes related with this disease, such as *SLC4A4* (Wu *et al.*, 2015), *GC* and *NPFFR2* (Sahana *et al.*, 2014) were retrieved.

Conclusion

Through the present study, we described some genetic differences between original and modern Brown populations. The different approaches used gave a picture of genetic relationships between original and modern breeds. As expected, larger heterogeneity was highlighted for the original breeds. Our study highlighted the presence of several genomic regions which vary between Brown populations, in line with their different breeding histories, and demonstrating that selection, other than genetic drift, contributed to the genetic differentiation among the original and modern breeds. The modern Brown has undergone the influences of intense artificial selection for milk production, losing some typical characteristics of the original Brown, like the dual propose attitude, establishing itself as a cosmopolitan breed for milk production. The original Brown maintains dual-purpose attitude, and have genetic characteristics linked to the marginal environments in which they are

reared and where continues to represent an important genetic resource for breeders. Most of the genes that were detected in our study were consistent with the phenotypic traits of these populations. Several genes and genomic regions here identified corroborate with previously reported studies carried out in other cattle breeds. However, further studies using the high-density array data, an increase in the number of genotyped animals and the collection of phenotypic records would be particularly relevant to refine and validate these results.

Conflict of interest

There are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

Bhati M., Kadri N.K., Crysnanto D. & Pausch H. (2020) Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. *BMC Genomics* 21(1), 1-14.

Buels R., Yao E., Diesh C.M. et *al*. (2016) JBrowse: a dynamic web platform for genome visualization and analysis. *Genome biology* 17(1), 66.

Cesarani A., Sorbolini S., Criscione A., Bordonaro S., Pulina G., Battacone G., Marletta D., Gaspa G. & Macciotta N.P.P. (2018) Genome-wide variability and selection signatures in Italian island cattle breeds. *Animal Genetics* 49(5), 371-383.

Cremonesi P., Capoferri R., Pisoni G. *et al*. (2012). Response of the goat mammary gland to infection with Staphylococcus aureus revealed by gene expression profiling in milk somatic and white blood cells. *BMC genomics*, 13(1), 540.

De Camargo G.M.F., Aspilcueta-Borquis R.R., Fortes M.R.S. *et al.* (2015). Prospecting major genes in dairy buffaloes. *BMC genomics* 16(1), 872.

Decker J.E., McKay S.D., Rolf M.M. *et al*. (2014) Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS genetics* 10(3).

Del Bo L., Polli M., Longeri M., Ceriotti G., Looft C., Barre-Dirie A., Dolf M. & Zanotti M. (2001) Genetic diversity among some cattle breeds in the Alpine area. *Journal of Animal Breeding and Genetics* 118(5), 317-325.

Elbeltagy A.R., Bertolini F., Fleming D.S., Van Goor A., Ashwell C.M., Schmidt C.J., Kugonza D.R., Lamont S.J. & Rothschild M.F. (2019) Natural Selection Footprints Among African Chicken Breeds and Village Ecotypes. *Frontiers in Genetics* 10, 376.

Excoffier L., Lischer H.E.L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10, 564-7.

Flori L., Fritz S., Jaffrézic F., Boussaha M., Gut I., Heath S., Foulley J.L. & Gautier M. (2009) The genome response to artificial selection: a case study in dairy cattle. *PloS one* 4(8).

Foll M. (2012) BayeScan v2. 1 user manual. *Ecology* 20, 1450-1462.

Gaspa G., Marras G., Sorbolini S., Ajmone-Marsan P., Williams J.L., Valentini A., Dimauro C. & Macciotta N.N.P. (2014) Genome-wide homozygosity in Italian Holstein cattle using HD panel. In

proceedings of the 10th World Congress of Genetics Applied to Livestock Production, 17 to 22 August, Vancouver, BC, Canada.

Gautier M., Laloë D. & Moazami-Goudarzi K. (2010) Insights into the genetic history of French cattle from dense SNP data on 47 worldwide breeds. *PloS one* 5(9).

Hagger C. (2005). Estimates of genetic diversity in the brown cattle population of Switzerland obtained from pedigree information. *Journal of Animal Breeding and Genetics* 122(6), 405-413.

