









RESEARCH ARTICLE

Open Access



Mendelian randomization analysis of 34,497 German Holstein cows to infer causal associations between milk production and health traits

Helen Schneider^{1*} , Valentin Haas¹ , Ana-Marija Krizanac² , Clemens Falker-Gieske² , Johannes Heise³ , Jens Tetens² , Georg Thaller⁴  and Jörn Bennewitz¹ 

Abstract

Background Claw diseases and mastitis represent the most important health issues in dairy cattle with a frequently mentioned connection to milk production. Although many studies have aimed at investigating this connection in more detail by estimating genetic correlations, they do not provide information about causality. An alternative is to carry out Mendelian randomization (MR) studies using genetic variants to investigate the effect of an exposure on an outcome trait mediated by genetic variants. No study has yet investigated the causal association of milk yield (MY) with health traits in dairy cattle. Hence, we performed a MR analysis of MY and seven health traits using imputed whole-genome sequence data from 34,497 German Holstein cows. We applied a method that uses summary statistics and removes horizontal pleiotropic variants (having an effect on both traits), which improves the power and unbiasedness of MR studies. In addition, genetic correlations between MY and each health trait were estimated to compare them with the estimates of causal effects that we expected.

Results All genetic correlations between MY and each health trait were negative, ranging from -0.303 (mastitis) to -0.019 (digital dermatitis), which indicates a reduced health status as MY increases. The only non-significant correlation was between MY and digital dermatitis. In addition, each causal association was negative, ranging from -0.131 (mastitis) to -0.034 (laminitis), but the number of significant associations was reduced to five nominal and two experiment-wide significant results. The latter were between MY and mastitis and between MY and digital phlegmon. Horizontal pleiotropic variants were identified for mastitis, digital dermatitis and digital phlegmon. They were located within or nearby variants that were previously reported to have a horizontal pleiotropic effect, e.g., on milk production and somatic cell count.

Conclusions Our results confirm the known negative genetic connection between health traits and MY in dairy cattle. In addition, they provide new information about causality, which for example points to the negative energy balance mediating the connection between these traits. This knowledge helps to better understand whether the negative genetic correlation is based on pleiotropy, linkage between causal variants for both trait complexes, or indeed on a causal association.

*Correspondence:

Helen Schneider

helen.schneider@uni-hohenheim.de

Full list of author information is available at the end of the article



© The Author(s) 2024, corrected publication 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

In farm animals, the negative side effects of a high production performance on animal health has raised concerns regarding the ecological footprint of livestock, animal welfare and the farmers' economy [1–3]. For dairy cattle, this implies that an increased milk yield is accompanied by increased incidences of diseases, such as mastitis or claw diseases and fertility problems. With the aim to identify whether this issue is genetically determined or related to management deficiencies, geneticists have repeatedly estimated the genetic correlations between milk production and health traits. Genetic correlations can arise because of several reasons, such as linkage between causal variants for different traits or horizontal pleiotropy [4, 5], i.e., that a genetic variant affects more than one trait. Thus, horizontal pleiotropy has to be separated from vertical pleiotropy, which describes the phenomenon that a genetic variant affects traits on the same causal pathway [6]. In addition, genetic correlations can indicate causality [4]. However, a detailed inference about causal associations from genetic correlations is not possible [7, 8].

The gold standard to assess the causal effect of an exposure (e.g., a treatment) on an outcome variable (e.g., a disease trait) is the randomized control trial (RCT). Here, individuals are randomly assigned (“randomized”) to groups categorized by exposure (e.g., treatment/control), which allows to differentiate between the effect of the exposure itself and the effect of potential confounding factors that can also evoke a causal association between the exposure and the outcome variable [4]. In humans, RCT allowed to confirm that low-density lipoprotein (LDL) cholesterol has an effect on coronary heart disease [9, 10], which was previously suggested in an observational study [11]. In dairy cattle, RCT were performed, e.g., to investigate the economic benefit of different voluntary waiting periods [12], the effect of homeopathy and teat sealers on mastitis [13] or the effects of different preventive methods for ketosis, such as the reduction in milking frequency [14]. Nevertheless, RCT are not feasible to assess causal associations in every setting. For some queries, they might be expensive, impracticable, or even unethical. In addition, it is impossible to test the association of multiple interventions on an outcome variable in a RCT [4, 15]. An alternative is the so-called Mendelian randomization (MR) analysis, where genetic variants [instrumental variables (IV)] are exploited for randomization. The advantage of using genetic variants is that they are present from conception and remain (approximately) unchanged during life. Hence, they are expected to be free from confounding factors [4, 15, 16]. Earlier studies performed MR analyses by using one or only a few

genetic variants as IV [17, 18], which might result in an underpowered MR if the variants have only small effect sizes. However, this can be outperformed by analysing multiple variants jointly in one analysis [15, 19], e.g., by performing a summary-statistics-based MR (SMR, [20]), where summary statistics from genome-wide association studies (GWAS) are used. Recently, this has also been extended to a generalized SMR (GSMR) [15], where, linkage disequilibrium (LD) among the variants is taken into account since LD results in biased MR estimates [19]. Selection of the genetic variants that will be adequate and valid IV for obtaining unbiased MR estimates relies on three assumptions [8]: (1) the relevance assumption, which implies that a variant is strongly associated with the exposure variable, (2) the independence assumption that defines a valid IV as sharing no common cause with the outcome variable, and (3) the exclusion restriction assumption implying that the IV affect the outcome only via the exposure, i.e., not via horizontal pleiotropy [8, 16].

A broad range of MR analyses have been performed in human genetics, where they are facilitated by the public availability of large sample size GWAS summary statistics. These MR were able to identify causal relationships e.g., between body mass index and coronary heart disease [21], vitamin D and mortality [22] or alcohol consumption and heart disease [23]. In contrast, to date only a few MR analyses have been performed in livestock genetics. These have investigated e.g., the causal association between average daily milk yield and resilience indicator traits [24] or between several stature and milk production traits and some functional traits, such as somatic cell count and fertility [25]. So far, no study has examined the causal association of milk production with health traits in dairy cows. Thus, our aim was to fill this gap by investigating the causal association of milk yield (MY) with six claw health traits and mastitis (MAS). Towards this aim, we used a large sample of German Holstein cows with imputed whole-genome sequence (WGS) data, consisting of 17 million sequence variants. The study was split into three parts. First, we estimated heritabilities and genetic correlations based on a 50K single nucleotide polymorphism (SNP) chip data, to compare the observational and causal associations that we expect to identify in our sample. Then, we performed a GWAS for each trait to generate summary statistics. Finally, these summary statistics were used in the last step, which consisted of a set of MR analyses with MY as the exposure variable and each of the health traits as the outcome variable. Here, our aim was to investigate the connection between these traits in more detail. It is thought that this knowledge will reveal whether the negative side effects of high milk production in dairy cattle are due to horizontal pleiotropy,

confounding factors or indeed to a causal effect of milk yield on health traits.

Methods

Material

In this study, we analyzed 34,497 German Holstein cows, which had their first lactation between 2015 and 2020. This dataset is a subset of a larger dataset, provided by the national computing center in Germany (VIT, Vereinigte Informationssysteme Tierhaltung w.V., Verden, Germany), that we filtered as described in Schneider et al. [26]. De-regressed proofs (DRP) with homogeneous accuracies of about 0.75 were available for seven health traits. They are based on on-farm recording by farmers, veterinarians and claw trimmers of disease cases. In addition to mastitis (MAS), the recorded health traits comprised the following claw diseases: claw ulcers (CU), digital dermatitis (DD), interdigital hyperplasia (IH), laminitis (LAM), white line disease (WL) and digital phlegmon (PH). It is important to mention that a lower DRP for a health trait indicates an unfavorable health status. In addition, DRP were available for milk yield (MY). An overview of the number of individuals with DRP for each trait is in Table 1.

