Journal of Animal Breeding and Genetics

Genome-wide detection of signatures of selection in three Valdostana cattle populations

Journal:	*Journal of Animal Breeding and Genetics*
Manuscript ID	JABG-19-0215.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Mastrangelo, Salvatore; University of Palermo, Scienze Agrarie, Alimentari e Forestali Ben Jemaa, Slim; Laboratoire des Productions Animales et Fourragères Strillacci, Maria; Universita degli Studi di Milano, ciani, elena Sottile, Gianluca; University of Palermo, Scienze Economiche, Aziendali e Statistiche Boussaha, Mekki; INRA Centre de Jouy-en-Josas, GA; COMUE Universite Paris-Saclay, Moscarelli, Angelo; University of Palermo, SAAF Portolano, Baldassare; University of Palermo, Scienze Agrarie e Forestali Montedoro, Marina; Instituto Sperimentale Italiano Lazzaro Spallanzani Pilla, Fabio; University of Molise, S.A.V.A. Cassandro, Martino; University of Padova, Department of Animal Science
Subject Area:	animal breeding, breed, cattle, selection, SNP, genetic variation

Genome-wide detection of signatures of selection in three Valdostana cattle populations

Salvatore Mastrangelo[1]*, Slim Ben Jemaa[2]*, Maria Giuseppina Strillacci[3], Elena Ciani[4], Gianluca Sottile[5], Mekki Boussaha[6], Angelo Moscarelli[1], Baldassare Portolano[1], Marina Montedoro[7], Fabio Pilla[8] and Martino Cassandro[9]

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

[2]Laboratoire des Productions Animales et Fourragères, Institut National de la Recherche Agronomique de Tunisie, Université de Carthage, 2049 Ariana, Tunisia.

[3]Dipartimento di Medicina Veterinaria, University of Milano, 20133 Milano, Italy

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

[5]Dipartimento Scienze Economiche, Aziendali e Statistiche, University of Palermo, 90128, Palermo, Italy

[6]GABI, INRA, AgroParisTech, Paris Saclay University, 78350 Jouy-en-Josas, France.

[7]Istituto Sperimentale Italiano Lazzaro Spallanzani, 26027 Rivolta d'Adda, Italy

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

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

Corresponding author: Salvatore Mastrangelo

Email: salvatore.mastrangelo@unipa.it

Tel: +39-09123896069

* Equal contributors

running title: Selection signatures in Valdostana cattle

Abstract

The Valdostana is a local dual-purpose cattle breed developed in Italy. Three populations are recognized within this breed, based on coat color, production level, morphology and temperament: Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA). Here, we investigated putative genomic regions under selection among these three populations using the Bovine 50K SNP array by combining three different statistical methods based either on allele frequencies (F_{ST}) or extended haplotype homozygosity (*iHS* and *Rsb*). In total, 8, 5 and 8 chromosomes harboring 13, 13 and 16 genomic regions potentially under selection were identified by at least two approaches in VPR, VPN and VCA, respectively. Most of these candidate regions were population-specific but we found one common genomic region spanning 2,38 Mb on BTA06 which either overlaps or is located close to runs of homozygosity islands detected in the three populations. This region included inter alia two well-known genes: *KDR*, a well-established coat color gene and *CLOCK,* which plays a central role in positive regulation of inflammatory response and in the regulation of the mammalian circadian rhythm. The other candidate regions identified harbored genes associated mainly with milk and meat traits as well as genes involved in immune response/inflammation or associated with behavioral traits. This last category of genes was mainly identified in VCA, which is selected for fighting ability. Overall, our results provide, for the first time, a glimpse into regions of the genome targeted by selection in Valdostana cattle. Finally, this study illustrates the relevance of using multiple complementary approaches to identify genomic regions putatively under selection in livestock.

KEYWORDS: local cattle populations, Bovine BeadChip50K, selection signatures, candidate genes

1. INTRODUCTION

Selection signatures are defined as regions of the genome that harbor functionally important sequence variations and therefore are, or have been, under either natural or artificial selection (Qanbari and Simianer, 2014). These regions are often characterized by high genetic differentiation across breeds and/or a strong reduction in genetic diversity in regions associated with traits under intense selection pressure (Onzima et al., 2018). This leads to a large phenotypic variation across populations related to several behavioral (Talenti et al., 2017) and economically relevant traits.

The identification of selection signatures involved in phenotypic variation is important to better understand the evolution process and the mechanisms that underlie traits that have been exposed to natural and artificial selection.

In cattle, artificial selection has resulted in divergent breeds that are specialized in either milk or meat production or raised as dual-purpose breeds. This genetic diversity is an economical and cultural inheritance that must be preserved. Finding links between phenotypical and genotypical changes is of great importance in order to ascertain a better understanding of genetic adaptation and presents the opportunity to improve breeding work through directed selection on favorable alleles (Rothammer et al., 2013).

The availability of single nucleotide polymorphism (SNP) arrays and the progress in statistical analysis have allowed the identification of genomic regions and genes that have been subjected to positive selection in livestock species (e.g. The Bovine HapMap Consortium, 2009; Fariello et al., 2014; Brito et al., 2017; Avila et al., 2018). The different methods developed for the detection of selection signatures are based either on the distribution of allelic frequencies or the properties of haplotypes segregating within a population, or on the distribution of genetic differentiation between populations (Gutiérrez-Gil et al., 2015).

Italy has a long history of cattle breeding and, despite a dramatic contraction in numbers, still several local breeds are raised, that represent a unique source of genetic diversity (Mastrangelo et al., 2018a). The Valdostana is an indigenous dual purpose Italian breed accounting for three populations with different coat color, production, morphology and temperament (Mazza et al.,

2015). These cattle are widespread in the Aosta Valley region (northwest of Italy) and are managed in two separated herd books. The first herd book is dedicated to the Valdostana Red Pied (VPR) while the second one includes both the Valdostana Black Pied (VPN) and the Valdostana Chestnut (VCA), considered to belong to a single group because of common characteristics and the practice of crosses that occurred in the past (Mazza et al., 2015). All these animals are perfectly adapted to the difficult Alpine mountain environment, such as rough climatic conditions and meagre food resources, and the main purpose is to produce milk (used for the production of cheese) and meat. At present, VPN and VCA selection goals are directed towards fighting ability (in particular for VCA), milk and meat while VPR breeding program are focused on the improvement of milk and meat production traits.

Studies that compare populations with similar production aptitudes can be considered highly informative to investigate their genetic variability for breeding purpose (Sorbolini et al., 2015; Mastrangelo et al., 2019). The main objective of the present study was to identify putative genomic regions under selection that may explain the phenotypic differences among the three Valdostana cattle populations. For this purpose, we used three genome scan approaches: the first one is based on a population differentiation index (F_{ST},) while the second and the third ones are extended haplotype homozygosity (EHH)-derived statistics (*iHS* and *Rsb*). We also checked if these putative selection signatures overlapped with regions of high-homozygosity (ROH).

