

ORIGINAL ARTICLE OPEN ACCESS

Single-Step Genome-Wide Association Study of Factors for Evaluated and Linearly Scored Traits in Swedish Warmblood Horses

Anahit Nazari-Ghadikolaei¹  | W. Freddy Fikse²  | Åsa Gelinder Viklund¹  | Sofia Mikko¹  | Susanne Eriksson¹ 

¹Department of Animal Biosciences, Swedish University of Agricultural Sciences, Uppsala, Sweden | ²Växa, Uppsala, Sweden

Correspondence: Susanne Eriksson (susanne.eriksson@slu.se)

Received: 28 June 2024 | **Revised:** 18 November 2024 | **Accepted:** 23 December 2024

Funding: The genotype data used in this study were previously funded by research grants from the Swedish-Norwegian Foundation for Equine Research (grant number: H1147215).

Keywords: conformation | gaits | jumping | sport horses | young horse test

ABSTRACT

Swedish Warmblood horses (SWB) are bred for show jumping and/or dressage with young horse test scores as indicator traits. This study aimed to investigate possible candidate genes and regions of importance for evaluated and linearly scored young horse test traits. A single-step genome-wide association study (ssGWAS) was done using the BLUPF90 suite of programs for factors scores from factor analysis of traits assessed at young horse tests together with height at withers. The ssGWAS included 20,814 SWB with factors scores for four factors for evaluated traits. A total of 6436 of these horses also had factor scores for 13 factors for linearly scored traits. Genotypes from a 670K SNP array were available for 380 of the horses in this study. All genotyped horses had factor scores for evaluated traits, and 379 also had factors scores for linearly scored traits. Significant SNPs associated with three factors related to size were located on ECA3 within or nearby a well-known region, including the genes *ligand dependent nuclear receptor corepressor like (LCORL)*, *non-SMC condensin I complex subunit G (NCAPG)*, *DDB1* and *CUL4 Associated Factor 16 (DCAF16)*, and the *Family with Sequence Similarity 184 Member B (FAM184B)*. Significant SNPs were also detected for two factors for evaluated traits representing conformation and jumping, and four factors for linearly scored traits related to body length, neck conformation, walk and trot (hindleg position and activity), respectively. Among nearby genes, *calcium/calmodulin-dependent protein kinase type 1D (CAMK1D)* for the factor for linearly scored traits related to neck conformation and *GLI Family Zinc Finger 2 (GLI2)* for the factor for evaluated jumping traits, were most promising. For these, top associated SNPs were detected within the genes, and the known gene functions seems to be related to the phenotypes. In conclusion, ssGWAS is beneficial to detect plausible candidate genes/regions for desired traits in warmblood horses.

1 | Introduction

Swedish Warmblood horses (SWB) are bred with the aim to produce noble, correct, durable and competitive horses at international level in show jumping and/or dressage (Swedish Warmblood Association 2021). Considerable genetic progress is seen for both show jumping and dressage performance (Bonow et al. 2023). However, there is limited knowledge about genomic

regions or genes of importance for the young horse test traits that, together with competition performance, constitute the basis for the genetic evaluation (Viklund et al. 2011).

Identification of genomic regions, and potentially genes, with a significant impact on traits related to the breeding goal can increase our understanding of the biology and may influence future breeding strategies (Stock et al. 2016). The quality of

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Journal of Animal Breeding and Genetics* published by John Wiley & Sons Ltd.

phenotypic data is of crucial importance for genomic studies. During recent years, the more refined traits recorded using linear scoring protocols have been suggested to be well suited also for genomic analysis (Wobbe et al. 2021). A protocol for linear trait profiling was introduced in 2013 at young horse tests for 3-year-old SWB (Viklund and Eriksson 2018). In this scoring system, the horse is described on a linear scale from one biological extreme to another, for a range of separate traits. The linear profiling aims to describe, rather than evaluate traits, in contrast to the traditional evaluating system that score horses with regard to the breeding goal (Duensing, Stock, and Krieter 2014). Still, both types of traits are visually assessed by judges.

At present, 50 linearly scored traits are recorded at tests for 3-year-old SWB horses, in addition to height at withers and eight traditional evaluated traits. Several of them can be assumed to be highly correlated, and it is desirable to reduce the number of traits when conducting genomic studies of young horse test traits. Factor analysis (FA) is one of the methodologies that is used to shrink the number of correlated traits to a lower number of underlying factors that can be helpful to understanding the nature and biology of the traits. This method has been used for young horse test results for 3-year-old SWB horses in a study by Nazari-Ghadikolaei et al. (2023), as well as in studies of other horse breeds and traits (Staiger, Albright, and Brooks 2016; Sigurðardóttir, Albertsdóttir, and Eriksson 2017).

Several genome-wide association studies (GWAS) have been conducted in horses of warmblood sport horse type for conformation (Signer-Hasler et al. 2012b; Frischknecht et al. 2016; Gmel et al. 2019; Reich et al. 2024), jumping (Schröder et al. 2012; Brard and Ricard 2015) and gaits (Legarra, Ricard, and Varona 2018; Ricard et al. 2020), including only genotyped horses. This has resulted in an increasing number of quantitative trait loci (QTLs), and gene(s), as reported in the Horse Quantitative Trait Locus Database (QTLdb 2023). However, far from all GWAS in horses have been successful in detecting significant associations and much work remains before a more complete understanding of the genetics behind traits related to show jumping and dressage performance is in place.

For horse populations where the recording of traits and pedigrees often has been extensive, whereas relatively few individuals are genotyped so far, the use of single-step methods for GWAS can be valuable. Single-step genomic best linear unbiased predictions (ssGBLUP) allows integration of genotype, phenotype and pedigree information simultaneously (Legarra et al. 2014). It can be implemented for conducting single-step GWAS (ssGWAS) in two main steps, described by Wang et al. (2012): (1) prediction of genomic estimated breeding values (GEBV) and (2) obtain individual prediction error variances of marker effects to use in back solving genomic breeding values to SNP effects and to obtain *p* values. This method has an advantage of including both genotyped and non-genotyped animals in the final model estimation. A few studies of traits in horses have used ssGWAS: Vosgerau et al. (2022) conducted ssGWAS to detect possible candidate loci or gene(s) associated with withers height in German Warmblood horses. Other examples of using ssGWAS include racing performance in Quarter horses (Pereira et al. 2018) and

French trotters (Ricard and Duluard 2021), as well as gaits and performance in Brazilian Mangalarga Marchador horses (de Oliveira Bussiman et al. 2020).

The aim of this study was to conduct a ssGWAS of factors for evaluated and linearly scored traits in three-year-old SWB horses, to explore and detect possible candidate region(s) or gene(s) related to traits of importance for the breeding goal.

2 | Material and Method

2.1 | Animals and Phenotypes

In total, the data included records for 20,935 SWB horses assessed at young horse tests for 3-year-olds between 1999 and 2020. The data were provided together with a seven-generation pedigree file by the Swedish Warmblood Association, via the Swedish University of Agricultural Sciences. All horses had evaluated trait records (scale from 1 to 10), and most (20,814) also had a measure of height at withers (cm). Among these horses, 6436 horses assessed between 2013 and 2020 also had linearly scored traits recorded (scale from 1 to 9). High density (670k) SNP data from Affymetrix Axiom Equine Genotyping Array, previously used by Ablondi, Viklund, et al. (2019), was available for 380 of the horses (182 males and 198 females) in the present study. The genotyped horses were all assessed at young horse tests for 3-year-olds during the years 2013 and 2014 and had complete information about height at withers and all evaluated traits. In addition, 379 of the genotyped horses had linearly scored trait records. Each horse was only assessed once, so there were no repeated scores. The 380 genotyped horses descended from 145 sires with 1–11 offspring each and 372 mares with 1–2 offspring each. As described by Ablondi, Viklund, et al. (2019), the horses were selected for genotyping on the basis of their young horse test results for evaluating traits. Selected horses had either high scores for jumping but lower scores for gaits ($N=48$), high scores for gaits but lower scores for jumping ($N=48$), high scores for both jumping and gaits ($N=143$), or low scores for both jumping and gaits ($N=141$). Efforts were made to match half-sibs across different trait profiles in this data set.

2.2 | Factor Analysis

Factor analyses were done using the Psych package (Revelle 2022) in R (R Core Team 2019) for complete records for evaluated traits together with height at withers, and for linearly scored traits together with height at withers, as described in detail by Nazari-Ghadikolaei et al. (2023). The correlation matrix used in the factor analysis was approximated by the product of a factor matrix and its transpose plus a diagonal matrix of uniqueness (proportion of variance not explained by the common factor) (Revelle 2022). Factors were extracted using maximum likelihood and subsequently rotated by the ‘varimax’ rotation to produce more easily interpreted factors that represented a smaller number of variables (Abdi 2003). The optimum number of factors was determined as the number of factors with Tucker Lewis Index larger than 0.9 and root mean square of residuals (RMSR index) value of close to zero (Tucker and Lewis 1973; Steiger 1990), as well as eigenvalues above 1.

To calculate factor scores, we used the factor loadings, the inverse of the data covariance matrix and the trait data after imputation and applied the regression method (Thurstone 1935). In contrast to the study by Nazari-Ghadikolaei et al. (2023), the first year of linear scoring (2013) was included in this study, and the traits related to conformation of pasterns (for which information was lacking in 2013) were removed, but this had a very small effect on the formation of factors, other than removing a specific factor for pasterns. To be able to predict factor scores in spite of a few missing linearly scored trait observations, missing values were imputed on the basis of correlated traits using the Mice package (Van Buuren and Groothuis-Oudshoorn 2011) in R as outlined in Nazari-Ghadikolaei et al. (2023). For all factors for which results are presented in this manuscript, the traits with significant loadings (≥ 10.31) were available for on average 99% of the genotyped horses (ranging from 94% to 100% for the different linearly scored traits).

Four factors for evaluated traits and 13 factors for linearly scored traits were used for the ssGWAS. More details about the factors are given when the results are presented and discussed later in this article and in Nazari-Ghadikolaei et al. (2023).

2.3 | Genotype Quality Control

Genotype quality control (QC) was performed in PLINK v.1.9 (Purcell et al. 2007), removing animals and SNPs with less than 90% genotyping call rates, SNPs with minor allele frequency (MAF) less than 0.05, and SNPs with p value for Hardy–Weinberg Equilibrium (HWE) test less than $1e-6$ from further analyses. In total, 30,712 SNPs were removed because of low call rate, 18,933 were removed on the basis of HWE, and 197,375 SNPs were removed on the basis of low MAF. Finally, 380 animals with 389,997 SNPs remained for the ssGWA study.