Genotypes

For all 34,497 cows, 50K chip and imputed WGS data were available. The 50K chip data were provided by the VIT and the imputation is described in Križanac et al. [27]. After filtering out the SNPs located on the sex chromosomes and those with a minor allele frequency (MAF) lower than 0.01, 44,126 SNPs of the 50K chip remained for the analysis. For the imputed WGS variants, we applied the same filtering criteria but set the MAF threshold to 0.05. Moreover, the dosage R-squared parameter (DR2) that measures the squared correlation between the true and the estimated allele dosage [28] was

Table 1 Overview of the traits and the number of individuals with deregressed proofs for each trait

Trait	Trait abbreviation	Number of individuals	h^2 (SE)
Claw ulcers	CU	27,012	0.153 (0.007)
Digital dermatitis	DD	30,056	0.175 (0.007)
Digital phlegmon	PH	26,437	0.098 (0.006)
Interdigital hyperplasia	IH	30,968	0.153 (0.007)
Laminitis	LAM	30,997	0.176 (0.007)
Mastitis	MAS	33,298	0.133 (0.006)
Milk yield	MY	34,497	0.436 (0.008)
White line disease	WL	30,973	0.139 (0.007)

h^2 : heritability estimates of the traits with their standard errors (SE).

applied to assess the quality of the imputed WGS data. Variants were filtered out if the DR2 was lower than 0.75 [27] and, in total, 16,882,734 imputed WGS variants remained for further analysis.

Statistical analysis

First, we performed a univariate variance component estimation for each trait with the following mixed linear model using the GCTA software version 1.92.3 beta3 [29]:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{g} + \mathbf{e}, \quad (1)$$

where, \mathbf{y} is the vector of the DRP of each animal, $\mathbf{1}$ is a vector of ones, μ denoting the mean, \mathbf{e} is the vector of the residuals and \mathbf{g} is the vector of the polygenic term with \mathbf{Z} as the corresponding incidence matrix that links the polygenic term to the trait records. It was assumed that both \mathbf{g} and \mathbf{e} follow a normal distribution with $\mathbf{g} \sim N(0, \mathbf{G}_{50K}\sigma_g^2)$ and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, where \mathbf{I} is the identity matrix and \mathbf{G}_{50K} the additive genetic relationship matrix (GRM) of the 50K chip. The construction of the GRM involved all 34,497 animals and its computation was performed using GCTA [30].

Then, we applied the following bivariate model to estimate the variance components for seven trait-combinations, which all consisted of MY (\mathbf{y}_1) and one health trait (\mathbf{y}_2):

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{1}\mu_1 \\ \mathbf{1}\mu_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_1 & 0 \\ 0 & \mathbf{Z}_2 \end{bmatrix} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix}, \quad (2)$$

where μ_1 and μ_2 denote the means of \mathbf{y}_1 and \mathbf{y}_2 with \mathbf{Z}_1 and \mathbf{Z}_2 being the corresponding incidence matrices. The variance-covariance-matrix of the respective polygenic (\mathbf{g}_1 and \mathbf{g}_2) and residual (\mathbf{e}_1 and \mathbf{e}_2) terms was modeled as:

$$\text{var} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{50K}\sigma_{g1}^2 & \mathbf{G}_{50K}\sigma_{g12} & 0 & 0 \\ \mathbf{G}_{50K}\sigma_{g12} & \mathbf{G}_{50K}\sigma_{g2}^2 & 0 & 0 \\ 0 & 0 & \mathbf{I}\sigma_{e1}^2 & \mathbf{I}\sigma_{e12} \\ 0 & 0 & \mathbf{I}\sigma_{e12} & \mathbf{I}\sigma_{e2}^2 \end{bmatrix}. \quad (3)$$

We calculated the heritabilities (h^2) and genetic correlations (r_g) by applying standard notations.

Genome-wide association study

The imputed WGS data were used to perform a GWAS for each trait using mixed linear models that are implemented in GCTA [29]. Three models were applied that used different methods to account for population structure. This was done to ensure sufficient power to detect trait-associated signals but also to avoid inflated type I errors. The first method included a polygenic term that

implemented the full GRM across all chromosomes (known as MLMA in GCTA applications). However, the MLMA method suffers from a reduced power due to the double-fitting of the candidate variants, both in the GRM and as fixed effects. Thus, we also applied the so-called “leave-one-chromosome-out” (LOCO) method. Here, the GRM is adjusted such that the chromosome that harbors the candidate variant was excluded from the calculation of the GRM [31]. However, it has been shown that the LOCO method can generate substantial genomic inflation due to the underestimated relationship between individuals because of the missing chromosome in the LOCO-adjusted GRM (e.g., [26, 32, 33]). Thus, in addition, we applied the PC_CHR method that is an extension of the LOCO approach with the aim to account for the reduced informativity of the LOCO-adjusted GRM. GCTA was applied to compute 20 principal components (PC) for each chromosome using the 50K chip data. Then, the PC were included as covariates in the model. We applied the PC on a chromosomal level since, in contrast to the PC on a genome-wide level, they do not include overlapping information with the GRM [33]. All three methods applied the following mixed linear model that is implemented in GCTA:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{g} + \mathbf{e}, \quad (4)$$

where \mathbf{y} is the vector of the DRP that were standardized to $\mathbf{y} \sim N(0, 1)$ to facilitate the interpretation of variant effects. \mathbf{b} is the vector of the mean and the fixed effects that include the effects of the variants and the 20 chromosomal PC for the PC_CHR method. \mathbf{X} denotes the incidence matrix that links \mathbf{b} to the number of allele copies of the variants and of trait records to the covariate effects. \mathbf{g} and \mathbf{e} are the vectors of the polygenic and residual terms, respectively, and followed a normal distribution with $\mathbf{g} \sim N(0, \mathbf{G}_{50K}\sigma_g^2)$ and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, where either the complete or the LOCO-adjusted GRM was applied. Then, we calculated the genomic inflation factor λ_{GC} to assess the inflation of type I errors by using the p-values of the variants to obtain the ratio of the median of the observed chi-square test statistics with one degree of freedom to the corresponding expected median [32].

Generalized summary-data based Mendelian randomization

In the subsequent MR analysis, we investigated the causal relationship between MY and each of the seven health traits. For the health traits, we used the summary statistics generated with the MLMA approach and for MY, we used the summary statistics generated with the PC_CHR method, because these summary statistics resulted in the inflation factor values that were closest to 1. Our decision relates to the narrow ridge that exists between a lack of

power to detect causal associations in our dataset due to a deflated summary statistic, and an increased type 1 error rate in the MR analysis due to a substantial inflation in the summary statistics [26].

The MR analysis was carried out using the GSMR method [15] that is also implemented in GCTA [29]. A detailed description of the method is in Zhu et al. [15], but we will give a brief explanation below. Following the suggestions of Zhu et al. [15], the top-associated ($p_{GWAS,MY} < 5 * 10^{-8}$) and quasi-independent variants after clumping ($r^2 < 0.05$) within a window of 1 Mb were selected as IV. We set the number of IV required to perform a MR analysis to a minimum of ten.