2. MATERIALS AND METHODS

2.1 Samples, genotyping and data quality control

Samples consisted of 24 individuals per population (for a total of 72 animals) from different farms. All animals were genotyped for 54,609 SNPs using Bovine SNP50K v2 BeadChip (Illumina Inc, San Diego, CA, USA). Chromosomal coordinates for each SNP were obtained from the ASR-UCD1.2 genome assembly. We excluded all SNPs not assigned to a *Bos taurus* chromosome (BTA) or assigned to chromosomes X and Y. Markers were filtered according to quality criteria that

included call frequency (⩾ 0.95) and minor allele frequency (MAF⩾0.01). Animals with more than 5% of missing genotypes were also removed.

2.2 Genetic relationships

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). The graphical representation was generated using the statistical *R* software (R Core Team, 2017).

2.3 F_{ST} analyses

The F_{ST} -outlier approach implemented in the BayeScan software (Foll & Gaggiotti, 2008) was adopted to identify loci under selection. The analyses were performed for each pairwise comparison (VPR *vs*. VPN, VPR *vs*. VCA and VPN *vs*. VCA). 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 SNPs were defined by applying a *q*-value threshold of 0.01.

2.4 Extended haplotype homozygosity-derived statistics (*iHS* and *Rsb*)

We used extended haplotype homozygosities (EHH) which is a measure for the breakdown of linkage disequilibrium with increasing distance from a SNP, to assess genome-wide signatures of positive selection. Two EHH-based metrics, *Rsb* between pairs of populations (Tang et al., 2007) and *iHS* (within population) (Voight et al., 2006) were computed using the Rehh package (Gautier & Vitalis, 2012). As a prerequisite to the *Rsb* computation, haplotypes were reconstructed from the genotyped SNPs using fastPHASE 1.4 (Scheet & Stephens, 2006). The following options were used for each chromosome: -T20 -Ku20 -Kl4 -Ki2. Considering that the *Rsb* values are normally distributed, a Z-test was applied to identify significant SNPs under selection between the three Valdostana cattle. One-sided *p*-values were derived as $pRsb=-\log_{10}[1-2|\Phi(Rsb)-0.5|]$ where Φ (x) represents the Gaussian cumulative distribution function. We used $-\log_{10}(p\text{-value}) = 4$ as a

threshold to define significant *Rsb* values. In *iHS* computation, the information on the ancestral and derived allele status is needed for each SNP because this statistic is based on the ratio of the EHH associated to each allele. In our analysis, the ancestral allele was inferred as the most common allele within 10 out-group species including yak, buffalo, sheep, horse, dog, rabbit, rat, mouse, dolphin and human. A large positive value indicates that an ancestral allele is under positive selection and has increased in frequency while a large negative value results from selection for the new derived allele. Genomic regions containing at least 4 neighboring SNPs (separated by less than 2 Mb) with an *iHS* score >2 ($p< 0.01$) were considered as putatively under selection. We have chosen to focus on clusters of neighboring SNPs because it has been demonstrated that it is more powerful to look for windows of consecutive SNPs that contain numerous extreme *iHS* scores rather than treating each SNP separately (Voight et al., 2006).

2.5 Runs of homozygosity islands

Runs of homozygosity (ROHs) were estimated for each sample using PLINK 1.07 (Purcell et al., 2007). These genomic regions were defined according to Mastrangelo et al. (2018b). To identify the genomic regions that were most commonly associated with ROH, the percentage of the occurrences of a SNP in ROH was calculated by counting the number of times the SNP was detected in those ROH across individuals. This percentage had to be higher than 25% to be an indication of a possible hotspot of ROH in the genome.

2.6 Functional characterization of regions identified as under selection

Genomic regions detected by at least two statistical approaches were interrogated for genes annotated to the *Bos taurus* genome assembly ASR-UCD1.2 using Genome Data Viewer (https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_002263795.1) provided by NCBI. We also established the list of genes covered by candidate regions identified by strong *iHS* signatures using BioMart tool of Ensembl (https://www.ensembl.org/biomart/martview/c8fe3a69961a4088a55b7a249db7e2fa). Such regions

were defined as those including a high number of consecutive outlier SNPs (>10) within small intervals (< 4 Mb). The Functional annotation clustering was performed for the list of genes located in putatively selected regions, in each of the three populations, using the Database for Annotation, Visualization and Integrated Discovery (DAVID) software. DAVID was used to identify Gene Ontology (GO) terms, Interpro protein domain and KEGG pathways with significant enrichment scores (ES) > 1.3 (which is equivalent to Fisher's exact test p-value of 0.05). To investigate the biological function and the phenotypes that are known to be affected by each annotated gene, we conducted a comprehensive literature search, including information from other species.

3. RESULTS

After filtering, the final number of animals and SNPs retained for analyses were 70 and 41,738, respectively. Only two individuals (belonging to VPN and VCA) were discarded because of low genotyping rate.

MDS analysis revealed that VPN and VCA populations are closely related, with a partial overlapping of some VPN individuals with VCA, suggesting a relatively high gene flow between them (Figure 1). For its part, VPR formed a distinct cluster.

3.1 Identifying selection signatures using F_{ST}

Manhattan plots of F_{ST} values for each pairwise comparison between the three Valdostana populations are reported in Figure 2. The results revealed a total of 39, 33 and 6 SNPs putatively under selection between VPR-VPN (Table S1), VPR-VCA (Table S2) and VPN-VCA (Table S3), respectively. Most of these were located far apart from each other. Interestingly, 6, 1 and 2 outlier SNPs, located on chromosomes BTA06, BTA18 and BTA28, respectively, were identified in both VPR-VPN and VPR-VCA comparisons (Tables S1 and S2, reported in italics). In particular, five of the six outliers SNPs detected on BTA06 were concentrated within a ~ 628 Kb interval (between positions 69,802,467 and 70,431,058 bp), while the two SNPs detected on BTA28 were 34 Kb away from each other (Tables S1 and S2).

3.2 Identifying selection signatures using *Rsb*

Plots of the three *Rsb* scores (*Rsb* VPR/VPN, *Rsb* VPR/VCA and *Rsb* VPN/VCA) over the bovine genome are presented in Figure 3. Chromosome BTA03 showed the highest number of SNPs putatively under selection with 88 markers. Among these, 43 SNPs detected by *Rsb* VPN/VCA, were located inside a 4,08 Mb region (40,19 - 44.27 Mb) and 16 SNPs identified by *Rsb* VPR/VPN, were located within the same genomic region (40.19 – 43.01 Mb) in a 2.82 Mb interval.

In total, 17, 23 and 17 candidate regions were detected for the VPR-VPN, VPR-VCA and VPN-VCA comparisons, respectively (Tables S4, S5 and S6, respectively). None of the candidate regions detected was identified concurrently by the three *Rsb* scores. However, 11 significant regions were identified simultaneously by two *Rsb* scores: 6 shared by *Rsb* VPR/VPN and *Rsb* VPR/VCA (Table S4 and S5, reported in bold), 4 by *Rsb* VPR/VPN and *Rsb* VPN/VCA (Table S4 and S6, reported in italic) and 1 region was simultaneously identified by *Rsb* VPR/VCA and *Rsb* VPN/VCA (Tables S5 and S6, reported in underlined). Candidate regions with the highest scores (-log10 *p*-value > 5.7) were located on chromosomes BTA08 and BTA16 for *Rsb* VPR/VPN, on chromosomes BTA02, BTA08 and BTA28 for *Rsb* VPR/VCA and on chromosomes BTA03, BTA06 and BTA27 for *Rsb* VPN/VCA (Figure 3).

