ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/j.1365-2052.2010.02064.x

Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle

G. Sahana, B. Guldbrandtsen, C. Bendixen and M. S. Lund

Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, Research Centre Foulum, Aarhus University, DK-8830, Tjele, Denmark

Summary

A genome-wide association study was conducted using a mixed model analysis for QTL for fertility traits in Danish and Swedish Holstein cattle. The analysis incorporated 2,531 progeny tested bulls, and a total of 36 387 SNP markers on 29 bovine autosomes were used. Eleven fertility traits were analyzed for SNP association. Furthermore, mixed model analysis was used for association analyses where a polygenic effect was fitted as a random effect, and genotypes at single SNPs were successively included as a fixed effect in the model. The Bonferroni correction for multiple testing was applied to adjust the significance threshold. Seventy-four SNP-trait combinations showed chromosome-wide significance, and five of these were significant genome-wide. Twenty-four QTL regions on 14 chromosomes were detected. Strong evidence for the presence of QTL that affect fertility traits were observed on chromosomes 3, 5, 10, 13, 19, 20, and 24. The QTL intervals were generally smaller than those described in earlier linkage studies. The identification of fertility trait-associated SNPs and mapping of the corresponding QTL in small chromosomal regions reported here will facilitate searches for candidate genes and candidate polymorphisms.

Keywords *Bos taurus*, fertility, genome-wide association, quantitative trait loci.

Introduction

Improving female fertility in cattle is increasingly important given observed declines in fertility, especially among high yielding cows. Low fertility can undermine the economic viability and competitiveness of dairy herds due to the increased costs of additional insemination, veterinary expenses, and replacement costs. Unfortunately, the low heritability of fertility-related traits makes genetic improvement by traditional animal breeding slow. The mean heritability for fertility traits among estimates given in 17 studies ranged from 0.017, for the interval from first to last insemination, to 0.05, for the interval from calving to first service (Pryce & Veerkamp 2001; Veerkamp & Beerda 2007). Further, the genetic correlations between fertility and milk yield traits are generally unfavourable (Pryce *et al.* 1997; Dematawewa & Berger 1998; Roxström *et al.* 2001), as evidenced by the fertility decline observed in recent decades in Holstein cattle populations, which are subject to intense selection for milk production traits (Sørensen *et al.* 2007). Unfortunately, a pronounced focus on female fertility in the breeding goals of Nordic countries has not been sufficient to offset this trend (Pedersen *et al.* 2008).

Low heritability, negative correlation with milk production, sex-limited expression, and late recording of phenotypes in life make female fertility a prime candidate for genetic improvement through marker-assisted selection. Selection efficiency could be improved by identification of QTL for fertility traits. Once suitable QTL have been mapped, the information can be used in several ways to identify causal genetic factors underlying the QTL, such as assessment of disease risk, prediction of quantitative traits, and selection of candidate genes.

The QTL findings for fertility so far do not overlap very well (Höglund *et al.* 2009). One of the reasons for this is that the traits used in many cases have been different. It should be noted that most of the studies reported to date have used linkage analysis for QTL mapping, and the QTL thus identified can typically only be utilized for within-family selection. An alternative approach to gene mapping, termed association or linkage disequilibrium (LD) mapping, is based on associations between a marker or marker haplotype and a causative gene at the population level. The LD mapping

Address for correspondence

G. Sahana, Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, Research Centre Foulum, Aarhus University, P.O. Box 50, 8830 Tjele, Denmark.
E-mail: goutam.sahana@agrsci.dk

Accepted for publication 16 March 2010

approach derives from the fact that because a disease-associated mutation arises in a particular haplotype background, individuals who inherit the mutation will also tend to inherit the same alleles at nearby marker loci. Because population genealogy is deeper than a family pedigree, LD mapping permits much finer-scale mapping than linkage analysis (Hästbacka *et al.* 1992). With the development of sequencing and high-throughput genotyping techniques, the number of SNP markers available for association studies has dramatically increased. A genome-wide association study (GWAS) enables a systematic search of an organism's entire genome for genetic factors underlying any inherited trait of interest. Thus, association studies have emerged as a powerful tool for revealing the genetic basis of diseases and quantitative traits (Devlin & Risch 1995; Collins *et al.* 1998) and have been used successfully to identify susceptibility loci for common diseases in humans (The Wellcome Trust Case Control Consortium [WTCCC] 2007).

