



HAL
open science

Genome-wide association studies and genomic prediction of breeding values for calving performance and body conformation traits in Holstein cattle

Mohammed K. Abo-Ismaïl, Luiz F. Brito, Stephen P. Miller, Mehdi Sargolzaei, Daniela A. Grossi, Steve S. Moore, Graham Plastow, Paul Stothard, Shadi Nayeri, Flavio S. Schenkel

► To cite this version:

Mohammed K. Abo-Ismaïl, Luiz F. Brito, Stephen P. Miller, Mehdi Sargolzaei, Daniela A. Grossi, et al.. Genome-wide association studies and genomic prediction of breeding values for calving performance and body conformation traits in Holstein cattle. *Genetics Selection Evolution*, 2017, 49 (1), pp.82. 10.1186/s12711-017-0356-8. hal-01631052

HAL Id: hal-01631052

<https://hal.science/hal-01631052>

Submitted on 8 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.


L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

Open Access



Genome-wide association studies and genomic prediction of breeding values for calving performance and body conformation traits in Holstein cattle

Mohammed K. Abo-Ismael^{1,2} , Luiz F. Brito¹, Stephen P. Miller^{1,3}, Mehdi Sargolzaei^{1,4}, Daniela A. Grossi¹, Steve S. Moore⁵, Graham Plastow⁶, Paul Stothard⁶, Shadi Nayeri⁶ and Flavio S. Schenkel^{1*}

Abstract

Background: Our aim was to identify genomic regions via genome-wide association studies (GWAS) to improve the predictability of genetic merit in Holsteins for 10 calving and 28 body conformation traits. Animals were genotyped using the Illumina Bovine 50 K BeadChip and imputed to the Illumina BovineHD BeadChip (HD). GWAS were performed on 601,717 real and imputed single nucleotide polymorphism (SNP) genotypes using a single-SNP mixed linear model on 4841 Holstein bulls with breeding value predictions and followed by gene identification and in silico functional analyses. The association results were further validated using five scenarios with different numbers of SNPs.

Results: Seven hundred and eighty-two SNPs were significantly associated with calving performance at a genome-wide false discovery rate (FDR) of 5%. Most of these significant SNPs were on chromosomes 18 (71.9%), 17 (7.4%), 5 (6.8%) and 7 (2.4%) and mapped to 675 genes, among which 142 included at least one significant SNP and 532 were nearby one (100 kbp). For body conformation traits, 607 SNPs were significant at a genome-wide FDR of 5% and most of them were located on chromosomes 5 (30%), 18 (27%), 20 (13%), 6 (6%), 7 (5%), 14 (5%) and 13 (3%). SNP enrichment functional analyses for calving traits at a FDR of 1% suggested potential biological processes including musculoskeletal movement, meiotic cell cycle, oocyte maturation and skeletal muscle contraction. Furthermore, pathway analyses suggested potential pathways associated with calving performance traits including tight junction, oxytocin signaling, and MAPK signaling ($P < 0.10$). The prediction ability of the 1206 significant SNPs was between 78 and 83% of the prediction ability of the BovineSNP50 SNPs for calving performance traits and between 35 and 79% for body conformation traits.

Conclusions: Various SNPs that are significantly associated with calving performance are located within or nearby genes with potential roles in tight junction, oxytocin signaling, and MAPK signaling. Combining the significant SNPs or SNPs within or nearby gene(s) from the HD panel with the BovineSNP50 panel yielded a marginal increase in the accuracy of prediction of genomic estimated breeding values for all traits compared to the use of the BovineSNP50 panel alone.

Background

The profitability of dairy production depends on the ability of cows for high milk production, good health and

fertility. Improving traits such as calving performance and body conformation reduces the culling rate, which in turn affects the profitability of the dairy cattle industry [1–4]. More than 50% of the first lactation heifers may require assistance at calving [5], which increases their culling risk by 18% and reduces their reproductive life [6]. In addition, dams with very difficult calving

*Correspondence: schenkel@uoguelph.ca

¹ Department of Animal Biosciences, Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, ON, Canada
Full list of author information is available at the end of the article

(score = 4) require eight more days to first service and 28 more days to conceive than those not needing assistance (score = 1) [2]. Furthermore, Dematawewa and Berger [7] reported that cows that experience extreme calving difficulty (score = 5) produce 703 kg of milk, 24 kg of fat and 21 kg of protein less per 305-d of lactation than unassisted cows (score = 1) over multiple parities. Body conformation traits are correlated to economically important traits in dairy cattle such as calving ease [8], longevity [9], lameness [10] and lifetime production efficiency [11]. Improving the accuracy of selection for calving performance and conformation traits would benefit the dairy industry as a whole and have a major impact on the profitability of individual farms and, thus, these traits are included in traditional and genomic breeding programs worldwide [12–15].

Since 2009, significant genetic improvement has been achieved for various dairy cattle traits through the use of high throughput single nucleotide polymorphism (SNP) panels for implementing genomic selection [16]. Increases in the rate of genetic progress were predicted to be as much as 100% [17], and data collected by the Canadian Dairy Network over the last few years support these predictions [18]. The Illumina Bovine SNP50 SNP chip (50 K; Illumina Inc., San Diego, USA) [19] has been used to identify significant regions that are associated with calving [20] and conformation traits [21]. Over the last four years, higher-density SNP chip panels have been developed, including the Illumina BovineHD BeadChip (HD; Illumina Inc., San Diego, USA) and, also, many animals have been sequenced [22]. These HD panels have not had a significant impact on genetic improvement programs, with very minor increases in prediction accuracies over the 50 K panel, either within or across breeds [23, 24]. Nonetheless, HD panels can be used to improve the detection of regions that harbor causal mutations and, in which, genes of interest and new SNPs can be identified.

The objectives of our study were to identify genomic regions associated with 10 calving performance and 28 body conformation traits in Holstein cattle, to retrieve information from associated regions to increase the understanding of the underlying biology of these traits, and to test the predictive ability of these regions in young bulls.

Methods

Description of traits

The evaluated traits included 10 calving performance and 28 body conformation traits (Table 1). Calving performance traits were: maternal calving ease at first (heifer; CEh) and later calvings (cow; CEc), maternal calf survival at first (heifer; CSh) and later calvings (cow; CSc), direct

calving ease at first (heifer; SCEh) and later calvings (cow; SCEc), direct calf survival at first (heifer; SCSh) and later calvings (cow; SCSc), calving ability index (CA), and daughter calving ability index (DCA). The phenotypes for calving ease were scored as 1 for unassisted, 2 for easy pull, 3 for hard pull and 4 for surgery. Calf survival was scored within 24 h after calving as 0 for dead and 1 for alive. The body conformation traits consisted of 23 linear descriptive traits scored on a 1 to 9 scale and four composite traits which were estimated from linear traits and overall score (www.holstein.ca). A detailed description of the traits and statistical methods for predicting the breeding values of the bulls was previously reported [9, 13, 25–27].

Genotypes and imputation

Genotypes for HD (774,605 SNPs) and 50 K (40,666 SNPs) SNPs were provided by the Canadian Dairy Network (CND, Guelph, ON, Canada) for 2387 and 11,926 bulls, respectively, both registered and used in North America. The genotyping data was coded as 0, 1, and 2 for BB, AB and AA genotypes, respectively. For purposes of this study, reference population will refer to the animals genotyped with the HD panel used to create the library of haplotypes for the imputation from 50 K to HD genotypes, while target population refers to the bulls genotyped with the 50 K panel and imputed to HD genotypes.

A genotyping quality control (QC) on the HD genotypes removed 39,366 SNPs located on the sex chromosomes, 5427 individual SNPs with a call-rate lower than 90% and 31 SNPs with a heterozygosity that deviates by more than 15% from the expected value under Hardy–Weinberg equilibrium or that are not in Hardy–Weinberg equilibrium ($P < 0.0000001$). In addition, 3679 SNPs with a high Mendelian error rate (> 0.05) and 6902 SNPs that were assumed to be misplaced and identified based on linkage disequilibrium (LD) decay (see Additional file 1: Figure S1) were excluded. A total of 719,200 SNPs from the HD panel distributed over the 29 *Bos taurus* autosomes (BTA) passed all QC criteria and were used for imputation of the target population. These same QC criteria were applied to the 50 K genotypes and 39,148 SNPs remained for further analyses.

Genotype imputation from the 50 K to HD panel was performed by using the FImpute software and applying a population imputation approach [28]. SNPs with more than 10% of genotypes imputed by random filling based on the allele frequencies in the reference population had these imputed genotypes set to missing using the *gp_thresh* option in the FImpute software and, in a further QC step, these SNPs were excluded from further analyses.

Table 1 Heritability and number of records used in the association and validation analyses

Trait	Heritability ^a	Number of records in the training population ^b	Number of records in validation population ^c
<i>Calving performance</i>			
Calving ability, CA ^d	0.300	4525	–
Daughter calving ability, DCA ^e	0.060	3892	726
Maternal calving ease at first calving (heifer), CEh	0.121	3908	726
Maternal calf survival at first calving (heifer), CSh	0.056	3873	726
Maternal calving ease at later calvings (cow), CEC	0.084	3879	726
Maternal calf survival at later calvings (cow), CSC	0.023	3530	725
Direct calving ease at first calving (heifer), SCEh	0.017	4523	–
Direct calf survival at first calving (heifer), SCSh	0.004	4361	–
Direct calving ease at later calvings (cow), SCEc	0.016	4516	–
Direct calf survival at later calvings (cow), SCSc	0.003	4315	–
<i>Rump</i>			
Rump, RUM	0.233	3573	927
Pin width, PW	0.340	3584	927
Pin setting, PS	0.087	3533	947
Rump angle, RAN	0.365	3573	927
Loin strength, LS	0.251	3576	947
<i>Body traits</i>			
Dairy strength, DS	0.359	3584	927
Stature, ST	0.529	3589	927
Height at front end, FE	0.259	3577	947
Chest width, CW	0.218	3582	927
Body depth, BD	0.320	3582	927
Angularity, ANG	0.257	3582	927
<i>Feet and legs</i>			
Feet and legs, FL	0.152	3572	927
Foot angle, FAN	0.109	3582	927
Heel depth, HD	0.076	3486	947
Bone quality, BQ	0.300	3584	947
Leg side view, LSV	0.244	3577	927
Set of rear legs, SRL	0.050	3564	947
Leg rear view, LRV	0.125	3555	927
<i>Udder</i>			
Mammary system, MS	0.247	3587	927
Udder depth, UD	0.415	3577	927
Udder texture, UT	0.141	3588	947
Median suspensory ligament, MSL	0.140	3581	927
Fore udder attachment, FA	0.282	3587	927
Front teat placement, FTP	0.313	3584	927
Rear attachment height, RAH	0.234	3585	927
Rear attachment width, RAW	0.200	3579	947
Rear teat placement, RTP	0.294	3580	927
Teat length, TL	0.293	3581	927
Overall conformation score, CONF	0.261	3586	927

^a Heritability = heritability estimates were provided by Canadian Dairy Network

^b Number of bulls in the training population with de-regressed EBV and used in the association analysis

^c Number of bulls in the validation population

^d CA = calving ability index = 0.64 SCEh + 0.16 SCEc + 0.16 SCSh + 0.04 SCSc

^e DCA = daughter calving ability index = 0.36 CSh + 0.24 CSC + 0.24 CEh + 0.16 CEC

To estimate the accuracy of imputation, 387 animals were randomly selected from the reference population of animals with HD genotypes and their genotypes were masked down to the 50 K panel (39,148 SNPs). Then, the genotypes for these animals were imputed to HD by using the remaining 2000 animals in the reference population. The squared correlation (r^2) between imputed genotypes and true genotypes was used as a measure of imputation accuracy [29]. For the genome-wide association studies (GWAS) and genomic prediction of breeding values, imputation was carried out using all animals in the reference population ($n = 2387$).

After imputation, an additional QC was performed on the real and imputed genotypes. This step excluded 117,482 SNPs that had a minor allelic frequency (MAF) lower than 0.01 and one SNP for which heterozygosity deviated by more than 15% from the expected value under Hardy–Weinberg equilibrium. There were no animals and no individual SNPs with a call rate lower than 90%. A total of 601,717 SNPs distributed over the 29 bovine autosomes passed the QC procedure (see Additional file 2: Figure S2). To examine population stratification among the bulls in the training population, the HD genotypes were used to perform a multidimensional scaling analysis based on identical-by-state (IBS) pairwise distances, as implemented in PLINK software [30].

De-regressed estimated breeding values for the training population

Domestic official evaluations of 10 calving performance and 28 body conformation traits, and their associated reliabilities from the April 2009 genetic evaluation, were provided by the Canadian Dairy Network (CDN; www.cdn.ca) for 4841 progeny-tested Holstein bulls born between 1956 and 2007. These progeny-tested Holstein bulls had either a real HD ($n = 287$) or an imputed HD ($n = 4554$) genotype. The de-regressed estimated breeding values of bulls were computed by CDN following VanRaden et al. [31]:

$$\text{De} - \text{EBV} = \text{PA} + (\text{EBV} - \text{PA}) / \text{Rel}_{\text{De} - \text{EBV}}, \quad (1)$$

where De – EBV is the bull’s de-regressed estimated breeding value for the trait of interest, PA is the parent average and $\text{Rel}_{\text{De} - \text{EBV}}$ is the reliability of the bull’s EBV adjusted to remove the contribution of PA. For purposes of this study, training population refers to the population of bulls that had both pseudo-phenotypes (De – EBV) and genotypes used for the estimation of the effects of SNPs, while validation population refers to the animals used for the validation of the effects of SNPs estimated by using the training population.

Pedigree data for 23,287 Holsteins individuals were obtained from CDN. The level of pedigree completeness

of the training population was assessed using the pedigree completeness index (PCI) proposed by MacCluer et al. [32] and implemented in the CFC package [33]. The PCI was calculated as:

$$\text{PCI} = \frac{K_{\text{dam}} K_{\text{sire}}}{K_{\text{dam}} + K_{\text{sire}}}, \quad (2)$$

where the K_{dam} and K_{sire} are the indexes for dams and the sires, respectively, given by:

$$K = \frac{1}{g} \sum_{i=1}^g a_i$$

where a_i is the proportion of known ancestors in generation i for either the dams or the sires, and g is the number of generations back (5 in our study).

Genome-wide association analyses

The GWAS were performed using an univariate single-SNP mixed linear model implemented in the snp1101 software [34]. The mixed model equations are described as:

$$\begin{bmatrix} \mathbf{1}'\mathbf{R}^{-1}\mathbf{1} & \mathbf{1}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{1}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{X}'\mathbf{R}^{-1}\mathbf{1} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} * \frac{\sigma_e^2}{\sigma_g^2} \end{bmatrix} \times \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}'\mathbf{R}^{-1}\mathbf{Y} \\ \mathbf{X}'\mathbf{R}^{-1}\mathbf{Y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Y} \end{bmatrix}, \quad (3)$$

where \mathbf{Y} is the vector of the De – EBV for each trait; $\mathbf{1}$ is a vector of ones, \mathbf{X} is the vector of the animals’ genotypes for a given SNP, coded as 0, 1 and 2; \mathbf{Z} is the design matrix that assigns animals to De – EBV; $\hat{\mu}$ is the overall mean, $\hat{\beta}$ is the linear regression coefficient (allele substitution effect) of the examined SNP; $\hat{\mathbf{u}}$ is a vector of direct genomic values (DGV). Assumptions for the model were: $\mathbf{u} \sim N(0, \mathbf{G}\sigma_g^2)$ where \mathbf{G} is the genomic relationship matrix, \mathbf{R} is a diagonal matrix containing weights for the residual variance based on the reliabilities of the De – EBV. The diagonal elements of \mathbf{R} are given by $\left(\frac{1}{1 - \text{Rel}_i}\right)$ [35], where Rel_i is the reliability of the i th De – EBV; σ_g^2 is the additive genetic variance; and, σ_e^2 is the residual variance. The values of σ_g^2 and σ_e^2 were estimated using the AI-REML algorithm implemented in the snp1101 software [34]. Allele frequencies used to calculate \mathbf{G} were estimated from the observed genotype data. The \mathbf{G} matrix was calculated as:

$$\mathbf{G} = \mathbf{CC}' / 2 \sum P_i(1 - P_i), \quad (4)$$

where \mathbf{C} is the genotypic coefficient matrix with dimensions equal to the number of genotyped animals by the number of SNPs and P_i is the allele frequency of the i th SNP. The relationship between individuals j and k is G_{jk} divided by the square root of the diagonal values of j (G_{jj}) and k (G_{kk}) individuals [35].