2.4 | Single-Step Genome-Wide Association Study

Prediction of genomic breeding values (GEBVs) was done as the first step in the ssGWAS. The estimation of genetic parameters (Table S1) that was used for this purpose is described in Nazari-Ghadikolaei et al. (2023). The GEBVs were predicted using BLUPF90 (Misztal et al. 2002) with a single-trait animal model, including the same fixed effects (sex and event) as was used for the estimation of genetic parameters. In the single-step prediction, the inverse of a hybrid relationship matrix, $\mathbf{H}^{-1}$, was used instead of the inverse numerator relationship matrix $\mathbf{A}^{-1}$ as follows (Aguilar et al. 2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

where $\mathbf{A}$ is the pedigree-based numerator relationship matrix for all animals, $\mathbf{G}$ is a genomic relationship matrix, and $\mathbf{A}_{22}$ is a numerator relationship matrix for genotyped animals. The default weight of 0.95 on genomic information was used. After prediction of GEBVs, POSTGSF90 was used for estimating SNP

effects (Wang et al. 2012; Aguilar et al. 2014), % variance explained by each SNP with OPTION windows_variance with a window size of 1 SNP, and p values using OPTION p value according to (Aguilar et al. 2019).

Quantile–quantile (QQ) and Manhattan plots were created using CMplot package in R (Yin 2022). The QQ-plots were used to interpret the inflation factor (lambda value) because of possible cryptic relationship among animals and to detect significant SNPs. Estimation of lambda value was done using the estlambda() function in the GenABEL package in R (Aulchenko et al. 2007; R Core Team 2022). Significant SNPs were identified on the basis of Bonferroni significant cut-off calculated as $(0.05/389,997 = 1.28E-07)$, and for all traits with such associated SNPs in this study, the lambda value was ≤ 1.0 . For SNPs that failed to pass the Bonferroni cut-off, a p value adjustment for a lambda value > 1.0 was done, and a false discovery rate (FDR) was estimated using Benjamini–Hochberg (BH) method in GenABEL package (Benjamini and Hochberg 1995; Aulchenko et al. 2007). SNPs with FDR values ≤ 0.3 were considered as a significant. This FDR threshold is inclusive, but was found to give similar p value levels for most factors to what has commonly been used for suggestive significance in GWA studies in horses for example by Gmel et al. (2019) and Frischknecht et al. (2016). Moreover, linkage disequilibrium between significant SNPs on the same chromosome was calculated using the Gaston package in R (Perdry and Dandine-Roulland 2022), with the r^2 method.

2.5 | Functional Annotation

Putative candidate QTL(s) or gene(s) were investigated on the basis of ± 500 -kb windows around the detected significant SNPs, using EquCab 3.0 (GCF_002863925.1) assembly (Kalbfleisch et al. 2018) in Genome Data Viewer on National Centre for Biotechnology Information (NCBI) and BioMart on Ensemble (Sayers et al. 2022; Martin et al. 2023).

2.6 | Least Squares Means

To illustrate the association of the top SNPs detected to be significant with factors in this study, least squares means were estimated using the mixed procedure in the SAS software (SAS Institute Inc. 2015). The effect of event ($N = 50$) and sex (male or female) and one top SNP at the time was included in the model for each trait. Three event classes with fewer than three genotyped horses each were merged with classes of events held at the same place but in a different year. This was only done for the estimation of least squares means and not for the ssGWAS, because the number of observations per event was reduced when including only genotyped horses.

3 | Results

Significant associations based on either Bonferroni correction or FDR after correction for the lambda value were found for three of the four factors for evaluated traits and for six of the 13 studied factors for linearly scored traits, and those are presented

below. Estimated least squares means are shown in Table S3 and Figure S2, and illustrate the size of the impact from genotypes for the detected top SNPs on the phenotypes in this data. In most cases, the least squares mean estimates points at substantial differences in relation to size of the standard deviations (Table S1) for the factor traits. However, likely due to low genotype frequency in some cases, not all genotypes had significantly different effects.

3.1 | Factors Related to Body Height and Size

For the factor E.size, where height at withers and the evaluated traits type and head–neck–body were the most important traits, two SNPs (AXE-104376260 and AXE-103658351) surpassed the Bonferroni cut-off significance level (Table 1, Table S2, Figure 1a [Manhattan plot] and Figure 2a [QQ-plot]). Both SNPs were located on ECA3, in genomic position 107,641,233 and 107,643,810bp, respectively. The LD between these two SNPs was estimated as 0.72 (Figure S1a). The top SNP for E.size (AXE-104376260) was located within 3'UTR of the *DDB1* and *CUL4 Associated Factor 16 (DCAF16)* gene and overlapping 5'UTR of the *Family with Sequence Similarity 184 Member B (FAM184B)* gene.

The same SNP AXE-104376260 also surpassed the Bonferroni significance cut-off for the factor L.height, dominated by height at withers and the linearly scored traits withers (high to low), and body proportions (long legged to short legged) (Table 1, Table S2 and Figures 1b and 2b). This SNP was in LD with two other close/neighbour SNPs: AXE-103658351 and AXE-103175342 with LD-values of 0.72 and 0.76, respectively, and those SNPs were also significantly associated with L.type (Figure S1a).

For the factor L.type for which height at withers and the linearly scored trait type (refined to heavy) loaded most strongly, 26 SNPs located in the genomic region ECA3:g.106,606,527–108,232,943, including both SNPs mentioned above detected for E.size and L.height, surpassed the Bonferroni significance cut-off (Table 1, Table S2 and Figures 1c and 2c). The LD between all 26 SNPs ranged from 0.04 to 0.99 (Figure S1a). The top SNP AX-103036686 for L.type was located at ECA3:g.107,115,309bp, equivalent to 318-kb upstream of the *ligand dependent nuclear receptor corepressor like (LCORL)* gene. Three other significantly associated SNPs for L.type (AX-102966003, AX-103096129 and AX-103175342) are located in introns within the *LCORL* gene (at positions ECA3:g.107,456,470, ECA3:g.107,496,015 and ECA3:g.107,528,272bp). Besides the SNP AX-104376260 within *DCAF16* (3'UTR) and *FAM184B* (5'UTR), three other significant SNPs were located within introns of the *FAM184B* gene at ECA3:g.107,643,810, ECA3:g.107,685,587 and ECA3:g.107,702,189bp. Also, the significant SNP AX-104783182 at ECA3:g.107,852,187bp was located in the 3'UTR of the gene *clarin 2 (CLRN2)* (Table S2).

Other genes within ± 500 kb to the significant SNPs for the three factors related to body size, especially height, were the *non-SMC condensin I complex subunit G (NCAPG)*, *mediator complex*

subunit 28 (MED28), *leucine aminopeptidase 3 (LAP3)* and *quinoid dihydropteridine reductase (QDPR)* (Table 1).

3.2 | Factor Related to Body Length

For the factor L.length, comprising the linearly scored traits body length (long to short), loins (long to short) and length of neck (long to short), the most significant SNP (AXE-104176662) was located on ECA15:g.5,021,795bp. This SNP surpassed the FDR significance level with a FDR value of 0.13 (Table 1, Table S2 and Figures 3a and 4a) and was upstream of *transforming growth factor-beta receptor associated protein 1 (TGFBRA1)*, *G protein-coupled receptor 45 (GPR45)*, *mitochondrial ribosomal protein S9 (MRPS9)* and *POU class 3 homeobox 3 (POU3F3)* (Table 1).

3.3 | Factors Related to Other Conformation Traits

For the factor E.conf, dominated by the evaluated traits type and head–neck–body, one SNP (AX-105004994) located on ECA26:g.30,188,501bp passed the FDR threshold with a FDR value of 0.29 (Table 1, Table S2 and Figures 3b and 4b). The top SNP for this factor was located intergenic, in a region comprising the genes *superoxide dismutase 1 (SOD1)*, *SR-related CTD associated factor 4 (SCAF4)*, *hormonally up-regulated Neu-associated kinase (HUNK9)*, *ENSECAG00000037230*, *MIS18 kinetochore protein A (MIS18A)*, *melanocortin 2 receptor accessory protein (MRAP)*, *URB1 ribosome biogenesis homologue (URB1)*, *eva-1 homologue C (EVA1C)*, *CFAP298-TCP10L readthrough (CFAP298-TCP10L)* and *synaptojanin 1 (SYNJ1)* (Table 1).

Three SNPs passed the FDR significance threshold for the factor L.neck dominated by the linearly scored traits position of neck (vertical to horizontal), shape of neck (arched to straight) and position of shoulder (sloping to straight). Two of the SNPs (AX-103450830 and AX-104958036) were located on ECA29: g.23,384,193 and ECA29: g.23,400,348bp, and had FDR values of 0.19 and 0.28, respectively (Table 1, Table S2 and Figures 3c and 4c). The LD value between these two SNPs was 0.89 (Figure S1b). These two SNPs were located within the *calcium/calmodulin-dependent protein kinase ID (CAMK1D)* gene. Besides that, nearby genes included *coiled-coil domain containing 3 (CCDC3)*, *cell division cycle 123 (CDC123)*, *nudix hydrolase 5 (NUDT5)*, *SEC61 translocon subunit alpha 2 (SEC61A2)*, *dehydrogenase E1 and transketolase domain containing 1 (DHTKD1)*, *UPF2 regulator of nonsense mediated mRNA decay (UPF2)*, *proline and serine rich 2 (PROSER2)* and *enoyl-CoA hydratase domain containing 3 (ECHDC3)* (Table 1).

The third SNP (AX-103936375) that passed the FDR threshold for L.neck was located in an intergenic region on ECA14:g.17,446,349bp (Table 1 and Figures 3c and 4c) and had an FDR value of 0.28 (Table S2). The nearest genes to this SNP were *gamma-aminobutyric acid type A receptor subunit alpha1 (GABRA1)*, *gamma-aminobutyric acid type A receptor subunit alpha6 (GABRA6)* and *gamma-aminobutyric acid type A receptor subunit beta2 (GABRB2)*, and the detected SNP was located upstream of the *GABRB2* gene (Table 1).

