

RESEARCH ARTICLE

A system-based analysis of the genetic determinism of udder conformation and health phenotypes across three French dairy cattle breeds

Andrew Marete^{1,2}*, Mogens Sandø Lund^{2†}, Didier Boichard^{1‡}, Yuliaxis Ramayo-Caldas¹

1 Génétique Animale et Biologie Intégrative, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France, **2** Center for Quantitative Genetics and Genomics, Aarhus University, Tjele, Denmark

✉ These authors contributed equally to this work.

‡ These authors also contributed equally to this work.

* andrew.marete@mbg.au.dk



OPEN ACCESS

Citation: Marete A, Lund MS, Boichard D, Ramayo-Caldas Y (2018) A system-based analysis of the genetic determinism of udder conformation and health phenotypes across three French dairy cattle breeds. PLoS ONE 13(7): e0199931. <https://doi.org/10.1371/journal.pone.0199931>

Editor: Juan J Loor, University of Illinois, UNITED STATES

Received: February 6, 2018

Accepted: June 15, 2018

Published: July 2, 2018

Copyright: © 2018 Marete et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: A. Marete is the recipient of an Erasmus Mundus Ph.D. grant. This work is funded by the European Graduate School in Animal Breeding and Genetics (France), Viking Genetics (Denmark) and the Center for Genomic Selection in Animals and Plants (GenSAP) funded by Innovation Fund Denmark (grant 0603-00519B). The funders had no role in study design, data collection and

Abstract

Using GWAS to identify candidate genes associated with cattle morphology traits at a functional level is challenging. The main difficulty of identifying candidate genes and gene interactions associated with such complex traits is the long-range linkage disequilibrium (LD) phenomenon reported widely in dairy cattle. Systems biology approaches, such as combining the Association Weight Matrix (AWM) with a Partial Correlation in an Information Theory (PCIT) algorithm, can assist in overcoming this LD. Used in a multi-breed and multi-phenotype context, the AWM-PCIT could aid in identifying udder traits candidate genes and gene networks with regulatory and functional significance. This study aims to use the AWM-PCIT algorithm as a post-GWAS analysis tool with the goal of identifying candidate genes underlying udder morphology. We used data from 78,440 dairy cows from three breeds and with own phenotypes for five udder morphology traits, five production traits, somatic cell score and clinical mastitis. Cows were genotyped with medium (50k) or low-density (7 to 10k) chips and imputed to 50k. We performed a within breed and trait GWAS. The GWAS showed 9,830 significant SNP across the genome ($p < 0.05$). Five thousand and ten SNP did not map a gene, and 4,820 SNP were within 10-kb of a gene. After accounting for 1SNP:1gene, 3,651 SNP were within 10-kb of a gene (set1), and 2,673 significant SNP were further than 10-kb of a gene (set2). The two SNP sets formed 6,324 SNP matrix, which was fitted in an AWM-PCIT considering udder depth/ development as the key trait resulting in 1,013 genes associated with udder morphology, mastitis and production phenotypes. The AWM-PCIT detected ten potential candidate genes for udder related traits: *ESR1*, *FGF2*, *FGFR2*, *GLI2*, *IQGAP3*, *PGR*, *PRLR*, *RREB1*, *BTRC*, and *TGFBR2*.

analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors, funding body, including Viking Genetics, declare that they have no competing interests. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

Introduction

Genetic architecture of complex phenotypes in cattle includes many loci affecting a given trait [1]. Most of these loci have small effects, but few segregating loci have moderate-to-large effects possibly due to epistatic effects, varying selection goals or recent selection for the favorable mutant allele. Moreover, markers collectively capture most but not all additive genetic variance for phenotypes. The incomplete variance capture may be due to causal mutations with low allele frequencies and therefore in incomplete linkage disequilibrium (LD) with markers [2]. To reduce this LD, we can either do a between breed analysis with a large sample of genotyped cows [3] or combine the results of a within breed GWAS in a multi-breed context. This is possible because the cost of genotyping is decreasing thus allowing many breeds, including those of medium population size, to be genotyped rapidly primarily for genomic selection purpose. These large populations of genotyped cows with own performances allows us to: (1) detect QTL for newly recorded traits or traits previously not studied; (2) carry out large confirmation studies for conventional traits. Previous studies have demonstrated that polymorphic sites that segregate within and across bovine populations can be studied using imputed low-to-dense genotypes [4,5]. Such genotypes have been used in model organisms and dairy cattle leading to the identification of candidate causal variants or closely neighboring variants that control complex phenotypes [6,7]. These studies have been useful for identifying QTL regions and probable genes associated with a phenotype. So far, however, there have been few validation studies of the vast number of putative variants across and between breeds and amongst multiple phenotypes. This study uses 50k SNP data to validate such variants using the Association Weight Matrix (AWM) [8] approach as a post GWAS analysis tool. The AWM is a systems biology approach for the genetic dissection of complex traits based on applying gene network theory to the results from GWAS. Hence, if the AWM SNP matrix is used in combination with a Partial Correlation (PC) in an Information Theory (IT) framework, and for correlated phenotypes, then it is possible to generate gene networks with regulatory and functional significance for udder related phenotypes.

Despite the limitations of the chip density, previous studies have shown the usefulness of the AWM to identify candidate genes in cattle, e.g. [9–10] and corroborated across different species, e.g. [11–12] in independent studies using the 50k marker density.

In this study, we report results based on GWAS analysis for mammary conformation, milk production and health phenotypes for 78,440 dairy cows. A multi-step validation by combining the results of single SNP, single phenotype, in a multiple-breed context using the AWM-P-CIT algorithm was performed. The aim was to identify the genes associated with mammary conformation and health phenotypes in Holstein, Montbeliarde, and Normande breeds, accounting for milk production as supportive traits. We further explored gene networks with the main gene ontology domains including biological processes, cellular component, and molecular functions.

Material and methods

Phenotypes

The cow sample was comprised of 46,732 Holstein, 20,141 Montbeliarde, and 11,965 Normande all with known parents. Phenotypes were yield deviations as produced by French national evaluation system [13]. A yield deviation is a performance adjusted for all non-genetic effects of the model [14]. In case of repeated records, a yield deviation is adjusted for the permanent environmental effect and averaged per animal. The 12 traits were: fore udder attachment (FUA), udder depth or development (UDD), udder cleft (UC), udder balance (UB),

Table 1. Characteristics of udder conformation, production and health traits in three French dairy breeds.

Trait	Cows with traits			Standard Deviation		
	MON ¹	NOR ²	HOL ³	MON ¹	NOR ²	HOL ³
Fore Udder Attachment (FUA)	17,330	7,671	32,491	1.18	1.26	1.19
Udder Depth or Development (UDD)	17,330	7,671	32,491	0.98	0.81	0.84
Udder Cleft (UC)	17,330	7,671	32,491	1.26	1.28	0.85
Udder Balance (UB)	17,330	7,671	32,491	0.79	0.82	1.39
Front Teat Placement (FTP)	17,330	7,671	32,491	1.11	1.14	0.96
Milk Yield (MY)	20,096	11,944	46,732	786	742	991
Fat Yield (FAT)	20,096	11,944	46,732	33.7	34.7	40.2
Protein Yield (PROT)	20,096	11,944	46,732	27.5	25.7	30.1
Fat Percent (FAT %)	20,096	11,944	46,732	1.26	1.28	0.85
Protein Percent (PROT %)	20,096	11,944	46,732	1.28	1.28	0.85
Clinical Mastitis (CM)	13,879	9,013	32,491	0.24	0.26	0.36
Somatic Cell Score (SCS)	20,141	11,965	46,732	0.93	0.90	0.96

¹Montbeliarde;

²Normande;

³Holstein

<https://doi.org/10.1371/journal.pone.0199931.t001>

front teat placement (FTP), milk (MY), fat (FAT and FAT%), protein (PROT and PROT%) clinical mastitis (CM) and somatic cell score (SCS). SCS are derived from Somatic Cell Counts (SCC). SCC are obtained at each monthly test-day, and SCS are defined as $SCS = 3 + \log_2(SCC/100,000)$. CM events are declared by the farmer and recorded by the technician at each test-day. The analysis trait is defined by 1 if the cow has at least one clinical mastitis event in the lactation, and 0 if there is no event. SCS and CM are repeated over lactations. The model for obtaining yield deviations was different according to the trait and all models included the genetic additive effect. The model for SCS and CM (recorded in the first three lactations), included the effect of herd x year, age at calving x parity x year, month of calving x parity x year, preceding days dry (for later parities), and a permanent environmental effect. All models assumed heterogeneous variances depending on herd-year. Each cow had one record for each trait. UDD can be either udder depth (Holstein and Normande) or udder development (Montbeliarde), but treated as the same trait because they have similar definitions. Depending on the trait, the number of animals with phenotypes ranged from 7,671 to 11,965 in Normande, 13,879 to 20,141 in Montbeliarde, and 32,491 to 46,732 in Holstein (Table 1). We estimated variance components by fitting a multiple trait REML animal model as implemented in the DMU software [15]. This study is based on already existing data hence we did not require ethical approval.

