Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds

Irene van den Berg,*^{†1} Didier Boichard,[†] and Mogens Sandø Lund*

*Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, DK-8830 Tjele, Denmark,

[†]GABI, INRA, AgroParisTech, Université Paris Saclay, 78350 Jouy-en-Josas, France

ABSTRACT

The objective of this study was to compare mapping precision and power of within-breed and multibreed genome-wide association studies (GWAS) and to compare the results obtained by the multibreed GWAS with 3 meta-analysis methods. The multibreed GWAS was expected to improve mapping precision compared with a within-breed GWAS because linkage disequilibrium is conserved over shorter distances across breeds than within breeds. The multibreed GWAS was also expected to increase detection power for quantitative trait loci (QTL) segregating across breeds. GWAS were performed for production traits in dairy cattle, using imputed full genome sequences of 16,031 bulls, originating from 6 French and Danish dairy cattle populations. Our results show that a multibreed GWAS can be a valuable tool for the detection and fine mapping of quantitative trait loci. The number of QTL detected with the multibreed GWAS was larger than the number detected by the within-breed GWAS, indicating an increase in power, especially when the 2 Holstein populations were combined. The largest number of QTL was detected when all populations were combined. The analysis combining all breeds was, however, dominated by Holstein, and QTL segregating in other breeds but not in Holstein were sometimes overshadowed by larger QTL segregating in Holstein. Therefore, the GWAS combining all breeds except Holstein was useful to detect such peaks. Combining all breeds except Holstein resulted in smaller QTL intervals on average, but this outcome was not the case when the Holstein populations were included in the analysis. Although no decrease in the average QTL size was observed, mapping precision did improve for several QTL. Out of 3 different multibreed meta-analysis methods, the weighted z-scores model resulted in the most similar results to the full multibreed GWAS and can be useful as an alternative

to a full multibreed GWAS. Differences between the multibreed GWAS and the meta-analyses were larger when different breeds were combined than when the 2 Holstein populations were combined.

Key words: genome-wide association studies (GWAS), multibreed, meta-analysis, whole genome sequence

INTRODUCTION

Genome-wide association studies (GWAS) are used to find associations between traits and polymorphisms. With the increasing number of sequences available, more causative mutations will be among the variants and can be detected directly in the GWAS. However, markers in high linkage disequilibrium (LD) with the causative mutations can show an equally high or even higher association than the true causative mutations. Genome-wide association studies often result in large regions associated with the same QTL because of long-range LD observed in dairy cattle breeds (de Roos et al., 2008). Across different breeds, LD is only conserved over short distances. Therefore, multibreed mapping could help improve the precision of GWAS. Furthermore, with the large number of variants studied in a sequence-based GWAS, a high detection threshold is necessary to avoid too many false positives. Quantitative traits can be influenced by many causative mutations with small individual effects, and unless large data sets are used, these effects could be too small to pass the thresholds. Assuming a mutation is indeed shared between different breeds, combining data of multiple breeds increases the sample size and thereby the detection power.

In addition to aiding the identification of causative mutations, improved GWAS precision could also help in selecting variants that are subsequently used for genomic prediction. Genomic prediction is widely used in dairy cattle, with high accuracies within breed, although accuracies of across-breed predictions are much lower (Hayes et al., 2009; Erbe et al., 2012; Lund et al., 2014). Unless only markers in full LD with the causative mutations are used, a loss in prediction reliability

Received February 22, 2016.

Accepted June 23, 2016.

¹Corresponding author: irene.vandenberg@unimelb.edu.au

occurs (de los Campos et al., 2013). This decrease is larger across breeds than within a breed (van den Berg et al., 2014). Using variants in high LD with causative mutations could therefore increase prediction accuracy. This increase could be especially beneficial for multi-breed prediction, if variants in high LD with causative mutations shared across breeds are used. A multibreed GWAS could be used to detect such variants.

Joining data of individual GWAS is not always possible, however. Alternatively, rather than using the full data, results of individual GWAS can be combined using meta-analysis (Begum et al., 2012; Evangelou and Ioannidis, 2013), as is commonly done in human genetics. Using both simulated and real data, Lin and Zeng (2010) found equal efficiency with a meta-analysis as when data from individual studies were used.

The objective of this study was to compare different strategies to perform a multibreed analysis using sequence data from 6 French and Danish dairy cattle populations. Results obtained by within-breed GWAS were compared with those of a multibreed GWAS. Furthermore, several meta-analysis methods were compared with a multibreed GWAS. Specifically, we tested the following hypotheses: (1) the power of detecting QTL is larger for a multibreed GWAS than a within-breed GWAS, (2) the QTL detection is more precise for a multibreed GWAS than a within-breed GWAS, and (3) a multibreed GWAS can be approximated by a meta-analysis of within-breed GWAS.

MATERIALS AND METHODS

Data

Imputed sequences of 4,993 Danish Holstein, 984 Jersey, 768 Danish Red, 5,626 French Holstein, 1,935 Montbéliarde, and 1,725 Normande bulls were used. The majority of sequences were obtained by imputation of bulls genotyped with the 50K and HD SNP chips (Illumina Inc., San Diego, CA). First, bulls genotyped with the 50K chip were imputed to the HD chip. For the French data (Hozé et al., 2013), this step was performed using Beagle 3.0.0 (Browning and Browning, 2007), while for the Danish breeds IMPUTE2 was used (Howie et al., 2012). Subsequent imputation to whole-genome sequence was done for all breeds, using IMPUTE2. Imputation was done within each country, imputing Danish Holstein, Jersey, and Danish Red together, and French Holstein, Montbéliarde, and Normande together. The reference used for imputation to sequences of the Danish bulls consisted of the bulls in run 4 of the 1,000 Bull Genomes project (Daetwyler et al., 2014), while for the imputation of the French bulls, a combined French and Danish reference set was

used. The latter consisted of 122 Holstein, 27 Jersey, 28 Montbéliarde, 23 Normande, and 45 Danish Red bulls. More details on the Danish imputation can be found in Höglund et al. (2015). For all bulls, sires were known and deregressed proofs (**DRP**) for milk, fat, and protein yields were used in the analysis.