TABLE 1 | Associated SNP(s) with FDR ≤ 0.3 for factors in 380 SWB horses and nearest annotated protein coding genes within ±500 kb from the SNP(s).

Factor	λ^a	ECA ^b	Top SNP			All significant SNP(s) in the region		
			Marker name	Pos. (bp)	FDR ^c	p^d	Position range (bp)	Protein coding gene(s) ^f
E.size	1.00	3	AX-104376260	107,641,233	0.02	4.56E-08 ^g	107,641,233–107,643,810	LCORL, NCAPG, DCAF16 , FAM184B , MED28, LAP3, CLRN2, QDPR
L.height	0.94	3	AX-104376260	107,641,233	0.01	1.13E-07 ^g	107,641,233	LCORL, NCAPG, DCAF16 , FAM184B , MED28, LAP3, CLRN2, QDPR
L.type	0.99	3	AX-103036686	107,115,309	3.4E-11	2.67E-16 ^g	106,606,527–108,232,943	LCORL , NCAPG, DCAF16 , FAM184B , MED28, LAP3, CLRN2 , QDPR, LDB2
L.length	1.10	15	AX-104176662	5,021,795	0.13	3.26E-07	5,021,795	TGFBRAP1, GPR45, MRPS9, POU3F3
E.conf	1.14	26	AX-105004994	30,188,501	0.29	7.50E-07	30,188,501	SOD1, SCAF4, HUNK, ENSECAG00000037230, MIS18A, MRAP, URBI, EVA1C, CFAP298-TCPI0L, SYNJ1
L.neck	1.17	29	AX-103450830	23,384,193	0.19	4.81E-07	23,384,193–23,400,348	CCDC3, CAMK1D , CDC123, NUDT5, SEC61A2, DHTKDI, UPF2, PROSER2, ECHDC3, USP6NL
E.jump	1.79	18	AX-103936375 AX-105015252	17,446,349 9,956,043	0.28 0.14	1.47E-06 3.62E-07	17,446,349 9,956,043	GABRA1, GABRA6, GABRB2 CLASPI, TFPC2LI, ENSECAG00000051759, GLI2
L.walk	1.18	17	AX-104203221; AX-105012856	75,955,220; 75,955,247	0.29	2.22E-06	75,955,220–75,955,247	NALF1, LIG4, ABHD13, TNFSF13B, MYO16
		22	AX-104667493	28,543,491	0.29	1.86E-06	28,543,491	RAB5IF, SLA2, NDRG3, SOGA1, TLDC2, SAMHDI, RBL1, MROH8, RPN2, GHRH, MANBAL, SRC, SOGA1, TLDC2, SAMHDI, RBL1, MROH8, RPN2, GHRH, MANBAL, SRC, ENSECAG00000052590, BLCAP, NNAT, CTNNBL1
L.trot.hind	1.35	15	AX-103369307	3,871,175	0.23	5.83E-07	3,871,175	UXS1, ECRG4, NCK2, ENSECAG00000051693

^aLambda value before correction.
^bEquine chromosome number.
^cFalse discovery rate.
^dAdjusted for lambda value when $\lambda > 1$.
^eNumber of significant SNPs in the region.
^fGenes in which any significant SNPs are located are marked in bold.
^gSurpassed the Bonferroni significant threshold.

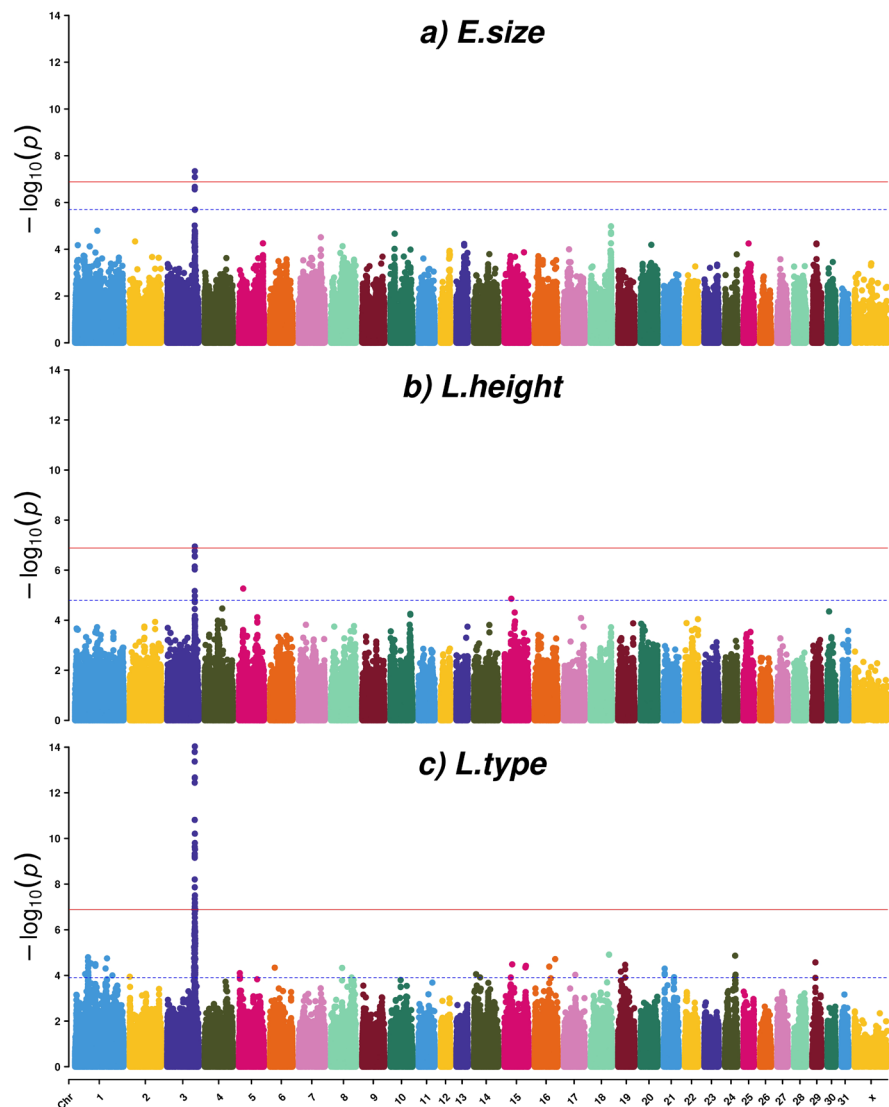


FIGURE 1 | Manhattan plots for three factors for evaluated (E) and linearly scored (L) traits related to body height and size in SWB horses: (a) E.size, (b) E.height and (c) L.type. The red line indicates the Bonferroni corrected significance threshold ($0.05/389,997 = 1.28\text{E-}07$), and the blue dotted line indicates the FDR (≤ 0.3) significance threshold. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

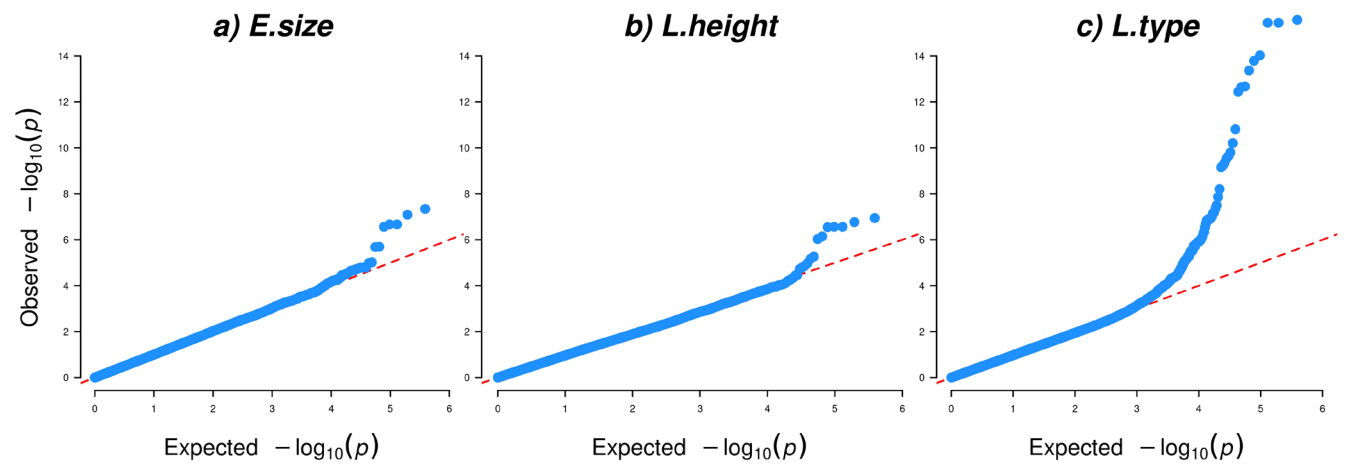


FIGURE 2 | Quantile–quantile (QQ) plots for three factors for evaluated (E) and linearly scored (L) traits related to body height and size in SWB horses: (a) E.size, (b) E.height and (c) L.type. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