We defined MY as the exposure (x) and each of the seven health traits as the outcome (y) variable. Then, $\hat{b}_{zx(i)}$ denotes the effect of the i -th IV (z) on MY and $\hat{b}_{zy(i)}$ the effect of z on y . The key-point and central output of the GSMR is the MR estimate \hat{b}_{xy} that indicates the causal effect of the exposure on the outcome variable. For the i -th variant, it is calculated as:

$$\hat{b}_{xy(i)} = \frac{\hat{b}_{zy(i)}}{\hat{b}_{zx(i)}}. \quad (5)$$

For all IV with m as the number of IV, $\hat{\mathbf{b}}_{xy}$ is then defined as:

$$\hat{\mathbf{b}}_{xy} = \left\{ \hat{b}_{xy(1)}, \hat{b}_{xy(2)}, \dots, \hat{b}_{xy(m)} \right\}, \quad (6)$$

and follows a normal distribution with $\hat{\mathbf{b}}_{xy} \sim N(\mathbf{b}_{xy}, \mathbf{V})$. Thereby, \mathbf{V} denotes the variance–covariance matrix of $\hat{\mathbf{b}}_{xy}$ that contains the LD correlation between the IV [15]. Including horizontal pleiotropic variants in the analysis would violate one of the initially mentioned IV assumptions, i.e., the exclusion restriction assumption [6]. Thus, the GSMR approach includes the heterogeneity-independent-instruments (HEIDI) method that detects and removes horizontal pleiotropic variants [15]. The HEIDI method is based on the idea that, in the absence of horizontal pleiotropy and if the effect of x on y is causal, \hat{b}_{xy} is only based on the effect of z on y via x and will be identical for any IV. Briefly, the method consists in computing for each IV z_i , the deviation d_i of $\hat{b}_{xy(i)}$ from $\hat{b}_{xy(i)}$ at a target variant ($\hat{b}_{xy(top)}$). A target IV is defined as the IV that shows the strongest association with the exposure variable (MY) in the third quintile of the ordered distribution $-\log_{10}(p\text{-value})$ of \hat{b}_{zx} . This procedure was applied to define the target IV, although the power to detect a horizontal pleiotropic outlier increases, as the association with the exposure is stronger. However, variants with a strong association are also likely to be horizontal pleiotropic variants [15]. Thus, it would not be straightforward

Table 3 Horizontal pleiotropic variants that were identified by the HEIDI method

Chr	Position	Milk yield		Health trait		
		<i>b</i> (SE)	<i>p</i> _{GWAS}	Trait	<i>b</i> (SE)	<i>p</i> _{GWAS}
6	82,842,199	0.055 (0.007)	<0.0001	MAS	−0.044 (0.012)	0.0002
6	85,526,144	−0.110 (0.015)	<0.0001	DD	−0.055 (0.021)	0.0087
6	87,304,351	0.112 (0.008)	<0.0001	MAS	−0.060 (0.011)	<0.0001
6	87,719,366	0.093 (0.013)	<0.0001	DD	0.047 (0.020)	0.0177
14	6,054,352	0.110 (0.011)	<0.0001	PH	0.038 (0.016)	0.0032
19	17,893,592	0.078 (0.013)	<0.0001	MAS	0.052 (0.018)	0.0200

Shown are the chromosome (Chr) and position (in bp) of the variant, together with the effect size (*b*), standard error (SE) and p-value (*p*_{GWAS}) from the association study for milk yield and the respective health trait

to use the IV that shows the strongest association with the exposure variable as target variant in order to detect horizontal pleiotropic outliers. Then, \hat{d}_i at the *i*-th IV is defined as $\hat{d}_i = \hat{b}_{xy(i)} - \hat{b}_{xy(top)}$ with:

$$\text{var}(\hat{d}_i) = \text{var}(\hat{b}_{xy(i)} - \hat{b}_{xy(top)}), \quad (7)$$

where LD among the IV is again considered, which implies that the HEIDI method filters for the LD that was not already removed by the clumping step [15]. In order to obtain the statistical significance of the MR estimate (\hat{b}_{xy}) and the deviation of $\hat{b}_{xy(i)}$ from $\hat{b}_{xy(top)}$, the test statistics T_{GSMR} and T_{HEIDI} were applied with:

$$T_{GSMR} = \frac{\hat{b}_{xy}^2}{\text{var}(\hat{b}_{xy})}, \quad (8)$$

and

$$T_{HEIDI} = \frac{\hat{d}_i^2}{\text{var}(\hat{d}_i)}. \quad (9)$$

They followed a Chi-square distribution with one degree of freedom. We chose two thresholds to declare

a causal effect \hat{b}_{xy} as significant, i.e., a Bonferroni corrected threshold of 0.05, corrected for seven tests, (this corresponds to an experiment-wide significance) and one, without correction, of 0.05 (nominal significance). Horizontal pleiotropic outliers were defined as IV that significantly ($p_{HEIDI} = 0.01$) deviate from homogeneity [15]. In this study, we applied two MR analyses, one that included the HEIDI method (HEIDI) and one that did not (noHEIDI). This was done to assess the contribution of horizontal pleiotropic variants to the putative causal association between the traits.

Results

Estimation of variance components

The heritability estimates of the DRP were moderate to low for the health traits and high ($h^2 = 0.439$) for MY. Concerning the health traits, LAM had the highest heritability ($h^2 = 0.176$) and PH the lowest ($h^2 = 0.098$). All health traits were negatively correlated with MY. We found a negative high correlation between MY and MAS ($r_g = -0.303$). The weakest correlation was between MY and DD ($r_g = -0.019$) (Table 2). A genetic correlation was declared to be significantly different from zero if the

Table 2 Genetic correlations and causal associations between milk yield and the respective health trait

Health trait	<i>r</i> _g (SE)	HEIDI			noHEIDI		
		\hat{b}_{xy} (SE)	<i>p</i> _{GSMR}	N	\hat{b}_{xy} (SE)	<i>p</i> _{GSMR}	N
Claw ulcers	−0.069 (0.028)	−0.041 (0.021)	0.0494	105	−0.041 (0.021)	0.0494	105
Digital dermatitis	−0.019 (0.026)	−0.054 (0.021)	0.0094	103	−0.048 (0.021)	0.0200	105
Digital phlegmone	−0.182 (0.032)	−0.098 (0.020)	<0.0001	104	−0.098 (0.020)	<0.0001	105
Interdigital hyperplasia	−0.106 (0.027)	−0.039 (0.020)	0.0577	105	−0.039 (0.020)	0.0577	105
Laminitis	−0.082 (0.040)	−0.034 (0.021)	0.0577	105	−0.034 (0.021)	0.0577	105
Mastitis	−0.303 (0.026)	−0.131 (0.020)	<0.0001	102	−0.133 (0.020)	<0.0001	105
White line disease	−0.088 (0.035)	−0.042 (0.020)	0.0381	105	−0.042 (0.020)	0.0381	105

Shown are the genetic correlations (*r*_g) and effect sizes (\hat{b}_{xy}) from the GSMR analysis together with their standard errors (SE) as well as the p-values (*p*_{GSMR}) and the number of instrumental variables (N) for the GSMR analysis. The results of the GSMR analysis are shown for the analysis with (HEIDI) and without (noHEIDI) applying the HEIDI method

absolute value of the estimate was at least twice the corresponding standard error. Following this definition, only the trait-combination consisting of DD and MY did not reach significance ($r_g = -0.019$, $SE = 0.026$).

Genome-wide association study

For the health traits, the genomic inflation factors were equal to 1.010 (MAS), 1.012 (IH), 1.015 (CU), 1.016 (DD), 1.031 (WL), 1.032 (PH) and 1.039 (LAM) using the MLMA method. These values were close to 1, which is why no other GWAS method was applied. In contrast, using the MLMA method, we found a strong deflation ($\lambda_{GC} = 0.900$) for MY that became a strong inflation ($\lambda_{GC} = 15.157$) when the LOCO method was applied. Thus, the summary statistics generated with the PC_CHR approach ($\lambda_{GC} = 1.569$) was subsequently used in the GSMR analysis.

Generalized summary-data based Mendelian randomization

After filtering for variants in high LD, 105 variants remained as IV for the GSMR analysis of MY and the seven health traits. In addition, three variants were removed by the HEIDI filtering step in the analysis of MY and MAS, two in the analysis of MY and DD and

one in the analysis of MY and PH. Four of these variants were located on *Bos taurus* (BTA) autosome 6, one on BTA14 and one on BTA19 (Table 3). The effects of almost all the horizontal pleiotropic variants were in the same direction for MY and the health trait, with only the two variants on BTA6 that affect MY and MAS showing an antagonistic effect.

