J. Dairy Sci. 102:503–510
https://doi.org/10.3168/jds.2018-14464

Genetic analysis of Fourier transform infrared milk spectra in Danish Holstein and Danish Jersey

R. M. Zaalberg,* N. Shetty, L. Janss, and A. J. Buitenhuis
Department of Molecular Biology and Genetics, Center for Quantitative Genetics and Genomics, Aarhus University, DK 8830 Tjele, Denmark

ABSTRACT

Fourier transform infrared milk spectral data are routinely used for milk quality control and have been revealed to be driven by genetics. This study aimed to (1) estimate heritability for 1,060 wavenumbers in the infrared region from 5,008 to 925 cm^{-1}, (2) estimate genomic correlations between wavenumbers with increased heritability, and (3) compare results between Danish Holstein and Danish Jersey cows. For Danish Holstein, 3,275 cows and 19,656 milk records were available. For Danish Jersey, 3,408 cows and 20,228 milk records were available. We used a hierarchical mixed model, with a Bayesian approach. Heritability of individual wavenumbers ranged from 0.00 to 0.31 in Danish Holstein, and from 0.00 to 0.30 in Danish Jersey. Genomic correlation was calculated between 15 selected wavenumbers, and varied from weak to very strong, in both Danish Holstein and Danish Jersey (0.03 to 0.97, and −0.11 to −0.97). Within the 15 selected wavenumbers, a subdivision into 2 groups of wavenumbers was observed, where genomic correlations were negative between groups, and positive within groups. Heritability and genomic correlations were higher in Danish Holstein compared with Danish Jersey, but followed a similar pattern in both breeds. Breed differences were most pronounced in the mid-infrared region that interacts with lactose and the spectral region that interacts with protein. In conclusion, heritability for individual wavenumbers of Fourier transform milk spectra was moderate, and strong genomic correlations were observed between wavenumbers across the spectrum. Heritability and genomic correlations were higher in Danish Holstein, with the strongest breed differences showing in spectral regions interacting with protein or lactose.
Key words: breed difference, genomic correlation, heritability, mid infrared

INTRODUCTION

Quality control of commercial milk is routinely performed with Fourier transform (mid) infrared [**FT-(M)IR**] milk spectral data. Milk spectra consist of wavenumbers, which provide insight in the chemical composition of milk (Karoui et al., 2010). Neighboring wavenumbers and wavenumbers across the MIR region are often phenotypically correlated (Soyeurt et al., 2010; Dagnachew et al., 2013). These correlations are used to apply FT-IR milk spectra in prediction of milk components, such as milk protein (Rutten et al., 2011), milk fat (Maurice-Van Eijndhoven et al., 2013; Ferragina et al., 2015), and specific fatty acids (**FA**; Rutten et al., 2009; Ferrand et al., 2010). The FT-IR milk spectra, additionally, have the potential to predict traits that are difficult or expensive to record, such as prevalence of metabolic diseases through milk β-BHB (Heuer et al., 2001; de Roos et al., 2007; Grelet et al., 2016). Attempts have been made to predict methane emissions (Dehareng et al., 2012; Vanlierde et al., 2015), energy balance (McParland et al., 2011, 2015), and residual feed intake (Shetty et al., 2017) from FT-IR milk spectra.

Genetic components have been revealed in FT-IR milk spectra, represented by moderate to high heritabilities (Soyeurt et al., 2010; Bittante and Cecchinato, 2013; Dagnachew et al., 2013; Wang et al., 2016). Genes such as diacylglycerol O-acyltransferase 1 (***DGAT1***), κ-casein (***CSN3***), and β-lactoglobulin (***LGB***) have been associated with individual wavenumbers (Wang et al., 2016).

Breed differences in milk traits have been observed, both phenotypically and genetically (Poulsen et al., 2015). Breed differences in genetics have been observed for *DGAT1* (Poulsen et al., 2015) and casein genes (Gustavsson et al., 2014). However, a knowledge gap still exists regarding the genetic basis of FT-IR milk spectra. Milk spectra could be applied more effectively in breeding programs if we have a better understanding of their genetic background. Breeders could, for example, predict new phenotypes, directly use wavenum-

Received January 17, 2018.
Accepted August 24, 2018.
*Corresponding author: roos.zaalberg@mbg.au.dk

bers in breeding programs, or use genetic correlations between wavenumbers and a trait of interest.

The aims of this study were to (1) calculate heritability (h^2) for 1,060 wavenumbers in the MIR region from 5,008 to 925 cm^{-1}, (2) calculate genomic correlations between transmittance values of wavenumbers with high h^2, and (3) compare results between Danish Holstein (**DH**) and Danish Jersey (**DJ**) cows.

MATERIALS AND METHODS

Study Population

Phenotypes. Milk records were provided by the Danish milk recording organization (RYK, Aarhus, Denmark). The study population consisted of 3,275 DH cows from 354 farms, and 3,408 DJ cows from 175 farms. On 4 farms, both DH and DJ cows were present. From farms with only DH, 1 to 112 cows were sampled, and from farms with only DJ, 1 to 217 cows were sampled. Milk records were sent to Eurofins-Steins laboratory (Vejen, Denmark) for FT-IR spectral analyses using MilkoScan FT+ (Foss, Hillerød, Denmark) and for analysis of the infrared region 5,008 to 925 cm^{-1}, including 1,060 individual wavenumbers.