The significance of the effects of SNPs was determined by using a genome-wide false discovery rate (FDR) of 5% [36]. The more lenient FDR threshold of 5% was used to increase the power to detect SNPs with small effects since traits of interest may be controlled by many QTL with a small effect [37]. The genomic inflation factor (λ) and quantile–quantile (Q–Q) plots were applied to assess the inflation of the test statistics due to population stratification by comparing the genome-wide distribution of the test statistic ($-\log_{10}$ of P -values) with the expected Chi squared distribution. The significant SNPs at a genome-wide FDR of 5% were aligned to the publicly available QTL in the Animal QTLdb (Release 26, access date: April 27th, 2015) [38].

In-silico functional analyses

Candidate SNP enrichment analyses were performed on the GWAS results using the SNP2GO R package [39] for gene ontology (GO). To infer overrepresentation of candidate SNPs (i.e., the list of significant SNPs at a P value < 0.05), the *Bos taurus* annotations from Ensembl version 75 for the UMD 3.1 assembly, the associated GO terms, and the list of candidate and non-candidate (i.e., non-significant) SNPs were used as input for SNP2GO. In the SNP2GO analyses, significant SNPs were assigned to plus or minus 50 nucleotides up- and down-stream of the corresponding genes. To account for multiple testing in the SNP enrichment analysis, a FDR of 1% was applied.

Pathway analyses were performed on the list of genes obtained by mapping the significant SNPs from the GWAS to their corresponding genes or nearby genes at a distance of 10 kbp using NGS-SNP [40], the bovine genome assembly UMD3.1 and the Ensembl database 75_31. The 10 kbp distance was used to exploit the expected LD (r^2) between pairs of syntenic SNPs, which had an average r^2 of 0.58 [41]. The pathway analyses were performed using the web-based Database for Annotation, Visualization, and Integrated Discovery (DAVID) version 6.8 [42].

Validation study

We performed a forward validation study, in which the Canadian domestic estimated breeding values from the April 2009 genetic evaluation were used to estimate SNP associations and direct genomic breeding values (DGV).

Unproven young bulls in the April 2009 genetic evaluation with estimated breeding values in December 2014 were considered for the validation population (Table 1). The DGV for the validation group were estimated by using the genomic best linear unbiased prediction method (GBLUP) [35, 43, 44] implemented in the *gebv* software [45]. In this method, the DGV were estimated using the following mixed model equations:

$$\begin{bmatrix} \mathbf{1}'\mathbf{R}^{-1}\mathbf{1} & \mathbf{1}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}_w^{-1}\frac{\sigma_e^2}{\sigma_g^2} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}'\mathbf{R}^{-1}\mathbf{Y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Y} \end{bmatrix}, \quad (5)$$

where \mathbf{Y} is the vector of De – EBV from the April 2009 genetic evaluation for the trait of interest for genotyped bulls, $\hat{\mu}$ is the overall mean, $\mathbf{1}$ is a vector of ones, \mathbf{Z} is the design matrix that assigns animals to De – EBV, $\hat{\mathbf{u}}$ is a vector of DGV. Assumptions for this model are: $\mathbf{u} \sim N(0, \mathbf{G}_w\sigma_g^2)$, where $\mathbf{G}_w = w^*\mathbf{G} + (1-w)^*\mathbf{A}$ and σ_g^2 is the additive genetic variance; \mathbf{G} is the genomic relationship matrix [35], \mathbf{A} is the numerator relationship matrix and w is the weight put on \mathbf{G} (0.8). As in Model (3), \mathbf{R} is

a diagonal matrix with elements defined as $R_{ii} = \left(\frac{1}{1 - \frac{1}{Rel_i}} \right)$

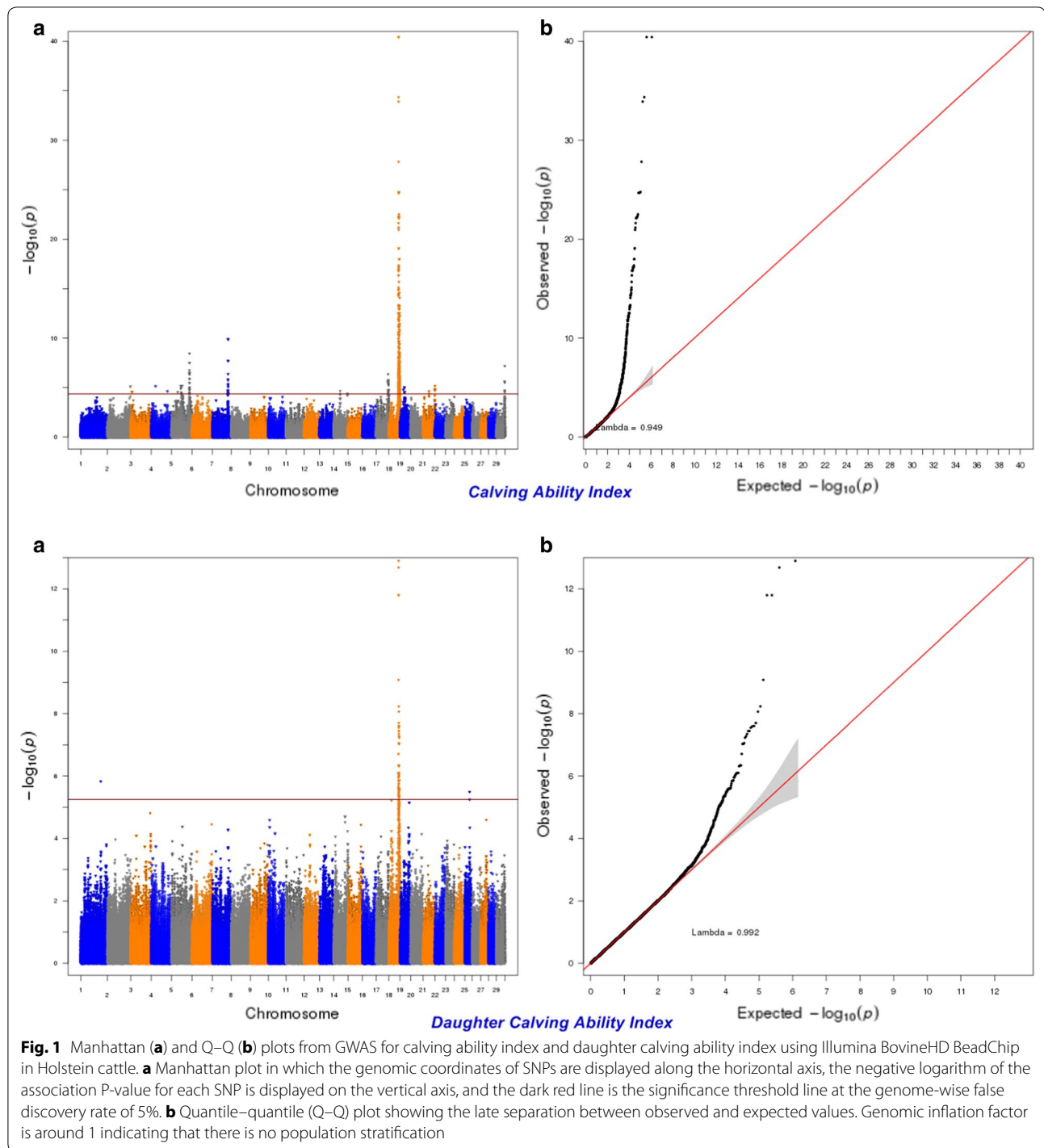
where Rel_i is the reliability of the i th De – EBV; σ_e^2 is the error variance.

Different scenarios for the validation analyses were tested depending on the number of SNPs used in the prediction. Scenario 1 used all the SNPs in the 50 K panel that passed the QC (38,724 SNPs), Scenario 2 used all the SNPs in the HD panel that passed the QC (607,794 SNPs), Scenario 3 used only a set of significant (genome-wide FDR of 5%) SNPs for all calving performance and conformation traits, Scenario 4 combined SNPs from the 50 K panel and the significant SNPs not in the 50 K panel, and Scenario 5 combined a set of significant SNPs for calving performance and conformation traits within or nearby the gene (± 100 kbp around the gene) plus the 50 K panel. The accuracy of genomic breeding values was calculated as the squared correlation (r^2) between DGV and official genetic evaluations from December 2014 for the validation group.

Results and discussion

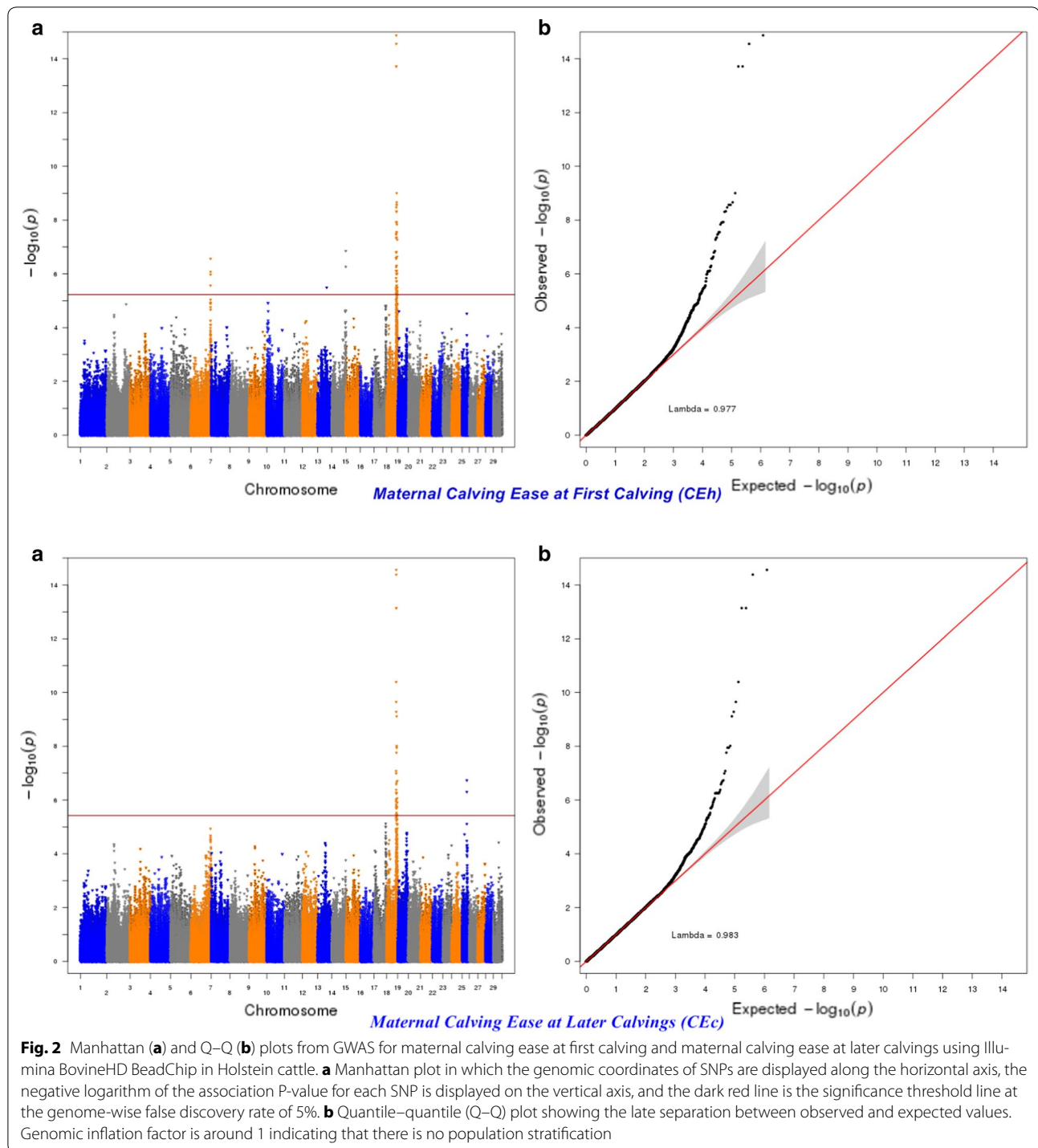
Population structure

The results from the multidimensional scaling analysis showed no divergent clusters within this population (see Additional file 3: Figure S3). In addition, the pedigree analysis indicated that the overall PCI considering five generations back was equal to 0.62. Although the PCI for the reference population was high, modeling the relationship



between animals using the pedigree information may not completely account for population structure because of missing pedigree records. The current Holstein cattle population was characterized by a high rate of inbreeding, genetic diversity loss and small effective population size

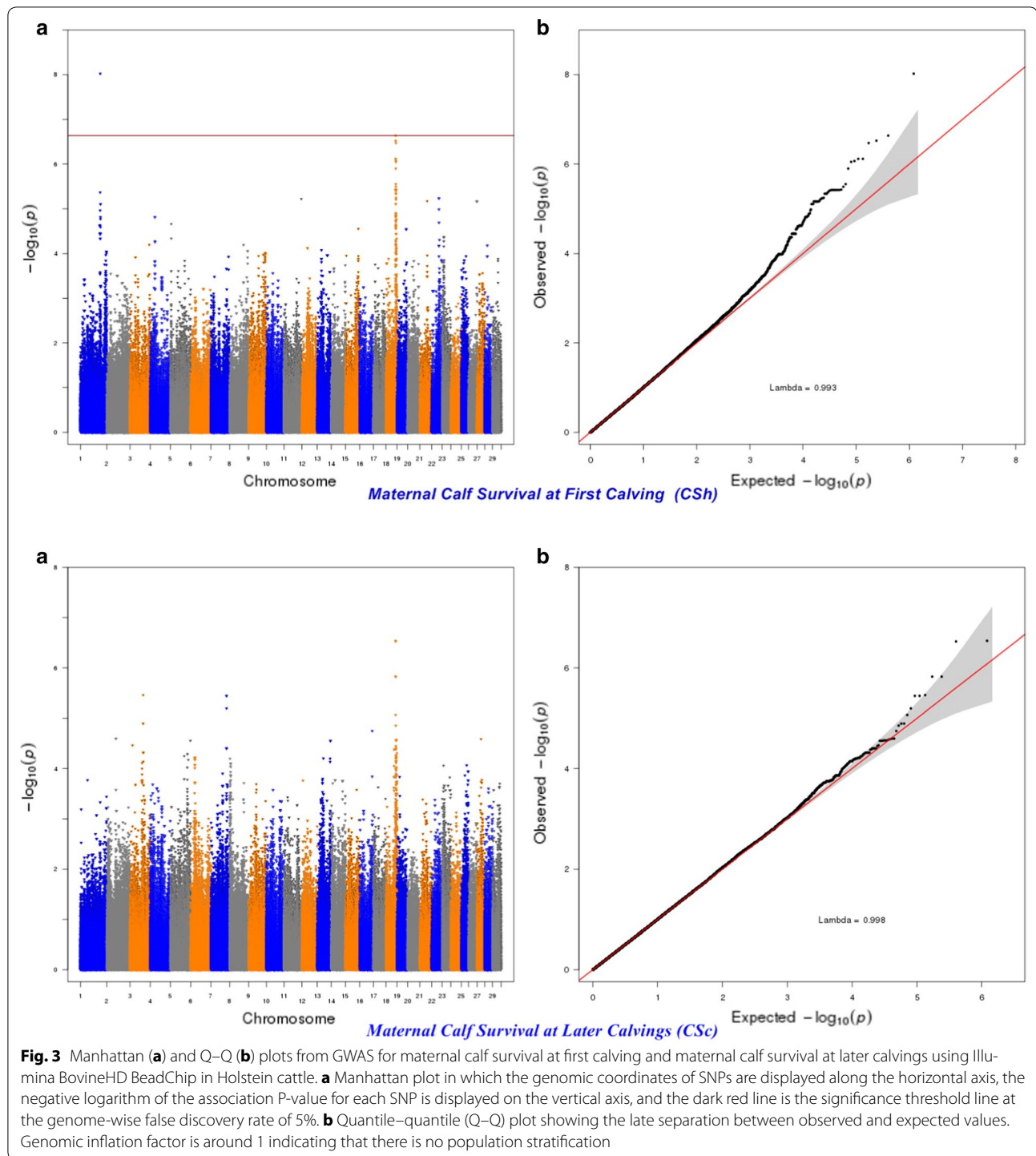
[46, 47] and would probably have cryptic relationships that are not described by the available pedigree. Thus, the population structure and cryptic relatedness not accounted for by an incomplete pedigree should be accounted for in GWAS to avoid spurious associations [48].



