Shiraz University

IJVR
ISSN: 1728-1997 (Print)
ISSN: 2252-0589 (Online)

Vol.26 No.1 Ser. No. 90 2025

IRANIAN JOURNAL OF VETERINARY RESEARCH

Original Article

Genome-wide association studies for conformation traits in the Turkish Holstein cattle population

Koncagül, S.[1]; Kasakolu, A.[2]; Yıldırır, M.[3*]; Ünay, E.[4]; Koyun, H.[5]; Karakoyunlu, İ.[6]; Yiğit, O.[1] and Şen, A. Ö.[1]

[1]Department of Animal Science, Faculty of Agriculture, Ankara University, Ankara, Turkey; [2]Ph.D. Student in Animal Breeding and Genetics, Department of Animal Science, Graduate School of Natural and Applied Sciences, Ankara University, Ankara, Turkey; [3]Ph.D. in Animal Science, International Center for Livestock Research and Training, Ankara, Turkey; [4]Ph.D. in Animal Nutrition, Ministry of Agriculture and Forestry, Ankara, Turkey; [5]Department of Animal Science, Faculty of Agriculture, Van Yüzüncü Yıl University, Van, Turkey; [6]MSc in Animal Breeding, Cattle Breeder Association of Turkey, Ankara, Turkey

[*]***Correspondence:*** *M. Yıldırır, Ph.D. in Animal Science, International Center for Livestock Research and Training, Ankara, Turkey. E-mail: mesutyildirir@hotmail.com*

10.22099/ijvr.2025.50205.7407

(Received 16 May 2024; revised version 19 Jan 2025; accepted 26 Jan 2025)

Abstract

Background: Conformation traits (CNTs) are part of the selection goals that significantly affect cow economic efficiency, health, welfare, and productive life in the dairy industry. **Aims:** This study focused on a genome-wide association study (GWAS) and genetic parameters estimation for 21 CNTs, including udder, leg-foot, body, type, and final classification traits in the Turkish Holstein (THol) dairy cattle population. **Methods:** A restricted maximum likelihood with a univariate model including the fixed effects of herd-year-season and days in milk was used. The total dataset consisted of CNTs records and Affymetrix BovineSNP54K data for 3,008 THol cows that calved from 2019 to 2022. The gene ontology and Kyoto Encyclopedia of genes and genome pathway databases were used to assign genes to functional categories. The biological pathways were performed in BioMart databases. **Results:** The heritability of these 21 CNTs ranged from 0.01 (udder index) to 0.133 (udder depth). A total of 16 significant single nucleotide polymorphisms (SNP) associated with 13 CNTs was identified. Significant SNP overlap in the candidate genes, which include *ITGB1*, *TNN*, and *SEMA3D*, have potential for researchers and breeders for CNTs in cattle breeding. **Conclusion:** These results provide valuable knowledge and contribute to the elucidation of the genetic factors responsible for conformation traits in dairy cattle populations.

Key words: GWAS, Heritability, Holstein, Linear type score, SNP

Introduction

Conformation traits (CNTs), including udder, feet and legs, body, rump framework, and milk production traits, are used to select animals for higher production and longevity in dairy cows (Miglior *et al*., 2017). Selection for linear type traits can affect cow welfare, health, production, economic efficiency, reproduction, and longevity extension (Miglior *et al*., 2017; Gutiérrez-Reinoso *et al*., 2023). A selection index that combines CNTs based on their economic importance is applied in several countries (VanRaden, 2004; Miglior *et al*., 2017). A descriptive classification program that recorded these CNTs in the Turkish Holstein (THol) population became part of the selection objectives. Genomic selection in the THol population has been carried out since 2017 and implemented in applied breeding since 2023. A reference population consisting mainly of cows has been established and continuously updated. In the genomic evaluation of the THol population, milk performance, reproductive, and CNTs are mainly considered.

CNTs are closely associated with milk yield, reproduction, lameness, mastitis, and longevity in dairy cows (Long *et al*., 2024), and could be useful as phenotypes, since they help improve economic efficiency in cattle (Nazar *et al*., 2022). Heritability values based on CNTs records and pedigree kinship data reported low 0.10 (rear legs set-side view) to medium values of 0.32 (stature) in the Serbian Holstein by Djedovic *et al*. (2023), and from low (0.09) for udder texture and foot angle to high (0.38) for teat length in Holstein cows in Brazil by Campos *et al*. (2015). The genomic heritability of udder traits was reported at 0.04 to 0.49 (Nazar *et al*., 2022).

Genome-wide association studies (GWAS) have proven to be an efficient tool to identify markers for single nucleotide polymorphisms (SNP) in the genome that are associated with phenotypes (Erdoğan *et al*., 2024). The determination of quantitative trait loci (QTL) is a crucial process for understanding the genetic

variations associated with the selected traits. Ashwell *et al*. (2005) identified QTLs affecting 22 CNTs and found 41 significant QTLs. Schrooten *et al*. (2004) performed a whole genome scan using microsatellite markers for QTLs that affect 18 CNTs. As a result, numerous GWAS results have been documented using SNP markers that specifically target CNTs and provide comprehensive insights into genetic improvement through gene-based selection (Cole *et al*., 2011).

In this study, GWAS was performed to investigate 21 CNTs within the contemporary first parity THol population. The SNP panel (Affymetrix-54K) was used to discover SNP markers and genes associated with CNTs useful for the THol cattle population.

Materials and Methods

Ethics statement

The whole process of blood sampling and data recording was completed according to the program of the Ministry of Agriculture and Forestry (MAF), and the Turkish Council on Animal Care. THol population genomic selection for higher milk production, milk fat, milk protein, conformation traits, and reproductive performance is conducted under the "National Cattle Genetic Selection Program". This study was approved by the Institutional Animal Care and Ethics Committee of the International Center for Livestock Research and Training (approval No.: 2021-196).

