Animal 19 (2025) 101540

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

Genomic insights into growth traits in German Black Pied cattle: a dual-purpose breed at risk

P. Korkuć [a,b,*], M. Reißmann [a], G.A. Brockmann [a]

[a] Humboldt University of Berlin, Albrecht Daniel Thaer- Institute for Agricultural and Horticultural Sciences, Animal Breeding and Molecular Genetics, Invalidenstr. 42, 10115 Berlin, Germany
[b] Leibniz Institute for Zoo and Wildlife Research (IZW), Alfred-Kowalke-Str. 17, 10315 Berlin, Germany

ARTICLE INFO

Article history:
Received 14 November 2024
Revised 28 April 2025
Accepted 29 April 2025
Available online 7 May 2025

Keywords:
Breed conservation
BW
Genome-wide association studies
Quantitative trait loci
Whole-genome sequencing

ABSTRACT

The German Black Pied cattle (**DSN**) is an endangered dual-purpose breed valued for its genetic diversity and high milk fat and protein content. However, due to competition with higher-yielding dairy breeds, the DSN population has declined, leading to its designation as an endangered breed. While previous research has focused on the milk production traits of DSN, this study aims to address meat traits to further understand the genetic determination of the dual-purpose characteristics of the breed. We conducted genome-wide association studies on 669 DSN bulls to identify genetic loci associated with birth weight, BW, and BW gain at different growth stages. Using imputed whole-genome sequencing data, we identified 14 quantitative trait loci across ten chromosomes. Significant associations were found for birth weight on chromosomes 5 and 18, for body weight at 3 weeks ($\mathbf{BW_{3w}}$) on chromosomes 3 and 16, for body weight at 7 months ($\mathbf{BW_{7m}}$) on chromosomes 3 and 10, and for body weight gain from birth or 3 weeks to 18 months ($\mathbf{BWG_{0d\text{-}18m}}$, $\mathbf{BWG_{3w\text{-}18m}}$) on chromosomes 4 and 7. Key positional candidate genes influencing muscle and fat tissue development included *RERGL* and *LMO3* (identified for birth weight), *MET* and *CAPZA2* (identified for $BWG_{0d\text{-}18m}$) which are essential for skeletal muscle development and actin filament regulation, respectively, *TLN2* (identified for BW_{7m}), *MYO1F* and *ADAMTS10* (identified for $BWG_{3w\text{-}18m}$) which are critical for actin filament assembly, cytoskeletal function, and skeletal development, respectively. Candidate genes such as *CPT2* (identified or BW_{3w}) and *VPS13C* (identified or BW_{7m}) are involved in lipid metabolism and mitochondrial function. Additionally, candidate genes such as *IGSF3* (identified for BW_{7m}), *KLRC1* and members of the C-type lectin family (identified for birth weight) are associated with immune regulation, and thus, suggest a potential interplay between metabolism, immune function, and growth efficiency. These findings highlight the distinct genetic mechanisms underlying growth at various developmental stages, underscoring the importance of breed-specific genetic evaluations. The identified loci also overlap with previously reported loci for meat and production traits in other cattle breeds, underscoring their relevance and potential utility in DSN breeding strategies. This study provides a foundation for conservation and genomic breeding strategies to maintain the dual-purpose characteristics of DSN through optimising both meat and milk production.

Implications

We explored the genetic factors that influence growth traits in the endangered dual-purpose German Black Pied cattle. We discovered that growth traits in German Black Pied cattle are controlled by distinct genetic mechanisms at various stages of development. Identified genes were involved in muscle and fat tissue development, immune response, and metabolism. Importantly, the genetic regions affecting growth are different from those influencing milk fat content. This knowledge allows targeted breeding strategies to enhance body weight and weight gain without compromising milk quality preserving the dual-purpose characteristics of the breed and aiding in the conservation and economic viability of the German Black Pied cattle.

Introduction

The German Black Pied (**DSN**) cattle, known in German as "Deutsches Schwarzbuntes Niederungsrind," is a traditional dual-purpose breed that has been historically valued for both milk and meat production. However, the DSN population has declined

* Corresponding author.
E-mail addresses: paula.korkuc@hu-berlin.de, korkuc@izw-berlin.de (P. Korkuć).

https://doi.org/10.1016/j.animal.2025.101540
1751-7311/

markedly as it was replaced by the more specialised and higher-yielding Holstein breed. Currently, there are only around 2 500 DSN cows remaining in herdbooks, a number that underscores the breed's critical status as endangered (Bundesanstalt für Landwirtschaft und Ernährung, 2023). The breed is mainly located in the federal state of Brandenburg, Germany. Given its considerable high genetic diversity (nucleotide diversity: 0.151% in DSN vs 0.147% in Holstein) and unique characteristics, such as higher milk fat (4.3%) and protein content (3.7%) compared to Holstein, preserving the DSN and exploring its genetic resources remain highly relevant (Mandel et al., 2025, Neumann et al., 2023).

Previous genetic studies on DSN have largely focused on milk production, fertility, and disease resistance traits (Korkuć et al., 2021, 2023; Wolf et al., 2021; Meier et al., 2020; May et al., 2019). Key findings include a quantitative trait locus (**QTL**) at the *MGST1* (microsomal glutathione S-transferase 1) gene associated with milk fat percentage in DSN cows (Korkuć et al., 2023), a locus that had been identified in multiple genome-wide association studies (**GWAS**) across different cattle breeds (Tribout et al., 2020; Weller et al., 2018; Pausch et al., 2017). Simultaneously, the K232A polymorphism in *DGAT1* (diacylglycerol O-acyltransferase 1) was found to be nearly fixed (frequency of 0.97) in DSN for the alanine protein variant, which is linked to higher milk and protein yield in other breeds (Grisart et al., 2002).

Recognising DSN as a dual-purpose breed underlines the importance of considering both milk yield and composition, as well as the genetic potential of the breed for growth and development, which are traits that should be taken into account for selection objectives related to meat production.

In this study, we pivot from examining milk yield and composition towards investigating growth traits in fattening bulls within the DSN breed, focusing on weight measurements at critical growth stages: at birth and at the average age of 3 weeks, 7 months, and 18 months. This study employs imputed whole-genome sequencing (**WGS**) data providing a detailed and breed-specific insight into the genetic factors affecting growth, analysing a cohort of 669 DSN bulls. Our goal was to identify not only unique genetic loci associated with growth traits but also to explore the interplay between growth or meat production and milk production to guide a conservation and breeding strategy that preserves the dual-purpose characteristics of DSN.

Material and methods

Animals and traits

BW measurements were collected for 669 DSN bulls, born between 2015 and 2023 at a single farm in Brandenburg, using the HERDEplus® herd management software (dsp-Agrosoft GmbH, Ketzin/Havel, Germany, 2023). This farm is the largest DSN farm in Germany, managing around 800 cows (~30% of all registered DSN herdbook cows) and 350 bulls. The genetic background of the DSN animals is representative for the whole DSN population since animals are bought particularly from this farm. It is also the only farm that fattens DSN bulls for slaughter, making it the primary source of phenotypic data. This farm systematically records weight measurements, unlike smaller DSN farms. Due to the standardised conditions, as bulls were housed and fed under uniform management, the environmental variability is reduced improving statistical robustness.

For each DSN bull, four different BW measurements in kilogram (kg) were recorded: Birth weight (BW_{0d}, 0 days ± 0 SD), BW at around 3 weeks (BW_{3w}, 23.2 days ± 5.9 SD), at around 7 months (BW_{7m}, 195.7 days ± 63.6 SD), and at around 18 months (BW_{18m}, 549.2 days ± 22.3 SD). BW gains were calculated as the difference between BW measurements resulting in the following traits: BW gain from birth to 3 weeks ($BWG_{0d\text{-}3w}$), birth to 7 months ($BWG_{0\text{-}d\text{-}7m}$), birth to 18 months ($BWG_{0d\text{-}18m}$), from 3 weeks to 7 months ($BWG_{3w\text{-}7m}$), 3 weeks to 18 months ($BWG_{3w\text{-}18m}$), and from 7 to 18 months ($BWG_{7m\text{-}18m}$). We applied quality control by excluding phenotypic values exceeding ± 3 SD from the mean for each trait. Table 1 presents the number of animals before and after filtering, along with summary statistics for the investigated traits in DSN bulls.

Genotypes

Among the 669 DSN bulls, 396 were genotyped using the custom DSN200K SNP chips (Neumann et al., 2021), and 273 with the Eurogenomics EuroG MD chip. Genotypes were imputed to WGS data density using Beagle v5.1 (Browning et al., 2018). The reference panel for imputation was representative of the whole DSN population. It consisted of 304 sequenced DSN cattle including selected cows from different farms and 47 DSN breeding bulls (Neumann et al., 2021). Regarding the size of the reference panel for imputation, in a previous study, we tested how much additional variation was found in the DSN population by subsequently adding the WGS data of additional animals (Korkuć et al., 2023). This followed a saturation curve. After reaching a certain number of animals, adding more animals to the reference panel did not significantly increase the number of newly identified variations. Based on this, we believe that having 304 DSN individuals in the reference panel for imputation was sufficient to capture the majority of variants in the investigated DSN population. Genotypes from Eurogenomics EuroG MD chip were initially imputed to match the DSN200K chip positions using a reference panel of 1 580 DSN cattle genotyped with the DSN200K chip (Neumann et al., 2021), and subsequently imputed to WGS level. Variant filtering for a minor allele frequency (**MAF**) of ≥ 5% and a variant call rate of ≥ 90% was conducted with vcftools v0.1.14 (Danecek et al., 2011) yielding 11 820 219 variants suitable for GWAS with genomic positions corresponding to the *Bos taurus* genome version ARS-UCD1.2 (Rosen et al., 2018). To ensure robustness of statistical test, genotype groups, defined as the different genotypes observed at a given variant (e.g., AA, AG, GG), were only included in the GWAS if each group was represented by at least 20 animals. Genotype groups with less than 20 individuals were excluded from the analysis to avoid biased estimates due to a small sample size.

Genome-wide association studies

GWAS was conducted using multiple linear regression models in R version 4.2.2. The model evaluated the additive effect of each variant on trait Y, incorporating fixed effects for population stratification (ps), birth year (by), birth season (bs), age at weighing in days for BW measurements or age difference in days for weight gain (ag), and the variant genotype (gt) and the residual error (e):

$$Y = ps + by + bs + a + gt + e \quad (1)$$

Population stratification was determined through the pairwise population concordance test using PLINK v1.9′s --cluster and --ppc parameters, based on the identity-by-state matrix across all 669 DSN bulls (Chang et al., 2015). A *P*-value threshold of 0.00001 identified 18 relatedness clusters. Population stratification analysis was necessary due to the small population size and the limited number of breeding bulls used per generation. As a result, complex family relationships exist within the population, which can introduce hidden substructures that may impact GWAS results. The fixed effect of age was not included for birth weight. For each investigated trait, the best-fit model was selected by including only fixed effects that

Table 1
Summary statistics for BW and BW gain traits in German Black Pied (DSN) bulls. The table includes the number of individuals before and after filtering (removing values ±3 SD from the mean of each investigated trait), heritabilities (h^2) with SE, and the summary statistics of fixed effects corrected traits including the mean, SD, CV, and skewness. BW_{0d}, BW_{3w}, BW_{7m}, BW_{18m} refer to BW at birth, 3 weeks, 7 months, and 18 months; $BWG_{0d-3w/7m/18m}$, $BWG_{3w-7m/18m}$, and BWG_{7m-18m} indicate BW gain over respective periods.

Trait	# animals (before filtering)	# animals (after filtering)	h^2	SE(h^2)	Mean (kg)	SD (kg)	CV	Skewness
BW_{0d}	667	661	0.67	0.08	42.41	4.90	0.12	0.11
BW_{3w}	624	621	0.50	0.09	66.70	6.72	0.10	−0.13
BW_{7m}	656	650	0.34	0.10	264.55	23.14	0.09	−0.43
BW_{18m}	519	512	0.31	0.10	740.46	49.99	0.07	−0.19
BWG_{0d-3w}	622	615	0.13	0.07	24.30	5.20	0.21	−0.13
BWG_{0d-7m}	654	648	0.22	0.09	222.35	22.30	0.10	−0.53
BWG_{0d-18m}	518	511	0.16	0.08	698.11	49.11	0.07	−0.22
BWG_{3w-7m}	614	608	0.27	0.10	194.01	20.63	0.11	−0.58
BWG_{3w-18m}	478	473	0.33	0.11	673.66	50.08	0.07	−0.24
BWG_{7m-18m}	513	512	0.24	0.10	461.01	50.32	0.11	−0.38

significantly contributed to the model, as determined using the Akaike Information Criterion (**AIC**) with the aictab() function from the R package AICcmodavg v2.3-3 (Mazerolle, 2023). All potential combinations of fixed effects were considered (whereas the population stratification was always included), and the final models are provided in Supplementary Table S1. The heritability (h^2) of traits was calculated using GCTA software v1.94.3 (Yang et al., 2011) by performing a restricted maximum likelihood (**REML**) analysis on the genomic relationship matrix. The genomic relationship matrix was constructed from autosomal genotypes using the --make-grm option. The REML analysis (parameter --reml) included the respective trait (parameter --pheno) and the relevant fixed effects (parameters --covar and –qcovar) as estimated previously (Supplementary Table S1). Similarly, pairwise genetic correlations between traits were estimated using bivariate GREML analysis implemented in GCTA (parameter --reml-bivar). This analysis also included the joint fixed effects of traits as estimated previously (Supplementary Table S1).