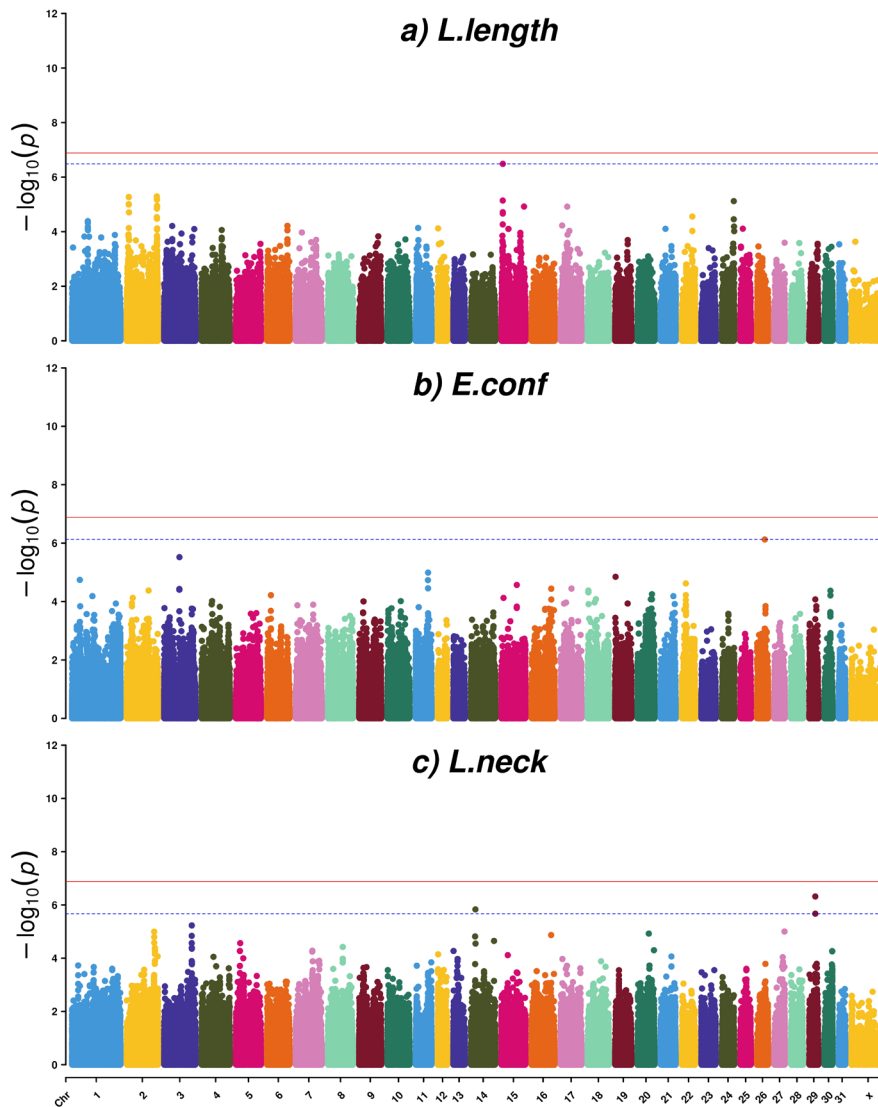


FIGURE 3 | Manhattan plots for factors for (a) linearly scored body length traits (L.length), (b) evaluated conformation traits (E.conf) and (c) linearly scored neck conformation traits (L.neck), in SWB horses. The red line indicates the Bonferroni corrected threshold ($0.05/389,997 = 1.28E-07$), and the blue dotted line indicates the FDR (≤ 0.3) significance threshold. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

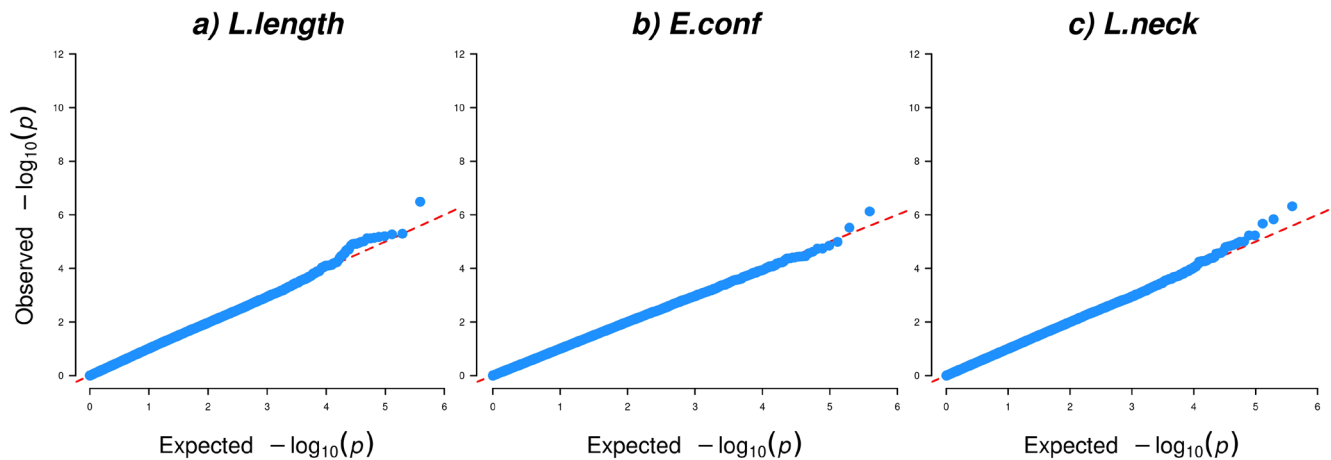


FIGURE 4 | Quantile–quantile (QQ) plots for factors for (a) linearly scored body length traits (L.length), (b) evaluated conformation traits (E.conf) and (c) linearly scored neck conformation traits (L.neck), in SWB horses. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

3.4 | Factor for Jumping

For the factor E.jump, collecting the two evaluated traits free-jumping technique and ability, and free-jumping temperament and general impression, one SNP (AX-105015252) in position ECA18:g.9,956,043 bp passed the FDR significance level, with a value of 0.14 (Table 1, Table S2 and Figures 5a and 6a). This SNP had a LD of 0.76 with its nearby SNP (AX-104534955) (Figure S1c). These SNPs were located within the *GLI family zinc finger 2 (GL2)* gene. Other genes close to the top SNP included *cytoplasmic linker associated protein 1 (CLASP1)* and *transcription factor CP2 like 1 (TFCP2L1)* and *ENSECAG00000051759* (Table 1).

3.5 | Factors for Walk and Trot

Three SNPs surpassed the FDR significance threshold with the FDR value of 0.29 for the factor L.walk that was dominated by the linearly scored traits walk stride length (long to

short), walk cadence (even to uneven) and walk suppleness (supple to stiff) (Table S2). Two of these SNPs (AX-104203221 and AX-105012856) were completely in LD with each other and located intergenic on ECA17: g.75,955,220 and ECA17: g.75,955,247 bp, respectively (Table 1, Figures 5b and 6b, and Figure S1d). The closest genes in the ± 500 kb window included *DNA ligase 4 (LIG4)*, *abhydrolase domain containing 13 (ABHD13)*, *TNF superfamily member 13b (TNFSF13B)* and *myosin XVI (MYO16)* (Table 1).

The third SNP that passed the FDR threshold for L.walk (AX-104667493) was located intergenic on ECA22:g.28,543,491 bp (Table 1 and Figures 5b and 6b). This chromosome region is quite gene dense and involves the genes *RAB5 interacting factor (RAB5IF)*, *Src like adaptor 2 (SLA2)*, *NDRG family member 3 (NDRG3)*, *suppressor of glucose autophagy (SOGA1)*, *TBC/LysM-associated domain containing 2 (TLDC2)*, *SAM and HD domain containing deoxynucleoside triphosphate triphosphohydrolase 1 (SAMHD1)*, *RB transcriptional corepressor like 1 (RBL1)*, *maestro heat like repeat family member 8 (MROH8)*, *ribophorin II (RPN2)*,

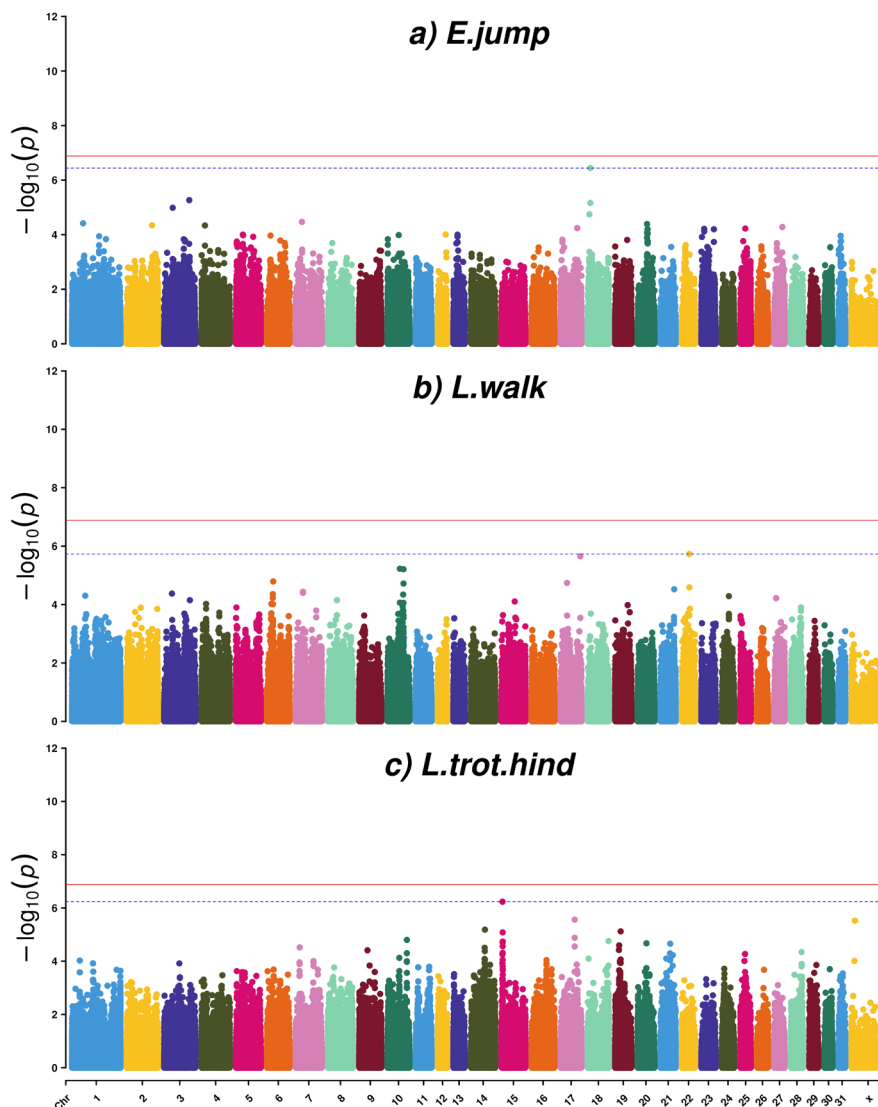


FIGURE 5 | Manhattan plots for factors for (a) evaluated jumping traits (E.jump), (b) linearly scored walk traits (L.walk) and (c) linearly scored trot traits with focus on the hind legs (L.trot.hind), in SWB horses. The red line indicates the Bonferroni corrected threshold ($0.05/389,997 = 1.28E-07$), and the blue dotted line indicates the FDR (≤ 0.3) significance threshold. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/jbg.12923)]