Accuracy of genome-wide imputation

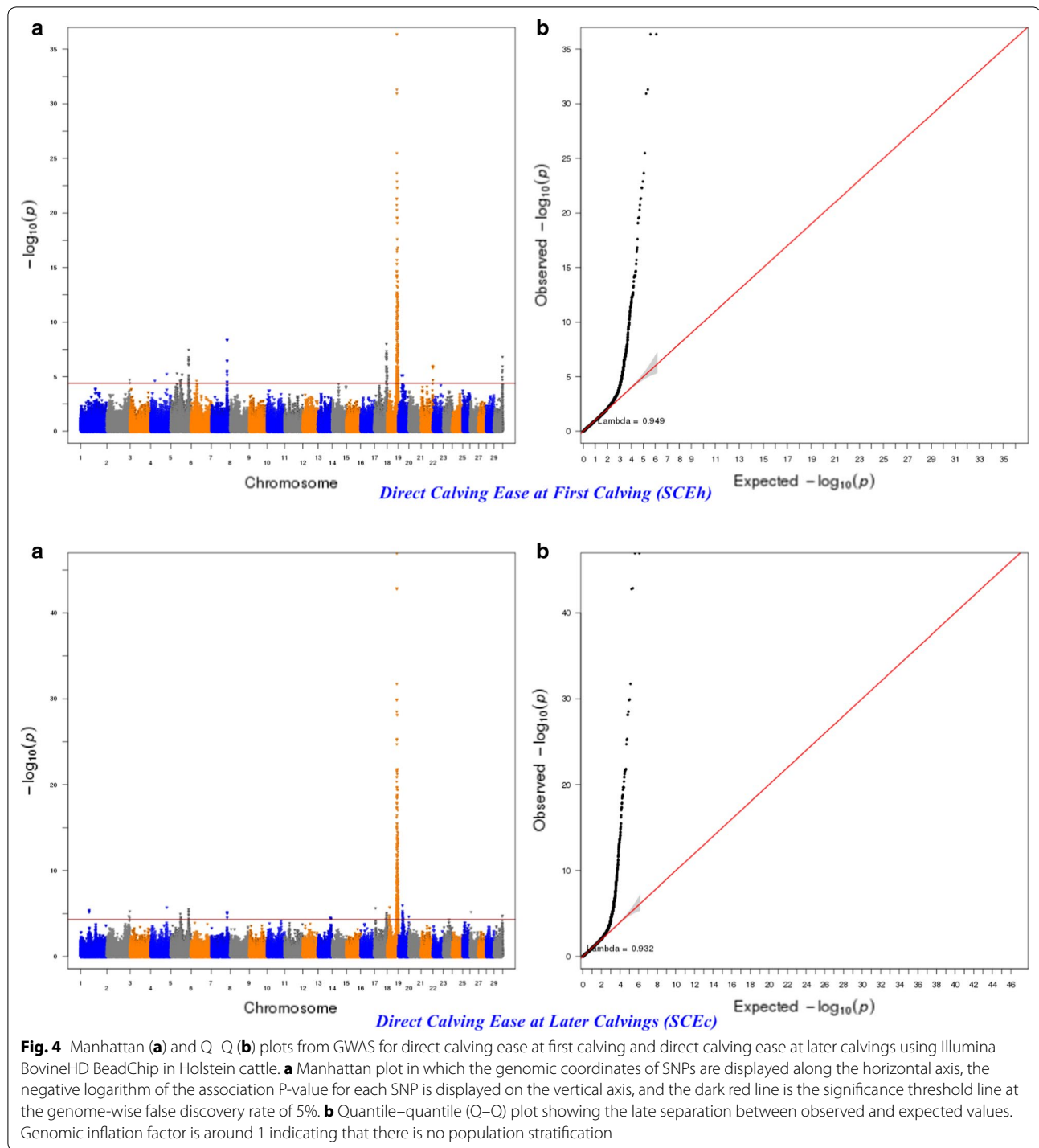
After applying QC criteria, 719,200 SNPs distributed over the 29 bovine autosomes in the reference population (2387 individuals) were used for genome-wide imputation of the target population (11,926 animals) for which 39,148 SNPs were available. The squared correlation

(r^2) between imputed genotypes and true genotypes of the target population was estimated at 99.29%, which indicated that the imputation was highly accurate. The allelic r^2 was used as a measure of imputation accuracy because it adequately reflects the imputation accuracy of SNPs with a low MAF [29]. In this study, the accuracy of



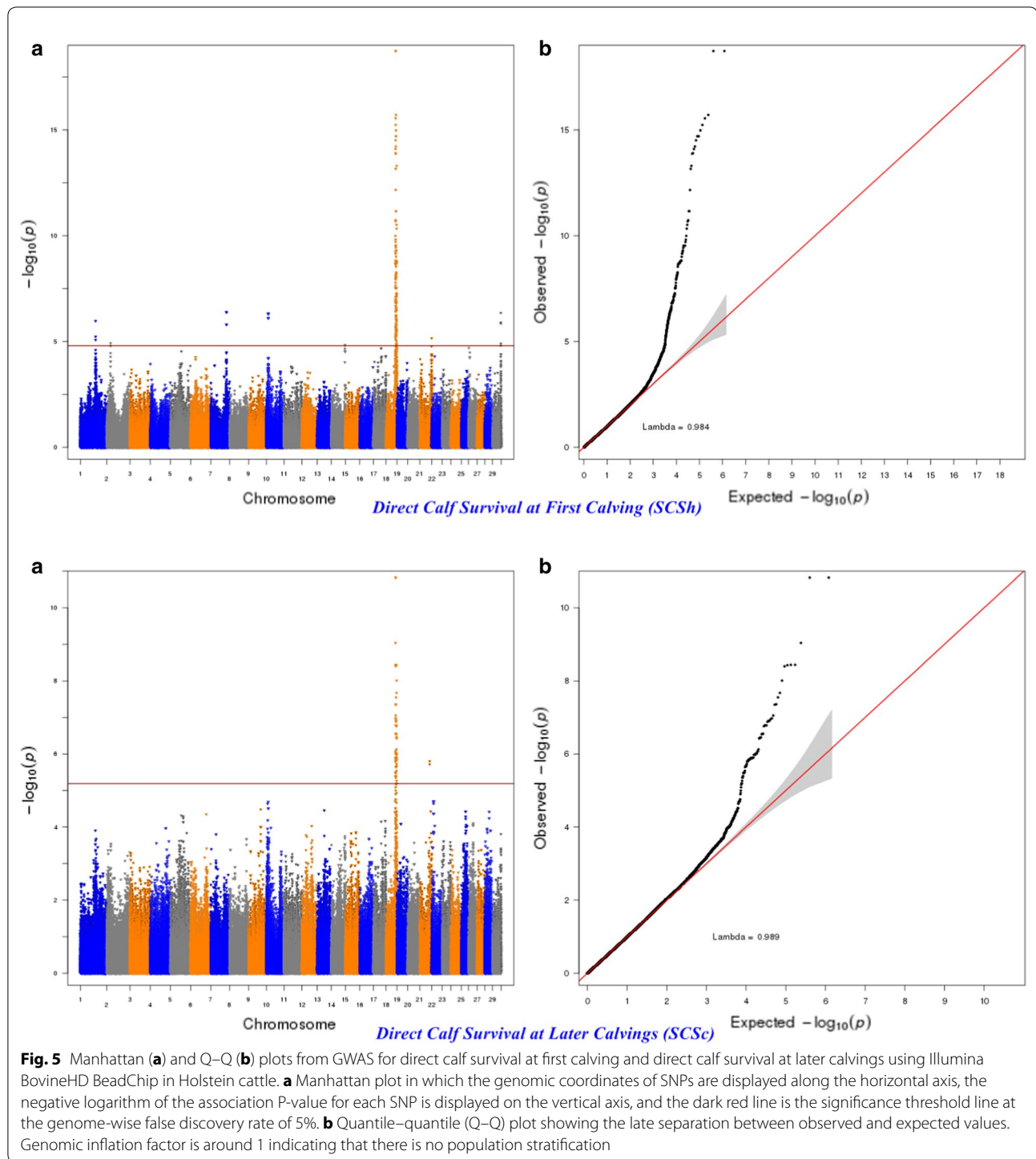
imputation was high for several reasons. First, the reference population with HD genotypes was relatively large ($n = 2387$) compared with other studies that reported high imputation accuracies by using the FImpute software and with less individuals in the reference group (e.g., 1733) than those used in our study [28]. In addition,

the accuracy of imputation from 50 K to HD using FImpute was also high (99.96%) in a Holstein reference population of limited size ($n = 1406$) [24]. Second, the reference population in our study shares relatively long-range haplotypes [28] with high phasing accuracy which also contributes to higher imputation accuracies [49].



Third, this reference population is also closely related to the target population which helped increase the accuracy of imputation, even for SNPs with a low MAF (≤ 0.05) [28]. Furthermore, an internal score based on the accuracy of imputation for each SNP allele was used to set less accurate imputed alleles to missing and the SNPs

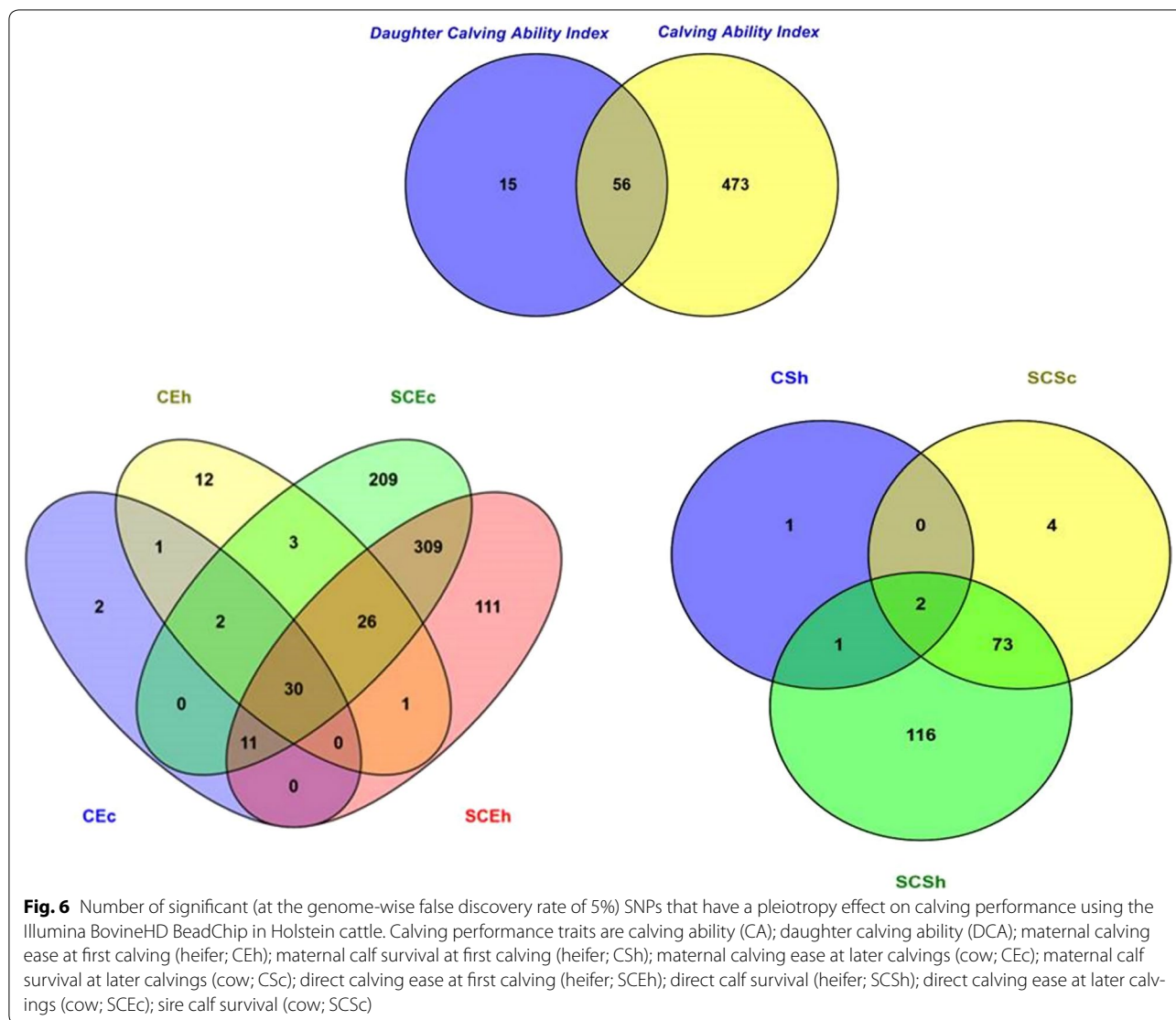
with a call rate lower than 90% were excluded in the QC before GWAS. In addition, SNPs with a MAF lower than 1% were initially excluded, which resulted in more accurate genotypes for the GWAS. The use of imputation increased both the numbers of SNPs and animals included in the association analysis, which increased the



power of QTL detection [50] and the accuracy of genomic prediction of breeding values [51, 52]. Combining SNP panels with different densities and using imputed genotypes in QTL mapping is a common approach in human GWAS [53] and in dairy cattle genomic evaluations [24, 54].

Genome-wide association for calving performance traits

In the GWAS, family and cryptic stratifications were accounted for by incorporating the full genomic covariance among individuals. The genomic inflation factor was used to assess bias in the test statistics. The average genomic inflation factor was equal to 0.98 and ranged



from 0.93 for SCEc to 1.00 for heel depth (see Additional file 4: Figure S4), which suggests that any potential bias due to population stratification was addressed [55–57]. Furthermore, accounting for the reliability of the de-regressed EBV allowed the use of more records in the estimation step by including de-regressed EBV with lower reliability, which subsequently improved the accuracy of GWAS [50, 52].

For calving performance, 782 SNPs significant at a genome-wide FDR of 5% were identified. Most of these significant SNPs were on BTA18 (71.87%), 17 (7.42%), 5 (6.78%), 7 (2.43%), 19 (1.92%), 21 (1.66%), 29 (1.66%), 1 (1.02%) and 10 (1.02%). The significant SNPs were mapped to 675 genes, among which 142 included at least one significant SNP and 532 were nearby (100 kbp) a SNP. A strong peak on BTA18 that affected maternal

and direct calving ease and calf survival in the first and later calvings was identified (Figs. 1, 2, 3, 4, 5). Previously, Sahana et al. [58] and Hoglund et al. [59] reported QTL on BTA18 that were associated with direct and maternal calving ease, calf size, stillbirth and birth index in Danish and Swedish Holstein cattle. In addition, Saatchi et al. [60] reported a QTL on BTA18 (at 54.37 Mb) that was associated with maternal calving ease in Simmental beef cattle. Also, the significant associations identified for most calving performance traits on BTA5, 6, 7, 9, 13, 17, 18, 19, 20 and 23 are in agreement with previous studies [58–60]. The significant associations were located within the confidence interval of previously detected QTL (see Additional file 5: Table S1). Thirty significant (FDR of 5%) SNPs were common between maternal and direct calving performance traits (Fig. 6).

