BMC Genomics

RESEARCH

Open Access

Assessment of genome complementarity in three beef-on-dairy crossbreds reveals sire-specific effects on production traits with comparable rates of genomic inbreeding reduction

Dorothea Lindtke[1*], Sylvain Lerch[2], Isabelle Morel[2] and Markus Neuditschko[1]

Abstract

Background Crossbreeding beef bulls with dairy cows can improve the economic value and fitness of calves not entering dairy production owing to increased meat yield and heterosis. However, outcrossing might reduce the dosage of alleles that confer local adaptation or result in a higher risk of dystocia due to increased calf size. Given the clear phenotypic differences between beef breeds, the varying phylogenetic distances between beef and dairy breeds, and the genomic variations within breeds, the attainable economic and fitness gains of calves will strongly depend on the selection of sires for crossing. Thus, the aim of this study was to assess genome complementarity between Angus (AAN), Limousin (LIM), or Simmental (SIM) beef bulls and Brown Swiss (BSW) dairy cows by quantifying genomic inbreeding reduction in F1 crosses and identifying genes potentially under BSW-specific selection that might be affected by outcrossing.

Results Low-pass sequencing data from 181 cows, 34 bulls, and 301 of their F1 progeny, and body weight and carcass composition measurements of 248 F1s were obtained. The high genomic inbreeding levels detected in the BSW cows were substantially reduced in the crossbreds, with only minor differences between the sire breeds. In the BSW cows, 585 candidate genes under selection were identified, overrepresenting genes associated with milk, meat and carcass, and production traits. Only a few genes were strongly differentiated at nonsynonymous variants between the BSW and beef breeds, including four tightly clustered genes (*FAM184B*, *NCAPG*, *DCAF16*, and *LCORL*) nearly fixed for alternate alleles in the BSW cows but mostly heterozygous or homozygous for the reference alleles in the AAN and LIM bulls. The alternate allele dosage at these genes significantly correlated with reduced carcass weight and protein mass in F1s.

Conclusion Some of the few genes that were highly divergent between the BSW and beef breeds at nonsynonymous variants were likely under strong selection for reduced carcass weight in the BSW breed, potentially due to trade-offs between beef and dairy productions. As alleles with opposing effects still segregate in beef cattle, marker-assisted selection of mating pairs may be used to modulate the desired phenotypes and simultaneously decrease genomic inbreeding.

*Correspondence:
Dorothea Lindtke
doro.lindtke@gmail.com
Full list of author information is available at the end of the article

Keywords Low-pass sequencing, *PPFIBP2*, *FAM184B–LCORL* complex, Brown Swiss, Dairy cattle, Homozygosity-by-descent

Background

The use of sexed semen in dairy herds to breed replacement cows enables the application of beef semen on the remaining dams that are solely intended for milk production, resulting in increased income from calves, a shared allocation of greenhouse gas emissions from dairy and beef production systems, and heterosis-related fitness gains in crossbred calves [1–3]. However, given the clear phenotypic differences between beef breeds, genetic variations within dairy females, and distinct requirements of diverse production systems, no single beef-on-dairy cross will provide the best value for all situations [4, 5]. In identifying appropriate beef bulls for mating to dairy females, the characteristics important to dairy production, such as easy calving and short gestation length, and beef production, such as carcass quality and growth efficiency, must be considered [6]. For example, using Angus (AAN) as beef sire on a dairy dam leads to fewer calving difficulties and reduced gestation length compared with using Limousin (LIM) or Simmental (SIM) beef sires [4, 6]. However, calves from AAN inseminations have lower birth weight, lower carcass weight, lower average daily carcass gain, and higher fat levels than calves from LIM or SIM inseminations, either in purebreds or beef-on-dairy crosses [4, 7–9].

High selection pressures on production traits have caused a reduction in effective population size and increased inbreeding in modern cattle breeds, which lead to increased homozygosity [10, 11]. Thus, recessive deleterious mutations are more likely to be expressed in inbred populations, which lower fitness-related traits in many species [12]. For example, increased homozygosity is associated with a significant reduction in human height [13] and negatively affects milk production and fertility traits in cattle [14, 15]. These associations were particularly pronounced for long continuous stretches of autozygous segments (i.e., long runs of homozygosity [ROH]) and suggest that fitness reduction is likely caused by recent inbreeding [16]. Crossbreeding between genetically distinct breeds or species can counter these negative effects by introducing alternative alleles at loci with recessive deleterious mutations, often resulting in the heterosis of F1 hybrids [12]. However, the extent of heterosis is affected by the combination of breeds used for crossing [8, 9]. While heterosis gains for production traits are expected to be small, lowly heritable traits such as fertility, survival, robustness, or vitality can substantially be improved through crossbreeding [1, 11, 17, 18]. However, crossbreeding can also undermine previous breed-specific selection by reducing the dosage of adaptive alleles. While this is not relevant for previously selected milk production traits for beef-on-dairy crosses that are intended only for nonreproductive growing and fattening beef production systems, crossing might nevertheless affect direct calving ease [19] or tolerance to extreme environments, poor-quality diets, or diseases [20, 21].

In this study, we used F1 crosses between Brown Swiss (BSW) dairy females from Switzerland and AAN, LIM, or SIM beef bulls to assess the degree of genome complementarity that can be achieved with the three types of crossbreds. The BSW breed originated from approximately 170 Original Braunvieh ancestors that were imported from Switzerland to the United States for around 40 years at the end of the nineteenth century [22, 23]. The small number of founders, together with intense artificial selection for milk production on the original dual-purpose breed, resulted in relatively high levels of genomic inbreeding, low genetic diversity, and a large number of ROH [10, 22, 24]. AAN is a British beef breed phylogenetically most distant to BSW, while LIM and SIM are French beef and Swiss dual-purpose breeds, respectively, at a phylogenetically similar distance to BSW [25, 26]. Thus, we expected a substantial reduction of genomic inbreeding by crossing AAN, LIM, or SIM to BSW, and varying levels of genome complementarity between the three crossbreds.

To investigate which beef breed is most suitable to complement the genome of BSW females, we (i) quantified genomic inbreeding across the BSW genome and the reduction that can be realized by crossing with AAN, LIM, or SIM; (ii) detected candidate genomic regions and genes putatively under selection in BSW, which will be affected by crossbreeding; (iii) assessed genomic divergences between BSW and the aforementioned beef breeds, specifically at candidate genes under selection; and (iv) evaluated the potential phenotypic consequences on carcass traits in the three types of crossbreds and implications for marker-assisted selection. Thus, our study advances the identification of appropriate beef bulls mated to BSW females to improve the fitness and economic value of calves.

Methods

Animals

F1 crosses between AAN, LIM, or SIM sires and BSW dams (96 BSW×AAN, 130 BSW×LIM, and 96 BSW×SIM) were reared in the course of three successive fattening trials (T1, T2, and T3) conducted at the

experimental farm of Agroscope (Posieux, 620-m altitude a.s.l., Switzerland) from November 7, 2018, to March 10, 2022. Calves were purchased from Swiss commercial dairy farms or commercial traders at 4 to 6 weeks of age with a body weight (BW) of 65 to 75 kg. The number of sires for each cross type were representative of the bulls available for insemination in Switzerland at that time (i.e., 4 to 7 AAN, 12 to 17 LIM, and 4 to 6 SIM per trial). The animals included 90 bulls in T1, 96 bulls in T2, and 104 heifers in T3 (each with approximately equal proportions of cross types), and 32 BSW×LIM steers in T3. After a common weaning and rearing period, the animals were allocated to different experimental treatments and subsequently slaughtered at predetermined intermediate or final BW, according to the experimental design. Blood for DNA extraction was collected from the jugular vein with an EDTA vacutainer tube during the experiment. Only animals from the final slaughters ($n=248$) were considered for the phenotypic evaluations described below, while sequencing data were obtained for all animals ($n=322$).

Rearing and diets

During the common weaning and rearing periods, the animals were kept in free-stall barns on deep straw, with 30 to 35 calves each. The calves received reconstituted milk for 7 weeks, a commercial concentrated rearing feed, first-cut hay and maize silage (trials T1 and T2), or grass silage (T3). This period lasted for 96, 76, and 113 days until an average BW of 169, 154, and 174 kg with an average daily gain of 1.00, 1.07, and 0.90 kg/d for T1, T2, and T3, respectively. Allocation to different treatments within each trial was then made in a balanced way according to F1 cross, sex (T3), BW, and average daily gain during the weaning and rearing periods.

For the experimental period, the animals were kept in free-stall barns with multiple areas: a feeding area equipped with a trough for two animals and automatic drinkers, a straw-bedded resting area, and a concrete outdoor exercise yard. In trial T3, half of the animals spent 111 days during the summer on a mountain pasture in the Swiss Jura at an altitude of 1,200 m a.s.l., with integral grazing on a system of rotational grazing on natural grasslands (PA treatment), while the other half remained in the barn (SP treatment). Outside this period, all T3 animals received identical diets. All diets fed in the barn were formulated on the basis of feed recommendations for fattening cattle [27] and distributed ad libitum as total mixed rations (TMRs; Additional file 1: Table S1). Within each trial, the TMRs were iso-energetic and calculated to cover the recommended intakes for protein digestible in intestines, with a minimum ratio of 19 g of crude protein (CP) per MJ of net energy for meat production (NEV; Additional file 1: Table S2).

Slaughtering took place either at Agroscope's experimental slaughterhouse in Posieux (SLP; $n=52$) or at a commercial slaughterhouse (COS) 27 km away from Agroscope ($n=196$), and was conducted by stunning with captive bolt followed by exsanguination, in accordance with legally defined procedures (Order 455.1 of Swiss federal laws, 2008). In T1, 29 bulls were slaughtered at a weight corresponding to the Swiss Quality Beef (SQB) quality label, lighter than the standard (ST) weight for Switzerland, including 6 at the SLP at 474 ± 7 kg BW and 23 at the COS at 478 ± 8 kg BW. The remaining cattle were slaughtered at ST weight, including 12 at SLP at an average of 525 ± 7 kg BW and 48 at the COS at an average of 530 ± 12 kg BW. In T2, the average BW of the 18 bulls slaughtered at the SLP was 508 ± 10 kg, and that of the 53 bulls slaughtered at the COS was 520 ± 8 kg. In T3, the average BWs of the 12 heifers and 4 steers slaughtered at the SLP were 525 ± 5 and 527 ± 3 kg for the heifers and steers, respectively, and those of the 54 heifers and 18 steers slaughtered at the COS were 528 ± 8 and 527 ± 5 kg, respectively.

Body weight, carcass weight, chemical composition, and intramuscular fat measurements

Upon arrival in the experimental facilities and at slaughter, the cattle were weighed from 0700 to 1100 h after feed deprivation from 0000 h. Hot carcass weight was recorded less than 1 h after the slaughter, before half-carcasses were chilled at 4 °C for 24 h. The carcass chemical composition was determined by chemical analyses after the left half-carcass was ground and homogenized, as described by Lerch et al. [28], for the 52 cattle slaughtered at the SLP. For the remaining 196 cattle slaughtered at COS, the carcass composition was estimated using a dual X-ray absorptiometry (DXA) scan of the 11th left rib, using the equations reported by Xavier et al. [29], where the DXA variables (lean, fat and bone mineral content) were combined with the left half-carcass weight to estimate the hot carcass protein and lipid masses and proportions. Immediately after the DXA scan, the 11th rib was dissected [28], and the *Longissimus thoracis* muscle was sampled and stored at −20 °C pending lyophilization (DM determination), grinding, and laboratory DM (105 °C, 3 h) and intramuscular fat (ISO 6492:1999, petroleum ether extraction after acid hydrolysis) determinations.