All genotypic and phenotypic data were obtained from preexisting routine genetic evaluation data for the dairy cattle populations and required no ethical approval.

Principal Component Analysis

To study genomic relationships between breeds, a genomic relationship was constructed using SNP from the 50K chip for 500 randomly selected individuals of each breed. For these individuals, genomic relationship matrix **G** was constructed following VanRaden (2008):

$$\mathbf{G} = \frac{\mathbf{Z}\mathbf{Z}'}{2\sum p_j(1-p_j)},$$

where **Z** was a standardized genotype matrix, and p_j was the allele frequency of the alternative allele for locus j . Subsequently, **G** was used for a principal component analysis using the `prcomp()` command in R (R Development Core Team, 2015).

GWAS Within Breed

Variants with a within-breed minor allele frequency below 0.005 or an IMPUTE2 info score below 0.60 were discarded. These exclusions resulted in 24,550,115 SNP and small indels (insertion-deletions). The following model was run for each of these polymorphisms within each breed:

$$y_{ij} = \mu + S_j + \beta g_{ij} + e_{ij},$$

where y_{ij} is the DRP of milk yield, fat yield, or protein yield for individual i with sire j , S_j the random effect of sire j , β the effect of the polymorphism, g_{ij} the allele dose (ranging from 0 to 2) of individual i with sire j , and e_{ij} a random residual.

Multibreed GWAS

Three multibreed GWAS were run, combining French and Danish Holstein (**HOL**), combining Jersey, Danish Red, Montbéliarde, and Normande (**REST**), or combining all 6 populations (**ALL**). To reduce the computing time, the multibreed GWAS was only run for variants with a P -value $< 10^{-5}$ in the within-breed

GWAS in French or Danish Holstein or $< 10^{-3}$ in one of the other breeds for at least one trait. These thresholds were chosen after visualizing the results of the within-breed GWAS, to exclude only variants that were not part of a peak. The model for the multibreed GWAS was identical to the within-breed GWAS, except for the addition of breed effect R of breed k :

$$y_{ijk} = \mu + S_j + R_k + \beta g_{ijk} + e_{ijk}.$$

QTL Detection

Quantitative trait loci regions were defined as following: first, to adjust for multiple testing, P -values were adjusted using the false discovery rate (**FDR**). The FDR was computed and used to adjust P -values (P_{FDR}) from both the within-breed and multibreed GWAS following Benjamini and Hochberg (1995) using the `p.adjust()` command in R (R Development Core Team, 2015), which uses the following procedure to adjust P -values:

- P -values were ranked from smallest to largest.
- Adjusted P -values were computed as $P_{\text{FDR},m} = \frac{P_m \times n}{m}$, where P is the raw P -value for variant m , $P_{\text{FDR},m}$ is the adjusted P -value of variant m , and n is the total number of variants.
- P -values were further adjusted to ensure that $P_{\text{FDR},m} \leq P_{\text{FDR},m+1}$.

Subsequently, for each chromosome, all variants with a $P_{\text{FDR}} \leq 10^{-5}$ were ranked based on their P -value. The QTL regions were identified and defined by the following procedure. At the beginning, the top variant on each chromosome was declared as a peak, and all variants within 0.5 Mb of the peak with $-\log_{10}(P) \geq -\log_{10}(P_{\text{peak}}) \times \frac{2}{3}$ were considered as part of the same QTL. The borders of the QTL region were defined as the variants with $-\log_{10}(P) \geq -\log_{10}(P_{\text{peak}}) \times \frac{2}{3}$ furthest away but within 0.5 Mb of the peak. Note that this $2/3$ coefficient was found to be more robust than a fixed reduction of $-\log_{10}(P)$ for high peaks. Then, the borders were expanded by including variants within 0.5 Mb of the borders with a $-\log_{10}(P) \geq -\log_{10}(P_{\text{peak}}) \times \frac{2}{3}$. The second step was repeated until no further variants fulfilled the criteria. If a QTL region contained fewer than 5 variants with a P -value with $-\log_{10}(P) \geq -\log_{10}(P_{\text{peak}}) \times \frac{2}{3}$, it was discarded. Otherwise, all variants within 5 Mb of the borders were removed from the list of variants with a $P_{\text{FDR}} \leq 10^{-5}$, and the remaining top variant was de-

clared as the next peak. This process was repeated until no further variants with a $P_{\text{FDR}} \leq 10^{-5}$ remained.

Meta-Analysis

All variants analyzed in the within-breed GWAS were also analyzed using 3 different meta-analysis methods: the weighted z-score, the fixed effects model, and the random effects model. For the weighted z-score, METAL software was used (Willer et al., 2010), and for the fixed and random effects model, META software was used (Liu et al., 2010). The following sections provide a brief description of the different models, and a more detailed comparison can be found elsewhere (Begum et al., 2012; Evangelou and Ioannidis, 2013).