Regarding the causal associations, the causal effects \hat{b}_{xy} of MY on any health trait were negative, which indicates a decrease in the DRP for the traits when the DRP for MY increases. In the HEIDI analysis, \hat{b}_{xy} ranged from -0.133 (MAS) to -0.034 (LAM) and in the noHEIDI analysis from -0.131 (MY) to -0.034 (LAM) (Table 2 and Figs. 1, 2, 3 and 4).

In both analyses, we found a nominal significant causal effect of MY on each health trait except for LAM and IH and an experiment-wide significant causal effect of MY on PH and MAS. The causal effect of MY on DD almost reached experiment-wide significance ($p_{GSMR} = 0.0094$) in the HEIDI analysis. Here, \hat{b}_{xy} was -0.054 and slightly lower ($\hat{b}_{xy} = -0.048$) when the horizontal pleiotropic variants were not removed (Table 2; Fig. 2). For MAS and PH, there was no obvious difference between the HEIDI and the noHEIDI analysis (Table 2 and Figs. 3 and 4).

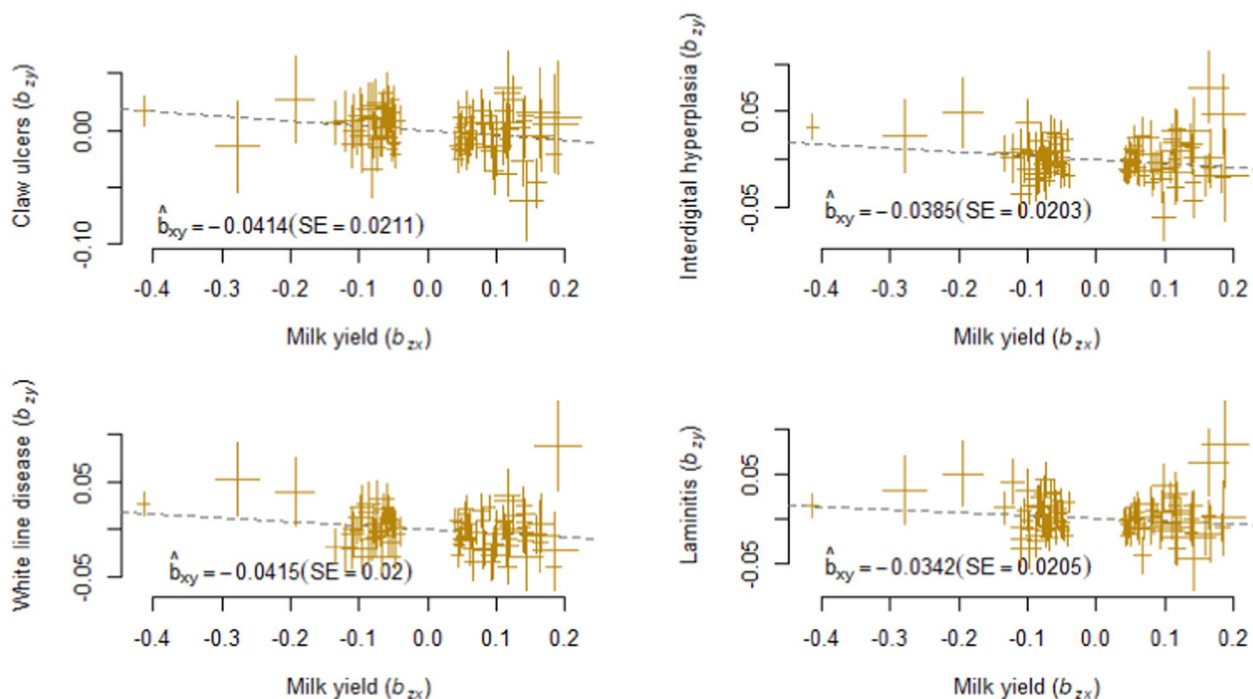


Fig. 1 GSMR analysis of the respective health trait and milk yield. Shown are the plots of the effect sizes. For these trait combinations, no difference was observed between the analyses that did and the one that did not exclude horizontal pleiotropic variants

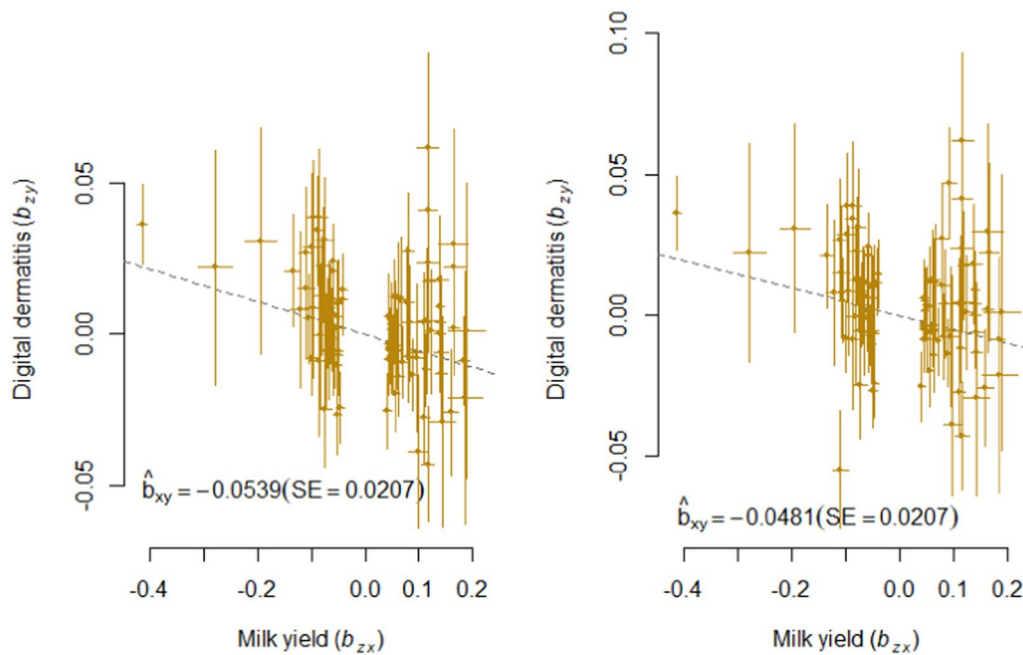


Fig. 2 GSMR analysis of digital dermatitis and milk yield. Shown are the plots of the effect sizes for the analysis with (left) and without (right) filtering for horizontal pleiotropic variants

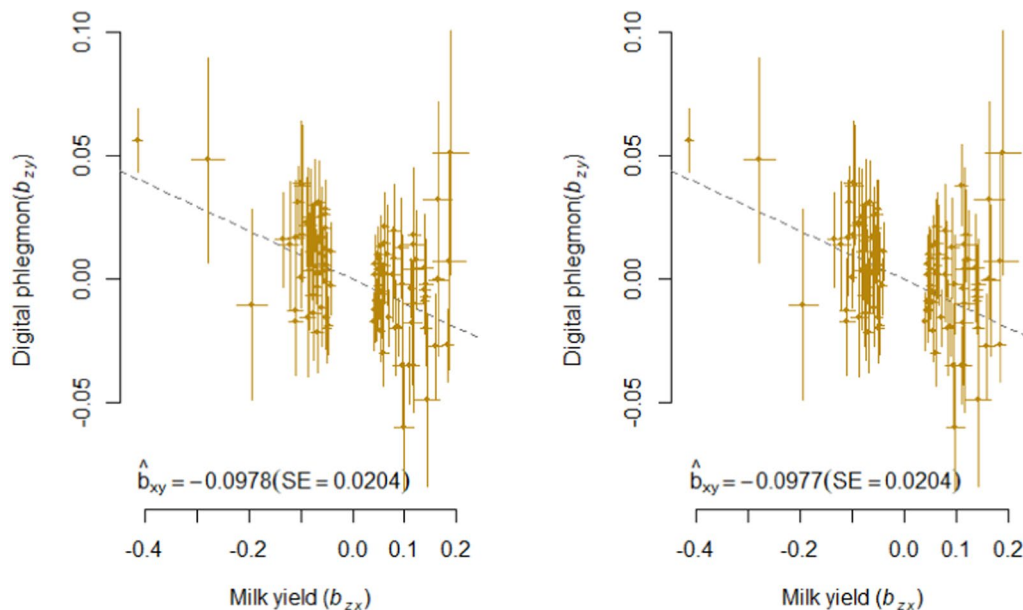


Fig. 3 GSMR analysis of digital phlegmon and milk yield. Shown are the plots of the effect sizes for the analysis with (left) and without (right) filtering for horizontal pleiotropic variants