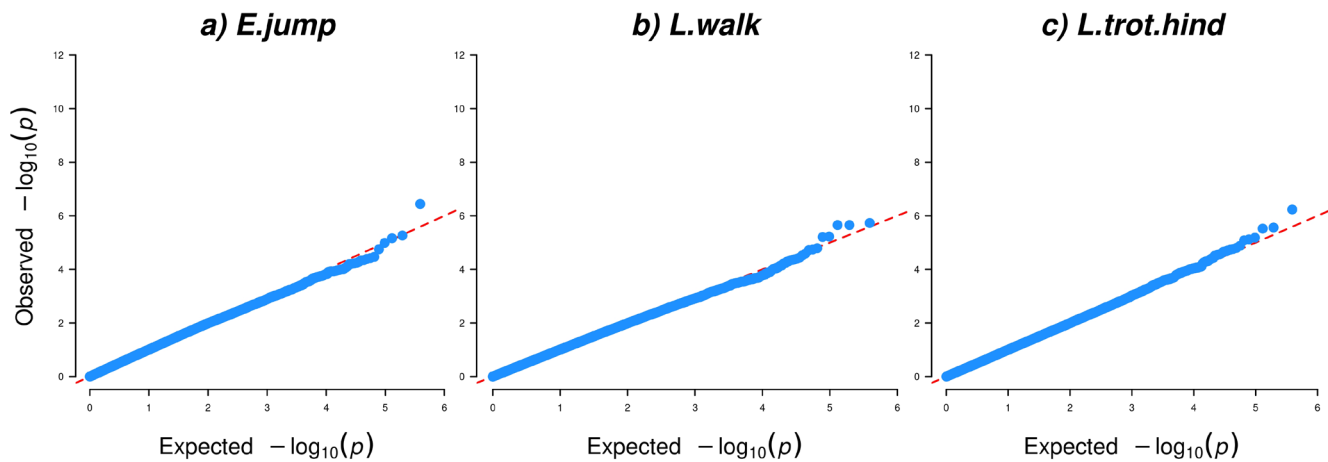


FIGURE 6 | Quantile–quantile (QQ) plots for factors for (a) evaluated jumping traits (E.jump), (b) linearly scored walk traits (L.walk) and (c) linearly scored trot traits with focus on the hind legs (L.trot.hind), in SWB horses. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

growth hormone releasing hormone (GHRH), *mannosidase beta like (MANBAL)*, *SRC proto-onco non-receptor tyrosine kinase (SRC)*, *ENECAG00000052590*, *BLCAP apoptosis inducing factor (BLCAP)*, *neuronatin (NNAT)* and *catenin beta like 1 (CTNNB1)* were the closest genes to this SNP (Table 1).

Finally, for the factor L.trot.hind on which the linearly scored traits trot hind leg activity (active to inactive), trot hind leg position (under the body to behind the body) and trot elasticity (elastic to inelastic) loaded most strongly. One FDR significant SNP (AX-103369307) located intergenic on ECA15:g.3,871,175bp with the FDR value of 0.23 was found (Table 1, Table S2 and Figures 5c and 6c). The genes *UDP-glucuronate decarboxylase 1 (UXS1)*, *ECRG4 augurin (ECRG4)* and *NCK adaptor protein 2 (NCK2)* and *ENSECAG00000051693* were found within ± 500 -kb window (Table 1).

4 | Discussion

In this study, single-step GWAS was used to identify plausible regions by estimating SNP effects and p values, including also non-genotyped horses. The analysis was based on factors instead of all individual evaluated and linearly scored traits. This was done to reduce the number of traits to a manageable number of latent traits with a meaningful interpretation (Nazari-Ghadikolaei et al. 2023). It has been demonstrated that estimated marker effect divided by its standard error obtained from GBLUP and standard single-marker GWAS methods are equivalent (Gualdrón Duarte et al. 2014; Aguilar et al. 2019; Legarra, Ricard, and Varona 2018). The single-step approach in which genomic breeding values were predicted and then back solved to SNP effects and p values, enabled adjustments of the phenotypes for effects in the model as well as correction for genetic structure in the population. It also made it possible to utilise phenotypic information from a large number of horses that were related to the relatively few genotyped individuals included in this study, which makes the use of single-step methodology attractive for this type of data.

It should be noted that all the studied traits are complex traits, based on subjective measures and likely influenced by many

genes and environmental influence, and that some missing factor scores were also obtained by imputation. This, in combination with the relatively small number of genotyped individuals, only allows for detection of genomic regions with large influence on the trait. In the end, the factors for which we found significant associations were not depending heavily on traits for which many observations had to be imputed, and thus, the effect of imputation for the presented results can be assumed to be minor.

The Bonferroni method is a simple and common correction for multiple testing, but it does not consider that some SNPs are in high LD with each other and thus not independent. Also, the method to minimise the family-wise error rate has been suggested to be too strict in many multiple test settings (Noble 2009). In contrast, to control the FDR has been proposed to be a reliable choice in the context of multiple testing, that reduces the risk of unnecessarily removing potentially important results (Glickman, Rao, and Schultz 2014). For this reason, we chose to use the less conservative FDR threshold in this study, in addition to the well-known Bonferroni corrected significance level. This FDR threshold gave similar suggestive associations as applying p value levels commonly used for suggestive significance in GWA studies in horses for example by Gmel et al. (2019) and Frischknecht et al. (2016).

The percentage of variance explained according to the results from the ssGWAS (Table S2) was in general considerably smaller for individual top SNPs than what has been indicated from previous studies, for example, for SNPs in the vicinity of the *LCORL* gene for height at withers by Vosgerau et al. (2022) who also used a single-step approach but another software. The difference is due to the different methods used and it should be noted that many SNPs in the region together explain more of the variation, and those were not considered in the calculation of the percentage of variance explained. To illustrate that the detected top SNPs did in fact seem to have a substantial influence on the studied factors, we included also estimations of least squares means where only genotypes for the top SNP were included for each trait, and then, a larger effect was as expected observed.

4.1 | Factors Related to Body Height and Size

SNPs in the same region on ECA3 were detected for the three factors related to body size, primarily height at withers: E.size, L.height and L.type. This region contains the *LCORL* and *NCAPG* genes, reported to influence height and size variation in horses and other species (Pryce et al. 2011; Signer-Hasler et al. 2012a; Tetens et al. 2013; Staiger et al. 2016; Reich et al. 2024). In horses, a lower expression level of *LCORL* may be associated with larger and heavier horses (Metzger et al. 2013). Moreover, a well-documented QTL close to *LCORL* gene has been identified for withers height using ssGWAS in German Warmblood horses (Vosgerau et al. 2022). A recent comprehensive study of German Warmblood horses by Reich et al. (2024) also reported a possible, albeit less likely, effect of the gene *DCAF16* in addition to *LCORL* and *NCAPG* for body size.

The *LCORL* and *NCAPG* genes in addition to *FAM184B* gene have also been associated with conformation and locomotive traits in Spanish purebred horses (Sevane et al. 2017). Also, a study of Simmental cattle showed possible associations between *LCORL*, *NCAPG* and *FAM184B* genes with bone weight in this breed (Xia et al. 2017). In another study, the *DCAF16*-*NCAPG* region was identified as potential candidate region responsible for average daily weight gain in Simmental beef cattle (Zhang et al. 2016).

4.2 | Factor Related to Body Length

Association mapping for the factor L.length, based on different traits describing length of the horse, could identify three genes on ECA15 within ± 500 kb from the FDR significantly associated SNP. One of the genes was *TGFBRAP1*, which encodes a protein that binds to *transforming growth factor-beta* (*TGF-beta*) receptors, and plays a key role in *TGF-beta* signalling pathway, which was one of the overrepresented gene ontology terms in SWB non-show jumping horses (Abbondi, Eriksson, et al. 2019). The *TGF-beta* superfamily plays important roles as morphogens during embryogenesis and is involved in tissue differentiation and in the establishment of body-axes (Messler et al. 2011), which makes it a possible candidate for a trait related to the length of the longitudinal axis. *TGF-beta* was previously found to play a role for muscle development and feed efficiency in beef cattle (Alexandre et al. 2019; Chen et al. 2021). Also, the *POU3F3* gene, located nearby, is important for embryo development (Safran et al. 2021).

4.3 | Factors Related to Other Conformation Traits

For the E.conf factor dominated by the evaluated traits type and head-neck-body, association mapping could detect nine genes on chromosome 26 within ± 500 kb from the FDR significantly associated SNP. Among detected genes, the *EVA1C* gene is related to X-Linked Intellectual Disability-Short Stature-Overweight Syndrome. Some of the symptoms of the syndrome include short stature, elevated body mass index, a pattern of truncal obesity (in older males), and variable neurologic features (Rappaport et al. 2013). As *EVA1C* has some association with short stature and body mass index, and E.conf is composed of

type and head-neck-body of the horse, it may be interesting to investigate whether *EVA1C* could be of importance for conformation in SWB horses.

For the L.neck factor based on traits describing the shape and position of neck and shoulders, nine genes were identified within the ± 500 kb window on ECA29. The two top SNPs were located within the *CAMK1D* gene, and a member of the calcium/calmodulin-dependent protein kinase 1 family, a subfamily of the serine/threonine kinases. The encoded protein is a component of the calcium-regulated calmodulin-dependent protein kinase cascade (Safran et al. 2021). *CAMK1D* has been suggested to play a role for liver gluconeogenesis, fat mass deposition, obesity and reduced insulin sensitivity (Rausch et al. 2018; Fromont et al. 2020). If fat deposition around the neck contributes to a more rounded/arched neck shape in SWB horses, then the *CAMK1D* could be a candidate gene for this trait. Enlarged fat depositions at the crest of the neck appear to be associated with insulin resistance and risk of laminitis in horses (Geor 2008; Carter et al. 2009).

The nearby *UPF2* and *ECHDC3* genes were suggested to be associated with pin width in Chinese Holstein cattle (Lu et al. 2021), and the *ECHDC3* gene plays a role for fatty acid biosynthesis and insulin sensitivity (Safran et al. 2021). Thus, this ECA29 region could be important for energy metabolism in the horse.

A potential genetic connection between the shape of the neck of horses and fat deposition and energy metabolism has been suggested previously. Gmel, Brem, and Neuditschko (2023) found in Lipizzan horses that a SNP within the gene, *Membrane Associated Guanylate Kinase, WW And PDZ Domain Containing 1* (*MAGI1*) on ECA16, was significantly associated with the shape and width of the neck. The *MAGI1* gene has also been suggested to be associated with insulin resistance and glucose response in humans (Palmer et al. 2010; Norris and Rich 2012; Ellis et al. 2015).