Visual inspection of spectra was performed using PLSR Hotelling T^2 versus Q-residual plots and leverage versus studentized residual plots, and outliers were removed. Milk records with outlying fat percentage (fat%), and protein percentage (protein%) were excluded from the analyses. For DH, records were excluded when $fat\%_{DH} > 8.0$ or when $2.5 > protein\%_{DH} > 5.0$. For DJ, records were excluded when $fat\%_{DJ} > 8.0$, or $2.5 > protein\%_{DJ} > 5.5$.

For DH cows, 19,656 milk records were available, and for DJ 20,228 milk records were available. Cows were either first parity ($n_{DH} = 3{,}001$; $n_{DJ} = 3{,}125$) or second parity ($n_{DH} = 273$; $n_{DJ} = 283$). Each cow had 1 to 20 milk records. Average time between records was 32 d for both DH and DJ cows. Milk records were collected from October 2015 through September 2016. Milk records that had 2 records on the same day were excluded from the data set. Milk records were considered from the grazing season when obtained from April 2016 through September 2016. Milk records were considered from the nongrazing season when obtained from October 2015 through March 2016. Milk records were collected from 1 through 400 DIM. Milk records with DIM ≤ 5 accounted for 299 records in DH and 270 records in DJ.

Genotypes. Studied cows were genotyped with EuroG10K custom SNP chip (LD-chip). This LD-chip consists of SNP selected from whole genome sequence of bulls as part of the 1,000-bull genome project. Genotypes of studied cows were imputed from LD-chip to 50K using BEAGLE 4 (Browning and Browning, 2016). The reference population used for imputation consisted of 4,000 cows for DH and 4,576 cows for DJ. Each reference cow was genotyped for the Illumina 50K BovineSNP50 v.2 BeadChip (Illumina Inc., San Diego, CA). Before quality control of SNP, SNP with duplicate locations but different names were merged; SNP that were present in both the DH reference population and the DJ reference population were selected. After the first SNP selection step, for each breed separately, an identical quality control was performed on genotypes of studied cows combined with the reference population. During quality control, all autosomes were selected, SNP with >40% missing genotypes were excluded, and SNP with a minor allele frequency of <0.01 were excluded. After quality control, DH study cows had genotypes for 10,353 SNP and DJ study cows for 9,749 SNP. After quality control, DH genotypes were imputed from 10,353 to 43,807 SNP, and DJ genotypes from 9,749 to 39,235 SNP. Median distance between SNP was 41 kb for DH and 43 kb for DJ. Each SNP used for analysis is present on the Illumina BovineSNP50 v.2 BeadChip (Illumina Inc.).

Estimation of Heritability

Estimation of h^2 followed a Bayesian approach and was performed with the Bayz software package (http://www.bayz.biz/; Krag et al., 2013). A Metropolis-Hastings sampler was used, with 50,000 iterations and a burn-in of 10,000 iterations. For each of the 1,060 wavenumbers, h^2 was determined with a hierarchical model:

$$y_{ijklmp} = \mu_i + Season_{ij} + Parity_{ik} + \beta_{1,i}DIM + \beta_{2,i}e^{-0.05DIM} + Herd_{il} + Cow_{im,A} + Cow_{im,PE} + E_{ijklmp}, \quad \text{[Model 1]}$$

where y_{ijklmp} is the transmittance value for wavenumber i ($i = 1, \ldots, 1{,}060$) for one milk record ($p = 1, \ldots, 20$); μ_i is mean transmittance value for wavenumber i; $Season_{ij}$ is a fixed seasonal effect for wavenumber i ($j =$ grazing, from April through September; or not grazing, from October through March); $Parity_{ik}$ is a fixed effect for parity for wavenumber i ($k = 1$ or 2); DIM and $e^{-0.05DIM}$, with regressors $\beta_{1,i}$ and $\beta_{2,i}$, correct wavenumber i for lactation stage with the Wilmink function (Wilmink, 1987). All fixed effects had a uniform prior distribution fixed effect $\sim U(0,\infty)$; $Herd_{il}$ is a random herd effect for wavenumber i ($l_{DH} = 1, \ldots, 354$, and $l_{DJ} = 1, \ldots, 175$) with a normal prior distribution $Herd_{il} \sim N\left(\mathbf{0},\sigma^2_{herd}\right)$; $Cow_{im,PE}$ is the permanent environ-

mental effect for cow m for wavenumber i, with a normal prior distribution $Cow_{im,PE} \sim N\left(\mathbf{0},\sigma^2_{PE}\right)$; and E_{ijklmp} is the residual variance of the model, with a normal prior distribution $E_{ijklmp} \sim N\left(\mathbf{0},\sigma^2_E\right)$. All variance parameters had a uniform prior $\sim U\left(0,+\infty\right)$. $Cow_{m,A}$ was modeled using SNP data with a hierarchical model:

$$Cow_{im,A} = \sum_p a_p g_{mp}, \qquad \text{[Model 2]}$$

where $Cow_{im,A}$ is the additive genetic value for cow m ($m_{DH} = 1,\ldots, 3{,}276$, and $m_{DJ} = 1,\ldots, 3{,}408$) for wavenumber i. $Cow_{im,A}$ is the sum of p additive SNP effects a_p, times allele dosage g_{mp}, for animal m at SNP p. Allele dosages were cantered, missing genotypes were replaced by the mean dosage, and SNP additive effects had a power LASSO (Gao et al., 2013) prior distribution with power parameter 0.5.