Discussion

Currently, knowledge about the relationship between milk production and health traits in dairy cattle is mostly acquired by estimating genetic correlations [2, 3]. However, there is a substantial difference between correlation

and causation. Thus, several human studies have performed MR studies to reveal the causal effect e.g., of risk factors on diseases based on genetic variants. To date, no MR study has investigated the causal effect of MY on health traits in dairy cattle. Thus, we performed

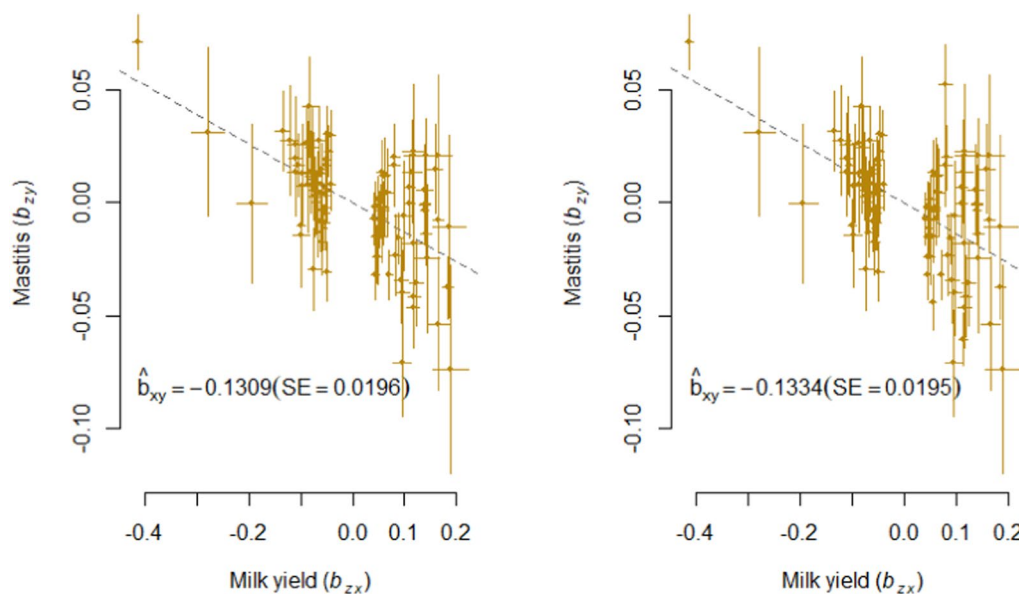


Fig. 4 GSMR analysis of mastitis and milk yield. Shown are the plots of the effect sizes for the analysis with (left) and without (right) filtering for horizontal pleiotropic variants

a MR analysis with MY as the exposure variable and the health traits as the outcome variables. To do so, we applied a method that uses summary statistics and that removes horizontal pleiotropic variants [15]. In addition, we estimated genetic correlations using the 50K chip data to obtain observational estimates of the connections between traits and to compare them with the estimates of causal effects from the MR analysis. Of seven trait-combinations, we found six significant genetic correlations as well as five nominal and two experiment-wide significant causal associations. The latter were between the traits with the highest genetic correlation. All estimates had a negative sign. Moreover, we identified six pleiotropic variants but found no substantial impact of these variants on the estimates of causal effects.

Biology

All the genetic correlations, except for that between MY and DD, were significant. On the opposite, only two experiment-wide significant causal associations were found, which were between MY and PH and MY and MAS. Interestingly, these traits were also those with the strongest genetic correlation with MY (Table 2). Regarding the causal effect of MY on the health traits, it is possible that the connection between the traits is mediated by another trait via vertical pleiotropy. Here, it is important to mention that, in contrast to horizontal pleiotropy, vertical pleiotropy does not result in instrumental invalidity. This is because in the case of vertical pleiotropy, the mediating variable is on the same causal pathway as

the exposure and the outcome variables [16]. A putative mediating variable is the negative energy balance (NEB), which is a physiological imbalance that occurs within two to three months after calving [34, 35]. In this period, the immune system of the animals is suppressed, the level of immunocytes is decreased and the level of inflammatory cytokines and reactive oxygen species that damage cellular tissue is increased [36]. This shift in the immune system can make individuals susceptible to infectious diseases, such as MAS or infectious claw diseases. Van der Spek et al. [37] defined DD, PH and IH as infectious and LAM, WL, and CU as non-infectious claw diseases. Indeed, the incidences of DD and PH were found to be highest in primiparous cows during the first months after calving [38]. Thus, we assume that the experiment-wide significant causal associations between MY and MAS and between MY and PH and the almost experiment-wide significant one with DD in this study might be due to the NEB as mediating variable. In addition, a nutritional imbalance, such as the NEB, is characterized by an insufficient intake of minerals and vitamins. This results in the weak formation of the claw horn, which makes the individual susceptible to infectious diseases [38]. Although IH also belongs to the group of infectious diseases, we did not find a significant causal effect of MY for this claw disease, neither experiment-wide nor nominal. A possible explanation would be that a genetic predisposition is causing IH rather than a causal effect of MY. This is based on the results of previous studies, which found chromosomal regions [39] and a missense mutation on BTA8

[40] that are associated with IH. Another potentially mediating variable is the gut microbiome. An effect of the gut microbiome on the immune system [41], udder [42, 43] and claw health [44] has already been mentioned by several studies. In addition, Monteiro et al. [45] reported an association between the level of milk production and the microbiome of the rumen as well as of the lower gut. Thus, it is possible that an elevated level of milk production leads to a change in the gut microbiome, which in turn could result in a deteriorated health status of the cows. Finally, a third variable that might mediate the effect of MY on MAS is the milking speed (MS). Several studies mentioned a positive genetic correlation between MS and MY [46–48] but also an unfavorable genetic correlation between udder health and MS [48–51]. Combining these findings, we suggest that a higher MY evokes an increased MS and thereby also an increased risk to develop mastitis.

We did not observe a large difference between the HEIDI and the noHEIDI analysis. This was surprising as horizontal pleiotropy has been reported to be common in humans and livestock [52, 53]. An explanation for our findings might be that only top associated variants were used as IV, which might not be the horizontal pleiotropic variants in this case. However, six horizontal pleiotropic variants were detected and removed by the HEIDI method, which were mostly (four out of six variants) on BTA6 (Table 3). Two of them are either located within (BTA6:87304531, MY-MAS) or nearby (BTA6:87719361, MY-DD) genes that have previously been identified as quantitative trait loci (QTL) for longevity [54, 55]. In addition, the variant on BTA6 at 87,304,531 bp is located within the *NPF2L4* gene, which is a QTL for mastitis resistance [56], somatic cell score, milk, and protein yield [57]. This variant is also located nearby the *GC* gene, which has a horizontal pleiotropic effect on milk production and mastitis [58]. A study on resilience indicator traits, conducted by Chen et al. [24], revealed two regions on BTA6 showing a noteworthy association with some resilience indicator traits. These regions are located close to the horizontal pleiotropic variants on BTA6. Moreover, the variant on BTA19 at 17,893,592 bp that affects MY and MAS, in our study, is located nearby the *RHBDL3* gene, which was mentioned to have an influence on a wide range of human diseases like cancer or inflammatory diseases [59]. These results from previous studies support our findings that the horizontal pleiotropic variants, which were identified by the HEIDI method, have an effect on MY and several health traits.