Also, the skeletal framework of the neck can be assumed to influence its conformation. In the present study, among the genes nearby the associated SNP on chromosome 14, the genes *GABRA1*, *GABRB2* and, especially, *GABRA6* have previously been suggested to be associated with osteochondrosis in Standardbreds (McCoy et al. 2016), and Hanoverian warmblood horses (*GABRA6*) (Naccache, Metzger, and Distl 2018). Osteochondral fragmentation of the cervical articular process joints in the neck of horses has been reported (Tucker et al. 2022). These genes are otherwise known for their importance in the central nervous system (Safran et al. 2021).

4.4 | Factor for Jumping

The SNP located within the *GLI2* gene on ECA18 passed FDR significance threshold for the factor E.jump, incorporating jumping ability, technique, attitude and general impression. This gene codes for the transcription factor Zinc finger protein GLI2, that is mediating the Sonic hedgehog (Shh) signalling, and is thought to play a role during embryogenesis (Safran et al. 2021).

Loss of function in the *GLI2* gene has been shown to cause microphthalmia and decreased head width in homozygous mice

(Heyne et al. 2016). The *GLI*-mediated *CyclinD2* is important for the development of the binocular circuit, when retinal ganglion cells (RGC) from the retina project axons to the brain (Slavi et al. 2023). The same study by Slavi et al. (2023) found that the depth perception was compromised in *CyclinD2*-deficient mice. It would be interesting to study further whether the *GLI2* gene could be of importance for distance estimation in show jumping horses.

The window of ± 500 kb from the significant SNP also included the gene *CLASPI*, for which mutations have been related to disorders, such as epiphyseal dysplasia that affect cartilage and bone development, especially in the long bones in the limbs, short stature, microcephaly and congenital nystagmus (Safran et al. 2021).

4.5 | Factors for Walk and Trot

For L.walk, based on traits describing different aspects of the gait walk, there were five genes in the studied window around the top SNPs on ECA17 and ECA22, and 25 genes in the ± 500 -kb window on ECA22. Among the detected genes on ECA17, *ABHD13* is related to the rare Williams syndrome in humans that cause growth delay, short stature, mental deficiency and some facial features developing with age (Safran et al. 2021). Another nearby gene *TNFSF13B* encodes the cytokine protein belonging to the tumour necrosis factor (TNF) ligand family, and has been associated with body weight gain in beef steers (Lindholm-Perry et al. 2017). Also, the *MYO16* gene in this region has been shown to play a role in cell proliferation and body size in cattle and sheep (Mastranestasis et al. 2016; Gonzalez et al. 2020).

Among the genes within ± 500 kb from the significant SNP on ECA22, the *RBL1*, *SRC* and *SAMHD1* genes were previously detected among overrepresented genes in a GO term analysis for show jumping SWB horses (Ablondi, Eriksson, et al. 2019). All three genes were identified as a potential candidate genes associated with mentality and postsynaptic signalling in show jumping SWB horses (Ablondi, Eriksson, et al. 2019). In addition to this, *SRC* has a role in osteoporosis in mice due to non-functional osteoclasts (Lowe et al. 1993).

Among other genes in the window on ECA22, the *SOGA1* gene encodes a protein that reduce glucose production and plays a role in autophagy (Cowerd et al. 2010). Other detected genes were *GHRH*, known for its role in growth in human and animals, such as carcass traits in Korean cattle (Cheong et al. 2006), and *CTNBL1* that is associated with body weight and height in humans (Andreasen et al. 2009), and muscle formation and body weight in cattle (Mancini et al. 2014).

An association of these genes with L.walk is not clear, except that the growth and conformation of the horse, for example, length of legs, may influence aspects like stride length. There are also other nearby genes in ECA22, like *NNAT*, that is involved in brain and nervous system development in humans (Safran et al. 2021), that could potentially influence gaits.

The top SNP for the factor L.trot.hind was located nearby three genes on ECA15. One of these was the *NCK2* gene that has

an impact on preosteoblastic and osteoblastic migration and bone mass in human by bone formation activity (Aryal and Noda 2015). As the L.trot.hind factor is related to hind leg position and activity, as well as elasticity during trot, it is possible that skeletal development influencing angles, and/or the area of muscle attachment, could be of importance for this trait.

4.6 | Additional Remarks

These findings need to be validated in future studies, and the function of the potential candidate genes related to the phenotypes remains to be clarified. The regions with SNPs associated with factors related to body height are well known from numerous previous studies in different species, which shows that the data quality and methods applied were sufficient to detect genomic regions of clear importance for the traits. Even for height at withers, the full genetic background of height in horses remains to be revealed. It should also be noted that besides protein coding genes, also, for example, regulatory regions could be of importance for the studied traits. Besides the value of learning more about the biology behind desired phenotypes, the results may also be of interest for future assessment and selection of horses. If, for example, a desired trait, such as an arched neck in SWB would be genetically associated with risk factors for developing laminitis, then that would need to be considered in the breeding goal.

5 | Conclusions

By the use of single-step genome-wide association analysis of factors for evaluated and linearly scored young horse test traits in SWB horses, we detected some plausible candidate regions and genes of interest to study further. These included a well-known region on ECA3 for factors related to body size, as well as more novel candidate genes for factors related to jumping and to the shape and position of the neck, both for which the detected top SNPs were found within genes.

Author Contributions

Susanne Eriksson, Sofia Mikko and Åsa Gelinder Viklund initiated the study. Susanne Eriksson and W. Freddy Fikse designed and supervised the study. Åsa Gelinder Viklund provided pre-edited data and trait information. Sofia Mikko was main applicant for the funding of the genotyping. Anahit Nazari-Ghadikolaei carried out the data analysis with contributions from W. Freddy Fikse, Susanne Eriksson and Anahit Nazari-Ghadikolaei drafted the manuscript. Sofia Mikko scrutinised the detected genomic regions. All authors read, edited and approved the final manuscript.

Acknowledgements

The authors wish to thank the Swedish Warmblood Association for providing phenotype and pedigree data.

Ethics Statement

Not applicable as the hair samples from Swedish Warmblood horses were originally collected for parentage testing and stored in the biobank at the Animal Genetics Laboratory, SLU. The Swedish Warmblood

Association approved the samples to be used in this research. No additional samples were collected for this study.

Conflicts of Interest

The Swedish Warmblood Association has provided the phenotype and pedigree data for this study, and Åsa Gelinder Viklund has regular commitments to Swedish Warmblood Association, regarding the routine genetic evaluation. We declare that there are no other conflicts of interest.

Data Availability Statement

Restrictions apply to the availability of these data, which were used under licence for this study. Data are available from the authors with the permission of the Swedish Warmblood Association.