Table 2 The 10 most significant SNPs associated with each calving performance trait at the genome-wide false discovery rate of 5%

Trait ^a	RefSNP (rs) ^b	c	Pos. (bp)	MAF	Estimated effect	SE	P-value	Location ^d	Gene name
CA	rs109478645	18	57589121	0.12	- 5.55	0.39	3.69E-41	5484 L	LOC100138951
CA	rs136283363	18	57548213	0.12	5.55	0.39	3.69E-41	5497L	LOC615876
CA	rs110801791	18	57516245	0.13	- 4.86	0.38	4.51E-35	1252 R	CTU1
CA	rs135253383	18	57520290	0.13	4.83	0.38	1.27E-34	Within	CTU1
CA	rs136514789	18	57486184	0.16	3.92	0.34	1.50E-28	1330 L	LOC540297
CA	rs109882115	18	58067310	0.28	- 3.15	0.29	1.71E-25	Within	ENSBTAG00000039491
CA	rs132734257	18	57511637	0.19	- 3.54	0.33	2.04E-25	5860 R	CTU1
CA	rs135508298	18	59602905	0.21	3.52	0.33	2.12E-25	75902 L	ENSBTAG00000048191
CA	rs134539659	18	59612379	0.21	- 3.52	0.33	2.12E-25		
CA	rs109540436	18	59890631	0.20	3.37	0.33	2.98E-23	3982 R	LOC100300095
CEc	rs110801791	18	57516245	0.13	- 3.93	0.49	2.72E-15	1252 R	CTU1
CEc	rs135253383	18	57520290	0.13	3.90	0.49	4.08E-15	Within	CTU1
CEc	rs109478645	18	57589121	0.12	- 3.86	0.51	7.21E-14	5484 L	LOC100138951
CEc	rs136283363	18	57548213	0.12	3.86	0.51	7.21E-14	5497L	LOC615876
CEc	rs137554975	18	56944300	0.15	- 2.98	0.45	4.04E-11	Within	MYH14
CEc	rs136113894	18	56959011	0.16	- 2.82	0.44	2.25E-10	Within	KCNC3
CEc	rs136514789	18	57486184	0.16	2.79	0.44	5.19E-10	1330 L	LOC540297
CEc	rs133544315	18	60019327	0.21	2.61	0.42	7.62E-10	47852 L	ZNF836
CEc	rs109540436	18	59890631	0.20	2.45	0.42	9.69E-09	3982 R	LOC100300095
CEc	rs135033267	18	60043680	0.20	- 2.45	0.43	1.13E-08	25523 R	ENSBTAG00000011844
CEh	rs110801791	18	57516245	0.13	- 3.73	0.46	1.33E-15	1252 R	CTU1
CEh	rs135253383	18	57520290	0.13	3.69	0.46	2.77E-15	Within	CTU1
CEh	rs109478645	18	57589121	0.12	- 3.73	0.48	1.92E-14	5484 L	LOC100138951
CEh	rs136283363	18	57548213	0.12	3.73	0.48	1.92E-14	5497L	LOC615876
CEh	rs133544315	18	60019327	0.21	2.45	0.40	9.84E-10	47852 L	ZNF836
CEh	rs109540436	18	59890631	0.20	2.41	0.40	2.17E-09	3982 R	LOC100300095
CEh	rs135033267	18	60043680	0.20	- 2.41	0.40	2.69E-09	25523 R	ENSBTAG00000011844
CEh	rs137436636	18	60006650	0.20	2.41	0.40	2.69E-09	35175 L	ZNF836
CEh	rs136514789	18	57486184	0.16	2.49	0.42	3.36E-09	1330 L	LOC540297
CEh	rs133144614	18	59242260	0.41	- 1.99	0.34	4.68E-09	13197 L	ZNF845
CSh	rs43264905	1	118222449	0.04	4.91	0.85	9.51E-09	15187 L	ENSBTAG00000038069
DCA	rs110801791	18	57516245	0.13	- 4.27	0.57	1.26E-13	1252 R	CTU1
DCA	rs135253383	18	57520290	0.13	4.23	0.57	2.06E-13	Within	CTU1
DCA	rs109478645	18	57589121	0.12	- 4.25	0.59	1.59E-12	5484 L	LOC100138951
DCA	rs136283363	18	57548213	0.12	4.25	0.59	1.59E-12	5497L	LOC615876
DCA	rs136514789	18	57486184	0.16	3.18	0.51	8.24E-10	1330 L	LOC540297
DCA	rs41601357	18	58800104	0.49	2.33	0.40	5.86E-09	7804 L	LOC101904049
DCA	rs133144614	18	59242260	0.41	- 2.41	0.42	8.66E-09	13197 L	ZNF845
DCA	rs109011289	18	59058582	0.48	- 2.27	0.40	1.99E-08	7587 L	LOC101904371
DCA	rs41582522	18	58666276	0.41	- 2.33	0.41	2.47E-08		
DCA	rs133544315	18	60019327	0.21	2.76	0.49	2.55E-08	47852 L	ZNF836
SCEc	rs109478645	18	57589121	0.12	- 6.19	0.40	1.07E-47	5484 L	LOC100138951
SCEc	rs136283363	18	57548213	0.12	6.19	0.40	1.07E-47	5497L	LOC615876
SCEc	rs110801791	18	57516245	0.13	- 5.66	0.39	1.36E-43	1252 R	CTU1
SCEc	rs135253383	18	57520290	0.13	5.65	0.39	1.66E-43	Within	CTU1
SCEc	rs136514789	18	57486184	0.16	4.38	0.36	1.81E-32	1330 L	LOC540297
SCEc	rs132734257	18	57511637	0.19	- 4.05	0.34	1.21E-30	5860 R	CTU1
SCEc	rs137554975	18	56944300	0.15	- 4.21	0.35	1.42E-30	Within	MYH14

Table 2 continued

Trait ^a	RefSNP (rs) ^b	c	Pos. (bp)	MAF	Estimated effect	SE	P-value	Location ^d	Gene name
SCEc	rs136113894	18	56959011	0.16	- 4.07	0.35	3.22E-29	Within	KCNC3
SCEc	rs135508298	18	59602905	0.21	3.90	0.34	7.48E-29	75902 L	ENSBTAG00000048191
SCEc	rs134539659	18	59612379	0.21	- 3.90	0.34	7.48E-29		
SCEh	rs109478645	18	57589121	0.12	- 6.13	0.46	4.12E-37	5484 L	LOC100138951
SCEh	rs136283363	18	57548213	0.12	6.13	0.46	4.12E-37	5497L	LOC615876
SCEh	rs110801791	18	57516245	0.13	- 5.42	0.44	4.98E-32	1252 R	CTU1
SCEh	rs135253383	18	57520290	0.13	5.38	0.44	1.17E-31	Within	CTU1
SCEh	rs136514789	18	57486184	0.16	4.39	0.40	3.24E-26	1330 L	LOC540297
SCEh	rs132734257	18	57511637	0.19	- 4.03	0.39	2.30E-24	5860 R	CTU1
SCEh	rs109882115	18	58067310	0.28	- 3.53	0.34	1.33E-23	Within	ENSBTAG00000039491
SCEh	rs135508298	18	59602905	0.21	3.90	0.39	5.03E-23	75902 L	ENSBTAG00000048191
SCEh	rs134539659	18	59612379	0.21	- 3.90	0.39	5.03E-23		
SCEh	rs43038601	18	57491196	0.22	- 3.59	0.36	4.60E-22	4128 R	KLK14
SCSc	rs109478645	18	57589121	0.12	- 5.59	0.82	1.50E-11	5484 L	LOC100138951
SCSc	rs136283363	18	57548213	0.12	5.59	0.82	1.50E-11	5497L	LOC615876
SCSc	rs136113894	18	56959011	0.16	- 4.41	0.71	9.17E-10	Within	KCNC3
SCSc	rs110801791	18	57516245	0.13	- 4.73	0.79	3.63E-09	1252 R	CTU1
SCSc	rs135253383	18	57520290	0.13	4.73	0.79	3.63E-09	Within	CTU1
SCSc	rs134066287	18	62844651	0.23	3.84	0.65	3.70E-09	Within	FCAR
SCSc	rs109882115	18	58067310	0.28	- 3.56	0.60	3.96E-09	Within	ENSBTAG00000039491
SCSc	rs109844326	18	63415091	0.16	4.19	0.72	9.77E-09	2950 L	MBOAT7
SCSc	rs134353106	18	62839617	0.26	3.47	0.61	2.13E-08	5900 L	KIR2DL5A
SCSc	rs133544315	18	60019327	0.21	3.74	0.67	2.82E-08	47852 L	ZNF836
SCSh	rs109478645	18	57589121	0.12	- 7.95	0.86	1.86E-19	5484 L	LOC100138951
SCSh	rs136283363	18	57548213	0.12	7.95	0.86	1.86E-19	5497L	LOC615876
SCSh	rs133544315	18	60019327	0.21	5.91	0.71	1.97E-16	47852 L	ZNF836
SCSh	rs110801791	18	57516245	0.13	- 6.94	0.83	2.84E-16	1252 R	CTU1
SCSh	rs135253383	18	57520290	0.13	6.86	0.83	5.87E-16	Within	CTU1
SCSh	rs109540436	18	59890631	0.20	5.76	0.71	1.05E-15	3982 R	LOC100300095
SCSh	rs135033267	18	60043680	0.20	- 5.72	0.71	1.98E-15	25523 R	ENSBTAG00000011844
SCSh	rs137436636	18	60006650	0.20	5.72	0.71	1.98E-15	35175 L	ZNF836
SCSh	rs136514789	18	57486184	0.16	6.02	0.75	3.07E-15	1330 L	LOC540297
SCSh	rs136766893	18	58494682	0.26	- 5.02	0.63	6.02E-15	Within	BOSTAUV1R416

^a Traits = evaluated calving traits included calving ability (CA), daughter calving ability (DCA), maternal calving ease at first calving (heifer; CEh), maternal calf survival at first calving (heifer; CSh), maternal calving ease at later calvings (cow; CEc), maternal calf survival at later calvings (cow; CSc), direct calving ease at first calving (heifer; SCEh), direct calf survival at first calving (heifer; SCSH), direct calving ease at later calvings (cow; SCEc), direct calf survival at later calvings (cow; SCSc) and sire conception rate (SCR). Calving performance traits are expressed as relative breeding values with a population average of 100 and standard deviation of 5

^b RefSNP (rs) = assigned reference SNP (rs) number for a SNP by the National Center for Biotechnology Information

^c BTA = *Bos taurus* autosomes

^d Location = gene location (within L left or R right) from the SNP of interest

Several genes that contain significant SNPs for calving performance traits were identified (Table 2) and are involved in biological processes such as lipid metabolism, immunity, reproduction and anatomical structure development. For example, *kallikrein-related peptidase 14 (KLK14)* is involved in mating and reproduction processes in a multicellular organism [61, 62]. *KLK14* was reported to play an important role in the synergistic effects between estrogens and androgens [61] and

at the onset of parturition [62]. Another important gene was *killer cell immunoglobulin like receptor, two Ig domains and long cytoplasmic tail 5A (KIR2DL5A)*, which is attributed to Graft-versus-host disease and natural killer cell mediated cytotoxicity [63]. The *membrane bound O-acyltransferase domain containing 7 (MBOAT7)* gene is involved in the metabolism of lipids and lipoproteins, including the pathway of glycerophospholipids metabolism, which is linked to energy balance

Table 3 The most highly significant SNPs associated with each body conformation trait at the genome-wide false discovery rate of 5%

Trait ^a	RefSNP (rs) ^b	BTA ^c	Pos. (bp)	MAF	Estimated effect	SE	P-value	Location ^d	Gene name
ANG	rs110434046	6	88919352	0.46829	- 2.46	0.40	1.35E-09		
ANG	rs137712965	6	88922396	0.46829	2.46	0.40	1.35E-09		
ANG	rs136287830	6	88865430	0.46168	- 2.43	0.40	2.21E-09		
ANG	rs108983037	6	88913092	0.46499	2.38	0.40	4.01E-09		
ANG	rs109512265	6	88485244	0.40611	- 2.19	0.38	1.51E-08	Within	SLC4A4
BD	rs109478645	18	57589121	0.1195	4.49	0.45	2.29E-22	Within	ENSBTAG00000037537
BD	rs136283363	18	57548213	0.1195	- 4.49	0.45	2.29E-22		
BD	rs110801791	18	57516245	0.12539	4.00	0.43	9.73E-20	1252 R	CTU1
BD	rs135253383	18	57520290	0.12549	- 3.99	0.43	1.03E-19	Within	CTU1
BD	rs132734257	18	57511637	0.19128	3.13	0.37	1.18E-16		
BQ	rs109901274	7	93244933	0.11506	- 2.81	0.43	9.10E-11	Within	ARRDC3
BQ	rs109618368	7	93254737	0.11506	2.81	0.43	9.10E-11	1643 L	ARRDC3
BQ	rs110680622	7	93287387	0.11506	- 2.81	0.43	9.10E-11		
BQ	rs110066813	7	93251138	0.13045	- 2.52	0.39	2.59E-10	Within	ARRDC3
BQ	rs109860522	7	93263102	0.13045	2.52	0.39	2.59E-10		
CW	rs109901274	7	93244933	0.11506	3.10	0.51	1.47E-09	Within	ARRDC3
CW	rs109618368	7	93254737	0.11506	- 3.10	0.51	1.47E-09	1643 L	ARRDC3
CW	rs110680622	7	93287387	0.11506	3.10	0.51	1.47E-09		
CW	rs110434046	6	88919352	0.46829	2.22	0.41	5.90E-08		
CW	rs137712965	6	88922396	0.46829	- 2.22	0.41	5.90E-08		
DS	rs109478645	18	57589121	0.1195	2.86	0.45	4.48E-10	Within	ENSBTAG00000037537
DS	rs136283363	18	57548213	0.1195	- 2.86	0.45	4.48E-10		
DS	rs136514789	18	57486184	0.15844	- 2.29	0.39	5.21E-09	1330 L	LOC540297
DS	rs110801791	18	57516245	0.12539	2.39	0.44	5.02E-08	1252 R	CTU1
DS	rs134321289	18	58053410	0.31698	- 1.73	0.32	5.49E-08		
FA	rs109517032	20	26955400	0.26761	2.31	0.40	1.11E-08		
FA	rs110521341	20	26957866	0.26761	- 2.31	0.40	1.11E-08		
FA	rs109222733	20	26960799	0.26761	- 2.31	0.40	1.11E-08		
FA	rs109747498	20	26978980	0.26761	- 2.31	0.40	1.11E-08		
FA	rs110486797	20	26990879	0.26761	- 2.31	0.40	1.11E-08		
FTP	rs133797664	8	79527489	0.1322	2.55	0.49	2.73E-07		
FTP	rs136251125	18	42320881	0.29168	1.67	0.32	3.28E-07		
FTP	rs136974352	18	42322043	0.29168	1.67	0.32	3.28E-07		
FTP	rs43571286	8	79497558	0.13468	2.42	0.47	3.66E-07		
FTP	rs137127919	5	12712534	0.06083	2.71	0.53	3.84E-07		
PS	rs42244554	5	1278211	0.3427	2.00	0.36	2.35E-08	694 L	ZFC3H1
PW	rs137169221	13	4626355	0.11558	- 2.73	0.46	5.63E-09		
PW	rs109478645	18	57589121	0.1195	2.53	0.43	6.48E-09	Within	ENSBTAG00000037537
PW	rs136283363	18	57548213	0.1195	- 2.53	0.43	6.48E-09		
PW	rs42382059	13	4518102	0.1165	- 2.60	0.46	2.12E-08		
PW	rs42382046	13	4524429	0.1165	2.60	0.46	2.12E-08		
RAH	rs109901274	7	93244933	0.11506	- 3.10	0.49	2.88E-10	Within	ARRDC3
RAH	rs109618368	7	93254737	0.11506	3.10	0.49	2.88E-10	1643 L	ARRDC3
RAH	rs110680622	7	93287387	0.11506	- 3.10	0.49	2.88E-10		
RAH	rs109343241	18	59615343	0.27577	1.93	0.35	4.08E-08		
RAH	rs135508298	18	59602905	0.20564	2.12	0.39	9.37E-08		
ST	rs109882115	18	58067310	0.2799	1.86	0.30	1.19E-09	Within	ENSBTAG00000039491
ST	rs109478645	18	57589121	0.1195	2.53	0.41	1.22E-09	Within	ENSBTAG00000037537

Table 3 continued

Trait ^a	RefSNP (rs) ^b	BTA ^c	Pos. (bp)	MAF	Estimated effect	SE	P-value	Location ^d	Gene name
ST	rs136283363	18	57548213	0.1195	- 2.53	0.41	1.22E-09		
ST	rs133960300	5	106250561	0.14439	- 2.34	0.39	2.94E-09	3346 R	CCND2
ST	rs109685956	5	106256616	0.14439	- 2.34	0.39	2.94E-09	Within	CCND2
TL	rs110895486	5	12438670	0.21287	3.00	0.41	4.27E-13		
TL	rs109194156	5	12444024	0.21318	2.99	0.41	4.92E-13		
TL	rs110137797	5	12464607	0.21328	2.93	0.41	1.63E-12	Within	TMTC2
TL	rs109264225	5	12468154	0.21328	2.93	0.41	1.63E-12	Within	TMTC2
TL	rs43435142	5	12468826	0.21328	2.93	0.41	1.63E-12	Within	TMTC2
UD	rs42192551	29	49182244	0.46055	- 1.48	0.27	6.23E-08		
UD	rs110432804	6	89036570	0.46932	- 2.00	0.37	8.93E-08		
UD	rs108983037	6	88913092	0.46499	- 2.00	0.37	9.11E-08		
UD	rs110434046	6	88919352	0.46829	2.00	0.37	1.04E-07		
UD	rs137712965	6	88922396	0.46829	- 2.00	0.37	1.04E-07		

Conformation trait EBV are standardized to have a mean value of zero and a standard deviation of 5

^a Trait = body depth (BD), bone quality (BQ), chest width (CW), pin width (PW), rear attachment height (RAH), stature (ST), angularity (ANG), body depth (BD), dairy strength (DS), fore udder attachment (FA), front teat placement (FTP), pin setting (PS), rear attachment height (RAH), teat length (TL)

^b RefSNP (rs) = Assigned Reference SNP (rs) number for a SNP by the National Center for Biotechnology Information