Genotyping, quality control, and imputation

All cows were genotyped using Illumina BovineSNP50 BeadChip (50k) or Illumina BovineLD v.2 BeadChip. We used the UMD3.1 assembly of the bovine genome [16] for SNP chromosomal positions. We did not consider mitochondrial, X-chromosomal and Y-chromosomal SNP, as well as unmapped SNP for further analyses. We examined 43,800 SNP currently used in French genomic evaluation procedure [13]. Selection criteria was: call rate higher than 99%, minor allele frequency (MAF) higher than 2% in at least one of the three breeds, lack of Hardy–Weinberg Equilibrium ($P < 10^{-4}$), technical quality assessed by their very low rate of Mendelian mismatch between parents and progeny and known position of genome assembly.

We imputed cows genotyped from BovineLD v.2 BeadChip to 50k to obtain a genotype without missing information. This step was performed within breed in the conventional pipeline for genomic selection [13]. We imputed using FImpute software [17], and reference population included all 50k genotyped male and female animals per breed. Imputation error rate, measured in routine evaluation situation (and not in this study), varied from 0.2 to 2.5% depending on whether parents were genotyped. Across breeds, best imputation accuracies were observed in Montbeliarde which has the highest proportion of sires and dams genotyped with 50k, while accuracies were lower in Holstein, a breed with a smaller portion of genotyped dams, a lower percentage of 50k genotyped parents, and even a non-zero proportion of non-genotyped (usually foreign) sires.

Statistical framework

GWAS was done within breed with the Mixed Linear Model Association (MLMA) method as implemented in GCTA [18]. Because phenotypes were yield deviations already adjusted for non-genetic effects, we did not consider additional fixed effects. For each phenotype and SNP i , the model used in each breed was the following

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{u} + \mathbf{w}_i s_i + \mathbf{e} \tag{1}$$

where \mathbf{y} is a vector of yield deviations, μ is a mean; \mathbf{u} is a vector of random additive polygenic effects and is $\sim N(0, \mathbf{G}\sigma_u^2)$ where \mathbf{G} is genomic relationship matrix based on all cows with phenotypes per breed and all autosomes. \mathbf{Z} is incidence matrix relating phenotypes \mathbf{y} to \mathbf{u} , \mathbf{w}_i is a vector of genotypes for SNP i , s_i is the effect of SNP i , and \mathbf{e} is a vector of random residual effects. We calculated the relationship between two individuals' j and k as

$g_{jk} = \frac{1}{w} \sum_{i=1}^W \frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)}$, x_{ij} being the number of alleles for individual j and SNP i and p_i is the observed allelic frequency, and, w was 43,800. We applied a genome-wide Bonferroni correction on all 43,800 tests to account for multiple testing.

Candidate variant discovery

We used the Association Weight Matrix (AWM) procedure to identify candidate genes per breed [8]. The AWM is a multiple trait approach that considers the genetic contribution of correlated traits allowing selection of pleiotropic SNP associated with numerous traits rather than a single trait. We classified trait information as either key or supportive trait, and the key trait in this study was udder depth or development (UDD) which is the most important type trait with the strongest relationship with mammary health and longevity. In addition, UDD is an aggregate trait, combining size, attachments, balance and strength of support. Populating the AWM starts with the selection of significant SNP from a GWAS [19]. The SNP additive effects are z-scored normalized by deviating the allele substitution effects from their mean and dividing by their standard deviation. We then created two matrices: (a) A z-scored additive values matrix (b) The GWAS p-values matrix. In both cases, rows represent SNP and columns represent traits. We then processed these matrices using the AWM algorithm, which includes five steps: (1) Primary SNP Selection: We select SNP associated with key trait using a P-value threshold ($P < 0.05$). (2): Exploring the dependency among traits: For the SNP selected in step (1), and, for the same threshold ($P < 0.05$), we register the average number of non-key traits to which the SNP are associated. In this study, that number was five traits. (3): Secondary SNP Selection: We select SNP from step (1) associated with at least five other traits including at least two udder traits. This step depends on correlation amongst traits and allows capturing most SNP associated with remaining traits. (4): Exploiting the genome map: We annotated the

SNP captured in step (1), and step (3) using the UMD3.1 Genome assembly [16]. We classified the SNP that (i) mapped a gene, (ii) <10-kb to known genes, and, (iii) >10-kb to any coding region. For genes represented by more than one SNP, we select the SNP associated with the highest number of traits and has the lowest P-value average across all traits as representative for that gene. (5): Populating the AWM: Each {i,t} cell value in the AWM matrix corresponds to the z-score normalized additive effect of the ith SNP on the tth trait. This allows exploration of trait correlations column-wise, and gene/SNP interactions row-wise. We then calculate the SNP-based correlations and compare the SNP-based and genetic correlations, the latter being calculated as pedigree-based restricted maximum likelihood (REML), established in the same cows' populations. We then use the AWM SNP matrix as the input for the PCIT algorithm [20] and for any trio of SNP; we estimate the first order partial correlation coefficients to identify meaningful gene-gene interactions. We annotated and clustered gene ontology (GO) annotations for significant PCIT gene-gene interactions using Cytoscape [21]. Finally, we compare the gene clusters amongst the three breeds and plot the most significant cluster.

Results

GWAS for all traits

Collectively for 78,440 cows, imputation resulted in a genotype density of 43,800 SNP. Of these, 38,827, 38,109, and 40,810 SNP had a MAF greater than 0.1 in Montbeliarde, Normande, and Holstein. Distribution of allele frequencies (MAF) of imputed genotypes was almost uniformly distributed across the MAF classes (Fig 1).

We performed GWAS for real and imputed SNP and yield deviations (YD) for 12 traits: five udder conformation traits, five milk production phenotypes, somatic cell score and clinical mastitis (Table 2). Fig 2 presents Manhattan plots for the key trait (Udder depth or development) for three breeds and S1 Fig for other traits. As expected, the number of significant SNP increased with breeds sample size. Holstein had 7,029 associated SNP, Montbeliarde had 1,762

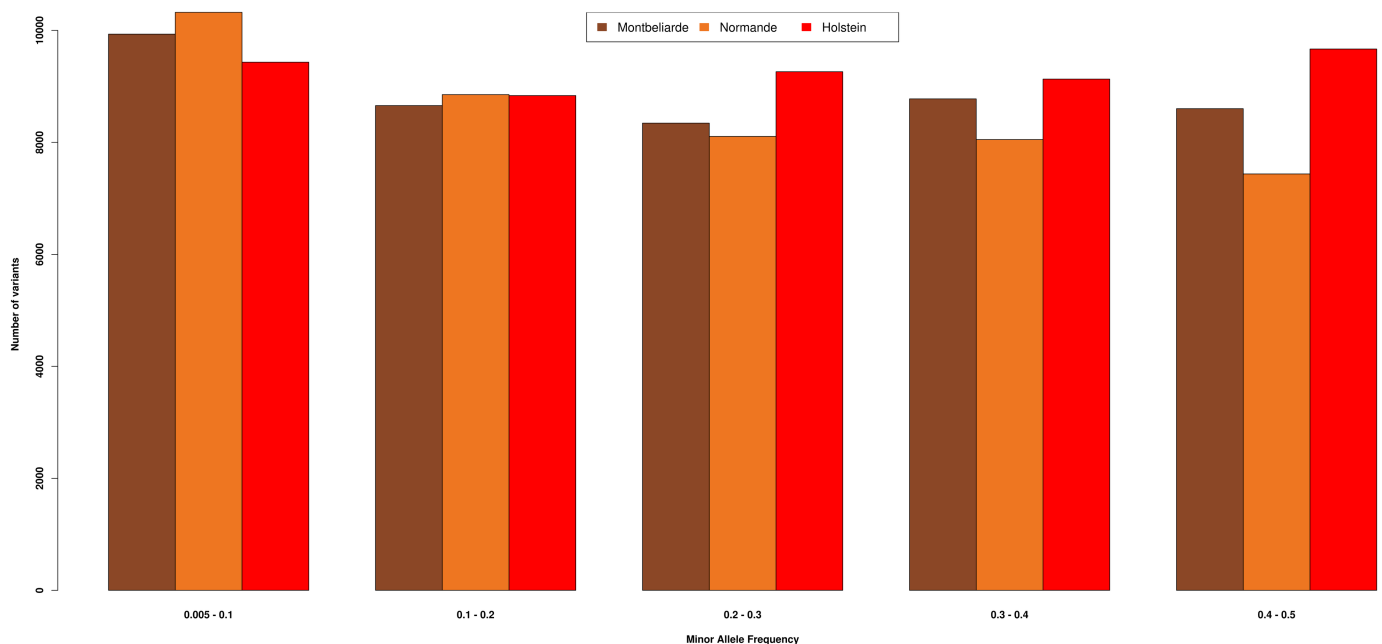


Fig 1. SNP variants as expressed in MAF in three French dairy breeds. The minor allele frequency of SNP variants in three French dairy breeds.