Hou G.Y., Yuan Z.R., Gao X., Li J.Y., Gao H.J., Chen J.B. & Xu S.Z. (2010) Genetic polymorphisms of the CACNA2D1 gene and their association with carcass and meat quality traits in cattle. *Biochemical genetics* 48(9-10), 751-759.

Hulsegge I., Calus M., Hoving-Bolink R., Lopes M., Megens H.J. & Oldenbroek K. (2019) Impact of merging commercial breeding lines on the genetic diversity of Landrace pigs. *Genetics Selection Evolution* 51(1), 60.

Huson D.H., Bryant D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology Evolution* 23, 254–67.

Iso-Touru T., Sahana G., Guldbrandtsen B., Lund M.S. & Vilkki J. (2016) Genome-wide association analysis of milk yield traits in Nordic Red Cattle using imputed whole genome sequence variants. *BMC genetics* 17(1), 55.

Lee J.H., Li Y. & Kim J.J. (2012) Detection of QTL for carcass quality on chromosome 6 by exploiting linkage and linkage disequilibrium in Hanwoo. *Asian-Australasian journal of animal sciences* 25(1), 17.

Lin S., Zhang H., Hou Y., Liu L., Li W., Jiang J., Han B., Zhang S. & Sun D. (2019) SNV discovery and functional candidate gene identification for milk composition based on whole genome resequencing of Holstein bulls with extremely high and low breeding values. *PloS one* 14(8).

Lopdell T. J., Tiplady K., Couldrey C. et *al.* (2019). Multiple QTL underlie milk phenotypes at the CSF2RB locus. *Genetics Selection Evolution* 51(1), 3.

Marete A., Lund M.S., Boichard D. & Ramayo-Caldas Y. (2018) A system-based analysis of the genetic determinism of udder conformation and health phenotypes across three French dairy cattle breeds. *PLoS One* 13(7).

Marras G., Gaspa G., Sorbolini S., Dimauro C., Ajmone-Marsam P., Valentini A., Williams J.L. & Macciotta N.P.P. (2014). Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Animal Genetics* 46, 110–121.

Marques E., Nkrumah J.D., Sherman E.L. & Moore S.S. (2009) Polymorphisms in positional candidate genes on BTA14 and BTA26 affect carcass quality in beef cattle. *Journal of animal science* 87(8), 2475-2484.

Mastrangelo S., Tolone M., Gerlando R.D., Fontanesi L., Sardina M.T. & Portolano B. (2016) Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. *Animal* 10, 746–54.

Mastrangelo S., Ciani E., Ajmone-Marsan P.A.,et al. (2018a) Conservation status and historical relatedness of Italian cattle breeds. *Genetics Selection Evolution* 50:35.

Mastrangelo S., Sardina M.T., Tolone M., Di Gerlando R., Sutera A.M., Fontanesi L. & Portolano B. (2018b) Genome-wide identification of runs of homozygosity islands and associated genes in local dairy cattle breeds. *Animal* 12(12), 2480-2488.

Meng Q., Wang K., Liu X., Zhou H., Xu L., Wang Z. & Fang M. (2017) Identification of growth trait related genes in a Yorkshire purebred pig population by genome-wide association studies. *Asian-Australasian journal of animal sciences* 30(4), 462.

Metzger J., Karwath M., Tonda R., Beltran S., Águeda L., Gut M., Gut I.G. & Distl O. (2015) Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. BMC genomics 16(1), 764.

Mi H., Muruganujan A. & Thomas P.D. (2013) PANTHER in: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research* 41, D377–D386.

Mudadu M.A., Porto-Neto L.R., Mokry F.B. *et al.* (2016) Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. *Bmc Genomics* 17(1), 235.

Onzima R.B., Upadhyay M.R., Doekes H.P., Brito L., Bosse M., Kanis E., Groenen M.A.M. & Crooijmans R.P. (2018) Genome-wide characterization of selection signatures and runs of homozygosity in Ugandan goat breeds. *Frontiers in genetics* 9, 318.

Pant S.D., Karlskov-Mortensen P., Salicio S.C. *et al.* (2014) Genome-wide linkage disequilibrium linkage analysis (LDLA) of body fat traits in an F2 porcine model for human obesity. In 10th World Congress of Genetics Applied to Livestock Production.