^c BTA = *Bos taurus* autosomes

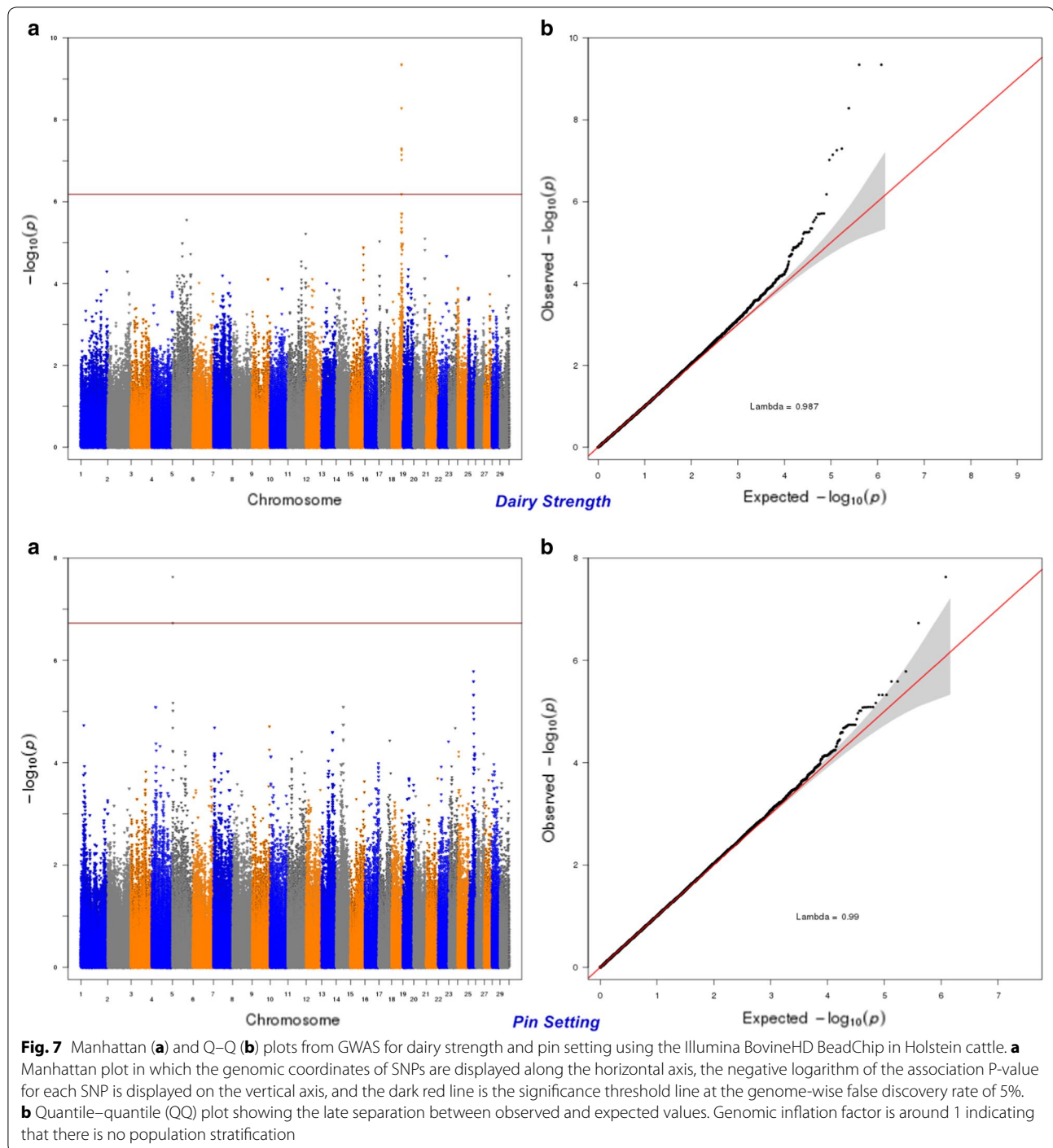
^d Location = gene location (within L left or R right) from the SNP of interest

metabolites including non-esterified fatty acids (NEFA), beta-hydroxybutyrate (BHBA) and glucose in animals [64]. It is well documented that the period around calving is associated with several biological processes related to fat, protein, glucose and mineral metabolism as well as complex hormonal changes [65]. Monitoring and nutritional modeling of the animal's energy balance to prevent negative energy balance during the transition period of dairy cows is an important management practice that can reduce the incidence of metabolic and infectious diseases [66]. Energy balance could be associated with the length of the dry period and subsequently with calving difficulty which is generally common under a short (< 60 days) compared to a long dry period [67].

In this study, we confirmed the relationship between the protective functions of the immune system and calving ease and survival by the detection of several genes such as *Fc fragment of IgA receptor (FCAR)*. *FCAR* includes a significant SNP (rs134066287) which is associated with calving ease and survival at first and later calving (Table 2). *FCAR* is also involved in several immune defense processes such as phagocytosis and *Staphylococcus aureus* infection pathways [68]. The expression profile of fetal membranes, myometrium and cervix tissues showed that inflammatory response was associated with labor [69].

Genome-wide association for body conformation traits

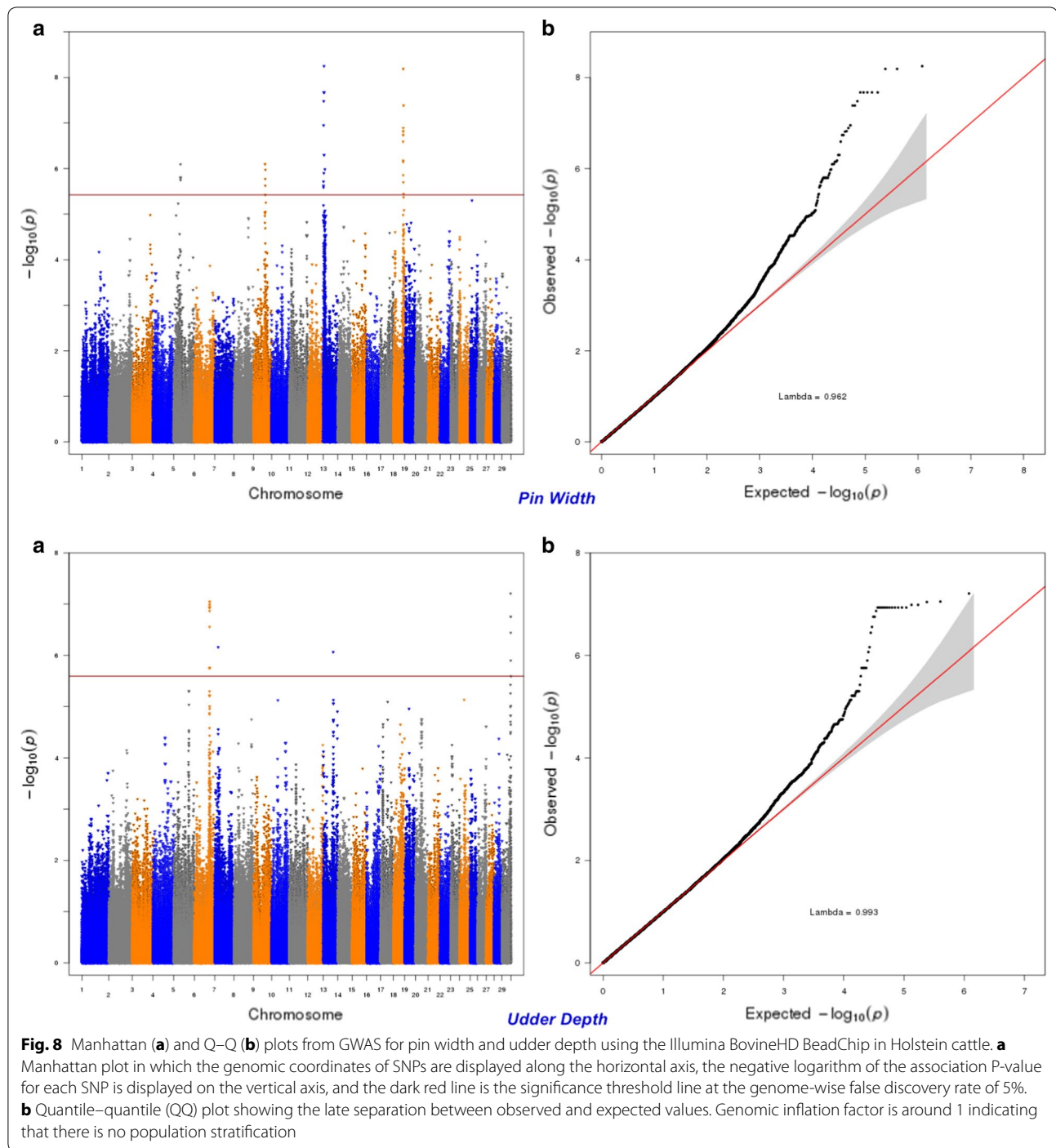
Six hundred and seven SNPs were statistically significant at a genome-wide FDR of 5% for body conformation traits (Table 3). These significant SNPs were mapped to 553 genes, among which 89 genes included at least one significant SNP and 464 were nearby (100 kbp) one. Peaks indicating association were found on BTA5, 6, 7, 13, 14, 18, and 20 (Figs. 7, 8, 9, 10, 11, 12, 13). Four significant SNPs (FDR of 5%) overlapped between bone quality and pin width (Fig. 14). For example, the *arresting domain-containing 3 (ARRDC3)* gene on BTA7 was associated with body conformation (e.g., bone quality and chest width) and calving performance traits. The *ARRDC3* gene was previously reported as a candidate gene with a pleiotropic effect on birth and weaning weights, direct and maternal calving ease and carcass traits in beef cattle [60]. Furthermore, this study identified other significant polymorphisms for body depth within the *carnitine palmitoyl transferase 1C (CPT1C)* gene, which is involved in PPAR signaling, adipocytokine signaling and fatty acid metabolism pathways linked to the animal's energy balance [70]. Interestingly, we identified a SNP within the *diacylglycerol O-acyltransferase 1 (DGAT1)* gene that was significantly associated with bone quality. Bone quality is an important trait that affects functional longevity of



Holstein cows [9]. Kaupé et al. [71] identified a SNP within *DGATI* that is associated with rump width and strength, which suggests a possible association between *DGATI* and functional longevity.

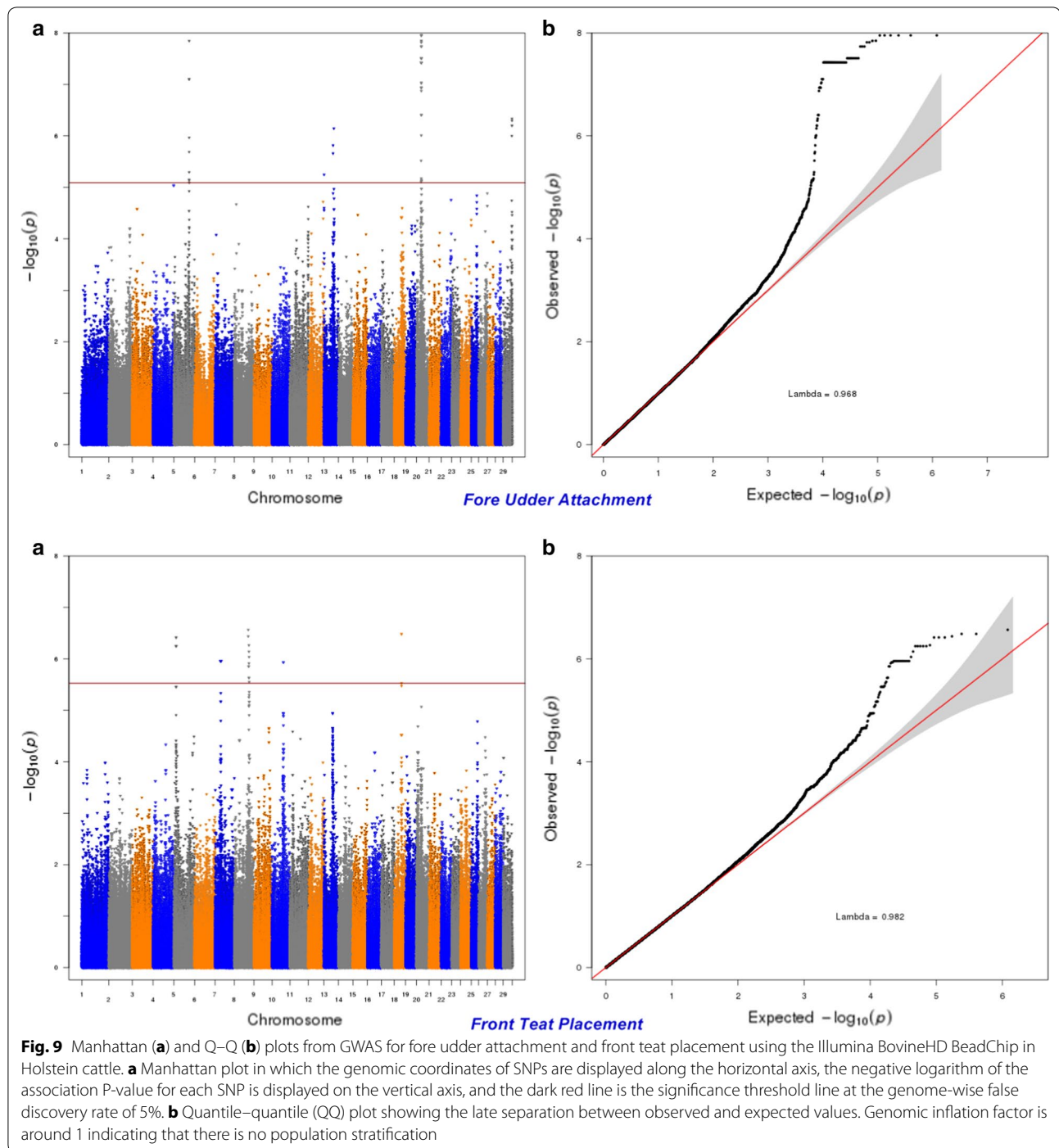
Pleiotropic polymorphisms for calving performance and body conformation traits

One hundred and eighty-three SNPs were associated with calving and conformation traits. Sixteen SNPs on



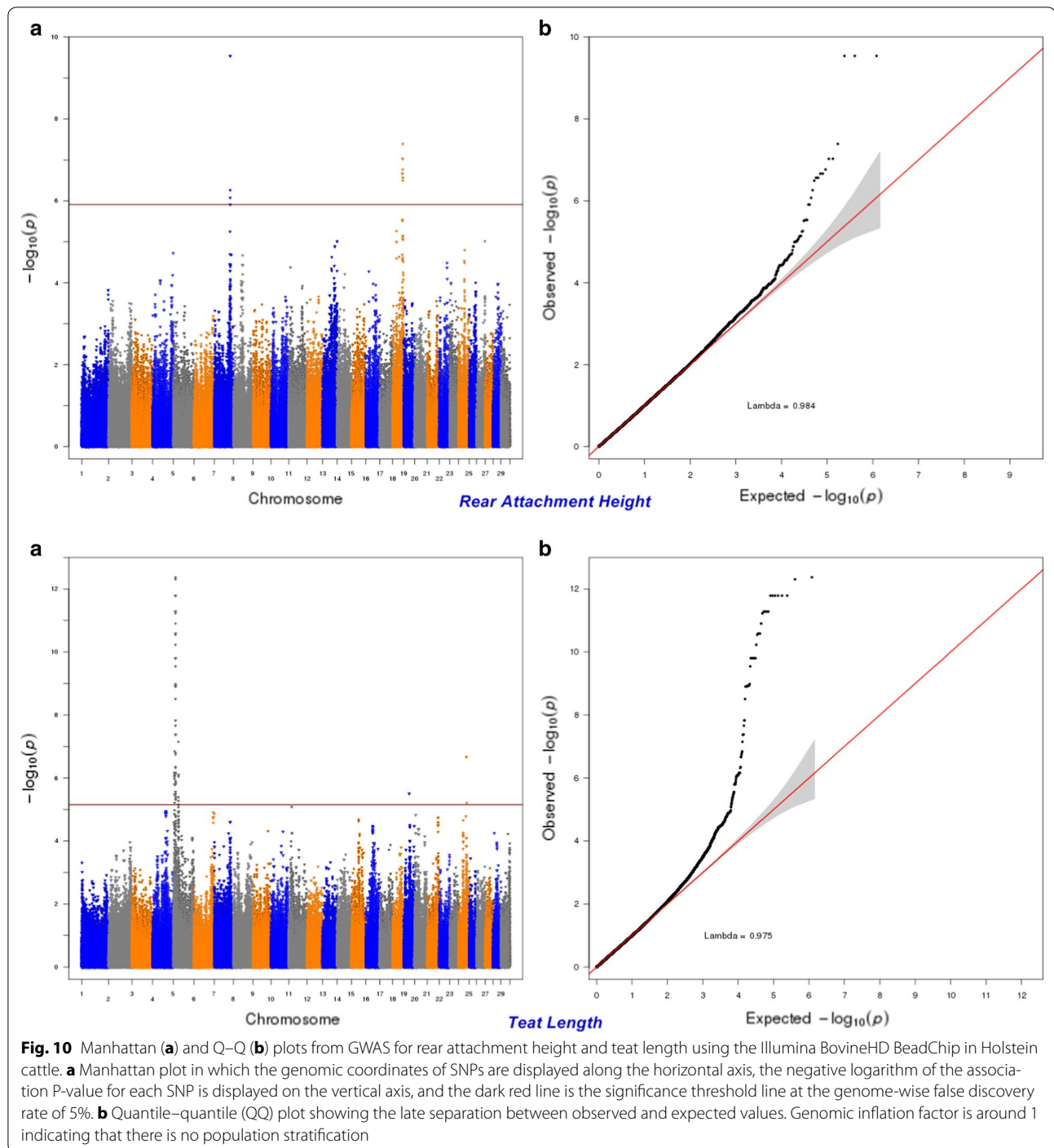
BTA18 in the region between 56.9 and 59.9 Mbp were associated with calving performance and rump traits particularly pin width (Fig. 15). This region harbours the *SH3 and multiple ankyrin repeat domains 1 (SHANK1)* gene, the *myosin heavy chain 14 (MYH14)* gene, and the *cytosolic thiouridylase subunit 1 (CTUI)* gene, which are involved in tight junction, viral myocarditis, regulation

of actin cytoskeleton, glutamatergic synapse, and sulfur relay system pathways [72]. The *MYH14* gene is associated with calving and body conformation and with one of the most significant SNPs (rs137554975) identified for eight calving traits. *MYH14* is also involved in the regulation of actin cytoskeleton and tight junction pathways which are critical for myometrial functions during



parturition [73]. The pleiotropic associations between calving performance and rump traits support the known genetic correlation between characteristics of the pelvis area (e.g., rump traits) and calving ease and calf survival [8]. A cow with a wide pin, long sloping rump, and slight