3.3 Identifying selection signatures using *iHS*

A total of 1,404 autosomal SNPs passed the threshold of *p*-value equal to 0.01 in the three populations. VPN showed the lowest number of outliers (468 SNPs) compared to VCA (490 SNPs) and VPR (523 SNPs) (Figure 4). These outliers defined 39, 34 and 37 candidate genomic regions putatively under positive selection in VPR, VPN and VCA, respectively (Table S7). All three populations showed several strong *iHS* signatures, defined as regions with a high number of consecutive outliers (>10) within small intervals (< 4 Mb) (Table S8). In contrast to VPN and VCA, these regions with stronger evidence of selection were distributed over a higher number of chromosomes in VPR. For instance, BTA01, BTA04, BTA05, BTA07 and BTA12 presented at least 10 outliers within intervals of less than 3 Mb in VPR while all SNP clusters with similar

characteristics (*i.e* clusters of at least 10 outlier SNPs located within less than 3 Mb) were distributed on two and three chromosomes in VPN and VCA, respectively (Table S8). Interestingly, among these candidate regions with a stronger signal, one common interval was shared between these two populations. This common genomic region, spanning less than 4 Mb, is located on BTA06 (between 68 and 72 Mb) (Table S8).

3.4 Runs of homozygosity islands

In total, six ROH islands were identified (Table S9). It is worth noting that all the three Valdostana populations showed a ROH island on BTA06 at neighboring positions (located between 65 and 71 Mb). The two ROH islands detected in VPR and VCA on BTA06 overlapped with selection signatures identified with *iHS* and F_{ST} approaches in each of these two populations (see Table 1). Otherwise, it is also interesting to note that two ROH islands identified on BTA06 (at position: 68,860,609 - 69,424,834 bp) and BTA08 (at position: 85,706,017 - 87,523,043 bp) in VPN and VPR, respectively, overlapped with regions showing strong *iHS* signatures (Table S8).

3.5 Overlap between selection signatures metrics

Only two candidate regions were jointly identified by the three statistical methods. This was in the VCA within regions located on BTA06 (at position: 77,787,020 - 78,566,538 bp) and on BTA28 (at position 23,860,572 - 24,527,896) (Table 1). Besides, we found 13, 13 and 14 putative selection signatures overlapping between two tests in VPR, VPN and VCA, respectively (Table 1). In VPN and VCA, all the 29 candidate regions were detected, inter alia, using *iHS* test. All three Valdostana populations showed a common selection signatures on BTA06 (70.56 – 70.90 Mb). Two ROH islands (those detected in VPR and VCA on BTA06) overlapped with these selection signatures.

3.6 Functional Annotation of Candidate Genes

A total of 31, 86 and 84 known genes were found within the candidate region intervals identified by at least two statistical approaches in VPR, VPN and VCA, respectively (Table 1). These candidate genes grouped into 17 functional term clusters (5, 7 and 5 clusters within VPR, VPN and VCA, respectively) (Table S10). Five among these 17 clusters were significantly enriched (enrichment

scores > 1.3): 3 for VPR regarding the PI3K-Akt signaling pathway (ES = 1.48), Tyrosine-protein kinase (ES = 1.34) and Immunoglobulin-like domain (ES = 1.31), 1 for VPN regarding Glutathione S-transferase (ES = 2.06) and 1 for VCA regarding GABA-A receptor activity (ES = 1.87). Likewise, a list of 146, 104, and 160 genes were retrieved from the candidate regions showing strong *iHS* signatures in VPR, VPN and VCA, respectively (Table S8). DAVID analysis shows that 2, 2 and 7 clusters were significantly enriched in VPR, VPN and VCA, respectively (Table S11). In the VPR, the two significantly enriched GO terms included the cysteine-type endopeptidase inhibitor activity (ES = 2.88) and ATP binding (ES = 2.06) while those found in VPN included genes involved in ligase acitivity (ES = 2.1) and others associated with the PI3K-Akt signaling pathway (ES = 1.35). The seven clustersthat were significantly over-represented in VCA were related to CXCR chemokine receptor binding (ES = 5.19), Serum albumin protein domain (ES = 4.69), response to estradiol, progesterone and dehydroepiandrosterone (ES = 3.97), glucuronosyltransferase activity (ES = 1.95), sulfation (ES = 1.63), growth factor activity (ES = 1.57) and GABA-A receptor activity (ES = 1.56).

4. DISCUSSION

Human-mediated selective processes, including within-breed selection to enhance productivity, have left noticeable genomic signatures surrounding numerous genes known for having a significant effect on economically relevant traits (Boitard & Rocha, 2013; Mancini et al., 2014; Fan et al., 2014; Gurgul et al., 2016). Uncovering these genomic footprints could give an insight for understanding the mechanisms of selection and could help to assign chromosomal regions related to important physiological and economical traits (Rothammer et al., 2013). In this study, we mapped, for the first time, selection signatures across the genome of three Italian Valdostana populations.

As a first step, we performed an MDS analysis to investigate the genetic relationships among these populations. The results were consistent with a previous study showing a closer genetic relationship between VPN and VCA (Del Bo et al., 2001), probably attributable to repeated crossbreeding

between Hérens cattle from Switzerland and VPN that have generated the VCA (Forabosco & Mantovani, 2011). This would explain also the lower proportion of outlier SNPs detected in VPN-VCA statistics (F_{ST} and Rsb) compared to VPN-VPR and VPR-VCA (Figures 1 and 2, Tables S1, S2 and S3).

As previously reported (Bahbahani et al., 2015), we found little overlap between the candidate regions identified by the three approaches. This is not surprising as there are differences in the statistics underlying each approach allowing to detect the signatures of different types of natural selection across different timescales. Differences between the two extended EHH-derived statistics might be explained by the fact that *Rsb* statistic detects more fixed and nearly fixed selective sweeps, whereas the *iHS* approach has higher power to detect partial sweeps (Tang et al., 2007). Notably, our results show that the overlap between candidate regions detected by the *Rsb* statistic to those identified by the *iHS* test is higher than the overlap observed between candidate regions detected by F_{ST} and each of these two EHH-derived statistics (Table 1). This is mainly due to the fact that F_{ST} is more efficient in identifying loci that are fixed or close to fixation for opposite alleles. This requires a large number of generations. Conversely, the EHH-derived statistics detect long-range haplotypes segregating at high frequency in the population. These latter are assumed to be recent because they persist for relatively short periods of time before being broken down by recombination.

The putatively selected genomic regions identified in the three Valdostana populations spanned a large number of candidate genes with diverse molecular, and cellular functions, which is more likely due to the fact that selection has targeted polygenic traits controlled by a complex network of genes acting simultaneously. Therefore, in our comparison with published literature, we mainly focused on genes located in candidate regions supported by at least two approaches (Table 1) which are mostly related to traits involved in livestock breeding.

