Animal (2020), **14:5,** pp 921–932 © The Animal Consortium 2020
doi:10.1017/S1751731119003318

animal

Comparative selection signature analyses identify genomic footprints in Reggiana cattle, the traditional breed of the Parmigiano-Reggiano cheese production system

F. Bertolini[1], G. Schiavo[2], S. Bovo[2], M. T. Sardina[3], S. Mastrangelo[3], S. Dall'Olio[2], B. Portolano[3] and L. Fontanesi[2†]

[1]*National Institute of Aquatic Resources, Technical University of Denmark, Kemitorvet, Building 202, 2800 Kongens Lyngby, Denmark;* [2]*Division of Animal Sciences, Department of Agricultural and Food Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy;* [3]*Dipartimento Scienze Agrarie, Alimentari e Forestali, University of Palermo, Viale delle Scienze, Ed. 4, 90128 Palermo, Italy*

(Received 29 June 2019; Accepted 11 October 2019; First published online 13 January 2020)

Reggiana is an autochthonous cattle breed reared mainly in the province of Reggio Emilia, located in the North of Italy. Reggiana cattle (originally a triple-purpose population largely diffused in the North of Italy) are characterised by a typical solid red coat colour. About 2500 cows of this breed are currently registered to its herd book. Reggiana is now considered a dual-purpose breed even if it is almost completely dedicated to the production of a mono-breed branded Protected Designation of Origin Parmigiano-Reggiano cheese, which is the main driver of the sustainable conservation of this local genetic resource. In this study, we provided the first overview of genomic footprints that characterise Reggiana and define the diversity of this local cattle breed. A total of 168 Reggiana sires (all bulls born over 35 years for which semen was available) and other 3321 sires from 3 cosmopolitan breeds (Brown, Holstein and Simmental) were genotyped with the Illumina BovineSNP50 panel. ADMIXTURE analysis suggested that Reggiana breed might have been influenced, at least in part, by the other three breeds included in this study. Selection signatures in the Reggiana genome were identified using three statistical approaches based on allele frequency differences among populations or on properties of haplotypes segregating in the populations (fixation index (F_{ST}); integrated haplotype score; cross-population extended haplotype homozygosity). We identified several regions under peculiar selection in the Reggiana breed, particularly on bovine chromosome (BTA) 6 in the KIT *gene region, that is known to be involved in coat colour pattern distribution, and within the region of the* LAP3, NCAPG and LCORL *genes, that are associated with stature, conformation and carcass traits. Another already known region that includes the* PLAG1 *gene (BTA14), associated with conformation traits, showed a selection signature in the Reggiana cattle. On BTA18, a signal of selection included the* MC1R *gene that causes the red coat colour in cattle. Other selection sweeps were in regions, with high density of quantitative trait loci for milk production traits (on BTA20) and in several other large regions that might have contributed to shape and define the Reggiana genome (on BTA17 and BTA29). All these results, overall, indicate that the Reggiana genome might still contain several signs of its multipurpose and non-specialised utilisation, as already described for other local cattle populations, in addition to footprints derived by its ancestral origin and by its adaptation to the specialised Parmigiano-Reggiano cheese production system.*

Keywords: autochthonous breed, *Bos taurus*, genome, selection signature, selection sweep

Implications

Reggiana cattle breed, once a multipurpose autochthonous breed, is now used to produce a mono-breed branded Parmigiano-Reggiano cheese, which is the main driver of the sustainable conservation of this local genetic resource. This study identified selection signatures in the Reggiana genome that provided information for both almost fixed breed-specific traits (e.g. coat colours) and several other more diluted signs of its re-adaptation and more recent production shifts. It was evident that this breed still contains signs of its multipurpose and non-specialised past utilisation suggesting the need to better define a tailored selection strategy for its current main use.

Introduction

Selection signature analyses based on single-nucleotide polymorphism (**SNP**) chip data have been carried out in cattle to

† E-mail: luca.fontanesi@unibo.it

identify loci under natural or artificial selection and peculiar genetic features that might be useful to describe breed-specific characteristics (e.g. Flori *et al.*, 2009; Zhao *et al.*, 2015). The statistical approaches that were used for these studies are based either on the evaluation of allele frequency differences among populations or on properties of haplotypes segregating in the populations. The fixation index F_{ST} (Wright, 1951) is one of the most used allele frequency difference approaches that quantifies population differentiation. F_{ST} provides an estimate of the amount of genetic variability that exists between populations relative to that within populations. This statistic assumes that different selective forces acting on different populations may favour divergent alleles. Therefore, allele frequency differences between populations may be more extreme in the chromosome regions in which these variants are located. Among the most frequently applied haplotype-based approaches, the integrated haplotype score (**iHS**) (Voight *et al.*, 2006) is an improvement of the extended haplotype homozygosity (**EHH**) method and compares EHH between derived and ancestral alleles within a population. The cross-population extended haplotype homozygosity (**XP-EHH**; Sabeti *et al.*, 2007) is based on both EHH and iHS but it is not calculated within populations but between populations and does not need to define ancestral and derived alleles as requested by iHS. According to their assumptions, these tests could be complementary to identify selection signatures (Gautier and Naves, 2011).

Reggiana is an autochthonous cattle breed reared mainly in the province of Reggio Emilia, located in the Emilia Romagna region, in the North of Italy. This breed is characterised by a typical red coat colour (referred as 'fromentino'). Tradition dates back the origin of the Reggiana ancestral population in the Barbaric invasion period after the fall of the Roman Empire (6th century). Historical records of the 12th century indicate that a red cattle population was used by the monks to produce in the same region a typical cheese from which subsequently originated the Parmigiano-Reggiano cheese, now a renowned and well-known worldwide Protected Designation of Origin dairy product. At that time, this population was not a specialised dairy cattle as it served for work and meat production as well.

Reggiana remained one of the most numerous cattle populations in the North of Italy till the mid of the 20th century (139 695 heads were recorded in 1954; Associazione Nazionale Allevatori Bovini di Razza Reggiana (ANABORARE), 2019). This number decreased progressively in the following decades due to the substitution of the Reggiana cattle with more specialised and productive Holstein cattle, and in the 1980s this local breed reached the minimum number of about 500 cows. Mean milk yield of Reggiana cows is about 30% lower than that of Holstein cows (Gandini *et al.*, 2007). Then a conservation programme, linked to a new brand of Parmigiano-Reggiano cheese made only of Reggiana milk, started in the 1990s. The economic advantage derived by selling this mono-breed cheese made it possible to fill the production gap in terms of economic income that the Reggiana farmers had compared to the farmers who raised more productive breeds. This branded Parmigiano-Reggiano cheese reverted the decreasing trend of the Reggiana population reaching, at present, the number of about 2500 cows reared in about 180 different farms.

A selection programme in Reggiana started in 1956 with the constitution of the National Association of Reggiana Cattle Breeders (ANABORARE), which officially could be considered the recognition of the Reggiana breed. The programme was organised in a modern way in the 1996 with the re-definition of the herd book of the breed which designed a breeding strategy aimed to reduce inbreeding. In addition, according to the use of the milk produced by Reggiana cows, a specific estimated breeding value for cheese making objectives (Parmigiano-Reggiano yield genetic index) has been implemented to improve both milk yield and milk quality for this production (including fat percentage and protein percentage, with a preference on casein variants positively associated with rennet coagulation properties; ANABORARE, 2019).