Heritability was calculated with

$$h^2 = \frac{\sigma^2_A}{\sigma^2_{Herd} + \sigma^2_E + \sigma^2_A + \sigma^2_{PE}}, \qquad \text{[Model 3]}$$

where σ^2_A is the additive genetic variance, which was evaluated as $E[var(Cow_{m,A})]$ over (50,000–10,000 burn-in) Markov chain Monte Carlo (MCMC) cycles; σ^2_{Herd} is the herd variance; σ^2_E is the residual variance; and σ^2_{PE} is the permanent environmental variance. Standard error of h^2 was calculated as the standard deviation of the mean h^2 of all iterations used for analysis.

Estimation of Genomic Correlations

Phenotypes. Wavenumbers were selected based on results of the estimation of h^2. For MIR regions that showed a peak in h^2, and where this peak $h^2 \geq 0.10$, the wavenumber with the highest h^2 was selected.

Phenotypes used for the estimation of the genomic correlation underwent an extra filtering step where records were selected when $-4 <$ transmittance value <4. This better approached normal distribution of response variables, which consequently resulted in faster computation of genomic correlation. For DH, 3,266 animals and 19,435 records remained for analysis, whereas for DJ, 3,402 animals and 20,067 records remained for analysis.

Analysis. Genomic correlations were calculated with a Bayesian 2-trait SNP-BLUP model, where SNP effects were correlated between traits. Model 1 and Model 2 were used for the analyses, where all parameters were the same as described before, except for $Cow_{m,A}$. For estimation of h^2, $Cow_{m,A}$ used a normal prior distribution for SNP additive effects. Bayz software (http://www.bayz.biz/) was used with a Metropolis-Hastings sampler, 100,000 iterations, and a burn-in of 50,000 iterations.

RESULTS

Heritability

Figure 1 presents h^2 for individual wavenumbers from the FT-IR region of 5,008 to 925 cm^{-1}. Maximum h^2 was 0.31 in DH and 0.30 in DJ. The range of h^2 and mean ratio between h^2_{DH} and h^2_{DJ} are shown in Table 1 for 5 MIR regions: (1) total MIR region (5,008–925 cm^{-1}), (2) first MIR region associated with fat (3,015–2,800 cm^{-1}; fat-region-I), (3) second MIR region associated with fat (1,770–1,680 cm^{-1}; fat-region-II), (4) MIR region associated with protein (1,600–1,240 cm^{-1}; protein-region), and (5) MIR region associated with lactose (1,200–925 cm^{-1}; lactose-region). We calculated genomic correlations between 15 selected wavenumbers with $h^2 > 0.10$ (Figure 1).

Genomic Correlation Analysis

Table 2 presents genomic correlations ($\boldsymbol{r_g}$) between the 15 selected wavenumbers for DH and DJ. Table 3 presents the range of r_g between wavenumbers of 3 MIR regions (lactose-, protein-, and fat-regions). Within the lactose-region (1,200–925 cm^{-1}), in both DH and DJ, wavenumber 964 cm^{-1} was negatively correlated with 1,106 cm^{-1} ($r_{g,DH} = -0.68$; $r_{g,DJ} = -0.33$) and 1,145 cm^{-1} ($r_{g,DH} = -0.93$; $r_{g,DJ} = -0.87$), whereas 1,106 and 1,145 cm^{-1} were strongly positively correlated with each other ($r_{g,DH} = 0.95$; $r_{g,DJ} = 0.91$). Wavenumber 1,106 cm^{-1}, furthermore, showed weaker r_g with wavenumbers from the fat-region, and protein-region compared with other wavenumbers from the lactose-region (Table 2). Within the protein-region, wavenumbers 1,392 and 1,542 cm^{-1} showed weaker r_g with wavenumbers from the lactose-, and fat-region compared with other wavenumbers from the protein-region (Table 2). Within the fat-region (3,015–1,680 cm^{-1}, fat-region-I and fat-region-II), a subdivision into 2 groups of wavenumbers was observed: **fat-i** (1,708, 2,097, and 3,015 cm^{-1}), and **fat-ii** (1,746, 2,884, and 2,964 cm^{-1}). Wavenumbers from fat-i showed strong positive r_g with wavenumbers from fat-i ($r_{g,DH} > 0.94$; $r_{g,DJ} > 0.95$), wavenumbers from fat-ii showed strong positive r_g with wavenumbers from fat-ii ($r_{g,DH} > 0.95$; $r_{g,DJ} > 0.93$), but wavenumbers from fat-i showed strong negative r_g with wavenumbers from fat-ii ($r_{g,DH} < -0.88$; $r_{g,DJ} < -0.88$).

Combining all 15 wavenumbers, a subdivision into 2 groups was visible (Table 2). Wavenumbers from group