The major overlap in genomic regions showing evidence of selection signatures identified by the three approaches (*iHS*, *Rsb*, and F_{ST}) occurred on the BTA06 (Table 1). The most striking result

concerns a 2,38 Mb interval (BTA06: 69,802,467 – 72,189,556 bp). This region showed strong divergent selection within the three populations, considering the outcome of the F_{ST} differentiation test (Figure 2). Lending further support to this hypothesis, this candidate region either overlaps or is located close to the ROH islands detected in the three populations (Table S9). ROH islands have been shown to be abundant in regions under positive selection because they are generally considered as consequence of selection of common ancestors that carried superior alleles at specific locations (Purfield et al., 2017). The region common to all three populations on BTA06 (70.56 – 70.90 Mb) included some known candidate genes, such as: *KDR*, associated with the white pattern in Hereford cattle (Whitacre et al. 2013) and *CLOCK*, which plays a central role in positive regulation of inflammatory response (GO:0050729) and in the regulation of the mammalian circadian rhythms. The circadian clock, an internal time-keeping system, regulates various physiological functions including metabolism, sleep, body temperature, blood pressure, endocrine, immune, cardiovascular, and behavior (Casey & Plaut, 2012). Moreover, two regions, partially overlapping and located on the BTA28, were also simultaneously identified in VPR and VCA (Table 1). The region detected in VCA spanned candidate genes such as *CTNNA3*, previously identified within a selection signature and associated with marbling score in cattle (Ryu & Lee, 2014) and *SIRT1,* involved in growth and meat quality traits (Gui et al., 2014). The region on BTA28 detected in VPR, spanned a 1,13 Mb interval and harbored 12 genes (including the two aforementioned genes) related to several meat traits such as *MYPN*, which is an important sarcomere protein with potential effects on meat quality traits in cattle (Jiao et al., 2010). Apart from the candidate regions listed earlier, most of the identified candidate regions were breed-specific.

In VPR, the *CXCR4* gene within a selection signature on BTA02, is reported as strong candidate gene for cattle trypanotolerance (Dayo et al., 2009), whereas *KIT* on BTA06 plays a key role in melanogenesis and is a major candidate for the spotting locus in cattle (Fontanesi et al., 2010), in agreement with the coat color of the breed. A genomic region identified on BTA08 included the

E4BP4 (also known as *NFIL3*). This gene is thought to play a role in the regulation of the mammalian circadian oscillatory mechanism responsible for adaptations to daily environmental changes (Cowell, 2002), in the resistance to intestinal nematodes (Araujo et al., 2009) and in the response to heat stress in cattle (Srikanth et al., 2017). In this sense, the VPR population shows excellent adaptability to local environments, sometimes with harsh conditions. Besides, this region displayed a strong signal of selection because it overlapped with a ROH island (Table S9).

VPN showed two highly significant signals throughout chromosomes BTA03 and BTA06 where *Rsb* and *iHS* scores were in good agreement (Table 1). On BTA03, we have identified a relevant genomic region spanning 3.5 Mb (between position 40,00 Mb and 43.50 Mb) that overlaps with two previously identified QTL regions. The first one harbors 3 candidate genes (*OLFM3, S1PR1, DPH5*), is associated with meat production in cattle (Lim et al., 2013). The second QTL region harbors 9 candidate genes (*RTCA, DBT, LRRC39, TRMT13, SASS6, MFSD14A, SCL35A3, AGL* and *FRRS1*) and was shown to be associated with several milk production traits in water buffalo (Liu et al., 2018). Similarly, the five selection signatures on BTA06 spanned several candidate genes associated with meat (*CHRNA9*) (Velez-Irizarry et al., 2019) and milk (*RBM47*, *NSUN7*, *APBB2* and *UCHL1*) traits in cattle (Hu et al., 2010). Some of these genes, such as *APBB2*, was also reported within selection signatures in previous studies (Qanbari et al., 2011; Porto-Neto et al., 2013). These results are in keeping with the breeding schemes objectives from which VPN was developed (i.e dual purpose). Other interesting genes are: *CYM* (on BTA03), involved to antigen recognition (Makina et al., 2015) and *RHOH* (on BTA06) which plays a role in the determination of the antibody response (Twomey et al., 2019). Finally, selection signatures were also detected in regions containing genes (e.g. *GSTM1, GSTM2* and *GSTM3* on BTA03 and *PCDH7* on BTA06) that have been already reported to be under selection in cattle and associated with feed intake (Chen et al., 2011; Tizioto et al., 2015).

In VCA, several candidate genes on the BTA06 are involved in immune response/inflammation or associated with behavioral traits. For instance, three chemokine genes (*CXCL9*, *CXCL10* and

CXCL11) were identified within a 1 Mb interval (BTA06: 90,723,593 – 91,723,593 bp). These genes are key players in many disease processes, including inflammation, autoimmune disease, infectious diseases (Zlotnik et al., 2006). It is worth noting that 5 other chemokine genes (*CXCL8, CXCL5, CXCL2, CXCL3, CXCL13*) were located within candidate genomic regions showing strong *iHS* signature on other chromosomes (Table S8). The presence of multiple chemokine genes, all of them identified exclusively in the VCA population, would seem to suggest that chemokine activity is under intense selection pressure in this breed. One of the most relevant genes that may influence behavioral traits in VCA cattle concerns *USP46*. This gene is involved in behavioral fear response (GO:0001662) and righting reflex (GO:0060013) (Tomida et al., 2009). Likewise, we identified a cluster of 4 GABA-A receptor subunits genes (*GABRA2*, *GABRA4*, *GABRB1* and *GABRG1*). GABA is the major inhibitory neurotransmitter in the mammalian brain where it acts at GABA-A receptors, which are ligand-gated chloride channels. These genes have been found to mediate anxiolytic activity, which plays a key role in emotional and behavioral control in human (Möhler, 2007) and in the level of neural excitation (Ray & Hutchison, 2009). This could be considered an intriguing result. In fact, these findings may be linked to the peculiar activity of the VCA, "Battailles de Reines", bloodless tournaments in which pairs of cows fight for the title of "Queen" in front of a huge public (Sartori & Mantovani, 2010). The VCA population is characterized by a lower milk production compared to VPR and VPN, but it is well-developed and very strong, lively and quite aggressive with counterparts on summer pasture (Mazza et al., 2015). The success of "Batailles de Reines" has recently led breeders to ask for the introduction of fighting ability within the selection index. In 2012, the index was introduced, making VCA a triple purpose breed (Sartori et al., 2014). Therefore, it is likely that this selection signature is the result of selection efforts on cow fighting ability. On BTA08, two relevant genes were detected simultaneously by *Rsb* and *iHS* statistics (with a strong *iHS* signature): *ZDHHC21*, associated with fertility in cattle (Kiser et al., 2019) and *TYRP1* (Berryere et al., 2003) which is most likely responsible for the brown coat color observed in this breed. We found also that several other genomic regions overlapped with

previously identified QTLs. For instance, the selection signature region on BTA17 overlaps with a QTL that harbors 4 candidate genes (*BRI3BP, DHX37, UBC, SCARB1*) associated with carcass trait (marbling) in cattle (Lee et al., 2013). Similarly, two candidate genes (*ANKH* and *CTNND2*) related with milk production traits (Sanchez et al., 2017; Du et al., 2019) were mapped within two significant regions on BTA20.