So far, few investigations were carried out in this breed to describe its genetic variability. After the pioneering studies of Mariani and Russo (1971) who evaluated the frequency distribution of k-casein protein variants, Caroli *et al.* (2004) analysed polymorphisms in three caseins and in β-lactoglobulin by isoelectrofocusing on milk. Then, 20 DNA markers were analysed in candidate genes to obtain information on their allele distribution and to identify polymorphisms associated with milk production and composition traits in Reggiana sires (Fontanesi *et al.*, 2015). Polymorphisms in coat colour genes were then investigated to identify markers useful for the authentication of Reggiana branded Parmigiano-Reggiano cheese (Russo *et al.*, 2007) and to study the genetic mechanisms differentiating solid coloured (i.e. Reggiana) from spotted patterns in cattle breeds (Fontanesi *et al.*, 2010b and 2012). Bertolini *et al.* (2015 and 2018) used SNP array data obtained from Reggiana and several other cattle breeds to identify population informative markers. Mastrangelo *et al.* (2016, 2018a and 2018b) used SNP chip data obtained in Reggiana cattle for a comparative analysis of genomic inbreeding parameters, runs of homozygosity (**ROH**) islands and population structure with other Italian local and commercial cattle breeds. The genetic structure of this breed reflects the small size of its population, with a contemporary effective population size of about 100 and a proportion of its autosomal genome covered by ROH of about 5%, similar to that of other local breeds of the North of Italy. The breed also clustered with several other cattle breeds of the North of Italy suggesting a general geographical influence of its genetic background (Mastrangelo *et al.*, 2018a).

In this study, we used Illumina SNP chip data and several statistical approaches based on allele frequency differences among populations and on properties of haplotypes segregating in the populations (F_{ST}, iHS and XP-EHH) to identify selection signatures in the Reggiana cattle genome that may distinguish this autochthonous breed from three cosmopolitan breeds (Holstein, Brown and Simmental) and

that might be indirectly derived by its ancestral origin and by its specialised use in the Parmigiano-Reggiano cheese production system.

Material and methods

Animals and genotyping data

A total of 3489 bulls of 4 cattle breeds (Reggiana, $n = 168$; Holstein, $n = 2093$; Brown, $n = 749$; and Simmental, $n = 479$) were genotyped with the Illumina BovineSNP50 v1 or v2 BeadChip arrays (Illumina, San Diego, CA, USA). Reggiana bulls were all sires born from 1975 to 2010 for which it was possible to obtain frozen semen in 2014. Considering that, on average, about 6 to 8 sires where available/approved per year over these 35 years, the analysed Reggiana bulls constituted about 70% of all bulls that were used for artificial insemination over this period in this autochthonous breed. The different numbers of analysed sires for the four breeds reflect the dimension of their respective populations.

Single-nucleotide polymorphisms were used with their coordinate position on the latest assembly of the bovine genome (ARS-UCD1.2; GCA_002263795.2). Basic SNP statistics were computed with PLINK software version 1.9 (Chang *et al.*, 2015). Only common SNPs across the two array versions and with a call rate ≥90% in each breed were retained for further analyses. All monomorphic SNPs across the dataset were removed. After filtering, all cattle had individual call rate of >0.90 and no animal was therefore discarded. The dataset was imputed using Beagle 3.3.2 (Browning and Browning, 2009) and phased for the haplotype-based analyses using fastPHASE (Scheet and Stephens, 2006) using default parameters. Imputation and phasing were carried out breed by breed.

Population structure analyses

Multidimensional scaling (**MDS**) plots were obtained with the cluster function of PLINK software version 1.9 (Chang *et al.*, 2015). Population stratification analysis was also performed with the ADMIXTURE software (Alexander *et al.*, 2009), with number of subpopulations (K) ranging from 1 to 29. As ADMIXTURE does not take linkage disequilibrium into consideration, and to reduce the computational time, the number of markers was reduced according to the observed sample correlation coefficient using the –indep-pairwise option of PLINK (Chang *et al.*, 2015).

Selection signature analyses

Detection of selection signatures in the Reggiana cattle genome was based either on the evaluation of allele frequency differences among populations and on properties of haplotypes segregating in the populations. The applied methods included within-population (iHS) and between-population (F_{ST} and XP-EHH) tests. Between-population tests were applied to identify potential sweeps that occurred in the Reggiana breed compared to the other three cosmopolitan breeds (Holstein, Brown and Simmental), which constitute the most numerous cattle populations in the North of Italy. The threshold selected for all these analyses was settled as the 99.5th percentile of the empirical distribution.

Integrated haplotype score. This statistic is applied to individual SNPs and was calculated following the procedures defined by Voight *et al.* (2006) and Sabeti *et al.* (2007). Information on the ancestral and derived alleles on all bovine SNPs was obtained from Rocha *et al.* (2014). The rehh R package v 2.0.4" (Gautier *et al.*, 2017) was used to calculate |iHS| for each autosomal SNP. Large positive or negative iHS values indicate unusually long haplotypes carrying the ancestral or derived alleles, respectively.

Fixation index. Three pairwise F_{ST} analyses were performed comparing each time the Reggiana breed with one of the other cosmopolitan breeds included in this study. Wright's F_{ST} for each SNP was calculated with PLINK 1.9 (Chang *et al.*, 2015). Average F_{ST} (mF_{ST}) was calculated in overlapping windows of 1 Mb with a step of 500 kb using an in-house script. All windows that contained at least four SNPs were then considered. Overall averaged F_{ST} was also calculated considering all SNPs in the pairwise comparisons.

Cross-population extended haplotype homozygosity. Three pairwise XP-EHH analyses were run. The XP-EHH scores were calculated using the rehh R package v 2.0.4 with default parameters (Gautier *et al.*, 2017) to detect alleles with increased frequency to the point of fixation or near-fixation in Reggiana compared to other analysed breed. In these pairwise analyses, the Reggiana breed was considered as the reference population. Therefore, only the extreme negative XP-EHH scores identified SNPs under selection in Reggiana but not in the other breeds. As XP-EHH searches for unusually long haplotypes, at least three consecutive SNPs should be above the threshold, rendering this analysis conservative. The threshold was determined using the log(*P*-value).

Annotation of candidate genome regions

Genes that were within the genome windows or haplotype regions identified as described above or that were ±500 kbp from iHS signals were retrieved from the *Bos taurus taurus* genome assembly ARS-UCD1.2 (https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1/) using the National Center of Biotechnology Information *Bos taurus* Annotation Release 106 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/). Identification of potential candidate genes for selection was obtained by comparing our results with those in the literature.

Gene enrichment analysis was performed with Enrichr (Chen *et al.*, 2013), via Fisher's exact test. Analyses run over the Gene Ontology – Biological Process (**GO**:**BP**; http://geneontology.org/), Kyoto Encyclopedia of Genes and Genomes (**KEGG**, http://www.kegg.jp/) and Reactome (https://reactome.org/) databases. As input, Enrichr took

the whole set of genes ($n = 52$) mapped within the genome regions identified by more than one method. We considered statistically enriched terms presenting: (i) at least two genes of the input set related to (at least) two different genome regions and (ii) an adjusted P-value < 0.05.

Results

Population descriptors

Supplementary Table S1 presents a descriptive summary of the genotyping data of the Reggiana and cosmopolitan cattle breeds. Reggiana cattle had intermediate values for both average minor allele frequency (**MAF**) and heterozygosity (**Het**), compared to all other breeds (MAF = 0.253 ± 0.145 and Het = 0.340 ± 0.153). Brown breed had the lowest values for these two measures (MAF = 0.232 ± 0.152 and Het = 0.313 ± 0.168) among the four analysed cattle breeds. Average Het distributed over all chromosomes in the four investigated breeds is reported in Supplementary Figure S1. No differences among chromosomes and breeds could be observed.