Genomic analyses

Low-pass sequencing data were obtained for 322 F1 crosses, 34 of their sires, and 184 of their dams, as previously described [30]. Briefly, genomic DNA was extracted

from blood (F1s), semen (sires), or hair bulbs (dams) and sequenced as 150-bp paired-end reads at Neogen Europe (Auchincruive, UK) and imputed to a cattle reference panel consisting of 946 animals of diverse beef and dairy breeds [31] using Gencove's loimpute pipeline v0.1.5 [32, 33]. The ARS-UCD1.2 genome [34], including the Btau5.0.1 Y chromosome [35], was used as a reference. After removing 21 F1s and three dams that were identified as potential blood chimeras [30], 516 samples remained for genomic data analysis (8 AAN, 18 LIM, 8 SIM, 181 BSW, 89 BSW × AAN, 121 BSW × LIM, and 91 BSW × SIM).

Variants were filtered using GATK v4.2.6.1 [36] to remove variants with depths > 700 (i.e., twice the mean depth), > 25 low-confidence calls (i.e., 5% of samples), and minor allele frequencies < 5%. Only bi-allelic SNPs on autosomes were kept for further analysis (8,296,799 SNPs). Chromosomes were phased with Beagle v4.0 [37, 38] using 10 burn-in and 10 phase iterations and by providing a pedigree with parent–offspring duos and trios. As the availability of parent sequences differed between the samples, potentially affecting phasing accuracy, phased data were used only when noted.

Detection of homozygosity-by-descent segments

Inbreeding was estimated by modeling homozygosity-by-descent (HBD) or autozygous segments, that is, pairs of identical haplotypes within an individual inherited from a common ancestor [39]. HBD segments are conceptually similar to ROH, which are consecutive stretches of homozygous genotypes, interpreted as autozygous segments resulting from the mating of related individuals [40, 41]. HBD segments were detected separately for BSW, BSW × AAN, BSW × LIM, and BSW × SIM using the R package RZooRoH [39, 42]. Genotype probabilities from the unphased data were used as input to take genotype uncertainty into account and to avoid varying genotyping error rates among samples due to differences in pedigree completeness that was available for phasing.

RZooRoH implements a hidden Markov model to partition the genome of an individual into HBD and non-HBD segments, taking allele frequencies and genotyping error rates into account. Multiple HBD classes can be modeled that relate to different numbers of generations back to the common ancestor, where longer HBD segments are expected to result from more recent inbreeding. Different HBD classes k were specified by setting the rate R_k of each class, which corresponds to approximately twice the number of generations back to the common ancestor. The expected lengths of the HBD segments of class k then follow an exponential distribution with mean $1/R_k$ Morgans. Estimates of autozygosity associated with each HBD class are obtained by averaging the probability of a marker belonging to this class over the whole genome. The sum of the autozygosity estimates over all HBD classes provides a measure of the genomic inbreeding coefficient F_G.

RZooRoH was run with 16 predefined HBD classes with rates R_k taken from an exponential distribution with a power of 2 (2^1 to 2^{16}), one non-HBD class with rate 2^{16}, and an assumed genotyping error rate of 0.5%. Genomic inbreeding coefficients were computed cumulatively with the increasing age of inbreeding events, that is, by summing autozygosity estimates of HBD classes with increasing rates. Considering all HBD classes, HBD segments were summarized for each SNP position by computing the average segment length and the proportion of samples with a segment ≥ 1 Mb in length at this position. For the F1 crossbreds, summaries across offspring per sire were computed first and subsequently averaged within each breed to reduce bias resulting from the variable number of progenies per bull.

Detection of selection signatures

Three complementary methods were used to identify candidate genomic regions under selection in BSW: (i) the integrated haplotype score (iHS) to detect recent incomplete hard sweeps [43]; (ii) the haplotype homozygosity score H12 [44] to detect recent incomplete or complete hard and soft sweeps; and (iii) a composite likelihood ratio (CLR) test [45] to detect recent complete hard sweeps.

The R package rehh v3.2.2 [46, 47] was used to compute the iHS [43] on phased data. Alleles were not polarized, which means that major and minor alleles were contrasted instead of ancestral and derived alleles [48]. First, for each autosome, the function scan_hh() was used to compute the integrated extended haplotype homozygosity (iHH) for the major and minor alleles at all SNPs. Gaps between SNPs > 20 kb were rescaled to 20 kb, and computation of iHH was stopped at gaps > 200 kb [43]. Second, the function ihh2ihs() was used to compute iHS from iHH jointly for all autosomes, discarding SNPs with a minor allele frequency < 0.05, and standardizing logratios of major compared with minor alleles without binning by allele frequency, as it is recommended for unpolarized data. To identify candidate windows, the genome was divided into non-overlapping 50 SNP windows, and the number of outlier SNPs with |iHS| values larger than the genome-wide 0.99 percentile (i.e., |iHS| > 2.537640) was computed. Candidate windows were defined as those with the number of outlier SNPs larger than the 0.995 percentile across all windows (i.e., > 27), and adjacent candidate windows were merged.

The VCFtools v0.1.17 [49] --hapcount was used to compute haplotype frequencies in overlapping windows

of 500 SNPs, with increments of 50 SNPs on phased data. A custom R script was applied to compute the haplotype homozygosity score $H12 = H1 + 2p_1p_2$, where haplotype homozygosity $H1 = \text{sum}(p_i^2)$, with p_i being the frequencies of all observed haplotypes in a window and p_1 and p_2 being the frequencies of the two most frequent haplotypes in a window [44]. Candidate windows were defined as those with H12 larger than the 0.995 percentile across all windows, and overlapping or adjacent candidate windows were merged.

For the CLR test, which is based on the shape of the allele frequency spectrum, we refiltered our raw data set to additionally include all low-frequency variants unless they were fully fixed. First, we selected only BSW samples. Then, we used GATK v4.2.6.1 [36] to remove variants with a sequencing depth > 179 (i.e., twice the mean depth) or > 9 low-confidence calls (i.e., 5% of samples). Only bi-allelic polymorphic SNPs on autosomes were kept for further analysis (13,947,934 SNPs). A scan for selective sweeps was performed with SweepFinder2 v1.0 [45, 50] separately for each autosome on a grid space of 10 kb, using a precomputed folded empirical allele frequency spectrum from all autosomal SNPs. Candidate windows were defined as those with CLR values larger than the 0.995 percentile across all windows, and adjacent candidate windows were merged.

Candidate genes under selection were defined as genes retrieved using the R package biomaRt [51, 52] from the *Bos taurus* Ensembl version 110 [53] that overlapped with the candidate windows from the selection scans.

Overrepresentation analysis

An overrepresentation analysis of candidate genes under selection was performed with traits from the Animal QTLdb [54, 55] release 51. Gene sets were created for all 501 unique trait IDs in the database. For each gene in the *Bos taurus* Ensembl version 110, the gene was assigned as a member of a gene set if the start and end positions of the gene, ± 3 bp, contained the start and end positions of a QTL, respectively. As the length of each QTL was exactly 5 bp in the database, this corresponds to the QTL midpoint being at most 1 bp outside the gene boundaries. Four hundred sixty-five gene sets with at least one member were retrieved.

An overrepresentation analysis was performed for candidate genes identified in each selection scan using the R package clusterProfiler [56]. For each scan, background gene lists were created by only including genes that had the possibility of being detected (i.e., being located on autosomes and overlapping with genome windows that were created to perform the scan). The function enricher() was run using default settings, that is, excluding gene sets with < 10 or > 500 genes. Significantly overrepresented gene sets were defined as those with Benjamini and Hochberg [57]-adjusted p-values ≤ 0.05.

Analysis of genomic divergence

Genomic variants were functionally annotated using ANNOVAR [58] with the gff3 file from the *Bos taurus* Ensembl version 110. Furthermore, variants were assigned as *cis*-regulatory based on *cis*-eQTL and *cis*-sQTL associations retrieved from the cattle GenotypeTissue Expression atlas (cGTEx) [59]. Variants with Benjamini and Hochberg [57]-adjusted nominal p-values ≤ 0.05 across tissues were classified as *cis*-e or *cis*-s regulatory.

Highly divergent SNPs between BSW and beef breeds (hereinafter termed outlier SNPs) were defined as those with at most three non-homozygous BSW samples and at most three bulls across two beef breeds homozygous for the major BSW allele. This genotype-based metric was chosen to specifically detect variants nearly fixed in the BSW breed while allowing bulls to be heterozygous and permitting one of the beef breeds to not deviate from BSW. To achieve a balanced comparison among the breeds, LIM bulls were down-sampled prior to these computations to eight individuals (five with the largest number of progenies plus three randomly selected bulls) to match the available number of eight samples for AAN and SIM. In addition, separately across the eight bulls from each beef breed, breed-specific outlier SNPs were identified. Analogous to the definition above, they were defined as variants with at most three BSW samples not homozygous and at most one bull within the focal breed homozygous for the major BSW allele.

Pearson chi-squared tests were performed using the chisq.test() function in R to test for an enrichment of functional classes within outlier SNPs and an enrichment of outlier SNPs within candidate windows under selection combined across all selection scans. Cells in the contingency table with expected counts < 5 were removed prior to the chi-square tests. Deviations between the observed and expected counts per cell were defined as significant when adjusted standardized residuals exceeded Bonferroni-corrected z criticals of ± 1.96 [60].

In-depth analysis of candidate genes under divergent selection

For each candidate gene under selection in the BSW cows that were identified in any of the three selection scans, SNPs were assigned to five functional classes: UTR3 or UTR5; splicing or ncRNA splicing; intronic or ncRNA intronic; synonymous SNP; or nonsynonymous, stopgain, or stoploss variants. The genotype probabilities of all 181 cows and 34 bulls in the data set were extracted for each candidate gene and functional class, and input files for

the software ngsDist [61] were created using BCFtools v1.17 [62] and PLINK v1.9b [63, 64]. Raw p-distances between all pairs of samples were computed from the extracted genotype probabilities with ngsDist.

The R package ape [65] was used to compute neighbor-joining trees from p-distances for each gene and functional class, employing the nj() function. The pam() function from the R package cluster [66] was used to perform the partitioning of the p-distances into k clusters around the medoids, setting $k=2$. For each gene and functional class, the cluster containing most BSW samples was determined, and the samples from each breed assigned to this main BSW cluster were counted. Genes with strong differentiation between BSW and beef bulls were defined as those with at most three BSW samples outside the main BSW cluster and at most three bulls across two beef breeds inside the main BSW cluster. For genes with strong differentiation between BSW and beef bulls, haplotypes for all samples at variants assigned as exonic or UTR were pulled from phased data using VCFtools [49] and unique haplotypes were plotted in R.