Association mapping is based on direct marker effects and utilizes population-wide LD of markers with causative genes. Results derived from association mapping of fertility traits would be of immediate use in breeding programs, because allelic variation present in the entire population can be investigated. In practice, GWAS results are used to prioritize chromosomal regions for further study or additional typing. In genomic selection in animals (Meuwissen *et al.* 2001; Schaeffer 2006), it is then possible to establish different priority weights for individual SNPs based on their established association with the phenotype of interest.

In the present study, association mapping was carried out for 11 female fertility traits in cattle using SNP markers in a GWAS. The objective was to identify QTL for female fertility traits in Danish and Swedish Holstein cattle populations.

Materials and methods

Animals and phenotypic data

A total of 2531 progeny tested Danish and Swedish Holstein bulls were genotyped. Out of the genotyped bulls, 2233 had phenotypic records available for fertility traits analyzed in this study, except for fertility treatments, where 2049 bulls had phenotypes.

Eleven fertility-related traits were analyzed for association with SNPs. Out of these traits, ten (all except the fertility index) were previously analyzed by Höglund *et al.* (2009) using regression-based linkage analysis. For details of the recorded phenotypes and models used in the breeding value prediction, see http://www.nordicebv.info, Ancker *et al.* (2006) and Höglund *et al.* (2009). Briefly, single-trait breeding values (STBVs) were calculated for each animal using the Nordic Cattle Genetic Evaluation criteria based on first to third parity data in cows, using the BLUP and sire models. The STBVs were adjusted for the same systematic environmental effects as in the official routine evaluations. The STBVs are available from national evaluations (separately for cows and heifers) for the number of inseminations (AIS), 56 day non-return rate (NRR), and interval from first to last insemination (IFL). Here, we use the suffixes C and H with trait abbreviations to denote cow and heifer traits respectively, where relevant. The interval from calving to first insemination (ICF) is only defined for cows. For veterinary treatments of reproductive disorders, STBVs were predicted separately for first, second, and third lactations (FRT1, FRT2, and FRT3). The details of the diseases included in the fertility treatments are presented in Höglund *et al.* (2009).

The fertility index (FTI) is a compound index describing how quickly and easily cows become pregnant. The FTI for female fertility is a joint Nordic index, based on insemination data from Denmark, Sweden, and Finland. Calculations were based on insemination data dating from 1985 (Denmark), 1982 (Sweden), and 1994 (Finland) to the present. The traits included in the FTI calculation are: AIS and IFL for cows and heifers, and ICF for cows. For details on the FTI, see Pedersen *et al.* (2008).

SNP chip and genotyping

We used bovineSNP50 beadchip (Illumina Inc.) for genotyping 2531 animals. The bovineSNP50 method assays 54 001 markers with a median interval of 37 kb between SNPs (Matukumalli *et al.* 2009). For these experiments, genomic DNA was extracted from blood samples, but if not available, semen samples were used for DNA extraction. The platform used was an Illumina® Infinium II Multisample assay device (Illumina Inc.). SNP chips were scanned using iScan (Illumina Inc.) and analyzed using BEADSTUDIO version 3.1 software (Illumina Inc.). The quality parameters used for selection of SNPs in the GWAS were minimum call rates of 85% for individuals and of 95% for loci. Marker loci with minor allele frequencies (MAFs) below 5% were excluded. The minimal acceptable GC score (Teo *et al.* 2007) was 0.60 for individual typings. Individuals with average GC scores below 0.65 were excluded. The SNP positions within a chromosome were based on the *Bos taurus* genome assembly (*Btau*_4.0, Liu *et al.* 2009).