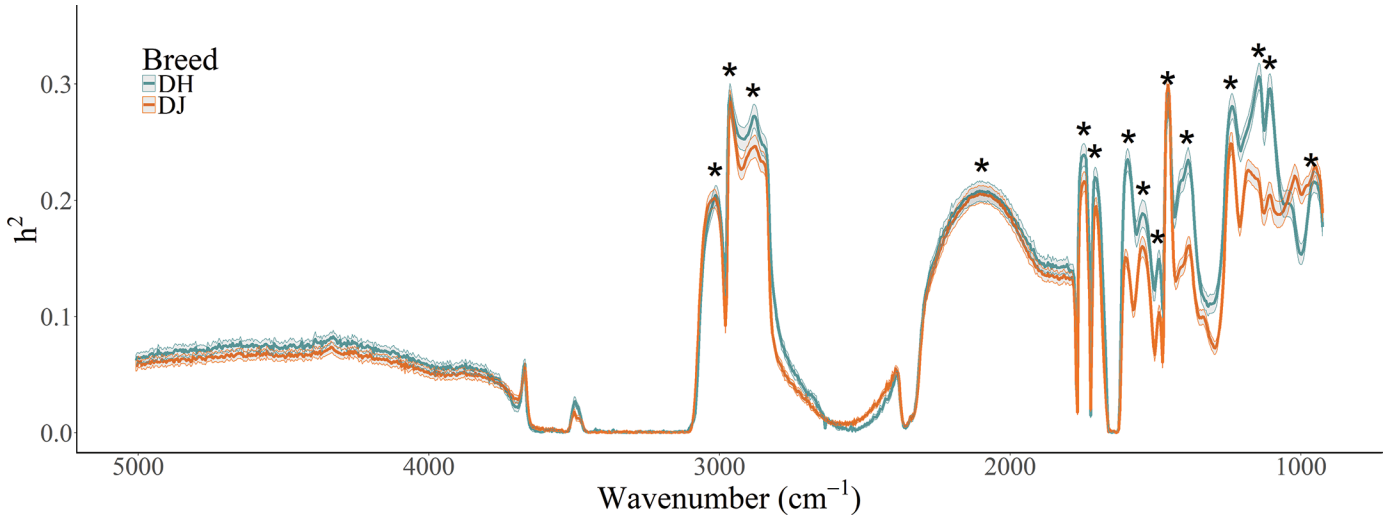

Figure 1. Heritability (h^2) with SE (gray) for 1,060 individual wavenumbers from the Fourier transform infrared region of 5,008 to 925 cm^{-1} in Danish Holstein (DH, blue line) and Danish Jersey (DJ, orange line). *Wavenumbers selected for calculation of genomic correlations are indicated with an asterisk (n = 15).

A showed strong positive r_g with wavenumbers from group A ($mean_{r_{g,DH}} = 0.94$; $mean_{r_{g,DJ}} = 0.92$), wavenumbers from group B showed strong positive r_g with wavenumbers from group B ($mean_{r_{g,DH}} = 0.92$; $mean_{r_{g,DJ}} = 0.88$), but wavenumbers from group A showed strong negative r_g with wavenumbers from group B ($mean_{r_{g,DH}} = -0.82$; $mean_{r_{g,DJ}} = -0.77$). Heritability of wavenumbers from group A was lower ($mean_{h^2_{DH}} = 0.20$, $mean_{h^2_{DJ}} = 0.18$) than h^2 of wavenumbers from group B ($mean_{h^2_{DH}} = 0.28$, $mean_{h^2_{DJ}} = 0.24$).

Breed Differences

Heritability of individual wavenumbers was on average 15% higher in DH than in DJ (Table 1). These breed differences were most pronounced in the protein-region, where individual wavenumbers had on average 39% higher h^2 in DH than in DJ (Table 1).

Genomic correlations between selected wavenumbers were on average 13% higher in DH than in DJ (Table 3). These breed differences were most pronounced for wavenumber 1,106 cm^{-1} from the lactose-region, and wavenumbers 1,542 and 1,392 cm^{-1} from the protein-region (Table 2). Direction of r_g between wavenumbers (positive or negative) was identical for DH and DJ, with the exception of r_g between 1,542 and 3,015 cm^{-1} (Table 2).

DISCUSSION

This study aimed at (1) calculating h^2 for 1,060 wavenumbers in the MIR region from 5,008 to 925 cm^{-1}, (2) calculating genomic correlations between transmittance values of wavenumbers with high h^2, and (3) comparing results between DH and DJ cows.

Breed differences in genetics of wavenumbers from the MIR region had not yet been studied directly. Differences between DH and DJ in h^2 were most pro-

Table 1. Range of heritability and ratio between h^2 for Danish Holstein (DH) and Danish Jersey (DJ) for regions within the spectral region of 5,008 to 925 cm^{-1}, in Danish Holstein and Danish Jersey

MIR[1] region		Range heritability		
Name	Range (cm^{-1})	Danish Holstein	Danish Jersey	Mean h^2_{DH}/h^2_{DJ}
Total	5,008–925	0.00–0.31	0.00–0.30	1.15
Fat-I	3,015–2,800	0.10–0.29	0.07–0.28	1.14
Fat-II	1,770–1,680	0.02–0.24	0.02–0.22	1.21
Protein	1,600–1,240	0.11–0.28	0.06–0.30	1.39
Lactose	1,200–925	0.15–0.31	0.19–0.23	1.10

[1]MIR = mid infrared.