<https://doi.org/10.1371/journal.pone.0199931.g001>

Table 2. Summary of GWAS for udder related traits in three French dairy cattle breeds.

Trait ²	Significant SNP ¹			Significant SNP close ³ to gene		
	MON ⁴	NOR ⁵	HOL ⁶	MON ⁴	NOR ⁵	HOL ⁶
Fore Udder Attachment	45	29	558	32	20	235
Udder Depth or Development	204	38	57	84	13	47
Udder Cleft	37	24	402	29	22	237
Udder Balance	100	17	199	58	16	98
Front Teat Placement	125	37	45	64	29	33
Milk Yield	70	53	684	30	42	182
Fat Yield	68	47	515	31	41	112
Protein Yield	5	10	362	2	6	137
Fat Percent	481	390	1,498	136	105	480
Protein Percent	613	382	2,362	214	168	816
Clinical Mastitis	14	12	336	5	6	113
Somatic Cell Score			11			8

¹Significant SNP = A SNP that has satisfied the Bonferroni threshold for a trait;

²Trait = A yield; deviation: phenotype corrected for environmental variances;

³Significant SNP within 10-kb of gene;

⁴Montbeliarde;

⁵Normande;

⁶Holstein

<https://doi.org/10.1371/journal.pone.0199931.t002>

associated SNP, and, Normande had 1,039 associated SNP (Table 2). There was an overlap of significant SNP across the breeds. Some 483 SNP were common between Holstein and Montbeliarde, 356 SNP were common between Holstein and Normande, 233 SNP were common between Montbeliarde and Normande, and 206 SNP were common among three breeds (Fig 3). We observed overlap of significant SNP between udder and milk production traits. Irrespective of the breed, 15 SNP overlapped in udder related traits, and 205 SNP overlapped in milk production phenotypes. When considering 1SNP:1Gene, 3,651 significant SNP were close to genes (<10-Kb) across three breeds. Of these, 1,017 were highly associated with udder conformation traits, 2,502 with production traits and 132 with somatic cell score and clinical mastitis. The 2,673 additional SNP satisfying step 4 of the AWM algorithm (as described in M&M) was augmented with the 3,651 significant SNP from GWAS forming the AWM matrix with 6,324 SNP (S1 Table). Of these, 1,309 SNP were common across three breeds, and they mapped 1013 genes for 12 traits.

Significant SNP associated with the key trait were evident for Holstein and Montbeliarde, and the most significant SNP that mapped a gene for the key trait is presented per chromosome in Table 3. In total, 17 SNP were most significant per chromosome SNP and mapped to a gene in Holstein and Montbeliarde. This included eight for Montbeliarde, nine for Holstein. Two lead SNP in Montbeliarde were rs41640614 (BTA16, *SOAT1* gene, $p = 3.47 \times 10^{-15}$, Effect size = -0.185(0.02), MAF = 0.08) and rs108972236 (BTA19, *ABCA5* gene, $p = 2.20 \times 10^{-11}$, Effect size = -0.112(0.01), MAF = 0.20). The two lead SNP in Holstein were rs41641987 (BTA19, *PAFAH1B1* gene, $p = 1.51 \times 10^{-13}$, Effect size = -0.128(0.02), MAF = 0.04) and rs110651226 (BTA29, *FOXRED1* gene, $p = 2.30 \times 10^{-12}$, Effect size = 0.059(0.01), MAF = 0.45).

We observed SNP associated with FUA and FTP in all breeds. The most significant of these signals were in Holstein and Montbeliarde and they include: BTA17 (rs41609100, *FGF2* gene, Effect size = -0.12(0.01), $p = 4.95 \times 10^{-12}$, MAF = 0.073, Holstein), BTA20 (rs109428015, *PRLR*

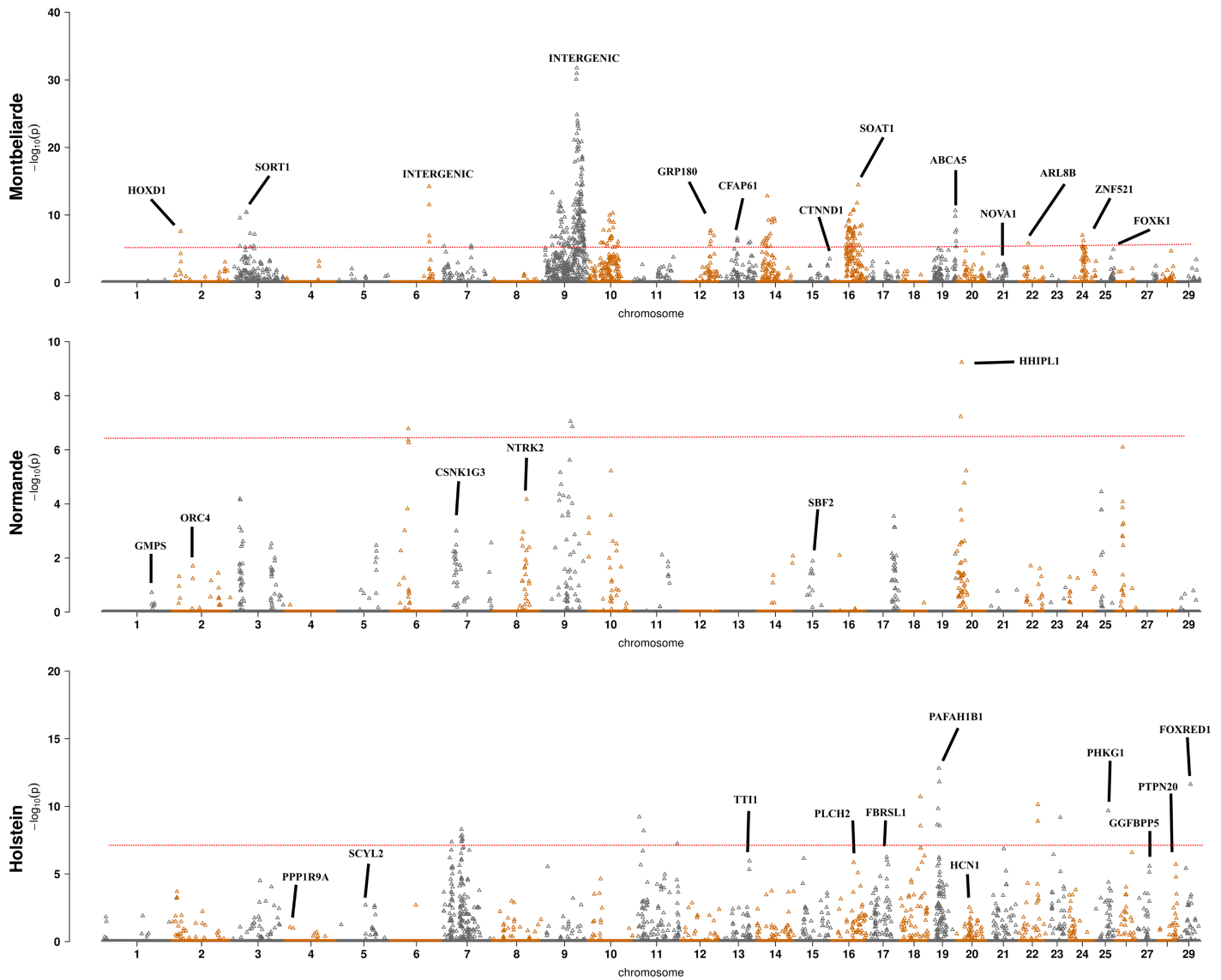


Fig 2. Manhattan plots for the key trait: Udder depth (Holstein and Normande) or development (Montbeliarde breed).

<https://doi.org/10.1371/journal.pone.0199931.g002>

gene, Effect size = -0.22(0.03), $p = 3.16 \times 10^{-16}$, MAF = 0.033, Holstein), and, BTA26 (rs42088948, *BTRC* gene, Effect size = -0.10(0.01), $p = 6.89 \times 10^{-4}$, MAF = 0.068, Montbeliarde).