Figure 1 reports two-dimensional MDS plots obtained using the SNP chip data of the four investigated breeds. All breeds were clearly separated by the first three coordinates (C). Reggiana sires were closer to the Brown and Simmental clouds than to the Holstein group.

The ADMIXTURE analysis plots are shown in Supplementary Figure S2. By inspecting the plot obtained with $K = 5$, a well-defined pattern could not be observed, suggesting that Reggiana breed can be considered a distinct genetic resource, compared to the other three breeds included in this study, and matching the MDS plot results. However, the plot obtained with $K = 3$ showed that Reggiana might be influenced by all three cosmopolitan breeds with a larger impact from the Simmental breed than from Brown or Holstein breeds.

Integrated haplotype score signatures in the Reggiana genome

The genome-wide distribution of |iHS| values in the Reggiana breed is shown in Figure 2. A total of 169 SNPs distributed over 18 out of 29 autosomes marked selection sweep regions

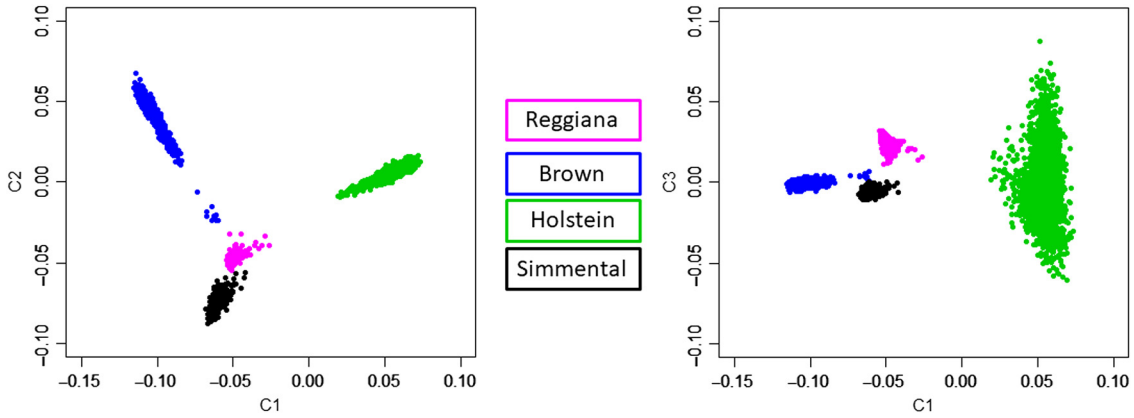

Figure 1 (colour online) Multidimensional scaling (MDS) plots of the four investigated cattle breeds obtained with the single-nucleotide polymorphism chip data. The plot on the left shows the distribution of the first (C1) and the second (C2) coordinates. The plot on the right shows the distribution of the first (C1) and the third (C3) coordinates.

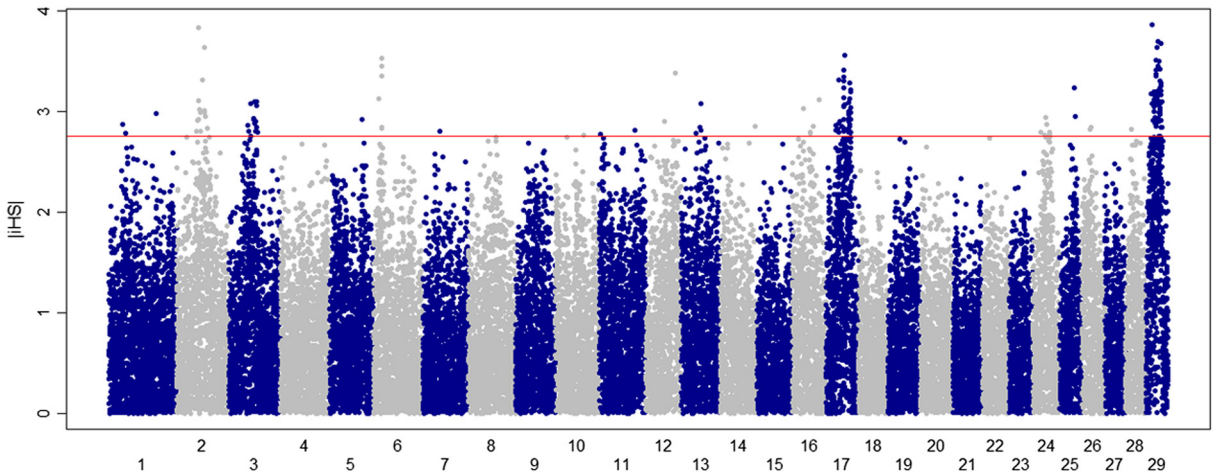

Figure 2 (colour online) Plot of the integrated haplotype score (iHS) analysis on the Reggiana breed. The |iHS| value corresponding to the bottom of the 99.5th percentile distribution was = 2.754 and is indicated with the red line in the Manhattan plot. The x-axis indicates the chromosome, the y-axis the |iHS| value.

Table 1 *List of the top 10 integrated haplotype score measures |iHS| for the single-nucleotide polymorphism (SNP) markers and their closest genes, with information on the* Bos taurus *chromosome (BTA) position*

SNP	BTA	Position	|iHS|	Closest genes on both SNP sides	Distance (bp)[1]
ARS-BFGL-NGS-39422	29	16996267	3.868	*TENM4*	0
BTB-01391891	2	54536305	3.833	*KYNU – HIGD1A*	647 722 – 443 471
Hapmap40017-BTA-65421	29	31971753	3.697	*ETS1 – FLI1*	21 920 – 212 503
ARS-BFGL-NGS-39172	29	36048959	3.679	*TMEM45B*	0
ARS-BFGL-NGS-52511	29	30103989	3.638	*KIRREL3 – ENSBTAG00000050013*	222 182 – 448 314
Hapmap49404-BTA-100549	2	70190436	3.637	*INSIG2 – ENSBTAG00000050695*	403 503 – 138 266
ARS-BFGL-NGS-9657	17	46895925	3.561	*PIWIL1 – FZD10*	133 818 – 36 824
BTB-00247622	6	16367079	3.532	*ENSBTAG00000049691 – ETNPPL*	451 494 – 186 958
ARS-BFGL-NGS-18412	29	28560818	3.516	*TMEM218 – PKNOX2*	11 801 – 249 050
Hapmap58618-rs29012371	29	32801728	3.502	*ENSBTAG00000055310 – JAM3*	87 554 – 192 247

A complete list is reported in Supplementary Table S2.
[1]Zero indicates that the SNP is within the reported gene. Two distances are reported when the SNP is between the indicated genes.

in the Reggiana genome (Supplementary Table S2). BTA17 and BTA29 were the chromosomes harboring the largest number of these SNPs (44 and 60, respectively; which included three and four regions of contiguous SNPs, respectively), followed by BTA2 and BTA3 (14 SNPs each). Among the top 10 |iHS| markers, 6 are located on BTA29, 2 on BTA2, 1 on BTA6 and 1 on BTA17 (Table 1). Some of these SNPs are within or close to genes already shown to be included in selection signature regions in the cattle genome (*TENM4* and *KIRREL3*; Bertolini *et al.*, 2018) or involved in key metabolic functions (e.g. *INSIG2* and *ETNPP*). Details and annotations for all 169 |iHS| markers are reported in Supplementary Table S2.