Methodology

Another popular method to assess causality are structural equation models (SEM) [60] that enable the identification

of causal relationships among traits and allow for the prediction of one trait by another “upstream” causal variable [61]. Recently, they have been extended to GWAS-SEM, where the total genetic effects for one trait are separated into direct and indirect genetic effects, i.e., genetic effects that are mediated by other upstream variables [62–64]. Thus, SEM provide detailed information about causal associations among traits. However, they are sensitive to erroneous model specifications, computationally very complex and require a temporal sequence of traits, i.e., that one trait has to occur prior to the other one [19]. Thus, we did not apply a SEM in this study.

As for all statistical analyses, a MR also relies on untestable assumptions and thus, the inference that one can draw depends strongly on the plausibility of the assumptions [6]. In the case of a MR, caution has to be paid to the choice of the variants as IV, according to the three initially mentioned assumptions. In order to meet the relevance assumption that relates to the association between a variant and the exposure trait, we used only strongly associated variants ($p_{GWAS,MY} < 5 * 10^{-8}$) as IV for our analyses. It has been shown that using these variants ensures the absence of inflation in the test statistics under the H_0 that $b_{xy} = 0$ [15]. In addition, applying only top-associated variants also helps to avoid weak instrument bias, i.e., that IV have a weak effect on the exposure variable, which would result in low power of the MR analysis. Moreover, a weak instrument bias is also attenuated by applying multiple variants instead of only one, since they are more likely to explain a larger proportion of the genetic variance [16]. Because of this, we applied summary statistics for our analysis. Nevertheless, quantitative traits are typically polygenic in nature, i.e., they are influenced by many variants that all have small effect sizes. Thus, applying only the top-associated variants as IV neglects this polygenicity, which might hamper the detection of causal exposure outcome associations [65]. Indeed, it is possible that the absence of causal associations among all the traits in our analysis is caused by the stringent threshold that we set for the definition of IV. However, if a variant has a larger effect on the outcome variable and a smaller effect on the exposure variable, we assume that this variant is less likely suitable for indicating causality. Instead, this variant is more likely having a horizontal pleiotropic effect on both the exposure and the outcome variables. Applying horizontal pleiotropic variants as IV would violate the third IV assumption, i.e., the exclusion restriction assumption [16]. Hence, we decided to apply only the top-associated variants as IV. Moreover, the power of a MR analysis depends strongly on the power of the underlying GWAS, the summary statistics of which are used in the MR. Thus, our aim was to apply a large sample size to our GWAS which results

in greater power to detect significantly associated variants. It has also been shown that the power of a GWAS benefits from the application of sequence rather than SNP chip data [66]. On the downside, an increased sample size can also evoke inflated false positive associations [67–69]. A metric to identify genomic inflation is the genomic inflation factor, where a value below 1 points to a deflation and a value above 1 points to an inflation in the GWAS signals. Acceptable values are below 1.5 [32]. In order to avoid both, a weak instrument bias and an inflation in the MR estimates caused by an inflated summary statistic, we decided to proceed with the summary statistics for each trait that had a genomic inflation factor closest to 1. Hence, we applied the summary statistics from the MLMA method for the health traits, whereas this method was deflated for MY. By applying the LOCO method, we found a strong inflation for MY, which made it necessary to apply the PC_CHR method to generate a summary statistic with a suitable genomic inflation. Including PC as covariates in addition to the GRM in GWAS has already been implemented in other studies on livestock to mitigate the effect of population stratification as a strong confounding bias [26, 32, 33].

An important aspect is the interpretation of the MR results in our study. MR estimates are meant to deliver credible causal associations between exposure and outcome variables [15] and relate to genetically-induced effects of the exposure on the outcome variable [6]. Genetic variants are fixed at conception. Hence, they indicate long-term effects of elevated levels of the exposure variable and cannot be understood as alterations in the outcome that one can expect by an intervention on the exposure variable at a specific time in life [6, 8, 16]. This hampers a quantitative inference from levels of e.g., LDL or vitamin D that are typically measured at a certain time in life. However, we used DRP as phenotypes in our study, which do not relate to a specific time. Thus, the causal effect of MY on MAS $\hat{b}_{xy} = -0.1309$ indicates that an increase by one standard deviation (SD) of the DRP for MY results in a decrease by 0.1309 SD of the DRP for MAS. Thus, it is important to keep in mind that MR relies on untestable assumptions and the credibility of the results depends strongly on the plausibility of these assumptions. As already mentioned, our aim was to pay attention to fulfilling these assumptions, e.g., by applying strongly associated genetic variants as IV or by removing horizontal pleiotropic variants. However, we believe that the inference one draws from \hat{b}_{xy} should be understood rather qualitatively than quantitatively, also because the DRP themselves are obtained from estimations that rely on statistical assumptions.

Several methods to perform a MR analysis have been developed and are reviewed for example in Davies et al.

[16], Slob and Burgess [70], and Burgess et al. [6]. They differ e.g., in the assumptions they make about instrumental invalidity and the way they account for horizontal pleiotropy. It has been mentioned that causal effects can be verified by applying various methods [6, 70]. However, we decided not to apply any other method than the GSMR because (1) it accounts for LD among variants, (2) it has been found to be not or only little affected by undetected horizontal pleiotropy, and (3) it outperforms other methods regarding the power to detect causal effects. The latter feature is based on the fact that GSMR takes the sampling variance in \hat{b}_{zy} and \hat{b}_{zx} into account as opposed to other methods [15]. Hence, we aimed at verifying the causal effects that were found in this study by means of a reverse MR analysis using the health traits as the exposure and MY as the outcome variable. Thus, the causal effects of MY on the health traits would have been verified if the corresponding reverse causal effects were negligible. Unfortunately, it was not possible to perform a reverse MR analysis because too few, i.e., less than ten, IV remained after filtering for LD among the variants. Thus, further studies are required to validate the causal association of MY especially with MAS and PH in other datasets.

Finally, a limitation of our study is that it is a one-sample MR, which means that the summary statistics for the outcome and exposure variables were generated using the same sample. This makes environmental confounding more likely to occur. Environmental confounding implies that genetic causal associations between the exposure and outcome variables are correlated in the direction of a confounding association between these two, which induces an increased type I error rate of the MR estimates [6]. However, whereas human geneticists can benefit from a huge range of publicly available summary statistics, this is not the case in animal genetics. Hence, we decided to proceed with the summary statistics that we generated ourselves and aimed at minimizing confounding by applying DRP in the GWAS. The DRP originated from the national breeding value estimation in Germany, where environmental factors are corrected for by using fixed effects as well as by applying a test-day random-regression model for MY [71]. Another limitation is that a one-sample setting might suffer from the effect of the winner's curse, i.e., that genetic associations tend to be overestimated in the dataset in which they were first discovered. This again might induce a weak instrument bias if the true association is much lower than the over-estimated one [6]. In general, these limiting factors can cause an inflation in MR estimates. However, we observed that only two of seven health traits showed an experiment-wide significant causal association with

milk yield and believe that this confirms the credibility of our results.

Conclusions

Exploiting the large sample size and sequence data to achieve sufficient statistical power in this study enabled us to more deeply investigate the genetic connection between milk yield and health traits in German Holstein. Regarding the generation of the summary statistics, our results confirm the need to include PC in the GWAS for milk yield to avoid substantial genomic inflation caused by population stratification. Moreover, an experiment-wide significant negative causal association was found between MY and digital phlegmon as well as between MY and mastitis and a nominal significant negative causal association between milk yield and almost each health trait except for interdigital hyperplasia and laminitis. This indicates the long-term negative impact of a genetically driven high milk yield on animal health. However, this causal association still requires to be validated using other samples. Not only the causal associations but also the genetic correlations were all negative. We assume that the causal association between MY and these health traits might be due to another mediating variable, such as the NEB. In the MR analysis, horizontal pleiotropic variants that affected both, MY and the health traits, were identified and removed using a filter step that is implemented in the GSMR method, which we used for our analysis. Interestingly, the difference between estimates of causal effects with and without horizontal pleiotropic variants was negligible. Our findings support the potential of investigating causal associations between economically important traits in livestock by applying a MR analysis in order to unravel these relationships between traits in more detail, and to better differentiate between their causal associations, the effect of horizontal pleiotropic variants affecting both traits and linkage between QTL.