Genetic parameters and genomic AWM-based correlations

Heritability coefficients (diagonal), pedigree-based genetic correlations (upper diagonal), and SNP correlations (lower diagonal) are presented in Table 4. Heritability values are close to reported values for type traits [22]. Fore udder attachment, (FUA) is 0.34, 0.26, 0.33 for Montbeliarde, Normande, and Holstein and higher for the other traits. For example, for Milk yield (MY), they reached 0.50, 0.61, and 0.54 in Montbeliarde, Normande and Holstein, respectively. These high values reflect the nature of the yield deviations (YD), which is a mean of records for repeated traits. YD is adjusted for permanent environment effect and thereby has a reduced non-genetic variability. As an example, assuming additive genetic variance of milk

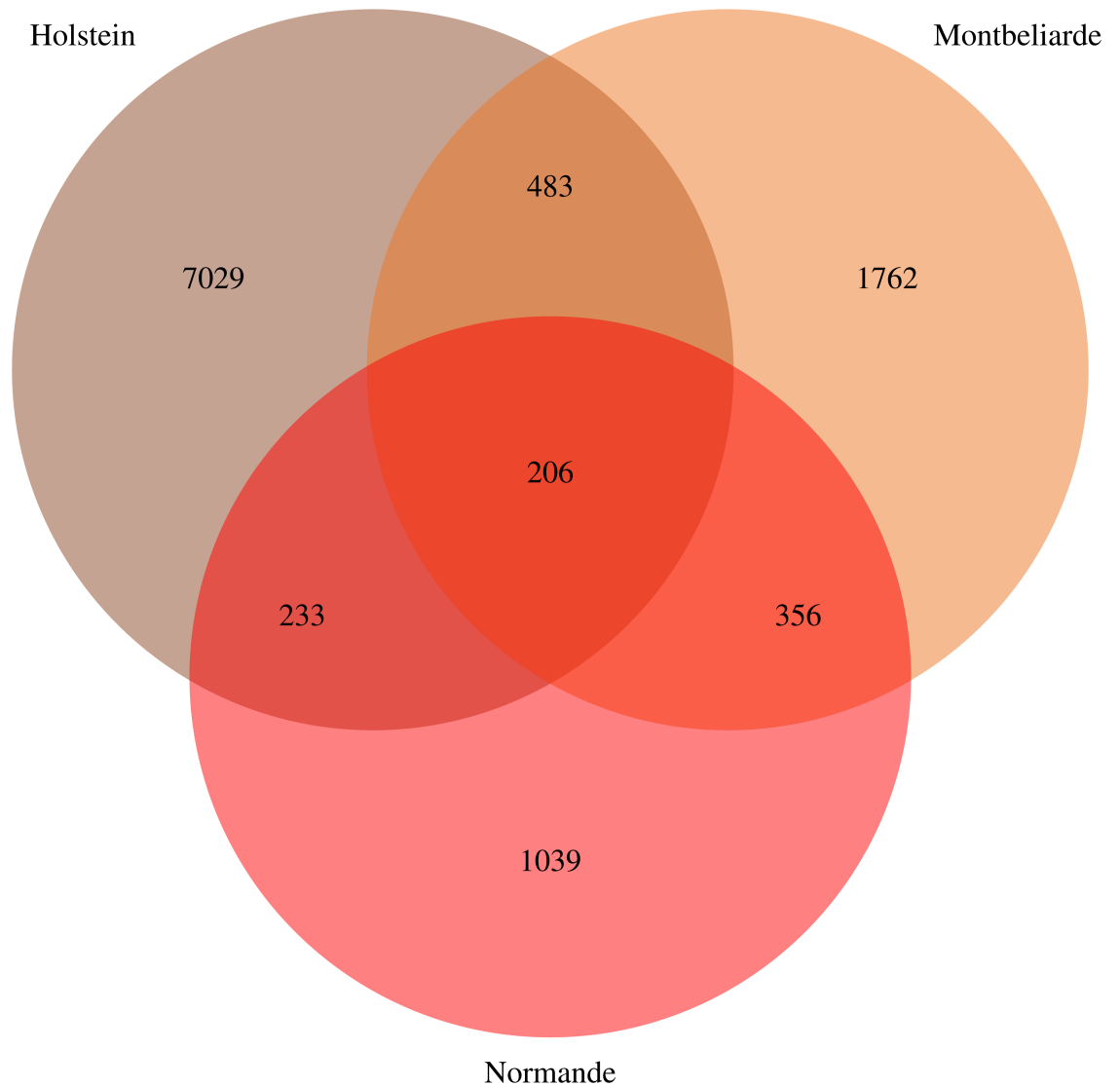


Fig 3. Summary of GWAS significant SNP and breed overlap for 12 traits in three French dairy breeds.

<https://doi.org/10.1371/journal.pone.0199931.g003>

yield is 0.3, permanent environment variance is 0.2, and residual variance is 0.5 (we divide the residual variance by 2.5 records on average for production), the heritability of the corresponding YD is $0.3 / (0.3 + 0.5/2.5) = 0.6$. These high values are very favorable for GWAS detection power.

Mammary morphology genetic correlations to production traits ranged from positive for fore udder attachment (FUA) and milk yield (MY) in Montbeliarde (0.42) and Normande (0.28) to medium negative for udder depth or development (UDD) and protein yield (PROT) in Normande (-0.56) and front teat placement (FTP) and fat yield (FAT) in Holstein (-0.33). SNP correlations were numerically different compared to genetic correlations; however, the correlation was generally in the same direction. For instance, the genetic correlation between fore udder attachment (FUA) and udder balance (UB) was 0.40, 0.63 and 0.39, whereas, SNP correlations for these two traits was 0.59, 0.16 and 0.16 for Montbeliarde, Normande and

Table 3. Most significant SNP per chromosome associated with udder depth/development and mapping a gene.

Breed	SNP ¹	BTA ²	Pos ³ (bp)	Effect allele	MAF ⁴	Effect size	SE Effect size	p ⁵	Gene ⁶
Montbeliarde	rs43293677	2	20760409	G	0.451	0.084	0.01	2.60X10 ⁻⁰⁸	HOXD1
Montbeliarde	rs29019267	3	34184021	A	0.367	-0.092	0.01	3.67X10 ⁻¹¹	SORT1
Montbeliarde	rs43704946	12	69648659	G	0.226	-0.095	0.01	1.97X10 ⁻⁰⁸	GPR180
Montbeliarde	rs109080985	13	40031719	A	0.46	0.075	0.01	2.63X10 ⁻⁰⁷	CFAP61
Montbeliarde	rs110761656	15	82317986	A	0.364	0.066	0.01	3.15X10 ⁻⁰⁴	CTNND1
Montbeliarde	rs41640614	16	62100110	A	0.08	-0.185	0.02	3.47X10 ⁻¹⁵	SOAT1
Montbeliarde	rs108972236	19	61919633	C	0.201	0.112	0.01	2.20X10 ⁻¹¹	ABCA5
Montbeliarde	rs41256881	22	21326038	G	0.315	0.078	0.01	1.49X10 ⁻⁰⁶	ARL8B
Montbeliarde	rs42049077	24	31765644	T	0.105	0.127	0.02	1.00X10 ⁻⁰⁷	ZNF521
Holstein	rs109049511	13	67557015	T	0.272	-0.051	0.01	1.07X10 ⁻⁰⁶	TTI1
Holstein	rs41808096	16	51621826	C	0.199	0.059	0.01	1.39X10 ⁻⁰⁶	PLCH2
Holstein	rs110859130	17	45680965	T	0.035	-0.128	0.02	5.57X10 ⁻⁰⁷	FBRSL1
Holstein	rs41641987	19	24136906	A	0.13	-0.088	0.01	1.51X10 ⁻¹³	PAFAH1B1
Holstein	rs42067431	25	28003780	C	0.413	-0.055	0.01	2.08X10 ⁻¹⁰	PHKG1
Holstein	rs41565991	27	27804403	A	0.194	-0.059	0.01	2.66X10 ⁻⁰⁶	GGFBPP5
Holstein	rs42147106	28	42881677	T	0.438	-0.047	0.01	1.94X10 ⁻⁰⁶	PTPN20
Holstein	rs110651226	29	30003729	A	0.45	0.059	0.01	2.30X10 ⁻¹²	FOXRED1

¹SNP = Reference SNP ID as assigned by NCBI;

²BTA = *Bos taurus* autosome;

³Position = base pair position in BTA (UMD3);

⁴MAF = minor allele frequency;

⁵p = Bonferroni corrected P-value;

⁶Gene = SNP annotation mapping a gene

<https://doi.org/10.1371/journal.pone.0199931.t003>

Holstein breeds, respectively. There were moderate to zero genetic correlations between milk yield (MY) and FUA for Montbeliarde (0.42), Normande (0.28), and Holstein (0), and low SNP correlation (Montbeliarde (0.10), Normande (0.16) and Holstein (0.16)) between MY and udder depth or development (UDD). Holstein breed had a highly positive genetic (0.49) and SNP correlation (0.59) between front teat placement (FTP) and clinical mastitis (CM), a trend that was not evident in other two breeds. However, Montbeliarde breed showed a strong SNP correlation between FTP and CM (0.33) and between UDD and CM (0.18). Other trends evident from genetic correlations were between FUA and PROT for Montbeliarde (0.39) and Normande (0.26), FUA and CM for Montbeliarde (0.19) and a moderate genetic correlation between FTP and CM for Holstein (0.49). Genetic and SNP correlations were also comparable between FTP and PROT for all breeds, with genetic/SNP correlation being from medium in Montbeliarde (0.21 / 0.12) and Holstein (-0.34 / -0.1) to low in Normande (-0.12 / -0.1). However, the trend deviated between udder balance (UB) and fat percent (FAT %) with minimal genetic correlation in Normande (0.22) and no genetic correlation in Montbeliarde and Holstein but with moderate SNP correlations in Montbeliarde (-0.41), Normande (0.38) and Holstein (0.38). In general, genetic correlations are in the range of usual values, with high correlations between milk, fat and protein, a moderately negative correlation between production and type traits, moderately positive correlations between conformation traits, and low correlations otherwise. However, though there were deviations between genetic and SNP correlations in some of the traits, most traits correlations were in the same direction thus drawing plausibility of SNP correlated traits.