Pant S.D., Schenkel F.S., Verschoor C.P., You Q., Kelton D.F., Moore S.S. & Karrow N.A. (2010) A principal component regression based genome wide analysis approach reveals the presence of a novel QTL on BTA7 for MAP resistance in holstein cattle. *Genomics* 95(3), 176-182.

Pausch H., Wang X., Jung S., Krogmeier D., Edel C., Emmerling R., Götz K.U. & Fries R. (2012) Identification of QTL for UV-protective eye area pigmentation in cattle by progeny phenotyping and genome-wide association analysis. *PloS one* 7(5).

Purcell S., Neale B., Todd-Brown K. *et al*. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81, 559–75.

Purfield D.C., Berry D.P., McParland S. & Bradley D.G. (2012) Runs of homozygosity and population history in cattle. *BMC genetics* 13(1), 70.

Purfield D.C., McParland S., Wall E. & Berry D.P. (2017) The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. PLoS ONE 12, e0176780.

Qanbari S., Gianola D., Hayes B., Schenkel F., Miller S., Moore S., Thaller G. & Simianer H. (2011) Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC genomics* 12(1), 318.

Qanbari S., Pausch H., Jansen S., Somel M., Strom T.M., Fries R., Nielsen R. & Simianer H. (2014) Classic selective sweeps revealed by massive sequencing in cattle. *PLoS genetics* 10(2).

Ramljak J., Bunevski G., Bytyqi H. *et al.*, (2018) Conservation of a domestic metapopulation structured into related and partly admixed strains. *Molecular ecology* 27(7), 1633-1650.

Raven L.A., Cocks B.G., Kemper K.E., Chamberlain A.J., Vander Jagt C.J., Goddard M.E. & Hayes B.J. (2016) Targeted imputation of sequence variants and gene expression profiling identifies twelve candidate genes associated with lactation volume, composition and calving interval in dairy cattle. *Mammalian genome* 27(1-2), 81-97.

Rothammer S., Seichter D., Förster M. & Medugorac I. (2013) A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC genomics* 14(1), 908.

Sahana G., Guldbrandtsen B., Thomsen B., Holm L.E., Panitz F., Brøndum R. F. & Lund M. S. (2014) Genome-wide association study using high-density single nucleotide polymorphism arrays

and whole-genome sequences for clinical mastitis traits in dairy cattle. *Journal of dairy science* 97(11), 7258-7275.

Senczuk G., Mastrangelo S., Ciani E. *et al*. (2020) The genetic heritage of Alpine local cattle breeds using genomic SNPs data. *Genetics Selection Evolution* Under review.

Signer-Hasler H., Burren A., Neuditschko M., Frischknecht M., Garrick D., Stricker C., Gredler B., Bapst B. & Flury C. (2017) Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genetics Selection Evolution* 49(1), 83.

Sorbolini S., Marras G., Gaspa G., Dimauro C., Cellesi M., Valentini A. & Macciotta N.P.P. (2015) Detection of selection signatures in Piemontese and Marchigiana cattle, two breeds with similar production aptitudes but different selection histories. *Genetics Selection Evolution* 47(1), 52.

Szmatoła T., Gurgul A., Jasielczuk I., Ząbek T., Ropka-Molik K., Litwińczuk Z. & Bugno-Poniewierska M. (2019) A Comprehensive Analysis of Runs of Homozygosity of Eleven Cattle Breeds Representing Different Production Types. *Animals* 9(12), 1024.

Valente T.S., Baldi F., Sant'Anna A.C., Albuquerque L.G. & da Costa M.J.R.P. (2016) Genome-Wide Association Study between single nucleotide polymorphisms and flight speed in Nellore cattle. *PloS one* 11(6).

Vallée A., Daures J., Van Arendonk J. A.M. & Bovenhuis H. (2016) Genome-wide association study for behavior, type traits, and muscular development in Charolais beef cattle. *Journal of animal science* 94(6), 2307-2316.