Weighted z-Scores

The weighted z-scores model uses P -values and the directions of marker effects as inputs and weights individual studies based on the sample size to compute a z-score:

$$Z_k = \Phi^{-1} \left(1 - \frac{P_k}{2} \right) \times \Delta_k,$$

where Z_k is the z-score for breed k , Φ is the cumulative distribution function, and P_k and Δ_k are the P -value and direction of the marker effect estimated within-breed k , respectively. Subsequently, within-breed z-scores were combined in an overall z-score $Z = \frac{\sum_k z_k w_k}{\sqrt{\sum_k w_k^2}}$,

where w_k equals $\sqrt{N_k}$ and N_k is the sample size for breed k . Finally, an overall P -value was estimated as $P = 2\Phi(|-Z|)$.

Fixed Effects Model

The fixed effects models weight within-breed effects based on their standard errors. A combined effect across breeds was computed as $B = \frac{\sum_k \beta_k w_k}{\sum_k w_k}$, where β_k is the effect obtained within breed k , and a weight $w_k = [\text{Var}(\beta_k)]^{-1}$. The variance of the combined effect across breeds equals $V = [\sum_k w_k]^{-1}$ with test statistic $\chi^2 = B^2/V$ and $\text{df} = 1$.

Random Effects Model

In contrast to the weighted z-scores and the fixed effects models, the random effects model accounts for heterogeneity between studies by assuming that the

mean effect in each study is different. First, Cochran's test for statistics was computed as $Q = \sum_k w_k (B - \beta_k)^2$. The between-study variance of heterogeneity equaled

$$\tau^2 = \frac{Q - N_k - 1}{\sum_k w_k - \left(\frac{\sum_k w_k^2}{\sum_k w_k} \right)}.$$

Subsequently, the weight for the random effects model for breed k equaled $w_k^* = [\tau^2 + \text{Var}(\beta_k)]^{-1}$ with combined effect across breeds

$$B^* = \frac{\sum_k \beta_k w_k^*}{\sum_k w_k^*},$$

across-breed variance $V^* = [\sum_k w_k^*]^{-1}$, test statistic $\chi^2 = B^{*2}/V^*$, and $\text{df} = 1$.

RESULTS

Principal Component Analysis

Figure 1 shows a principal component analysis of the genomic relationships between the different breeds used for the studies. French and Danish Holstein populations were very similar, and Danish Red was more similar to Holstein than Montbéliarde and Normande. Jersey was the most distinct from the other breeds.

Variants Selected for Multibreed GWAS

Figure 2 shows the number of breeds in which a variant is below 10^{-5} in Holstein or 10^{-3} in the other breeds. All variants for which this threshold is passed in at least one breed were used for the multibreed GWAS. All chromosomes contained a large number of variants passing the threshold in 1 or 2 breeds, but only a few variants had P -values below the threshold in more than 2 breeds.

Comparison of Within-Breed GWAS and Multibreed GWAS

Table 1 compares the average QTL size, peak size, and number of variants per QTL of the within-breed GWAS and the multibreed GWAS. The number of unique QTL detected in within-breed GWAS and the number of QTL detected in the multibreed GWAS can be found in Table 2. Combining the 2 Holstein populations in a joint analysis did not reduce the size of the QTL intervals. Adding the other breeds decreased the

size of the QTL intervals compared with HOL, but the size of the QTL detected by ALL was not lower than that of the QTL detected within Danish Holstein and French Holstein. REST resulted in smaller intervals than detected within breed, except for protein yield. On average, QTL detected by ALL were smaller than those detected within breed for milk and fat, but larger for protein.

Quantitative trait loci peaks had similar heights in the within-breed and multibreed GWAS. The average number of variants per QTL was lower for QTL detected with REST and ALL than detected within Jersey, Montbéliarde, Normande, and Danish Red. For QTL detected in HOL and ALL, the average number of QTL per region was in between the number of QTL detected within Danish Holstein and French Holstein.

Table 2 compares the number of QTL regions detected by the within-breed GWAS with the number QTL regions detected by multibreed GWAS. The multibreed GWAS resulted in an increased number of QTL regions compared with the within-breed GWAS for milk and fat. The largest increases were found for fat in Holstein,

Figure 1. Principal component analysis of genomic relationships. Showing principal components (PC) 1 and 2. Color version available online.

Figure 2. Variants used for multibreed genome-wide association study (GWAS), showing number of breeds in which variants had a P -value $< 10^{-3}$ in Holstein or 10^{-3} in Jersey, Montbéliarde, Normande, or Danish Red. Color version available online.

Table 1. Comparison of peak size and number of variants per QTL in the within-breed and multibreed genome-wide association studies (GWAS)

Analysis ¹	Milk				Fat				Protein			
	nQTL ²	Size ³	Max ⁴	No. ⁵	nQTL	Size	Max	No.	nQTL	Size	Max	No.
HDK	48	1.03	15.6	122	50	0.91	16.6	160	81	1.46	12.9	220
HFR	50	1.58	15.1	202	35	1.79	17.7	287	116	1.17	13.3	89
JER	11	1.30	13.5	222	1	1.66	13.1	170	0	—	—	—
MON	5	2.39	12.6	311	10	1.59	12.8	204	8	0.99	12.6	188
NOR	1	0.76	13.4	7	1	1.55	14.3	86	0	—	—	—
RDC	0	—	—	—	1	4.44	18.0	509	0	—	—	—
HOL	128	1.40	15.2	165	154	1.28	15.5	184	172	1.58	13.7	208
REST	20	1.28	13.6	121	16	0.85	14.1	85	5	1.89	14.3	107
ALL	146	1.08	15.3	135	152	1.12	15.8	125	166	1.34	14.0	120

¹HDK = within-breed GWAS, Danish Holstein; HFR = within-breed GWAS, French Holstein; JER = within-breed GWAS, Jersey; MON = within-breed GWAS, Montbéliarde; NOR = within-breed GWAS, Normande; RDC = within-breed GWAS, Danish Red; HOL = multibreed GWAS, combining Danish and French Holstein; REST = multibreed GWAS, combining Jersey, Montbéliarde, Normande and Danish Red; ALL = multibreed GWAS, combining all 6 populations.