Table 4. Heritability (diagonal), genetic¹ (upper-diagonal), and SNP² (lower-diagonal) correlations of udder conformation, milk production and health traits in three French dairy breeds.

Breed	Trait ³	FUA	UDD	FTP	UB	UC	MY	FAT	PROT	FAT%	PROT%	CM	SCS
Montbeliarde	FUA	0.34	-0.11	-0.28	0.40	0.37	0.42	0.43	0.39	0.03	-0.05	0.19	-0.05
	UDD	0.54	0.37	-0.39	-0.38	0.37	0.33	0.32	0.31	-0.02	-0.08	0.09	0.03
	FTP	0.04	-0.01	0.43	-0.24	-0.15	0.20	0.22	0.21	0.06	0.04	0.06	0.02
	UB	0.59	0.66	-0.21	0.36	-0.03	0.02	0.01	0.00	0.00	-0.03	-0.09	-0.02
	UC	0.65	0.61	-0.05	0.4	0.30	-0.01	-0.01	-0.01	-0.05	-0.06	0.00	-0.05
	MY	0.33	0.1	-0.05	0.03	0.17	0.50	0.88	0.94	-0.22	-0.13	0.02	0.02
	FAT	-0.02	-0.07	0.08	-0.1	-0.06	0.31	0.44	0.90	0.26	0.09	-0.01	0.02
	PROT	-0.01	-0.09	0.12	-0.05	-0.08	0.42	0.15	0.44	-0.06	0.20	0.01	0.05
	FAT%	-0.32	-0.35	0.09	-0.41	-0.32	-0.26	0.25	0.11	0.68	0.46	-0.03	-0.02
	PROT%	0.01	0.09	0.19	0.09	0.02	-0.23	0.29	0.15	0.15	0.66	-0.03	0.00
CM	0.25	0.18	0.33	0.08	0.18	0.75	-0.1	-0.01	-0.17	0.18	0.03	0.20	
SCS	0.54	0.75	-0.08	0.55	0.7	0.03	-0.11	-0.05	-0.26	0.01	0.04	0.39	
Normande	FUA	0.26	0.42	0.39	0.63	0.15	0.28	0.27	0.26	-0.01	-0.03	-0.07	-0.04
	UDD	0.21	0.33	0.27	0.51	0.64	-0.53	-0.54	-0.56	-0.07	-0.23	0.00	-0.01
	FTP	-0.04	0.03	0.39	0.38	0.49	-0.13	-0.10	-0.12	0.06	0.02	-0.02	-0.02
	UB	0.16	0.11	-0.19	0.30	0.35	0.01	0.10	0.06	0.22	0.23	-0.01	-0.02
	UC	0.2	0.21	-0.08	0.72	0.33	0.03	0.00	0.02	-0.04	-0.31	0.00	-0.01
	MY	-0.53	0.16	0.06	0.1	0.06	0.61	0.90	0.95	-0.16	0.04	0.00	-0.10
	FAT	0.08	0.12	0.09	0.18	0.15	0.78	0.60	0.93	0.28	0.25	0.00	0.00
	PROT	0.01	0.37	-0.1	0.4	0.38	0.27	0.37	0.62	-0.01	0.33	0.02	-0.05
	FAT%	0.28	0.29	-0.21	0.38	0.44	0.21	0.31	0.43	0.63	0.50	-0.03	0.06
	PROT%	0.19	0.19	0.02	-0.09	0.01	0.12	0.12	0.09	0.16	0.61	-0.03	0.03
CM	0.07	0.37	0.09	-0.01	0.05	0.01	-0.01	-0.06	0.04	0.30	0.03	0.37	
SCS	-0.01	0.06	0.00	-0.27	-0.25	0.00	-0.1	-0.18	-0.11	0.33	0.10	0.40	
Holstein	FUA	0.33	0.49	-0.33	0.39	0.09	0.00	-0.02	-0.01	-0.04	-0.05	-0.07	-0.02
	UDD	0.21	0.46	-0.16	0.42	0.19	-0.07	-0.09	-0.09	-0.02	-0.06	-0.01	0.01
	FTP	-0.04	0.03	0.48	-0.29	-0.35	-0.34	-0.33	-0.34	0.06	-0.05	0.49	-0.04
	UB	0.16	0.11	-0.19	0.23	0.27	0.05	0.05	0.07	-0.01	0.09	0.12	-0.02
	UC	0.2	0.21	-0.08	0.72	0.33	-0.02	-0.01	0.00	-0.05	0.02	-0.05	0.01
	MY	0.09	0.16	0.06	0.1	0.06	0.54	0.85	0.96	-0.43	-0.03	-0.01	0.13
	FAT	0.08	0.12	0.09	0.18	0.15	0.78	0.49	0.90	0.10	0.26	0.00	-0.03
	PROT	0.01	0.37	-0.1	0.4	0.38	0.27	0.37	0.48	-0.27	0.25	-0.03	0.07
	FAT%	0.28	0.29	-0.21	0.38	0.44	0.21	0.31	0.43	0.65	0.51	0.01	-0.10
	PROT%	0.19	0.19	0.02	-0.09	0.01	0.12	0.12	0.09	0.16	0.58	-0.01	-0.10
CM	0.07	0.37	0.59	-0.01	0.05	0.01	-0.01	-0.06	0.12	-0.03	0.02	-0.07	
SCS	-0.01	0.06	0.00	-0.27	-0.25	0.00	-0.1	-0.18	-0.04	0.04	0.33	0.38	

¹Genetic correlation = correlations based on phenotypic data as estimated using a pedigree based REML;

²SNP correlation = correlations estimated from AWM matrix;

³Trait = A yield deviation: phenotype corrected for environmental variances: FUA: fore udder attachment, UDD: udder depth development, FTP: front teat placement, UB: udder balance, UC: udder cleft, MY: milk yield, FAT and FAT%: fat yield and content, PROT and PROT%: protein yield and content, CM: clinical mastitis, SCS: Somatic cell score

<https://doi.org/10.1371/journal.pone.0199931.t004>

Gene ontology (GO) for AWM selected genes

We considered main ontology domains including, biological processes, cellular component, and molecular functions. We observed 39 gene clusters (S2 Table) and Table 5 presents the

Table 5. Top five clusters of transcription factors/genes and associated GO terms associated with udder morphology in three French dairy breeds.