Table 2. Genomic correlations for 15 wavenumbers selected from 15 spectral regions with a high h^2 in Danish Holstein (top) and Danish Jersey (bottom)[1]

		Wavenumber (cm^{-1})														
Wavenumber (cm^{-1})		3,015	2,964	2,884	2,097	1,746	1,708	1,592	1,542	1,492	1,457	1,392	1,241	1,145	1,106	964
		F-I	F-I	F-I	–	F-II	F-II	P	P	P	P	P	P	L	L	L
Danish Holstein																
3,015	F-I	**0.20**	−0.88[B]	−0.92[B]	0.95	−0.96[B]	0.94	0.94	*0.01*[B]	0.95	−0.78[B]	*0.55*	−0.83[B]	−0.88[B]	−0.56[B]	0.96
2,964	F-I		**0.29**[B]	0.97	−0.96[B]	0.95	−0.95[B]	−0.78[B]	0.74	−0.73[B]	0.96	*0.18*	0.97	0.95	0.87	−0.94[B]
2,884	F-I			**0.27**[B]	−0.96[B]	0.95	−0.96[B]	−0.87[B]	*0.57*	−0.83[B]	0.94	*0.08*	0.96	0.97	0.91	−0.94[B]
2,097	–				**0.21**	−0.96[B]	0.97	0.91	−*0.50*[B]	0.92	−0.94[B]	*0.17*	−0.95[B]	−0.95[B]	−0.81[B]	0.96
1,746	F-II					**0.24**[B]	−0.96[B]	−0.94[B]	*0.15*	−0.94[B]	0.86	−*0.43*	0.92	0.94	0.72	−0.96[B]
1,708	F-II						**0.22**	0.93	−*0.30*[B]	0.93	−0.92[B]	*0.29*	−0.95[B]	−0.96[B]	−0.82[B]	0.96
1,592	P							**0.23**	*0.37*	0.95	−0.55[B]	0.79	−0.72[B]	−0.87[B]	−0.61[B]	0.92
1,542	P								**0.19**[B]	0.61	0.89	0.79	0.83	0.40	0.53	−0.31[B]
1,492	P									**0.15**	−0.43[B]	0.78	−0.60[B]	−0.84[B]	−0.48[B]	0.92
1,457	P										**0.28**[B]	0.53	0.97	0.92	0.86	−0.90[B]
1,392	P											**0.23**	0.35	0.09	0.47	0.31
1,241	P												**0.28**[B]	0.95	0.88	−0.93[B]
1,145	L													**0.31**[B]	0.95	−0.93[B]
1,106	L														**0.30**[B]	−0.68[B]
964	L															**0.21**
Danish Jersey																
3,015	F-I	**0.20**	−0.88[B]	−0.92[B]	0.95	−0.96[B]	0.96	0.90	−*0.23*[B]	0.94	−0.84[B]	*0.38*	−0.84[B]	−0.84[B]	−0.20[B]	0.96
2,964	F-I		**0.28**[B]	0.96	−0.95[B]	0.93	−0.96[B]	−0.74[B]	0.84	−0.75[B]	0.96	*0.31*	0.96	0.94	0.73	−0.93[B]
2,884	F-I			**0.25**[B]	−0.96[B]	0.94	−0.96[B]	−0.83[B]	*0.71*	−0.86[B]	0.96	*0.20*	0.96	0.95	0.76	−0.94[B]
2,097	–				**0.20**	−0.96[B]	0.97	0.85	−*0.67*[B]	0.91	−0.95[B]	*0.03*	−0.96[B]	−0.91[B]	−0.53[B]	0.96
1,746	F-II					**0.22**[B]	−0.95[B]	−0.89[B]	*0.37*	−0.93[B]	0.89	−*0.28*	0.91	0.91	0.41	−0.97[B]
1,708	F-II						**0.19**	0.88	−*0.54*[B]	0.92	−0.94[B]	*0.14*	−0.95[B]	−0.94[B]	−0.57[B]	0.96
1,592	P							**0.14**	*0.39*	0.94	−0.47[B]	0.84	−0.64[B]	−0.84[B]	−0.37[B]	0.84
1,542	P								**0.16**[B]	0.49	0.94	0.85	0.91	*0.58*	0.56	−*0.53*[B]
1,492	P									**0.10**	−0.48[B]	0.76	−0.60[B]	−0.81[B]	−0.11[B]	0.92
1,457	P										**0.30**[B]	0.63	0.97	0.92	0.71	−0.92[B]
1,392	P											**0.16**	0.50	*0.23*	0.58	*0.12*
1,241	P												**0.25**[B]	0.94	0.77	−0.92[B]
1,145	L													**0.22**[B]	0.91	−0.87[B]
1,106	L														**0.20**[B]	−0.33[B]
964	L															**0.22**

[1]F-I = fat-region-I (3,015–2,800 cm^{-1}); F-II = fat-region-II (1,770–1,618 cm^{-1}); P = protein-region (1,600–1,240 cm^{-1}); L = lactose-region (1,200–925 cm^{-1}). Bold: diagonal with h^2 of individual wavenumbers. Wavenumbers in group B are labeled with a superscript B. All other wavenumbers with no superscript are in group A. Underlined: genomic correlation (r_g) between wavenumbers where breed differences are most pronounced. Italic: r_g between wavenumber 1,392 and 1,542 cm^{-1}, and wavenumbers from the fat and lactose region, where r_g is reduced compared with other wavenumbers of the protein region. Standard errors for genomic correlations ranged from 0.00 to 0.07.

Table 3. Overview of minimum (Min) and maximum (Max) genomic correlations individual wavenumbers for 3 spectral regions in Danish Holstein and Danish Jersey[1]

	Danish Holstein				Danish Jersey				
	−		+		−		+		
Region[2]	Min	Max	Min	Max	Min	Max	Min	Max	Mean ($r_{g,DH}/r_{g,DJ}$)[3]
Total	−0.16	−0.97	0.01	0.97	−0.11	−0.97	0.03	0.97	1.13
F × F	−0.88	−0.96	0.94	0.97	−0.88	−0.96	0.93	0.97	1.00
P × F	−0.30	−0.97	0.01	0.97	−0.23	−0.96	0.03	0.96	1.10
P × P	−0.16	−0.72	0.15	0.97	−0.47	−0.64	0.39	0.97	0.98
L × F	−0.56	−0.96	0.72	0.97	−0.20	−0.97	0.41	0.96	1.22
L × P	−0.31	−0.93	0.09	0.95	−0.11	−0.92	0.12	0.94	1.27
L × L	−0.68	−0.93	—	0.95	−0.33	−0.87	—	0.91	1.40

[1]The far right column presents mean ratio of genomic correlations for Danish Holstein and Danish Jersey.