Van den Berg I., Fritz S., Rodriguez S., Rocha D., Boussaha M., Lund M.S. & Boichard D. (2014) Concordance analysis for QTL detection in dairy cattle: a case study of leg morphology. *Genetics Selection Evolution* 46(1), 31.

Wei C., Wang H., Liu G. *et al.* (2016) Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. *Scientific reports* 6, 26770.

Wu X., Lund M.S., Sahana G., Guldbrandtsen B., Sun D., Zhang Q. & Su G. (2015) Association analysis for udder health based on SNP-panel and sequence data in Danish Holsteins. *Genetics Selection Evolution* 47(1), 50.

Zhang Q., Guldbrandtsen B., Bosse M., Lund M.S. & Sahana G. (2015) Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genomics* 16(1), 542.

Zhang T., Gao H., Sahana G. *et al.* (2019) Genome-wide association studies revealed candidate genes for tail fat deposition and body size in the Hulun Buir sheep. *Journal of Animal Breeding and Genetics* 136(5), 362-370.

Zhao F., McParland S., Kearney F., Du L. & Berry D.P. (2015) Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genetics Selection Evolution* 47(1), 49.

Yoder D.M. & Lush J.L. (1937) A genetic history of the Brown Swiss cattle in the United States. *Journal of Heredity* 28(4), 154-160.

Table 1. Population type, name and code, number of animals (N), country of origin and source for the genotypic data of the Brown breeds used in the contrasted groups.

Group	Breed	Breed code	N	Country	Data source
1	Original Braunvieh	ORBR	20	Switzerland	Decker *et al.*, 2014
	Braunvieh	BRHV	35	Switzerland/Germany	Ramljak *et al.*, 2018
2	Brown Swiss	BRSW	19	Americas	Gautier *et al.*, 2010
	Italian Brown	ITBR	32	Italy	Mastrangelo *et al.*, 2018

Table 2. Genetic diversity indices in the four Brown populations: observed (Ho) and expected (He) heterozygosity, the genomic inbreeding (F_{HOM}) and standard deviation (SD).

Breed	Ho±SD	He±SD	F_{HOM}±SD
ORBR	0.360 ± 0.151	0.353 ± 0.134	0.014 ± 0.018
BRHV	0.359 ± 0.165	0.351 ± 0.137	0.020 ± 0.048
BRWS	0.343 ± 0.179	0.328 ± 0.153	0.055 ± 0.040
ITBR	0.332 ± 0.173	0.323 ± 0.155	0.085 ± 0.035

Table 3. Results of the BayeScan analysis showing outlier SNPs identified between original (ORBR and BRHV) and modern (BRSW and ITBR) Brown cattle populations.

BTA	Pos	SNP name	F_{ST}	Gene
2	113,621,322	ARS-BFGL-NGS-40251	0.305	
2	113,650,536	Hapmap51816-BTA-20000	0.296	*NYAP2*
2	130,360,932	ARS-BFGL-NGS-43912	0.285	*ZBTB40*
4	38,606,134	BTB-00178575	0.266	*CACNA2D1*
4	75,566,919	ARS-BFGL-NGS-2431	0.279	-
5	15,943,853	ARS-BFGL-NGS-20187	0.251	*MGAT4C*
5	96,563,734	Hapmap28380-BTA-74649	0.245	-
6	18,195,932	ARS-BFGL-NGS-91848	0.262	*DKK2*
6	31,738,269	Hapmap33117-BTC-032493	0.253	*GRID2*
6	31,761,997	Hapmap30962-BTC-032558	0.248	*GRID2*
6	39,437,165	Hapmap33744-BTC-050901	0.245	-
6	39,573,948	BTB-00252917	0.275	-
6	39,667,998	BTB-00406718	0.273	-
6	46,364,510	ARS-BFGL-NGS-31387	0.271	*STIM2*
6	47,983,800	Hapmap52342-rs29016265	0.254	*STIM2*
6	70,387,975	Hapmap32219-BTC-042322	0.249	-
6	90,103,125	BTA-86242-no-rs	0.248	-
6	109,282,605	Hapmap32834-BTA-149103	0.255	-
8	111,318,888	ARS-BFGL-NGS-89567	0.258	-
11	27,312,691	Hapmap23233-BTA-88497	0.257	-
13	22,505,137	Hapmap35027-BES1_Contig330_454	0.284	-
14	4,534,257	Hapmap30091-BTC-005211	0.244	*FAM135B*
14	48,435,848	ARS-BFGL-NGS-70879	0.253	-
17	47,697,987	Hapmap38389-BTA-17091	0.284	*TMEM132D*
17	47,740,843	Hapmap42713-BTA-87646	0.283	*TMEM132D*
18	41,475,285	Hapmap42198-BTA-39980	0.246	-
18	54,783,476	ARS-BFGL-NGS-70161	0.246	-
20	5,665,364	BTB-01104222	0.295	*CPEB4*
20	70,694,677	ARS-BFGL-NGS-98971	0.255	-
22	11,818,533	ARS-BFGL-NGS-24627	0.252	*XYLB*
22	32,711,432	Hapmap51256-BTA-54100	0.282	*FAM19A4*
27	43,602,071	Hapmap47470-BTA-101224	0.268	*ZNF385D*
28	17,475,241	BTA-63612-no-rs	0.286	-
28	42,481,492	ARS-BFGL-NGS-81842	0.259	*PTPN20*
29	13,770,679	BTB-01423078	0.247	-