References

- Abdi, H. 2003. *Factor Rotations in Factor Analyses, Encyclopedia for Research Methods for the Social Sciences*, 792–795. Thousand Oaks, CA: Sage.
- Ablondi, M., S. Eriksson, S. Tetu, A. Sabbioni, Å. Viklund, and S. Mikko. 2019. “Genomic Divergence in Swedish Warmblood Horses Selected for Equestrian Disciplines.” *Genes* 10, no. 12: 976. <https://doi.org/10.3390/genes10120976>.
- Ablondi, M., Å. Viklund, G. Lindgren, S. Eriksson, and S. Mikko. 2019. “Signatures of Selection in the Genome of Swedish Warmblood Horses Selected for Spurt Performance.” *BMC Genomics* 20: 1–12. <https://doi.org/10.1186/s12864-019-6079-1>.
- Aguilar, I., A. Legarra, F. Cardoso, Y. Masuda, D. Lourenco, and I. Misztal. 2019. “Frequentist p-Values for Large-Scale-Single Step Genome-Wide Association, With an Application to Birth Weight in American Angus Cattle.” *Genetics Selection Evolution* 51, no. 1: 1–28. <https://doi.org/10.1186/s12711-019-0469-3>.
- Aguilar, I., I. Misztal, D. Johnson, A. Legarra, S. Tsuruta, and T. Lawlor. 2010. “Hot Topic: A Unified Approach to Utilize Phenotypic, Full Pedigree, and Genomic Information for Genetic Evaluation of Holstein Final Score.” *Journal of Dairy Science* 93, no. 2: 743–752. <https://doi.org/10.3168/jds.2009-2730>.
- Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. “PREGSF90–POSTGSF90: Computational Tools for the Implementation of Single-Step Genomic Selection and Genome-Wide Association With Ungenotyped Individuals in BLUPF90 Programs.” In *10. World Congress on Genetics Applied to Livestock Production (WCGALP)*.
- Alexandre, P. A., M. Naval-Sanchez, L. R. Porto-Neto, J. B. S. Ferraz, A. Reverter, and H. Fukumasu. 2019. “Systems Biology Reveals NR2F6 and TGFB1 as Key Regulators of Feed Efficiency in Beef Cattle.” *Frontiers in Genetics* 10: 230. <https://doi.org/10.3389/fgene.2019.00230>.
- Andreasen, C. H., M. S. Mogensen, K. Borch-Johnsen, et al. 2009. “Studies of CTNBL1 and FDFT1 Variants and Measures of Obesity: Analyses of Quantitative Traits and Case-Control Studies in 18,014 Danes.” *BMC Medical Genetics* 10: 1–9. <https://doi.org/10.1186/1471-2350-10-17>.
- Aryal, A. C. S., K. Miyai, Y. Izu, et al. 2015. “Nck Influences Preosteoblastic/Osteoblastic Migration and Bone Mass.” *Proceedings of the National Academy of Sciences* 112, no. 50: 15432–15437. <https://doi.org/10.1073/pnas.1518253112>.
- Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. Van Duijn. 2007. “GenABEL: An R Library for Genome-Wide Association Analysis.” *Bioinformatics* 23, no. 10: 1294–1296. <https://doi.org/10.1093/bioinformatics/btm108>.
- Benjamini, Y., and Y. Hochberg. 1995. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.” *Journal of the Royal Statistical Society: Series B: Methodological* 57, no. 1: 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Bonow, S., S. Eriksson, E. Thorén Hellsten, and Å. Gelinder Viklund. 2023. “Consequences of Specialized Breeding in the Swedish Warmblood Horse Population.” *Journal of Animal Breeding and Genetics* 140, no. 1: 79–91. <https://doi.org/10.1111/jbg.12731>.
- Brard, S., and A. Ricard. 2015. “Genome-Wide Association Study for Jumping Performances in F Rench Sport Horses.” *Animal Genetics* 46, no. 1: 78–81. <https://doi.org/10.1111/age.12245>.
- Carter, R. A., K. Treiber, R. Geor, L. Douglass, and P. A. Harris. 2009. “Prediction of Incipient Pasture-Associated Laminitis From Hyperinsulinaemia, Hyperleptinaemia and Generalised and Localised Obesity in a Cohort of Ponies.” *Equine Veterinary Journal* 41, no. 2: 171–178. <https://doi.org/10.2746/042516408X342975>.
- Chen, W., P. A. Alexandre, G. Ribeiro, et al. 2021. “Identification of Predictor Genes for Feed Efficiency in Beef Cattle by Applying Machine Learning Methods to Multi-Tissue Transcriptome Data.” *Frontiers in Genetics* 12: 619857. <https://doi.org/10.3389/fgene.2021.619857>.
- Cheong, H. S., D.-H. Yoon, L. H. Kim, et al. 2006. “Growth Hormone-Releasing Hormone (GHRH) Polymorphisms Associated With Carcass Traits of Meat in Korean Cattle.” *BMC Genetics* 7, no. 1: 1–6. <https://doi.org/10.1186/1471-2156-7-35>.
- Cowder, R. B., M. M. Asmar, J. M. Alderman, et al. 2010. “Adiponectin Lowers Glucose Production by Increasing SOGA.” *American Journal of Pathology* 177, no. 4: 1936–1945. <https://doi.org/10.2353/ajpath.2010.100363>.
- de Oliveira Bussiman, F., B. A. dos Santos, B. D. C. A. Silva, et al. 2020. “Genome-Wide Association Study: Understanding the Genetic Basis of the Gait Type in Brazilian Mangalarga Marchador Horses, A Preliminary Study.” *Livestock Science* 231: 103867. <https://doi.org/10.1016/j.livsci.2019.103867>.
- Duensing, J., K. F. Stock, and J. Krieter. 2014. “Implementation and Prospects of Linear Profiling in the Warmblood Horse.” *Journal of Equine Veterinary Science* 34, no. 3: 360–368. <https://doi.org/10.1016/j.jevs.2013.09.002>.
- Ellis, K., Y. Zhou, J. Beshansky, et al. 2015. “Genetic Modifiers of Response to Glucose–Insulin–Potassium (GIK) Infusion in Acute Coronary Syndromes and Associations With Clinical Outcomes in the IMMEDIATE Trial.” *Pharmacogenomics Journal* 15, no. 6: 488–495. <https://doi.org/10.1038/tpj.2015.10>.
- Frischknecht, M., H. Signer-Hasler, T. Leeb, S. Rieder, and M. Neuditschko. 2016. “Genome-Wide Association Studies Based on Sequence-Derived Genotypes Reveal New QTL Associated With Conformation and Performance Traits in the Franches–Montagnes Horse Breed.” *Animal Genetics* 47, no. 2: 227–229. <https://doi.org/10.1111/age.12406>.
- Fromont, C., A. Atzori, D. Kaur, et al. 2020. “Discovery of Highly Selective Inhibitors of Calmodulin-Dependent Kinases That Restore Insulin Sensitivity in the Diet-Induced Obesity In Vivo Mouse Model.” *Journal of Medicinal Chemistry* 63, no. 13: 6784–6801. <https://doi.org/10.1021/acs.jmedchem.9b01803>.
- Geor, R. J. 2008. “Metabolic Predispositions to Laminitis in Horses and Ponies: Obesity, Insulin Resistance and Metabolic Syndromes.” *Journal of Equine Veterinary Science* 28, no. 12: 753–759. <https://doi.org/10.1016/j.jevs.2008.10.016>.
- Glickman, M. E., S. R. Rao, and M. R. Schultz. 2014. “False Discovery Rate Control Is a Recommended Alternative to Bonferroni-Type Adjustments in Health Studies.” *Journal of Clinical Epidemiology* 67, no. 8: 850–857. <https://doi.org/10.1016/j.jclinepi.2014.03.012>.
- Gmel, A., G. Brem, and M. Neuditschko. 2023. “New Genomic Insights Into the Conformation of Lipizzan Horses.” *Scientific Reports* 13, no. 1: 8990. <https://doi.org/10.1038/s41598-023-36272-4>.

- Gmel, A. I., T. Druml, R. von Niederhäusern, T. Leeb, and M. Neuditschko. 2019. "Genome-Wide Association Studies Based on Equine Joint Angle Measurements Reveal New QTL Affecting the Conformation of Horses." *Genes* 10, no. 5: 370. <https://doi.org/10.3390/genes10050370>.
- Gonzalez, M., R. Villa, C. Villa, et al. 2020. "Inspection of Real and Imputed Genotypes Reveled 76 SNPs Associated to Rear Udder Height in Holstein Cattle." *Journal of Advanced Veterinary and Animal Research* 7, no. 2: 234–241. <https://doi.org/10.5455/javar.2020.g415>.
- Gualdrón Duarte, J., R. Cantet, R. Bates, C. Ernst, N. Raney, and J. Steibel. 2014. "Rapid Screening for Phenotype-Genotype Associations by Linear Transformations of Genomic Evaluations." *BMC Bioinformatics* 15, no. 1: 246. <http://www.biomedcentral.com/1471-2105/15/246>.
- Heyne, G. W., J. L. Everson, L. J. Ansen-Wilson, et al. 2016. "Gli2 Gene-Environment Interactions Contribute to the Etiological Complexity of Holoprosencephaly: Evidence From a Mouse Model." *Disease Models & Mechanisms* 9, no. 11: 1307–1315. <https://doi.org/10.1242/dmm.026328>.
- Kalbfleisch, T. S., E. S. Rice, M. S. DePriest, et al. 2018. "Improved Reference Genome for the Domestic Horse Increases Assembly Contiguity and Composition." *Communications Biology* 1, no. 1: 1–8. <https://doi.org/10.1038/s42003-018-0199-z>.
- Legarra, A., O. F. Christensen, I. Aguilar, and I. Misztal. 2014. "Single Step, a General Approach for Genomic Selection." *Livestock Science* 166: 54–65. <https://doi.org/10.1016/j.livsci.2014.04.029>.
- Legarra, A., A. Ricard, and L. Varona. 2018. "GWAS by GBLUP: Single and Multimarker EMMAX and Bayes Factors, With an Example in Detection of a Major Gene for Horse Gait." *G3: Genes, Genomes, Genetics* 8, no. 7: 2301–2308. <https://doi.org/10.1534/g3.118.200336>.
- Lindholm-Perry, A. K., V. M. Artegoin, J. R. Miles, and A. P. Foote. 2017. "Expression of Cytokine Genes and Receptors in White Blood Cells Associated With Divergent Body Weight Gain in Beef Steers." *Agri Gene* 6: 37–39. <https://doi.org/10.1016/j.aggene.2017.09.005>.
- Lowe, C., T. Yoneda, B. F. Boyce, H. Chen, G. R. Mundy, and P. Soriano. 1993. "Osteopetrosis in Src-Deficient Mice Is due to an Autonomous Defect of Osteoclasts." *Proceedings of the National Academy of Sciences* 90, no. 10: 4485–4489. <https://doi.org/10.1073/pnas.90.10.4485>.
- Lu, X., I. M. Abdalla, M. Nazar, et al. 2021. "Genome-Wide Association Study on Reproduction-Related Body-Shape Traits of Chinese Holstein Cows." *Animals* 11, no. 7: 1927. <https://doi.org/10.3390/ani11071927>.
- Mancini, G., M. Gargani, G. Chillemi, et al. 2014. "Signatures of Selection in Five Italian Cattle Breeds Detected by a 54K SNP Panel." *Molecular Biology Reports* 41: 957–965. <https://doi.org/10.1007/s11033-013-2940-5>.
- Martin, F. J., M. R. Amode, A. Aneja, et al. 2023. "Ensembl 2023." *Nucleic Acids Research* 51, no. D1: D933–D941. <https://doi.org/10.1093/nar/gkac958>.
- Mastranestasis, I., A. Kominakis, A. Hager-Theodorides, L. Ekateriniadou, C. Ligda, and K. Theodorou. 2016. "Associations Between Genetic Polymorphisms and Phenotypic Traits in the Lesvos Dairy Sheep." *Small Ruminant Research* 144: 205–210. <https://doi.org/10.1016/j.smallrumres.2016.09.020>.
- McCoy, A. M., S. K. Beeson, R. K. Splan, et al. 2016. "Identification and Validation of Risk Loci for Osteochondrosis in Standardbreds." *BMC Genomics* 17: 1–11. <https://doi.org/10.1186/s12864-016-2385-z>.
- Messler, S., S. Kropp, V. Episkopou, et al. 2011. "The TGF- β Signaling Modulators TRAP1/TGFBRAPI and VPS39/Vam6/TLP Are Essential for Early Embryonic Development." *Immunobiology* 216, no. 3: 343–350. <https://doi.org/10.1016/j.imbio.2010.07.006>.
- Metzger, J., R. Schrimpf, U. Philipp, and O. Distl. 2013. "Expression Levels of LCORL Are Associated With Body Size in Horses." *PLoS One* 8, no. 2: e56497. <https://doi.org/10.1371/journal.pone.0056497>.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. Lee. 2002. "BLUPF90 and Related Programs (BGF90)." In *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*.
- Naccache, F., J. Metzger, and O. Distl. 2018. "Genetic Risk Factors for Osteochondrosis in Various Horse Breeds." *Equine Veterinary Journal* 50, no. 5: 556–563. <https://doi.org/10.1111/evj.12824>.
- Nazari-Ghadikolaei, A., F. Fikse, Å. Gelinder Viklund, and S. Eriksson. 2023. "Factor Analysis of Evaluated and Linearly Scored Traits in Swedish Warmblood Horses." *Journal of Animal Breeding and Genetics* 140: 366–375. <https://doi.org/10.1111/jbg.12764>.
- Noble, W. S. 2009. "How Does Multiple Testing Correction Work?" *Nature Biotechnology* 27, no. 12: 1135–1137. <https://doi.org/10.1038/nbt1209-1135>.
- Norris, J. M., and S. S. Rich. 2012. "Genetics of Glucose Homeostasis: Implications for Insulin Resistance and Metabolic Syndrome." *Arteriosclerosis, Thrombosis, and Vascular Biology* 32, no. 9: 2091–2096. <https://doi.org/10.1161/ATVBAHA.112.255463>.
- Palmer, N., C. Langefeld, J. Ziegler, et al. 2010. "Candidate Loci for Insulin Sensitivity and Disposition Index From a Genome-Wide Association Analysis of Hispanic Participants in the Insulin Resistance Atherosclerosis (IRAS) Family Study." *Diabetologia* 53: 281–289. <https://doi.org/10.1007/s00125-009-1586-2>.
- Perdry, H., and C. Dandine-Roulland. 2022. "gaston: Genetic Data Handling (QC, GRM, LD, PCA) & Linear Mixed Models." <https://CRAN.R-project.org/package=gaston>.
- Pereira, G., L. Chardulo, J. I. Silva, R. Faria, and R. Curi. 2018. "Genomic Regions Associated With Performance in Racing Line of Quarter Horses." *Livestock Science* 211: 42–51. <https://doi.org/10.1016/j.livsci.2018.02.015>.
- Pryce, J. E., B. J. Hayes, S. Bolormaa, and M. E. Goddard. 2011. "Polymorphic Regions Affecting Human Height Also Control Stature in Cattle." *Genetics* 187, no. 3: 981–984.
- Purcell, S., B. Neale, K. Todd-Brown, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81, no. 3: 559–575. <https://doi.org/10.1086/519795>.
- QTLdb, H. 2023. <https://www.animalgenome.org/cgi-bin/QTLdb/EC/index> (Accessed May 2023).
- R Core Team. 2019. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>.
- R Core Team. 2022. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>.
- Rappaport, N., N. Nativ, G. Stelzer, et al. 2013. "MalaCards: An Integrated Compendium for Diseases and Their Annotation." *Database* 2013: bat018. <https://doi.org/10.1093/database/bat018>.
- Rausch, J. C., J. E. Lavine, N. Chalasani, et al. 2018. "Genetic Variants Associated With Obesity and Insulin Resistance in Hispanic Boys With Nonalcoholic Fatty Liver Disease." *Journal of Pediatric Gastroenterology and Nutrition* 66, no. 5: 789–796. <https://doi.org/10.1097/MPG.0000000000001926>.
- Reich, P., S. Möller, K. F. Stock, et al. 2024. "Genomic Analyses of Withers Height and Linear Conformation Traits in German Warmblood Horses Using Imputed Sequence-Level Genotypes." *Genetics Selection Evolution* 56, no. 1: 45. <https://doi.org/10.1186/s12711-024-00914-6>.
- Revelle, W. 2022. "How to: Use the Psych Package for Factor Analysis and Data Reduction." <https://cran.r-project.org/web/packages/psychTools/vignettes/factor.pdf>.
- Ricard, A., and A. Duluard. 2021. "Genomic Analysis of Gaits and Racing Performance of the French Trotter." *Journal of Animal Breeding and Genetics* 138, no. 2: 204–222. <https://doi.org/10.1111/jbg.12526>.