[2]F = fat-region-I and fat-region-II (3,015–2,800 cm^{-1} and 1,618–1,770 cm^{-1}); P = protein-region (1,600–1,240 cm^{-1}); L = lactose-region (1,200–925 cm^{-1}).

[3]r_g = genomic correlation; *DH* = Danish Holstein; *DJ* = Danish Jersey.

nounced in the lactose-region, and in the protein-region (Figure 1). Increased h^2 of FT-IR milk spectra predicted milk components (protein, fat, and FA content) was observed in Holstein Friesians compared with other breeds (Maurice-Van Eijndhoven et al., 2015). Differences in FT-IR milk-spectra between DH, and DJ have been observed, but were, in contrast with our findings, mainly caused by differences in MIR regions associated with fat (Shetty et al., 2017). Breed differences in h^2 for milk fat predicted from FT-IR milk-spectra, and milk protein predicted from FT-IR milk spectra were explained by differences in allele frequencies in major milk genes, such as *DGAT1* (Maurice-Van Eijndhoven et al., 2015), and casein genes (study on DH and DJ, Gustavsson et al., 2014). For the study population of the current study, the SNP explaining most additive genetic variation for all wavenumbers was ARS-BFGL-NGS-4939. This SNP has been strongly associated with *DGAT1* on chromosome 14 and multiple production traits (Wang et al., 2012; Li et al., 2014; Bovenhuis et al., 2015). In both DH and DJ of the current study, alleles *A* and *G* were observed, yet allele frequencies (***af***) deviated [$af\ (A_{DH}) = 0.76$, and $af\ (A_{DJ}) = 0.25$]. This difference in allele frequency could be the underlying cause of differences in h^2 and r_g. Genomic correlations were 13% higher in DH compared with DJ (Table 3). The direction and pattern of r_g, however, was the same for both DH and DJ. Differences in allele frequency for *DGAT1* between DH and DJ, as explained before, could be the underlying cause of differences in genomic correlations.

Within the lactose-region, the MIR region surrounding 1,106 cm^{-1} (1,020–1,110 cm^{-1}) has been strongly associated with lactose (Hashimoto and Kameoka, 2008; Bajdik et al., 2009). This association could explain why r_g to wavenumbers from the protein- and fat-region was weaker for wavenumber 1,106 cm^{-1} than r_g observed for 964 or 1,145 cm^{-1} (Table 2). Within the protein-region, wavenumber 1,392 cm^{-1} has been strongly associated with *LGB*, weakly with *DGAT1*, but not with *CSN* (Wang et al., 2016). These associations could explain weaker r_g to wavenumbers from the protein-, fat-, and lactose-region (Table 2). Consistent weaker genetic correlations have also been observed for LGB content in milk, when studying genetic correlations between milk contents of LGB, α-lactose, α_{S1}-CN, α_{S2}-CN, β-CN, and κ-CN (Sanchez et al., 2017). Within the fat-regions, wavenumbers in fat-i were associated with fat A (Kaylegian et al., 2009), transfat (Hewavitharana and van Brakel, 1997), and *DGAT1* (Wang et al., 2016). Wavenumbers in fat-ii were associated with both Fat-A and Fat-B (Kaylegian et al., 2009), C–H stretching (Safar et al., 1994), and *DGAT1* (Wang et al., 2016). Negative r_g between wavenumbers from fat-i and fat-ii could be explained by differences in FA synthesis during early lactation (body fat mobilization), and mid to late lactation (de novo synthesis; Palmquist and Beaulieu, 1993). An interaction between lactation stage and *DGAT1* has been observed, where *DGAT1* had positive effects on milk lactose%, protein%, and fat% at the start of lactation, and negative effects at the end of lactation (Bovenhuis et al., 2015). Negative correlations between FA, furthermore, have been observed for FA traits expressed in FA% of milk fat (Soyeurt et al., 2007; Mele et al., 2009). The subdivision of the fat region seems to radiate through into the full MIR region (5,008–925 cm^{-1}; Table 2), which could be explained by a strong association between *DGAT1* to the full MIR region (Wang et al., 2016).

Heritability for wavenumbers of FT-IR milk spectra varied from 0.00 to 0.27 in Brown Swiss (Bittante, and Cecchinato, 2013), from 0.00 to 0.42 in first-parity Hol-

stein (Soyeurt et al., 2010), and from 0.00 to 0.60 in first-parity Holstein Friesians (Wang et al., 2016). In the current study, h^2 of individual wavenumbers was on the lower side, ranging up to 0.31 in DH, and 0.30 in DJ. A reason for the low h^2 observed in the current study could be the use of different breeds. A more likely reason for low h^2 would be the different methods used to calculate h^2 in the different studies. We used a power LASSO distribution for the additive SNP effects. The LASSO distribution fits SNP effects into a distribution where the majority of SNP have no effect, and a small number of SNP a large effect (Tibshirani, 1996). This can lead to underestimation of additive genetic variation. The observed moderate h^2 in combination with strong r_g, however, can be useful in breeding programs.