Data collection

The data of the THol population in this study included 3,008 first parity Holstein cows. All cows came from 49 dairy farms in the population. The CNTs were collected by experts from the Cattle Breeder Association of Turkey (CBAT) as part of the Holstein Genomic Selection Program. From 2019 to 2022, the phenotypic measurement of CNTs was completed for each cow. According to the linear classification system, 16 linear traits were scored from 1 to 9, and 5 composite traits were measured with an index ranging from 65 to 90. The 16 traits were fore udder attachment (FUA), udder support (US), udder depth (UD), front teat placement (FTP), front teat length (FTL), rear teat placement (FTP), rear udder height (RUH), hock quality (HQ), rear leg view (RLRW), rear leg angle (RLA), heel depth (HD), body depth (BD), chest width (CW), rump width (RW), and rump angle (RA). The five traits of the conformation index were udder (UD), leg-foot (LF), body (BO), type (TY), and general (GN). The scores for the five composite traits were calculated based on linear point value weights. Description of the linear-type traits and composite index traits were given in Table 1.

Genotyping data and quality control

Blood samples from the tail vein of 3008 cattle were collected for DNA extraction and genotyping. Genotyping (using the Axiom Bovine 54K SNP, Affymetrix) was performed in the Laboratory of the International Center for Livestock Research and Training (ICLRT) in Ankara, Turkey.

Quality control (QC) was carried out with the code Markers function of the statgen GWAS package in the R program (R core team, 2021). SNPs were not included in the analysis if the call rate was less than 90%, a call rate for animal genotypes of less than 90%, a minor allele frequency (MAF) of less than 1%, and an extreme deviation from Hardy-Weinberg equilibrium (HWE; 1E-13). In addition, single nucleotide polymorphisms (SNPs) mapped on sex-linked and mitochondrial DNA

Table 1: Description of the linear-type traits and composite index traits

Traits	Abrv.	Definitions	Score[2] 1/65	Score 9/90	Ideal score
Udder	**UI[2]**	**Udder index**	**Poor**	**Excellent**	**90**
Fore udder attachment	FUA[3]	Angle formed between the fore udder and abdomen, side view	Weak	Strong	9
Udder support	US[3]	Depth of cleft at base of rear udder	Weak	Strong	9
Udder depth	UD[3]	Depth of udder from hock to udder floor	Deep	Shallow	5-6
Front teat placement	FTP[3]	The position of the front teat from center of quarter	Wide	Close	5
Front teat length	FTL[3]	Length of the front teats from side view	Short	Long	5
Reat teat placement	RTP[3]	Teat placement from center of quarter	Wide	Close	5-6
Rear udder height	RUH[3]	The distance between the vulva and the milk secreting tissue	Low	High	9
Dairy character	DC[3]	Angle of cidago bones	Coarse	Thin	9
Leg-Foot	**LF[2]**	**Linear leg and foot index**	**Poor**	**Excellent**	**90**
Hock quality	HQ[3]	Quality of the hocks	Coarse	Flat	5-6
Rear leg rear view	RLRV[3]	Turn of hock when viewed from rear	Hocked-in	Straight	9
Rear leg angle	RLA[3]	Rear leg side view (rear leg hock angle)	Steep	Low	5
Heel depth	HD[3]	Depth of heel on outside claw	Shallow	Deep	7-8
Body	**BO[2]**	**Linear body index**	**Poor**	**Excellent**	**90**
Body depth	BD[3]	Depth of body at the rear rib	Shallow	Deep	7
Chest width	CW[3]	Width of chest (narrow to wide)	Narrow	Wide	9
Rump width	RW[3]	Distance between inside pins	Narrow	Wide	8-9
Rump angle	RA[3]	Height of the pin bones relative to height of hook bones	High	Low	5-6
Type	**TY[2]**	**Body harmony**	**Poor**	**Excellent**	**90**
General	**GN[2]**	**Total index: TY × 0.15 + BO × 0.20 + LF × 0.25 + UI × 0.40**	**Poor**	**Excellent**	**90**

1: UI: Udder index, FUA: Fore udder attachment, US: Udder support, UD: Udder depth, FTP: Front teat placement, FTL: Front teat length, RTP: Rear teat placement, RUH: Rear udder height, DC: Dairy character, LF: Leg foot, HQ: Hock quality, RLRV: Rear leg rear view, RLA: Rear leg angle, HD: Heel depth, BO: Body, BD: Body depth, CW: Chest width, RW: Rump width, RA: Rump angle, TY: Type, and GN: General. 2: Traits calculated 65-90, and 3: Traits measured in scores 1-9

chromosomes, as well as SNPs with unclear positions (SNPs without information on chromosomal position), were excluded from further analysis.

Following the QC phase (stage), SNP imputation assignment was performed for the missing SNP genotypes. There are many algorithms among these imputation methods, and the imputation was performed randomly in the R program. After SNP imputation, the criterion of an MAF of less than 1% of loci with SNPs was also checked and removed. The HWE was then re-evaluated to identify genotyping errors. Loci with probability values below 1E-13 were excluded based on the χ-squared test. After all these evaluations, 31,944 SNPs belonging to 3,008 genotyped cows were selected for further analysis.

Genome-wide association analysis

The association analysis was performed with mixed model equations in an R environment statgen GWAS package with EMMA-based REML (Kang *et al*., 2008; Kang *et al*., 2010). SNP effects were estimated using the following statistical model:

$$y = 1\mu + Za + e$$

Where,

y: A vector of phenotype adjusted for known environmental factors (heard-year-season of calving, days in milk)

1: A vector of ones

μ: The overall mean

Z: An incidence matrix, which holds the genotypes

a: A vector of random polygenic effects and assumed as normally distributed [$a \sim N(0, G\sigma_a^2)$, G is the genomic relationship matrix (GRM) (VanRaden, 2008) built from SNP genotypes, σ_a^2 is the additive genetic variance]

e: The vector of the random residual effect

To identify the significance of the SNPs, the p-values were calculated and transformed into a -log scale to create a Manhattan plot. The Manhattan plot visually represents the association between SNP markers and the observations to identify potential genomic regions associated with additive variation in the traits. The significance level and effect of each SNP were calculated with GLS solutions (Segura *et al*., 2012). The additive variance explanation rate (A-VER) or SNP heritability with low significance for each variant was calculated as follows (Kimura *et al*., 1970):

$$\frac{\text{var}(SNP)}{\text{var}(phenotype)} = \frac{2pq\beta^2}{\text{var}(phenotype)}$$

Where,

p: The major allele frequency of each SNP

q: The minor allele frequency of each SNP

β: SNP effect to a related trait

The descriptive statistics of CNT were generated using SPSS for data analysis (mean, standard error, minimum and maximum values, standard deviation, and coefficient of variation in %).