²Number of detected QTL regions.

³Average QTL size (Mb).

⁴Average $-\log_{10}(P)$ of the peak.

⁵Average number of variants per QTL.

Table 2. Number of QTL detected with within-breed (WB) and multibreed (MB) genome-wide association studies

Breed ¹	Milk		Fat		Protein	
	WB	MB	WB	MB	WB	MB
HOL	82	128	72	154	175	172
REST	16	20	11	16	8	5
ALL	95	146	82	152	183	166

¹HOL = French and Danish Holstein; REST = Jersey, Montbéliarde, Normande, and Danish Red; ALL = French and Danish Holstein, Jersey, Montbéliarde, Normande, and Danish Red.

where HOL more than doubled the number of QTL regions compared with the within-breed GWAS.

Comparison of Multibreed GWAS and Meta-Analyses

Figure 3 compares the $-\log_{10}$ of the P -values obtained with the multibreed GWAS with those of the meta-analyses. The weighted z-scores model gave the most similar results to the multibreed GWAS, while the random effects model yielded results that were the most different. A large proportion of variants with low P -values in the multibreed GWAS had much larger P -values in the random effects model. For example, in the ALL analysis, 20% of the variants had a P -value at least 100 times larger in the random effects model than in the multibreed GWAS, while this was only 1% and 5% for the weighted z-scores model and the fixed effects model, respectively. For these variants, the heterogeneity computed by the random effects model was highly significant. For all models, combining Holstein yielded more similar results to the meta-analysis than combining more distinct breeds. For fat and protein, the meta-analyses were more similar to the multibreed GWAS than for milk.

Comparison of Multibreed GWAS, Within-Breed GWAS, and Meta-Analysis at Peaks

Three peaks were selected to illustrate the improved power (all examples) and precision (example 3) observed at certain peaks, as well as how the larger Holstein data sets dominated the results and overshadowed some of the smaller peaks observed in the smaller breeds (example 2). Figures 4 to 6 compare the results obtained by the within-breed GWAS and the multibreed GWAS for these examples. A comparison of the meta-analyses and the multibreed GWAS can be found in Supplemental Figures S1, S2, and S3 (<http://dx.doi.org/10.3168/jds.2016-11073>). For each example, P -values and annotations of the most significant variants detected in the within-breed and multibreed GWAS can be found in Tables 3 and 4.

Example 1, Chromosome 2, 105–107 Mb, Milk Yield. The first example demonstrates the increase in power observed at certain peaks in the multibreed GWAS. A peak around 105.8 Mb was visible in Danish Holstein, Danish Red, and Montbéliarde (Figure 4), and it was most pronounced in Danish Red and Montbéliarde. For HOL, a peak occurred around 105.8 Mb, and the P -values were lower than within either of the populations. A similar peak was observed for REST, and for ALL the P -value further decreased.

Example 2, Chromosome 5, 93–113 Mb, Fat Yield. The second example shows that, although detection power was increased, the multibreed GWAS was dominated by the large Holstein populations, and some of the peaks observed in the other breeds were overshadowed in the analysis combining all populations. Within the breed, a clear peak was observed in both Holstein populations (Figure 5). A similar but smaller peak was visible in Normande. Combining Holsteins or all breeds resulted in a larger peak, with the most significant variant located in an intron in *MGST1*. Around 112 Mb, another peak was observed in Normande, Jersey, and Holstein. When both Holstein populations were combined, the most significant variant was an intergenic variant with a $-\log_{10}(P)$ of 14.0. An intron variant in *MKL1* was the most significant variant when all breeds except Holstein were combined, with a $-\log_{10}(P)$ of 17.3. The most distinct peak was observed when all breeds except Holstein were combined.

Example 3, Chromosome 6, 60–100 Mb, Protein Yield. The third example shows how the multibreed GWAS improved mapping precision for a peak observed in all breeds. In all breeds, high $-\log_{10}(P)$ were observed on chromosome 6, especially within the Holsteins (Figure 6). When both Holsteins were combined, the most significant variant was located in an intron in *CSN1S1* with a $-\log_{10}(P)$ of 34. When all breeds except Holstein were combined, the highest $-\log_{10}(P)$ observed were smaller than within Holstein, but the peak was narrower, with an intergenic variant with a $-\log_{10}(P)$ of 22 being the most significant. This

variant was also the most significant when all breeds were combined with a $-\log_{10}(P)$ of 49.9.

DISCUSSION

Comparison of Within-Breed and Multibreed GWAS

Combining different breeds in a multibreed GWAS could either increase or decrease the number of detected variants. With an increased number of individuals, the detection power increases. This pattern assumes, however, that QTL are segregating across breeds. Raven et al. (2014) found an increase in power when combining

different populations of the same breed, but a decrease in power in a multibreed GWAS combining Holstein and Jersey compared with within-breed GWAS; they attributed this finding to the limited amount of QTL shared across breed.

In our study, all multibreed GWAS resulted in an increase in power compared with the within-breed GWAS for milk and fat, but a small decrease for protein yield. Not surprisingly, the largest increase in power was observed when the 2 Holstein populations were joined. These 2 closely related populations are the same breed rather than different breeds, and most QTL are expected to be shared between the breeds. Consequently,

Figure 3. Comparison of P -values obtained by multibreed genome-wide association study (GWAS; x-axis) and meta-analysis method (y-axis; z-scores, random, fixed). HOL = Danish and French Holstein; REST = Jersey, Montbéliarde, Normande, and Danish Red; ALL = Danish and French Holstein, Jersey, Montbéliarde, Normande, and Danish Red.