Significance threshold and quantitative trait locus definition

Due to the limited number of breeding bulls, DSN animals are highly related, even within a single farm. This high relatedness may contribute to *P*-value inflation in GWAS. To minimise false positive associations to traits, *P*-values from GWAS underwent adjustment to ensure a maximum inflation factor λ of 1.2 whenever it exceeded this threshold (Supplementary Table S2). This adjustment was applied jointly to all autosomes and separately to chromosome X as recombination on the sex chromosomes differs from those on autosomes (Zhang et al., 2020). The Q-Q plots in Supplementary Fig. S1 illustrate the *P*-value distributions before and after correction for λ = 1.2. Prior to correction, the mean inflation factor across all traits was 1.18 (± 0.10 SD) for autosomes and 1.47 (± 0.40 SD) for chromosome X.

Significance thresholds were determined using a Bonferroni correction, which accounts for multiple testing by dividing the genome-wide α-level threshold by the number of independent variants. To estimate the number of independent variants, linkage disequilibrium (**LD**)-based pruning was performed in PLINK v1.9 using a window size of 2 000 kb, a step size of 200, and a r^2 value of 0.6, resulting in 423 453 independent variants. Based on this, the significance thresholds were set as $-\log_{10}(p) \geq 7.6$ for highly significant ($P < 0.01$), $-\log_{10}(p) \geq 6.9$ for significant ($P < 0.05$), and $-\log_{10}(p) \geq 6.6$ for suggestive ($P < 0.1$) associations, respectively. QTL regions were coarsely defined by grouping associated variants ($P < 0.1$) within ± 2.0 Mb on a chromosome. Afterwards, QTL regions were defined by a 1.5 LOD-drop including additional variants with a LD of $r^2 \geq 0.75$ to the top variant (Dupuis and Siegmund, 1999). When a resulting QTL region was shorter than 200 kb, it was defined as a region ± 100 kb from the top variant.

Quantitative trait loci annotation

Detected QTL regions were investigated for candidate genes considering 21 880 protein-coding genes obtained from Ensembl release 110 (Howe et al., 2021). Genes were analysed separately in every QTL region for Gene Ontology (**GO**) term enrichment using g:Profiler with the R package gprofiler2 v0.2.3 (Raudvere et al., 2019). The options for annotated genes and a g:SCS threshold of $P < 0.1$ for suggestive and $P < 0.05$ for significant GO term enrichments were used.

Variants in QTL regions were investigated for impact consequence on gene transcripts using Variant Effect Predictor (**VEP**) from Ensembl (McLaren et al., 2016). For this, also rare variants (MAF ≥ 1%) from the imputed WGS genotypes were considered. Categories included variants with high, moderate, or low impact. Additionally, VEP predicted if missense variants were tolerant or deleterious by using the Sorting Intolerant From Tolerant (**SIFT**) algorithm (Kumar et al., 2009). Variants with impact defined as intergenic, intron, non-coding transcript, or non-coding transcript exon variant were not considered. Significantly associated variants located in intron or promoter regions were tested whether they are located in transcription factor binding sites of vertebrates as obtained from JASPAR CORE database release 9 (Castro-Mondragon et al., 2022) using the R package TFBSTools v1.36.0 (Tan and Lenhard, 2016).

Identified QTL regions were investigated for an overlap with previously published associations and QTLs for "Meat and Carcass" and "Production" from cattleQTLdb release 54 (Hu et al., 2022). Entries with the trait names "Docosahexaenoic acid content", "Meat colour", "PTA type", "DM intake", "Residual feed intake", and "Net merit" were excluded. These traits were excluded because our comparison focused specifically on traits related to BW and BW gain. Additionally, we tested if QTLs for "Milk" were located in the identified region for growth. PubMed IDs of corresponding publications were obtained using R package easyPubMed v2.13 (Fantini, 2019). LD between variants as squared Pearson correlation coefficient r^2 was calculated using the default cor() function in R using pairwise complete observations:

$$r^2 = \left(\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \right)^2 \quad (2)$$

where x_i are the genotypes and $\bar{x}$ the mean of the genotypes of DSN bulls at the first variant, y_i are the genotypes and $\bar{y}$ the mean of the genotypes of DSN bulls at the second variant. Genotypes were coded as 0, 1, and 2 corresponding to homozygous reference, heterozygous, and homozygous alternative alleles.

Additionally, LD as D' was calculated using the LD() function from the R package genetics v.1.3.8.1.3 (Warnes et al., 2022). This includes the calculation of the raw difference in frequency between the observed and expected number of AB haplotype pairs:

$$D = p_{AB} - p_A p_B \quad (3)$$

The measure D' was then computed to scale D to the range between −1 and 1:

$$D' = \frac{D}{D_{max}} \quad (4)$$

where D_{max} is defined as $min(p_A p_b, p_a p_B)$ if $D > 0$ or as $max - p_A p_B - p_a p_b$ if $D < 0$. Here, p_A and p_a represent the observed frequencies of alleles "A" and "a" (with $p_a = 1\text{-}p_A$) at the first variant, respectively, while p_B and p_b represent the observed frequencies of alleles "B" and "b" (with $p_b = 1\text{-}p_B$) at the second variant. The value p_{AB} denotes the observed haplotype frequency of the allele pair "AB".

R version 4.2.2 was used for analyses. Figures were produced using the R package ggplot2 v3.5.1 (Wickham, 2016). Gene arrows and names were added to plots using the R packages gggenes v0.5.1 and ggrepel v0.9.5 (Slowikowski, 2021; Wilkins, 2020). *P*-values between genotype groups in variant effect plots were estimated using pairwise t-tests and displayed using R package ggpubr v0.6.0 (Kassambara, 2020).

Results

Genetic correlations

The genetic correlation analysis revealed high positive correlations between several BW and BW gain traits, particularly at later growth stages (Supplementary Table S3). Only genetic correlations with SE < 0.1 were considered for interpretation. A high and positive genetic correlation was found between birth weight and BW_{3w} (r = 0.92, SE = 0.05), suggesting that genetic factors influencing birth weight strongly impact BW at 3 weeks. In later growth phases, BW_{7m} showed a very high genetic correlation with $BWG_{0\text{-}d\text{-}7m}$ (r = 0.95, SE = 0.02), and $BWG_{3w\text{-}7m}$ (r = 0.95, SE = 0.03). Similarly, $BWG_{0d\text{-}7m}$ was highly correlated with $BWG_{3w\text{-}7m}$ (r = 0.99, SE = 0.01). For the late growth phase, BW_{18m} showed very high genetic correlation with $BWG_{0d\text{-}18m}$ (r = 0.99, SE = 0.01), $BWG_{3w\text{-}18m}$ (r = 0.99, SE = 0.01), and $BWG_{7m\text{-}18m}$ (r = 0.98, SE = 0.03). Additionally, high correlations were observed between $BWG_{0d\text{-}18m}$ and $BWG_{3w\text{-}18m}$ (r = 1.00, SE = 0.00), as well as between $BWG_{7m\text{-}18m}$ and $BWG_{0d\text{-}18m}$ (r = 0.96, SE = 0.04) or $BWG_{3w\text{-}18m}$ (r = 0.98, SE = 0.03).

Genomic regions associated with BW and BW gain

We identified 71 sequence variants above the significant threshold ($-\log_{10}(p) \geq 6.9$) for birth weight, the BW traits BW_{3w} and BW_{7m}, and for the BW gain traits $BWG_{0d\text{–}7m}$, $BWG_{0d\text{–}18m}$, $BWG_{3w\text{–}18m}$, and $BWG_{7m\text{–}18m}$ (Supplementary Table S4). Those variants could be grouped into 14 QTL regions on ten chromosomes (3, 4, 5, 6, 7, 10, 13, 16, 18, and X) (Fig. 1), each represented by its top variant (Table 2). Of those 14 QTL regions, four were highly significantly associated genomic loci ($-\log_{10}(p) \geq 7.6$): two for birth weight on chromosome 5 next to each other, one for BW_{3w} on chromosome 3, and one for $BWG_{0d\text{-}18m}$ on chromosome 4. In the following, we explain in detail all highly significant loci (Fig. 1, red circles) and additionally five significant loci interesting for birth weight on chromosome 18; for BW_{3w} on chromosome 16; for BW_{7m} on chromosomes 3 and 10 (this locus was also significant for $BWG_{0d\text{-}7m}$); and for $BWG_{3w\text{-}18m}$ on chromosome 7 (Fig. 1, blue circles).

Birth weight – significant loci on chromosomes 5 and 18

On chromosome 5, two highly significant and one significant QTL regions associated with birth weight were identified in close proximity to each other (Table 2, Fig. 2a). The first highly significant top variant rs519831213 was located at 92 970 239 bp ($-\log_{10}(p)$ = 8.00, Fig. 2b); the second highly significant top variant was a novel variant identified in DSN which was located 6.3 Mb downstream at 99 255 711 bp (no rsID, $-\log_{10}(p)$ = 8.32, Fig. 2c). The third significant top variant rs135635336 was located another 2.5 Mb downstream at 101 741 125 bp ($-\log_{10}(p)$ = 7.59, Fig. 2d). The minor alleles of the respective top variants had frequencies between 0.28 and 0.43 and reduced birth weight by 1.33–2.35 kg depicting additive allele effects (Fig. 2e). The LD between the three top variants of these QTL regions ranged between r^2 = 0.14 (D'=0.37) for the first and second top variant (5:92 970 239 and 5:99 255 711) to r^2 = 0.33 (D'=0.81) for the second and third top variant (5:99 255 711 and 5:101 741 125), suggesting a low and moderate joint inheritance, respectively.

The top variant rs519831213 at 92 970 239 bp was an intergenic variant between the genes *RERGL* (Ras-related and oestrogen-regulated growth inhibitor-like) and *LMO3* (LIM domain only 3), which were located 1.0 Mb upstream and 0.3 Mb downstream of the top variant, respectively (Table 2, Fig. 2b). The QTL itself spanning 0.2 Mb contained no gene (Table 3).

The top variant at 99 255 711 bp was an intergenic variant located around 2.5 kb upstream of ENSBTAG00000054018 and 30 kb downstream of *KLRC1* (killer cell lectin-like receptor subfamily C) on the negative strand (Fig. 2c). *KLRC1* is a protein-coding gene known to be involved in metabolic pathways related to the immune response. The other genes in this QTL region between 99 155 711 and 99 355 711 bp, spanning 0.2 Mb, were ENSBTAG00000049367 and ENSBTAG00000052486 (Table 3). The genes *KLRC1*, ENSBTAG00000054018, and ENSBTAG00000049367 contained eleven variants with moderate impact on gene transcripts. ENSBTAG00000049367 harboured two tolerated (5:99 171 008, rs110189296 and 5:99 171 139, rs211288086) and one deleterious missense variant (5:99 169 358, rs135507591), ENSBTAG00000054018 harboured five tolerated missense variants (5:99 213 847, rs475014238; 5:99 213 865, rs444872274; 5:99 213 887, rs442536873; 5:99 252 995, rs3423601692; 5:99 253 154, no rsID), and *KLRC1* harboured two tolerated (5:99 289 124, rs209818105 and 5:99 290 856, rs136588070) and one deleterious missense variant (5:99 290 937, rs209617608) (Supplementary Table S5). The deleterious variant in ENSBTAG00000049367 changed the codon in exon 4 out of 6 exons and the amino acid from threonine to alanine at protein position 117 of a 219 amino acid long protein. The deleterious variant in *KLRC1* changed the codon in exon 1 out of 6 exons and the amino acid from lysine to glutamic acid at protein position 18 of a 236 amino acid long protein, not affecting the functional domain "C-type lectin" of KLRC1 which is located at the protein positions 127–231 (UniProtKB: A0A2Z2AM52_BOVIN). Moreover, those three genes were involved in the significantly enriched immune response pathways "Antigen processing and presentation" (KEGG:04612) and "Natural killer cell mediated cytotoxicity" (KEGG:04650, Supplementary Table S6).