Identification of SNP locations and gene annotation

Using the NCBI database (Genome Data Viewer, NCBI, www.ncbi.nlm.nih.gov), the SNP locations were compared with their positions in the UMD 3.1.1 bovine reference genome. The genes and QTLs closest to the significant SNPs were reported by the National Animal Genome Research Program (Animal QTLdb of Cattle QTLdb, www.animalgenome.org) and the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov).

Enrichment of functional pathways and gene network analysis of candidate genes

In this study, to better understand the biological information between and among candidate genes, all candidate genes from the GWAS analysis were entered into Gene Ontology (GO), and a KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis was performed using ShinyGO (v8.0) (http://bioinformatics.sdstate.edu/go/). In addition, a GSEA (gene-set enrichment analysis) was performed with BIOMART (https://www.ensembl.org/info/data/biomart/index.html).

Results

Descriptive analysis

The 21 CNT of 3,008 cattle, including udder index (UI), leg-foot index (LF), body index (BO), type index (TY), and general index (GN), are shown in Table 2. The variation values of the 16 linear scores for the body CNT ranged from 5.16 ± 0.027 (HQ) to 6.73 ± 0.024 (BD), as shown in Table 2. For the general index (GN) traits, the mean value of the total score was 80.81 ± 0.078, while the mean values of the composite index traits in the 5 components ranged from 78.91 ± 0.105 (LF) to 83.21 ± 0.127 (UI). The coefficient of variation (CV%) ranged from 5.27% to 32.52%.

Genomic variation and estimation of heritability

The genomic, residual, and phenotypic variances, as well as the SNP heritabilities of these traits, are shown in Table 3. The estimated heritabilities were low for all traits and ranged from 0.001 to 0.133, 0.014 to 0.064, 0.030 to 0.111, 0.078, and 0.036 for udder, leg-foot, body, type, and general traits, respectively.

Genome-wide association study

Quantile-quantile plots (QQ plots) indicated that this study's GWAS model was appropriate (Supplementary Figure 1 (SF1)). The lambda values (λ) were all close to 1.0, and the Manhattan plots show the results of GWAS significance levels (-log10 of the p-value of each SNP) by chromosomal position (Supplementary Figure 1 (SF1)). The single SNP regression analysis shows significantly associated SNP at the Bonferroni-corrected level of 5% for the CNTs (Table 4). FTP, FTL, RUH, HQ, BD, and RW shared a common SNP (rs109459144) that was 94 kb away from *LOC100139826* on Bos taurus autosome (BTA11). Twelve SNPs were identified in

Table 2: Phenotypic means, standard errors (SE), coefficient of variation (CV) and ideal means for the evaluated traits

Trait[1]	n	Min	Max	Mean	SE	SD	CV (%)
Udder							
UI[2]	3008	65	90	83.21	0.127	6.981	8.39
FUA[3]	3008	1	9	5.86	0.028	1.540	26.28
US[3]	3008	1	9	6.25	0.026	1.433	22.92
UD[3]	3008	1	9	5.22	0.031	1.698	32.52
FTP[3]	3008	1	9	5.30	0.027	1.499	28.30
FTL[3]	3008	1	9	5.54	0.020	1.113	20.07
RTP[3]	3008	1	9	6.33	.027	1.503	23.74
RUH[3]	3008	1	9	6.12	0.027	1.457	23.79
DC[3]	3008	1	9	6.19	0.029	1.588	25.63
Leg-foot							
LF[2]	3008	65	90	78.91	0.098	5.387	6.83
HQ[3]	3008	1	8	5.16	0.027	1.474	28.56
RLRV[3]	3008	1	9	5.29	0.025	1.345	25.44
RLA[3]	3008	1	8	5.50	0.020	1.110	20.18
HD[3]	3008	1	9	5.30	0.025	1.365	25.76
Body							
BO	3008	65	90	79.52	0.164	9.006	11.33
BD[3]	3008	1	9	6.73	0.024	1.321	19.63
CW[3]	3008	2	9	6.21	0.024	1.302	20.95
RW[3]	3008	1	9	6.20	0.023	1.255	20.24
RA[3]	3008	1	9	5.20	0.030	1.651	31.75
Type							
TY	3008	65	90	79.31	0.144	7.889	9.95
General							
GN	3008	65	89	80.81	0.078	4.260	5.27

1: UI: Udder index, FUA: Fore udder attachment, US: Udder support, UD: Udder depth, FTP: Front teat placement, FTL: Front teat lenght, RTP: Rear teat placement, RUH: Rear udder height, DC: Dairy character, LF: Leg foot, HQ: Hock quality, RLRV: Rear leg rear view, RLA: Rear leg angle, HD: Heel depth, BO: Body, BD: Body depth, CW: Chest width, RW: Rump width, RA: Rump angle, TY: Type, and GN: General. 2: Traits calculated 65-90, and 3: Traits measured in scores 1-9

Table 3: Estimates of variance components and heritabilities for linear type traits