Figure 4. Example 1, within-breed (top and middle) and multibreed genome-wide association study (GWAS; bottom) for milk yield. Chromosome 2, between 105 and 107 Mb, HOLDK = Danish Holstein, JER = Jersey, RDC = Danish Red, HOLFR = French Holstein, NOR = Normande, MON = Montbéliarde, HOL = Danish and French Holstein, REST = all breeds except Holstein, ALL = all breeds.

approximately doubling the sample size led to large increases in power for milk and fat yield. Interestingly, this increase was not the case for protein yield. The total number of QTL detected for protein yield within Danish Holstein and French Holstein was much larger than for milk and fat. Because protein yield is more polygenic than milk and fat, each QTL explains a

smaller part of the variance; therefore, detection power is lower and leads to a lower sharing level across populations. Joining all breeds in the ALL analysis resulted in a further increase in the number of QTL regions for milk and fat. This increase indicates that some QTL are shared between Holstein and the other breeds, and that the power of the analyses without Holstein was

insufficient to detect a portion of these QTL. Joining Jersey, Montbéliarde, Normande, and Danish Red without Holstein also resulted in more QTL regions than detected within these breeds, although the increase was smaller than for HOL and ALL. Because of the much larger number of Holsteins in the data than any other

breed, the results of ALL were dominated by Holstein. As shown in example 2, QTL segregating in the smaller breeds but not in Holstein can be overshadowed in ALL by large Holstein peaks nearby. Therefore, running a joined GWAS both with and without Holstein was useful. Reducing the sample size of the Holstein

Figure 5. Example 2, within-breed (top and middle) and multibreed genome-wide association study (GWAS; bottom) for fat yield. Chromosome 5, between 93 and 113 Mb, HOLDK = Danish Holstein, JER = Jersey, RDC = Danish Red, HOLFR = French Holstein, NOR = Normande, MON = Montbéliarde, HOL = Danish and French Holstein, REST = all breeds except Holstein, ALL = all breeds.

Table 3. Functional annotation and within breed $-\log_{10}(P)$ for most significant variants in the regions of examples 1 to 3

Example	Trait	BTA	Position	rs-id ¹	Ann ²	Gene	Breed ³					
							HDK	HFR	JER	MON	NOR	RDC
1	Milk	2	105783655	rs382743604	intg	—	10.8	3.4	—	0.7	1.6	2.0
			105832735	rs210733747	intg	—	1.2	1.3	2.0	7.3	2.0	5.0
			105864893	rs382506988	intg	—	3.7	6.0	0.6	4.2	2.9	3.0
2	Fat	5	93305237	rs380052783	intg	—	7.6	1.0	0.3	5.6	5.6	1.8
			93945738	rs211210569	intr	<i>MGST1</i>	33.7	33.0	—	4.7	2.1	4.5
			111662916	—	intg	—	11.5	5.6	0.2	0.0	7.9	0.3
			112374194	rs110294643	intr	<i>MKL1</i>	3.9	0.0	7.1	2.4	14.2	0.7
			112397190	rs210164055	intg	—	13.0	3.3	2.8	0.6	9.6	0.6
3	Protein	6	87155863	rs383429876	intr	<i>CSN1S1</i>	15.4	20.5	1.4	6.9	4.5	0.4
			87296809	rs134776019	intg	—	13.2	18.1	7.5	6.0	8.7	2.5

¹rs-id = reference SNP ID, <http://www.ncbi.nlm.nih.gov/projects/SNP/>.

²Ann = functional annotation: intg = intergenic variant; intr = intron.

³HDK = Danish Holstein; HFR = French Holstein; JER = Jersey; MON = Montbéliarde; NOR = Normande; RDC = Danish Red.

populations to numbers comparable to the other breeds would have been possible to make the analysis more balanced and to reduce the influence of Holstein on the ALL analysis. However, in practice, an important objective of joining data of multiple breeds is to enable smaller breeds to take advantage of the large data sets available for some breeds. Therefore, we decided to use all Holstein data and then run a separate analysis that completely excluded Holstein to detect QTL that were overshadowed by Holstein.

Besides a reduction in power, a reduction in the number of QTL variants could also reflect a reduction in the number of false positives. With millions of variants to analyze, independently testing them in a GWAS inevitably results in false positives. Furthermore, due to LD, multiple variants may be associated with the same QTL, and several nearby regions could be associated with the same QTL rather than representing different QTL. We tried to minimize the number of

false positives by correcting the P -values using the false discovery rate, and reduced the number of regions per QTL by excluding new QTL within 5 Mb of a border of a previously defined QTL. The combination of high power and long-range LD within Holstein could, however, still result in the selection of multiple variants associated with the same QTL. The number of QTL reported by the multibreed GWAS is therefore likely to overestimate the true number of QTL. Contrarily, the high thresholds required for Holstein were too strict for the smaller breeds, and hardly any QTL were detected in these breeds. Consequently, the number of QTL regions detected in the smaller breeds is an underestimation of the true number of QTL. Our aim was not to detect all QTL segregating in all breeds in the study, but to compare the mapping power and precision of the multibreed GWAS. Comparing the power of different analyses required use of the same thresholds for all studies, but because of the large differences in sample

Table 4. Multibreed $-\log_{10}(P)$ for most significant variants in the regions of examples 1 to 3