CONCLUSIONS

This study showed that transmittance of individual FT-IR wavenumbers in bovine milk was heritable. A large difference was observed in h^2 between the DH and DJ breeds, especially for wavenumbers interacting with protein and lactose. This study, furthermore, showed strong genomic correlations between the individual IR wavenumbers across the spectrum, which were higher in DH compared with DJ. The results of this study open up the possibility of exploring the underlying genetic architecture of different individual FT-IR wavenumbers. This would improve understanding of genetics of milk composition. Milk FT-IR spectra, so far, are used mainly indirectly to predict phenotypes. Showing that there is a genetic component underlying individual IR wavenumbers would suggest a possibility to use IR wavenumbers directly. Genomic correlations to (new) phenotypes, which are expensive to measure, could help to generate a source of information to improve breeding value estimation.

ACKNOWLEDGMENTS

This study was financially supported by the project: "FT-IR spektre i mælk: Genetisk variation og effekt på sundhed, frugtbarhed og energibalance" of the Milk Levy Fund, Denmark (2015-2017). RMZ was financially supported by The Graduate School of Science and Technology (GSST) of Aarhus University.

REFERENCES

Bajdik, J., Z. Makai, O. Berkesi, K. Süvegh, T. Marek, I. Eros, and K. Pintye-Hódi. 2009. Study of the effect of lactose on the structure of sodium alginate films. Carbohydr. Polym. 77:530–535.

Bittante, G., and A. Cecchinato. 2013. Genetic analysis of the Fourier-transform infrared spectra of bovine milk with emphasis on individual wavelengths related to specific chemical bonds. J. Dairy Sci. 96:5991–6006.

Bovenhuis, H., M. H. P. W. Visker, H. J. F. van Valenberg, A. J. Buitenhuis, and J. A. M. van Arendonk. 2015. Effects of the DGAT1 polymorphism on test-day milk production traits throughout lactation. J. Dairy Sci. 98:6572–6582.

Browning, B. L., and S. R. Browning. 2016. Genotype imputation with millions of reference samples. Am. J. Hum. Genet. 98:116–126.

Dagnachew, B. S., A. Kohler, and T. Ådnøy. 2013. Genetic and environmental information in goat milk Fourier transform infrared spectra. J. Dairy Sci. 96:3973–3985.

de Roos, A. P. W., H. J. C. M. van den Bijgaart, J. Hørlyk, and G. de Jong. 2007. Screening for subclinical ketosis in dairy cattle by Fourier transform infrared spectrometry. J. Dairy Sci. 90:1761–1766.

Dehareng, F., C. Delfosse, E. Froidmont, H. Soyeurt, C. Martin, N. Gengler, A. Vanlierde, and P. Dardenne. 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. Animal 6:1694–1701.

Ferragina, A., G. de los Campos, A. I. Vazquez, A. Cecchinato, and G. Bittante. 2015. Bayesian regression models outperform partial least squares methods for predicting milk components and technological properties using infrared spectral data. J. Dairy Sci. 98:8133–8151.

Ferrand, M., M. Ferrand, B. Huquet, S. Barbey, F. Barillet, F. Faucon, H. Larroque, O. Leray, J. M. Trommenschlager, and M. Brochard. 2010. Determination of fatty acid profile in cow's milk using mid-infrared spectrometry: Interest of applying a variable selection by genetic algorithms before a PLS regression. Chemometr. Intell. Lab. 106:183–189.

Gao, H., M. S. Lund, Y. Zhang, and G. Su. 2013. Accuracy of genomic prediction using different models and response variables in the Nordic Red cattle population. J. Anim. Breed. Genet. 130:333–340.

Grelet, C., C. Bastin, M. Gelé, J.-B. Davière, M. Johan, A. Werner, R. Reding, J. A. Fernandez Pierna, F. G. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, and F. Dehareng. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, beta-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99:4816–4825.

Gustavsson, F., A. J. Buitenhuis, M. Johansson, H. P. Bertelsen, M. Glantz, N. A. Poulsen, H. Lindmark Månsson, H. Stålhammar, L. B. Larsen, C. Bendixen, M. Paulsson, and A. Andrén. 2014. Effects of breed and casein genetic variants on protein profile in milk from Swedish red, Danish Holstein, and Danish Jersey cows. J. Dairy Sci. 97:3866–3877.

Hashimoto, A., and T. Kameoka. 2008. Applications of infrared spectroscopy to biochemical, food, and agricultural processes. Appl. Spectrosc. Rev. 43:416–451.

Heuer, C., H. J. Luinge, E. T. G. Lutz, Y. H. Schukken, J. H. van der Maas, H. Wilmink, and J. P. T. M. Noordhuizen. 2001. Determination of acetone in cow milk by Fourier Transform infrared spectroscopy for the detection of subclinical ketosis. J. Dairy Sci. 84:575–582.

Hewavitharana, A. K., and V. van Brakel. 1997. Fourier transform infrared spectrometric method for the rapid determination of casein in raw milk. Analyst (Lond.) 122:701–704.

Karoui, R., G. Downey, and C. Blecker. 2010. Mid-infrared spectroscopy coupled with chemometrics: A tool for the analysis of intact food systems and the exploration of their molecular structure-quality relationships - A review. Chem. Rev. 110:6144–6168.