Traits	σ^2_g	σ^2_e	σ^2_p	h^2
Udder				
UI	0.041	39.943	39.984	0.001
FUA	0.144	1.710	1.853	0.078
US	0.059	1.779	1.838	0.032
UD	0.265	1.729	1.994	0.133
FTP	0.095	0.924	1.019	0.093
FTL	0.167	1.623	1.789	0.093
RTP	0.175	1.927	2.102	0.083
RUH	0.111	1.318	1.429	0.078
DC	0.097	1.413	1.510	0.064
Leg-foot				
LF	0.279	20.071	20.350	0.014
HQ	0.082	1.186	1.268	0.064
RLRV	0.047	1.010	1.056	0.044
RLA	0.014	0.990	1.003	0.014
HD	0.029	0.940	0.969	0.030
Body				
BO	3.831	30.558	34.390	0.111
BD	0.042	0.921	0.964	0.044
CW	0.034	1.116	1.150	0.030
RW	0.089	0.982	1.071	0.083
RA	0.163	1.935	2.098	0.078
Type				
TY	2.952	35.128	38.080	0.078
General				
GN	0.445	11.786	12.231	0.036

1: UI: Udder index, FUA: Fore udder attachment, US: Udder support, UD: Udder depth, FTP: Front teat placement, FTL: Front teat lenght, RTP: Rear teat placement, RUH: Rear udder height, DC: Dairy character, LF: Leg foot, HQ: Hock quality, RLRV: Rear leg rear view, RLA: Rear leg angle, HD: Heel depth, BO: Body, BD: Body depth, CW: Chest width, RW: Rump width, RA: Rump angle, TY: Type, and GN: General. 2: Traits calculated 65-90, and 3: Traits measured in scores 1-9

previously documented genes or QTL, linked to conformation-related traits (Table 4). More than one CNT was influenced by the three significant SNP (rs41653166, BTA11: rs109459144, and BTA18: rs1163262).

Udder traits

A total of 4 SNPs (on BTA1, BTA6, and BTA11) were significant for five udder traits. One significant SNP on BTA11: 65353177 (rs109459144) was significant for RUH, FTP, and FTL. For UI, one significant SNP was determined on BTA6: 95988438 (rs41653166), 79 kb away from gene *LOC784058*. For FTP, two significant SNPs were identified on BTA6: 84475210 (rs41567590), and BTA11: 65353177 (rs109459144). For FUA, only one SNP on BTA1 was significant: 11603310 (rs41586703).

Leg and foot traits

A total of six significant SNPs were identified for leg and foot traits. Two important SNPs for RLRW were found in the *RAMP1* and *FASTK* genes. These are BTA3: 117929861 (rs109552830) and BTA4: 114451421 (rs438343186). Within the *TSPAN18* gene,

94 kb of *LOC100139826* (shared with RUH, FTL, FTP, BD, and RW), BTA11: 65353177 (rs109459144), and BTA15: 75821904 (rs108964424), two significant SNP for HQ were identified. Two significant SNPs were identified in the *ITGB1* gene and close to the *LOC104974792* (1,287 bp) and *LOC104974793* (14,849 bp) genes on BTA13: 20282525 (rs109184865) and BTA18: 25067681 (rs41632062).

Body traits

For trait BO, one SNP on BTA16 was significant: 57762562 (rs109089868). For BD, one SNP on BTA11: 65353177 (rs109459144) was significant, together with traits HQ, RW, and udder. For the traits of RW and RA, a total of seven SNPs were found to be significant. Three SNPs were located within the *IKBIP*, *TMTC1*, and *SHTN1* genes for RW, and one SNP within the *SEMA3D* gene. The significant SNP for RW and RA were not significant for other CNTs.

General traits

For GN traits, two SNP were significant on BTA4: 35605434 (rs109547262) and BTA6: 95988438 (rs41653166). One significant SNP (rs109547262) was located within the *SEMA3D* gene. One SNP (rs41653166) was also significant for UI.

The gene enrichment (GO, KEGG, and GSEA) analyses

Eleven bovine genes associated with CNTs entered into the ShinyGO (v. 0.80) program are listed in Table 5. The gene ontology (GO) enrichment analysis showed that 49.7% of the analyzed genes were involved in the interaction between the extracellular matrix (ECM) and the receptors. 22.5% of the genes were involved in axon guidance, while 20.7% were associated with focal adhesion pathways (Tables 5 and 6 and Fig. 1).

The significant pathways based on the KEGG analysis are shown in the Supplementary Figures 2-4 (SF2-SF4). According to these results, genes play a role in cell biological processes (BP), such as ECM-receptor interaction, axon guidance, and focal adhesion pathways.

Gene enrichment analysis of eleven bovine genes was performed using Biomart software (https://www.ensembl.org/info/data/biomart/index.html). Based on the software analyses, all genes, except for the *SLC4A2* gene, showed no phenotype description. The *SLC4A2* gene is responsible for the phenotypic manifestation of osteopetrosis, a rare disease characterized by abnormal bone growth and excessive bone density (www.ncbi.nih.gov). Only three of the eleven genes showed an interaction in biological processes in cells. As shown in Table 7, the FDR values represent the p-values (P≤0.05)