Statistical methods for association analysis

Mixed model: Mixed model analysis, as proposed by Yu *et al.* (2006), was used for association analyses. In this approach, a polygenic genetic effect is fitted as a random effect, and genotypes at single SNPs are successively included as fixed effects. The model was as below:

$$y = \mu + S\alpha + Zu + e$$

where **y** is a vector of observed phenotypes, μ is the general mean, **Z** is a matrix relating additive polygenic effects to

doi: 10.1111/j.1365-2052.2010.02064.x

individuals, **u** is a vector of additive polygenic effects, α is a vector of SNP effects, **S** is an incidence matrix relating α to the individuals, and **e** is a vector of random residual effects. The random variables **u** and **e** are assumed to be normally distributed. Specifically, **u** is normally distributed with (**0**, σ_g^2**A**), where σ_g^2 is the polygenic genetic variance, and **A** is the additive relationship matrix derived from the pedigree. The analyses were carried out using the DMU software package (available at http://dmu.agrsci.dk).

Significance test: The Bonferroni correction was applied to control the family-wise error rate (FWER) when testing multiple hypotheses. If an experiment tests **n** independent hypotheses on a set of data, then one way of maintaining the FWER is to test each individual hypothesis at a statistical significance level of 1/**n** times what it would be if only one hypothesis was tested. The Bonferroni adjusted controls FWER (α) = 1-(1-α_i)^m $\approx$ α_im, where α_i is the individual test rejection level, and **m** is the number of tests. Bonferroni correction ignores correlation between makers and leads to an overly conservative correction, which is exacerbated as the marker density increases (Han *et al.* 2009). This results in high false negatives if a genome-wide Bonferroni corrected threshold is set, especially when the power to identify QTL is not high. Therefore, we tested associations at the chromosome-wide significance level. The 5% chromosome-wide significance thresholds ranged from the point-wise *P*-value of 2.16×10^{-5} on chromosome 1 to 7.41×10^{-5} on chromosome 28, or 4.67 and 4.13 in the $-\log_{10}$ transformed scale respectively. Because we used the chromosome-wide significance level as the criteria to call significant associations, a significant association indicates chromosome-wide significance, and a suggestive association corresponds to a *P*-value less than 10^{-4}.

Marking the QTL region

Normally, multiple SNPs in the vicinity of a QTL are expected to yield significant results in a single SNP analysis. This is because sets of SNPs that are physically located near the causal factor will tend to be in linkage disequilibrium; this effect declines with genetic distance and also depends on minor allele frequencies. Considering this, the boundaries of a QTL region can be marked by looking for markers flanking a SNP hit where the –log(p) value returns to background noise levels (WTCCC 2007). In this study, QTL regions were demarcated subjectively. Starting at the most significant SNP, the QTL region was extended left and right until a region was reached where all markers had –log(p) values below 3. That is the QTL thus demarcated may contain one or more non-significant markers. To compare results from the present study with the earlier one, we took the maker position from Btau_4.0. If the maker location was not available in Btau_4.0, we have reported marker and given the position in cM from MARC table (http://www.marc.usda.gov/genome/cattle/cattle.html).

Results and discussion

A total of 36 387 SNPs on 29 bovine autosomes (BTAs) were selected for association analyses. These SNPs passed the quality control criteria defined in the Materials and methods and had MAFs equal to or higher than 0.05. The number of SNPs included for analysis varied from 675 on BTA28 to 2320 on BTA1.

The genome-wide scan revealed significant association between SNPs and FTI, ICF, AISC, IFLC, FERT1, FERT2, and FERT3 (Table 1, Fig. 1). For the other phenotypes analyzed, the associations did not cross the significance threshold, though there was suggestive evidence for association, especially with SNPs in the QTL regions identified for the previously mentioned traits. A total of 74 SNP-trait combinations on 18 chromosomes crossed the chromosome-wide significance threshold. Out of these, a large number were for FTI (23) and ICF (22), followed by AISC (12) and IFLC (10). Two SNPs on BTA5, and one each on BTA13 and BTA19, showed genome-wide significant association with fertility traits. Table 2 presents 24 QTL regions detected in the present study. A chromosomal region was defined as a QTL only when multiple SNPs in that region showed significant and/or suggestive association with one or more traits. If a SNP displayed chromosome-wide or genome-wide significance (Table 1), but no other SNP showing suggestive evidence of association with the same trait or other traits was found nearby, that SNP was not included when defining QTL regions. Single, isolated, significant markers were excluded from QTL definition because they have a high risk of representing a false positive.