Transcription factor ¹ /genes GO ² Term	Gene count	p	Top Associated Genes ³
cluster1 Enrichment Score ⁴ : 9.17			
mammary gland epithelium development	134	4.64 x10 ⁻¹⁶	[BTRC, ESRI, FGF2, FGFR2, GLI2, IQGAP3, PGR, PRLR, RREB1, TGFB2]
regulation of response to stimulus	3653	7.14 x10 ⁻⁶	[BMP7, GCH1]
gland development	429	6.50 x10 ⁻¹²	[ROR2, RORA, SYCP2, TCF7L2]
epithelium development	1057	6.64 x10 ⁻⁹	[FGFR2, HHIP, RDH10, RSPO2, TNC]
tissue development	1714	2.10 x10 ⁻⁷	[AKT3, CTNNA2, CTNNA3, GNAI3, INADL, MAGI3, PARD3, PPP2R2B, PRKCB, PRKCE, PRKCH, PTEN, RAB3B]
animal organ development	3022	2.66 x10 ⁻⁷	[ADGRB1, ADGRB3, ANGPT1, BMPER, CALCRL, COL8A1, CSPG4, EPAS1, FGF2, FGFR2, FN1, GJA5, HDAC9, HSPG2, LOXL2, MAP2K5, MTDH, NOV, NRP2, NRXN3, PDCL3, PRKCB, PTEN, PTK2, ROCK2, RORA, SH2D2A, SHB, SPI1, STAB1, TGFB2, THSD7A, TIE1, VAV3]
system development	4309	5.38 x10 ⁻⁵	[HMGCLL1, HMGCS2, OXCT1]
multicellular organism development	4855	2.00 x10 ⁻⁴	[GALNT13, GALNT14, GALNT18, GALNTL6, GCNT3]
cluster2 Enrichment Score ⁴ : 8.92			
mammary gland alveolus development	68	5.16 x10 ⁻¹⁶	[ESRI, PRLR]
intracellular signal transduction	2681	6.97 x10 ⁻⁵	[ABCA1, SYCP2]
regulation of intracellular signal transduction	1632	4 x10 ⁻⁴	[ANXA4, CD58]
cluster3 Enrichment Score ⁴ : 8.41			
positive regulation of macromolecule metabolic process	2863	2.77 x10 ⁻⁹	[CHRN2, GPAM, IL7, PELI1, SPTA1, STAT5B, VAV3, VTCN1]
positive regulation of cellular metabolic process	2899	2.62 x10 ⁻⁹	[ARPC2, AXIN2, BAIAP2, CDH4, CUX1, CUX2, EPB41L5, EPHA4, FBXW8, FN1, FYCO1, LPAR3, LRP8, LRRC16A, MAP2K2, NTRK2, PTPRD, RREB1, RUFY3, SEMA4D, TCF7L2, TGFB3, TGFB2]
positive regulation of metabolic process	3563	1.46 x10 ⁻⁷	[PIWIL2, PLCB1]
negative regulation of biological process	4579	8.50 x10 ⁻⁶	[DAB1, FYCO1, PTEN, PTK2, RTN4, RUFY3, SEMA3B, SEMA4D, SYNGAP1]
cluster4 Enrichment Score ⁴ : 8.26			
branching morphogenesis of an epithelial tube	146	5.32 x10 ⁻⁹	[CD44, EYA1, FAT4, FGF2]
morphogenesis of a branching epithelium	178	1.29 x10 ⁻⁸	[ADORA1, ARRB1, CACNA1A, CACNA1B, GABBR2, GABRA1, GNAI3, GNB5, GNG7, GNGT2, PDE11A, PDE1A, PDE2A, PRKACB, PRKCB]
epithelial tube morphogenesis	305	1.65 x10 ⁻⁷	[FGFR2, HHIP, RDH10, RSPO2, TNC]
tube morphogenesis and development	341	2.28 x10 ⁻⁷	[AARS2, ALKBH8, CDK5RAP1, CDKAL1, NDC1, NUP210, NUP93, SARS, TSEN54]
organ morphogenesis	905	2.63 x10 ⁻⁵	[CHRN2, KCNA2]
cluster5 Enrichment Score ⁴ : 7.99			
cellular response to organic substance	2352	2.89 x10 ⁻⁹	[LBP, NDUFA2]
cellular response to chemical stimulus	2845	1.28 x10 ⁻⁷	[ABCC9, AKAP6, CACNA1D, CASQ2, HCN1, KCNA2, KCNAB2, KCNB2, KCNC1, KCND3, KCNH1, KCNK2, KCNMA1, KCNMB1, KCNN2, KCNN3, KCNT1, KCNT2, KCNV2, STK39, YWHAE]
response to organic substance	2981	2.47 x10 ⁻⁷	[ASS1, SMYD3, SRD5A1, TRIM63]
response to chemical	4373	6.51 x10 ⁻⁵	[MCM3, POLD1, POLE3, PRIM2, SMARCAL1]

¹Transcription factor = a protein that controls rate of transcription of genetic information from DNA to messenger RNA, by binding to a specific DNA sequence.

Sometimes its homonymous to gene;

²GO = Gene Ontology as expressed in main domains: Biological, Cellular and Molecular;

³Top Associated Genes = most probable associated candidate gene of interest for udder morphology;

⁴Enrichment Score = measure of over-represented (or under-represented) GO terms using AWM gene annotations

<https://doi.org/10.1371/journal.pone.0199931.t005>

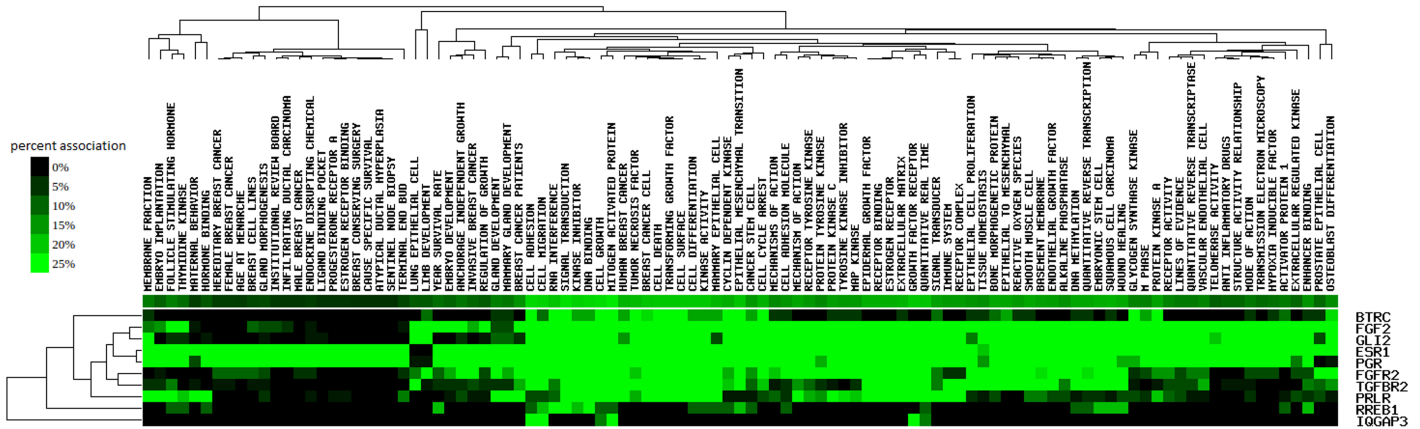


Fig 4. Ten candidate gene-GO term interactions with percentage association in three French dairy breeds.

<https://doi.org/10.1371/journal.pone.0199931.g004>

top five biological processes that are relevant for udder morphology and health traits (S3 Table contains complete gene ontology list for this study). Top cluster had eight GO terms with the most significant GO term being “mammary gland epithelium development” ($p = 4.64 \times 10^{-16}$). Among the AWM-PCIT, genes ten were transcription factors (TF) directly associated with the terms “mammary gland development”, “mammary gland duct morphogenesis”, “mammary gland alveolus development”, “tissue development”, and “epithelial tube morphogenesis”. These ten TF include *GLI2* (BTA2:72.98Mb), *IQGAP3* (BTA3:14.31Mb), *PGR* (BTA4:62.53Mb), *ESR1* (BTA9:89.97Mb), *FGF2* (BTA17:35.23Mb), *PRLR* (BTA20:39.13Mb), *TGFB2* (BTA22:5.14Mb), *RREB1* (BTA23:47.90Mb), *BTRC* (BTA26:22.06Mb), and *FGFR2* (BTA26:41.82Mb). Fig 4 presents their GO term interactions with percentage association. Other GO terms in the top cluster included “gland development” ($p = 6.50 \times 10^{-12}$) and “system development” ($p = 5.38 \times 10^{-5}$). Top GO term for the second cluster was “mammary gland epithelium development” ($p = 5.16 \times 10^{-16}$) while “regulation of intracellular signal transduction” was the least significant term in this cluster ($p = 4 \times 10^{-4}$). There was an enrichment for “neuropathic pain-signaling in dorsal horn neurons pathway” ($p = 2.57 \times 10^{-8}$), “G-protein coupled receptor signaling” ($p = 1.07 \times 10^{-7}$) as well as the “CREB signaling in neurons” ($p = 7.24 \times 10^{-7}$). Other pathways detected were “cAMP-mediated signaling” ($p = 2.88 \times 10^{-5}$), “synaptic long-term depression ($p = 8.71 \times 10^{-7}$), “axonal guidance signaling” ($p = 1.15 \times 10^{-6}$), and “synaptic long-term potentiation” ($p = 1.58 \times 10^{-5}$).

Pathway enrichments detected (S3 File) included “Calcium: cation antiporter activity” and the “Calcium-activated Potassium channel activity” for molecular functions ($p < 10^{-3}$), whereas, for biological processes, “dendrite development” and “putrescine biosynthetic” process were most represented ($P < 10^{-3}$) while “postsynaptic density” and “presynaptic membrane” were top GO terms for cellular components.

Discussion

Previously, Fortes et al. [8] and Ramayo-Caldas et al. [23] suggested the Association Weight Matrix’s (AWM) as an alternative tool to identify genes that would otherwise be missed by traditional single-trait GWAS. This study further supports that suggestion by focusing on GWAS for other kinds of traits, such as udder morphology and health traits common across three French dairy breeds. Though single-trait-single-SNP GWAS focus is on most significant SNP, they can, aid in identifying lead SNP for QTL associated with a given trait. Our study identified three lead SNP associated with front teat placement (FTP) and fore udder attachment (FUA).

These SNP mapped *FGF2*, *PRLR*, and *BTRC* genes. *PRLR* gene (prolactin receptor) was previously associated with milk production traits in Finnish Ayrshire dairy cows [24]. Wang et al. [25] reported the association of *FGF2* gene (Fibroblast growth factor 2) to fat yield and percentage and somatic cell score in US Holstein. Coleman-Krncak et al. [26] reported the expression of *FGF2* gene in the bovine mammary gland and uterine endometrium (UE). In the mammary gland, the *FGF2* gene may play a role in development and reorganization of the mammary gland, while in UE, *FGF2* gene is mainly expressed throughout estrous cycle and early pregnancy. *BTRC* gene (Beta-Transducing Repeat Containing E3 Ubiquitin Protein Ligase) is an F-box protein involved in Wnt/ β -Catenin signaling pathway and indirectly activates nuclear factor kappa-B (NF- κ B) [27]. Raven et al. reported these pathways to be highly relevant during mammary development and pregnancy, and as such, could have a major functional role in lactation [27]. The AWM-PCIT algorithm identified interacting candidate genes for udder conformation traits by first establishing SNP based correlation and fixing udder depth or development (UDD) as the key trait when constructing the AWM SNP matrix. These genes were represented by several biological processes involved in positive regulation of cellular biosynthetic processes and cell development, suggesting an endogenous characterization linked to udder morphology [28]. Fortes et al. [10] reported more similar SNP and genetic correlations for traits with moderate to high heritability and less similar correlations between traits with low heritability (Table 4). In our study, most traits had a heritability >10% thus aiding both GWAS detection power and AWM SNP detection.