The top variant rs135635336 at 101 741 125 bp was an intergenic variant located between the genes *CLEC4E* (C-type lectin domain family 4, member E) and *CD163* (CD163 molecule) (Fig. 2d). This QTL spanned 1.15 Mb and contained 14 genes (Table 3) including three C-type lectin domain genes *CLEC4E*, *CLEC6A*, and *CLEC4D* that were contained in the significantly enriched GO term "Antifungal innate immune response" (GO:0061760) and other immune and signalling pathways (Supplementary Table S6). Altogether, 202 missense variants were found in 11 of the 14 genes in the QTL associated with the top variant rs135635336 at 101 741 125 bp (*CLEC6A*, *CLEC4E*, ENSBTAG00000053242, ENSBTAG00000053486, *WC–7*, ENSBTAG00000048875, *WC1-12*, ENSBTAG00000024318, ENSBTAG00000052511, *WC1.3*, *WC1*), whereof 52 were deleterious and 150 were tolerated, showing a high variability of genes in this QTL (Supplementary Table S5). Additionally, seven genes

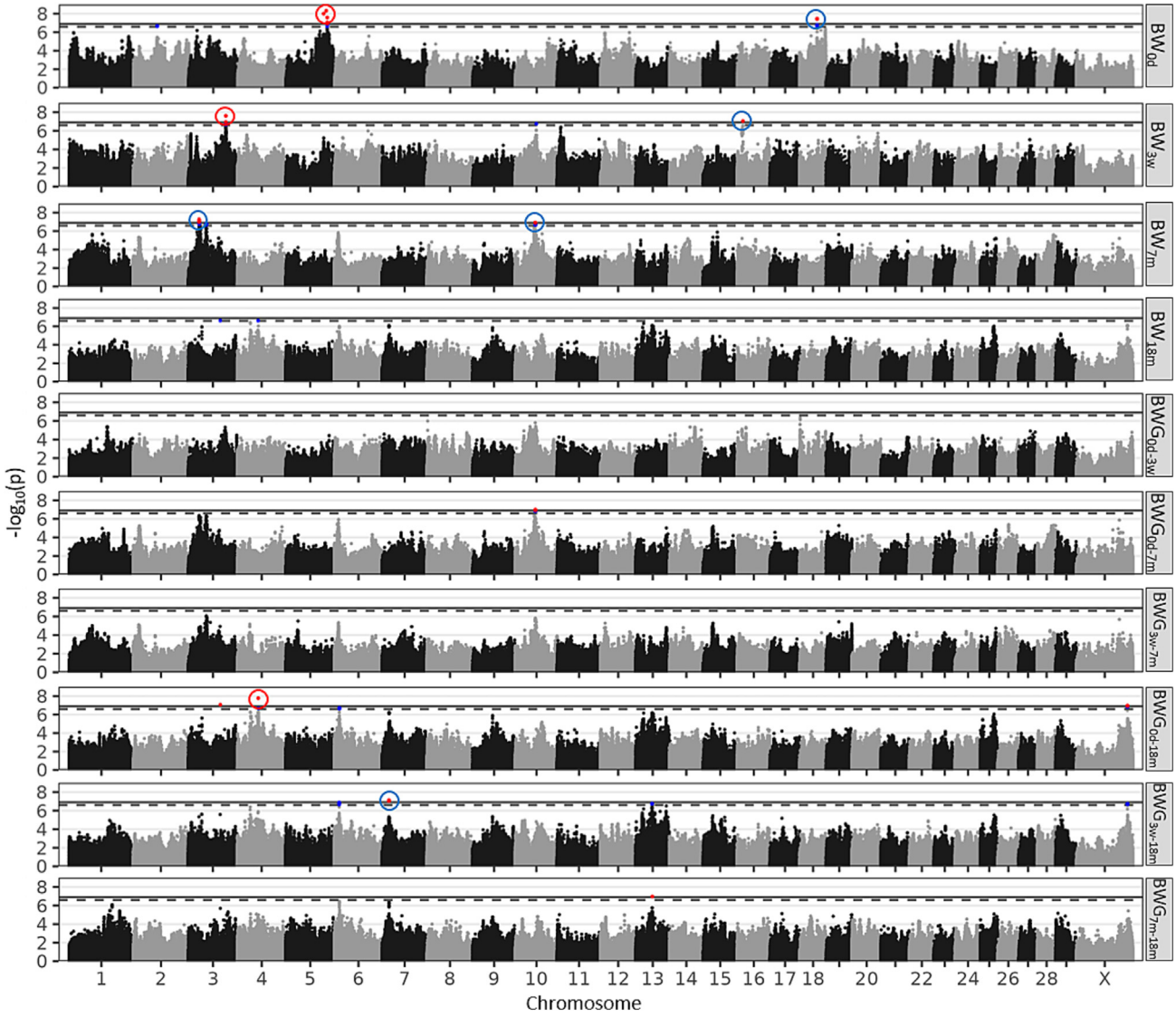

Fig. 1. Manhattan plots for all investigated traits in German Black Pied (DSN) bulls. BW_{0d}, BW_{3w}, BW_{7m}, BW_{18m} refer to BW at birth, 3 weeks, 7 months, and 18 months; $BWG_{0d-3w/7m/18m}$, $BWG_{3w-7m/18m}$, and BWG_{7m-18m} indicate BW gain over respective periods. Red circles highlight highly significant quantitative trait loci (QTL) regions ($-\log_{10}(p) \geq 7.6$) and blue circles interesting significant QTL regions ($-\log_{10}(p) \geq 6.9$) which are explained in detail in the results section. Dashed and solid lines mark genome-wide suggestive and significant thresholds, respectively.

(ENSBTAG00000053486, *WC-7*, ENSBTAG00000048875, *WC1-12*, ENSBTAG00000052511, *WC1.3*, *WC1*) contained twelve high–impact variants like frameshift, stop gained, and splice donor/acceptor variants.

Another QTL was significant for birth weight on chromosome 18 with the top variant rs209682285 at 42 909 805 bp ($-\log_{10}(p) = 7.45$) (Table 2, Fig. 3a). The minor allele A, with a frequency of 0.15, increased birth weight by 2.72 kg (Fig. 3b). The top variant was located between the genes *ZNF507* (zinc finger protein 507) and *DPY19L3* (dpy-19 like C-mannosyltransferase 3), approximately 7 kb downstream of *ZNF507* and 9 kb upstream of *DPY19L3*. The QTL contained a total of nine genes (Table 3), but only the genes *ZNF507*, *ANKRD27* (ankyrin repeat domain 27), *RGS9BP* (regulator of G protein signalling 9 binding protein), *NUDT19* (nudix hydrolase 19) and *SLC7A9* (solute carrier family 7 member 9) contained altogether ten missense variants (Supplementary Table S5). Of those, three missense variants were predicted as deleterious. The deleterious missense variants rs437638954 (18:43 154 706) and rs481740327 (18:43 155 021) in *RGS9BP* were both located in the only exon, with rs437638954 changing the amino acid cysteine to glycine at amino acid position 19 and with rs481740327 changing the amino acid valine to leucine at position 124 out of 237. The deleterious missense variant rs109419027 (18:43 266 263) in *SLC7A9* changed the codon in exon 12 out of 14 exons and the amino acid tyrosine to alanine at amino acid position 418 out of 487. With regard to GO enrichment, "broad specificity neutral L–amino acid:basic L-amino acid antiporter activity" (GO:0180009) containing *SLC7A9* was found significantly enriched (Supplementary Table S6).

BW at 3 weeks – Significant loci on chromosomes 3 and 16

We identified two significant QTL regions for BW_{3w} on chromosomes 3 and 16. The top variant rs43353145 on chromosome 3 was located at 92 613 103 bp ($-\log_{10}(p) = 7.61$, Table 2, Fig. 4a). Its minor allele C, with a frequency of 0.32, increased BW_{3w} by 2.51 kg (Fig. 4b). This variant was located in intron 2 of the *GLIS1* gene (GLIS family zinc finger 1), which encodes a transcription factor that regulates gene expression during cell differentiation. Six other candidate genes, *DMRTB1* (DMRT–like family B with

Table 2
Top variants for BW and BW gain traits in German Black Pied (DSN) bulls. Top variants (chromosome:position) are listed with their rs-ID, reference (Ref) and alternative (Alt) alleles, minor allele (MA) and its frequency (MAF), as well as the allele substitution effect of MA (β_{MA}) and its SE. Test statistics with $-\log_{10}(p) \geq 6.9$ were significant and with $-\log_{10}(p) \geq 7.6$ highly significant. See Table 1 for trait abbreviation explanations.

Trait and top variants	rs-ID	Ref/Alt	MA (MAF)	β_{MA} (SE)	$-\log_{10}(p)$
BW_{0d}					
5:92 970 239	rs519831213	C/T	T (0.43)	−1.33 (0.54)	8.00
5:99 255 711	–	A/G	A (0.42)	−1.49 (0.45)	8.32
5:101 741 125	rs135635336	C/T	C (0.28)	−2.35 (0.42)	7.59
18:42 909 805	rs209682285	G/A	A (0.15)	2.72 (0.46)	7.45
BW_{3w}					
3:92 613 103	rs43353145	C/T	C (0.32)	2.51 (0.59)	7.61
16:14 033 067	rs41590750	C/T	C (0.10)	−4.50 (0.81)	7.05
BW_{7m}					
3:26 611 023	rs481113165	C/T	T (0.12)	−13.6 (2.47)	7.28
10:49 410 011	rs135276827	ATTTGT/A	ATTTGT (0.47)	7.44 (2.21)	6.91
BWG_{0d-7m}					
10:49 409 323	rs133466503	C/CT	CT (0.49)	−6.59 (2.25)	6.98
BWG_{0d-18m}					
3:78 993 067	rs207494601	T/C	C (0.06)	−42.34 (7.78)	7.07
4:51 589 752	rs449334628	A/C	C (0.15)	−29.74 (5.18)	7.78
X:126 226 630	rs133366378	C/T	T (0.25)	−15.92 (2.55)	6.96
BWG_{3w-18m}					
7:17 246 788	rs385757234	G/A	G (0.45)	21.86 (5.82)	7.11
BWG_{7m-18m}					
6:12 147 927	rs109832541	T/C	C (0.43)	−28.28 (6.6)	6.97
13:41 013 268	rs208007513	A/G	G (0.11)	−34.78 (6.18)	7.52

proline-rich C-terminal 1), ENSBTAG00000039540, *LRP8* (LDL receptor–related protein 8), *CPT2* (carnitine palmitoyl transferase 2), *CZIB* (CXXC motif containing zinc binding protein), and *MAGOH* (mago homolog, exon junction complex subunit), were also located in this QTL (Table 3). Five of the seven candidate genes, including *GLIS1*, possessed missense variants that could affect protein function (Supplementary Table S5). Moreover, the genes *GLIS1*, ENSBTAG00000039540, and *CPT2* contained each one deleterious missense variant. In *GLIS1*, rs43351481 (3:92 739 495) changed a codon in the last exon 11 of the gene, leading to the exchange of leucine to methionine at the amino acid position 758 out of 796; none of the known functional domains of *GLIS1* was affected. The variant rs382616052 (3:92 806 607) changed a codon in the only exon altering the amino acid from aspartic acid to histidine at protein position 260 out of 330. The variant was located in a repeat region of the protein Solute carrier family 25 member 3 encoded by ENSBTAG00000039540 (UniProtKB: A6QQ54_BOVIN). The variant rs211475725 (3:93 035 737) changed the codon in exon 4 out of 5 exons, substituting glutamic acid with aspartic acid at protein position 172 out of 658, which is located in the choline/carnitine acyltransferase domain of *CPT2* (UniProtKB: F1N1M7_BOVIN). The gene *LRP8* was included in the significant GO enrichment term "reelin receptor activity" (GO:0038025) (Supplementary Table S6).

The top variant rs41590750 associated with BW_{3w} on chromosome 16 was located at 14 033 067 bp ($-\log_{10}(p) = 7.05$) (Table 2, Fig. 4c). The minor allele C, with a frequency of 0.10, decreased BW_{3w} by 4.5 kg (Fig. 4d). This variant was located between the genes *RGS18* (regulator of G protein signalling 18) and ENSBTAG00000053584 which were 0.88 Mb upstream and 0.77 Mb downstream of the top variant, respectively (Table 2). The QTL spanned from 11 725 340 to 14 133 067 bp on chromosome 16, covering 2.41 Mb, and contained ten candidate genes: *CDC73* (cell division cycle 73), *B3GALT2* (beta-1, 3-galactosyltransferase 2), *GLRX2* (glutaredoxin 2), *RO60* (Ro60, Y RNA binding protein), *UCHL5* (ubiquitin C-terminal hydrolase L5), ENSBTAG00000034366, and four regulator of G protein signalling genes *RGS13*, *RGS1*, *RGS21*, and *RGS18* (Table 3). In these genes, five missense variants with a moderate VEP impact were identified: rs136040928 (16:12 176 541), rs135094480 (16:12 426 398), rs135275219 (16:12 700 837), rs210279648 (16:12 893 371), and rs109098798 (16:13 126 197) (Supplementary Table S5). SIFT scores, which predict the *P* of missense variants being tolerated or deleterious, indicated that all identified variants were tolerated. The four regulator of G protein signalling genes and ENSBTAG00000034366 of this region contributed to significant signalling pathways, particularly the "G alpha (q) signalling events" pathway (REAC:R-BTA-416476) (Supplementary Table S6).