Fixation index signals in the Reggiana v. *cosmopolitan breed comparisons*

Average F_{ST} values including all tested SNPs in the three between-breeds comparisons, that is Reggiana *v.* Brown, Reggiana *v.* Holstein and Reggiana *v.* Simmental, were 0.0938, 0.0972 and 0.0533, respectively. Figure 3 reports the Manhattan plots of the window-based pairwise genome-wide F_{ST} analyses of the Reggiana breed against all other breeds. It is worth mentioning that, as the pairwise F_{ST} analyses cannot distinguish the direction of the signals, we regarded the identified signals obtained with this test as derived by regions that can differentiate the compared breeds. F_{ST} signals were identified on 19 autosomes (Supplementary Table S3). The highest total number of 1 Mbp outlier regions (considering all three comparisons; partially or completely overlapping or independent) was observed on BTA6 ($n = 21$) and BTA5 ($n = 10$).

On BTA6, 11, 1 and 9 regions were identified against the Brown, Holstein and Simmental breeds, respectively. Among them, two partially overlapping windows indicated a region (from positions 69.0 to 70.5 Mbp) that was in common in the Brown and Simmental comparisons. This BTA6 region contains the *KIT* gene that is well known to be involved in coat colour pattern distribution (e.g. Fontanesi *et al.*, 2010b).

In the Reggiana *v.* Brown comparison, the genomic windows with the highest mean F_{ST} (mF_{ST}) values were on BTA11, from 67.5 to 69.0 Mbp (two partially overlapping regions, $mF_{ST} = 0.47$ and 0.43, respectively), and on BTA6, from 69.5 to 70.5 Mbp ($mF_{ST} = 0.39$) and from 78.0 to 79.0 Mbp ($mF_{ST} = 0.38$). The BTA11 region corresponds to one of the most extended signatures reported by Rothammer *et al.* (2013) in a Swiss dual-purpose (dairy-beef) cattle breed (i.e. Original Braunvieh) and includes a few genes affecting meat and carcass traits (*CAPN14* and *PCBP1*). The first BTA6 region overlaps or is contiguous with other four windows with mF_{ST} above the threshold. As already mentioned, the *KIT* gene is contained in this large window, whereas in the second region, no gene is annotated.

The chromosome regions having the highest mF_{ST} values against the Holstein breed were located on BTA14 (positions 22.5 to 23.5 Mbp) and on BTA20 (positions 43.5 to 45.0 Mbp, in which no genes are annotated), with $mF_{ST} = 0.49$ and 0.48, respectively. The BTA14 region (which was also detected in the Simmental comparison) contains the *PLAG1* gene (23.33 to 23.38 Mbp) that has been already shown to determine pleiotropic quantitative trait loci (**QTL**) affecting BW, stature, reproduction traits and milk production in several cattle populations (e.g. Utsunomiya *et al.*, 2017). Another region was identified on BTA4 (76.0 to 77.0 Mbp, $mF_{ST} = 0.42$) and contains SNPs that have been already reported to differentiate cattle breeds, including Reggiana *v.* Holstein, using a random forest classification method (Bertolini *et al.*, 2015). The signal on BTA6 (windows from 37.0 to 38.0 Mbp) against the Holstein breed contains other genes (*LAP3, NCAPG and LCORL*) already associated with conformation and carcass traits, stature of the animals and calving easy (e.g. Takasuga, 2016). A signal was also observed on BTA18 with two overlapping regions (13.5 to 14.5 Mbp and 14.0 to 15.0 Mbp, mFst = 0.35 and 0.32, respectively) that include the *MC1R* gene, determining different coat colours in cattle. Two overlapping regions on BTA26 (22.0 to 23.0 and 21.5 to 22.5 Mbp, $mF_{ST} = 0.43$ and 0.41, respectively) were also identified. This chromosome portion include genes (*PAX2*, *FGF8*, *KCNIP2*, *BTRC*, *HPS6*, *ELOVL3* and MGEA5) already suggested to be involved in several processes determining coat colour and QTLs for meat and carcass traits, milk

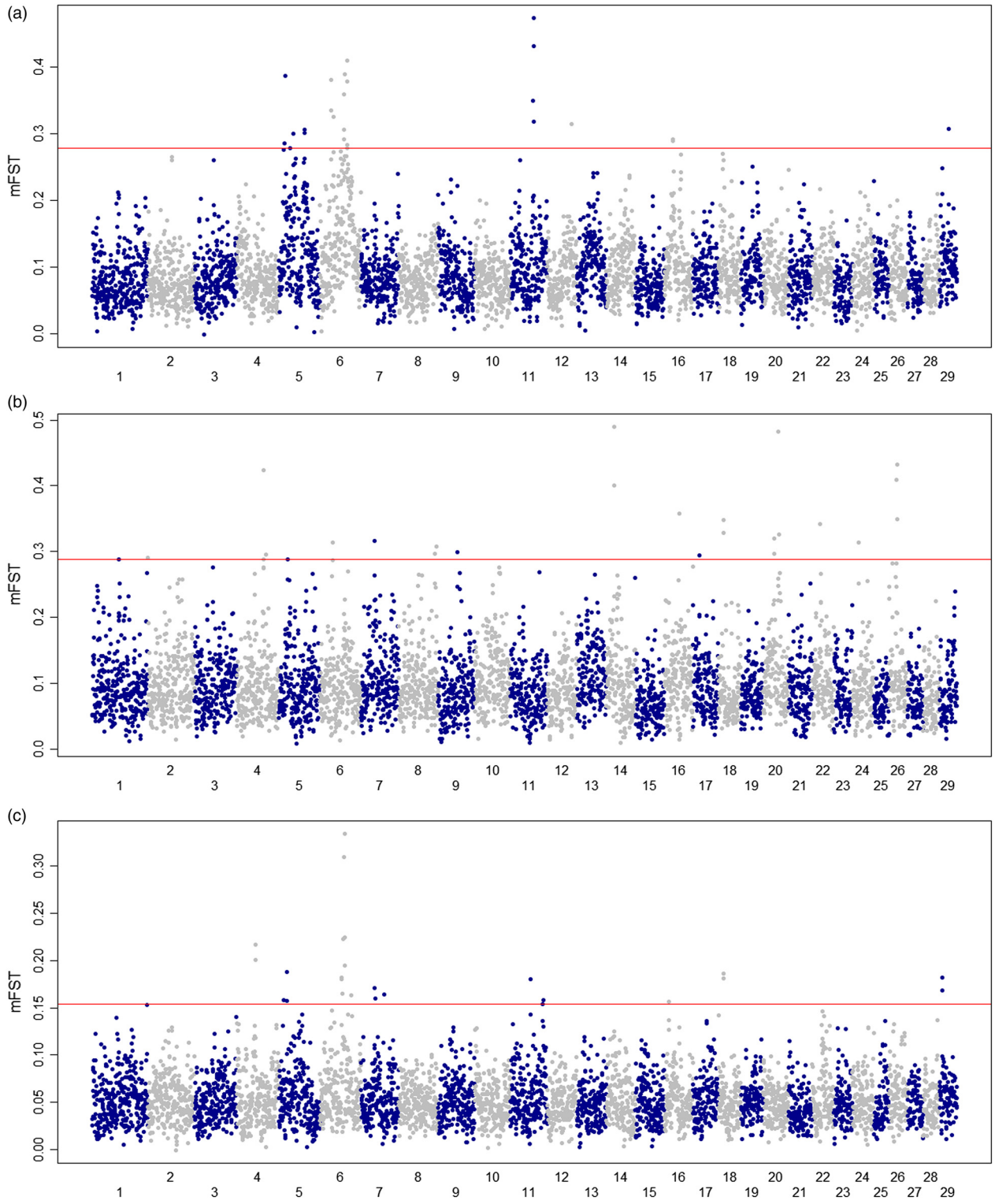

Figure 3 (colour online) Manhattan plots showing the results of the mean fixation index F_{ST} (mF_{ST}) analyses against the Brown (a), Holstein (b) and Simmental (c) breeds. The x-axis indicates the chromosome, the y-axis the mF_{ST} value. The red line in each plot represents the bottom of the 99.5th percentile distribution that is equal to 0.287, 0.279 and 0.154 for the comparisons against the Brown, Holstein and Simmental breeds, respectively.

production traits and heat tolerance (e.g. Macciotta *et al.*, 2017; Hu *et al.*, 2019).