Phenotype associations

For candidate genes with strong differentiation between BSW and beef bulls, genotypes were extracted at nonsynonymous outlier SNPs from phased data, keeping only one SNP in case of complete linkage disequilibrium (LD) with other variants. These genotypes were then used to perform linear mixed-model based association tests with BW, hot carcass weight, chemical composition, and *Longissimus thoracis* intramuscular fat in F1s. To prevent false positive associations due to relatedness structure, a genomic relationship matrix (GRM) was included in all models [67, 68]. GRMs were built from autosomal SNPs in F1s with a minor frequency > 5% and which were then LD-pruned with PLINK v2.00a [69, 70] --indep-pairwise 200 10 0.4 (i.e., using sliding windows of 200 SNPs, steps of 10 SNPs, and pruning variants within a window with squared correlation > 0.4). GRMs were computed from the 230,568 retained SNPs with GCTA v1.94.1 [71, 72] --make-grm-gz, either including all variants or excluding the chromosome of the focal locus. Further covariates included sire breed, sex, BW at slaughter (scaled to tons), and the age of the calf at the time of measurement (scaled to years) as fixed effects, and trial or ration as random effects (Additional file 1: Table S3). Different subsets of covariates were chosen as appropriate for each trait, and null models including the full GRM but excluding any single marker regressions were fitted using the fitNullModel() function from the GENESIS [73] package in R. The variance components of the random effects were then estimated using the average information REML procedure. The null model with the lowest AICc, or a simpler model with a difference in AICc < 2 from the model with the lowest AICc, was chosen for each trait. The null model was then refitted with the GRM that excluded the chromosome of the focal locus, and supplied to the assocTestSingle() function from the GENESIS package, together with the focal SNP, to test for an association between the trait and each SNP genotype. Separately for each SNP, the association between genotype and trait was defined as significant when the Benjamini and Hochberg [57]-adjusted p-value was ≤ 0.05.

Results

Detection of HBD segments

The computation of genomic inbreeding coefficients F_G from 16 HBD classes with rates R_k ranging from 2 to 65,536 (where R_k approximately corresponds to twice the number of generations back to the common ancestor) revealed that the distribution of ancient inbreeding equal to or higher than rates $R_k=512$ is very similarly shaped for BSW and F1 crossbreds, with two peaks at approximately $R_k=1024$ and $R_k=8192$ (Fig. 1, Additional file 2: Fig. S1). Equal or below $R_k=128$, relevant amounts of inbreeding were detected only for BSW, with a marked peak at $R_k=64$, which led to an increase in total genomic inbreeding in BSW compared with crossbreds (BSW mean $F_G=0.461$). No noticeable inbreeding occurred prior to $R_k=16$. Only minor differences in timing and total levels of inbreeding were observed among crossbreds, with BSW × SIM having slightly more recent inbreeding and BSW × AAN having slightly lower levels of total genomic inbreeding than the other crosses (BSW × AAN, mean $F_G=0.381$; BSW × LIM, mean $F_G=0.392$; BSW × SIM, mean $F_G=0.390$; Additional file 2: Fig. S1).

Likewise, the genomic distributions of the lengths of HBD segments (Fig. 2), and the proportion of samples with a SNP part of an HBD segment ≥ 1 Mb (Additional file 2: Fig. S2), were much higher in BSW than in crossbreds. In particular, long HBD segments and high proportions of samples carrying long segments were observed in BSW on BTA5 and BTA6, while these two chromosomes were not particularly conspicuous in crossbreds. Nevertheless, crosses showed some peaks of long or common HBD segments at some genomic regions, which were sometimes shared among breeds (e.g., on BTA7; Fig. 2 and Additional file 2: Fig. S2).

Detection of selection signatures

Three complementary approaches were used to identify the candidate genomic windows under selection in BSW. As these methods rely on different assumptions regarding the completeness and softness of selective sweeps, the detected genomic windows differed among the methods

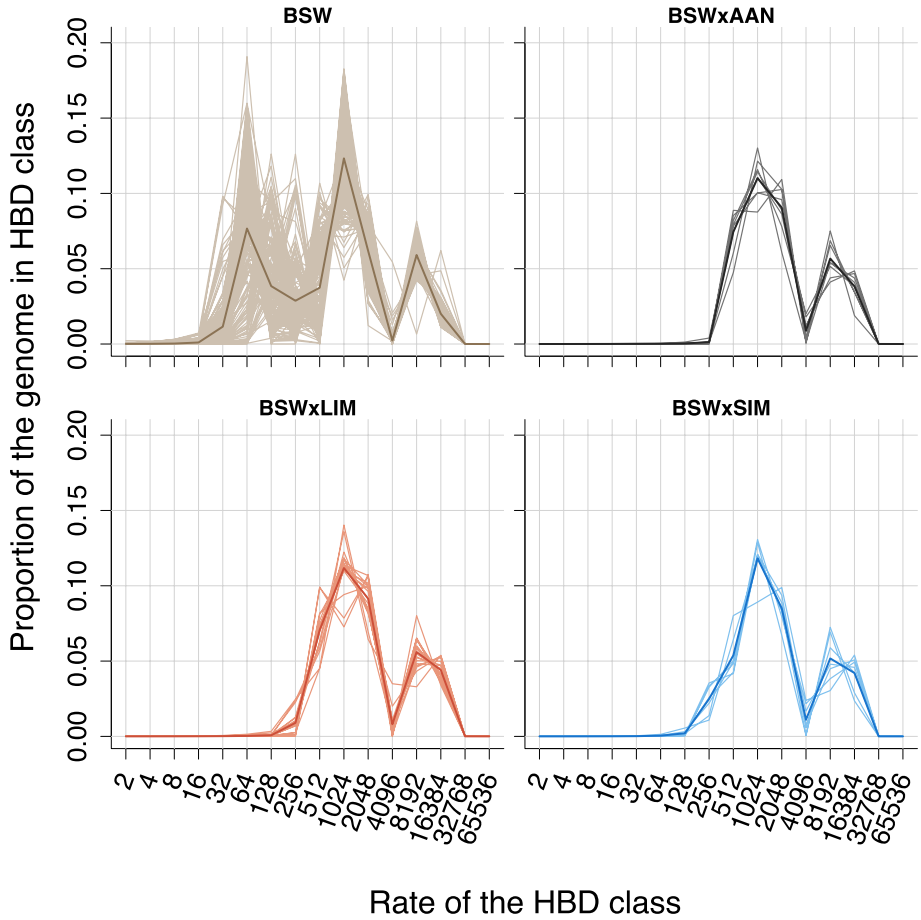

Fig. 1 Contribution of each HBD class to genome-wide inbreeding estimates per breed. Higher rates model shorter HBD segments and thus inbreeding events that are more ancient. The thin lines indicate individual values for BSW or averages across F1s per sire for crossbreds, and the bold lines indicate the averages of these values

but also showed overlaps (Fig. 3). For example, H12 identified a large genomic region on BTA6 with unusually high levels of haplotype homozygosity. Owing to the complete lack of genetic variation in large parts of this region, no iHS could be computed, as the score requires at least low frequencies of the minor variant. In addition, candidate windows were identified on all autosomes by any of the three scans except on BTA23, where only the CLR test showed a strong signal (Additional file 1: Tables S4–S6). With the iHS, H12, and CLR tests, the candidate windows contained 163, 232, and 255 candidate genes, respectively, with 585 genes across all scans, where only three genes were common to all scans (*UGGT2*, *PPFIBP2*, and *MTARC2*).

Overrepresentation analysis

Genes within candidate windows under selection were overrepresented in several gene sets obtained from trait associations listed in the Animal QTLdb [54]. For iHS candidate genes, significant gene sets were mainly related to milk traits (including milk β-lactoglobulin content, milk phosphorylated α_{S2}-casein percentage, milk unglycosylated κ-casein percentage, milk rennet coagulation time, and cheese protein recovery), one meat and carcass trait (shear force), one production trait (BW gain), and one health trait (ketosis; Additional file 2: Fig. S3). For the H12 and CLR candidate genes, significant gene sets were predominantly related to meat and carcass or production traits (including metabolic body weight, bone weight, *longissimus* muscle area, carcass weight, subcutaneous fat thickness, and average daily gain), which were all strongly affected by genes located on BTA6 (Additional file 2: Figs. S4 and S5). One significant gene set of each scan was further related to reproduction (first service conception for H12, calving index for CLR). For H12 candidate genes, additional gene sets were related to milk traits (including milk lactose content and κ-casein percentage; Additional file 2: Fig. S4).

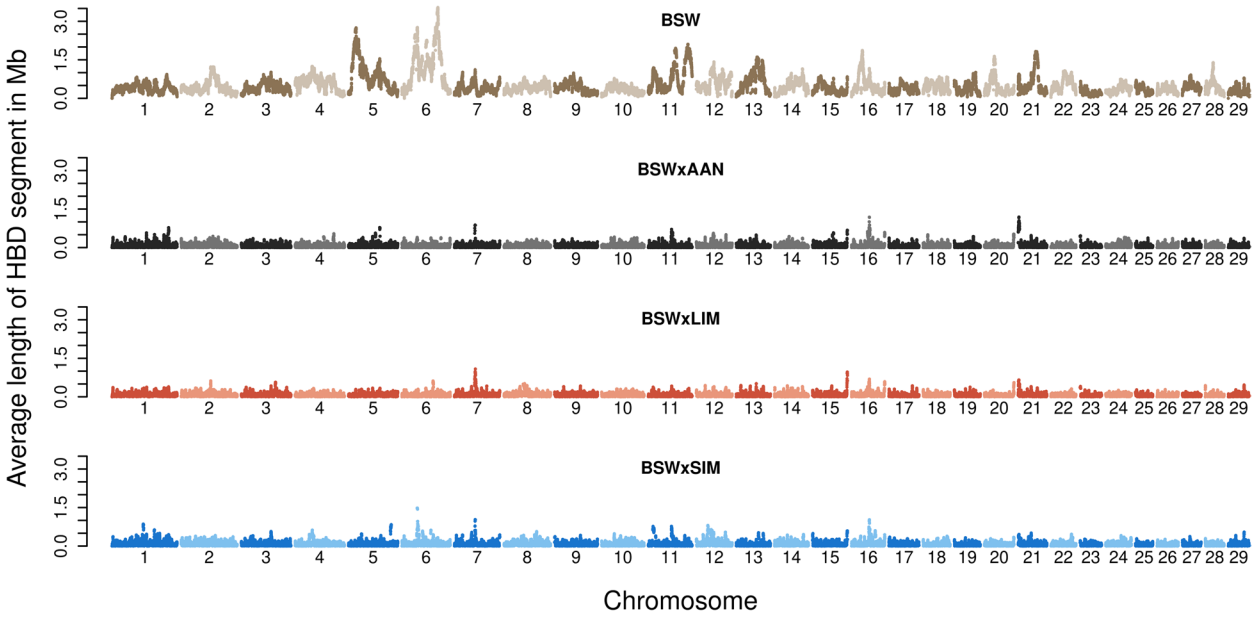

Fig. 2 Average HBD segment lengths at each SNP position along the genome for BSW and crossbreds. For crossbreds, averages were computed across F1s from each sire and subsequently averaged for each sire breed