5. CONCLUSION

Selection has left important footprints throughout the Valdostana cattle genome. Our study highlighted for the first time the presence of several selective sweeps which vary between the three Valdostana populations which is in line with their different breeding histories. We identified genomic regions putatively under selection harboring genes with molecular functions that might be associated with traits under natural and/or human-mediated selection, such as coat color, milk and meat production, immune response and fighting ability. Several signals identified here corroborate with previously reported studies carried out in other cattle breeds. Our results illustrate the complementarities of the three approaches we used to detect footprints of selection. It is possible that some important genomic regions involved in the phenotypic differentiation might not have been identified. In future studies, the use of high-density array data, an increase in the number of genotyped animals and the collection of phenotypes would be particularly relevant to refine and validate these results using other analytic approaches.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

ACKNOWLEDGEMENTS

We thanks Dr. Gianluca Sottile for graphical representation in R. The authors would also like to thank two anonymous referees for valuable comments, which helped to improve the manuscript.

ORCID

Salvatore Mastrangelo http://orcid.org/0000-0001-6511-1981

Slim Ben Jemaa https://orcid.org/0000-0002-7103-3315

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author.

REFERENCES

Araujo, R. N., Padilha, T., Zarlenga, D., Sonstegard, T., Connor, E. E., Van Tassel, C., ... & Gasbarre, L. C. (2009). Use of a candidate gene array to delineate gene expression patterns in cattle selected for resistance or susceptibility to intestinal nematodes. *Veterinary parasitology*, *162*(1-2), 106-115.

Avila, F., Mickelson, J. R., Schaefer, R. J., & McCue, M. E. (2018). Genome-wide signatures of selection reveal genes associated with performance in American quarter horse subpopulations. *Frontiers in Genetics*, *9*, 249.

Bahbahani, H., Clifford, H., Wragg, D., Mbole-Kariuki, M. N., Van Tassell, C., Sonstegard, T., ... & Hanotte, O. (2015). Signatures of positive selection in East African Shorthorn Zebu: A genome-wide single nucleotide polymorphism analysis. *Scientific Reports*, *5*, 11729.

Berryere, T. G., Schmutz, S. M., Schimpf, R. J., Cowan, C. M., & Potter, J. (2003). TYRP1 is associated with dun coat colour in Dexter cattle or how now brown cow?. *Animal Genetics*, *34*(3), 169-175.

Boitard, S., & Rocha, D. (2013). Detection of signatures of selective sweeps in the Blonde d'Aquitaine cattle breed. *Animal Genetics*, *44*(5), 579-583.

Bovine HapMap Consortium. (2009). Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, *324*(5926), 528-532.

Brito, L. F., Kijas, J. W., Ventura, R. V., Sargolzaei, M., Porto-Neto, L. R., Cánovas, A., ... & Schenkel, F. S. (2017). Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics*, *18*(1), 229.

Casey, T. M., & Plaut, K. (2012). Lactation Biology Symposium: circadian clocks as mediators of the homeorhetic response to lactation. *Journal of Animal Science*, *90*(3), 744-754.

Chen, Y., Quinn, K., Herd, R. M., Parnell, P. F., & Arthur, P. F. (2011). Quantitative real-time PCR revealed differentially expressed genes between high and low residual feed intake in Angus cattle. In *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* (Vol. 19, pp. 251-254).

Cowell, I. G. (2002). E4BP4/NFIL3, a PAR-related bZIP factor with many roles. *Bioessays*, *24*(11), 1023-1029.

Dayo, G. K., Thevenon, S., Berthier, D., Moazami-Goudarzi, K., Denis, C., Cuny, G., ... & Gautier, M. (2009). Detection of selection signatures within candidate regions underlying trypanotolerance in outbred cattle populations. *Molecular Ecology*, *18*(8), 1801-1813.

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

Du, C., Deng, T., Zhou, Y., Ye, T., Zhou, Z., Zhang, S., ... & Yang, L. (2019). Systematic analyses for candidate genes of milk production traits in water buffalo (Bubalus Bubalis). *Animal Genetics*, *50*(3), 207-216.

Fan, H., Wu, Y., Qi, X., Zhang, J., Li, J., Gao, X., ... & Gao, H. (2014). Genome-wide detection of selective signatures in Simmental cattle. *Journal of Applied Genetics*, *55*(3), 343-351.

Fariello, M. I., Servin, B., Tosser-Klopp, G., Rupp, R., Moreno, C., San Cristobal, M., ... & International Sheep Genomics Consortium. (2014). Selection signatures in worldwide sheep populations. *PLoS One*, *9*(8), e103813.

Foll, M., & Gaggiotti, O. E. (2008). A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*, 977-993.

Fontanesi, L., Tazzoli, M., Russo, V., & Beever, J. (2010). Genetic heterogeneity at the bovine KIT gene in cattle breeds carrying different putative alleles at the spotting locus. *Animal Genetics*,*41*(3), 295-303.

Forabosco, F., & Mantovani, R. (2011). European and indigenous cattle breeds in Italy. *Schiel & Denver Publishing Limited, Houston, TX*.

Gautier, M., & Vitalis, R. (2012). rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, *28*(8), 1176-1177.

Gui, L., Wang, H., Wei, S., Zhang, Y., & Zan, L. (2014). Molecular characterization, expression profiles, and analysis of Qinchuan cattle SIRT1 gene association with meat quality and body measurement traits (Bos taurus). *Molecular Biology Reports*, *41*(8), 5237-5246.

Gurgul, A., Szmatoła, T., Ropka-Molik, K., Jasielczuk, I., Pawlina, K., Semik, E., & Bugno-Poniewierska, M. (2016). Identification of genome-wide selection signatures in the L imousin beef cattle breed. *Journal of Animal Breeding and Genetics*, *133*(4), 264-276.

Gutiérrez-Gil, B., Arranz, J. J., & Wiener, P. (2015). An interpretive review of selective sweep studies in Bos taurus cattle populations: identification of unique and shared selection signals across breeds. *Frontiers in Genetics*, *6*, 167.

Hu, F., Liu, J. F., Zeng, Z. B., Ding, X. D., Yin, C. C., Gong, Y. Z., & Zhang, Q. (2010). QTL identification using combined linkage and linkage disequilibrium mapping for milk production traits on BTA6 in Chinese Holstein population. *Asian-Australasian Journal of Animal Sciences*, *23*(10), 1261-1267.

Jiao, Y., Zan, L. S., Liu, Y. F., Wang, H. B., & Guo, B. L. (2010). A novel polymorphism of the MYPN gene and its association with meat quality traits in Bos taurus. *Genetics and Molecular Research*, *9*(3), 1751-1758.