Kaylegian, K. E., J. M. Lynch, J. R. Fleming, and D. M. Barbano. 2009. Influence of fatty acid chain length and unsaturation on mid-infrared milk analysis. J. Dairy Sci. 92:2485–2501.

Krag, K., N. A. Poulsen, M. K. Larsen, L. B. Larsen, L. L. Janss, and B. Buitenhuis. 2013. Genetic parameters for milk fatty acids in Danish Holstein cattle based on SNP markers using a Bayesian approach. BMC Genet. 14:79.

Li, C., S. Zhang, S. Wang, X. Wu, Q. Zhang, L. Liu, Y. Li, and L. Qiao. 2014. Genome wide association study identifies 20 novel promising genes associated with milk fatty acid traits in Chinese Holstein. PLoS One 9:e96186.

Maurice-Van Eijndhoven, M. H. T., H. Bovenhuis, H. Soyeurt, and M. P. L. Calus. 2013. Differences in milk fat composition predicted by mid-infrared spectrometry among dairy cattle breeds in the Netherlands. J. Dairy Sci. 96:2570–2582.

Maurice-Van Eijndhoven, M. H. T., R. F. Veerkamp, H. Soyeurt, and M. P. L. Calus. 2015. Heritability of milk fat composition is considerably lower for Meuse-Rhine-Yssel compared to Holstein Friesian cattle. Livest. Sci. 180:58–64.

McParland, S., G. Banos, E. Wall, M. P. Coffey, H. Soyeurt, R. F. Veerkamp, and D. P. Berry. 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. J. Dairy Sci. 94:3651–3661.

McParland, S., E. Kennedy, E. Lewis, S. G. Moore, B. McCarthy, M. O'Donovan, and D. P. Berry. 2015. Genetic parameters of dairy cow energy intake and body energy status predicted using mid-infrared spectrometry of milk. J. Dairy Sci. 98:1310–1320.

Mele, M., R. Dal Zotto, M. Cassandro, G. Conte, A. Serra, A. Buccioni, G. Bittante, and P. Secchiari. 2009. Genetic parameters for conjugated linoleic acid, selected milk fatty acids, and milk fatty acid unsaturation of Italian Holstein-Friesian cows. J. Dairy Sci. 92:392–400.

Palmquist, D. L., and A. D. Beaulieu. 1993. Feed and animal factors influencing milk fat composition. J. Dairy Sci. 76:1753–1771.

Poulsen, N. A., A. J. Buitenhuis, and L. B. Larsen. 2015. Phenotypic and genetic associations of milk traits with milk coagulation properties. J. Dairy Sci. 98:2079–2087.

Rutten, M. J. M., H. Bovenhuis, J. M. L. Heck, and J. A. M. van Arendonk. 2011. Predicting bovine milk protein composition based on Fourier transform infrared spectra. J. Dairy Sci. 94:5683–5690.

Rutten, M. J. M., H. Bovenhuis, K. A. Hettinga, H. J. F. van Valenberg, and J. A. M. van Arendonk. 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. J. Dairy Sci. 92:6202–6209.

Safar, M., D. Bertrand, P. Robert, M. F. Devaux, and C. Genot. 1994. Characterization of edible oils, butters and margarines by Fourier transform infrared spectroscopy with attenuated total reflectance. J. Am. Oil Chem. Soc. 71:371–377.

Sanchez, M. P., M. Ferrand, M. Gelé, D. Pourchet, G. Miranda, P. Martin, M. Brochard, and D. Boichard. 2017. Short communication: Genetic parameters for milk protein composition predicted using mid-infrared spectroscopy in the French Montbéliarde, Normande, and Holstein dairy cattle breeds. J. Dairy Sci. 100:6371–6375.

Shetty, N., P. Løvendahl, M. S. Lund, and A. J. Buitenhuis. 2017. Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid-infrared spectroscopy of milk. J. Dairy Sci. 100:253–264.

Soyeurt, H., A. Gillon, S. Vanderick, P. Mayeres, C. Bertozzi, and N. Gengler. 2007. Estimation of heritability and genetic correlations for the major fatty acids in bovine milk. J. Dairy Sci. 90:4435–4442.

Soyeurt, H., I. Misztal, and N. Gengler. 2010. Genetic variability of milk components based on mid-infrared spectral data. J. Dairy Sci. 93:1722–1728.

Tibshirani, R. 1996. Regression shrinkage and selection via the Lasso. J. R. Stat. Soc. B 58:267–288.

Vanlierde, A., M.-L. Vanrobays, F. Dehareng, E. Froidmont, H. Soyeurt, S. McParland, E. Lewis, M. H. Deighton, F. Grandl, M. Kreuzer, B. Gredler, P. Dardenne, and N. Gengler. 2015. Hot topic: Innovative lactation-stage-dependent prediction of methane emissions from milk mid-infrared spectra. J. Dairy Sci. 98:5740–5747.

Wang, Q., A. Hulzebosch, and H. Bovenhuis. 2016. Genetic and environmental variation in bovine milk infrared spectra. J. Dairy Sci. 99:6793–6803.

Wang, X., C. Wurmser, H. Pausch, S. Jung, F. Reinhardt, J. Tetens, G. Thaller, and R. Fries. 2012. Identification and dissection of four major QTL affecting milk fat content in the German Holstein-Friesian population. PLoS One 7:e40711.

Wilmink, J. B. M. 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. Livest. Prod. Sci. 16:335–348.