BW at 7 months – Significant loci on chromosomes 3 and 10

We identified one significant QTL region for body BW_{7m} on chromosome 3 and one which was significant for both BW_{7m} and BWG_{0d-7m} on chromosome 10. For BW_{7m}, the top variant rs481113165 on chromosome 3 was located at 26 611 023 bp ($-\log_{10}(p) = 7.28$) (Table 2, Fig. 5a). The minor allele T, with a frequency of 0.12, decreased BW_{7m} by 13.6 kg (Fig. 5b). This variant was located within intron 2 of 11 of the gene *IGSF3* (immunoglobulin superfamily member 3) and the corresponding QTL spanned from 26 389 879 to 26 711 023 bp on chromosome 3, covering 0.32 Mb (Table 3). This region encompassed two more candidate genes: *CD2* (CD2 molecule) and *CD58* (CD58 molecule). Notably, *CD2* and *CD58* were involved in important biological pathways, including "cell adhesion molecules" (KEGG:04514) and "cell surface interactions at the vascular wall" (REAC: R-BTA-202733). In addition, nine missense variants with moderate VEP impact were identified within this QTL: rs133747802 (3:26 477 535), rs135279784 (3:26 477 546), rs136558757 (3:26 477 547), rs134192317 (3:26 477 574), rs136618819 (3:26 477 646), rs133131840 (3:26 477 662), rs137143925 (3:26 649 097), rs136136244 (3:26 731 425), and rs384518130 (3:26 731 431)

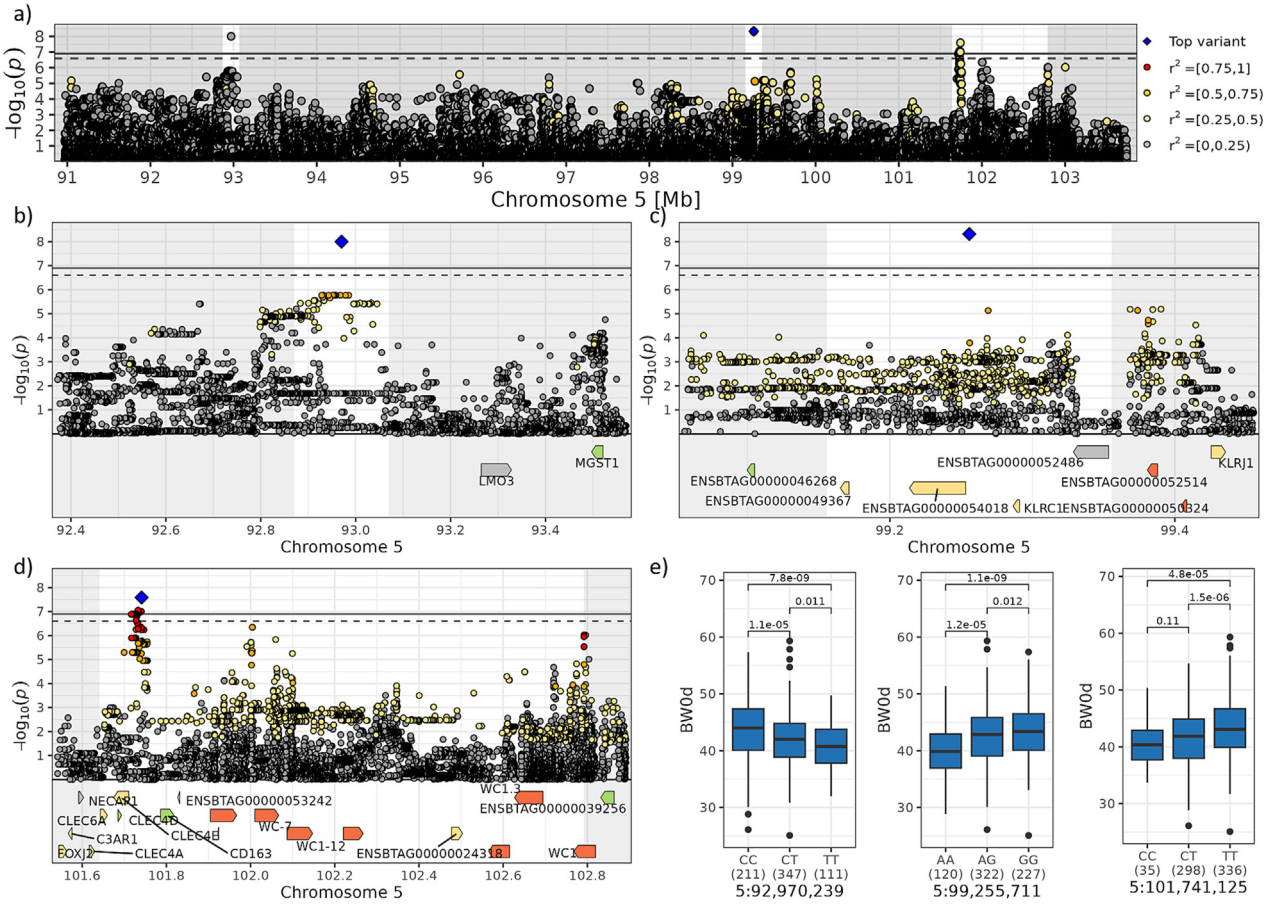

Fig. 2. (a) Three loci for birth weight (BW_{0d}) on chromosome 5 in German Black Pied (DSN) bulls. (b–d) Detailed presentation of the three loci including genes. Suggestive ($P < 0.1$, dashed line) or significant ($P < 0.05$, solid line) thresholds are shown. Genes containing variant(s) with low impact, moderate, or high impact are coloured green, yellow, and orange, respectively, otherwise grey. (e) Corresponding variant effect plots of the three loci. For each locus, the top variant position, the number of animals per genotype and the *P*-values for the pairwise genotype comparisons are given.

(Supplementary Table S5). Of these, only rs135279784 (3:26 477 546) which was located in *CD2* was scored as deleterious by the Sorting Intolerant From Tolerant score, changing the codon in exon 3 out of 5 exons, and thus substituting lysine with asparagine at position 189 out of 338 amino acids in the protein. This variant did not affect the Immunoglobulin V-set domain which is located at protein positions 28–120 (UniProtKB: Q148M9_BOVIN). Moreover, the significantly associated variant rs438644529 (3:26 569 653) which was located 4 745 bp upstream of *IGSF3* was found to be in the putative transcription factor binding sites of *NR2C1* (nuclear receptor subfamily two group C member 1) (JASPAR: MA1535.1) and *NR2C2* (nuclear receptor subfamily two group C member 1) (JASPAR: MA1536.1). The top variant rs481113165 (3:26 611 023) of this QTL was located within the putative transcription factor binding site (JASPAR: MA1522.1) of *MAZ* (myc-associated zinc finger), a transcriptional regulator that potentially has dual roles in transcription initiation and termination, and within the putative transcription factor binding site (JASPAR: MA1965.1) of *SP5* (transcription factor Sp5), a probable transcriptional activator that plays a role in the coordination of changes in transcription required to generate pattern in the developing embryo.

For the QTL on chromosome 10, the top variants for BW_{7m} and $BWG_{0d\text{-}7m}$ were rs135276827 at 49 410 011 bp ($-\log_{10}(p) = 6.91$) and rs133466503 at 49 409 323 bp ($-\log_{10}(p) = 6.98$), respectively (Table 2, Fig. 5c). These top variants were only 688 bp apart and thus were in high LD ($r^2 = 0.82$, D'=0.98). The minor allele ATTTGT of rs135276827, with a frequency of 0.47, increased BW_{7m} by 7.44 kg, while the minor allele CT of rs133466503, with a frequency of 0.49, decreased $BWG_{0d\text{-}7m}$ by 6.59 kg (Fig. 5d). The substitution effects were of opposite directions which was possible since the minor allele frequency was close to 0.50 for both minor alleles. Both top variants were located in intron 1 out of 10 of the gene *RORA* (RAR-related orphan receptor alpha). The QTL contained the four additional candidate genes *TLN2* (talin 2), *C2CD4B* (C2 calcium-dependent domain-containing protein 4B), ENSBTAG00000010952, and *VPS13C* (vacuolar protein sorting 13 homolog C) (Table 3). Altogether, 17 tolerated and 9 deleterious missense variants were identified in those genes (Supplementary Table S5). Notably, three high-impact variants were identified in the *VPS13C* gene. These included the frameshift variants rs137793231 (10:48 112 397), which inserted a nucleotide into the DNA sequence, and rs133133999 (10:48 140 665), which deleted a nucleotide. Additionally, the splice donor variant rs210308889 (10:48 116 721) was detected. The *VPS13C* gene itself was included in the enriched GO terms "negative regulation of parkin-mediated stimulation of mitophagy in response to mitochondrial depolarisation" (GO:1905090) and "dense core granule membrane" (GO:0032127) (Supplementary Table S6). The significantly associated variants rs137408179 (10:49 410 017) and rs132911489 (10:49 410 019) which were also located in intron 1 of *RORA* were found to be in the in the putative transcription factor binding site of *MEIS1* (homeobox protein Meis1) (JASPAR: MA0498.2), which is required for hematopoiesis, megakaryocyte

Table 3
Quantitative trait loci (QTL) regions associated with BW and BW gain traits in German Black Pied (DSN) bulls. The table includes the associated trait, QTL location (chromosome: start–end), QTL length (Mb), and positional candidate genes within the QTL region. See Table 1 for trait abbreviation explanations.

Trait and QTL location	Length	Genes within QTL region (Number of genes)
BW_{0d}		
5: 92 870 239 - 93 070 239	0.20	–
5: 99 155 711 - 99 355 711	0.20	ENSBTAG00000049367, ENSBTAG00000054018, *KLRC1*, ENSBTAG00000052486 (4)
5: 101 641 125 - 102 792 163	1.15	*CLEC6A*, *CLEC4E*, *CLEC4D*, *CD163*, *WC-7*, *WC1-12*, ENSBTAG00000053242, ENSBTAG00000053486, *WC1.3*, ENSBTAG00000054128, ENSBTAG00000048875, ENSBTAG00000024318, ENSBTAG00000052511, *WC1* (14)
18: 42 809 805 - 43 299 928	0.49	*ZNF507*, *DPY19L3*, *PDCD5*, *ANKRD27*, *RGS9BP*, *NUDT19*, ENSBTAG00000049809, *TDRD12*, *SLC7A9* (9)
BW_{3w}		
3: 92 513 103 - 93 086 162	0.57	*GLIS1*, *DMRTB1*, ENSBTAG00000039540, *LRP8*, *MAGOH*, *CZIB*, *CPT2* (7)
16: 11 725 340 - 14 133 067	2.41	*CDC73*, *B3GALT2*, *GLRX2*, *RO60*, *UCHL5*, ENSBTAG00000034366, *RGS13*, *RGS1*, *RGS21*, *RGS18* (10)
BW_{7m}		
3: 26 389 879 - 26 711 023	0.32	*CD2*, *IGSF3*, *CD58* (3)
10: 47 128 191 - 49 509 323	2.38	*TLN2*, *C2CD4B*, ENSBTAG00000010952, *VPS13C*, *RORA* (5)
BWG_{0d-7m}		
10: 47 128 191 - 49 509 323	2.38	*TLN2*, *C2CD4B*, ENSBTAG00000010952, *VPS13C*, *RORA* (5)
BWG_{0d-18m}		
3: 78 893 067 - 79 093 067	0.20	*PDE4B* (1)
4: 51 489 752 - 52 146 870	0.66	*CAPZA2*, *MET*, *CAV1*, *CAV2* (4)
X: 126 082 924 - 126 771 255	0.69	*NHS*, ENSBTAG00000031256, *REPS2*, *RBBP7*, *TXLNG* (5)
BWG_{3w-18m}		
7: 17 146 788 - 17 346 788	0.20	*MYO1F*, *ADAMTS10*, *ACTL9*, *OR2Z1*, ENSBTAG00000054873, ENSBTAG00000053612 (6)
BWG_{7m-18m}		
6: 12 047 927 - 12 851 659	0.80	*CAMK2D*, ENSBTAG00000002392, ENSBTAG00000048956, ENSBTAG00000051025 (4)
13: 40 652 358 - 41 411 387	0.76	*XRN2*, *NKX2-4*, *NKX2-2*, ENSBTAG00000035643, *PAX1* (5)

lineage development and vascular patterning. Furthermore, rs132911489 (10:49 410 019) was located in the putative transcription factor binding site of *RBPJ* (recombining binding protein suppressor of hairless) (JASPAR: MA1116.1), a transcriptional regulator that plays a central role in Notch signalling, a pathway that regulates a broad spectrum of cell-fate determinations. Additionally, the significantly associated variant rs133379398 (10:49 413 424) which was also located in intron 1 of *RORA* was found to be in the putative transcription factor binding site for *Arid3a* (AT-rich interactive domain-containing protein 3A) (JASPAR: MA0151.1), a transcription factor that is involved in B-cell differentiation.

BW gain to 18 months – Significant loci on chromosomes 4 and 7

We identified one significant QTL region for BWG_{0d-18m} on chromosome 4 and BWG_{3w-18m} on chromosome 7. The top variant rs449334628 on chromosome 4 was located at 51 589 752 bp ($-\log_{10}(p)$ = 7.78) (Table 2, Fig. 6a). The minor allele C, with a frequency of 0.15, decreased BWG_{0d-18m} by 29.74 kg (Fig. 6b). The whole QTL spanned the region from 51 489 752 to 52 146 870 bp, covering 0.66 Mb, and encompassed four candidate genes: *CAPZA2* (capping actin protein of muscle Z-line alpha subunit 2), *MET* (MET proto-oncogene, receptor tyrosine kinase), *CAV1* (caveolin 1), and *CAV2* (caveolin 2) (Table 3). Within this QTL region, three missense variants were identified in *MET*: rs476903927 (4:51 640 753), rs439124431 (4:51 640 844), and rs383187998 (4:51 656 645). Among these, rs439124431 was predicted to be deleterious by SIFT, whereas the other two variants were classified as tolerated (Supplementary Table S5). Furthermore, *CAPZA2* and *MET* were involved in "regulation of actin filament organisation" (GO:0110053), which is essential for cytoskeletal stability and cellular signalling, particularly in muscle cells (Supplementary Table S6).