The AWM was assessed for gene ontology (GO). The top GO terms were Calcium cation antiporter and Calcium-activated Potassium channel activity. As reported by Paulsen et al. [29], the former is a member of the cation diffusion facilitator (CDF) superfamily which are integral membrane proteins that increase tolerance to divalent metal ions, whereas, the latter is involved in ionic signaling in cells, a critical function for hormonal control of cell proliferation and differentiation [30]. Control of calcium signaling is likely to have profound effects on mammary physiology and pathophysiology. Their high significance level can be explained by the fact that mammary glands extract large quantities of calcium from the plasma during lactation [31], to ensure sufficient calcium concentration in milk. Dendrite development and putrescine biosynthetic processes were top biological GO terms. Dendritic cells (DC) are accessory cells of the mammalian immune system whose work is availing antigen material to T cells of the immune system [32–33]. This ability to stimulate native T-cells makes this result significant because it can directly be applied to improve udder health. DC has also been reported to play a pivotal role in the initiation of an adaptive immune response [34]. Putrescine, as a member of polyamine pathway, is regulated by the periovulatory endocrine milieu [35].

The central pathways that showed enrichment were “Neuropathic pain-signaling in dorsal horn neurons pathway,” and, “CREB signaling pathway.” Neuropathic pain is the pain after nerve injury whereas the dorsal horn (tip of the spinal cord) pathways have been shown to offer substantial overlap with spinal projections from adjacent mammary glands in model organisms [36]. Reflex milk ejection may result from the strong integration of sensory input from mammary glands afferents that terminate in the dorsal horn. CREB (cyclic-AMP response element-binding) protein family of transcription factors (TF—a protein that controls the rate of transcription of genetic information from DNA to messenger RNA, by binding to a specific DNA sequence [37]) play a crucial role in supporting the survival of sensory neurons [38]. The intervention of sensory neurons stimulates cells to secrete nerve growth factor (NGF) via the sympathetic nervous system (SNS) that maintains homeostasis [39]. NGF mediate its functions through ligation of tyrosine kinase (Trk) receptors [40]. Regulation of Trk signaling is by a variety of intracellular signaling cascades, which include MAPK

pathway, which promotes cell continuation and growth [41]. Previous studies indicate that Trk could be multifunctional growth factors that exert various effects through their receptors on non-neuronal cells such as mammary ducts [30].

Fontanesi et al. [8] reported that long-term transcriptomic adaptations of tissues depend on the action of external stimuli that induce action of cellular functions on transcription factors (TF) [42]. In this study, we identified 39 interacting gene clusters and the most significant cluster had ten TF directly involved with mammary gland development and mammary gland duct morphogenesis: *BTRC*, *ESR1*, *FGF2*, *FGFR2*, *GLI2*, *IQGAP3*, *PGR*, *PRLR*, *RREB1*, and *TGFBR2*. These TF are homonymous with genes encoding them. This top cluster is directly involved in mammary gland development, regulation of response to a stimulus, gland development, epithelium development, tissue development, animal organ development, system development and multicellular organism development. This was partly in agreement to works reported by Yang et al. [43] on QTL associated with follicle stimulating hormone production in Chinese Holstein cattle.

Conclusion

Our study suggests the usefulness of system-based approaches to identify candidate genes from interacting gene networks in a multi-breed context. We achieved this by exploiting associations between correlated traits. The reluctant inclusion of intergenic SNP leaves the possibility that the AWM approach was not capturing significant SNP such as trans-activators. Nonetheless, the AWM proved to be more efficient for integrating related complex traits and analyzing thousands of SNP and therefore appropriate for the analysis of these complex traits. When applied to our dataset, it predicted gene interactions that are consistent with the known biology of udder morphology and health captured known TF (e.g., *ESR1*, *FGF2*, *GLI2*, *PGR* and *BTRC*), and provided new candidate genes for udder morphology. This experiment may be replicated using whole genome sequence data and other independent datasets.

Supporting information

S1 Fig. Manhattan plots for milk production, udder morphology and udder health traits in three French dairy cattle breeds.

(PDF)

S1 Table. Associated Weight Matrix for three French dairy cattle breeds.

(XZ)

S2 Table. Enrichment scores for udder morphology genes.

(XLSX)

S3 Table. Gene ontology for AWM-PCIT selected genes.

(XLSX)

Acknowledgments

The authors are grateful to the French National Animal Breeding Database for providing phenotypic data, Valogene for access to genotypes and INRA for providing the study environment.

Author Contributions

Conceptualization: Yulixaxis Ramayo-Caldas.

Data curation: Andrew Marete.

Formal analysis: Andrew Marete, Yuliaxis Ramayo-Caldas.

Methodology: Andrew Marete.

Software: Andrew Marete.

Supervision: Mogens Sandø Lund, Didier Boichard.

Validation: Andrew Marete, Yuliaxis Ramayo-Caldas.

Visualization: Andrew Marete, Yuliaxis Ramayo-Caldas.

Writing – original draft: Andrew Marete.

Writing – review & editing: Andrew Marete, Yuliaxis Ramayo-Caldas.

References

1. Goddard ME, Hayes BJ. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rev Genet.* 2009; 10: 381–391. <https://doi.org/10.1038/nrg2575> PMID: 19448663
2. Mackay TFC. The Genetic Architecture of Quantitative Traits. *Annu Rev Genet.* 2001; 35: 303–339. <https://doi.org/10.1146/annurev.genet.35.102401.090633> PMID: 11700286
3. van den Berg I, Boichard D, Lund MS. 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. *J Dairy Sci.* 2016; 99: 8932–8945.
4. Druet T, Schrooten C, de Roos APW. Imputation of genotypes from different single nucleotide polymorphism panels in dairy cattle. *J Dairy Sci.* 2017; 93: 5443–5454. <https://doi.org/10.3168/jds.2010-3255> PMID: 20965360
5. Marchini J, Howie B, Myers S, Mcvean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet.* 2007; 39: 906–913. <https://doi.org/10.1038/ng2088> PMID: 17572673
6. Pausch H, Emmerling R, Schwarzenbacher H, Fries R. A multi-trait meta-analysis with imputed sequence variants reveals twelve QTL for mammary gland morphology in Fleckvieh cattle. *Genet Sel Evol.* 2016; 48: 14. <https://doi.org/10.1186/s12711-016-0190-4> PMID: 26883850
7. Tiezzi F, Parker-Gaddis KL, Cole JB, Clay JS, Maltecca C, Tiezzi F, et al. A genome-wide association study for clinical mastitis in first parity US Holstein cows using single-step approach and genomic matrix re-weighting procedure. *PLoS One.* 2015; 10: e0114919. <https://doi.org/10.1371/journal.pone.0114919> PMID: 25658712
8. Fortes MRS, Reverter A, Zhang Y, Collis E, Nagaraj SH, Jonsson NN, et al. Association weight matrix for the genetic dissection of puberty in beef cattle. *Proc Natl Acad Sci. National Academy of Sciences;* 2010; 107: 13642–13647. <https://doi.org/10.1073/pnas.1002044107> PMID: 20643938
9. Widmann P, Reverter A, Fortes MRS, Weikard R, Suhre K, Hammon H, et al. A systems biology approach using metabolomic data reveals genes and pathways interacting to modulate divergent growth in cattle. *BMC Genomics.* 2013; 14: 798. <https://doi.org/10.1186/1471-2164-14-798> PMID: 24246134
10. Fortes MRS, Reverter A, Nagaraj SH, Zhang Y, Jonsson NN, Barris W, et al. A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. *J Anim Sci.* 2011; 89: 1669–1683. <https://doi.org/10.2527/jas.2010-3681> PMID: 21357453
11. Ramayo-Caldas Y, Ballester M, Fortes MR, Esteve-Codina A, Castelló A, Noguera JL, et al. From SNP co-association to RNA co-expression: Novel insights into gene networks for intramuscular fatty acid composition in porcine. *BMC Genomics.* 2014; 15: 232. <https://doi.org/10.1186/1471-2164-15-232> PMID: 24666776
12. García-Gámez E, Reverter A, Whan V, McWilliam SM, Arranz JJ, International Sheep Genomics Consortium, et al. Using regulatory and epistatic networks to extend the findings of a genome scan: Identifying the gene drivers of pigmentation in Merino sheep. *PLoS One.* 2011; 6. <https://doi.org/10.1371/journal.pone.0021158> PMID: 21701676
13. Boichard D, Guillaume F, Baur A, Croiseau P, Rossignol MN, Boscher MY, et al. Genomic selection in French dairy cattle. *Anim Prod Sci.* 2012; 52: 115–120. <https://doi.org/10.1071/AN11119>
14. VanRaden PM, Wiggans GR. Derivation, calculation and use of national animal model information. *J Dairy Sci.* 1991; 74.