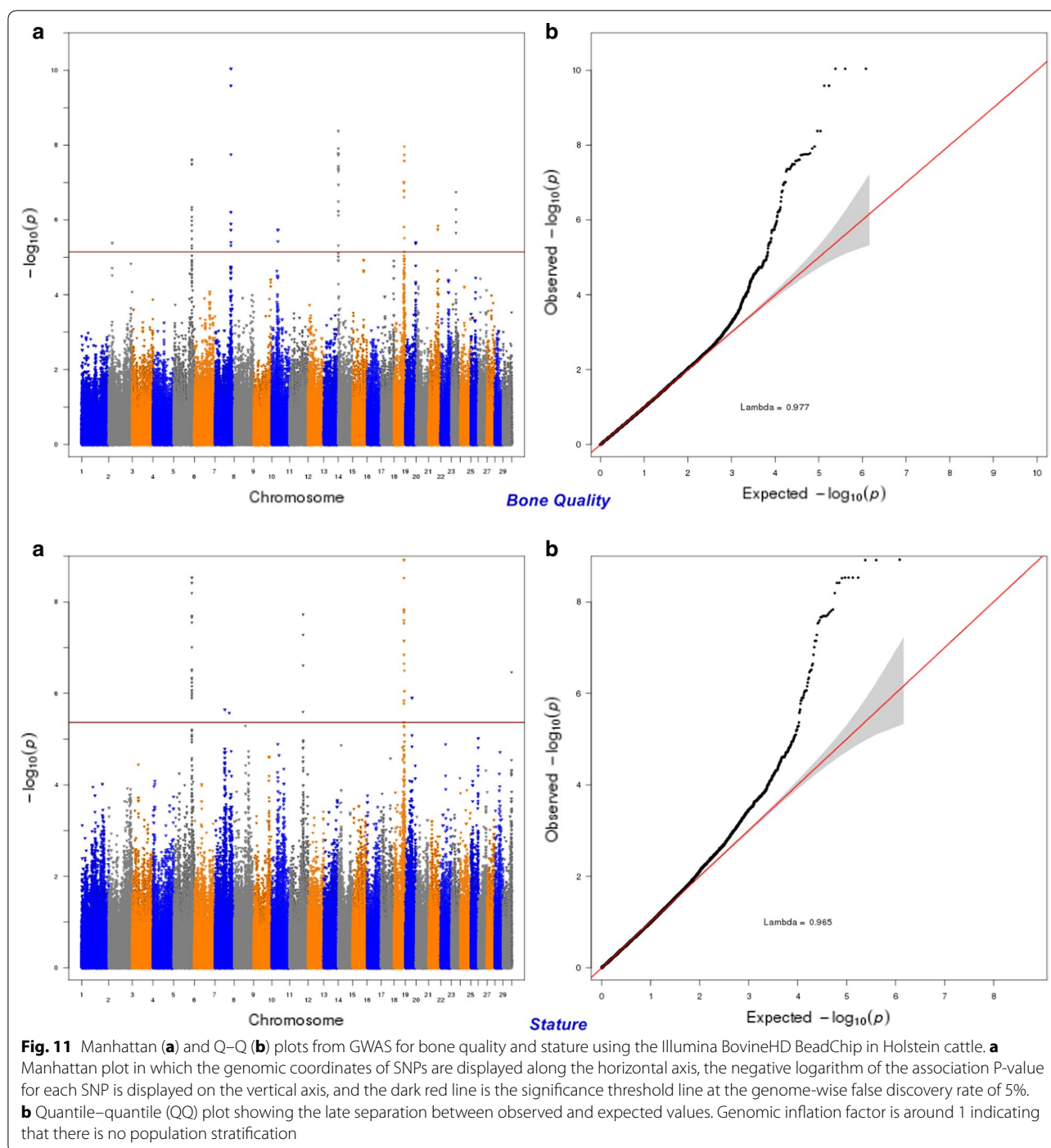
slope from pin bone to thurl is known to be able to calve easily [8, 74]. An unfavorable genetic association between height of pin bones and slope to the cow's birth canal was observed, which suggests that higher pin bones may result in a tight birth canal causing calving difficulty [75].



Functional consequences of significant SNPs and in silico functional analyses

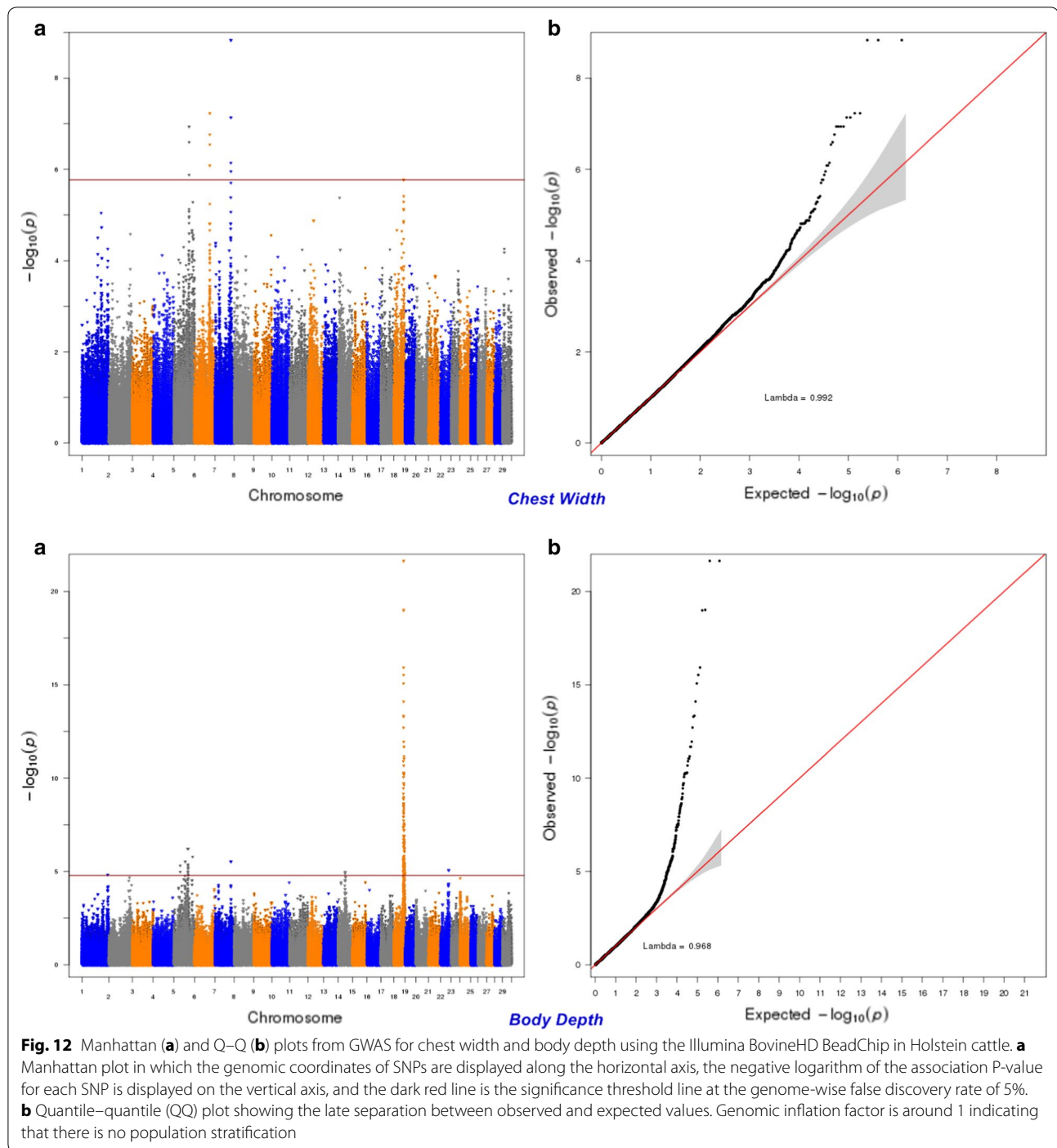
Most of the identified SNPs for calving and/or body conformation traits were intergenic (44.4%), intronic (31%), downstream (11.4%), or upstream (9.8%) of a gene. As shown in Table 4, we identified several SNPs with high functional interest (e.g., non-synonymous coding and

splice site intronic variants). For instance, a stop-lost mutation was identified within the *HORMA domain containing 2* (*HORMAD2*) gene and was significantly associated with CA, SCEc and SCEh. *HORMAD2* is involved in the M-phase of the mitotic cell cycle and, in humans, a polymorphism within this gene is known to cause meiotic arrest leading to human azoospermia [76].



The SNP enrichment analyses provided potential GO terms for calving and conformation traits including biological mechanisms, molecular function and cellular component (see Additional file 6: Table S2). Biological processes including musculoskeletal movement (GO:0050881), meiotic cell cycle (GO:0051321), oocyte maturation (GO:0001556), protein localization

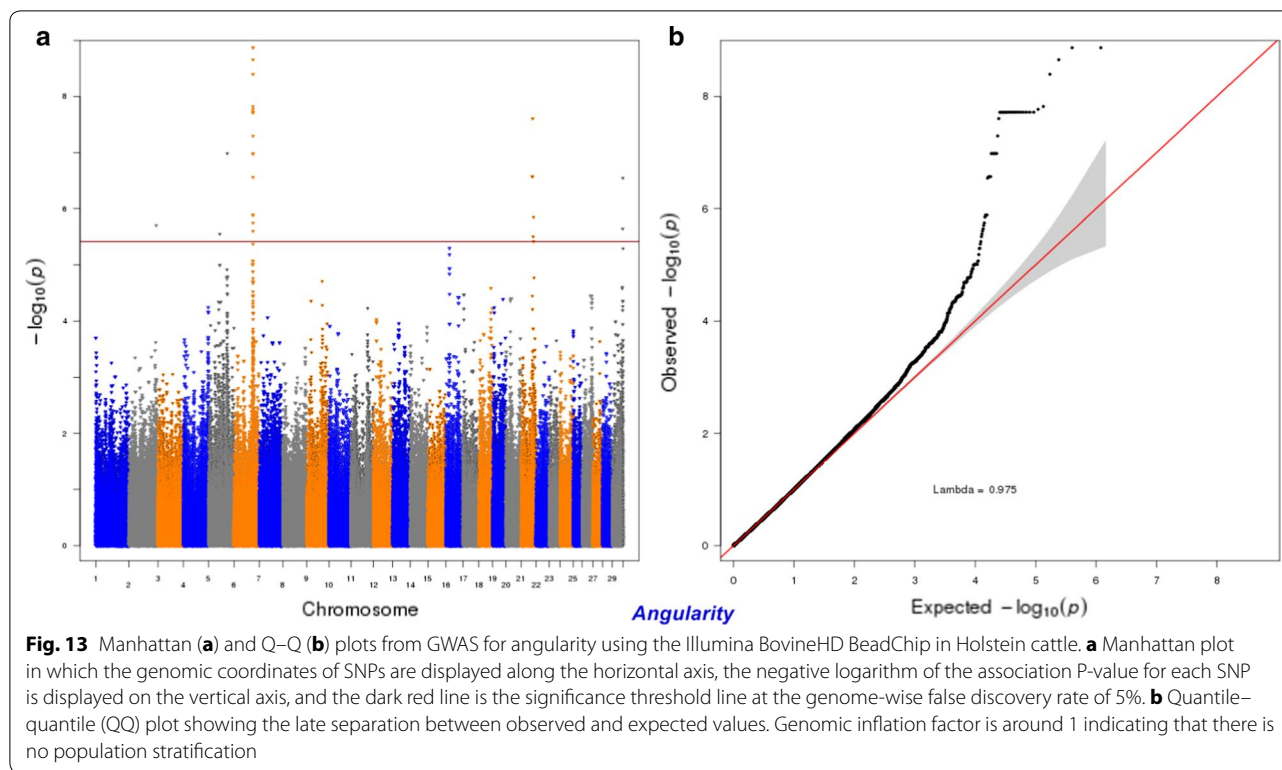
to synapse (GO:0035418) and skeletal muscle contraction (GO:0003009) were enriched (FDR of 1%) for calving ease and survival traits. Also, biological processes such as steroid dehydrogenase activity (GO:0016229) and 3-beta-hydroxy-delta5-steroid dehydrogenase activity (GO:0016229) were over-represented (FDR of 1%) for calving ease in heifers. The steroids are very important



for the control of the synthesis of oxytocin receptors and can facilitate the initiation of parturition [62].

Thirty-three genes were enriched ($P < 0.10$) in 13 pathways, in which tight junction, oxytocin signaling and MAPK signaling are associated with calving performance traits (Table 5). The tight junction proteins are highly expressed in the final stages of cervical ripening and

dilation in preparation for parturition [76, 77]. Furthermore, mitogen-activated protein kinase (MAPK) signaling pathway was enriched ($P = 0.065$) based on six genes. MAPK signaling cascades are critical in the regulation of several mechanisms including uterine contractility and myometrial cell proliferation [73]. In agreement with our findings, previous GWAS in Angus and Hereford cattle



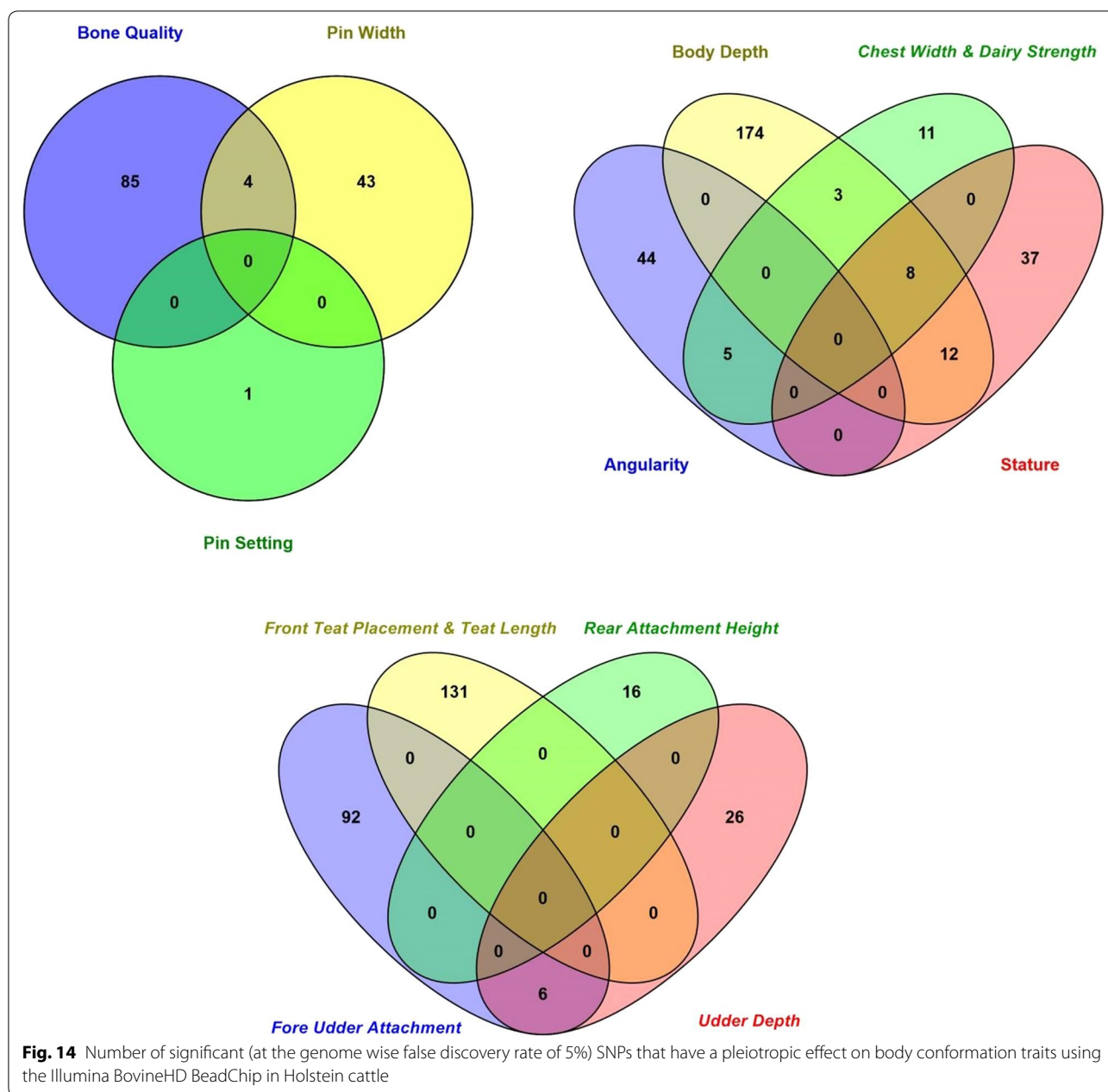
detected the MAPK signaling pathway as having a pleiotropic effect on birth weight, calving ease (direct and maternal) [60]. In addition, oxytocin signaling pathway which comprises the main drivers of calving [77] was also enriched ($P < 0.05$) for five genes (see Additional file 7: Figure S5).