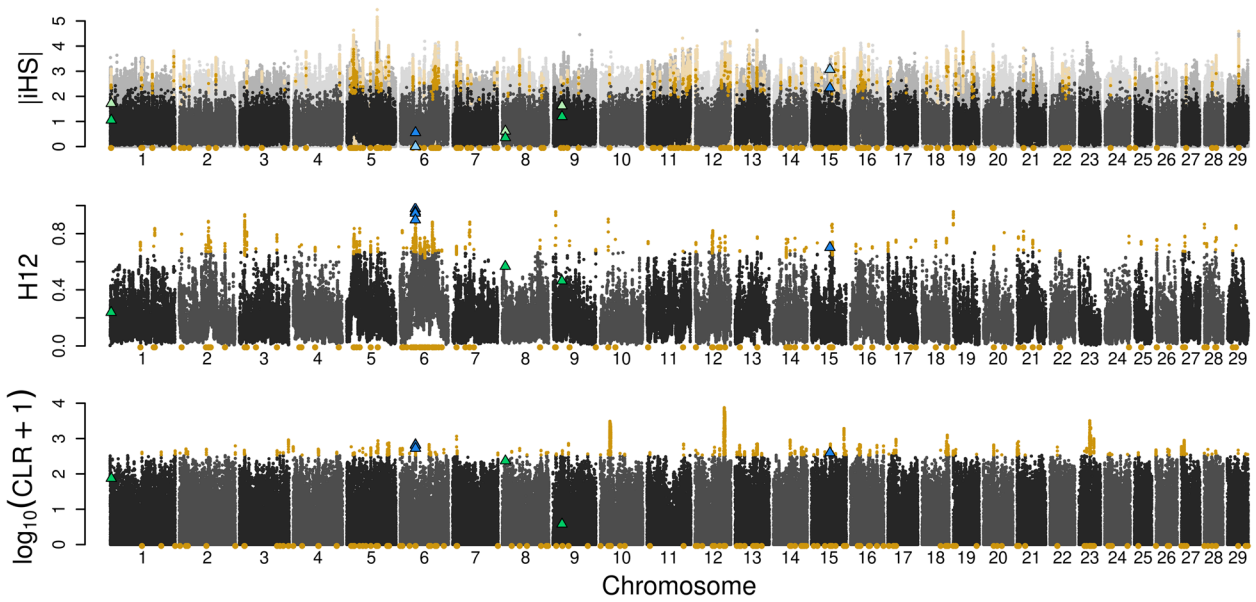

Fig. 3 Scores from three selection scans along autosomes in the BSW. In the top panel, dots in pale colors give |iHS| for each SNP, superimposed by the mean |iHS| of each non-overlapping 50-SNP window in dark colors. In the middle panel, each dot indicates H12 computed in overlapping windows of 500 SNPs, with increments of 50 SNPs. In the bottom panel, the dots show $\log_{10}$-transformed CLR scores on a grid space of 10 kb. In each panel, yellow dots indicate SNPs, windows, or grid midpoints that fell within the candidate genomic windows under selection for each selection scan. These regions are highlighted in yellow on the *x*-axis. The blue and green triangles highlight the top scores of the regions overlapping the coordinates of the candidate genes under selection within or outside the candidate windows combined across all scans

Analysis of genomic divergence

When investigating all autosomal SNPs, 3,758 of 8,296,799 variants (0.05%) were identified as outliers, with high levels of genomic divergence between BSW and beef breeds (i.e., divergence outliers). Compared with non-outliers, outliers were underrepresented in intergenic regions and overrepresented in intronic, upstream, exonic, ncRNA intronic, and UTR3 regions

(Additional file 1: Table S7; Pearson $\chi^2 = 646.44$, $df = 6$, N = 8,347,169, $p < 2.2e\text{-}16$). Outliers were also overrepresented in variants assigned as *cis*-regulatory (Additional file 1: Table S8; *cis*-eQTL, Pearson $\chi^2 = 437.32$, $df = 1$, N = 8,296,799, $p < 2.2e\text{-}16$; *cis*-sQTL, Pearson $\chi^2 = 291.92$, $df = 1$, N = 8,296,799, $p < 2.2e\text{-}16$). Furthermore, outliers were strongly enriched in candidate windows under selection (Additional file 1: Table S9; Pearson $\chi^2 = 18{,}292$, $df = 1$, N = 8,296,799, $p < 2.2e\text{-}16$).

Of 19 nonsynonymous outlier SNPs, 15 were within the candidate windows. These SNPs were part of four tightly clustered genes *FAM184B*, *NCAPG*, *DCAF16*, and *LCORL* on BTA6, and *PPFIBP2* on BTA15 (Additional file 1: Tables S10 and S11). The remaining four nonsynonymous outlier SNPs were outside the candidate windows within the genes *URB1* on BTA1, *NEIL2* on BTA8, and *IBTK* on BTA9 (Additional file 1: Tables S10 and S11). Of 897 *cis*-eQTL and 974 *cis*-sQTL assigned as outlier SNPs, 79 *cis*-eQTL and 180 *cis*-sQTL SNPs were within candidate windows and clustered in three narrow regions on BTA6, BTA13, and BTA15, and one larger genomic region on BTA6 (Additional file 1: Table S12). These regions largely overlapped with the gene boundaries of *FAM184B*, *NCAPG*, *DCAF16*, *LCORL*, and *PPFIBP2*. A large proportion of *cis*-regulatory outlier SNPs within *FAM184B* and *NCAPG* were annotated as being involved in the regulation of the upstream genes *HERC6* (19 SNPs) and *ABCG2* (56 SNPs), respectively, and were mainly assigned as *cis*-sQTL (Additional file 1: Table S12).

When investigating breed-specific outlier SNPs separately, AAN showed the highest level of differentiation from BSW, with 20,164 breed-specific outlier SNPs, while LIM and SIM showed much fewer differences, with 6,738, and 8,026 breed-specific outlier SNPs, respectively (Additional file 1: Table S13a). Most of these variants were annotated as intergenic or intronic (Additional file 1: Table S13b), followed by variants assigned as *cis*-regulatory, of which 2,071, 1,023, and 353 were assigned as *cis*-eQTL; and 2,511, 1,076, and 545, as *cis*-sQTL for AAN, LIM, and SIM, respectively (Additional file 1: Table S14). Furthermore, 71, 31, and 22 variants were assigned as nonsynonymous (Additional file 1: Table S13c), which were located in 50, 15, and 19 genes for AAN, LIM, and SIM, respectively (Additional file 1: Tables S10 and S15). Four of the genes intersected between AAN and LIM (*FAM184B*, *DCAF16*, *NCAPG*, and *LCORL*), two genes intersected between AAN and SIM (*URB1* and *NEIL2*), and none intersected between LIM and SIM.

In-depth analysis of the candidate genes under divergent selection

K-medoids clustering applied to 585 candidate genes putatively under selection in BSW identified five genes that exhibited strong differentiation between BSW and beef bulls for nonsynonymous, stopgain, or stoploss variants (Fig. 4). These genes are exactly the same as identified above (*FAM184B*, *NCAPG*, *DCAF16*, and *LCORL* on BTA6, and *PPFIBP2* on BTA15). For any of the other functional classes, only subsets of these genes and no additional ones were identified. For the genes on BTA6, most SIM bulls shared the main cluster with the BSW cows, while for *PPFIBP2*, mainly LIM and some SIM bulls were assigned to the same cluster as most BSW samples. For all five genes, the BSW samples showed markedly reduced genetic variation (Fig. 4).

The frequencies of unique haplotypes composed of exonic and UTR variants from the five strongly differentiated genes between the BSW and beef breeds are given in Additional file 1: Tables S16 and S17 and Additional file 2: Figs. S6 and S7. Genes *FAM184B*, *NCAPG*, *DCAF16*, and *LCORL* are tightly clustered on BTA6, where *NCAPG* partially overlaps with *FAM184B* and *DCAF16* owing to their location on opposite strands (Additional file 2: Fig. S6). The BSW and SIM samples were nearly fixed for one single haplotype, which was largely composed of alternative alleles, while haplotypes mainly found in AAN and LIM carried almost only reference alleles. Directly upstream *of FAM184B*, the peaks from the H12 and CLR tests covered two additional genes, *LAP3* and *MED2*8, which were located on a haplotype that was nearly fixed in BSW (Additional file 2: Figs. S8, S9). However, these two genes did not include nonsynonymous SNPs, and variants of other functional classes within these genes were not particularly differentiated between the BSW and beef breeds. For *PPFIBP2* on BTA15, the two main haplotypes, differing by one UTR5 substitution, were most common in BSW (Additional file 2: Fig. S7). These haplotypes were composed of a mix of reference and alterative alleles.

Phenotype associations

As a large number of bulls were heterozygous for nonsynonymous outlier SNPs at candidate genes with strong differentiation between BSW and beef bulls (Additional file 2: Fig. S10), genotypes likewise differed among F1s, which allowed testing for associations with the measured phenotypes. Owing to the complete LD among variants, 7 of 15 divergence outliers within candidate genes were chosen as representative for association analysis. Based on AICc, null models that included BW at slaughter provided a substantially better fit to hot carcass weight, hot carcass protein mass, and hot carcass lipid mass than models excluding BW, and age at slaughter improved the fit of null models for hot carcass lipid mass, hot carcass lipid proportion, and *Longissimus thoracis* intramuscular fat (Additional file 1: Tables S3 and S18). Age at slaughter