Table 4: Genome-wide significant SNPs for linear conformation traits

Traits[1]	SNP	Chr.	Pos.	MAF	P-value	Nearest gene	Distance (bp)	A_VAR
Udder								
UI	rs41653166*	6	95988438	0.19	2.48×10^{-7}	*LOC784058*	79,442	0.007
FTP	rs41567590	6	84475210	0.17	6.71×10^{-8}	-	-	0.009
FTP	rs109459144*	11	65353177	0.19	5.31×10^{-9}	*	*	0.014
FTL	rs109459144*	11	65353177	0.19	7.90×10^{-7}	*	*	0.010
FUA	rs41586703	1	11603310	0.08	3.24×10^{-7}	-	-	0.008
RUH	rs109459144*	11	65353177	0.19	5.71×10^{-8}	*	*	0.010
Leg-foot								
RLRV	rs109552830	3	117929861	0.49	9.28×10^{-8}	*RAMP1*	Within	0.007
RLRV	rs438343186	4	114451421	0.31	1.28×10^{-6}	*FASTK*	Within	0.005
HQ	rs109459144*	11	65353177	0.19	6.81×10^{-9}	*LOC100139826*	94,823	0.010
HQ	rs108964424	15	75821904	0.41	8.71×10^{-7}	*TSPAN18*	Within	0.006
RLA	rs109184865	13	20282525	0.19	1.22×10^{-6}	*ITGB1*	Within	0.007
RLA	rs41632062*	18	25067681	0.17	2.93×10^{-7}	*LOC104974792*	1,287	0.008
						LOC104974793	14,849	
Body								
BO	rs109089868	16	57762562	0.37	3.10×10^{-7}	*MRPS14*	31,415	0.008
						TNN	13,862	
BD	rs109459144*	11	65353177	0.19	4.39×10^{-8}	*	*	0.007
RW	rs109417275	5	63113033	0.08	1.22×10^{-7}	*IKBIP*	within	0.007
RW	rs41602734	5	80277740	0.15	2.94×10^{-7}	*TMTC1*	within	0.014
RW	rs109459144*	11	65353177	0.19	5.10×10^{-10}	*	*	0.014
RW	rs41666756	12	5413602	0.40	8.06×10^{-7}	*PCDH17*	210,339	0.006
RW	rs41632062*	18	25067681	0.17	2.25×10^{-8}	*	*	0.008
RW	rs41583184	26	37621755	0.27	3.72×10^{-8}	*SHTN1*	within	0.008
RA	rs437402990	4	114446132	0.03	9.62×10^{-7}	*SLC4A2*	within	0.010
General								
GN	rs109547262	4	35605434	0.23	3.30×10^{-7}	*SEMA3D*	within	0.006
GN	rs41653166*	6	95988438	0.19	3.14×10^{-8}	*ANTXR2*	358294	0.007

1: UI: Udder index, FTP: Front teat placement, FTL: Front teat length, FUA: Front udder attachment, RUH: Rear udder height, RLRW: Rear leg rear view, HQ: Hock quality, RLA: Rear leg angle, BO: Body, BD: Body depth, RW: Rump width, RA: Rump angle, and GN: General. * Same SNps have found significant for different traits

Table 5: The bovine genes related to conformation traits

Gene symbol	Ensembl gene ID	Entrez	Type	Chr.	Position (Mbp)	Description
TSPAN18	ENSBTAG00000013320	506550	Coding	3	99.8959	Tetraspanin 1
RAMP1	ENSBTAG00000011534	617017	Coding	3	117.2795	Receptor activity modifying protein 1
SEMA3D	ENSBTAG00000024394	536417	Coding	4	35.3902	Semaphorin 3D
SLC4A2	ENSBTAG00000011226	404084	Coding	4	113.6364	Solute carrier family 4 member 2
FASTK	ENSBTAG00000011228	509781	Coding	4	113.6515	Fas activated serine/threonine kinase
IKBIP	ENSBTAG00000021660	540640	Coding	5	62.7575	IKBKB interacting protein
TMTC1	ENSBTAG00000005370	533191	Coding	5	79.8077	Transmembrane O-mannosyltransferase targeting cadherins 1
ANTXR2	ENSBTAG00000014324	510080	Coding	6	94.5799	ANTXR cell adhesion molecule 2
ITGB1	ENSBTAG00000015910	281876	Coding	13	19.9716	Integrin subunit beta 1
TNN	ENSBTAG00000015650	517433	Coding	16	56.3254	Tenascin N
SHTN1	ENSBTAG00000007578	532603	Coding	26	37.2709	Shootin 1

Table 6: Top pathways among the submitted genes influencing conformational traits

Enrichment FDR	nGenes	Pathway genes	Fold enrichment	Pathways
2.1E-02	2	80	49.7	ECM-receptor interaction
4.0E-02	2	177	22.5	Axon guidance
4.0E-02	2	192	20.7	Focal adhesion

Table 7: The gene-set enrichment (GSEA) analysis of the candidate genes

Enrichment FDR	nGenes in interactions	Pathway (gene functions)	Interactions between genes
0.01752	2	BTA04512 ECM-receptor interaction	ITGB1 TNN (these two gene function in ECM-receptor interaction)
0.03297	2	BTA04360 Axon guidance	ITGB1 SEMA3D (these two genes function in Axon guidance)
0.03297	2	BTA04510 Focal adhesion	ITGB1 TNN (these two gene function in focal adhesion)

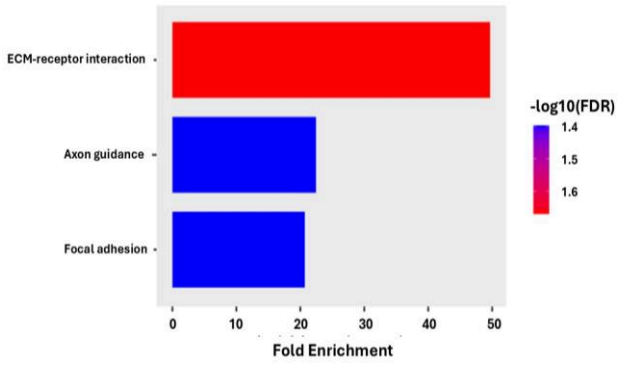

Fig. 1: A bar plot of genes associated biological pathways

and gene interactions in the pathways. The genes *ITGB1* and *SEMA3A* from the *SEMA3* gene group were found to interact in the repulsion and attraction of axons. Accordingly, the genes are *ITG1*, *SEMA3D*, and *TNN* on BTA4, and the *ITGB1* gene was found to interact with both *SEMA3D* and *TNN* genes, resulting in a statistically significant molecular function (P≤0.05). The *ITG1* and *TNN* genes function within cells and play a role in the interaction with extracellular matrix (ECM) receptors. The pathway of ECM receptor interaction influences whether some related genes are transcribed more or less at the beginning of lactation (P<0.01). In addition, the *ITG1* and *TNN* genes together have a focal adhesion function. It is hypothesized that they form large, dynamic protein complexes whose scaffold is associated with the ECM (P<0.05). It was found that the genes *ITG1* and *SEMA3D* play a role in the emission of axons (nerve cell processes or fibers) that bring the neurons to their proper destinations (P<0.05).