Acknowledgements

Not applicable.

Author contributions

HS performed the statistical analyses. JH provided the dataset. AMK and CFG performed the imputation. HS and JB wrote the paper. VH contributed to the writing. JB supervised and JB, JT and GT initiated the study. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This study is part of the project QTCC: *From Quantitative Trait Correlation to Causation in dairy cattle* and was supported by the Deutsche Forschungsgemeinschaft (DFG, BE3703/15-1, TE622/6-1).

Availability of data and materials

Restrictions apply to the availability of the genotype and phenotype data analyzed in this study since they are property of the German Holstein breeding organisations organized within the Bundesverband Rind und Schwein e.V.

We thank all participating organisations. Thus, the data have commercial value and are not publicly available. The corresponding author can be contacted for a reasonable request.

Declarations

Ethics approval and consent to participate

For this study, no biological samples were collected and no animals were handled. Consent for use of the data was obtained where required prior to conducting this study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Institute of Animal Science, University of Hohenheim, 70599 Stuttgart, Germany. ²Department of Animal Sciences, University of Göttingen, 37077 Göttingen, Germany. ³Vereinigte Informationssysteme Tierhaltung w.V. (VIT), 27283 Verden, Germany. ⁴Institute of Animal Breeding and Husbandry, Christian-Albrechts University of Kiel, 24098 Kiel, Germany.

Received: 26 November 2023 Accepted: 22 March 2024

Published: 8 April 2024

References

- Rauw WM, Kanis E, Noordhuizen-Stassen EN, Grommers FJ. Undesirable side effects of selection for high production efficiency in farm animals: a review. *Livest Prod Sci.* 1998;56:15–33.
- Martin P, Barkema HW, Brito LF, Narayana SG, Miglior F. Symposium review: novel strategies to genetically improve mastitis resistance in dairy cattle. *J Dairy Sci.* 2018;101:2724–36.
- Heringstad B, Egger-Danner C, Charfeddine N, Pryce JE, Stock KF, Kofler J, et al. Invited review: genetics and claw health: opportunities to enhance claw health by genetic selection. *J Dairy Sci.* 2018;101:4801–21.
- Pickrell JK. Fulfilling the promise of Mendelian randomization. *bioRxiv.* 2015. <https://doi.org/10.1101/018150v1>.
- Cai Z, Dusza M, Guldbbrandtsen B, Lund MS, Sahana G. Distinguishing pleiotropy from linked QTL between milk production traits and mastitis resistance in Nordic Holstein cattle. *Genet Sel Evol.* 2020;52:19.
- Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations: update for summer 2023. *Wellcome Open Res.* 2023;4:186.
- Gratten J, Visscher PM. Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome Med.* 2016;8:78.
- Walker VM, Zheng J, Gaunt TR, Smith GD. Phenotypic causal inference using genome-wide association study data: Mendelian randomization and beyond. *Annu Rev Biomed Data Sci.* 2022;5:1–17.
- The lipid research clinics coronary primary prevention trial results. I. Reduction in incidence of coronary heart disease. *JAMA.* 1984;251:351–64.
- Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet.* 2005;366:1267–78.
- Kannel WB, Castelli WP, Gordon T, McNamara PM. Serum cholesterol, lipoproteins, and the risk of coronary heart disease. The Framingham study. *Ann Intern Med.* 1971;74:1–12.
- Burgers EEA, Kok A, Gosalink RMA, Hogeveen H, Kemp B, van Knegsel ATM. Revenues and costs of dairy cows with different voluntary waiting periods based on data of a randomized control trial. *J Dairy Sci.* 2022;105:4171–88.
- Klocke P, Ivmeyer S, Butler G, Maeschli A, Heil F. A randomized controlled trial to compare the use of homeopathy and internal Teat Sealers for the