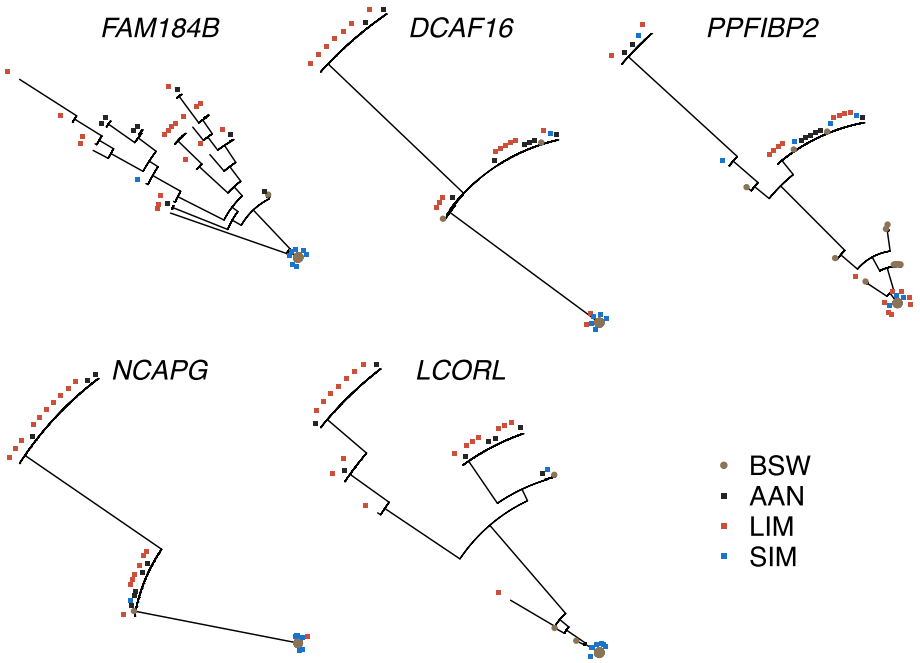

Fig. 4 Neighbor-joining trees from p-distances. The p-distances were computed using nonsynonymous variants for five candidate genes with strong differentiation between the BSW and beef breeds. The tip symbols are colored by breed and slightly offset for better visibility. In each tree, the large majority of BSW samples cluster at the bottom right node, as indicated with an enlarged symbol

and BW at slaughter showed a moderately positive correlation (Pearson's $r=0.46$, $p=4.8 \times 10^{-13}$). After correcting for multiple testing, significant associations were detected for SNPs within *FAM184B*, *DCAF16*, *NCAPG*, and *LCORL* on hot carcass weight and for SNPs within *NCAPG* and *LCORL* on hot carcass protein mass, with negative correlations between the dosage of the BSW-like allele and the two phenotypes (Table 1; Additional file 1: Tables S3, S18, and S19; Additional file 2: Figure S11). In addition, some associations (unadjusted p-value ≤ 0.05) were found for SNPs within *FAM184B* and *DCAF16* on hot carcass protein mass, a SNP within *FAM184B* on hot carcass lipid mass, hot carcass lipid proportion, and *Longissimus thoracis* intramuscular fat (again with the dosage of the BSW-like allele being negatively correlated with the phenotypes), and SNPs within *PPFIBP2* on hot carcass protein proportion (with the dosage of the BSW-like allele being positively correlated with the phenotype; Table 1). No significant associations were detected for the remaining traits (BW at arrival and age at slaughter).

Discussion

We used low-pass sequencing data from BSW dairy cows, three breeds of beef bulls, and their progeny to assess the effect of sire on genome complementarity in beef-on-dairy crosses. Our results show a substantial reduction in recent genomic inbreeding in F1s, while few genes putatively under selection in the BSW breed feature new genotypes in crosses that are not already segregating within the BSW breed. The absence of completely fixed nonsynonymous genetic variation between the BSW and all three beef breeds, and the presence of genetic variation among and within beef breeds at loci putatively under selection in the BSW breed suggest that the desired phenotypes of BSW calves intended for beef production can be modulated through marker-assisted selection of mating pairs.

HBD segments

The estimation of HBD segments revealed higher levels of genomic inbreeding in the BSW breed than in the F1s, as expected for crosses between divergent breeds. Nevertheless, the total genomic inbreeding resulting from very short segments was considerable, even in crossbreds, and likely reflects shared ancient inbreeding predating breed formation. A similar increase in the number of very short HBD segments at $R_k=1024$ to $R_k=2048$ has been observed previously with >500,000 SNPs in 11 cattle breeds [74]. Owing to the lower SNP density, the previous study likely lacked the resolution to detect the second peak of extremely short HBD segments at $R_k=8192$, which was evident in our results. In contrast to the crossbreds, the BSW breed showed a peak of longer HBD segments from more recent inbreeding at $R_k=64$ (Fig. 1

Table 1 Results from association tests between phenotypic measurements in F1 crosses and seven nonsynonymous outlier SNPs

SNP ID Position Gene name	Trait	Est	Est SE	Score	Score SE	*P*-value	*P* adj
rs110867784 6:37,234,954 *FAM184B*	BW at arrival	0.23	0.98	0.24	1.02	0.8122	0.8122
	Age at slaughter	0.01	0.02	62.86	65.60	0.3380	0.3863
	HC weight	2.19	1.41	1.11	0.71	0.1187	0.2281
	HC protein mass	0.48	0.33	4.40	3.01	0.1444	0.2281
	HC lipid mass	−1.83	0.87	−2.42	1.15	**0.0355**	0.1180
	HC protein proportion	0.10	0.07	19.40	14.18	0.1711	0.2281
	HC lipid proportion	−0.65	0.30	−7.30	3.34	**0.0291**	0.1180
	LT intramuscular fat	−0.23	0.11	−17.86	8.88	**0.0442**	0.1180
rs110281825 6:37,300,810 *FAM184B*	BW at arrival	0.62	1.02	0.60	0.98	0.5426	0.8755
	Age at slaughter	−0.01	0.01	−90.11	91.98	0.3273	0.8727
	HC weight	−3.35	1.05	−3.05	0.95	**0.0014**	**0.0111**
	HC protein mass	−0.55	0.28	−7.19	3.61	**0.0467**	0.1867
	HC lipid mass	−0.15	0.72	−0.29	1.39	0.8337	0.8755
	HC protein proportion	0.01	0.06	2.90	16.99	0.8646	0.8755
	HC lipid proportion	0.11	0.25	1.86	4.03	0.6447	0.8755
	LT intramuscular fat	0.01	0.08	1.85	11.79	0.8755	0.8755
rs109126348 6:37,300,872 *FAM184B*	BW at arrival	1.30	1.03	1.22	0.97	0.2085	0.5561
	Age at slaughter	−0.01	0.01	−87.51	83.08	0.2922	0.5844
	HC weight	−3.58	1.16	−2.67	0.86	**0.0020**	**0.0159**
	HC protein mass	−0.62	0.31	−6.56	3.26	**0.0441**	0.1764
	HC lipid mass	−0.37	0.79	−0.58	1.26	0.6445	0.9866
	HC protein proportion	0.01	0.07	2.42	15.31	0.8743	0.9866
	HC lipid proportion	0.09	0.27	1.13	3.64	0.7568	0.9866
	LT intramuscular fat	0.00	0.09	−0.18	10.68	0.9866	0.9866
rs210190561 6:37,315,480 *DCAF16* *LCORL*	BW at arrival	0.59	1.03	0.55	0.97	0.5673	0.9076
	Age at slaughter	−0.01	0.01	−52.07	90.28	0.5641	0.9076
	HC weight	−3.30	1.07	−2.88	0.93	**0.0021**	**0.0165**
	HC protein mass	−0.55	0.28	−6.91	3.54	0.0507	0.2026
	HC lipid mass	0.00	0.73	0.00	1.37	0.9993	0.9993
	HC protein proportion	0.00	0.06	0.72	16.63	0.9657	0.9993
	HC lipid proportion	0.16	0.25	2.49	3.95	0.5286	0.9076
	LT intramuscular fat	0.01	0.09	1.79	11.57	0.8770	0.9993
rs109570900 6:37,343,379 *NCAPG* *LCORL*	BW at arrival	0.85	1.02	0.81	0.98	0.4062	0.8965
	Age at slaughter	0.00	0.01	9.37	85.03	0.9122	0.9226
	HC weight	−3.69	1.13	−2.90	0.89	**0.0011**	**0.0085**
	HC protein mass	−0.75	0.29	−8.68	3.40	**0.0107**	**0.0427**
	HC lipid mass	−0.21	0.76	−0.37	1.32	0.7816	0.9226
	HC protein proportion	−0.01	0.06	−1.55	15.96	0.9226	0.9226
	HC lipid proportion	0.15	0.26	2.21	3.80	0.5603	0.8965
	LT intramuscular fat	0.06	0.09	6.83	10.99	0.5345	0.8965
rs109696064 6:37,403,795 *LCORL*	BW at arrival	0.79	1.09	0.66	0.91	0.4703	0.7525
	Age at slaughter	−0.01	0.03	−7.48	31.85	0.8143	0.8143
	HC weight	−0.96	3.03	−0.10	0.33	0.7517	0.8143
	HC protein mass	0.37	0.82	0.56	1.22	0.6477	0.8143
	HC lipid mass	−1.79	2.11	−0.40	0.47	0.3944	0.7525
	HC protein proportion	0.15	0.17	5.09	5.76	0.3769	0.7525
	HC lipid proportion	−0.55	0.73	−1.02	1.37	0.4552	0.7525
	LT intramuscular fat	0.30	0.25	5.04	4.08	0.2163	0.7525

Table 1 (continued)

SNP ID Position Gene name	Trait	Est	Est SE	Score	Score SE	*P*-value	*P* adj
rs133933788 15:45,066,565 *PPFIBP2*	BW at arrival	−0.71	0.92	−0.83	1.08	0.4412	0.6240
	Age at slaughter	0.00	0.01	26.83	98.67	0.7857	0.8979
	HC weight	1.02	0.98	1.07	1.02	0.2962	0.5924
	HC protein mass	0.45	0.26	6.81	3.90	0.0808	0.3234
	HC lipid mass	−0.48	0.66	−1.10	1.51	0.4680	0.6240
	HC protein proportion	0.11	0.05	35.98	18.33	**0.0497**	0.3234
	HC lipid proportion	−0.24	0.23	−4.58	4.36	0.2935	0.5924
	LT intramuscular fat	0.00	0.08	0.76	12.62	0.9521	0.9521

Position indicates BTA:position in bp of the test SNP, and gene name gives candidate genes containing nonsynonymous SNPs for which the test SNP was in full linkage disequilibrium with. Est and Est SE provide an approximation of the effect size estimate for each additional copy of the BSW-like allele and an approximation of the standard error of the effect size estimate, Score and Score SE give the value and estimated standard error of the score function, *P*-value gives the score *p*-value, *P* adj. reports the Benjamini and Hochberg-adjusted *p*-value. Bold text indicates *p*-values ≤ 0.05.

The following abbreviations were used: *BW* Body weight, *HC* Hot carcass, *LT Longissimus thoracis*

and Additional file 2: Fig. S1). This rate is larger than that detected in a previous work at $R_k = 16$ [74], possibly due to the merging of adjacent segments in panels with lower SNP densities or the breaking up of longer segments with higher genotyping error rates in low-pass sequencing data. Assuming a generation time of 5 years [75], $R_k = 64$ corresponds to inbreeding occurring at around 160 years ago and is thus coincident with the time of BSW breed formation at the end of the nineteenth century [23]. By comparing the genome-wide distribution of HBD segments among crossbreds (Fig. 2 and Additional file 2: Fig. S2), we observed multiple peaks of the remaining long homozygous stretches at some genomic regions, which might result from shared selection pressures between the BSW and beef breeds or shared genome properties such as reduced recombination or mutation rates. Overall, we observed no marked differences in the number, length, and location of HBD segments among the three crossbreds, indicating that all three sire breeds should be similarly suitable to complement the genetic variation in the BSW breed that was reduced by inbreeding.

Selection scans and genomic divergence between the breeds

By using three distinct methods to detect genomic regions potentially under selection in BSW (i.e., candidate windows), 585 candidate genes were identified. Genes detected with the iHS scan were particularly overrepresented for QTL associated with milk traits (Additional file 2: Fig. S3). By contrast, the genes detected using the H12 or CLR tests were mainly associated with meat and carcass or production traits (Additional file 2: Figs. S4 and S5). As the iHS scan requires lower frequencies of the beneficial allele than the H12 and CLR tests, where the beneficial allele can even be fixed [76], these results suggest more recent or weaker selection pressures for milk than for meat and carcass or production traits.