Discussion

Accurate heritability estimates are critical for predicting expected selection responses and estimating breeding values. In this study, the heritability of the CNTs was low, ranging from 0.04 to 0.13 (Table 3). Long *et al*. (2024) recently reviewed CNT studies and reported that the heritability of CNTs generally ranges from low to moderate. Several studies based on the pedigree kinship heritabilities of the CNTs reported low to medium, ranging from 0.09 to 0.38 by Campos *et al*. (2015), 0.10 to 0.32 by Djedovic *et al*. (2023), 0.04 to 0.23 by Olasege *et al*. (2019). A study by Roveglia *et al*. (2019) on the Italian Jersey breed estimated heritabilities ranging from low 0.04 (legs and locomotion) and 0.07 (foot angle), whereas stature shows moderate heritability at 0.32. Wu *et al*. (2013) reported a GWAS study heritability ranging from 0.07 (rump width) to 0.37 (stature). Pedigree-based heritability covers the whole genome and is expected to be higher than genomic heritability. Efforts have been made to enhance CNTs in THol dairy populations, but the low heritability of this trait may indicate smaller genetic gains.

Because of the research carried out to date on this topic, many researchers have investigated QTL on BTA6, indicating the presence of multiple QTLs on this chromosome. In a study conducted by Schrooten *et al*. (2000), two significant single nucleotide polymorphisms (SNPs) (rs41653166) were discovered on BTA6 at position 95988438. One is located 79 kb away from

LOC784058 and is linked to UI and GN. The other is located in a QTL (10286) that has been associated with udder height (Schrooten *et al*., 2000), milking speed (Boichard *et al*., 2003), clinical mastitis (Lund *et al*., 2008), carcass weight (McClure *et al*., 2010), and teat length (Ashwell *et al*., 2005). Ashwell *et al*. (2005) identified another important SNP for FTP on BTA6: 84475210. This had a peak of 106.4 centiMorgan (cM) in a QTL associated with teat length (1569). SNPs and QTL affecting udder and general index traits were both significant in this region. The region in question was associated with candidate genes, suggesting that it warrants further investigation.

The two significant SNPs associated with RLRV (BTA3: 117929861 and BTA4: 114451421) are located in the *RAMP1* and *FASTK* genes, respectively. The *RAMP1* gene encodes a protein called receptor activity modifying protein 1, while the *FASTK* gene encodes a serine/threonine kinase that is activated by Fas. Cells with *RAMP1* gene over-expression show typical calcium deposits in the bone matrix structure (Zhao *et al*., 2013). In addition, several QTLs in this specific genomic region have been reported to have effects on body weight (McClure *et al*., 2010). These genes are proposed as functional candidate genes for RLRV.

In this study, six significant SNPs for RW were found on BTA5, BTA11, BTA12, BTA18, and BTA26. A significant relationship between aspects of fertility and rump traits has been shown in studies. Animals with wide pins, long sloping rumps, and a low slope from the thurl to the pin bone are favored for easy calving (Ali *et al*., 1984; Cue *et al*., 1990). Similarly, Lu *et al*. (2021) reported that the rump angle is closely related to the reproductive performance of dairy cows. Cows with high and narrow pin bones had an increased risk of retained placenta (Van Dorp *et al*., 1998). Makgahlela *et al*. (2009) reported longer calving intervals associatied with deep angular bodies and steep rump angles. In this study, two notable SNPs were discovered for RW on BTA5: 63113033 and BTA5: 80277740, located in the *IKBIP* and *TMTC1* genes, respectively. Weller *et al*. (2018) reported OTL for thurl width on BTA5. Boichar *et al*. (2003) also identified a significant quantitative trait locus (QTL) on BTA5 at position 103.1 (cM) for traits related to RW and rump length. Rump breeding values indicated that animals with an average rump angle (from 4 to 6 on a scale of 1 to 9) had a lower culling rate than animals at the extremes (high and low pin bones) (Wall *et al*., 2005). The relative risk of unwanted culling is smallest at intermediate rump angles (Caraviello *et al.*, 2004). The use of RW can help overcome management issues that may be present in fertility performance. The RA can provide additional information for the estimation of the breeding value of fertility in dairy cattle breeding. Further studies for RW and RA may be beneficial for estimating genomic breeding values. Therefore, this study has shown there is also some potential in the RW and RA traits in helping to estimate genomic breeding values of fertility.

Several QTLs for the rump angle and body weight were detected on BTA11 (Boichard *et al*., 2003; McClure *et al*., 2010). McClure *et al*. (2010) found an important QTL for calving ease on the BTA12 peak located at 6.04 (cM), which covers the locus where we detected a significant SNP (rs41666756; BTA12: 5413602) for RW. The significant SNP (rs41666756) at BTA12: 5413602 for RW is 210,339 bp away from the *PCDH17* gene. McClure *et al.* (2010) found a QTL (10914) on BTA12 peak 6.04 (cM), affecting calving ease. Müller *et al*. (2017) found an important QTL (126847) on BTA5 peak at 65.85 (cM) for calving ease and stillbirths. Consistently, a very prominent QTL for different fertility traits was identified on BTA5 in one region. In this study, a significant SNP (rs109459144) on BTA11: 65353177 affected five CNTs (HQ, RW, BD, RUH, FTL, and FUA) 98 kb near *LOC100139826*, an uncharacterized protein-coding gene. This SNP within the QTL region is associated with rump angle (Boichard *et al*., 2003) and body weight (Michenet *et al*., 2016). In this study, a notable SNP (rs41583184) was identified on BTA26: 37621755 for RW, which is consistent with the results of Thomasen *et al*. (2008), who reported a prominent quantitative trait locus (QTL) on BTA26 affecting stillbirths. The RLA and RW traits share a common SNP (rs41632062) on BTA18: 25067681, which is probably the most important conformation trait for calving ease (Dadati *et al*., 1985). The results of the study indicate that the major SNP for RW is within QTL associated with calving ease, stillbirth, RW, and rump angle phenotypes (Boichard *et al*., 2003; Thomasen *et al*., 2008; McClure *et al*., 2010). Therefore, this study provides additional evidence for a QTL or several closely related QTLs.