- Ricard, A., B. Dumont Saint Priest, M. Chassier, M. Sabbagh, and S. Danvy. 2020. "Genetic Consistency Between Gait Analysis by Accelerometry and Evaluation Scores at Breeding Shows for the Selection of Jumping Competition Horses." *PLoS One* 15, no. 12: e0244064. <https://doi.org/10.1371/journal.pone.0244064>.
- Safran, M., N. Rosen, M. Twik, et al. 2021. "The Genecards Suite, Practical Guide to Life Science Databases, 27–56."
- SAS Institute Inc. 2015. *SAS/STAT® 14.1 User's Guide*. Cary, NC: SAS Institute Inc.
- Sayers, E. W., E. E. Bolton, J. R. Brister, et al. 2022. "Database Resources of the National Center for Biotechnology Information." *Nucleic Acids Research* 50, no. D1: D20–D26. <https://doi.org/10.1093/nar/gkab112>.
- Schröder, W., A. Klostermann, K. Stock, and O. Distl. 2012. "A Genome-Wide Association Study for Quantitative Trait Loci of Show-Jumping in Hanoverian Warmblood Horses." *Animal Genetics* 43, no. 4: 392–400. <https://doi.org/10.1111/j.1365-2052.2011.02265.x>.
- Sevane, N., S. Dunner, A. Boado, and J. Cañon. 2017. "Polymorphisms in Ten Candidate Genes Are Associated With Conformational and Locomotive Traits in Spanish Purebred Horses." *Journal of Applied Genetics* 58: 355–361. <https://doi.org/10.1007/s13353-016-0385-y>.
- Signer-Hasler, H., C. Flury, B. Haase, et al. 2012a. "A Genome-Wide Association Study Reveals Loci Influencing Height and Other Conformation Traits in Horses." *PLoS One* 7, no. 5: e37282. <https://doi.org/10.1371/journal.pone.0037282>.
- Signer-Hasler, H., C. Flury, B. Haase, et al. 2012b. "A Genome-Wide Association Study Reveals Loci Influencing Height and Other Conformation Traits in Horses." *PLoS One* 7, no. 5: e37282.
- Sigurðardóttir, H., E. Albertsdóttir, and S. Eriksson. 2017. "Analysis of New Temperament Traits to Better Understand the Trait Spirit Assessed in Breeding Field Tests for Icelandic Horses." *Acta Agriculturae Scandinavica Section A Animal Science* 67, no. 1–2: 46–57. <https://doi.org/10.1080/09064702.2017.1383507>.
- Slavi, N., R. Balasubramanian, M. A. Lee, et al. 2023. "CyclinD2-Mediated Regulation of Neurogenic Output From the Retinal Ciliary Margin Is Perturbed in Albinism." *Neuron* 111, no. 1: 49–64.e45. <https://doi.org/10.1016/j.neuron.2022.10.025>.
- Staiger, E., M. Al Abri, K. M. Pflug, et al. 2016. "Skeletal Variation in Tennessee Walking Horses Maps to the LCORL/NCAPG Gene Region." *Physiological Genomics* 48, no. 5: 325–335. <https://doi.org/10.1152/physiolgenomics.00100.2015>.
- Staiger, E., J. Albright, and S. Brooks. 2016. "Genome-Wide Association Mapping of Heritable Temperament Variation in the Tennessee Walking Horse." *Genes, Brain and Behavior* 15, no. 5: 514–526. <https://doi.org/10.1111/gbb.12290>.
- Steiger, J. H. 1990. "Structural Model Evaluation and Modification: An Interval Estimation Approach." *Multivariate Behavioral Research* 25, no. 2: 173–180. https://doi.org/10.1207/s15327906mbr2502_4.
- Stock, K. F., L. Jönsson, A. Ricard, and T. Mark. 2016. "Genomic Applications in Horse Breeding." *Animal Frontiers* 6, no. 1: 45–52. <https://doi.org/10.2527/af.2016-0007>.
- Swedish Warmblood Association. 2021. "Avelsplan för SWB 2021–2026." swb.org/wp-content/uploads/2016/11/Avelsplan_2021-2026_publ.pdf (Accessed May 2023).
- Tetens, J., P. Widmann, C. Kühn, and G. Thaller. 2013. "A Genome-Wide Association Study Indicates LCORL/NCAPG as a Candidate Locus for Withers Height in German Warmblood Horses." *Animal Genetics* 44, no. 4: 467–471. <https://doi.org/10.1111/age.12031>.
- Thurstone, L. L. 1935. *The Vectors of Mind: Multiple-Factor Analysis for the Isolation of Primary Traits*. Chicago, IL: University of Chicago Press.
- Tucker, L. R., and C. Lewis. 1973. "A Reliability Coefficient for Maximum Likelihood Factor Analysis." *Psychometrika* 38, no. 1: 1–10. <https://doi.org/10.1007/BF02291170>.
- Tucker, R., Y. Hall, T. Hughes, and R. Parker. 2022. "Osteochondral Fragmentation of the Cervical Articular Process Joints; Prevalence in Horses Undergoing CT for Investigation of Cervical Dysfunction." *Equine Veterinary Journal* 54, no. 1: 106–113. <https://doi.org/10.1111/evj.13410>.
- Van Buuren, S., and K. Groothuis-Oudshoorn. 2011. "Mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45: 1–67. <https://doi.org/10.18637/jss.v045.i03>.
- Viklund, Å., and S. Eriksson. 2018. "Genetic Analyses of Linear Profiling Data on 3-Year-Old Swedish Warmblood Horses." *Journal of Animal Breeding and Genetics* 135, no. 1: 62–72. <https://doi.org/10.1111/jbg.12311>.
- Viklund, Å., A. Näsholm, E. Strandberg, and J. Philipsson. 2011. "Genetic Trends for Performance of Swedish Warmblood Horses." *Livestock Science* 141, no. 2–3: 113–122. <https://doi.org/10.1016/j.livsci.2011.05.006>.
- Vosgerau, S., N. Krattenmacher, C. Falker-Gieske, et al. 2022. "Genetic and Genomic Characterization Followed by Single-Step Genomic Evaluation of Withers Height in German Warmblood Horses." *Journal of Applied Genetics* 63: 369–378. <https://doi.org/10.1007/s13353-021-00681-w>.
- Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. Muir. 2012. "Genome-Wide Association Mapping Including Phenotypes From Relatives Without Genotypes." *Genetics Research* 94, no. 2: 73–83. <https://doi.org/10.1017/S0016672312000274>.
- Wobbe, M., H. Alkhoder, K. F. Stock, et al. 2021. "Single-Step Genomic Evaluation in German Riding Horses." In *72th Annual Meeting of the EAAP*, Davos, Switzerland, 172.
- Xia, J., H. Fan, T. Chang, et al. 2017. "Searching for New Loci and Candidate Genes for Economically Important Traits Through Gene-Based Association Analysis of Simmental Cattle." *Scientific Reports* 7, no. 1: 1–9. <https://doi.org/10.1038/srep42048>.
- Yin, L. 2022. "CMplot: Circle Manhattan Plot. R Package Version 4.2.0." <https://CRAN.R-project.org/package=CMplot>.
- Zhang, W., J. Li, Y. Guo, et al. 2016. "Multi-Strategy Genome-Wide Association Studies Identify the DCAF16-NCAPG Region as a Susceptibility Locus for Average Daily Gain in Cattle." *Scientific Reports* 6, no. 1: 38073. <https://doi.org/10.1038/srep38073>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.