Outlier SNPs with high levels of genomic divergence between the BSW and beef breeds showed enrichment in candidate windows (Additional file 1: Table S9). This means that at least some of the detected sweep signatures contain alleles that are nearly fixed in the BSW breed but are segregating in the beef breeds, suggesting BSW-specific selection. Furthermore, outlier SNPs were enriched for variants within or near genes, particularly for SNPs located in introns (Additional file 1: Table S7), and were enriched for variants assigned as *cis*-regulatory (Additional file 1: Table S8). This indicates that the overall genomic divergence between BSW and beef breeds involves both coding and regulatory mutations, where regulatory variants clearly outnumber amino acid-changing mutations.

In-depth analysis of candidate genes under BSW-specific selection

From 585 genes within the candidate windows under selection in BSW, only five genes showed strong genetic differentiation at 15 nonsynonymous variants between the BSW breed and the investigated beef breeds. The *FAM184B–LCORL* complex on BTA6, containing *FAM184B*, *NCAPG*, *DCAF16*, and *LCORL*, is already known as a pleiotropic QTL that has been associated with several production (e.g., dry matter intake, average daily gain, and BW), morphological (e.g., stature and

body conformation), and meat and carcass traits (e.g., backfat thickness, ribeye area, lean meat yield, and hot carcass weight) in various studies and across diverse cattle breeds [77–81]. In addition, associations with direct calving ease [82, 83] or increased calving difficulties [79] have been reported. Some associations were also found for milk protein and fat percentages [84]. In humans, *LCORL* has been associated with birth weight [85] and height [86–89], sometimes together with *NCAPG* [86, 89].

Strong LD between variants within and adjacent to *NCAPG* and *LCORL* resulted in discordance in the determination of the causal gene [78, 90–92]. A mutation in *NCAPG* has been suggested to affect fetal growth, although additional loci in the region could not be excluded [93]. A missense variant in *NCAPG* has been associated with carcass weight, live weight, and net meat weight [94]. Likewise, the most significant associations with growth have been found for *NCAPG* [95]. By contrast, increased calving difficulties and increased calf size have been found to be associated with a haplotype encompassing the *LCORL* gene [79], and *LCORL* has been suggested as a causal gene for stature [78]. Other studies have found associations between feed efficiency and carcass traits with both *NCAPG* and *LCORL* [80, 81]. Finally, QTL and lead SNPs on BTA6 were not found to be unique across breeds or traits [83]. Here, we found that nonsynonymous SNPs within the *FAM184B–LCORL* complex, including those with significant associations to hot carcass weight and protein mass, are in complete LD between *DCAF16* and *LCORL*, or between *NCAPG* and *LCORL* (Additional file 1: Table S19; Additional file 2: Fig. S6), rendering the identification of causal variants impossible. In contrast to previous studies, we could not confirm an association of the *FAM184B–LCORL* complex to the BW of young calves, although this might be caused by the unstandardized rearing conditions prior to arrival or their advanced age.

We detected very low levels of haplotype diversity and strong signatures of selection in the *FAM184B–LCORL* complex in the BSW breed in the H12 and CLR tests, while iHS was not applicable because of the below-threshold minor allele frequencies in this region. With the ARS-UCD1.2 assembly as a reference, many nonsynonymous SNPs were fixed for the alternate allele and corresponded mostly to the putative ancestral alleles for *FAM184B*, *DCAF16*, and the first part of *NCAPG*, but derived alleles for the second part of *NCAPG* and *LCORL* (Ensembl version 110) [53]. Signatures of selection additionally covered the genes *LAP3* and *MED28* directly upstream of *FAM184B* (Additional file 2: Figs. S8 and S9), which have likewise been associated with production [81] or meat and carcass traits [80]. However, these two genes were not particularly differentiated between the BSW and investigated beef breeds.

Low levels of haplotype diversity and signatures of selection in the *FAM184B–LCORL* region have previously been reported for Original Braunvieh [96, 97]. As in our study on the BSW breed, in Original Braunvieh, most SNPs were found to be fixed for the alternate allele, and the CLR test revealed the highest peak at the *NCAPG* gene [96]. The reduced haplotype diversity in this region is restricted to a subset of breeds, while others do not show such reduced variation [78, 97]. This suggests very strong and shared selection pressure in some breeds but not others. Our results that show that almost all SIM and BSW animals share the same haplotype in this region are in line with the previous findings. However, owing to the strong LD and pleiotropy at the locus, the selected phenotype remains uncertain. Given our detected effects of BSW-like alleles on diminishing hot carcass weight and protein mass, direct selection might have occurred on these or other traits correlated with them. Alternatively, trade-offs between nutrient allocation for milk production and muscle and adipose tissue growth might have resulted in indirect selection at this locus. A large number of SNPs within *FAM184B* and *NCAPG* were reported as *cis*-sQTL, affecting the upstream genes *HERC6* and *ABCG2*, respectively [59] (Additional file 1: Table S12). These two genes have been associated with milk fat and protein percentages in dairy cattle [98] and cheese-making properties in Italian BSW cows [99]. Thus, the *FAM184B–LCORL* complex is a strong candidate for further studies that elucidate the genetic pathways and trade-offs involved between dairy and beef productions in cattle.

To date, very little is known about the role of the fifth highly differentiated gene, *PPFIBP2* on BTA15, a protein-binding gene found in the presynaptic active zone and involved in neuromuscular junction development [100, 101]. In mice, it has been associated with morphine-induced locomotor activity and rearing [102]. In rats, it has been identified as a potential target of miRNA regulation after muscle reinnervation, resulting in recovery from muscle atrophy [103]. The gene was also found within a QTL for milk yield, fat percentage, and protein percentage in North American Holstein cattle [104]. It was detected with all three selection scans and includes several *cis*-regulatory variants and two nonsynonymous mutations that were highly differentiated between BSW and most beef bulls. Similar to the genes in the *FAM184B–LCORL* complex, BSW-like variants still segregate in the beef breeds, which could, together with the reported association to both neuromuscular and milk production traits, likewise point to an involvement in a trade-off between dairy and beef productions.

Only three genes located outside the candidate windows contained nonsynonymous outlier SNPs. *URB1* on BTA1 was reported to show a minor association with direct herd life in Holstein [105], is located within a QTL for the polled phenotype in Nellore beef cattle [106], and was found to be under selection in Swedish cattle breeds [107]. *NEIL2* on BTA8 is expressed in bovine ovaries during gestation [108] and is critical for DNA repair processes and immune responses in mammals [109, 110]. *IBTK* on BTA9 has been suggested to regulate B-cell function and cell survival [111, 112]. Of those three genes, only *NEIL2* showed a relatively high signal in the H12 and CLR tests (Fig. 3) and exhibited clustering for UTR3 and UTR5 variants with differentiation between BSW and AAN or SIM, while clustering for nonsynonymous variants was just below the threshold. By contrast, for *URB1* and *IBTK*, neighbor-joining trees revealed high diversity within BSW, with the genotypes of beef bulls being only incompletely separated, suggesting that differentiation at these two genes is not due to strong or recent selection. Nevertheless, all three genes, especially *NEIL2*, possibly have BSW-specific properties associated with local adaptation and immune function that might likewise be beneficial for crossbred calves.

Implications for genome complementarity in beef-on-dairy crossbreds

The number of breed-specific outlier SNPs was more than twice as high for AAN as for LIM and SIM. This reflects the phylogenetic positioning of AAN furthest away from BSW and the closer positioning of LIM and SIM at approximately similar distances from BSW [25, 26]. However, this increased genetic distance did not result in a notable additional reduction of genomic inbreeding in BSW × AAN compared with the other crosses, indicating that all the investigated beef breeds are similarly suitable to reduce the potentially negative effects of recent inbreeding across the BSW genome. Nevertheless, the number of potentially functionally important breed-specific outlier SNPs (i.e., nonsynonymous or *cis*-regulatory SNPs) was highest for AAN, intermediate for LIM, and lowest for SIM. The higher number of breed-specific outlier SNPs for LIM than for SIM disappeared when BTA6 was excluded, indicating a strong role of this chromosome in differentiation among beef breeds. The reason for the associated similarity between BSW and SIM on BTA6, including almost perfect haplotype sharing along the *FAM184B–LCORL* complex, might go back to the descent of BSW from Original Braunvieh, which, alike SIM, is used as a dual-purpose breed and origins in Switzerland.

From the candidate genes detected as potentially under selection in BSW, very few nonsynonymous SNPs were nearly fixed in BSW and simultaneously divergent from the beef breeds, indicating that most selective sweeps on amino acid changing variation were likely not BSW-specific or not strong enough for fixation of the beneficial variant. Furthermore, differentiated genes showed very little overlap among the beef breeds; no gene was found to be divergent between the BSW and all three beef breeds, and genetic variation still segregates within beef breeds. Thus, for each of the candidate genes under selection, the desired genotype of the calf can be arranged by choosing the most suitable individual beef bull for mating a specific BSW female.

Given the large phenotypic differences between the investigated breeds and the very limited number of divergent nonsynonymous variants that we detected, most phenotypic divergence is most likely attributable to a high number of coding or regulatory mutations that are only incompletely differentiated among the breeds. As this variation still segregates within breeds, genomic selection of mating pairs based on crossbred performance should enable the improvement of production and fitness traits in beef-on-dairy crosses.

A caveat to the results is that only a limited number of sires was available for crossing and sequencing at the time of the experiment. Inclusion of additional, unrelated bulls would have likely resulted in more genetic variation within breeds and thus fewer outlier SNPs, which would further increase the options for genomic selection of mating pairs. Inclusion of additional sires in the assessment of genomic inbreeding reduction in crossbreds should have led to more power to detect differences among sire breeds. However, these differences are expected to be small given that we already found very similar patterns of genomic inbreeding reduction despite the varying phylogenetic distances between sires and BSW. Another caveat is that our results relate to animals bred in Switzerland, which might differ in allele frequencies and selective pressures from cattle of other geographic areas. This needs to be taken into account before applying our results to animals that differ genetically from the ones investigated here.

Conclusions

Using AAN, LIM, or SIM beef bulls for mating to BSW dams is expected to improve the economic gain and fitness of calves owing to increased carcass yield and diminished genomic inbreeding. We assessed genome complementarity and its effects on economically relevant phenotypes in beef-on-dairy crosses at high resolution with cost-efficient low-pass sequencing data. Our results show that divergent haplotypes within the *FAM184B–LCORL* complex, which mainly segregate within AAN and LIM, significantly increased hot carcass weight and

protein mass in the F1 crosses, while all three beef breeds were similarly suitable to reduce recent genomic inbreeding in the BSW breed. Given the available genetic variation between and within the investigated beef breeds, selecting or specifically breeding beef bulls most suitable to complement the available haplotypes of BSW females should be possible. However, the trade-offs between the carcass merit and direct calving ease, and the detected BSW-specific alleles potentially involved in local adaptation and immune function must be considered. Our study shows that the evaluation of genome complementarity from sequencing data can facilitate the selection of suitable breeds used for crossing prior to performing costly experiments.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-11029-z.

Additional file 1: Supplementary Tables.

Additional file 2: Supplementary Figures.