Gene set enrichment analysis has been a successful extension of genome-wide association analysis (Abdalla *et al*., 2016). Considering GO and GSEA analyses, *ITGB1* and *TNN* genes were found to play a role in extracellular matrix (ECM) receptor interactions, *ITGB1* and *SEMA3D* in axon guidance interactions, and *ITGB1* and *TNN* genes in focal adhesion interactions (Tables 6 and 7). The ECM is crucial for the development of animal tissues and organs. Enzymes involved in the ECM signaling pathway promote cell proliferation, cell differentiation, skeletal development, and morphogenesis (Jeong *et al*., 2017). The ECM has been found to promote the functional recovery of several tissue components, including musculoskeletal tissues (Zhao and Bass, 2018). In addition to the interactions, the influence of these genes on CNTs was also determined. In this regard, *ITGB1*, *TNN*, and *SEMA3D* genes were found to influence the leg-foot criterion, the body trait, and the general CNTs, respectively (Table 4). These genomic markers overlap with potential candidate genes that are involved in biological pathways such as skeletal development, morphogenesis, and adipogenesis. This study confirmed previous results from GWAS of CNTs and also identified additional regions in the cattle genome associated with these economically important traits. These results can be used to increase genetic progress in breeding programs.

Efforts have been made to improve CNTs in THol dairy populations, but the low heritability of this trait may indicate smaller genetic gains. A reference population of reasonable size is crucial for genomic selection; the amount of both genotyped individuals of CNTs is smaller. Our results showed that the accuracy of genomic prediction for CNTs could be improved by GWAS. Therefore, both genotyped and phenotype animals in the reference population need to be enhanced. The present GWAS identified 16 significant SNPs associated with 13 CNTs in the THol population. The significant SNP markers associated with the *ITGB1*, *TNN*, and *SEMA3D* genes discovered in this research could have potential for researchers and breeders in the genomic selection of CNTs in dairy cattle breeding. These results contribute to the elucidation of the genetic factors responsible for conformation traits in the THol cattle population.

Acknowledgements

This work was supported by the National Holstein Cattle Genomic Breeding Program of Turkey. The authors would like to thank the Cattle Breeder Association for providing phenotype data.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Abdalla, EA; Peñagaricano, F; Byrem, TM; Weigel, KA and Rosa, GJ (2016). Genome-wide association mapping and pathway analysis of leukosis incidence in a US Holstein cattle population. Anim. Genet., 47: 395-407.

Ali, TE; Burnside, EB and Schaeffer, LR (1984). Relationship between external body measurements and calving difficulties in Canadian Holstein-Friesian cattle. J. Dairy Sci., 67: 3034-3044.

Ashwell, MS; Heyen, DW; Weller, JI; Ron, M; Sonstegard, TS; Van Tassell, CP and Lewin, HA (2005). Detection of quantitative trait loci influencing conformation traits and calving ease in Holstein-Friesian cattle. J. Dairy Sci., 88: 4111-4119.

Boichard, D; Grohs, C; Bourgeois, F; Cerqueira, F; Faugeras, R; Neau, A; Rupp, R; Amigues, Y; Yvonne, MB and Levéziel, H (2003). Detection of genes influencing economic traits in three French dairy cattle breeds. Genet. Sel. Evol., 35: 77-101.

Campos, RV; Cobuci, JA; Kern, EL; Costa, CN and McManus, CM (2015). Genetic parameters for linear type traits and milk, fat, and protein production in Holstein cows in Brazil. Asian-Aus. J. Anim. Sci., 28: 476-484.

Caraviello, DZ; Weigel, KA and Gianola, D (2004). Analysis of the relationship between type traits and functional survival in US Holstein cattle using a Weibull proportional hazards model. J. Dairy Sci., 87: 2677-2686.

Cole, JB; Wiggans, GR; Ma, L; Sonstegard, TS; Lawlor, TJ; Crooker, BA; Van Tassel, CP; Yang, J; Wang, S; Matukumalli, LK and Da, Y (2011). Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary US Holstein cows. BMC Genomics, 12: 1-17.

Cue, RI; Monardes, HG and Hayes, JF (1990). Relationships of calving ease with type traits. J. Dairy Sci., 73: 3586-3590.

Dadati, E; Kennedy, BW and Burnside, EB (1985). Relationships between conformation and reproduction in Holstein cows: type and calving performance. J. Dairy Sci., 68: 2639-2645.

Djedovic, R; Vukasinovic, N; Stanojević, D; Bogdanović, V; Ismael, H; Janković, D; Gligović, N; Brka, M and Štrbac, L (2023). Genetic parameters for functional longevity, type traits, and production in the Serbian Holstein. Animals, 13: 534.

Erdoğan, M; Çinkaya, S; Brenig, B; Çelikeloğlu, K; Demirtaş, M; Sarıbrahimoğlu, S and Tekerli, M (2024). Genome-wide association studies for milk production traits and persistency of first calving Holstein cattle in Turkey. Front. Vet. Sci., 11: 1461075.

Gutiérrez-Reinoso, MA; Aponte, PM and García-Herreros, M (2023). Genomic and phenotypic udder evaluation for dairy cattle selection: A review. Animals. 13: 1588.

Jeong, JY; Ibrahim, M; Kim, MJ; So, K; Jeong, YD; Park, S; Kim, M and Lee, HJ (2017). Comparisons of extracellular matrix-related gene expression levels in different adipose tissues from Korean cattle. Liv. Sci., 198: 138-146.

Kang, HM; Sul, JH; Service, SK; Zaitlen, NA; Kong, SY; Freimer, NB; Sabatti, C and Eskin, E (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet., 42: 348-354.