Example	BTA	Position	Multi ¹			z-Scores ²			Fixed ³			Random ⁴		
			HOL	REST	ALL	HOL	REST	ALL	HOL	REST	ALL	HOL	REST	ALL
1	2	105783655	12.2	0.8	2.1	12.1	0.6	7.2	11.0	1.9	10.8	0.5	0.6	0.6
		105832735	1.9	12.7	11.2	2.1	13.1	10.2	1.3	6.2	5.9	0.4	1.0	1.1
		105864893	7.2	8.0	14.2	9.1	8.7	16.6	3.8	2.6	5.9	0.5	0.7	0.9
2	5	93305237	7.7	9.2	0.4	6.4	10.4	0.1	8.2	5.3	1.6	8.2	1.9	0.9
		93945738	62.6	8.7	71.6	65.6	9.3	70.4	56.8	9.4	65.4	1.7	4.9	4.8
		111662916	14.0	3.6	17.2	15.6	3.7	18.0	15.9	2.5	17.8	5.3	0.7	2.5
		112374194	1.9	17.3	12.5	2.2	18.1	12.7	3.4	12.4	13.3	0.9	2.1	2.8
		112397190	12.7	7.0	18.9	13.6	8.8	21.6	15.7	5.8	20.7	15.7	1.6	3.8
3	6	87155863	34.0	6.9	38.5	34.8	6.2	38.1	33.4	3.3	33.5	4.7	0.8	2.6
		87296809	28.8	22.0	49.9	30.3	21.9	50.9	28.3	22.0	49.2	3.4	15.4	15.5

¹Multi = multibreed genome-wide association study (GWAS); HOL = multibreed GWAS combining Danish and French Holstein; REST = multibreed GWAS combining Jersey, Montbéliarde, Normande, and Danish Red; ALL = multibreed GWAS combining all 6 populations.

²Weighted z-scores model.

³Fixed effects model.

⁴Random effects model.

size, it was not possible to use a threshold appropriate for all breeds. Therefore, the number of QTL regions reported here should be interpreted cautiously and not seen as an estimate of the true number of QTL. Similarly, the table with overlapping numbers of QTL does not necessarily reflect the true number of QTL

segregating across breeds. The table only illustrates the number of QTL regions detected in one breed that overlap with a region detected in another breed. Appropriate thresholds to truly estimate the portion of QTL shared across breeds would have to be lower for the breeds with the smaller data sizes.

Figure 6. Example 3, within-breed (top and middle) and multibreed genome-wide association study (GWAS; bottom) for protein yield. Chromosome 11, between 50 and 70 Mb. HOLDK = Danish Holstein, JER = Jersey, RDC = Danish Red, HOLFR = French Holstein, NOR = Normande, MON = Montbéliarde, HOL = Danish and French Holstein, REST = all breeds except Holstein, ALL = all breeds.

Smaller QTL regions with fewer variants per region indicate the QTL mapping is more precise. Because LD is conserved over long distances within breeds and short distances across breeds (de Roos et al., 2008), a single marker model tends to result in many variants associated with the same QTL over relatively long distances within breeds but fewer variants in smaller regions across breeds. Because of the faster decay in LD across breeds, we expected by joining different breeds, the size of the QTL regions would be reduced. This outcome was true for REST, but HOL and ALL did not on average result in smaller regions than within Danish Holstein and French Holstein. The absence of a decrease in QTL size in HOL is not surprising because HOL was not a true multibreed analysis. ALL did result in smaller regions than HOL, indicating that adding the other breeds to the Holstein analysis did improve mapping precision. Similar to the size of QTL regions, the number of SNP per region was lower for REST than within breed and lower for ALL than for HOL. Therefore, our results show that adding data of different breeds reduces QTL intervals, in agreement with results from a study by Raven et al. (2014).

Averaged across the genome, the P -values of the peaks detected in the multibreed GWAS were similar to those of peaks detected within breed. Our examples show, however, that P -values of certain peaks decrease substantially. For example, the $-\log_{10}(P)$ of the peak detected within Holstein described in example 2 approximately doubled by joining the 2 Holstein populations. This result was not the case for all detected QTL regions, however. For some regions, a QTL may have different effects in different breeds, and joining these breeds may result in a decreased peak. This finding could explain why on average, genome wide, no increase in the $-\log_{10}(P)$ of the peak was observed. The 2 Holstein populations were highly related though, and most peaks would be expected to increase, but this pattern was not always the case. At the peaks in the examples, the $-\log_{10}(P)$ did increase. A possible explanation could be that when the 2 Holstein populations were combined, the P -value of many small QTL decreased below the threshold, reducing the average $-\log_{10}(P)$.

Defining QTL Regions

Different thresholds and strategies to define QTL regions were tested (results not shown), and although the absolute numbers of number of QTL and QTL size were influenced by the threshold and strategy chosen, the overall conclusion was the same: joining data increases power when closely related populations are combined, while mapping precision is improved when more distinct populations are added. Nevertheless, our

approach to defining QTL regions relied on arbitrarily chosen thresholds. A more objective way of defining QTL would be desirable. In practice, however, GWAS and QTL detection studies always rely on some kind of detection threshold. In theory, a simulation study could be useful to choose a threshold. The appropriate detection threshold for a trait is, however, likely to depend on genetic architecture underlying the trait, and so the number of QTL influencing a trait would have to be known to be able to design a simulation study. Some studies have attempted to estimate the number of QTL for production traits in dairy cattle (Hayes and Goddard, 2001; Chamberlain et al., 2007), but these estimates rely on arbitrarily chosen QTL detection thresholds. Therefore, in the absence of an objective approach to define QTL regions, we had to use arbitrarily chosen thresholds.