The highest mF_{ST} values (0.33 and 0.31) in the Reggiana *v.* Simmental comparison were again on BTA6 for the common *KIT* region (four partially overlapping regions spanning from 69.0 to 71.5 Mbp). Other mF_{ST} signals in the Simmental breed comparison were also observed on BTA7 (three regions, two of which partially overlapping), on BTA11 (three windows), on BTA16 (one region), on the same BTA18 region reported for the Holstein breed and on two overlapping windows of BTA29.

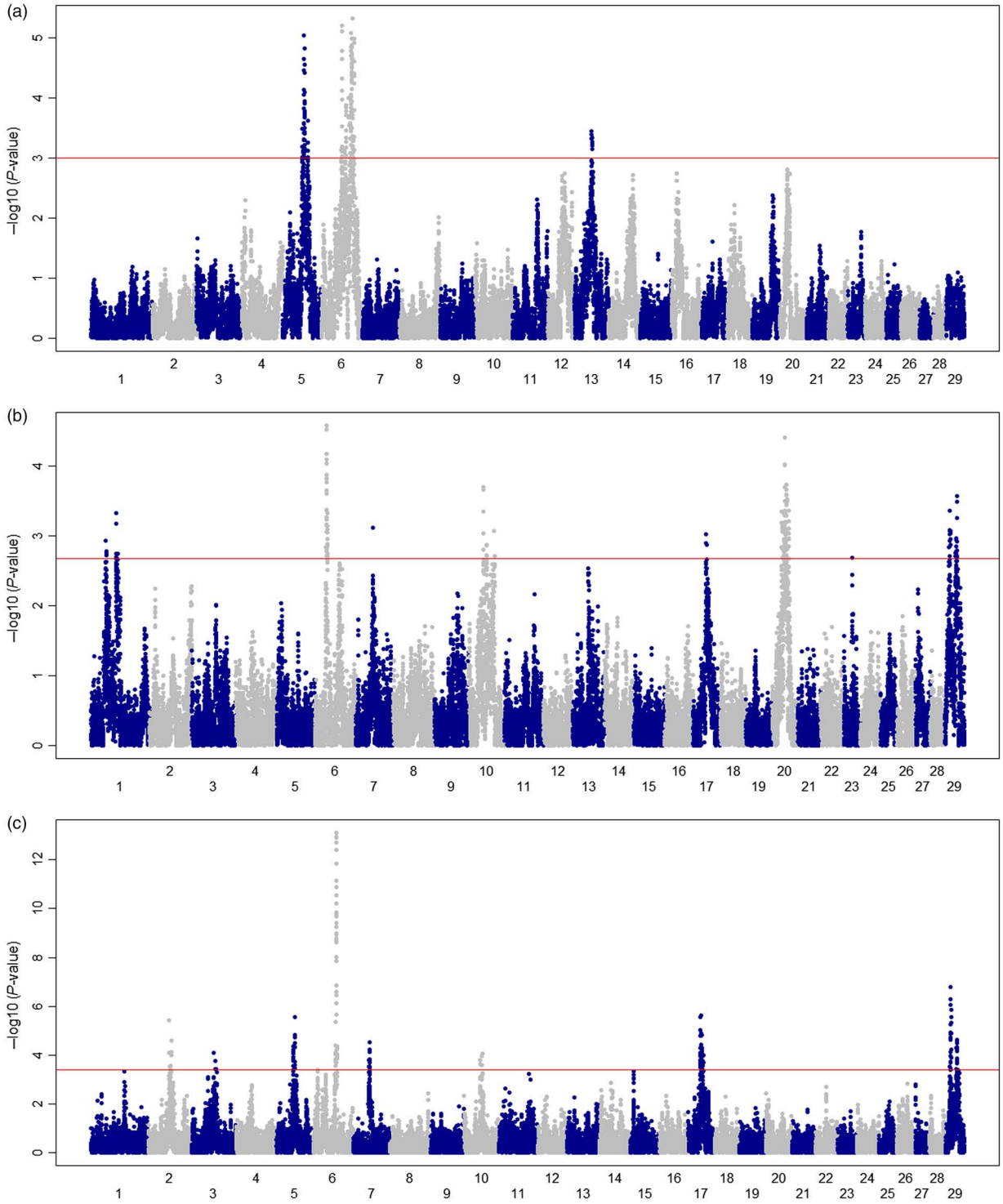

Figure 4 (colour online) Cross-population extended haplotype homozygosity (XP-EHH) analyses for Brown (a), Holstein (b) and Simmental (c) against the Reggiana. The x-axis indicates the chromosome, the y-axis the negative logarithm to base 10 of the *P*-value. For each figure, the red line represents the bottom of the 99.5th percentile distribution.

Cross-population extended haplotype homozygosity signatures in the Reggiana genome

Results of the pairwise genome-wide XP-EHH analyses between the Reggiana breed and all other three cosmopolitan breeds are shown in the Manhattan plots included in Figure 4. Signals of selection were reported on 12 out of 29 autosomes but only for 9 of these chromosomes (BTA2, BTA5, BTA6, BTA7, BTA10, BTA13, BTA17, BTA20 and BTA29) at least 3 consecutive SNPs were identified. Of these signals, negative XP-EHH values (indicating a selection on the Reggiana breed genome) were identified on the following chromosomes on the three comparisons (see Supplementary Table S4 for details):

(i) against the Brown breed, on BTA5 (three close regions separated by less than 1 Mbp), BTA6 (nine regions divided in five blocks separated by more than 1 Mbp) and BTA13 (two regions separated by more than 1 Mbp), for a total of about 5.9 Mbp;
(ii) against the Holstein breed, on BTA10 (one region) and BTA20 (nine regions divided in six blocks separated by more than 1 Mbp), for a total of about 3.4 Mbp;
(iii) against the Simmental breed, on BTA5 (four regions divided in two main blocks separated by more than 1 Mbp), BTA6 (one region) and BTA7 (three close regions, separated by less than 1 Mbp), for a total of about 4.0 Mbp.

The signals of selection on BTA5 identified against the Brown and the Simmental breeds (located in QTL regions for feed efficiency and selection signature reported in other studies) did not overlap. BTA6, summing up what observed in the different comparisons, again, showed the largest number of selection signature regions ($n = 10$). This chromosome contained the region with the highest XP-EPP log value of this study (7.608, against the Simmental breed; positions from about 68.3 to 71.4 Mbp), which encompasses the *KIT* gene. The BTA20 region detected in the Reggiana *v.* Holstein analysis contained several signals of selection in regions that have a high density of QTL for several milk production traits (Hu *et al.*, 2019).