For the QTL on chromosome 7, the top variant rs385757234 was located at 17 246 788 bp ($-\log_{10}(p)$ = 7.11) (Table 2, Fig. 6c). The minor allele G, with a frequency of 0.45, increased BWG_{3w-18m} by 21.86 kg (Fig. 6d). This QTL spanned the region from 17 146 788 to 17 346 788 bp, covering 0.20 Mb, and contained six candidate genes: *MYO1F* (myosin IF), *ADAMTS10* (ADAM metallopeptidase with thrombospondin type 1 motif 10), *ACTL9* (actin-like 9), *OR2Z1* (olfactory receptor family 2 subfamily Z member 1), ENSBTAG00000054873, and ENSBTAG00000053612 (Table 2). Within this QTL region, six tolerated missense variants were identified, including rs135525700 (7:17 252 201) in *ACTL9*, several variants (rs210381561, 7:17 264 520; rs385820627, 7:17 264 589; rs209434743, 7:17 264 923) in *OR2Z1*, and two in ENSBTAG00000054873 (rs210113204, 7:17 277 265; and rs211445474, 7:17 277 823) (Supplementary Table S5). Additionally, two deleterious missense variants were identified in *OR2Z1* (rs380843465, 7:17 264 389; no rs-ID, 7:17 264 742). The variant rs380843465 in exon 1 out of 2 exons changed arginine to leucine at protein position 17 of a 311 amino acids long protein but did not affect any domain of *OR2Z1*. The second deleterious variant in *OR2Z1* at 17 264 742 bp changed the codon in exon 1 from arginine to cysteine at protein position 135, which affected the functional domain "G-protein coupled receptors family 1 profile" located at protein positions 38–287 (UniProtKB: F1N170_BOVIN).

Discussion

The analysis of genetic associations in this study provided valuable insights into the genetic architecture of growth traits in DSN bulls, with differences observed across various developmental stages. Despite the sample size of the investigated DSN population being relatively small, we were able to identify 14 QTLs associated with BW and BW gain traits, which was supported in part by the moderate to high heritability of the examined traits.

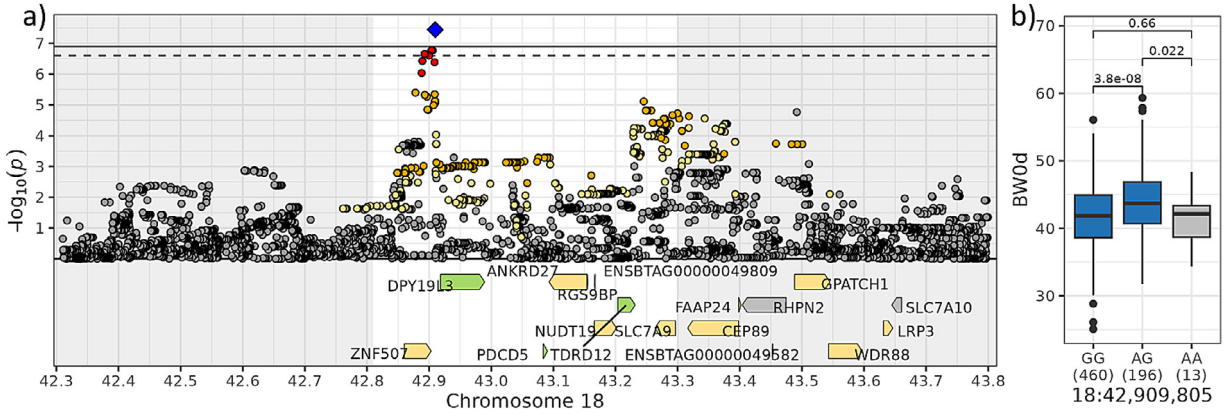

Fig. 3. (a) Locus for birth weight (BW_{0d}) on chromosome 18 including genes in German Black Pied (DSN) bulls. Suggestive ($P < 0.1$, dashed line) or significant ($P < 0.05$, solid line) thresholds are shown. Genes containing variant(s) with low impact or moderate are coloured green or yellow, otherwise grey. (b) Corresponding variant effect plot including the top variant position, the number of animals per genotype and the *P*-values for the pairwise genotype comparisons.

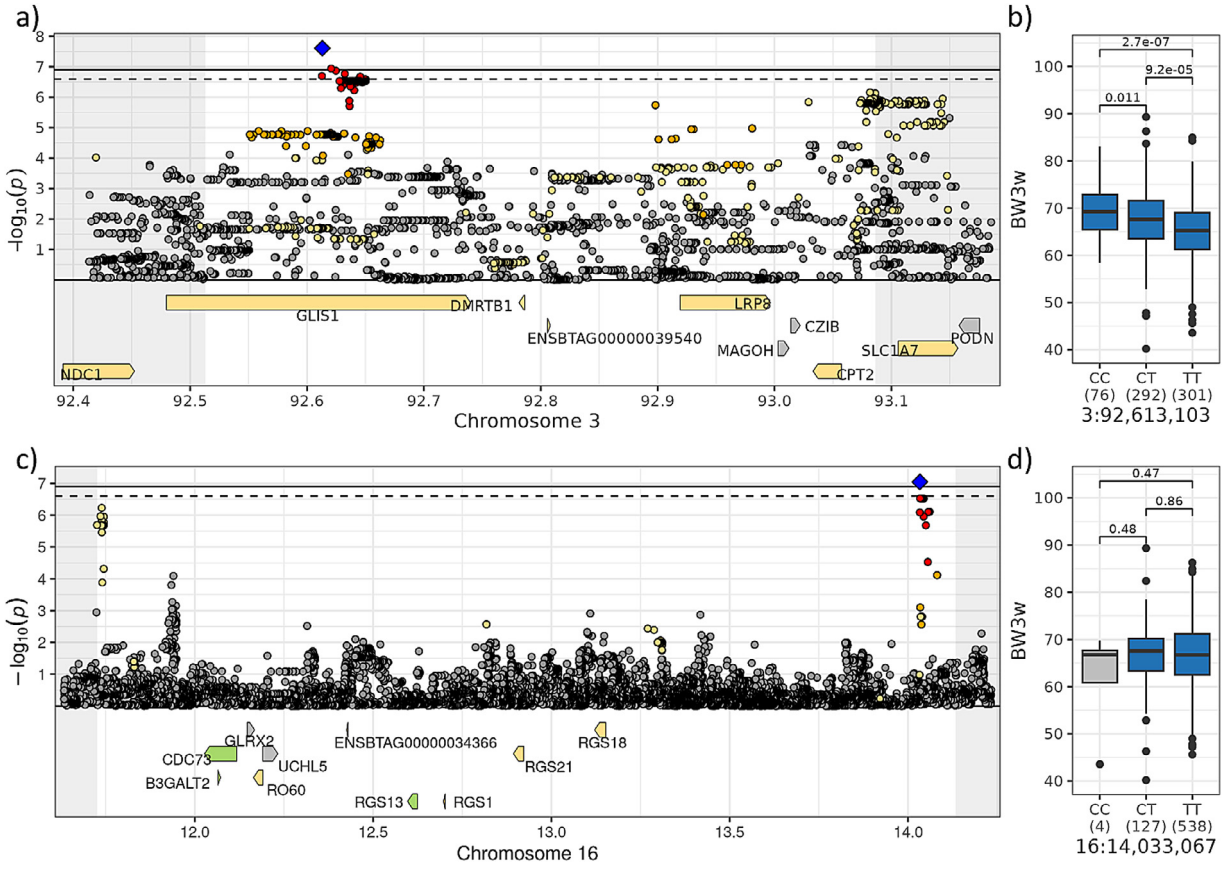

Fig. 4. (a) Locus for BW at 3 weeks (BW_{3w}) on chromosome 3 including genes in German Black Pied (DSN) bulls. (b) Corresponding variant effect plot. (c) Locus for BW_{3w} on chromosome 16 including genes. (d) Corresponding variant effect plot. Suggestive ($P < 0.1$, dashed line) or significant ($P < 0.05$, solid line) thresholds are shown. Genes containing variant(s) with low impact or moderate are coloured green or yellow, otherwise grey. Variant effect plots contain the top variant position, the number of animals per genotype and the *P*-values for the pairwise genotype comparisons.

The most significant genetic associations were identified for birth weight. This can be attributed to the highest number of data points for birth weight, as not all DSN bulls had reached older ages by the time of data evaluation. Additionally, birth weight is reliable to measure and exhibits the highest heritability among the traits studied (Table 1), indicating that a larger proportion of its variation is due to genetic factors rather than environmental influences. The traits measured later in life, such as BW at 7 or 18 months and corresponding BW gain traits, were more affected by environmental factors like feeding practices and management conditions. The combination of high heritability and greater data availability made birth weight more genetically predictable, facilitating the identification of specific genomic loci and yielding the most reliable results. Although all investigated traits are complex and polygenic, the presence of major gene effects or high-effect loci can further facilitate QTL detection, even in small populations (Mancin et al.,

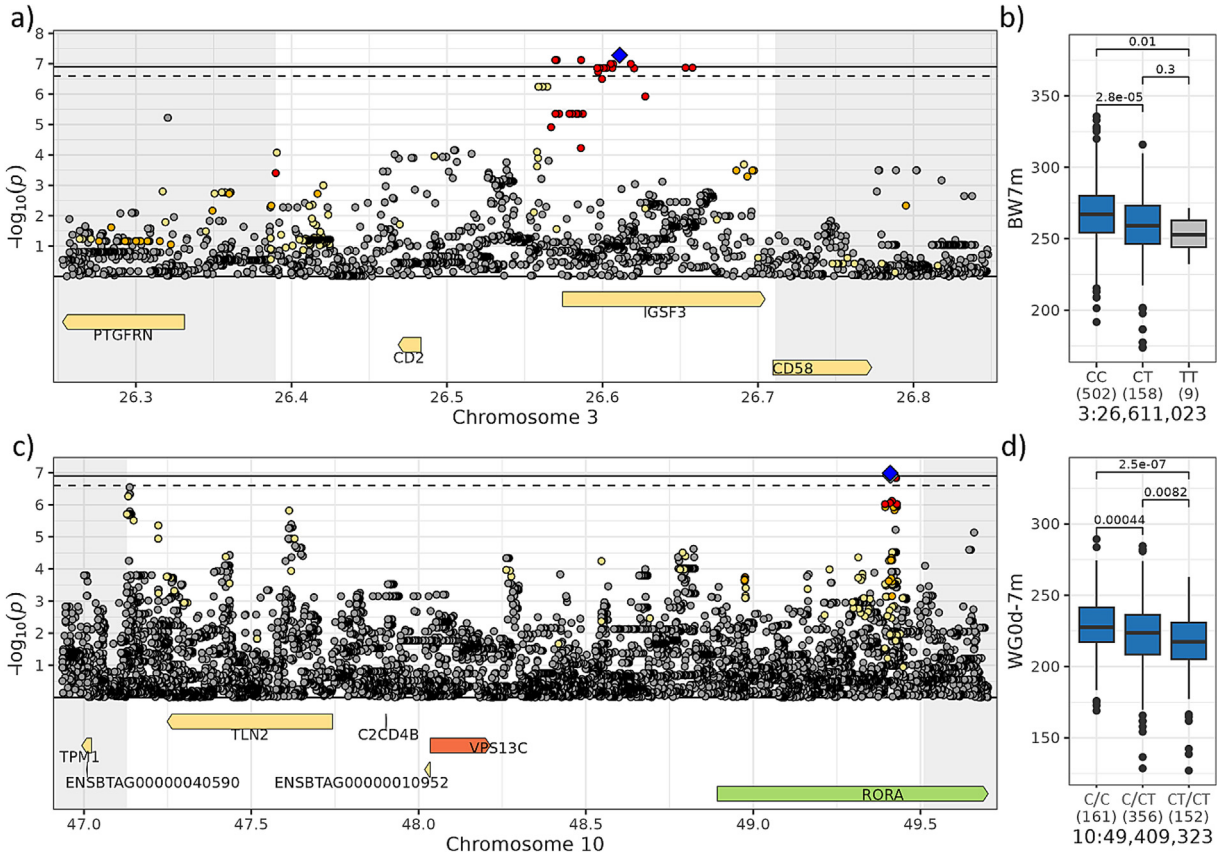

Fig. 5. (a) Locus for BW at 7 months (BW_{7m}) on chromosome 3 including genes in German Black Pied (DSN) bulls. (b) Corresponding variant effect plot. (c) Locus for BW gain from birth to 7 months ($BWG_{0d\text{-}7m}$) on chromosome 10 including genes. (d) Corresponding variant effect plot. Suggestive ($P < 0.1$, dashed line) or significant ($P < 0.05$, solid line) thresholds are shown. Genes containing variant(s) with low impact, moderate, or high impact are coloured green, yellow, and orange, respectively, otherwise grey. Variant effect plots contain the top variant position, the number of animals per genotype and the *P*-values for the pairwise genotype comparisons.

2022). For example, in the case of birth weight, an effect size of 2.7 kg for a single locus, which is 6.75% of a birth weight of 40 kg, is already considered a large genetic effect.