Validation of genomic predictions

In this study, there was no substantial difference between prediction accuracies of the DGV from the 50 K or HD panels for young bulls (Fig. 16). This finding is in agreement with the study of Erbe et al. [23], in which the use of 624,213 SNPs provided no extra gain in accuracy than the 50 K panel when using GBLUP in Jersey and Holstein cattle. The reason why the use of both panels yielded similar results may be that the imputed HD panel captures the same loci with a large effect than those captured by the 50 K panel and that other QTL with a small effect captured by the imputed HD panel explain just a small proportion of the genetic variance and, consequently, result in only a small increase in accuracy of prediction of DBV [24]. A slight improvement in the accuracy of DGV when combining the 50 K panel with the significant SNPs

or the SNPs located nearby genes compared to using only the SNPs from the 50 K panel may indicate that the HD SNP panel contains SNPs that are potentially linked to causal mutations. However, a decrease of about 6.5% in prediction accuracy on average for calving traits was observed when only significant SNPs (1206) were used compared to the 50 K panel. Nonetheless, the small number of significant SNPs (1206) with a prediction accuracy up to 0.34 supported the informativeness of the candidate genes and the biological mechanisms associated with the traits.

In general, the validation of significantly associated SNPs in a different population is a necessary step before application in breeding programs in order to estimate the accuracy of the DGV properly and to test for potentially spurious associations in the original findings. The calving ability index and direct calving ease and survival traits were not evaluated in the validation analyses because of the limited number of young bulls. The GBLUP method was used because it is the official method for genomic evaluations in Canadian dairy cattle [35], it was recommended by earlier studies [78, 79], and has been shown to be as accurate as other statistical methods [31]. The



current study also identified several SNPs with pleiotropic effects, which are associated with known biological pathways involved in calving performance and conformation traits. The correlations between the breeding values of calving performance ranged from 0.11 to 0.82 [80], which suggests that common genes control more than one calving trait, as observed in this study.

Conclusions

We identified various SNPs that are significantly associated with calving performance on chromosomes 5, 18, and 19 and are located within or nearby genes with potential roles in tight junction, oxytocin signaling, and MAPK signaling. Sixteen SNPs within or nearby the *SHANK1*, *MYH14*, and *CTU1* genes on BTA18 showed

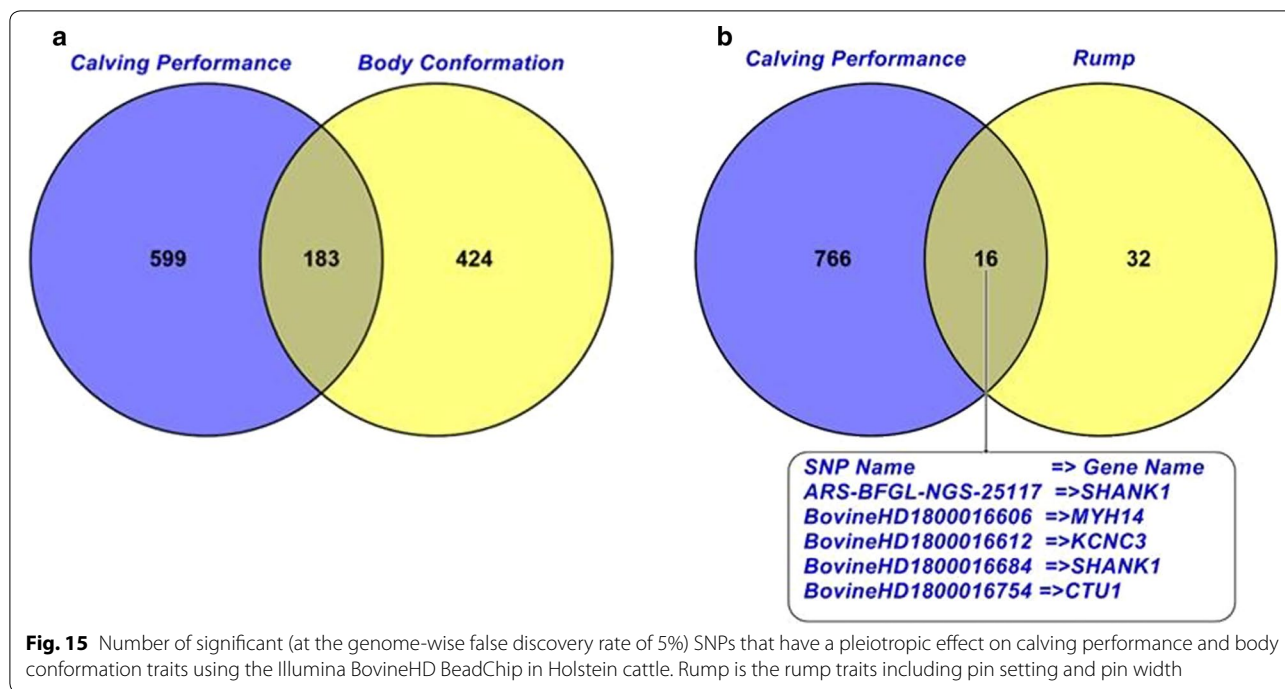


Table 4 Significant SNPs and their genes associated with calving and conformation traits in Holstein cattle and causing potential functional consequences

SNP accession number	BTA ^a	Position (bp)	Consequence ^b	Amino acid change ^c	Entrez gene name	Type traits	Calving traits
rs134519015	5	58464570	NSC	F/S	LOC782296	–	CA
rs110862689	5	105933019	NSC	F/L	AKAP3	BQ, ST	CA, SCEc, SCEh
rs109901274	7	93244933	NSC	F/S	ARRDC3	BD, BQ, CW, RAH	CA, SCEc, SCEh, SCSh
rs134432442	14	1736599	NSC	N/S	CPSF1	BQ	–
rs133271979	14	2019390	NSC	I/M	GRINA	BQ	–
rs110073506	18	56740979	NSC	Q/P	LOC511180	–	CA, SCEc, SCEh
rs136766893	18	58494682	NSC	L/P	BOSTAUV1R416	BD, BQ	CA, SCEc, SCEh, SCSc, SCSh
rs43098627	18	58590005	NSC	Y/S	BOSTAUV1R420	–	DCA, SCEc
rs110283226	18	60233963	NSC	I/R	LOC788599	BD	CA, DCA, SCEc, SCEh, SCSh
rs110104114	18	60342959	NSC	V/G	ZNF677	–	CA, SCEh
rs110302968	18	62832344	NSC	S/R	KIR2DL5A	–	CA, SCEc, SCEh, SCSc, SCSh
rs132815594	18	62843699	NSC	E/G	FCAR	–	CA, SCEc, SCEh, SCSh
rs41902720	18	63588968	NSC	H/R	BOSTAUV1R427	–	CA, SCEc, SCEh
rs41898987	18	65754940	NSC	W/G	ZNF274	–	SCEc
rs41892719	18	56183207	SSI	–	LOC785907	–	CA, SCEc, SCEh
rs41852917	17	71314935	SL	*/Q	HORMAD2	–	CA, SCEc, SCEh

Evaluated traits include calving ability (CA), daughter calving ability (DCA), maternal calving ease (heifer; CEh), maternal calf survival at first calving (heifer; CS h), maternal calving ease at later calvings (cow; CEc), maternal calf survival at later calvings (cow; CSc), direct calving ease at first calving (heifer; SCEh), direct calf survival at first calving (heifer; SCSh), direct calving ease at later calvings (cow; SCEc), direct calf survival at later calvings (cow; SCSc) and sire conception rate (SCR)

^a BTA = *Bos taurus* autosomes

^b Consequence = non-synonymous coding (NSC), splice site intronic (SSI), stop lost (SL)

^c Amino acid change = glutamic acids (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), valine (V), tryptophan (W), tyrosine (Y)

Table 5 KEGG biological pathways enriched with the identified candidate genes for calving performance and body conformation traits

Pathway name	P-value	Gene name
<i>Calving performance</i>		
Tight junction	0.016	MYH13, MYH14, PPP2R1A, PRKCG, MYH10
Oxytocin signaling	0.040	CACNG6, CACNG7, KCNJ14, PPP1R12C, PRKCG
MAPK signaling	0.065	CACNG6, FGF21, CACNG7, FGF6, FGF23, PRKCG
<i>Body conformation</i>		
Glutamatergic synapse	0.013	SHANK1, CACNA1C, PLCB1, SLC17A7, ITPR2
MAPK signaling	0.017	FGF21, CACNA1C, FGF23, RRAS, LOC615727, RASGRF2, NTRK2
Aldosterone synthesis and secretion	0.026	ATF1, CACNA1C, PLCB1, ITPR2
Oxytocin signaling	0.035	OXT, CACNA1C, PLCB1, KCNJ14, ITPR2
Retrograde endocannabinoid signaling	0.047	CACNA1C, PLCB1, SLC17A7, ITPR2
Cholinergic synapse	0.059	CACNA1C, PLCB1, KCNJ14, ITPR2
Staphylococcus aureus infection	0.072	FCAR, C2, CFB
Dopaminergic synapse	0.077	CACNA1C, PLCB1, PPP2R1A, ITPR2
Long-term depression	0.082	PLCB1, PPP2R1A, ITPR2
Ribosome	0.090	RPS11, RPS9, RPL36, RPL13A
Long-term potentiation	0.096	CACNA1C, PLCB1, ITPR2
Renin secretion	0.099	CACNA1C, PLCB1, ITPR2

a pleiotropic effect on calving performance and body conformation traits. In total, eight, six and four QTL were confirmed for direct calving ease, maternal calving ease, and direct stillbirth, respectively. Combining significant SNPs from the GWAS with SNPs in the 50 K panel for genomic evaluation slightly increased

the prediction accuracy of genomic breeding values for calving and body conformation traits. Validation of the SNPs in independent populations, followed by the possible identification of the causal mutations within the validated candidate genes are the logical next research steps.

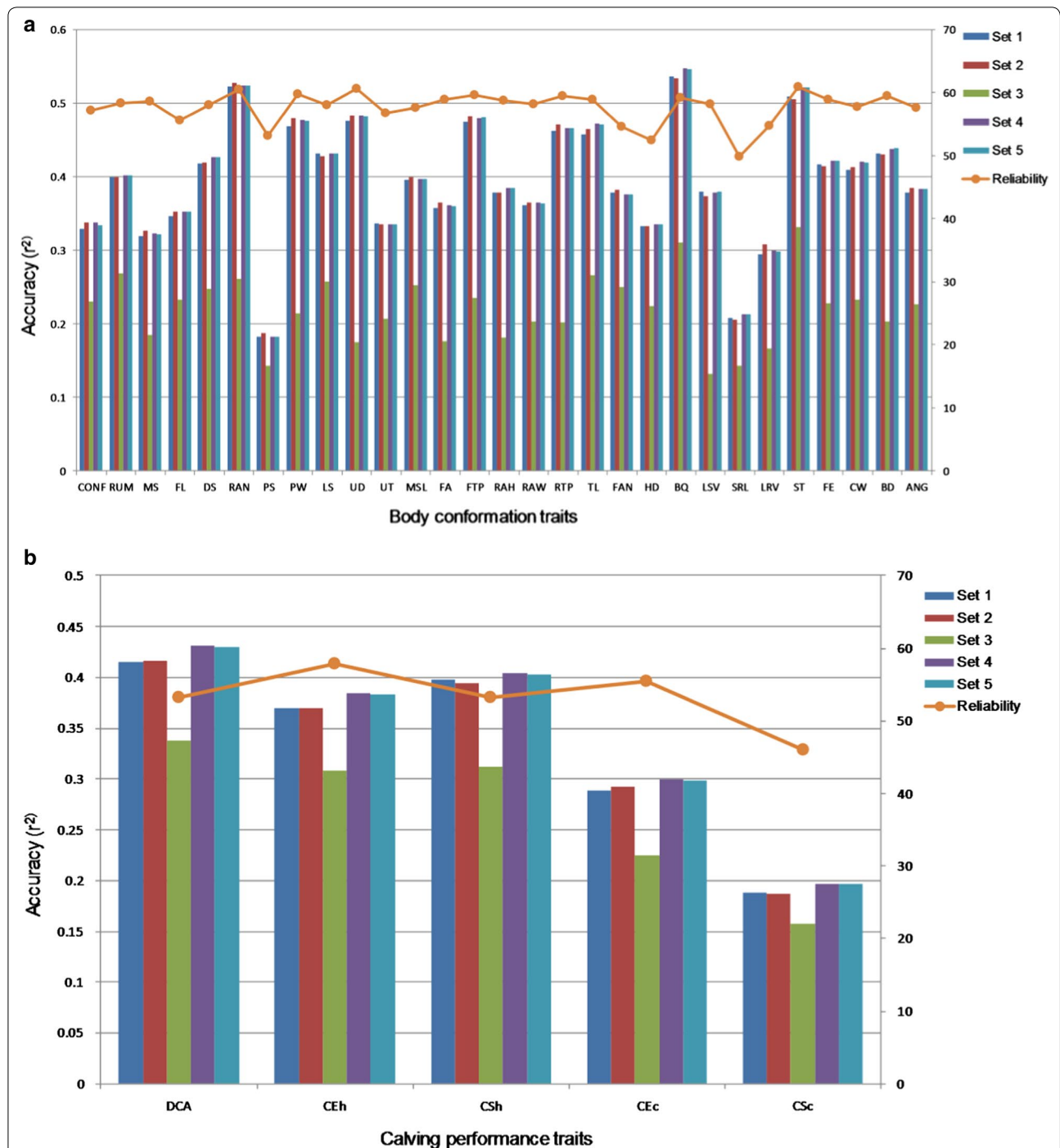


Fig. 16 Squared correlation (r^2) between direct genomic breeding values (DGV) and the corresponding de-regressed EBV for validation bulls using different SNP sets. **a** Calving performance traits are daughter calving ability (DCA); maternal calving ease at first calving (heifer; CEh); maternal calf survival at first calving (heifer; CSh) and maternal calf survival at later calvings (cow; CSc). **b** Body conformation traits are conformation score (CONF); rump (RU); mammary system (MS); feet & legs (FL); dairy strength (DS); rump angle (RAN); pin setting (PS); pin width (PW); loin strength (LS); udder depth (UD); udder texture (UT); median suspensory ligament (MSL); fore udder attachment (FA); front teat placement (FTP); teat length (TL); foot angle (FAN); heel depth (HD); bone quality (BQ); leg side view (LSV); set of rear legs (SRL); leg rear view (LRV); stature (ST); height at front end (FE); chest width (CW); body depth (BD) and angularity (ANG). Set 1 (38,405 SNPs) = set of SNPs from the 50K SNP panel; set 2 (601,717 SNPs) = set of SNPs from the HD SNP panel; set 3 (1206 SNPs) = set of significant (genome-wide false discovery rate of 5%) SNPs for calving and body conformation traits; set 4 (39,564 SNPs) = set of combined SNPs from sets 1 and 3; set 5 (39,408 SNPs) = set of combined SNPs from set 1 and a set of significant SNPs for all traits within or nearby genes (± 100 kbp around the gene); reliability = average reliabilities of the EBV from the December 2014 genetic evaluation for validation population

Additional files

Additional file 1: Figure S1. Distribution of linkage disequilibrium (r^2) calculated between pairs of SNPs on each chromosome before (A) and after (B) exclusion of the misplaced SNPs. The physical distances between pairs of SNPs are displayed along the horizontal axis, while the r^2 value for each pair is displayed on the vertical axis. The mean r^2 over successive intervals of 0.1 Mb is plotted in white

Additional file 2: Figure S2. Distribution of 601,717 SNPs in the high-density panel across the bovine genome. Genomic coordinates of SNPs are displayed along the horizontal axis (Mb) and chromosome numbers are displayed on the vertical axis

Additional file 3: Figure S3. The first three principal components of genetic co-ancestry based on Illumina BovineHD BeadChip (611,146 SNPs) genotypes for the 4848 bulls in the training population.