Table 4. List of genomic regions of extended homozygosity (ROHs islands) identified in original (ORBR and BRHV) (group 1) and modern (BRSW and ITBR) (group 2) Brown cattle populations.

Group	BTA	n of SNPs	n of Genes	Start bp	End bp	Length (Mb)	Genes
1	11	68	59	65,681,864	71,263,018	5.58	*ETAA1, LOC101904865, C1D, WDR92, PNO1, PPP3R1, CNRIP1, LOC104973409, PLEK, FBXO48, APLF, PROKR1, ARHGAP25, BMP10, LOC509961, LOC104973410, GKN2, GKN1, ANTXR1, GFPT1, NFU1, AAK1, LOC787229, LOC104973411, ANXA4, GMCL1, SNRNP27, MXD1, ASPRV1, LOC104973412, LOC101905499, PCBP1, LOC107132938, C11H2orf42, TIA1, PCYOX1, LOC107132939, SNRPG, EHD3, CAPN14, GALNT14, CAPN13, LCLAT1, LOC101905676, LBH, YPEL5, LOC101905873, LOC104968430, ALK, CLIP4, C11H2orf71, FAM179A, LOC107132940, WDR43, TRMT61B, SPDYA, PPP1CB, PLB1*
	26	17	36	21,539,987	22,954,453	1.41	*LOC104975965, PAX2, SLF2, SEMA4G, MRPL43, C26H10orf2, LZTS2, PDZD7, SFXN3, KAZALD1, TLX1, LOC107131879, LBX1, LOC100847491, BTRC, LOC104975969, LOC783067, POLL, DPCD, FBXW4, LOC101908075, FGF8 NPM3, MGEA5, KCNIP2, C26H10orf76, HPS6, LOC104975987, LDB1, PPRC1, NOLC1, LOC101902227, LOC785229, ELOVL3, PITX3, GBF1*
2	5	56	43	75,086,818	78,560,464	3.49	*CACNG2, IFT27, LOC101906363, LOC101906435, PVALB, LOC107132513, NCF4, CSF2RB, LOC788541, TEX33, TST, MPST, KCTD17, TMPRSS6, IL2RB, LOC510185, C1QTNF6, SSTR3, RAC2, MIR1835, CYTH4, LOC107132508, ELFN2, LOC107132514, MFNG, CARD10, USP18, ALG10, SYT10, LOC107132509, PKP2, YARS2, DNM1L, LOC101907810, FGD4, BICD1, LOC782092, KIAA1551, LOC100137780, LOC104972515, AMN1, ETFBKMT, DENND5B*
	5	9	3	79,334,574	80,235,852	0.90	*TMTC1, LOC518980, ERGIC2*
	6	21	7	86,399,795	87,818,793	1.42	*LOC112447099, SLC4A4, LOC782958, GC, NPFFR2, ADAMTS3, TRNAC-GCA*

Figures

Figure 1 Principal components analysis for the genetic differentiation among the four Brown cattle breeds. Original Braunvieh = ORBR; Braunvieh = BRHV; Brown Swiss = BRSW; Italian Brown = ITBR.