The association signal plots shown in Fig. 1 present an overview of the location of the associated SNPs across the genome and provide a visualization of whether the QTL are affecting more than one trait. Signal plots are only shown for those traits which had chromosome-wide SNP association. In many cases, QTL at the same chromosomal location were observed for several fertility traits. For example, QTL for AISC, FTI, and IFLC were detected within the same region on BTA5. Other chromosomes with QTL associated with several fertility traits were BTA10, 13, 20, and 24. Strong association signals in the same chromosomal region for more than one correlated trait give confidence about the presence of a true QTL.

QTL regions

On BTA1, four SNPs showed chromosome-wide significant association with FTI (72.1 and 92.4 Mb), ICF (136.5 Mb), and IFLC (148.8 Mb) (Table 1). The most significant *P*-value was observed at 148.8 cM for IFLC with SNP ss61522035. This QTL region extended from 148.3 to 154.2 Mb, and it showed association with IFLC, FTI, and AISC (Table 2). There was another QTL region close by (ranging from 135.2 to 142.3 Mb); it had effects on ICF,

doi: 10.1111/j.1365-2052.2010.02064.x

Table 1 SNPs showing significant chromosome-wide association with fertility-related traits.

Chromosome	SNP	Trait*	Position (bp)	Allele substitution effect	S.E.	*P*-value
1	ss61480552	FTI	72076618	0.7365	0.1636	1.64E-05
1	rs29019866	FTI	92433977	−1.8709	0.4125	1.42E-05
1	ss86338186	ICF	136499200	0.7182	0.1584	1.42E-05
1	ss61522035	IFLC	148765346	−0.6072	0.1301	7.80E-06
3	ss86299093	FTI	81303069	−2.0262	0.4151	2.81E-06
3	ss86299093	IFLC	81303069	−1.6611	0.3482	4.78E-06
3	ss86309132	FTI	81446820	−0.691	0.1563	2.37E-05
3	ss86287773	ICF	87102164	0.6453	0.1464	2.48E-05
4	ss61489726	FERT3	36814032	0.5415	0.1119	3.43E-06
4	ss61489996	AISC	63995740	−0.7157	0.1513	5.78E-06
4	rs29020693	IFLC	90418033	−0.5984	0.131	1.22E-05
5	ss86274830	FTI	116338745	−1.3087	0.2958	2.33E-05
5	**ss86274830**	**AISC**	**116338745**	−1.3912	**0.2651**	**4.48E-07**
5	ss86274830	IFLC	116338745	−1.1849	0.248	4.60E-06
5	ss86299524	FTI	116362061	−0.6773	0.1497	1.48E-05
5	**ss86299524**	**AISC**	**116362061**	−0.764	**0.1338**	**3.69E-08**
5	**ss86299524**	**IFLC**	**116362061**	−0.6441	**0.1252**	**7.62E-07**
6	ss86337596	ICF	89774923	−0.5911	0.1302	1.39E-05
7	ss117968986	IFLC	90305805	−0.8586	0.1934	2.18E-05
9	rs42466129	AISC	51661514	0.6513	0.1446	1.64E-05
9	ss117964048	FTI	67688020	−0.6985	0.16	2.98E-05
9	rs43732223	FTI	83133779	0.756	0.1623	8.11E-06
9	rs43732223	FERT1	83133779	0.5374	0.1187	1.47E-05