- prevention of mastitis in organically farmed dairy cows during the dry period and 100 days post-calving. *Homeopathy*. 2010;99:90–8.
14. Williamson M, Couto Serrenho R, McBride BW, LeBlanc SJ, DeVries TJ, Duffield TF. Reducing milking frequency from twice to once daily as an adjunct treatment for ketosis in lactating dairy cows—a randomized controlled trial. *J Dairy Sci*. 2022;105:1402–17.
 15. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun*. 2018;9:224.
 16. Davies NM, Holmes MV, Davey SG. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*. 2018;362: k601.
 17. Palmer TM, Nordestgaard BG, Benn M, Tybjaerg-Hansen A, Davey Smith G, Lawlor DA, et al. Association of plasma uric acid with ischaemic heart disease and blood pressure: mendelian randomisation analysis of two large cohorts. *BMJ*. 2013;347: f4262.
 18. C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC), Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, et al. Association between C reactive protein and coronary heart disease: Mendelian randomisation analysis based on individual participant data. *BMJ*. 2011;342:548.
 19. Ziegler A, Mwambi H, König IR. Mendelian randomization versus path models: making causal inferences in genetic epidemiology. *Hum Hered*. 2015;79:194–204.
 20. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016;48:481–7.
 21. Afzal S, Brøndum-Jacobsen P, Bojesen SE, Nordestgaard BG. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. *BMJ*. 2014;349: g6330.
 22. Dale CE, Fatemifar G, Palmer TM, White J, Prieto-Merino D, Zabaneh D, et al. Causal associations of adiposity and body fat distribution with coronary heart disease, stroke subtypes, and type 2 diabetes mellitus: a Mendelian randomization analysis. *Circulation*. 2017;135:2373–88.
 23. Holmes MV, Dale CE, Zuccolo L, Silverwood RJ, Guo Y, Ye Z, et al. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. *BMJ*. 2014;349: g4164.
 24. Chen S-Y, Gloria LS, Pedrosa VB, Doucette J, Boerman JP, Brito LF. Unraveling the genomic background of resilience based on variability in milk yield and milk production levels in North American Holstein cattle through GWAS and Mendelian randomization analyses. *J Dairy Sci*. 2024;107:1035–53.
 25. Xiang R, van den Berg I, MacLeod IM, Daetwyler HD, Goddard ME. Effect direction meta-analysis of GWAS identifies extreme, prevalent and shared pleiotropy in a large mammal. *Commun Biol*. 2020;3:88.
 26. Schneider H, Segelke D, Tetens J, Thaller G, Bennewitz J. A genomic assessment of the correlation between milk production traits and claw and udder health traits in Holstein dairy cattle. *J Dairy Sci*. 2023;106:1190–205.
 27. Križanac A-M, Reimer C, Heise J, Liu Z, Pryce J, Bennewitz J, et al. Sequence-based GWAS in 180 000 German Holstein cattle reveals new candidate genes for milk production traits. *bioRxiv*. 2023. <https://doi.org/10.1101/2023.12.06.570350>.
 28. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet*. 2009;84:210–23.
 29. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76–82.
 30. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 2010;42:565–9.
 31. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet*. 2014;46:100–6.
 32. van den Berg S, Vandenplas J, van Eeuwijk FA, Lopes MS, Veerkamp RF. Significance testing and genomic inflation factor using high-density genotypes or whole-genome sequence data. *J Anim Breed Genet*. 2019;136:418–29.
 33. Yin T, König S. Genome-wide associations and detection of potential candidate genes for direct genetic and maternal genetic effects influencing dairy cattle body weight at different ages. *Genet Sel Evol*. 2019;51:4.
 34. Buttchereit N, Stamer E, Junge W, Thaller G. Evaluation of five lactation curve models fitted for fat:protein ratio of milk and daily energy balance. *J Dairy Sci*. 2010;93:1702–12.
 35. Krattenmacher N, Thaller G, Tetens J. Analysis of the genetic architecture of energy balance and its major determinants dry matter intake and energy-corrected milk yield in primiparous Holstein cows. *J Dairy Sci*. 2019;102:3241–53.
 36. Esposito G, Irons PC, Webb EC, Chapwanya A. Interactions between negative energy balance, metabolic diseases, uterine health and immune response in transition dairy cows. *Anim Reprod Sci*. 2014;144:60–71.
 37. van der Spek D, van Arendonk JAM, Vallée AAA, Bovenhuis H. Genetic parameters for claw disorders and the effect of preselecting cows for trimming. *J Dairy Sci*. 2013;96:6070–8.
 38. Garvey M. Lameness in dairy cow herds: disease aetiology, prevention and management. *Dairy*. 2022;3:199–210.
 39. Croué I, Michenet A, Leclerc H, Ducrocq V. Genomic analysis of claw lesions in Holstein cows: opportunities for genomic selection, quantitative trait locus detection, and gene identification. *J Dairy Sci*. 2019;102:6306–18.
 40. Zhang X, Swalve HH, Pijl R, Rosner F, Wensch-Dorendorf M, Brenig B. Interdigital hyperplasia in Holstein cattle is associated with a missense mutation in the signal peptide region of the tyrosine-protein kinase transmembrane receptor gene. *Front Genet*. 2019;10:1157.
 41. Liu L, Wu P, Guo A, Yang Y, Chen F, Zhang Q. Research progress on the regulation of production traits by gastrointestinal microbiota in dairy cows. *Front Vet Sci*. 2023;10:1206346.
 42. Ma C, Sun Z, Zeng B, Huang S, Zhao J, Zhang Y, et al. Cow-to-mouse fecal transplantations suggest intestinal microbiome as one cause of mastitis. *Microbiome*. 2018;6:200.
 43. Hu X, Li S, Mu R, Guo J, Zhao C, Cao Y, et al. The rumen microbiota contributes to the development of mastitis in dairy cows. *Microbiol Spectr*. 2022;10: e0251221.
 44. Bay V, Gillespie A, Ganda E, Evans NJ, Carter SD, Lenzi L, et al. The bovine foot skin microbiota is associated with host genotype and the development of infectious digital dermatitis lesions. *Microbiome*. 2023;11:4.
 45. Monteiro HF, Zhou Z, Gomes MS, Peixoto PMG, Bonsaglia ECR, Canisso IF, et al. Rumen and lower gut microbiomes relationship with feed efficiency and production traits throughout the lactation of Holstein dairy cows. *Sci Rep*. 2022;12:4904.
 46. Samoré AB, Román-Ponce SI, Vacirca F, Frigo E, Canavesi F, Bagnato A, et al. Bimodality and the genetics of milk flow traits in the Italian Holstein-Friesian breed. *J Dairy Sci*. 2011;94:4081–9.
 47. Gray KA, Vacirca F, Bagnato A, Samoré AB, Rossoni A, Maltecca C. Genetic evaluations for measures of the milk-flow curve in the Italian Brown Swiss population. *J Dairy Sci*. 2011;94:960–70.
 48. Berry DP, Coyne J, Coughlan B, Burke M, McCarthy J, Enright B, et al. Genetics of milking characteristics in dairy cows. *Animal*. 2013;7:1750–8.
 49. Boettcher PJ, Dekkers JC, Kolstad BW. Development of an udder health index for sire selection based on somatic cell score, udder conformation, and milking speed. *J Dairy Sci*. 1998;81:1157–68.
 50. Rupp R, Boichard D. Genetic parameters for clinical mastitis, somatic cell score, production, udder type traits, and milking ease in first lactation Holsteins. *J Dairy Sci*. 1999;82:2198–204.
 51. Zwald NR, Weigel KA, Chang YM, Welper RD, Clay JS. Genetic evaluation of dairy sires for milking duration using electronically recorded milking times of their daughters. *J Dairy Sci*. 2005;88:1192–8.
 52. Visscher PM, Yang J. A plethora of pleiotropy across complex traits. *Nat Genet*. 2016;48:707–8.
 53. Xiang R, MacLeod IM, Daetwyler HD, Jong de G, O'Connor E, Schrooten C, et al. Genome-wide fine-mapping identifies pleiotropic and functional variants that predict many traits across global cattle populations. *Nat Commun*. 2021;12:860.
 54. Zhang Q, Guldbandsen B, Thomassen JR, Lund MS, Sahana G. Genome-wide association study for longevity with whole-genome sequencing in 3 cattle breeds. *J Dairy Sci*. 2016;99:7289–98.
 55. Mészáros G, Eaglen S, Waldmann P, Sölkner J. A Genome wide association study for longevity in cattle. *Open J Genet*. 2014;04:46–55.
 56. Sahana G, Guldbandsen B, Thomsen B, Holm L-E, Panitz F, Brøndum RF, et al. Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle. *J Dairy Sci*. 2014;97:7258–75.

57. Jiang J, Ma L, Prakapenka D, VanRaden PM, Cole JB, Da Y. A large-scale genome-wide association study in U.S. Holstein cattle. *Front Genet.* 2019;10:412.
58. Olsen HG, Knutsen TM, Lewandowska-Sabat AM, Grove H, Nome T, Svendsen M, et al. Fine mapping of a QTL on bovine chromosome 6 using imputed full sequence data suggests a key role for the group-specific component (GC) gene in clinical mastitis and milk production. *Genet Sel Evol.* 2016;48:79.
59. Bergbold N, Lemberg MK. Emerging role of rhomboid family proteins in mammalian biology and disease. *Biochim Biophys Acta.* 2013;1828:2840–8.
60. Gianola D, Sorensen D. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics.* 2004;167:1407–24.
61. Valente BD, Rosa GJM, de Los CG, Gianola D, Silva MA. Searching for recursive causal structures in multivariate quantitative genetics mixed models. *Genetics.* 2010;185:633–44.
62. Pegolo S, Momen M, Morota G, Rosa GJM, Gianola D, Bittante G, Cecchinato A. Structural equation modeling for investigating multi-trait genetic architecture of udder health in dairy cattle. *Sci Rep.* 2020;10:7751.
63. Momen M, Campbell MT, Walia H, Morota G. Utilizing trait networks and structural equation models as tools to interpret multi-trait genome-wide association studies. *Plant Methods.* 2019;15:107.
64. Haas V, Rodehutsord M, Camarinha-Silva A, Bennewitz J. Inferring causal structures of gut microbiota diversity and feed efficiency traits in poultry using Bayesian learning and genomic structural equation models. *J Anim Sci.* 2023;101:skad044.
65. Shi H, Mancuso N, Spendlove S, Pasaniuc B. Local genetic correlation gives insights into the shared genetic architecture of complex traits. *Am J Hum Genet.* 2017;101:737–51.
66. MacLeod IM, Bowman PJ, Vander Jagt CJ, Haile-Mariam M, Kemper KE, Chamberlain AJ, et al. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics.* 2016;17:144.
67. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature.* 2010;467:832–8.
68. Yang J, Weedon MN, Purcell S, Lettre G, Estrada K, Willer CJ, et al. Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet.* 2011;19:807–12.
69. Liu L, Zhou J, Chen CJ, Zhang J, Wen W, Tian J, et al. GWAS-based identification of new loci for milk yield, fat, and protein in Holstein cattle. *Animals (Basel).* 2020;10:2048.
70. Slob EAW, Burgess S. A comparison of robust Mendelian randomization methods using summary data. *Genet Epidemiol.* 2020;44:313–29.
71. Vereinigte Informationssysteme Tierhaltung w.V. Beschreibung der Zuchtwertschätzung für alle Schätzmerkmale bei den Milchrinderrassen für die vit mit der Zuchtwertschätzung beauftragt ist. 2023. https://www.vit.de/fileadmin/DE/Zuchtwertschaetzung/Zws_Bes_deu.pdf. Accessed 16 Oct 2023.

Publisher's Note

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