Acknowledgements
We thank Raphael Siegenthaler (Agroscope, Posieux, Switzerland) for selecting the calves and compiling associated information; the team from the experimental barn of Agroscope for conducting the animal experimentations; Marmy Viande en Gros SA (Estavayer le Lac, Switzerland) and the Animal Biology research group (Agroscope) for the slaughtering activities and collection of the 11th ribs at the commercial slaughterhouse; Bertrand Egger, Fabrice Sansonnens, Bernard Dougoud, and Guy Maikoff (Agroscope) for performing the slaughter and carcass mincing procedures at the experimental slaughterhouse of Agroscope; Caroline Xavier (Agroscope) for the DXA scans of ribs; and Cord Drögemüller (University of Bern) for providing DNA samples from the calves. We acknowledge Braunvieh Schweiz and IFN Schönow e.V. for sending DNA samples from the dams, Braunvieh Schweiz for sending hair samples from the dams and providing parentage information, birth records, and database access; Swissgenetics for sending semen samples of sires and providing statistics of the number of bulls used for insemination; and Qualitas AG for facilitating the sample exchange. We thank our partners in the fattening trials (Swissgenetics, Swiss Herdbook, Vache mère Suisse, Proviande, Braunvieh Schweiz) for their support. We thank the associate editor and two anonymous reviewers for their valuable comments on the manuscript. Open access funding was provided by Agroscope.

Authors' contributions
IM and SL designed the fattening experiment and collected and preanalyzed all phenotypic data. DL performed the genomics data and association analyses. DL wrote the manuscript, with contributions from IM and SL. MN commented on and revised the manuscript. All authors read and approved the final manuscript.

Funding
Not applicable.

Data availability
The low-pass sequencing data of the calves are accessible at the NCBI Sequence Read Archive under BioProject PRJNA930154. Sequencing data of cows and bulls are not openly available due to reasons of confidentiality and are available from the authors upon reasonable request and with permissions from Braunvieh Schweiz and Swissgenetics. Phenotypes are available at https://doi.org/10.5281/zenodo.12805438, and parts of the phenotype data on body and carcass composition were previously published at https://doi.org/10.57745/EK4FFP.

Declarations

Ethics approval and consent to participate
All the procedures performed on the animals were approved by the ethics committee of the canton Fribourg in Switzerland (2016_48_FR; 2020_03_FR; 2020–45-FR).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Animal GenoPhenomics, 1725 Posieux Agroscope, Switzerland. [2]Ruminant Nutrition and Emissions, 1725 Posieux Agroscope, Switzerland.

Received: 1 August 2024 Accepted: 11 November 2024
Published online: 20 November 2024

References

1. Buckley F, Lopez-Villalobos N, Heins BJ. Crossbreeding: implications for dairy cow fertility and survival. Animal. 2014;8(Suppl 1):122–33.
2. Cabrera VE. Economics of using beef semen on dairy herds. JDS Communications. 2022;3(2):147–51.
3. Holden SA, Butler ST. Review: Applications and benefits of sexed semen in dairy and beef herds. Animal. 2018;12:s97–103.
4. Eriksson S, Ask-Gullstrand P, Fikse WF, Jonsson E, Eriksson J-Å, Stålhammar H, et al. Different beef breed sires used for crossbreeding with Swedish dairy cows - effects on calving performance and carcass traits. Livest Sci. 2020;232: 103902.
5. Berry DP, Twomey A, Ring S. Mean breed performance of the progeny from beef-on-dairy matings. J Dairy Sci. 2023;106(12):9044–54.
6. Berry DP, Amer PR, Evans RD, Byrne T, Cromie AR, Hely F. A breeding index to rank beef bulls for use on dairy females to maximize profit. J Dairy Sci. 2019;102(11):10056–72.
7. Kause A, Mikkola L, Strandén I, Sirkko K. Genetic parameters for carcass weight, conformation and fat in five beef cattle breeds. Animal. 2015;9(1):35–42.
8. Wetlesen MS, Åby BA, Vangen O, Aass L. Estimation of breed and heterosis effects for cow productivity, carcass traits and income in beef × beef and dairy × beef crosses in commercial suckler cow production. Acta Agricult Scand, Section A — Animal Sci. 2020;69(3):137–51.
9. Williams JL, Aguilar I, Rekaya R, Bertrand JK. Estimation of breed and heterosis effects for growth and carcass traits in cattle using published crossbreeding studies. J Anim Sci. 2010;88(2):460–6.
10. Moscarelli A, Sardina MT, Cassandro M, Ciani E, Pilla F, Senczuk G, et al. Genome-wide assessment of diversity and differentiation between original and modern Brown cattle populations. Anim Genet. 2021;52(1):21–31.
11. Sørensen MK, Norberg E, Pedersen J, Christensen LG. Invited Review: Crossbreeding in Dairy Cattle: A Danish Perspective. J Dairy Sci. 2008;91(11):4116–28.
12. Charlesworth D, Willis JH. The genetics of inbreeding depression. Nat Rev Genet. 2009;10(11):783–96.
13. McQuillan R, Eklund N, Pirastu N, Kuningas M, McEvoy BP, Esko T, et al. Evidence of inbreeding depression on human height. Plos Genet. 2012;8(7):e1002655.
14. Pacheco HA, Rossoni A, Cecchinato A, Peñagaricano F. Identification of runs of homozygosity associated with male fertility in Italian Brown Swiss cattle. Front Genet. 2023;14:1227310.
15. Pryce JE, Haile-Mariam M, Goddard ME, Hayes BJ. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genet Sel Evol. 2014;46(1):71.
16. Szpiech ZA, Xu J, Pemberton TJ, Peng W, Zöllner S, Rosenberg NA, Li JZ. Long runs of homozygosity are enriched for deleterious variation. Am J Hum Genet. 2013;93(1):90–102.

17. Kenny D, Carthy TR, Murphy CP, Sleator RD, Evans RD, Berry DP. The Association Between Genomic Heterozygosity and Carcass Merit in Cattle. Front Genet. 2022;13: 789270.
18. Berry DP. Invited review: Beef-on-dairy—The generation of crossbred beef × dairy cattle. J Dairy Sci. 2021;104(4):3789–819.
19. Mee JF. Prevalence and risk factors for dystocia in dairy cattle: a review. Vet J. 2008;176(1):93–101.
20. Mylostyvyi R, Lesnovskay O, Karlova L, Khmeleva O, Kalinichenko O, Orishchuk O, et al. Brown Swiss cows are more heat resistant than Holstein cows under hot summer conditions of the continental climate of Ukraine. Journal of Animal Behaviour and Biometeorology. 2021;9(4):2134.
21. Hoffmann I. Adaptation to climate change - exploring the potential of locally adapted breeds. Animal. 2013;7(Suppl 2):346–62.
22. Hagger C. Estimates of genetic diversity in the brown cattle population of Switzerland obtained from pedigree information. J Anim Breed Genet. 2005;122(6):405–13.
23. Yoder DM, Lush JL. A genetic history of the Brown Swiss cattle in the United States. J Hered. 1937;28(4):154–60.
24. Signer-Hasler H, Burren A, Neuditschko M, Frischknecht M, Garrick D, Stricker C, et al. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. Genet Sel Evol. 2017;49:83.
25. Decker JE, McKay SD, Rolf MM, Kim J, Molina Alcalá A, Sonstegard TS, et al. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. Plos Genet. 2014;10(3): e1004254.
26. Decker JE, Pires JC, Conant GC, McKay SD, Heaton MP, Chen K, et al. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. Proc Natl Acad Sci U S A. 2009;106(44):18644–9.
27. Morel I, Oberson J-L, Schlegel P, Chassot A, Lehmann E, Kessler J. Apports alimentaires recommandés pour le bovin à l'engrais. 2017. In: Apports alimentaires recommandés pour les ruminants. Posieux: Agroscope; [1–27]. Available from: https://www.agroscope.admin.ch/agroscope/fr/home/services/soutien/aliments-pour-animaux/apports-alimentaires-recommandes-ruminants.html.
28. Lerch S, Morel I, Dohme-Meier F, Le Cozler Y, Xavier C. Estimation of body and carcass composition of crossbred growing bulls from 11th rib dissection. Animal - Open Space. 2023;2: 100030.
29. Xavier C, Morel I, Dohme-Meier F, Siegenthaler R, Le Cozler Y, Lerch S. Estimation of carcass chemical composition in beef-on-dairy cattle using dual-energy X-ray absorptiometry (DXA) scans of cold half-carcass or 11th rib cut. J Animal Sci. 2023;101:skad380.
30. Lindtke D, Seefried FR, Drögemüller C, Neuditschko M. Increased heterozygosity in low-pass sequencing data allows identification of blood chimeras in cattle. Anim Genet. 2023;54(5):613–8.
31. Snelling WM, Hoff JL, Li JH, Kuehn LA, Keel BN, Lindholm-Perry AK, Pickrell JK. Assessment of Imputation from Low-Pass Sequencing to Predict Merit of Beef Steers. Genes (Basel). 2020;11(11);1312.
32. Wasik K, Berisa T, Pickrell JK, Li JH, Fraser DJ, King K, Cox C. Comparing low-pass sequencing and genotyping for trait mapping in pharmacogenetics. BMC Genomics. 2021;22(1):197.
33. loimpute-public. Available from: https://gitlab.com/gencove/loimpute-public. Accessed 22 Dec 2021.
34. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. GigaScience. 2020;9(3):giaa021.
35. Elsik CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM, Weinstock GM, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science. 2009;324(5926):522–8.
36. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–303.
37. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet. 2009;84(2):210–23.
38. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81(5):1084–97.
39. Druet T, Gautier M. A model-based approach to characterize individual inbreeding at both global and local genomic scales. Mol Ecol. 2017;26(20):5820–41.
40. Broman KW, Weber JL. Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. Am J Hum Genet. 1999;65(6):1493–500.
41. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, et al. Runs of homozygosity in European populations. Am J Hum Genet. 2008;83(3):359–72.
42. Bertrand AR, Kadri NK, Flori L, Gautier M, Druet T. RZooRoH: An R package to characterize individual genomic autozygosity and identify homozygous-by-descent segments. Methods Ecol Evol. 2019;10(6):860–6.
43. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. Plos Biol. 2006;4(3): e72.
44. Garud NR, Messer PW, Buzbas EO, Petrov DA. Recent selective sweeps in North American Drosophila melanogaster show signatures of soft sweeps. Plos Genet. 2015;11(2):e1005004
45. Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. Genomic scans for selective sweeps using SNP data. Genome Res. 2005;15(11):1566–75.
46. Gautier M, Klassmann A, Vitalis R. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. Mol Ecol Resour. 2017;17(1):78–90.
47. Gautier M, Vitalis R. rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics. 2012;28(8):1176–7.
48. Klassmann A, Gautier M. Detecting selection using extended haplotype homozygosity (EHH)-based statistics in unphased or unpolarized data. PLoS ONE. 2022;17(1): e0262024.
49. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156–8.
50. DeGiorgio M, Huber CD, Hubisz MJ, Hellmann I, Nielsen R. SweepFinder2: increased sensitivity, robustness and flexibility. Bioinformatics. 2016;32(12):1895–7.
51. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics. 2005;21(16):3439–40.
52. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc. 2009;4(8):1184–91.
53. Harrison PW, Amode MR, Austine-Orimoloye O, Azov Andrey G, Barba M, Barnes I, et al. Ensembl 2024. Nucleic Acids Res. 2024;52(D1):D891–9.
54. Hu Z-L, Park CA, Reecy JM. A combinatorial approach implementing new database structures to facilitate practical data curation management of QTL, association, correlation and heritability data on trait variants. Database. 2023;2023:baad024.
55. The Cattle Quantitative Trait Locus (QTL) Database. https://www.animalgenome.org/cgi-bin/QTLdb/BT/index. Accessed 1 Nov 2023.
56. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–7.
57. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc: Ser B (Methodol). 1995;57(1):289–300.
58. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164-e.
59. Liu S, Gao Y, Canela-Xandri O, Wang S, Yu Y, Cai W, et al. A multi-tissue atlas of regulatory variants in cattle. Nat Genet. 2022;54(9):1438–47.
60. Sharpe D. Chi-Square Test is Statistically Significant: Now What? Practical Assessment. Research, and Evaluation. 2015;20(1):8.
61. Vieira FG, Lassalle F, Korneliussen TS, Fumagalli M. Improving the estimation of genetic distances from Next-Generation Sequencing data. Biol J Lin Soc. 2016;117(1):139–49.
62. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. GigaScience. 2021;10(2):giab008.

63. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
64. PLINK 1.90 beta. https://www.cog-genomics.org/plink/1.9/. Accessed 26 July 2022.
65. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019;35(3):526–8.
66. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. cluster: Cluster Analysis Basics and Extensions. R package; 2023.
67. Kennedy BW, Quinton M, van Arendonk JA. Estimation of effects of single genes on quantitative traits. J Anim Sci. 1992;70(7):2000–12.
68. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. Nat Genet. 2014;46(2):100–6.
69. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015;4(1):7.
70. PLINK 2.00 alpha. https://www.cog-genomics.org/plink/2.0. Accessed 26 July 2022.
71. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42(7):565–9.
72. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88(1):76–82.
73. Gogarten SM, Sofer T, Chen H, Yu C, Brody JA, Thornton TA, et al. Genetic association testing using the GENESIS R/Bioconductor package. Bioinformatics. 2019;35(24):5346–8.
74. Solé M, Gori A-S, Faux P, Bertrand A, Farnir F, Gautier M, Druet T. Age-based partitioning of individual genomic inbreeding levels in Belgian Blue cattle. Genet Sel Evol. 2017;49(1):92.
75. Guinan FL, Wiggans GR, Norman HD, Dürr JW, Cole JB, Van Tassell CP, et al. Changes in genetic trends in US dairy cattle since the implementation of genomic selection. J Dairy Sci. 2023;106(2):1110–29.
76. Weigand H, Leese F. Detecting signatures of positive selection in non-model species using genomic data. Zool J Linn Soc. 2018;184(2):528–83.
77. Akanno EC, Chen L, Abo-Ismail MK, Crowley JJ, Wang Z, Li C, et al. Genome-wide association scan for heterotic quantitative trait loci in multi-breed and crossbred beef cattle. Genet Sel Evol. 2018;50(1):48.
78. Bouwman AC, Daetwyler HD, Chamberlain AJ, Ponce CH, Sargolzaei M, Schenkel FS, et al. Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. Nat Genet. 2018;50(3):362–7.
79. Sahana G, Höglund JK, Guldbrandtsen B, Lund MS. Loci associated with adult stature also affect calf birth survival in cattle. BMC Genet. 2015;16:47.
80. Wang Y, Zhang F, Mukiibi R, Chen L, Vinsky M, Plastow G, et al. Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. BMC Genomics. 2020;21(1):38.
81. Zhang F, Wang Y, Mukiibi R, Chen L, Vinsky M, Plastow G, et al. Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: I: feed efficiency and component traits. BMC Genomics. 2020;21(1):36.
82. Bongiorni S, Mancini G, Chillemi G, Pariset L, Valentini A. Identification of a short region on chromosome 6 affecting direct calving ease in Piedmontese cattle breed. PLoS ONE. 2012;7(12): e50137.
83. Saatchi M, Schnabel RD, Taylor JF, Garrick DJ. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. BMC Genomics. 2014;15(1):442.
84. Olsen HG, Knutsen TM, Lewandowska-Sabat AM, Grove H, Nome T, Svendsen M, et al. Fine mapping of a QTL on bovine chromosome 6 using imputed full sequence data suggests a key role for the group-specific component (GC) gene in clinical mastitis and milk production. Genet Sel Evol. 2016;48(1):79.
85. Horikoshi M, Beaumont RN, Day FR, Warrington NM, Kooijman MN, Fernandez-Tajes J, et al. Genome-wide associations for birth weight and correlations with adult disease. Nature. 2016;538(7624):248–52.
86. Weedon MN, Frayling TM. Reaching new heights: insights into the genetics of human stature. Trends Genet. 2008;24(12):595–603.
87. Soranzo N, Rivadeneira F, Chinappen-Horsley U, Malkina I, Richards JB, Hammond N, et al. Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. Plos Genet. 2009;5(4): e1000445.
88. Sovio U, Bennett AJ, Millwood IY, Molitor J, O'Reilly PF, Timpson NJ, et al. Genetic determinants of height growth assessed longitudinally from infancy to adulthood in the northern Finland birth cohort 1966. Plos Genet. 2009;5(3): e1000409.
89. Lin E, Tsai SJ, Kuo PH, Liu YL, Yang AC, Conomos MP, Thornton TA. Genome-wide association study in the Taiwan Biobank identifies four novel genes for human height: NABP2, RASA2, RNF41 and SLC39A5. Hum Mol Genet. 2021;30(23):2362–9.
90. Lindholm-Perry AK, Sexten AK, Kuehn LA, Smith TP, King DA, Shackelford SD, et al. Association, effects and validation of polymorphisms within the NCAPG - LCORL locus located on BTA6 with feed intake, gain, meat and carcass traits in beef cattle. BMC Genet. 2011;12:103.
91. Liu Y, Long H, Feng S, Ma T, Wang M, Niu L, et al. Trait correlated expression combined with eQTL and ASE analyses identified novel candidate genes affecting intramuscular fat. BMC Genomics. 2021;22(1):805.
92. Niu Q, Zhang T, Xu L, Wang T, Wang Z, Zhu B, et al. Identification of Candidate Variants Associated With Bone Weight Using Whole Genome Sequence in Beef Cattle. Front Genet. 2021;12: 750746.
93. Eberlein A, Takasuga A, Setoguchi K, Pfuhl R, Flisikowski K, Fries R, et al. Dissection of genetic factors modulating fetal growth in cattle indicates a substantial role of the non-SMC condensin I complex, subunit G (NCAPG) gene. Genetics. 2009;183(3):951–64.
94. Niu Q, Zhang T, Xu L, Wang T, Wang Z, Zhu B, et al. Integration of selection signatures and multi-trait GWAS reveals polygenic genetic architecture of carcass traits in beef cattle. Genomics. 2021;113(5):3325–36.
95. Snelling WM, Allan MF, Keele JW, Kuehn LA, McDaneld T, Smith TP, et al. Genome-wide association study of growth in crossbred beef cattle. J Anim Sci. 2010;88(3):837–48.
96. Bhati M, Kadri NK, Crysnanto D, Pausch H. Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. BMC Genomics. 2020;21(1):27.
97. Signer-Hasler H, Casanova L, Barenco A, Maitre B, Bagnato A, Vevey M, et al. Genomic regions underlying positive selection in local. Alpine cattle breeds Anim Genet. 2023;54(3):239–53.
98. Cohen-Zinder M, Seroussi E, Larkin DM, Loor JJ, Everts-van der Wind A, Lee JH, et al. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res. 2005;15(7):936–44.
99. Cecchinato A, Chessa S, Ribeca C, Cipolat-Gotet C, Bobbo T, Casellas J, Bittante G. Genetic variation and effects of candidate-gene polymorphisms on coagulation properties, curd firmness modeling and acidity in milk from Brown Swiss cows. Animal. 2015;9(7):1104–12.
100. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000;25(1):25–9.
101. The Gene Ontology C, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, et al. The Gene Ontology knowledgebase in 2023. Genetics. 2023;224(1):iyad031.
102. Crusio WE, Dhawan E, Chesler EJ, Delprato A. Analysis of morphine responses in mice reveals a QTL on Chromosome 7. F1000Res. 2016;5:2156.
103. Jeng SF, Rau CS, Liliang PC, Wu CJ, Lu TH, Chen YC, et al. Profiling muscle-specific microRNA expression after peripheral denervation and reinnervation in a rat model. J Neurotrauma. 2009;26(12):2345–53.
104. Pedrosa VB, Schenkel FS, Chen SY, Oliveira HR, Casey TM, Melka MG, Brito LF. Genomewide Association Analyses of Lactation Persistency and Milk Production Traits in Holstein Cattle Based on Imputed Whole-Genome Sequence Data. Genes (Basel). 2021;12(11):1830.
105. Nayeri S, Sargolzaei M, Abo-Ismail MK, Miller S, Schenkel F, Moore SS, Stothard P. Genome-wide association study for lactation persistency, female fertility, longevity, and lifetime profit index traits in Holstein dairy cattle. J Dairy Sci. 2017;100(2):1246–58.
106. Stafuzza NB, Silva RMO, Peripolli E, Bezerra LAF, Lôbo RB, Magnabosco CU, et al. Genome-wide association study provides insights into genes related with horn development in Nelore beef cattle. PLoS ONE. 2018;13(8): e0202978.

107. Ghoreishifar SM, Eriksson S, Johansson AM, Khansefid M, Moghaddaszadeh-Ahrabi S, Parna N, et al. Signatures of selection reveal candidate genes involved in economic traits and cold acclimation in five Swedish cattle breeds. Genet Sel Evol. 2020;52(1):52.
108. Liu M, Hummitzsch K, Hartanti MD, Rosario R, Bastian NA, Hatzirodos N, et al. Analysis of expression of candidate genes for polycystic ovary syndrome in adult and fetal human and fetal bovine ovaries†. Biol Reprod. 2020;103(4):840–53.
109. Sarker AH, Cooper PK, Hazra TK. DNA glycosylase NEIL2 functions in multiple cellular processes. Prog Biophys Mol Biol. 2021;164:72–80.
110. Pan L, Xue Y, Wang K, Zheng X, Islam A, Tapryal N, et al. Nei-like DNA glycosylase 2 selectively antagonizes interferon-β expression upon respiratory syncytial virus infection. J Biol Chem. 2023;299(8): 105028.
111. Liu W, Quinto I, Chen X, Palmieri C, Rabin RL, Schwartz OM, et al. Direct inhibition of Bruton's tyrosine kinase by IBtk, a Btk-binding protein. Nat Immunol. 2001;2(10):939–46.
112. Vecchio E, Golino G, Pisano A, Albano F, Falcone C, Ceglia S, et al. IBTK contributes to B-cell lymphomagenesis in Eμ-myc transgenic mice conferring resistance to apoptosis. Cell Death Dis. 2019;10(4):320.

Publisher's Note