Kang, HM; Ye, C and Eskin, E (2008). Accurate discovery of expression quantitative trait loci under confounding from spurious and genuine regulatory hotspots. Genetics, 180: 1909-1925.

Long, M; Wang, B; Yang, Z and Lu, X (2024). Genome-Wide association study as an efficacious approach to discover candidate genes associated with body linear type traits in dairy cattle. Animals, 14: 2181.

Lu, X; Abdalla, IM; Nazar, M; Fan, Y; Zhang, Z; Wu, X; Xu, T and Yang, Z (2021). Genome-wide association study on reproduction-related body-shape traits of Chinese Holstein cows. Animals, 11: 1927.

Lund, MS; Guldbrandtsen, B; Buitenhuis, AJ; Thomsen, B and Bendixen, C (2008). Detection of quantitative trait loci in Danish Holstein cattle affecting clinical mastitis, somatic cell score, udder conformation traits, and assessment of associated effects on milk yield. J. Dairy Sci., 91: 4028-4036.

Makgahlela, MI; Mostert, BE and Banga, CB (2009). Genetic relationships between calving interval and linear type traits in South African Holstein and Jersey cattle. SA. Soc. Anim. Sci., 39: 90-92.

McClure, MC; Morsci, NS; Schnabel, RD; Kim, JW; Yao, P; Rolf, MM; McKay, SD; Gregg, SJ; Chapple, RH; Northcutt, SL and Taylor, JF (2010). A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Anim. Genet., 41: 597-607.

Michenet, A; Barbat, M; Saintilan, R; Venot, E and Phocas, F (2016). Detection of quantitative trait loci for maternal traits using high-density genotypes of Blonde d'Aquitaine beef cattle. BMC Genet., 17: 1-13.

Miglior, F; Fleming, A; Malchiodi, F; Brito, LF; Martin, P and Baes, CF (2017). A 100-Year review: Identification

and genetic selection of economically important traits in dairy cattle. J. Dairy Sci., 100: 10251-10271.

Müller, MP; Rothammer, S; Seichter, D; Russ, I; Hinrichs, D; Tetens, J; Thaller, G and Medugorac, I (2017). Genome-wide mapping of 10 calving and fertility traits in Holstein dairy cattle with special regard to chromosome 18. J. Dairy Sci., 100: 1987-2006.

Nazar, M; Abdalla, IM; Chen, Z; Ullah, N; Liang, Y; Chu, S; Xu, T; Mao, Y; Yang, Z and Lu, X (2022). Genome-Wide association study for udder conformation traits in Chinese Holstein cattle. Animals. 12: 2542.

Olasege, B; Zhang, S; Zhao, Q; Liu, D; Sun, H; Wang, Q; Ma, P and Pan, Y (2019). Genetic parameter estimates for body conformation traits using composite index, principal component, and factor analysis. J. Dairy Sci., 102: 5219-5229.

R Core Team (2021). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. https://www.R-project.org/.

Roveglia, C; Niero, G; Bobbo, T; Penasa, M; Finocchiaro, R; Visentin, G; Lopez-Villalobos, N and Cassandro, M (2019). Genetic parameters for linear type traits including locomotion in Italian Jersey cattle breed. Livest. Sci., 229: 131-136.

Schrooten, C; Bink, MCAM and Bovenhuis, H (2004). Whole genome scan to detect chromosomal regions affecting multiple traits in dairy cattle. J. Dairy Sci., 87: 3550-3560.

Schrooten, C; Bovenhuis, H; Coppieters, W and Van Arendonk, JAM (2000). Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. J. Dairy Sci., 83: 795-806.

Segura, V; Vilhjálmsson, BJ; Platt, A; Korte, A; Seren, Ü; Long, Q and Nordborg, M (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat. Genet., 44: 825-830.

Thomasen, JR; Guldbrandtsen, B; Sørensen, P; Thomsen, B and Lund, MS (2008). Quantitative trait loci affecting calving traits in Danish Holstein cattle. J. Dairy Sci., 91: 2098-2105.

Van Dorp, TE; Dekkers, JCM; Martin, SW and Noordhuizen, JPTM (1998). Genetic parameters of health disorders, and relationships with 305-day milk yield and conformation traits of registered Holstein cows. J. Dairy Sci., 81: 2264-2270.

VanRaden, PM (2004). Invited review: Selection on net merit to improve lifetime profit. J. Dairy Sci., 87: 3125-3131.

VanRaden, PM (2008). Efficient methods to compute genomic predictions. J. Dairy Sci., 91: 4414-4423.

Wall, E; White, IMS; Coffey, MP and Brotherstone, S (2005). The relationship between fertility, rump angle, and selected type information in Holstein-Friesian cows. J. Dairy Sci., 88: 1521-1528.

Weller, JI; Bickhart, DM; Wiggans, GR; Tooker, ME; O'Connell, JR; Jiang, J; Ron, M and VanRaden, PM (2018). Determination of quantitative trait nucleotides by concordance analysis between quantitative trait loci and marker genotypes of US Holsteins. J. Dairy Sci., 101: 9089-9107.

Wu, X; Fang, M; Liu, L; Wang, S; Liu, J; Ding, X; Shengli, Z; Qin, Z; Yuan, Z; Lv, Q; Mogens, SL; Guosheng, S and Sun, D (2013). Genome wide association studies for body conformation traits in the Chinese Holstein cattle population. BMC Genet., 14: 1-10.

Zhao, JY and Bass, KD (2018). Skeletal muscle regeneration by extracellular matrix biological scaffold: a case report. J. Wound Care. 27: 11-14.

Zhao, Z; Fu, X; Zhang, G; Li, Y; Wu, M and Tan, Y (2013). The influence of RAMP1 overexpression on CGRP-induced osteogenic differentiation of MG-63 cells *in vitro*: An experimental study. J. Cell. Biochem., 114: 314-322.

Supporting Online Material

Refer to web version on PubMed Central® (PMC) for Supplementary Material.