Compromises Made to Reduce the Computation Time

Because the model had to be run almost 442 million times (the number of variants $\times$ 6 populations $\times$ 3 traits), compromises were made in the analyses to reduce the computation time. To simplify the model, we used a sire effect to account for family structure rather than a genomic relationship matrix or full pedigree. Most of the family structure can be accounted for by a sire effect in dairy cattle, but using only a sire effect may have resulted in spurious associations. Fitting a relationship matrix would, however, have resulted in a major increase in computational time because of the large number of individuals in the data set (Zhang et al. 2010).

A second compromise was made by preselecting the variants used in the multibreed GWAS based on their P -values in the within-breed GWAS. An important advantage of a joint GWAS is that variants with a P -value above a detection threshold in a single GWAS could have a P -value below the threshold in a joint GWAS. By preselecting variants rather than using all variants for the multibreed GWAS, we may have underestimated the advantage of a multibreed GWAS. The thresholds used for preselection were, however, lenient, and having a substantial number of variants with P -values above these thresholds in all breeds and P -values below the stringent detection threshold used in the multibreed GWAS would have been unlikely.

Comparison of Multibreed GWAS and Meta-Analyses

Three different meta-analysis methods were compared. The weighted z-scores model used P -values and directions as well as weighting based on the sample size

of individual studies, while the fixed and random effects model used the estimated effects and standard errors. The random effects differed from the fixed effects model by computing Cochran's statistic to account for heterogeneity. The weighted z-score model gave the most similar results to the multibreed GWAS. Since this model uses the *P*-values and the other models the effects, the difference in scaling of DRP between the French and Danish breeds likely explains why the z-score model performed most similarly to the multibreed GWAS. The random effects model gave the most different results. A large part of variants with a high $-\log_{10}(P)$ in the multibreed GWAS had a much lower $-\log_{10}(P)$ in the random effects model. For these variants, heterogeneity detected by the random effects model was highly significant. The random effects model is underpowered when a small number of studies are compared (Begum et al., 2012), and combining 2 to 6 breeds might not be enough to accurately estimate heterogeneity, which may indicate the random effects model is inappropriate with such a low number of studies.

The differences between results of the meta-analyses and those of the multibreed GWAS were larger when results of different breeds were combined than when the 2 Holstein populations were combined. The more the combined breeds differ, the smaller the expected proportion of QTL that are segregating across them. Furthermore, QTL segregating across breeds can still have a different effect in different breeds, for example, due to epistasis. This circumstance makes it more difficult to estimate an across-breed effect and can explain some of the differences between the meta-analyses and the multibreed GWAS.

CONCLUSIONS

We have shown that a multibreed GWAS improved detection power when closely related populations that share many QTL were combined, whereas mapping precision improved when more distinct populations were joined. Furthermore, some QTL were detected in the multibreed GWAS that were not detected within the smaller breeds. The multibreed GWAS was, however, dominated by Holstein because of the much larger data size compared with the other breeds. Therefore, some peaks segregating in the smaller breeds were overshadowed by a nearby larger Holstein peak. Therefore, performing the multibreed GWAS both with and without Holstein was valuable. Alternatively, a meta-analysis can be used to approximate a multibreed GWAS. Out of 3 different meta-analysis methods, the weighted z-scores model gave the most similar results to the true multibreed analysis. Our results show that a meta-anal-

ysis can be used as an alternative for a full multibreed GWAS. Differences between the meta-analyses and the multibreed GWAS were larger when less related breeds were combined. A full multibreed GWAS is therefore preferable if more distinct breeds are combined.

ACKNOWLEDGMENTS

The first author benefited from an Erasmus-Mundus fellowship and a grant from ApisGene, within the framework of the European Graduate School in Animal Breeding and Genetics, Paris, France. French sequence data originated from the Cartoseq project funded by ANR and ApisGene (ANR10-GENM-0018), Paris, France. This research was supported by the center for Genomic Selection in Animals and Plants (GenSAP) funded by The Danish Council for Strategic Research). We acknowledge the 1,000 bull genomes project (<http://www.1000bullgenomes.com>) for sharing data to impute the genome sequence.

REFERENCES

- Begum, F., D. Ghosh, G. C. Tseng, and E. Feingold. 2012. Comprehensive literature review and statistical considerations for GWAS meta-analysis. *Nucleic Acids Res.* 40:3777–3784. <http://dx.doi.org/10.1093/nar/gkr1255>.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J.R. Stat. Soc.* 57:289–300. <http://dx.doi.org/10.2307/2346101>.
- Browning, S. R., and B. L. Browning. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81:1084–1097. <http://dx.doi.org/10.1086/521987>.
- Chamberlain, A. J., H. C. McPartlan, and M. E. Goddard. 2007. The number of loci that affect milk production traits in dairy cattle. *Genetics* 177:1117–1123.
- Daetwyler, H. D., A. Capitan, H. Pausch, P. Stothard, R. van Binsbergen, R. F. Brøndum, X. Liao, A. Djari, S. C. Rodriguez, C. Grohs, D. Esquerré, O. Bouchez, M.-N. Rossignol, C. Klopp, D. Rocha, S. Fritz, A. Eggen, P. J. Bowman, D. Coote, A. J. Chamberlain, C. Anderson, C. P. VanTassell, I. Hulsege, M. E. Goddard, B. Guldbrandtsen, M. S. Lund, R. F. Veerkamp, D. A. Boichard, R. Fries, and B. J. Hayes. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat. Genet.* 46:858–865. <http://dx.doi.org/10.1038/ng.3034>.
- de los Campos, G., A. I. Vazquez, R. Fernando, Y. C. Klimentidis, and D. Sorensen. 2013. Prediction of complex human traits using the genomic best linear unbiased predictor. *PLoS Genet.* 9:e1003608. <http://dx.doi.org/10.1371/journal.pgen.1003608>.
- de Roos, A. P. W., B. J. Hayes, R. J. Spelman, and M. E. Goddard. 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. *Genetics* 179:1503–1512. <http://dx.doi.org/10.1534/genetics.107.084301>.
- Erbe, M., B. J. Hayes, L. K. Matukumalli, S. Goswami, P. J. Bowman, C. M. Reich, B. A. Mason, and M. E. Goddard. 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *J. Dairy Sci.* 95:4114–4129. <http://dx.doi.org/10.3168/jds.2011-5019>.
- Evangelou, E., and J. P. A. Ioannidis. 2013. Meta-analysis methods for genome-wide association studies and beyond. *Nat. Rev. Genet.* 14:379–389. <http://dx.doi.org/10.1038/nrg3472>.