Figure 2 Manhattan plot of the pairwise genome-wide autosomal F_{ST} analysis generated by BayeScan between group 1 (ORBR/BRHV) and group 2 (BRSW/ITBR). The red line indicates the threshold of F_{ST} value (0.243) corresponding to 0.999 of the F_{ST} percentile distribution.

Figure 3 Manhattan plots of regions of homozygosity for (A) the group 1 (ORBR/BRHV) and (B) the group 2 (BRSW/ITBR). Thresholds used to detect high-homozygosity regions are indicated with a red line.

Supporting information

Figure S1 Neighbor-Joining tree relating all the individuals. The tree was constructed using allele-sharing distances.

Figure S2 Box plots of the distribution of ROH into five classes of length (1 to ≤2 Mb, 2 to ≤4 Mb, 4 to ≤8 Mb, 8 to ≤16 Mb and >16 Mb) for the two Brown cattle groups.

Table S1. Genetic differentiation among the four Brown cattle breeds measured using F_{ST}

Table S2. Descriptive statistics of the number and the frequency distribution of runs of homozygosity (ROH) in five different size classes (Mb) for the two Brown cattle groups.

Table S3. Comparison among overlapped runs of homozygosity (ROHs) islands here detected and those reported in previous studies.

Table S4. Gene ontology (GO) terms enriched (P<0.05) based on runs of homozygosity islands and number of involved genes (n) in the two Brown cattle groups.

Table S5. Bovine QTL mapped within the ROHs islands identified within the two Brown cattle groups.

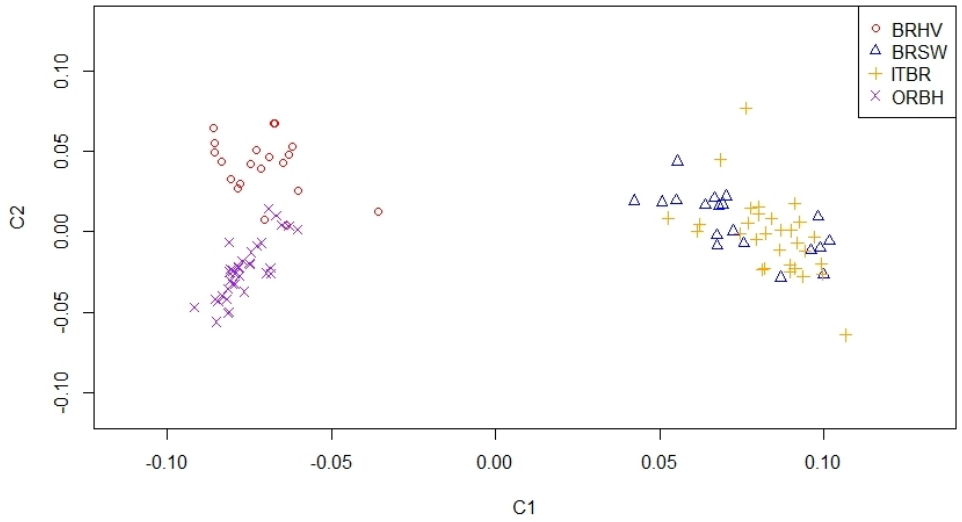


Figure 1

228x142mm (96 x 96 DPI)