15. Madsen P, Jensen J. A user's guide to DMU. Center for Quantitative Genetics and Genomics Dept. of Molecular Biology and Genetics, University of Aarhus Research Centre Foulum Box 50, 8830 Tjele Denmark. 2013. <http://dmu.agrsci.dk>
16. Elsik CG, Unni DR, Diesh CM, Tayal A, Emery ML, Nguyen HN, et al. Bovine Genome Database: new tools for gleaning function from the *Bos taurus* genome. *Nucleic Acids Res*. 2016; 44: D834–D839. <https://doi.org/10.1093/nar/gkv1077> PMID: 26481361
17. Sargolzaei M, Chesnais JP, Schenkel FS, Nejati-Javaremi A, Smith C, Gibson J, et al. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*. 2014; 15: 478. <https://doi.org/10.1186/1471-2164-15-478> PMID: 24935670
18. Yang J, Ferreira T, Morris AP, Medland SE, Madden PAF, Heath AC, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet*. 2012; 44: 369–375. <https://doi.org/10.1038/ng.2213> PMID: 22426310
19. Reverter A, Fortes MRS. Association weight matrix: A network-based approach towards functional genome-wide association studies. In: Gondro C, van der Werf J, Hayes B, editors. *Methods in Molecular Biology*. 2013. pp. 437–447.
20. Watson-Haigh NS, Kadarmideen HN, Reverter A. PCIT: An R package for weighted gene co-expression networks based on partial correlation and information theory approaches. *Bioinformatics*. 2009; 26: 411–413. <https://doi.org/10.1093/bioinformatics/btp674> PMID: 20007253
21. Shannon P, Markiel A, Owen Ozier, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; 2498–2504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658
22. Rupp R, Boichard D. Genetic Parameters for Clinical Mastitis, Somatic Cell Score, Production, Udder Type Traits, and Milking Ease in First Lactation Holsteins. *J Dairy Sci*. 1999; 82: 2198–2204.
23. Ramayo-Caldas Y, Renand G, Ballester M, Saintilan R, Rocha D. Multi-breed and multi-trait co-association analysis of meat tenderness and other meat quality traits in three French beef cattle breeds. *Genet Sel Evol*. 2016; 48: 37. <https://doi.org/10.1186/s12711-016-0216-y> PMID: 27107817
24. Viitala S, Szyda J, Blott S, Schulman N, Lidauer M, Mäki-Tanila A, et al. The role of the bovine growth hormone receptor and prolactin receptor genes in milk, fat and protein production in Finnish Ayrshire dairy cattle. *Genetics*. 2006; 173: 2151–2164. <https://doi.org/10.1534/genetics.105.046730> PMID: 16751675
25. Wang X, Maltecca C, Tal-Stein R, Lipkin E, Khatib H. Association of bovine fibroblast growth factor 2 (FGF2) gene with milk fat and productive life: An example of the ability of the candidate pathway strategy to identify quantitative trait genes. *J Dairy Sci*. 2008; 91: 2475–2480. <https://doi.org/10.3168/jds.2007-0877> PMID: 18487671
26. Coleman-Krnacik S, Rosen JM. Differential temporal and spatial gene expression of fibroblast growth factor family members during mouse mammary gland development. *Mol Endocrinol*. 1994; 8: 218–229. <https://doi.org/10.1210/mend.8.2.8170478> PMID: 8170478
27. Raven LA, Cocks BG, Kemper KE, Chamberlain AJ, Vander Jagt CJ, Goddard ME, et al. Targeted imputation of sequence variants and gene expression profiling identifies twelve candidate genes associated with lactation volume, composition and calving interval in dairy cattle. *Mamm Genome*. 2016; 27: 81–97. <https://doi.org/10.1007/s00335-015-9613-8> PMID: 26613780
28. Luscombe NM, Babu MM, Yu H, Snyder M, Teichmann S a, Gerstein M. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature*. 2004; 431: 308–312. <https://doi.org/10.1038/nature02782> PMID: 15372033
29. Paulsen IT, Saier M.H. Jr. A Novel Family of Ubiquitous Heavy Metal Ion Transport Proteins. *J Membr Biol*. 1997; 156: 99–103.
30. Lee WJ, Monteith GR, Roberts-Thomson SJ. Calcium transport and signaling in the mammary gland: Targets for breast cancer. *Biochim Biophys Acta—Rev Cancer*. 2006; 1765: 235–255. <https://doi.org/10.1016/j.bbcan.2005.12.001> PMID: 16410040
31. Shennan DB, Peaker M. Transport of milk constituents by the mammary gland. *Physiol Rev*. 2000; 80: 925–51. <https://doi.org/10.1152/physrev.2000.80.3.925> PMID: 10893427
32. Banchereau J, Steinman RM. Dendritic cells and the control of immunity. *Nature*. 1998; 392: 245–252. <https://doi.org/10.1038/32588> PMID: 9521319
33. Maxymiv NG, Bharathan M, Mullarky IK. Bovine mammary dendritic cells: A heterogeneous population, distinct from macrophages and similar in phenotype to afferent lymph veiled cells. *Comp Immunol Microbiol Infect Dis*. 2012; 35: 31–38. <https://doi.org/10.1016/j.cimid.2011.09.009> PMID: 22019401
34. Bridges GA, Mussard ML, Pate JL, Ott TL, Hansen TR, Day ML. Impact of preovulatory estradiol concentrations on conceptus development and uterine gene expression. *Anim Reprod Sci*. 2012; 133: 16–26. <https://doi.org/10.1016/j.anireprosci.2012.06.013> PMID: 22789700

35. Ramos R dos S, Mesquita FS, D’Alexandri FL, Gonella-Diaza AM, Papa P de C, Binelli M. Regulation of the polyamine metabolic pathway in the endometrium of cows during early diestrus. *Mol Reprod Dev.* 2014; 81: 584–594. <https://doi.org/10.1002/mrd.22323> PMID: 24659573
36. Tasker JG, Theodosios DT, Poulain DA. Afferent projections from the mammary glands to the spinal cord in the lactating rat-I. A neuroanatomical study using the transganglionic transport of horseradish peroxidase-wheatgerm agglutinin. *Neuroscience.* 1986; 19: 495–509. [https://doi.org/10.1016/0306-4522\(86\)90276-9](https://doi.org/10.1016/0306-4522(86)90276-9) PMID: 3774151
37. Latchman DS. Transcription factors: An overview. *Int J Biochem Cell Biol.* 1997; 29: 1305–1312. [https://doi.org/10.1016/S1357-2725\(97\)00085-X](https://doi.org/10.1016/S1357-2725(97)00085-X) PMID: 9570129
38. Parlato R, Otto C, Begus Y, Stotz S, Schütz G. Specific ablation of the transcription factor CREB in sympathetic neurons surprisingly protects against developmentally regulated apoptosis. *Development.* 2007; 134: 1663–1670. <https://doi.org/10.1242/dev.02838> PMID: 17376811
39. Riccio A, Pierchala BA, Ciarallo CL, Ginty DD. An NGF-TrkA-Mediated Retrograde Signal to Transcription Factor CREB in Sympathetic Neurons. *Science.* 1997; 277: 1097 LP–1100. <https://doi.org/10.1126/science.277.5329.1097>
40. Bothwell M. NGF, BDNF, NT3, and NT4. Lewin GR, Carter BD, editors. *Handb Exp Pharmacol.* Berlin, Heidelberg, Heidelberg: Springer Berlin Heidelberg; 2014; 220: 3–15. https://doi.org/10.1007/978-3-642-45106-5_1 PMID: 24668467
41. Reichardt LF. Neurotrophin-regulated signalling pathways. *Philos Trans R Soc B Biol Sci.* London: The Royal Society; 2006; 361: 1545–1564. <https://doi.org/10.1098/rstb.2006.1894> PMID: 16939974
42. Fontanesi L, Calo DG, Galimberti G, Negrini R, Marino R, Nardone A, et al. A candidate gene association study for nine economically important traits in Italian Holstein cattle. *Anim Genet.* 2014; 45: 576–580. <https://doi.org/10.1111/age.12164> PMID: 24796806
43. Yang SH, Bi XJ, Xie Y, Li C, Zhang SL, Zhang Q, et al. Validation of PDE9A gene identified in GWAS showing strong association with milk production traits in Chinese holstein. *Int J Mol Sci.* 2015; 16: 26530–26542. <https://doi.org/10.3390/ijms161125976> PMID: 26556348