The high genetic correlation between BW_{7m} and $BWG_{0d\text{-}7m}$ suggests shared genetic influences, while the unique loci for other traits imply stage-specific gene regulation. The fact that most QTLs significantly identified for different traits did not overlap, except for one on chromosome 10 for BW_{7m} and $BWG_{0d\text{-}7m}$, suggests that different developmental stages are controlled by distinct genetic mechanisms. However, a closer look at significant QTL regions, e.g. QTL with the top variant rs449334628 at 51 589 752 bp on chromosome 4, reveals that undetected effects on correlated traits may exist, although we could not provide significant evidence. The causes for missing significance can be the changing gene effect during ontogenesis on one hand, but also insufficient statistical power to detect small effects due to a small sample size that was too small, insufficient accuracy of BW measurements inherent to field data, and undetected environmental influence on the other hand. With more data, additional overlaps between traits with high genetic correlations are expected, potentially revealing common genetic factors underlying multiple growth phases.

The 14 identified QTLs across ten chromosomes contained a total of 77 candidate genes. Among the positional candidate genes for birth weight were genes influencing muscle and fat tissue development, intracellular transport, and immune response on chromosome 5 and metabolism, nutrient transport, and cellular signalling on chromosome 18. In detail, *RERGL* and *LMO3* have functions related to muscle and fat tissue development. *RERGL* is predicted to enable G protein activity and GTP binding, with enrichment in smooth muscle cells (https://www.proteinatlas.org), while *LMO3* is involved in transcriptional regulation, neural development, and has been implicated in the reprogramming of adipose tissue during obesity (Wagner et al., 2021). The genes coding for CD163 and the WC family members were associated with cargo receptor activity highlighting the importance of intracellular transport mechanisms for the rapid growth and development occurring around birth. Additionally, several genes related to immune function were identified, including *KLRC1*, a killer cell lectin-like receptor, and members of the C-type lectin family (*CLEC4E*, *CLEC6A*, and *CLEC4D*) (https://www.proteinatlas.org). These genes were associated with immune response pathways related to lectin. Lectins, which are involved in pathogen recognition and immune activation by binding to specific sugars on the surfaces of bacteria, viruses, and fungi, suggest a potential role for immune function in early development. The repeated detection of immune-related genes in loci associated with birth weight suggests a link between proper immune function and a healthy early development. While immune-related pathways are frequently implicated in GWAS across various traits, their specific role in regulating neonatal growth and adaptation to early environmental challenges needs further investigation. Further, *DPY19L3* stands out due to its proximity to the top variant in the QTL on chromosome 18. Knockout studies in mice have shown that loss of *DPY19L3* leads to decreased circulating free fatty acids, haemoglobin content, and erythrocyte cell number, suggesting a potential influence on energy metabolism and blood parameters (https://www.mousephenotype.org). *SLC7A9* is crucial for amino acid transport, which is fundamental for cellular homeostasis and metabolic

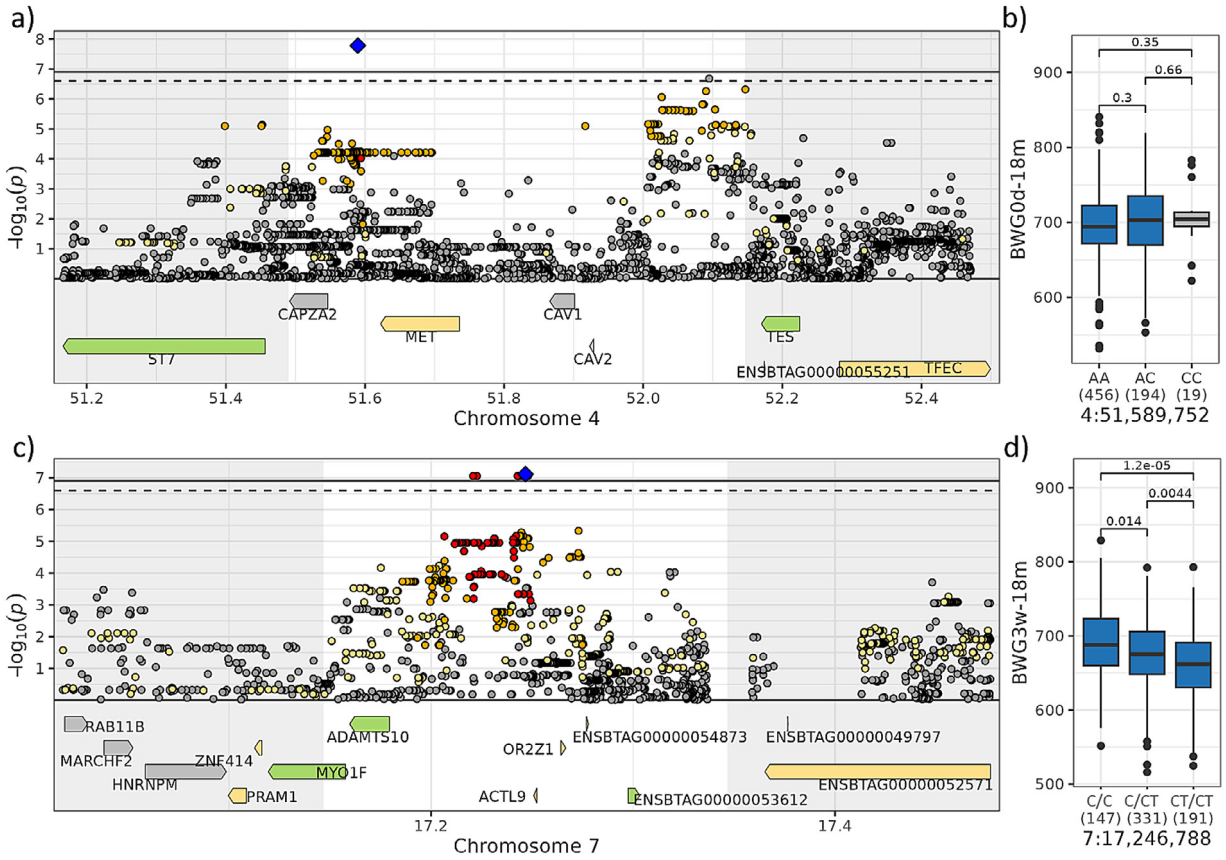

Fig. 6. (a) Locus for BW gain from birth to 18 months ($BWG_{0d\text{-}18m}$) on chromosome 4 including genes in German Black Pied (DSN) bulls. (b) Corresponding variant effect plot. c) Locus for BW gain from 3 weeks to 18 months ($BWG_{3w\text{-}18m}$) on chromosome 7 including genes. (d) Corresponding variant effect plot. Suggestive ($P < 0.1$, dashed line) or significant ($P < 0.05$, solid line) thresholds are shown. Genes containing variant(s) with low impact or moderate are coloured green or yellow, respectively, otherwise grey. Variant effect plots contain the top variant position, the number of animals per genotype and the *P*-values for the pairwise genotype comparisons.

balance. *ANKRD27* and *RGS9BP* are less likely candidate genes affecting birth weight since their functions are more associated with the development of the eye. *ANKRD27* was associated with guanyl-nucleotide exchange factor activity and endosomal trafficking. Knockout mice for *ANKRD27* exhibited decreased retina thickness and short tibiae. *RGS9BP* is involved in regulating G protein-coupled receptor signalling, particularly in phototransduction, and is exclusively expressed in the retina (https://www.proteinatlas.org).

Positional candidate genes for BW_{3w} are involved in transcriptional regulation, energy metabolism, and signalling pathways on chromosome 3 and G protein-coupled receptor signalling on chromosome 16. In detail, on chromosome 3, *GLIS1* functions as an activator and repressor of transcription and is essential for regulating gene expression during cell differentiation as it inhibits the lineage commitment of multipotent cells (Maekawa et al., 2011) (https://www.proteinatlas.org). Thus, *GLIS1* directly impacts tissue formation and growth. *DMRTB1*, a transcription factor involved in germ cell development, sex differentiation, and transcription regulation, likely influences growth patterns indirectly through its effects on hormonal pathways and developmental timing (https://www.proteinatlas.org). *CPT2*, which is responsible for the mitochondrial oxidation of long-chain fatty acids, provides the necessary energy for growth, e.g. during developmental stages. Among its related pathways are "Fatty acid metabolism" and "Transcriptional regulation of white adipocyte differentiation" (https://www.pathcards.genecards.org). The importance of *CPT2* is underscored by knockout studies in mice, where its absence led to severe developmental issues, including preweaning lethality and abnormal vertebrae morphology (https://www.mousephenotype.org). *LRP8*, a low-density lipoprotein receptor involved in signal transduction, lysosomal degradation, and neuron migration (NCBI gene summary), was linked to reelin receptor activity. While reelin signalling is crucial for brain development, it also has broader implications for tissue development and metabolic regulation. Together, these genes highlight the intricate network of transcriptional regulation, energy metabolism, and signalling pathways that may collectively influence early growth. On chromosome 16, no genes were directly located near the significantly associated variants, but the region between the significant variants contained four regulator of G protein signalling genes associated with the significant signalling pathway "G alpha (q) signalling events" (REAC:R-BTA-416476). G protein-coupled receptor signalling plays a critical role in various cellular processes, including those related to growth and development.

Positional candidate genes for BW_{7m} are involved in immune response, cell adhesion, and metabolic regulation on chromosome 3 and growth, development, and energy metabolism on chromosome 10. On chromosome 3, *IGSF3* is involved in immune response processes and is specifically enriched in skeletal muscle fibroblasts (https://www.proteinatlas.org), suggesting a link between immune function and muscle development. The significantly associated variants upstream of *IGSF3* were found in the putative binding sites for the transcription factors NR2C1 and NR2C2. Both transcription factors are primarily repressors of a broad range of genes and were reported to repress embryonic and foetal globin transcription or may be involved in stem cell proliferation and differentiation. *CD2* encodes a surface antigen found on T-cells and is crucial for

triggering T-cell responses, playing a vital role in immune regulation (https://www.proteinatlas.org). *CD58*, a ligand for the CD2 protein, is essential for the adhesion and activation of T lymphocytes. The involvement of these genes in immune response and cell adhesion pathways suggests a potential link between immune system regulation and growth traits since only resilient and healthy animals are likely to reach their genetic potential for growth. On chromosome 10, *RORA* encodes a nuclear hormone receptor that regulates a wide range of biological processes, including embryonic development, cellular differentiation, immunity, circadian rhythm, and various metabolic pathways such as lipid, steroid, and glucose metabolism (https://www.proteinatlas.org). Knockout studies in mice show that the absence of *RORA* leads to severe phenotypes, including preweaning lethality, underscoring its importance in early development (https://www.mousephenotype.org). *TLN2* encodes a protein related to talin 1, which is crucial for actin filament assembly and cell migration, processes vital for tissue development and repair (https://www.proteinatlas.org). This gene is particularly important in cell types such as fibroblasts and osteoclasts, suggesting a role in skeletal growth and maintenance. *C2CD4B* is involved in the regulation of acute inflammatory responses and cell adhesion, both of which are essential for normal growth and immune function (https://www.proteinatlas.org). In knockout mice, this gene is associated with decreased haemoglobin content and cholesterol levels, indicating its potential impact on blood parameters and metabolic health (https://www.mousephenotype.org). *VPS13C* is necessary for proper mitochondrial function and lipid transport, with knockout mice exhibiting altered levels of circulating amylase, lipase, and increased red blood cell distribution width (https://www.mousephenotype.org). This gene's role in mitochondrial maintenance and lipid metabolism suggests it may influence energy balance and cellular health during growth.

Positional candidate genes for $BWG_{0d\text{-}18m}$ on chromosome 4 suggest a direct influence on muscle development but also on signal transduction and cellular stability, all of which are critical for growth and development. In detail, *CAPZA2* encodes the capping actin protein of muscle Z-line alpha subunit 2, which is crucial for actin filament regulation. Given its involvement in GO term "Regulation of actin filament organisation", *CAPZA2* may contribute to muscle cell integrity and stability, both of which are essential for skeletal muscle function and growth. Actin cytoskeleton remodelling is a key component of muscle fibre development, and its disruption can impair muscle mass accumulation, potentially affecting BW gain. Further, *MET*, a proto-oncogene encoding a receptor tyrosine kinase, plays a pivotal role in cellular growth, differentiation, and migration. It is a key component of hepatocyte growth factor (HGF) signalling, which influences skeletal muscle development, tissue regeneration, and metabolic regulation (https://www.proteinatlas.org). *CAV1* and *CAV2*, two members of the caveolin family, are integral membrane proteins involved in caveolae formation, lipid homeostasis, and signal transduction. The involvement of these genes in metabolic regulation suggests a link between lipid metabolism and BW gain, potentially affecting energy utilisation and storage.

Positional candidate genes identified for $BWG_{3w\text{-}18m}$ on chromosome 7 can be linked to muscle development, extracellular matrix organisation, and cellular signalling, all of which are crucial for growth and BW regulation. Among these, *ACTL9* is involved in actin filament formation, a process fundamental for muscle growth and maintenance. *MYO1F*, a member of the myosin superfamily, plays a role in intracellular transport and cytoskeletal organisation, which are essential for muscle function and cell mobility. *ADAMTS10* is involved in extracellular matrix remodelling, influencing connective tissue integrity and skeletal growth. Disruptions in *ADAMTS10* have been linked to abnormal musculoskeletal development in humans (Mularczyk et al., 2018) suggesting its role in long-term BW gain. With respect to *OR2Z1*, an olfactory receptor gene, its functional role in growth traits remains unclear. Olfactory receptors have been implicated in metabolic regulation and energy balance.