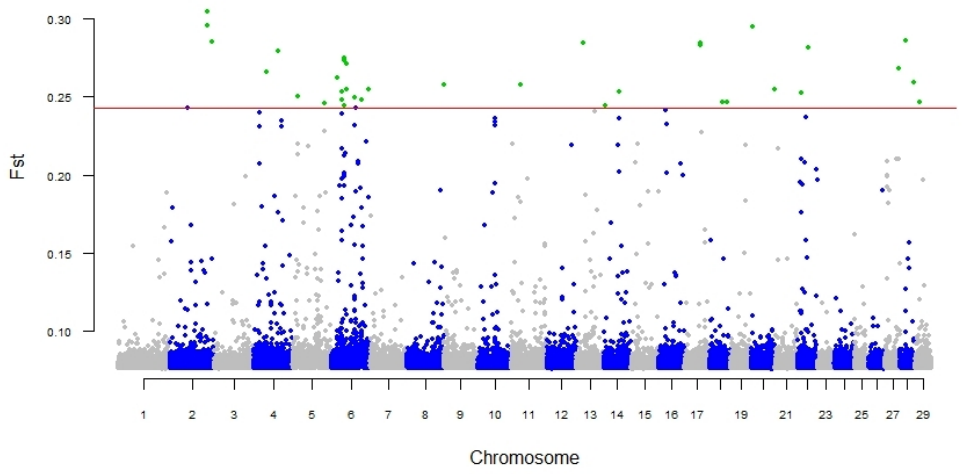


Figure 2

228x142mm (96 x 96 DPI)

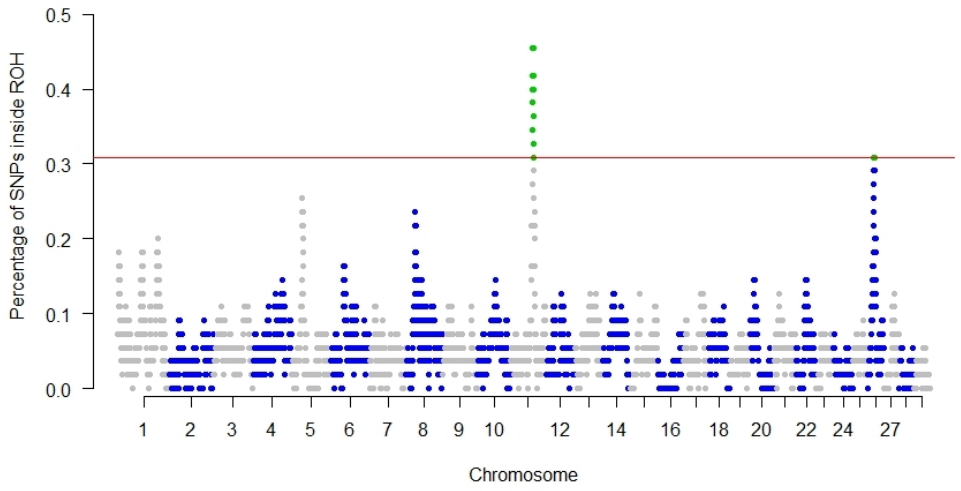


Figure 3 A

229x142mm (96 x 96 DPI)

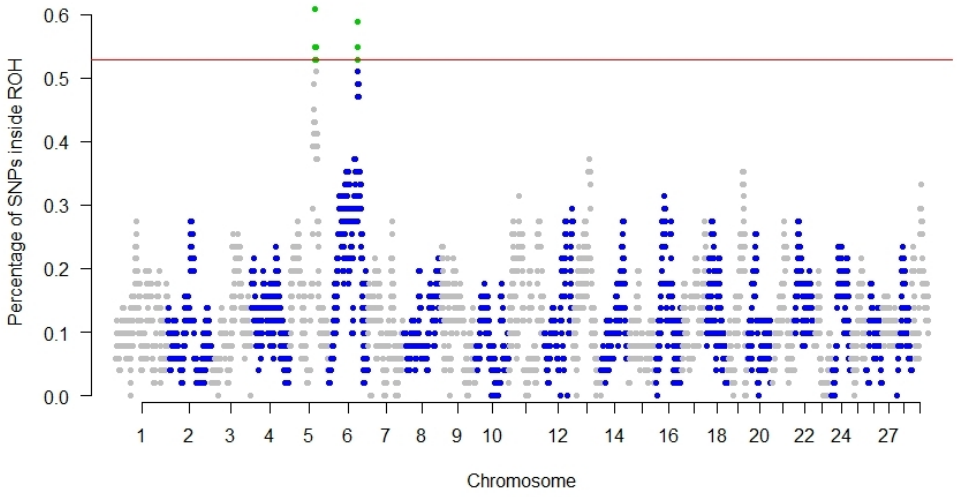


Figure 3 B

229x142mm (96 x 96 DPI)