Kiser, J. N., Clancey, E., Moraes, J. G., Dalton, J., Burns, G. W., Spencer, T. E., & Neibergs, H. L. (2019). Identification of loci associated with conception rate in primiparous Holstein cows. *BMC Genomics*, *20*(1), 840.

Lee, S. H., Van Der Werf, J., Lee, S. H., Park, E. W., Gondro, C., Yoon, D., ... & Thompson, J. (2012). Genome wide QTL mapping to identify candidate genes for carcass traits in Hanwoo (Korean Cattle). *Genes & Genomics*, *34*(1), 43-49.

Lim, D., Gondro, C., Park, H. S., Cho, Y. M., Chai, H. H., Seong, H. H., ... & Lee, S. H. (2013). Identification of recently selected mutations driven by artificial selection in Hanwoo (Korean cattle). *Asian-Australasian Journal of Animal Sciences*, *26*(5), 603.

Liu, J. J., Liang, A. X., Campanile, G., Plastow, G., Zhang, C., Wang, Z., ... & Yang, L. G. (2018). Genome-wide association studies to identify quantitative trait loci affecting milk production traits in water buffalo. *Journal of Dairy Science*, *101*(1), 433-444.

Makina, S. O., Muchadeyi, F. C., van Marle-Köster, E., Taylor, J. F., Makgahlela, M. L., & Maiwashe, A. (2015). Genome-wide scan for selection signatures in six cattle breeds in South Africa. *Genetics Selection Evolution*, *47*(1), 92.

Mancini, G., Gargani, M., Chillemi, G., Nicolazzi, E. L., Marsan, P. A., Valentini, A., & Pariset, L. (2014). Signatures of selection in five Italian cattle breeds detected by a 54K SNP panel. *Molecular Biology Reports*, 41(2), 957-965.

Mastrangelo, S., Ben Jemaa, S., Sottile, G., Casu, S., Portolano, B., Ciani, E., & Pilla, F. (2019). Combined approaches to identify genomic regions involved in phenotypic differentiation between low divergent breeds: Application in Sardinian sheep populations. *Journal of Animal Breeding and Genetics*,*136*, 526–534.

Mastrangelo, S., Ciani, E., Marsan, P. A., Bagnato, A., Battaglini, L., Bozzi, R., ... & Ciampolini, R. (2018a). Conservation status and historical relatedness of Italian cattle breeds. *Genetics Selection Evolution*, *50*(1), 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.

Mazza, S., Sartori, C., & Mantovani, R. (2015). Genetic parameters of type traits in two strains of dual purpose autochthonous Valdostana cattle. *Livestock Science*, *178*, 35-42.

Möhler, H. (2007). Functional Relevance of GABA A-Receptor Subtypes. In *The GABA receptors* (pp. 23-39). Humana Press.

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

Porto-Neto, L. R., Sonstegard, T. S., Liu, G. E., Bickhart, D. M., Da Silva, M. V., Machado, M. A., ... & Van Tassell, C. P. (2013). Genomic divergence of zebu and taurine cattle identified through high-density SNP genotyping. *BMC Genomics*, 14(1), 876.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559-575.

Purfield, D. C., 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*(5), e0176780.

Qanbari, S., & Simianer, H. (2014). Mapping signatures of positive selection in the genome of livestock. *Livestock Science*, *166*, 133-143.

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

Ray, L. A., & Hutchison, K. E. (2009). Associations among GABRG1, level of response to alcohol, and drinking behaviors. *Alcoholism: Clinical and Experimental Research*, *33*(8), 1382-1390.

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.

Ryu, J., & Lee, C. (2014). Identification of contemporary selection signatures using composite log likelihood and their associations with marbling score in Korean cattle. *Animal Genetics*, 45(6), 765-770.

Sanchez, M. P., Govignon-Gion, A., Croiseau, P., Fritz, S., Hozé, C., Miranda, G., ... & Brochard, M. (2017). Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. *Genetics Selection Evolution*, *49*(1), 68.

Sartori, C., & Mantovani, R. (2010). Genetics of fighting ability in cattle using data from the traditional battle contest of the Valdostana breed. *Journal of Animal Science, 88*(10), 3206-3213.

Sartori, C., Vevey, M., & Mantovani, R. (2014). Triplice attitudine in valdostana pezzata neracastana: introduzione della combattività nell'indice di selezione. *Quaderno SOZOOALP* n° 8, 133-142.

Scheet, P., & Stephens, M. (2006). A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *The American Journal of Human Genetics*, *78*(4), 629-644.

Sorbolini, S., Marras, G., Gaspa, G., Dimauro, C., Cellesi, M., Valentini, A., & Macciotta, N. 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.

Srikanth, K., Kwon, A., Lee, E., & Chung, H. (2017). Characterization of genes and pathways that respond to heat stress in Holstein calves through transcriptome analysis. *Cell Stress and Chaperones*, *22*(1), 29-42.

Talenti, A., Bertolini, F., Pagnacco, G., Pilla, F., Ajmone-Marsan, P., Rothschild, M. F., ... & Italian Goat Consortium. (2017). The Valdostana goat: a genome-wide investigation of the distinctiveness of its selective sweep regions. *Mammalian genome*, *28*(3-4), 114-128.

Tang, K., Thornton, K. R., & Stoneking, M. (2007). A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biology*, *5*(7), e171.

Tizioto, P. C., Coutinho, L. L., Decker, J. E., Schnabel, R. D., Rosa, K. O., Oliveira, P. S., ... & Lanna, D. P. (2015). Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. *BMC Genomics*, *16*(1), 242.

Tomida, S., Mamiya, T., Sakamaki, H., Miura, M., Aosaki, T., Masuda, M., ... & Imai, S. (2009). Usp46 is a quantitative trait gene regulating mouse immobile behavior in the tail suspension and forced swimming tests. Nature Genetics, *41*(6), 688.

Twomey, A. J., Berry, D. P., Evans, R. D., Doherty, M. L., Graham, D. A., & Purfield, D. C. (2019). Genome-wide association study of endo-parasite phenotypes using imputed whole-genome sequence data in dairy and beef cattle. *Genetics Selection Evolution*, *51*(1), 15.

Velez-Irizarry, D., Casiro, S., Daza, K. R., Bates, R. O., Raney, N. E., Steibel, J. P., & Ernst, C. W. (2019). Genetic control of longissimus dorsi muscle gene expression variation and joint analysis with phenotypic quantitative trait loci in pigs. *BMC Genomics, 20*(1), 3.

Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS Biology*, *4*(3), e72.

Whitacre, L., Decker, J., Kim, J.W., Schnabel, R. & Taylor, J. (2013) Kinase insert domain receptor - a candidate for the Hereford "spotted" gene. Plant and Animal Genome XXI, San Diego, CA, P0543.

564 Zlotnik, A., Yoshie, O., & Nomiyama, H. (2006). The chemokine and chemokine receptor superfamilies and their molecular evolution. *Genome Biology,* 7, 243.

FIGURES

FIGURE 1 Genetic relationship defined with multidimensional scaling analysis for the three Valdostana cattle. Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA).