Interestingly, the three top variants identified for birth weight on chromosome 5 in this study were located near the *MGST1* locus (top variant rs211210569 at 93 516 066 bp) which was associated with milk fat content in DSN cows (Korkuć et al., 2023). Despite the proximity of the *MGST1* locus to the identified birth weight loci, linkage disequilibrium analysis revealed only low LD between the top variant of *MGST1* and the top variants on chromosome 5 from this study ($r2 \leq 0.07$, D'≤0.68). Further analysis of haplotype frequencies for the four top variants (92 970 239 bp – 93 516 066 bp – 99 255 711 bp – 101 741 125 bp) revealed that the haplotype carrying favourable alleles for both high birth weight and high milk fat content (C-C-G-T) occurred with a frequency of 25.5%. This was followed by the haplotype associated with low birth weight and high milk fat content (T-C-A-C) at 17.4%, and the haplotype for high birth weight and low milk fat content (C-T-G-T) at 16.2%. The findings suggest that while the *MGST1* locus is closely associated with traits related to milk fat content, the loci identified for birth weight on chromosome 5 appear to operate independently, as indicated by the low LD between them. The distinct haplotype patterns observed imply that different genetic mechanisms may be driving the regulation of birth weight and the production of fat in the milk later in life.

Some of the loci identified for DSN in this study were already identified in GWAS with meat and production traits of diverse dairy and beef cattle breeds, but also with milk production traits. For birth weight, the locus on chromosome 5 overlapped with loci for "Length of productive life" in Holstein cattle (Cole et al., 2011), "Marbling score" in Korean Hanwoo cattle (Li et al., 2016), "Connective tissue amount" in Angus (Leal-Gutiérrez et al., 2020), but also "Milk yield", "Milk fat yield", "Milk fat percentage", "Milk fat percentage" and "Milk protein percentage" in Holstein (Pedrosa et al., 2021; Wang et al., 2019; Jiang et al., 2019; Ning et al., 2018; van den Berg et al., 2016; Nayeri et al., 2016; Cole et al., 2011) and Brown Swiss (Frischknecht et al., 2017). The QTL on chromosome 18 overlapped with loci for "Average daily gain" in diverse breeds (Akanno et al., 2018), and "Tenderness score" and "Connective tissue amount" in Angus (Leal-Gutiérrez et al., 2020). For BW_{3w}, the QTL on chromosome 3 overlapped as well with a locus for "Length of productive life" in Holstein cattle (Cole et al., 2011). The QTL on chromosome 16 overlapped with loci for "BW" and "BW gain" in diverse breeds (Snelling et al., 2010), and "Marbling score", "Shear force", and "Connective tissue amount" in Angus (Leal-Gutiérrez et al., 2020). For BW_{7m}, the QTL on chromosome 3 overlapped with loci for "Shear force" in Angus cattle (Leal-Gutiérrez et al., 2020), and "Milk fat percentage" and "Milk protein percentage" in Holstein (Pedrosa et al., 2021; Cochran et al., 2013). The QTL on chromosome 10 which was identified not only for BW_{7m} but also $BWG_{0d\text{-}7m}$ overlapped with loci for "BW" and "BW gain" in diverse breeds including Simmental cattle (Zhuang et al., 2020; Snelling et al., 2010), "Marbling score" in Angus (Leal-Gutiérrez et al., 2020), but also with loci for "Milk fat yield", "Milk protein yield" and "Milk protein percentage" in Holstein (Pedrosa et al., 2021; Marete et al., 2018). The overlap between loci identified in this study and those previously reported in GWAS for various meat, production, and milk traits across diverse cattle breeds highlights the consistency and relevance of these genetic markers across different populations and traits. Additionally, the identification of these loci across various breeds emphasises their importance in cattle genetics and their potential utility as targets for marker-assisted or genomic selection. The full list of overlapping GWAS including the investigated trait, cattle

breed, and publication information is provided in Supplementary Table S7.

This study was conducted using data from the largest DSN farm, which is also the only DSN farm fattening for milk and beef production, rearing approximately 350 fattening bulls. While using a single farm may raise concerns about the generalisability of the findings, it is important to consider the structure of the DSN population. As an endangered breed, DSN has a low number of well-selected breeding bulls, which are widely used across multiple farms. This results in a high degree of genetic connectedness among DSN animals within the whole DSN population, reducing the likelihood that our findings are farm-specific. Although small populations are more susceptible to genetic drift and higher mutation rates, the use of breeding bulls helps to ensure genetic improvement, maintains genetic consistency and diversity across farms, and minimises farm-specific divergence. While environmental factors unique to the sampled farm cannot be excluded, specific undetected gene-environment-effects may exist. Nevertheless, the genetic basis of the observed traits is expected to be representative of the entire DSN population. On the other hand, as our study includes only males from a single farm, the scope and applicability of the findings must be carefully considered, particularly with respect to yet unknown interactions with milk production in cows. Future studies conducted across farms would help validate these results and provide a more comprehensive understanding of genetic variation in DSN.

When breeding for both high milk and high beef traits, it is important to consider and exclude loci with opposing effects on the two traits, where a variant improves one trait while the same or a genetically linked variant may negatively impact another trait, potentially conflicting with breeding goals. A comparison of the effects of the loci for birth weight, BW and BW gain from this study with loci previously identified for milk production (Korkuć et al., 2023) revealed that only a QTL region on chromosome 5 affected both birth weight and milk fat content. A closer look at the top variants for both traits of that QTL region showed only minimal LD ($r^2 < 0.07$) suggesting that they are very unlikely to be inherited together. Although some associations were found on the same chromosomes, the distances between the loci were too large for significant linkage. For instance, the QTL for BW_{7m} on chromosome 3 at 26.6 Mb was located 12.7 Mb downstream of the milk protein content QTL at 13.9 Mb ($r2$ = 0.01, D'=0.09). Similarly, the QTL for BW_{7m} and $BWG_{0d\text{-}7m}$ on chromosome 10 at 49.4 Mb was 4.7 Mb downstream of the milk protein content QTL at 44.7 Mb ($r2$ = 0.07, D'=0.42). As a result, most of the significant variants contributing to variation in the dual-purpose DSN cattle can be used independently to select for either high BW or high milk yield and content, thereby enhancing the economic value of the animals. Further, while the lack of overlapping QTLs suggests largely independent genetic regulation, this does not exclude the possibility of interactions between BW and milk traits through small-effect loci that have not been discovered in this study, pleiotropic effects or shared biological pathways. Caution is also needed when selecting for high birth weight, as larger calves can lead to more difficult births, even though they may be more resilient during the critical postnatal period.

The specific goal of this study was to identify genetic loci contributing to growth, which are indirect measures for an effective fattening period of young bulls. Higher growth rates ultimately affect feed efficiency and meat production indirectly. This study builds upon our previous research on milk loci in DSN. The current breeding goals in DSN are the conservation of its genetic diversity and the genetic improvement of the breed for milk yield while simultaneously maintaining the dual-purpose characteristics by ensuring muscularity. The variants identified in this study can therefore be used to support the current selection of advantageous breeding bulls and guide mating decisions to maintain or even improve the beef characteristics of DSN, while considering improvement of milk production.

Conclusions

To preserve the dual-purpose characteristics of the DSN breed, it is essential to consider both milk and meat performance in breeding strategies. In this study, we identified 14 QTLs associated with BW or BW gain, which are key traits for beef production. It is crucial to avoid any negative impact on milk production traits during this process. Our analysis revealed that most identified QTLs were not linked to milk production QTLs, with only minimal linkage observed between a QTL on chromosome 5 and previously identified *MGST1*-associated loci for milk fat content.

A key finding is that several positional candidate genes directly influence muscle tissue development and fat metabolism, emphasising their role in skeletal growth and body composition. Notably, *RERGL* and *LMO3* regulate muscle and fat tissue development, with *LMO3* implicated in adipose tissue reprogramming and nutrient homeostasis. *MET* and *CAPZA2* are involved in actin filament organisation, muscle cell integrity, and skeletal muscle function. Additionally, *TLN2*, *ADAMTS10*, and *MYO1F,* which are crucial for muscle fibre assembly, extracellular matrix remodelling, and intracellular transport, further reinforce the role of cytoskeletal integrity and tissue maintenance in long-term BW gain. Other candidate genes, such as *KLRC1*, *IGSF3*, and members of the C-type lectin family, are involved in immune response suggesting a potential interplay between growth and health. Furthermore, *GLIS1*, directly influences tissue formation and growth, while genes like *DPY19L3*, *SLC7A9*, regulators of G protein signalling, and *RORA* may affect growth and development through their roles in metabolic processes, cellular signalling, and nutrient transport. The importance of these identified loci is further supported by their identification in GWAS studies of other breeds and traits. The number of candidate genes discussed in our study is still high. Further reduction of genes that may causally affect growth in DSN requires an enlarged sample size which remains a challenge in endangered populations.

Despite these challenges, the significant variants identified in this study could be used to enhance the economic value of DSN cattle in the market. However, due to the limited population size, genomic breeding values are not yet available for DSN. As an alternative, genetic scores based on the cumulative effects of all significant loci could be calculated to aid breeders in making informed selection decisions.

Supplementary material

Supplementary Material for this article (https://doi.org/10.1016/j.animal.2025.101540) can be found at the foot of the online page, in the Appendix section.

Ethics approval

Not applicable.

Data and model availability statement

The dataset generated and analysed during the current study is publicly available in the European Nucleotide Archive (Project number: PRJEB82246, https://www.ebi.ac.uk/eva/?eva-study=PRJEB82246). Information can be made available from the authors upon request.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT to check the grammar and spelling of the text. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

Author ORCIDs

Paula Korkuć: https://orcid.org/0000-0002-3772-9562.
Monika Reißmann: https://orcid.org/0000-0003-3418-6344.
Gudrun A. Brockmann: https://orcid.org/0000-0002-4387-2947.

CRediT authorship contribution statement

P. Korkuć: Writing – review & editing, Writing – original draft, Visualisation, Methodology, Investigation, Formal analysis, Data curation, Conceptualisation. **M. Reißmann:** Writing – review & editing, Investigation. **G.A. Brockmann:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualisation.

Declaration of interest

None.

Acknowledgements

The RBB Rinderproduktion Berlin-Brandenburg GmbH and DSN farms supported this project with their expertise in animal selection, collecting ear tags, and providing weight data. We acknowledge support by the Open Access Publication Fund of Humboldt-Universität zu Berlin.

Financial support statement

The project was supported by funds from the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the Federal Program for Ecological Farming and Other Forms of Sustainable Agriculture (Funding number: 2815NA010).

References

Akanno, E.C., Chen, L., Abo-Ismail, M.K., Crowley, J.J., Wang, Z., Li, C., Basarab, J.A., MacNeil, M.D., Plastow, G.S., 2018. Genome-wide association scan for heterotic quantitative trait loci in multi-breed and crossbred beef cattle. Genetics Selection Evolution 50, 48. https://doi.org/10.1186/s12711-018-0405-y.

Browning, B.L., Zhou, Y., Browning, S.R., 2018. A one-penny imputed genome from next-generation reference panels. The American Journal of Human Genetics 103, 338–348. https://doi.org/10.1016/j.ajhg.2018.07.015.

Bundesanstalt für Landwirtschaft und Ernährung (BLE), 2023. Einheimische Nutztierrassen in Deutschland und Rote Liste gefährdeter Nutztierrassen 2023. BLE, Bonn, Germany.

Castro-Mondragon, J.A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Manosalva Pérez, N., Fornes, O., Leung, T.Y., Aguirre, A., Hammal, F., Schmelter, D., Baranasic, D., Ballester, B., Sandelin, A., Lenhard, B., Vandepoele, K., Wasserman, W.W., Parcy, F., Mathelier, A., 2022. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Research 50, D165–D173. https://doi.org/10.1093/nar/gkab1113.

Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4, 7. https://doi.org/10.1186/s13742-015-0047-8.

Cochran, S.D., Cole, J.B., Null, D.J., Hansen, P.J., 2013. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. BMC Genetics 14, 49. https://doi.org/10.1186/1471-2156-14-49.

Cole, J.B., Wiggans, G.R., Ma, L., Sonstegard, T.S., Lawlor, T.J., Crooker, B.A., Van Tassell, C.P., Yang, J., Wang, S., Matukumalli, L.K., Da, Y., 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. BMC Genomics 12, 408. https://doi.org/10.1186/1471-2164-12-408.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 2011. The variant call format and VCFtools. Bioinformatics 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Dupuis, J., Siegmund, D., 1999. Statistical methods for mapping quantitative trait loci from a dense set of markers. Genetics 151, 373–386. https://doi.org/10.1093/genetics/151.1.373.

Fantini, D., 2019. EasyPubMed: Search and Retrieve Scientific Publication Records from PubMed. R package version 2.13. Published 29 March 2019 on https://doi.org/10.32614/CRAN.package.easyPubMed.