Comparative analysis of selection signatures

The diagram of Figure 5 visualises the distribution of selection signatures obtained with the three used approaches (i.e. iHS, F_{ST} and XP-EHH) across all chromosomes. Only a small proportion of all signals overlapped among these tests. In all cases, overlapping signatures derived only by two tests. A total of 13 regions on 6 chromosomes (BTA6, BTA7, BTA13, BA17, BTA26 and BTA29) were identified by more than one method (Table 2). BTA6 contained the largest number of overlapping regions ($n = 6$), followed by BTA13 and BTA26, with two regions each. Seven regions were detected by both F_{ST} and XP-EHH tests. Three of all these overlapping regions were congruent, that means that the pairwise results were obtained against the same breed, whereas in four cases the pairwise tests identified overlapping regions derived by the comparison of different breeds. It is, however, worth to mention that in the first part of the overlapping regions of BTA6 (from about 68.3 to 70.7 Mbp; Table 2), the signals observed for the Brown (F_{ST} test) and Simmental (XP-EHH) seems parts of a broader region actually captured by both methods on each breed, as deduced from Figures 3 and 4, Supplementary Tables S3 and S4. Annotation of these regions identified several candidate genes already reported by other studies to be included in selection sweeps or to be associated with several production traits in cattle (e.g. Hu *et al.*, 2019), as also mentioned above for the description of the single methods (Table 2).

Functional analysis was carried out with Enrichr among all genes ($n = 52$) mapped in the genomic regions detected with at least two different approaches. This analysis over-represented a total of six functional terms when run over the BP branch of the GO hierarchy (Supplementary Table S5). These terms outline different processes involving the androgen metabolic process (putatively linked to fertility) and melanocyte differentiation (linked to coat colour). Other processes were related to the vesicle-mediated transport, the regulation of kinase activity and the regulation of transcription factor activity. Analyses over the KEGG and Reactome databases did not highlight any over-represented pathway.

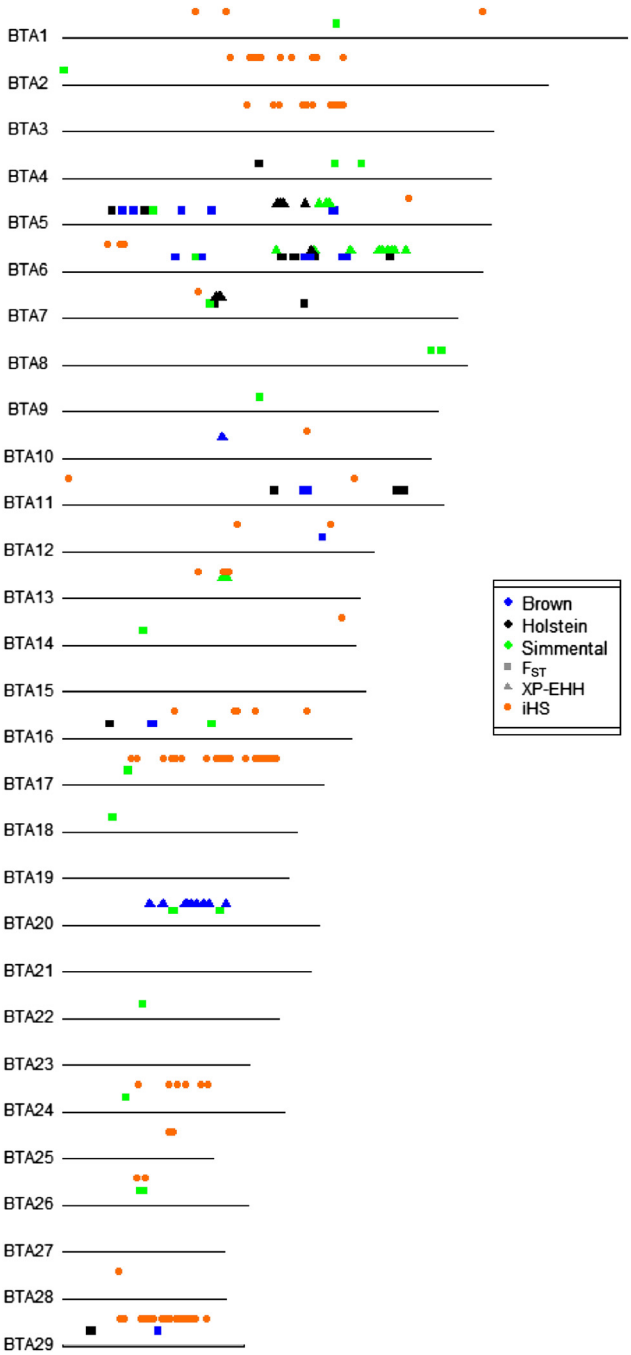

Figure 5 (colour online) Genomic footprint map reported on *Bos taurus* chromosomes (BTA) of the Reggiana breed, including selection signatures obtained with the three used approaches. iHS = integrated haplotype score; F_{ST} = fixation index; XP-EHH = cross-population extended haplotype homozygosity).

Table 2 *Selection sweeps identified by more than one test in the Reggiana chromosomes (BTA) and annotated genes in these regions*

Tests[1]	BTA (Start–End)[2]	Annotated genes
F_{ST} (Brown); XP-EHH (Simmental)	6 (68331252–70500000)	*RF00568, GSX2, RF00026, RF00026, USP46, RASL11B, CHIC2, KIT, SCFD2, FIP1L1, LNX1, PDGFRA*
F_{ST} (Brown); XP-EHH (Brown)	6 (70431058–70500000)	–
F_{ST} (Simmental); XP-EHH (Simmental)	6 (68331252–71428675)	*RF00568, GSX2, RF00026, USP46, RASL11B, CHIC2, KIT, KDR, SRD5A3, TMEM165, PDCL2, EXOC1L, CEP135, SCFD2, FIP1L1, LNX1, PDGFRA, CLOCK, NMU, EXOC1*
F_{ST} (Simmental); XP-EHH (Brown)	6 (70431058–70716954)	*RF00026, KDR*
F_{ST} (Simmental); XP-EHH (Brown)	6 (91500000–91800062)	*SOWAHB, SEPT11, SHROOM3*
F_{ST} (Simmental); XP-EHH (Brown)	6 (92383295–92430962)	*CNOT6L*
F_{ST} (Simmental); XP-EHH (Simmental)	7 (43047351–43105247)	*C2CD4C, MIER2, THEG*
XP-EHH (Holstein); iHS	13 (44978611–45082428)	*PITRM1*
XP-EHH (Holstein); iHS	13 (45936412–46049129)	*ADARB2*
F_{ST} (Holstein); iHS	17 (18961407–19000000)	–
F_{ST} (Holstein); iHS	26 (21500000–21534491)	–
F_{ST} (Holstein); iHS	26 (22661478–23500000)	*HPS6, RF00099, PITX3, NFKB2, FBXL15, TRIM8, CYP17A1, CYP17A1, LDB1, NOLC1, ELOVL3, PSD, CUEDC2, MFSD13A, ACTR1A, ARL3, WBP1L, CYP17A1, ARMH3, PPRC1, GBF1, SUFU, SFXN2*
F_{ST} (Brown); iHS	29 (27381130–27500000)	–

[1]Selection signature detection methods are reported, including the breed used in the fixation index (F_{ST}) or cross-population extended haplotype homozygosity (XP-EHH) comparisons.
[2]Chromosome positions are given in bp on the cattle reference genome for that chromosome (BTA). Regions are identified by combining positions of selection signatures derived by the different approaches. Integrated haplotype score (iHS) signal borders were defined as ±500 kb from the detected single-nucleotide polymorphisms.