FIGURE 2 Manhattan plot of the pairwise genome-wide autosomal F_{ST} analyses generated by BayeScan (VPR_VPN, VPR_VCA and VPN_VCA). The red lines indicate the threshold of significance set at 0.01. Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA).

FIGURE 3 Manhattan plot of the pairwise genome-wide autosomal Rsb analyses. The red lines indicate the threshold of significance for the Rsb values (*p*-value of 0.0001). (VPR_VPN, VPR_VCA and VPN_VCA) Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA).

FIGURE 4 Manhattan plot of the genome-wide iHS analyses for Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA). The red lines indicate the threshold of significance for the iHS values (*p*-value of 0.01).

SUPPORTING INFORMATION

TABLE S1 Outlier single nucleotide polymorphisms (SNPs) found between Valdostana Red Pied (VPR) and Valdostana Black Pied (VPN) with the F_{ST}-based method implemented in BayeScan. SNPs that were consecutive, or separated by less than 3 markers, are reported in bold. SNPs identified in at least two comparisons are underlined and reported in italics.

TABLE S2 Outlier single nucleotide polymorphism (SNP) found between Valdostana Red Pied (VPR) and Valdostana Chestnut (VCA) with the F_{ST}-based method implemented in BayeScan. SNPs that were consecutive, or separated by less than 3 markers, are reported in bold. SNPs identified in at least two comparisons are underlined and reported in italics.

TABLE S3 Outlier single nucleotide polymorphism (SNP) found between Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA) with the F_{ST}-based method implemented in BayeScan.

TABLE S4 Putative selection signatures identified between Valdostana Red Pied (VPR) and Valdostana Black Pied (VPN) in the *Rsb* analysis. Candidate regions reported in bold are shared with those identified by *Rsb* VPR-VCA. Candidate regions reported in italic are shared with those identified by *Rsb* VPN-VCA.

TABLE S5 Putative selection signatures identified between Valdostana Red Pied (VPR) and Valdostana Chestnut (VCA) in the Rsb analysis. Candidate regions reported in bold are shared with those identified by *Rsb* VPR-VPN. Candidate regions reported underlined are shared with those identified by *Rsb* VPN-VCA.

TABLE S6 Putative selection signatures identified between Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA) in the *Rsb* analysis. Candidate regions reported in italic are shared with those identified by *Rsb* VPR-VPN. Candidate regions reported underlined are shared with those identified by *Rsb* VPR-VCA.

TABLE S7 Candidate genomic regions under selection based on *iHs* in Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA) populations.

TABLE S8 Candidate genomic regions under selection based on extreme *iHs* values (p-value < 0.01) and including a cluster of at least 10 consecutive SNPs in Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA) breeds.

TABLE S9 Run of homozygosity (ROH) islands identified in Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA) breeds.

TABLE S10 DAVID functional annotation clustering analysis of the candidate region intervals identified by at least two statistical approaches in Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA).

TABLE S11 DAVID functional annotation clustering analysis of the candidate regions showing strong *iHS* signatures in Valdostana Red Pied (VPR), Valdostana Black Pied (VPN) and Valdostana Chestnut (VCA).