Frischknecht, M., Pausch, H., Bapst, B., Signer-Hasler, H., Flury, C., Garrick, D., Stricker, C., Fries, R., Gredler-Grandl, B., 2017. Highly accurate sequence imputation enables precise QTL mapping in Brown Swiss cattle. BMC Genomics 18, 999. https://doi.org/10.1186/s12864-017-4390-2.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M., Snell, R., 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Research 12, 222–231. https://doi.org/10.1101/gr.224202.

Howe, K.L., Achuthan, P., Allen, James, Allen, Jamie, Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., El Houdaigui, B., Fatima, R., Gall, A., Garcia Giron, C., Grego, T., Guijarro-Clarke, C., Haggerty, L., Hemrom, A., Hourlier, T., Izuogu, O.G., Juettemann, T., Kaikala, V., Kay, M., Lavidas, I., Le, T., Lemos, D., Gonzalez Martinez, J., Marugán, J.C., Maurel, T., McMahon, A.C., Mohanan, S., Moore, B., Muffato, M., Oheh, D.N., Paraschas, D., Parker, A., Parton, A., Prosovetskaia, I., Sakthivel, M.P., Salam, A.I.A., Schmitt, B. M., Schuilenburg, H., Sheppard, D., Steed, E., Szpak, M., Szuba, M., Taylor, K., Thormann, A., Threadgold, G., Walts, B., Winterbottom, A., Chakiachvili, M., Chaubal, A., De Silva, N., Flint, B., Frankish, A., Hunt, S.E., IIsley, G.R., Langridge, N., Loveland, J.E., Martin, F.J., Mudge, J.M., Morales, J., Perry, E., Ruffier, M., Tate, J., Thybert, D., Trevanion, S.J., Cunningham, F., Yates, A.D., Zerbino, D.R., Flicek, P., 2021. Ensembl 2021. *Nucleic Acids Research* 49, D884–D891. https://doi.org/10.1093/nar/gkaa942.

Hu, Z.-L., Park, C.A., Reecy, J.M., 2022. Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. Nucleic Acids Research 50, D956–D961. https://doi.org/10.1093/nar/gkab1116.

Jiang, J., Ma, L., Prakapenka, D., VanRaden, P.M., Cole, J.B., Da, Y., 2019. A large-scale genome-wide association study in U.S. Holstein cattle. Frontiers in Genetics 10, 412. https://doi.org/10.3389/fgene.2019.00412.

Kassambara, A., 2020. ggpubr: "ggplot2" Based Publication Ready Plots. R package version 0.6.0. Accessed 10 February 2023 on https://doi.org/10.32614/CRAN.package.ggpubr.

Korkuć, P., Arends, D., May, K., König, S., Brockmann, G.A., 2021. Genomic loci affecting milk production in German Black Pied cattle (DSN). Frontiers in Genetics 12, 640039. https://doi.org/10.3389/fgene.2021.640039.

Korkuć, P., Neumann, G.B., Hesse, D., Arends, D., Reißmann, M., Rahmatalla, S., May, K., Wolf, M.J., König, S., Brockmann, G.A., 2023. Whole-genome sequencing data reveal new loci affecting milk production in German Black Pied cattle (DSN). Genes 14, 581. https://doi.org/10.3390/genes14030581.

Kumar, P., Henikoff, S., Ng, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature Protocols 4, 1073–1081. https://doi.org/10.1038/nprot.2009.86.

Leal-Gutiérrez, J.D., Rezende, F.M., Reecy, J.M., Kramer, L.M., Peñagaricano, F., Mateescu, R.G., 2020. Whole genome sequence data provides novel insights into the genetic architecture of meat quality traits in beef. Frontiers in Genetics 11, 538640. https://doi.org/10.3389/fgene.2020.538640.

Li, Y., Gao, Y., Kim, Y.-S., Iqbal, A., Kim, J.-J., 2016. A whole genome association study to detect additive and dominant single nucleotide polymorphisms for growth and carcass traits in Korean native cattle, Hanwoo. Asian-Australasian Journal of Animal Sciences 30, 8–19. https://doi.org/10.5713/ajas.16.0170.

Maekawa, M., Yamaguchi, K., Nakamura, T., Shibukawa, R., Kodanaka, I., Ichisaka, T., Kawamura, Y., Mochizuki, H., Goshima, N., Yamanaka, S., 2011. Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. Nature 474, 225–229. https://doi.org/10.1038/nature10106.

Mancin, E., Tuliozi, B., Pegolo, S., Sartori, C., Mantovani, R., 2022. Genome wide association study of beef traits in local alpine breed reveals the diversity of the pathways involved and the role of time stratification. Frontiers in Genetics 12, 746665. https://doi.org/10.3389/fgene.2021.746665.

Mandel, A., Reißmann, M., Brockmann, G.A., Korkuć, P., 2025. Whole-genome insights into the genetic basis of conformation traits in German Black Pied (DSN) cattle. Genes 16, 445. https://doi.org/10.3390/genes16040445.

Marete, A.G., Guldbrandtsen, B., Lund, M.S., Fritz, S., Sahana, G., Boichard, D., 2018. A meta-analysis including pre-selected sequence variants associated with seven traits in three french dairy cattle populations. Frontiers in Genetics 9, 399480. https://doi.org/10.3389/FGENE.2018.00522/BIBTEX.

May, K., Scheper, C., Brügemann, K., Yin, T., Strube, C., Korkuć, P., Brockmann, G.A., König, S., 2019. Genome-wide associations and functional gene analyses for endoparasite resistance in an endangered population of native German Black

Pied cattle. BMC Genomics 20, 277. https://doi.org/10.1186/s12864-019-5659-4.
Mazerolle, M.J., 2023. Model selection and multimodel inference using the AICcmodavg package. R package version 2.3-3. Published 16 November 2023 on https://doi.org/10.32614/CRAN.package.AICcmodavg.
McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S.S., Thormann, A., Flicek, P., Cunningham, F., 2016. The ensembl variant effect predictor. Genome Biology 17, 122. https://doi.org/10.1186/s13059-016-0974-4.
Meier, S., Arends, D., Korkuć, P., Neumann, G.B., Brockmann, G.A., 2020. A genome-wide association study for clinical mastitis in the dual-purpose German Black Pied cattle breed. Journal of Dairy Science 103, 10289–10298. https://doi.org/10.3168/jds.2020-18209.
Mularczyk, E.J., Singh, M., Godwin, A.R.F., Galli, F., Humphreys, N., Adamson, A.D., Mironov, A., Cain, S.A., Sengle, G., Boot-Handford, R.P., Cossu, G., Kielty, C.M., Baldock, C., 2018. ADAMTS10-mediated tissue disruption in Weill–Marchesani syndrome. Human Molecular Genetics 27, 3675–3687. https://doi.org/10.1093/hmg/ddy276.
Nayeri, S., Sargolzaei, M., Abo-Ismail, M.K., May, N., Miller, S.P., Schenkel, F., Moore, S.S., Stothard, P., 2016. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. BMC Genetics 17, 75. https://doi.org/10.1186/s12863-016-0386-1.
Neumann, G.B., Korkuć, P., Arends, D., Wolf, M.J., May, K., Reißmann, M., Elzaki, S., König, S., Brockmann, G.A., 2021. Design and performance of a bovine 200 k SNP chip developed for endangered German Black Pied cattle (DSN). BMC Genomics 22, 905. https://doi.org/10.1186/s12864-021-08237-2.
Neumann, G.B., Korkuć, P., Arends, D., Wolf, M.J., May, K., König, S., Brockmann, G.A., 2023. Genomic diversity and relationship analyses of endangered German Black Pied cattle (DSN) to 68 other taurine breeds based on whole-genome sequencing. Frontiers in Genetics 13, 3532. https://doi.org/10.3389/fgene.2022.993959.
Ning, C., Wang, D., Zheng, X., Zhang, Q., Zhang, S., Mrode, R., Liu, J.-F., 2018. Eigen decomposition expedites longitudinal genome-wide association studies for milk production traits in Chinese Holstein. Genetics Selection Evolution 50, 12. https://doi.org/10.1186/s12711-018-0383-0.
Pausch, H., Emmerling, R., Gredler-Grandl, B., Fries, R., Daetwyler, H.D., Goddard, M. E., 2017. Meta-analysis of sequence-based association studies across three cattle breeds reveals 25 QTL for fat and protein percentages in milk at nucleotide resolution. BMC Genomics 18, 853. https://doi.org/10.1186/S12864-017-4263-8.
Pedrosa, V.B., Schenkel, F.S., Chen, S.-Y., Oliveira, H.R., Casey, T.M., Melka, M.G., Brito, L.F., 2021. Genomewide association analyses of lactation persistency and milk production traits in Holstein cattle based on imputed whole-genome sequence data. Genes 12, 1830. https://doi.org/10.3390/genes12111830.
Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., Vilo, J., 2019. g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Research 47, W191–W198. https://doi.org/10.1093/NAR/GKZ369.
Rosen, B.D., Bickhart, D.M., Schnabel, R.D., Koren, S., Elsik, C.G., Zimin, A., Dreischer, C., Schultheiss, S., Hall, R., Schroeder, S.G., Van Tassell, C.P., Smith, T.P.L., Medrano, J.F., 2018. Modernizing the bovine reference genome assembly. Proceedings of the 11th World Congress on Genetics Applied to Livestock Production, 11–16 February 2018, Auckland, New Zealand, p. 802.
Slowikowski, K., 2021. ggrepel: Automatically position non-overlapping text labels with 'ggplot2'. R package version 0.9.1. Published 15 January 2021 on https://doi.org/10.32614/CRAN.package.ggrepel.
Snelling, W.M., Allan, M.F., Keele, J.W., Kuehn, L.A., McDaneld, T., Smith, T.P.L., Sonstegard, T.S., Thallman, R.M., Bennett, G.L., 2010. Genome-wide association study of growth in crossbred beef cattle. Journal of Animal Science 88, 837–848. https://doi.org/10.2527/jas.2009-2257.
Tan, G., Lenhard, B., 2016. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. Bioinformatics 32, 1555–1556. https://doi.org/10.1093/bioinformatics/btw024.
Tribout, T., Croiseau, P., Lefebvre, R., Barbat, A., Boussaha, M., Fritz, S., Boichard, D., Hoze, C., Sanchez, M.P., 2020. Confirmed effects of candidate variants for milk production, udder health, and udder morphology in dairy cattle. Genetics Selection Evolution 52, 1–26. https://doi.org/10.1186/S12711-020-00575-1/FIGURES/5.
van den Berg, I., Boichard, D., Lund, M.S., 2016. Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. Journal of Dairy Science 99, 8932–8945. https://doi.org/10.3168/jds.2016-11073.
Wagner, G., Fenzl, A., Lindroos-Christensen, J., Einwallner, E., Husa, J., Witzeneder, N., Rauscher, S., Gröger, M., Derdak, S., Mohr, T., Sutterlüty, H., Klinglmüller, F., Wolkerstorfer, S., Fondi, M., Hoermann, G., Cao, L., Wagner, O., Kiefer, F.W., Esterbauer, H., Bilban, M., 2021. LMO3 reprograms visceral adipocyte metabolism during obesity. Journal of Molecular Medicine 99, 1151–1171. https://doi.org/10.1007/S00109-021-02089-9/FIGURES/10.
Wang, D., Ning, C., Liu, J.-F.-F., Zhang, Q., Jiang, L., 2019. Short communication: Replication of genome-wide association studies for milk production traits in Chinese Holstein by an efficient rotated linear mixed model. Journal of Dairy Science 102, 2378–2383. https://doi.org/10.3168/jds.2018-15298.
Warnes, G., Gorjanc, G., Leisch, F., Man, M., 2022. genetics: Population Genetics. R package version 1.3.8.1.3. Published 1 March 2021 on https://doi.org/10.32614/CRAN.package.genetics.
Weller, J.I., Bickhart, D.M., Wiggans, G.R., Tooker, M.E., O'Connell, J.R., Jiang, J., Ron, M., VanRaden, P.M., 2018. Determination of quantitative trait nucleotides by concordance analysis between quantitative trait loci and marker genotypes of US Holsteins. Journal of Dairy Science 101, 9089–9107. https://doi.org/10.3168/JDS.2018-14816.
Wickham, H., 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York, NY, USA.
Wilkins, D., 2020. gggenes: Draw gene arrow maps in 'ggplot2'. R package version 0.5.1. Accessed 5 September 2021 on https://doi.org/10.32614/CRAN.package.gggenes.
Wolf, M.J., Yin, T., Neumann, G.B., Korkuć, P., Brockmann, G.A., König, S., May, K., 2021. Genome-wide association study using whole-genome sequence data for fertility, health indicator, and endoparasite infection traits in German Black Pied cattle. Genes 12, 1163. https://doi.org/10.3390/genes12081163.
Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011. GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics 88, 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.
Zhang, J., Kadri, N.K., Mullaart, E., Spelman, R., Fritz, S., Boichard, D., Charlier, C., Georges, M., Druet, T., 2020. Genetic architecture of individual variation in recombination rate on the X chromosome in cattle. Heredity 125, 304–316. https://doi.org/10.1038/s41437-020-0341-9.
Zhuang, Z., Xu, L., Yang, J., Gao, H., Zhang, L., Gao, X., Li, J., Zhu, B., 2020. Weighted single-step genome-wide association study for growth traits in Chinese Simmental beef cattle. Genes 11, 189. https://doi.org/10.3390/genes11020189.