Discussion

Reggiana breed is a small cattle population that can be considered a unique example of conservation of an animal genetic resource in an advanced agricultural production system, represented by the specialised dairy sector of the North of Italy. Reggiana cattle are, at present, almost completely dedicated to the production of Parmigiano-Reggiano cheese. The past unspecialised purpose of this cattle population (Reggiana was a triple-purpose breed, dairy-beef-work, till the 1960s; ANABORARE, 2019) has been redefined after the constitution of its first herd book. However, signs of its undifferentiated purposes could be left behind in the genome of these animals. Then, this red breed passed through a recent genetic bottleneck that may have further contributed to shape its current genetic make-up. Oral traditions and historical records indicate that a few genetic introgressions might have occurred in the past from Brown, Simmental and Danish Red (ANABORARE, 2019). ADMIXTURE analysis and MDS plots, however, indicate that this breed could be considered a distinct genetic pool, compared to the most important cattle breeds that constitute the backbone of the North of Italy dairy industry. Reggiana breed is however clearly closer to Simmental cattle, a dual-purpose breed. Genetic variability of Reggiana population is similar to that of the other analysed cosmopolitan breeds (Supplementary Table S1) and its estimated effective population size is larger or very close to that of the Holstein and Brown breeds, as previously determined (Marras *et al.*, 2015; Mastrangelo *et al.*, 2016 and 2018a).

In this study, we wanted to identify the unique genetic patterns that characterise the Reggiana breed genome, compared to that of the three most diffused cosmopolitan breeds in the same geographic area. Therefore, we genotyped with the Illumina BovineSNP50 panel all Reggiana sires for which we could get semen samples. The sires were born over a period of about 35 years and constitute the most active bulls that have been used since the recovery of the breed that started in the 1980s. Selection signatures were detected using three methods (i.e. iHS, F_{ST} and XP-EHH tests) which can potentially capture different selection sweep events or structures (González-Rodríguez *et al.*, 2016). Considering the complementarity of the applied methods, as expected, a small proportion of signals overlapped between these tests. It is also clear that the signals determined by the mFST tests cannot completely be assigned to an effect originated from the Reggiana breed only. Extreme mFST values might be also derived by forces acting on opposite direction on the compared cosmopolitan breed, thus this test could contain, in part, signatures not only present in the Reggiana genome. Therefore, a combination of signals derived by other methods was also used for the general interpretation of the results, particularly when F_{ST} signals were involved.

A strong selection signal, detected with both pairwise approaches, was identified in the *KIT* gene region (well known to affect coat colour patterns, e.g. Fontanesi *et al.*, 2010b), in the comparisons against the Simmental and Brown breeds. The signal in this region against the Holstein was just below the applied threshold. This is in agreement to what we already reported by comparing a few *KIT* haplotypes in several cattle breeds having different

coat colours and patterns (including Reggiana and the other three cosmopolitan breeds included in this study; Fontanesi *et al.*, 2010a). Another signal associated with different coat colour phenotypes detected by the F_{ST} pairwise analysis between Reggiana and Holstein was observed in the *MC1R* gene region, on BTA18. In this case, even if this signal was detected only with the F_{ST} analysis, it is obvious that these two breeds in this region have extreme allele frequency differences. Holstein cattle are expected to carry the E^D allele (determining the dominant black coat colour) at high frequency, whereas Reggiana cattle are fixed for the recessive *e* allele (determining the red coat colour) at the *Extension* locus (Russo *et al.*, 2007). The same BTA18 region reported a signal of selection in the F_{ST} analysis against the Simmental breed. As Reggiana and Simmental cattle have the same red coat colour (even if the latter has a spotted phenotype) and carry the same almost fixed genotype at the *MC1R* gene (allele *e* frequency in Simmental is >96%; Russo *et al.*, 2007), it seems plausible to suppose that other genetic factors may contribute to differentiate this genomic region between these two red breeds.

Other selection signatures were detected in regions containing genes (e.g. *LAP3*, *NCAPG* and *LCORL* on BTA6 and *PLAG1* on BTA14) that have been already reported to be under strong selection in cattle and shown to affect several morphological traits (Takasuga, 2016; Utsunomiya *et al.*, 2017). These regions were also described to differentiate dairy, dual-purpose and beef breeds (Gutiérrez-Gil *et al.*, 2015).

In addition to the selection signatures identified using the methods reported in this study, other regions of the Reggiana genome might have been under selection pressure. Mastrangelo *et al.* (2018b) analysed the Reggiana genome and identified ROH islands in a total of eight windows of six different chromosomes (BTA1, BTA3, BTA6, BTA17, BTA26 and BTA29; see Supplementary Tables S2 and S3). Three of these ROH islands overlap with the iHS signals we detected on BTA3 (positions from about 75.0 to 78.0 Mbp), BTA17 (from about 54.9 to 59.6 Mbp) and BTA29 (from about 16.1 to 22.6 Mbp) and another ROH island overlaps with an F_{ST} signal we reported against the Holstein breed on BTA6 (from about 37.0 to 38.0 Mbp).

Reggiana cows have, on average, a lower milk yield compared to that of Holstein and Brown. The dual-purpose Simmental breed has a similar average milk yield to that of the Reggiana breed. Simmental *v.* Reggiana has also an almost halved overall averaged F_{ST} value than that obtained in the Brown and Holstein breed comparisons (0.0533 against Simmental; 0.0938 against Brown; 0.0972 against Holstein). This lower differentiation level against the Simmental breed is also evident from the window-based mF_{ST} analysis that showed that the regions over the 99.5th percentile had a lower average value ($mF_{ST} = 0.179$) than that observed against the Brown ($mF_{ST} = 0.313$) and Holstein ($mF_{ST} = 0.338$) breeds.

Several selection sweeps detected in the Reggiana genome are located in QTL regions for milk and production efficiency traits. It is plausible to suggest that Reggiana might have a higher frequency of the less efficient and productive haplotypes for most of these regions, in addition to a general genomic background favouring heavy carcasses and high statures (as also inferred from the iHS analysis and the XP-EHH results). Taking together all these results, it could be possible to deduce that the Reggiana breed genome might still contain several signs of its multipurpose and non-specialised utilisation, as already described for other local cattle populations (Gutiérrez-Gil *et al.*, 2015). The signatures that might address the adaptation (or re-adaptation) to the Parmigiano-Reggiano production system (which cannot be simplified or summarised with few genetic determinants) are therefore mixed and then diluted with other signatures that should have been derived by the history of the Reggiana cattle breed. It will be interesting to further evaluate the genetic background of the Reggiana ancestral genome architecture in comparisons with other autochthonous breeds of similar ancestry or with other local selection goals.

Conclusion

This study provided the first overview of genomic footprints in the Reggiana cattle breed. Several signatures that have been probably left behind from the ancestral unspecialised purpose of Reggiana have contributed to differentiate this breed and testify the diversity of this cattle genetic resource. Selection sweeps were located in a few chromosome regions already known to affect coat colour and morphological traits. Several other signatures might be the results of the slow re-adaptation of this breed to its peculiar production system, at present dominated by the Parmigiano-Reggiano cheese. Being constituted by a small and close population, genetic progress of Reggiana breed towards milk yield has been limited and its genomic footprint might reflect, in general, this productive weakness even if only indirect proof could be detected with the applied methods. Other studies are needed to evaluate what could be the achievable genetic progress on milk production traits in this breed.