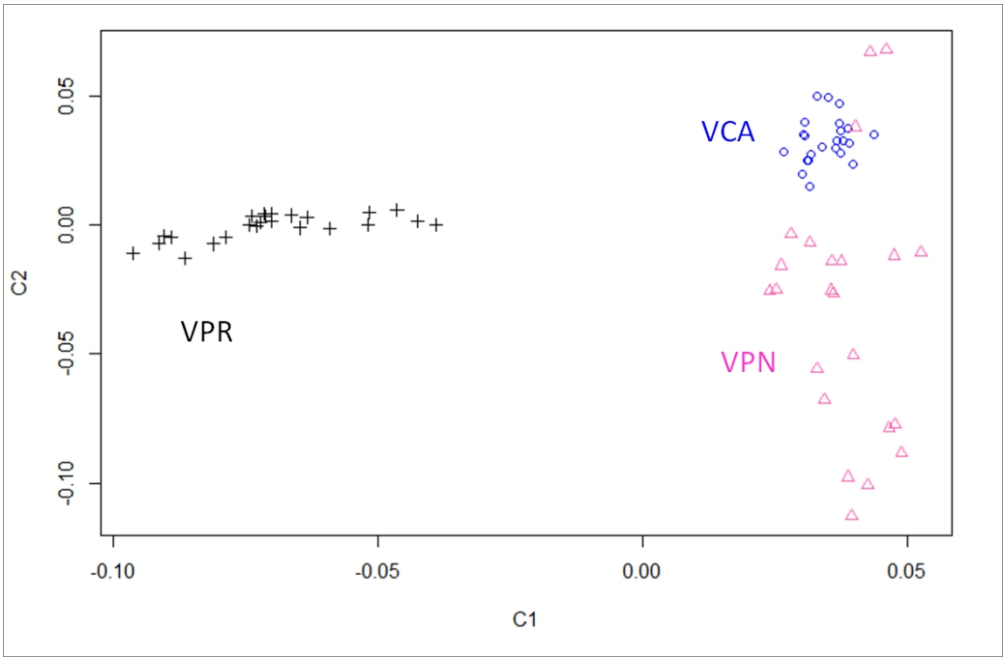
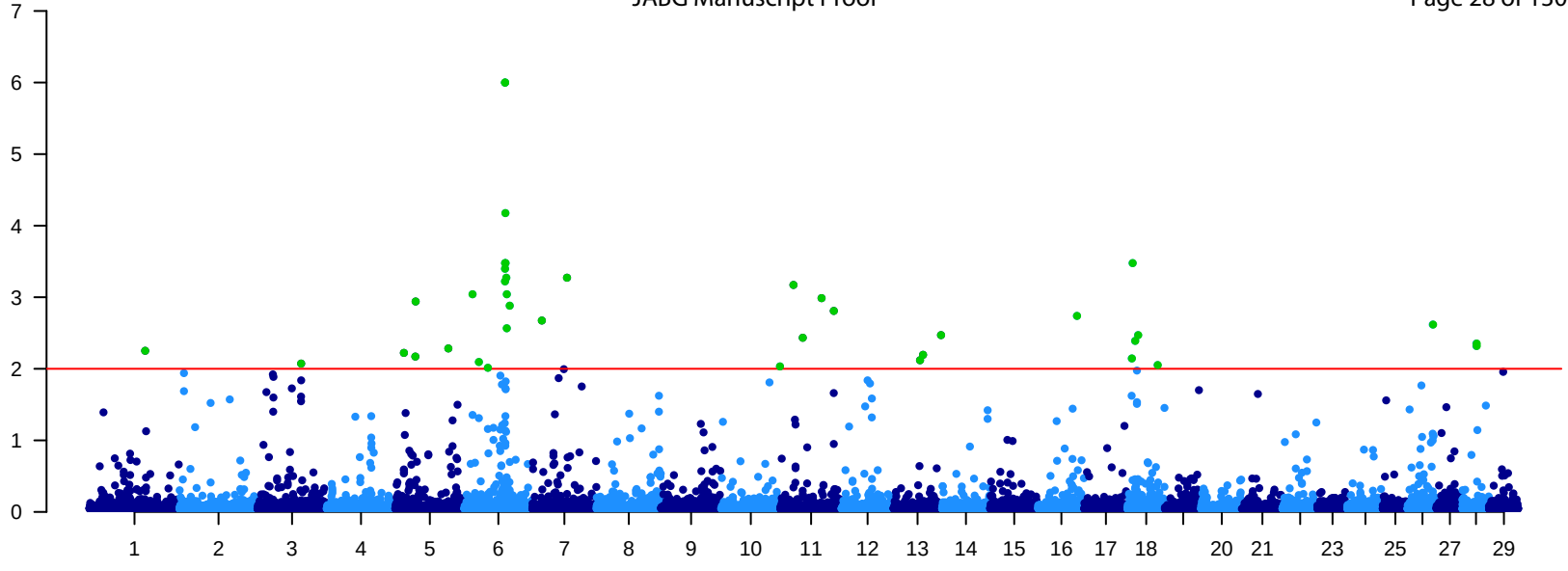
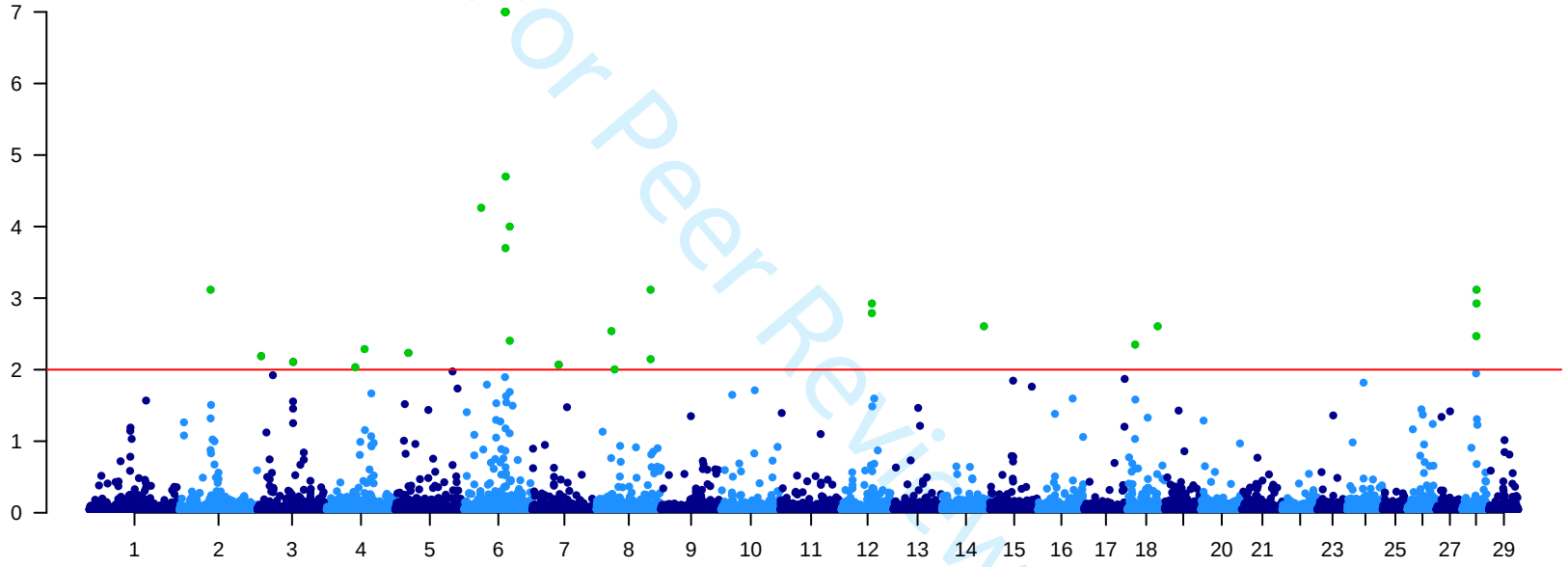


Figure 1

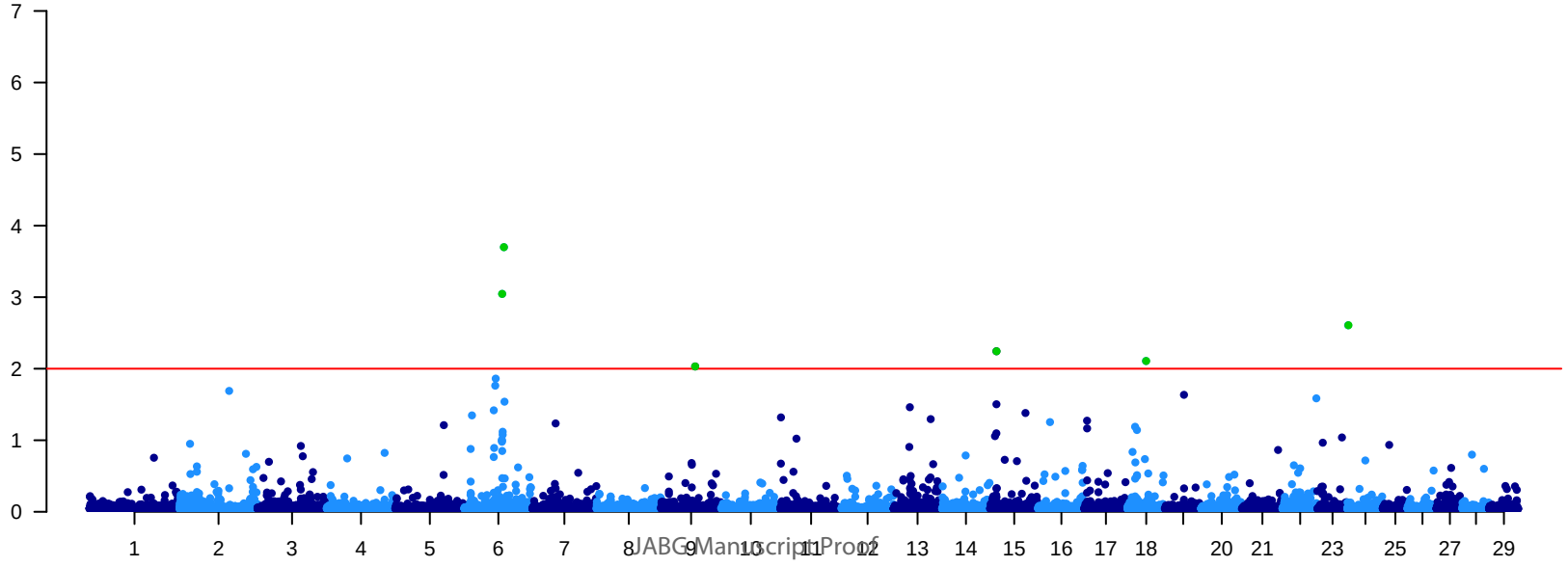
VPR_VPN



VPR_VCA



VPN_VCA



Chromosome