Additional file 4: Figure S4. Genomic inflation factor (λ) used for the test statistic (P-value) from the GWAS methods using the generalized linear mixed model accounting for reliability values associated with the de-regressed EBV. The traits are direct calf survival (cow) (SCSc); direct calf survival (heifer) (SCSh); direct calving ease (cow) (SCEC); direct calving ease (heifer) (SCEh); maternal calf survival (cow) (CSc); set of rear legs (SRL); calf survival (heifer) (CSh); daughter calving ability (DCA); heel depth (HD); calving ease (cow) (CEC); pin setting (PS); foot angle (FAN); calving ease (heifer) (CEh); leg rear view (LRV); median suspensory ligament (MSL); udder texture (UT); feet & legs (FL); rear attachment width (RAW); chest width (CW); rump (RU); rear attachment height (RAH); leg side view (LSV); mammary system (MS); loin strength (LS); angularity (ANG); height at front end (FE); conformation (CONF); fore udder attachment (FA); teat length (TL); rear teat placement (RTP); calving ability (CA); bone quality (BQ); front teat placement (FTP); body depth (BD); pin width (PW); dairy strength (DS); rump angle (RAN); udder depth (UD); and stature (ST).

Additional file 5: Table S1. List of SNPs associated with calving performance and conformation traits at a FDR of 5% and located within the confidence interval of a previously reported quantitative trait loci (QTL).

Additional file 6: Table S2. Gene ontology terms enriched at a FDR of 1% using SNP enrichment analysis for various calving performance and body conformation traits.

Additional file 7: Figure S5. The oxytocin signaling pathway enriched for genes (in red boxes) located near the identified significant SNPs for calving performance and body conformation traits.

Authors' contributions

MKA designed the studies, performed the statistical analyses, and drafted the manuscript. MS helped to collect data, performed analyses and edited the manuscript. FSS, SPM, SSM, GP and PS, developed the project, designed the studies, helped perform analyses and edited the manuscript. DAG and SN helped perform analyses and edit the manuscript. LFB edited and helped to finalize the manuscript. All authors read and approved the final manuscript.

Author details

¹ Department of Animal Biosciences, Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, ON, Canada. ² Department of Animal and Poultry Production, Damanhour University, Damanhour, Egypt. ³ The Angus Genetics Inc, Saint Joseph, MO, USA. ⁴ The Semex Alliance, Guelph, ON, Canada. ⁵ University of Queensland, Brisbane, QLD, Australia. ⁶ University of Alberta, Edmonton, AB, Canada.

Acknowledgements

This research was funded by a Collaborative Research and Development grant from NSERC and L'Alliance Boviteq (LAB). We thank the Canadian Dairy Network for providing some of the data used in this study.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

All the data supporting the results of this article are included within the article and in its supplementary files. The raw data cannot be made available, as it is property of the Canadian dairy cattle producers and this information is commercially sensitive.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 14 October 2016 Accepted: 23 October 2017

Published online: 07 November 2017

References

1. Beaudeau F, Seegers H, Ducrocq V, Fourichon C, Bareille N. Effect of health disorders on culling in dairy cows: a review and a critical discussion. *Ann Zootech.* 2000;49:293–311.
2. Eaglen SAE, Coffey MP, Woolliams JA, Mrode R, Wall E. Phenotypic effects of calving ease on the subsequent fertility and milk production of dam and calf in UK Holstein-Friesian heifers. *J Dairy Sci.* 2011;94:5413–23.
3. McGuirk BJ, Forsyth R, Dobson H. Economic cost of difficult calvings in the United Kingdom dairy herd. *Vet Rec.* 2007;161:685–7.
4. Eaglen SAE, Coffey MP, Woolliams JA, Wall E. Direct and maternal genetic relationships between calving ease, gestation length, milk production, fertility, type, and lifespan of Holstein-Friesian primiparous cows. *J Dairy Sci.* 2013;96:4015–25.
5. Lombard JE, Garry FB, Tomlinson SM, Garber LP. Impacts of dystocia on health and survival of dairy calves. *J Dairy Sci.* 2007;90:1751–60.
6. López de Maturana E, Ugarte E, González-Recio O. Impact of calving ease on functional longevity and herd amortization costs in Basque Holsteins using survival analysis. *J Dairy Sci.* 2007;90:4451–7.
7. Dematawewa CM, Berger PJ. Effect of dystocia on yield, fertility, and cow losses and an economic evaluation of dystocia scores for Holsteins. *J Dairy Sci.* 1997;80:754–61.
8. Ali TE, Burnside EB, Schaeffer LR. Relationship between external body measurements and calving difficulties in Canadian Holstein-Friesian cattle. *J Dairy Sci.* 1984;67:3034–44.
9. Sewalem A, Kistemaker GJ, Miglior F, Van Doormaal BJ. Analysis of the relationship between type traits and functional survival in Canadian Holsteins using a Weibull proportional hazards model. *J Dairy Sci.* 2004;87:3938–46.
10. Chapinal N, Koeck A, Sewalem A, Kelton DF, Mason S, Cramer G, et al. Genetic parameters for hoof lesions and their relationship with feet and leg traits in Canadian Holstein cows. *J Dairy Sci.* 2013;96:2596–604.
11. Sawa A, Bogucki M, Krężel-Czopek S, Neja W. Relationship between conformation traits and lifetime production efficiency of cows. *ISRN Vet Sci.* 2013;2013:124690.
12. VanRaden PM, Sanders AH, Tooker ME, Miller RH, Norman HD, Kuhn MT, et al. Development of a national genetic evaluation for cow fertility. *J Dairy Sci.* 2004;87:2285–92.
13. Jamrozik J, Fatehi J, Kistemaker GJ, Schaeffer LR. Estimates of genetic parameters for Canadian Holstein female reproduction traits. *J Dairy Sci.* 2005;88:2199–208.
14. Weigel KA, Rekaya R. Genetic parameters for reproductive traits of Holstein cattle in California and Minnesota. *J Dairy Sci.* 2000;83:1072–80.
15. Van Doormaal BJ, Kistemaker G, Fatehi J, Miglior F, Jamrozik J, Schaeffer LR. Genetic evaluation of female fertility in Canadian dairy breeds. *Interbull Bull.* 2004;32:86–9.
16. Bouquet A, Juga J. Integrating genomic selection into dairy cattle breeding programmes: a review. *Animal.* 2013;7:705–13.
17. Schaeffer LR. Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet.* 2006;123:218–23.
18. Van Doormaal BJ. Impact of genomics on genetic selection and gain. Canada: Canadian Dairy Network; 2014. <https://www.cdn.ca/document.php?id=340>.

19. Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, et al. Development and characterization of a high-density SNP genotyping assay for cattle. *PLoS One*. 2009;4:e5350.
20. Purfield DC, Bradley DG, Kearney JF, Berry DP. Genome-wide association study for calving traits in Holstein-Friesian dairy cattle. *Animal*. 2014;8:224–35.
21. Cole JB, Wiggans GR, Ma L, Sonstegard TS, Lawlor TJ Jr, Crooker BA, et al. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics*. 2011;12:408.
22. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brondum RF, et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat Genet*. 2014;46:858–65.
23. Erbe M, Hayes BJ, Matukumalli LK, Goswami S, Bowman PJ, Reich CM, et al. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *J Dairy Sci*. 2012;95:4114–29.
24. VanRaden PM, Null DJ, Sargolzaei M, Wiggans GR, Tooker ME, Cole JB, et al. Genomic imputation and evaluation using high-density Holstein genotypes. *J Dairy Sci*. 2013;96:668–78.
25. Sewalem A, Miglior F, Kistemaker GJ, Sullivan P, Van Doormaal BJ. Relationship between reproduction traits and functional longevity in Canadian dairy cattle. *J Dairy Sci*. 2008;91:1660–8.
26. Van Doormaal BJ, Kistemaker GJ, Miglior F. Implementation of reproductive performance genetic evaluations in Canada. *Interbull Bull*. 2007;37:129–33.
27. Miglior F, Muir BL, Van Doormaal BJ. Selection indices in Holstein cattle of various countries. *J Dairy Sci*. 2005;88:1255–63.
28. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*. 2014;15:478.
29. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet*. 2009;84:210–23.
30. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–75.
31. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, et al. Invited review: reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci*. 2009;92:16–24.
32. MacCluer JW, Boyce AJ, Dyke B, Weitkamp LR, Pfenning DW, Parsons CJ. Inbreeding and pedigree structure in Standardbred horses. *J Heredity*. 1983;74:394–9.
33. Sargolzaei M, Iwaisaki H, Colleau J. CFC: a tool for monitoring genetic diversity. In: Proceedings of the 8th world congress on genetics applied to livestock production, 13–18 August 2006; Belo Horizonte; 2006.
34. Sargolzaei M. *snp1101 User's Guide*; 2014.
35. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci*. 2008;91:4414–23.
36. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 1995;57:289–300.
37. Purfield DC, Bradley DG, Evans RD, Kearney FJ, Berry DP. Genome-wide association study for calving performance using high-density genotypes in dairy and beef cattle. *Genet Sel Evol*. 2015;47:47.
38. Hu ZL, Park CA, Wu XL, Reecy JM. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Res*. 2013;41:D871–9.
39. Szkiba D, Kapun M, von Haeseler A, Gallach M. SNP2GO: functional analysis of genome-wide association studies. *Genetics*. 2014;197:285–9.
40. Grant JR, Arantes AS, Liao X, Stothard P. In-depth annotation of SNPs arising from resequencing projects using NGS-SNP. *Bioinformatics*. 2011;27:2300–1.
41. Sargolzaei M, Schenkel FS, Jansen GB, Schaeffer LR. Extent of linkage disequilibrium in Holstein cattle in North America. *J Dairy Sci*. 2008;91:2106–17.
42. da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44–57.
43. Nejati-Javaremi A, Smith C, Gibson JP. Effect of total allelic relationship on accuracy of evaluation and response to selection. *J Anim Sci*. 1997;75:1738–45.
44. Hayes BJ, Visscher PM, Goddard ME. Increased accuracy of artificial selection by using the realized relationship matrix. *Genet Res (Camb)*. 2009;91:47–60.
45. Sargolzaei M, VanRaden PM, Kistemaker G, Schenkel F, Chesnais J. *gebv User's Guide*, Release 1.0; 2013.
46. Melka MG, Stachowicz K, Miglior F, Schenkel FS. Analyses of genetic diversity in five Canadian dairy breeds using pedigree data. *J Anim Breed Genet*. 2013;130:476–86.
47. Ma L, Wiggans GR, Wang S, Sonstegard TS, Yang J, Crooker BA, et al. Effect of sample stratification on dairy GWAS results. *BMC Genomics*. 2012;13:536.
48. Kadri NK, Gulbrandsen B, Sørensen P, Sahana G. Comparison of genome-wide association methods in analyses of admixed populations with complex familial relationships. *PLoS One*. 2014;9:e88926.
49. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet*. 2010;11:499–511.
50. Goddard ME. Uses of genomics in livestock agriculture. *Anim Prod Sci*. 2012;52:73–7.
51. VanRaden PM, O'Connell JR, Wiggans GR, Weigel KA. Genomic evaluations with many more genotypes. *Genet Sel Evol*. 2011;43:10.
52. de Roos APW, Hayes BJ, Goddard ME. Reliability of genomic predictions across multiple populations. *Genetics*. 2009;183:1545–53.
53. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genom Hum Genet*. 2009;10:387–406.
54. Segelke D, Chen J, Liu Z, Reinhardt F, Thaller G, Reents R. Reliability of genomic prediction for German Holsteins using imputed genotypes from low-density chips. *J Dairy Sci*. 2012;95:5403–11.
55. Wu C, DeWan A, Hoh J, Wang Z. A comparison of association methods correcting for population stratification in case-control studies. *Ann Hum Genet*. 2011;75:418–27.
56. Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. *Nat Rev Genet*. 2010;11:459–63.
57. Bermingham ML, Bishop SC, Woolliams JA, Pong-Wong R, Allen AR, McBride SH, et al. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. *Heredity (Edinb)*. 2014;112:543–51.
58. Sahana G, Gulbrandsen B, Lund MS. Genome-wide association study for calving traits in Danish and Swedish Holstein cattle. *J Dairy Sci*. 2011;94:479–86.
59. Höglund JK, Gulbrandsen B, Lund MS, Sahana G. Analyses of genome-wide association follow-up study for calving traits in dairy cattle. *BMC Genet*. 2012;13:71.
60. Saatchi M, Schnabel RD, Taylor JF, Garrick DJ. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics*. 2014;15:442.
61. Paliouras M, Diamandis EP. Androgens act synergistically to enhance estrogen-induced upregulation of human tissue kallikreins 10, 11, and 14 in breast cancer cells via a membrane bound androgen receptor. *Mol Oncol*. 2008;1:413–24.
62. Ventolini G. Activation of simultaneous pathways in the initiation of parturition in humans. *Med Hypotheses*. 2013;1:11.
63. Storset AK, Slettedal IÖ, Williams JL, Law A, Dissen E. Natural killer cell receptors in cattle: a bovine killer cell immunoglobulin-like receptor multigene family contains members with divergent signaling motifs. *Eur J Immunol*. 2003;33:980–90.
64. Ha NT, Gross JJ, van Dorland A, Tetens J, Thaller G, Schlather M, et al. Gene-based mapping and pathway analysis of metabolic traits in dairy cows. *PLoS One*. 2015;10:e0122325.
65. Ingvarsen KL. Feeding- and management-related diseases in the transition cow: physiological adaptations around calving and strategies to reduce feeding-related diseases. *Anim Feed Sci Technol*. 2006;126:175–213.
66. Goff JP, Horst RL. Physiological changes at parturition and their relationship to metabolic disorders. *J Dairy Sci*. 1997;80:1260–8.
67. El-Tarabany MS. Effects of non-lactating period length on the subsequent calving ease and reproductive performance of Holstein, Brown Swiss and the crosses. *Anim Reprod Sci*. 2015;158:60–7.

68. Wehrli M, Cortinas-Elizondo F, Hlushchuk R, Daudel F, Villiger PM, Miescher S, et al. Human IgA Fc receptor FcαRI (CD89) triggers different forms of Neutrophil death depending on the inflammatory microenvironment. *J Immunol*. 2014;193:5649–59.
69. Bollopragada S, Youssef R, Jordan F, Greer I, Norman J, Nelson S. Term labor is associated with a core inflammatory response in human fetal membranes, myometrium, and cervix. *Am J Obstet Gynecol*. 2009;200:104.e1–11.
70. Dowell P, Hu Z, Lane MD. Monitoring energy balance: Metabolites of fatty acid synthesis as hypothalamic sensors. *Annu Rev Biochem*. 2005;74:515–34.
71. Kaupé B, Brandt H, Prinzenberg EM, Erhardt G. Joint analysis of the influence of CYP11B1 and DGAT1 genetic variation on milk production, somatic cell score, conformation, reproduction, and productive lifespan in German Holstein cattle. *J Anim Sci*. 2007;85:11–21.
72. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res*. 2010;38:D355–60.
73. Breuiller-Fouche M, Germain G. Gene and protein expression in the myometrium in pregnancy and labor. *Reproduction*. 2006;131:837–50.
74. Sawa A, Bogucki M, Krężel-Czopek Neja W. Association between rump score and course of parturition in cows. *Anim Breed*. 2013;56:816–22.
75. Atkins G. Using conformational anatomy to identify functionality in dairy cows. In: Proceedings of the 24th South Western Ontario Dairy Symposium, 22 February 2007; Woodstock; 2007. p. 37–48.
76. Miyamoto T, Tsujimura A, Miyagawa Y, Koh E, Namiki M, Horikawa M, et al. Single-nucleotide polymorphisms in *HORMAD1* may be a risk factor for azoospermia caused by meiotic arrest in Japanese patients. *Asian J Androl*. 2012;14:580–3.
77. Gimpl G, Fahrenholz F. The oxytocin receptor system: structure, function, and regulation. *Phys Rev*. 2001;81:629–83.
78. Goddard M. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica*. 2009;136:245–57.
79. Goddard ME, Hayes BJ, Meuwissen THE. Using the genomic relationship matrix to predict the accuracy of genomic selection. *J Anim Breed Genet*. 2011;128:409–21.
80. Miglior F. Genetic evaluation of reproductive performance in Canadian dairy cattle. *Ital J Anim Sci*. 2010;6:29–37.

Submit your next manuscript to BioMed Central
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