Acknowledgements

We thank ANABORARE and the INNOVAGEN Consortium for the collaboration in this study. This study was funded by the PSRN Dual Breeding project and it received financial support from the Italian Ministry of Agriculture, Food and Forestry (MiPAAF) under the project INNOVAGEN, from the PON02_00451_3133441 project, CUP: B61C1200076005 funded by MIUR and from the University of Bologna RFO 2018 programme.

L. Fontanesi 0000-0001-7050-3760

Declaration of interests

The authors declare that they do not have competing interests.

Ethics statement

No ethical approval was required since only genotyping data were used in the study and data were provided by the previous research programmes.

Software and data repository resources

None of the data were deposited in an official repository.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1751731119003318

References

Alexander DH, Novembre J and Lange K 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research 19, 1655–6164.

Associazione Nazionale Allevatori Bovini di Razza Reggiana (ANABORARE) 2019. Retrieved on 1 June 2019 from https://www.razzareggiana.it/.

Bertolini F, Galimberti G, Calò DG, Schiavo G, Matassino D and Fontanesi L 2015. Combined use of principal component analysis and random forests identify population-informative single nucleotide polymorphisms: application in cattle breeds. Journal of Animal Breeding and Genetics 132, 346–356.

Bertolini F, Galimberti G, Schiavo G, Mastrangelo S, Di Gerlando R, Strillacci MG, Bagnato A, Portolano B and Fontanesi L 2018. Preselection statistics and Random Forest classification identify population informative single nucleotide polymorphisms in cosmopolitan and autochthonous cattle breeds. Animal 12, 12–19.

Browning BL and Browning SR 2009. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. American Journal of Human Genetics 84, 210–223.

Caroli A, Chessa S, Bolla P, Budelli E and Gandini GC 2004. Genetic structure of milk protein polymorphisms and effects on milk production traits in a local dairy cattle. Journal of Animal Breeding and Genetics 121, 119–127.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM and Lee JJ 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4, 7.

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR and Ma'ayan A 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 128, 14.

Flori L, Fritz S, Jaffrezic F, Boussaha M, Gut I, Heath S, Foulley JL and Gautier M 2009. The genome response to artificial selection: a case study in dairy cattle. PLoS One 4, e6595.

Fontanesi L, Scotti E and Russo V 2010a. Analysis of SNPs in the *KIT* gene of cattle with different coat colour patterns and perspectives for using these markers for breed traceability and authentication of beef and dairy products. Italian Journal of Animal Science 9, e42.

Fontanesi L, Scotti E and Russo V 2012. Haplotype variability in the bovine MITF gene and association with piebaldism in Holstein and Simmental cattle breeds. Animal Genetics 43, 250–256.

Fontanesi L, Scotti E, Samorè AB, Bagnato A and Russo V 2015. Association of 20 candidate gene markers with milk production and composition traits in sires of Reggiana breed, a local dairy cattle population. Livestock Science 176, 14–21.

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

Gandini G, Maltecca C, Pizzi F, Bagnato A and Rizzi R 2007. Comparing local and commercial breeds on functional traits and profitability: the case of Reggiana dairy cattle. Journal of Dairy Science 90, 2004–2011.

Gautier M, Klassmann A and Vitalis R 2017. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. Molecular Ecology Resources 17, 78–90.

Gautier M and Naves M 2011. Footprints of selection in the ancestral admixture of a New World Creole cattle breed. Molecular Ecology 20, 3128–3143.

González-Rodríguez A, Munilla S, Mouresan EF, Cañas-Álvarez JJ, Díaz C, Piedrafita J, Altarriba J, Baro JÁ, Molina A and Varona L 2016. On the performance of tests for the detection of signatures of selection: a case study with the Spanish autochthonous beef cattle populations. Genetics Selection Evolution 48, 81.

Gutiérrez-Gil B, Arranz JJ and 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 Z-L, Park CA and Reecy JM 2019. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic Acids Research 47, D701–D710.

Macciotta NPP, Biffani S, Bernabucci U, Lacetera N, Vitali A, Ajmone-Marsan P and Nardone A 2017. Derivation and genome-wide association study of a principal component-based measure of heat tolerance in dairy cattle. Journal of Dairy Science 100, 4683–4697.

Mariani P and Russo V 1971. Distribuzione delle varianti genetiche delle caseine e della b-lattoglobulina nelle vacche di razza Reggiana. Rivista di Zootecnia 44, 310–322.

Marras G, Gaspa G, Sorbolini S, Dimauro C, Ajmone-Marsan P, Valentini A, Williams JL and Macciotta NP 2015. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Animal Genetics 46, 110–121.

Mastrangelo S, Ciani E, Ajmone-Marsan P, Bagnato A, Battaglini L, Bozzi R, Carta A, Catillo G, Cassandro M, Casu S, Ciampolini R, Crepaldi P, D'Andrea M, Di Gerlando R, Fontanesi L, Longeri M, Macciotta NPP, Mantovani R, Marletta D, Matassino D, Mele M, Pagnacco G, Pieramati C, Portolano B, Sarti MF, Tolone M and Pilla F 2018a. Conservation status and historical relatedness of Italian cattle breeds. Genetics Selection Evolution 50, 35.

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

Mastrangelo S, Tolone M, Di Gerlando R, Fontanesi L, Sardina MT and Portolano B 2016. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. Animal 10, 746–754.

Rocha D, Billerey C, Samson F, Boichard D and Boussaha M 2014. Identification of the putative ancestral allele of bovine single-nucleotide polymorphisms. Journal of Animal Breeding and Genetics 131, 483–486.

Rothammer S, Seichter D, Förster M and Medugorac I 2013. A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. BMC Genomics 14, 908.

Russo V, Fontanesi L, Scotti E, Tazzoli M, Dall'Olio S and Davoli R 2007. Analysis of melanocortin 1 receptor (*MC1R*) gene polymorphisms in some cattle breeds: their usefulness and application for breed traceability and authentication of Parmigiano Reggiano cheese. Italian Journal of Animal Science 6, 257–272.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES; International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallée C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Johnson TA, Mullikin JC, Sherry ST, Feolo M, Skol A,

Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archevêque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R and Stewart J. 2007. Genome-wide detection and characterization of positive selection in human populations. Nature 449, 913–918.

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

Takasuga A 2016. PLAG1 and NCAPG-LCORL in livestock. Animal Science Journal 87, 159–167.

Utsunomiya YT, Milanesi M, Utsunomiya ATH, Torrecilha RBP, Kim ES, Costa MS, Aguiar TS, Schroeder S, do Carmo AS, Carvalheiro R, Neves HHR, Padula RCM, Sussai TS, Zavarez LB, Cipriano RS, Caminhas MMT, Hambrecht G, Colli L, Eufemi E, Ajmone-Marsan P, Cesana D, Sannazaro M, Buora M, Morgante M, Liu G, Bickhart D, Van Tassell CP, Sölkner J, Sonstegard TS and Garcia JF 2017. A PLAG1 mutation contributed to stature recovery in modern cattle. Scientific Reports 7, 17140.

Voight BF, Kudaravalli S, Wen X and Pritchard JK 2006. A map of recent positive selection in the human genome. PLoS Biology 4, e72.

Wright S 1951. The genetical structure of populations. Annals of Eugenetics 15, 323–354.

Zhao F, McParland S, Kearney F, Du L and Berry DP 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. Genetics Selection Evolution 47, 49.