- Hayes, B., and M. E. Goddard. 2001. The distribution of the effects of genes affecting quantitative traits in livestock. *Genet. Sel. Evol.* 33:209–229.
- Hayes, B. J., P. J. Bowman, A. C. Chamberlain, K. Verbyla, and M. E. Goddard. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* 41:51 <http://dx.doi.org/10.1186/1297-9686-41-51>.
- Höglund, J. K., B. Guldbrandtsen, M. S. Lund, and G. Sahana. 2015. Identification of genomic regions associated with female fertility in Danish Jersey using whole genome sequence data. *BMC Genet.* 16:60 <http://dx.doi.org/10.1186/s12863-015-0210-3>.
- Howe, B., C. Fuchsberger, M. Stephens, J. Marchini, and G. R. Abecasis. 2012. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* 44:955–959. <http://dx.doi.org/10.1038/ng.2354>.
- Hozé, C., M.-N. Fouilloux, E. Venot, F. Guillaume, R. Dassonneville, S. Fritz, V. Ducrocq, F. Phocas, D. Boichard, and P. Croiseau. 2013. High-density marker imputation accuracy in sixteen French cattle breeds. *Genet. Sel. Evol.* 45:33 <http://dx.doi.org/10.1186/1297-9686-45-33>.
- Lin, D. Y., and D. Zeng. 2010. Meta-analysis of genome-wide association studies: No efficiency gain in using individual participant data. *Genet. Epidemiol.* 34:60–66. <http://dx.doi.org/10.1002/gepi.20435>.
- Liu, J. Z., F. Tozzi, D. M. Waterworth, S. G. Pillai, P. Muglia, L. Middleton, W. Berrettini, C. W. Knouff, X. Yuan, G. Waeber, P. Vollenweider, M. Preisig, N. J. Wareham, J. H. Zhao, R. J. F. Loos, I. Barroso, K.-T. Khaw, S. Grundy, P. Barter, R. Mahley, A. Kesaniemi, R. McPherson, J. B. Vincent, J. Strauss, J. L. Kennedy, A. Farmer, P. McGuffin, R. Day, K. Matthews, P. Bakke, A. Gulsvik, S. Lucae, M. Ising, T. Brueckl, S. Horstmann, H.-E. Wichmann, R. Rawal, N. Dahmen, C. Lamina, O. Polasek, L. Zgaga, J. Huffman, S. Campbell, J. Kooner, J. C. Chambers, M. S. Burnett, J. M. Devaney, A. D. Pichard, K. M. Kent, L. Satler, J. M. Lindsay, R. Waksman, S. Epstein, J. F. Wilson, S. H. Wild, H. Campbell, V. Vitart, M. P. Reilly, M. Li, L. Qu, R. Wilensky, W. Matthai, H. H. Hakonarson, D. J. Rader, A. Franke, M. Wittig, A. Schäfer, M. Uda, A. Terracciano, X. Xiao, F. Busonero, P. Scheet, D. Schlessinger, D. St Clair, D. Rujescu, G. R. Abecasis, H. J. Grabe, A. Teumer, H. Völzke, A. Petersmann, U. John, I. Rudan, C. Hayward, A. F. Wright, I. Kolcic, B. J. Wright, J. R. Thompson, A. J. Balmforth, A. S. Hall, N. J. Samani, C. A. Anderson, T. Ahmad, C. G. Mathew, M. Parkes, J. Satsangi, M. Caulfield, P. B. Munroe, and M. Farrall., et al. 2010. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat. Genet.* 42:436–440. <http://dx.doi.org/10.1038/ng.572>.
- Lund, M. S., G. Su, L. Janss, B. Guldbrandtsen, and R. F. Brøndum. 2014. Invited review: Genomic evaluation of cattle in a multi-breed context. *Livest. Sci.* 166:101–110. <http://dx.doi.org/10.1016/j.livsci.2014.05.008>.
- R Development Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org/>.
- Raven, L.-A., B. G. Cocks, and B. J. Hayes. 2014. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics* 15:62 <http://dx.doi.org/10.1186/1471-2164-15-62>.
- van den Berg, I., B. Guldbrandtsen, C. Hozé, R. F. Brøndum, D. A. Boichard, and M. S. Lund. 2014. Across breed QTL detection and genomic prediction in French and Danish dairy cattle breeds. Page 490 in *Proc. of the 10th World Congr. Genet. Appl. Livest. Prod.*, Vancouver, BC, Canada.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <http://dx.doi.org/10.3168/jds.2007-0980>.
- Willer, C. J., Y. Li, and G. R. Abecasis. 2010. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 26:2190–2191. <http://dx.doi.org/10.1093/bioinformatics/btq340>.
- Zhang, Z., E. Ersoz, C.-Q. Lai, R. J. Todhunter, H. K. Tiwari, M. A. Gore, P. J. Bradbury, J. Yi, D. K. Arnett, J. M. Ordoñez, and E. S. Buckler. 2010. Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42:355–360. <http://dx.doi.org/10.1038/ng.546>.