10	ss61524214	AISC	40769608	−2.2287	0.5017	2.14E-05
10	rs43087086	AISC	52660587	−0.7052	0.1521	8.90E-06
10	ss61556155	FTI	69177539	0.8255	0.189	2.95E-05
10	ss86306178	FTI	69199384	−0.7739	0.1755	2.48E-05
10	ss105241610	FTI	69247532	0.9484	0.1896	1.55E-06
10	rs42227324	FTI	70287191	0.9317	0.1975	6.12E-06
10	ss61559544	AISC	93016335	0.965	0.2073	8.16E-06
12	ss86339393	FERT2	81888393	−0.3888	0.0902	3.83E-05
13	ss86273520	ICF	28502960	−0.7351	0.1608	1.20E-05
13	ss61523084	ICF	29012170	−0.7166	0.1603	1.88E-05
13	ss61533466	ICF	29057593	−0.7914	0.1651	4.29E-06
13	ss86285310	FTI	30172946	−1.448	0.3121	8.83E-06
13	ss61533497	ICF	30386996	0.5873	0.1325	2.23E-05
13	ss117964280	ICF	32984105	−0.6254	0.1368	1.19E-05
13	ss86341075	ICF	33029140	0.6367	0.1367	8.10E-06
13	ss105239828	FTI	33117305	−0.6904	0.1542	1.83E-05
13	ss105239828	ICF	33117305	−0.641	0.1338	4.34E-06
13	ss86288462	FTI	33517748	−0.7145	0.1576	1.42E-05
13	**ss86288462**	**ICF**	**33517748**	−0.7372	**0.1365**	**2.01E-07**
13	rs29020401	ICF	33741174	0.7306	0.1552	6.43E-06
13	ss86336415	FERT3	37854558	−0.4964	0.1108	1.83E-05
13	ss61474875	ICF	44121804	1.3439	0.2911	9.82E-06
13	rs42517187	ICF	50071315	−0.6101	0.1341	1.33E-05
13	ss38325211	ICF	50414985	0.5901	0.1339	2.50E-05
13	ss105238449	ICF	58056235	−1.4052	0.2982	6.31E-06
13	ss86273802	ICF	59366358	−0.6026	0.1387	3.28E-05
13	ss86288836	FTI	60775960	0.6541	0.1492	2.75E-05
13	ss86288836	ICF	60775960	0.6384	0.1287	1.93E-06
13	ss86336399	FTI	68060079	0.7497	0.1568	4.58E-06
19	**rs43721195**	**FERT3**	**41676084**	−1.5952	**0.3123**	**9.18E-07**
19	ss86336183	FTI	45775576	1.7088	0.3827	1.94E-05
19	ss86336183	IFLC	45775576	1.4877	0.3248	1.15E-05

doi: 10.1111/j.1365-2052.2010.02064.x

Table 1 Continued.

Chromosome	SNP	Trait*	Position (bp)	Allele substitution effect	S.E.	*P*-value
19	ss86332973	FTI	45797094	−1.3543	0.3018	1.76E-05
19	ss86330159	ICF	54108418	−0.8511	0.1947	2.91E-05
20	ss61543090	FTI	26416375	1.4581	0.3209	1.36E-05
20	ss61543090	AISC	26416375	1.3132	0.2924	1.72E-05
20	ss61543090	IFLC	26416375	1.2441	0.2733	1.31E-05
20	ss61543090	FERT3	26416375	0.7364	0.1665	2.32E-05
20	rs42289232	ICF	45344421	1.1453	0.2652	3.66E-05
21	rs29017183	ICF	15643297	−0.607	0.1399	3.35E-05
22	ss61545188	ICF	24997615	0.6305	0.1361	9.13E-06
23	ss105257856	FERT2	1694991	−0.688	0.163	5.55E-05
23	ss86308584	FERT2	28890489	−0.7603	0.1753	3.40E-05
24	ss86302405	FTI	29730649	−0.6367	0.1492	4.57E-05
24	ss61486559	FTI	36746469	1.4573	0.3338	2.99E-05
24	ss86301009	AISC	48680176	−0.6484	0.1379	6.62E-06
24	rs29019756	AISC	48700454	1.2431	0.2772	1.78E-05
26	ss86336536	AISC	19918015	−0.5914	0.1403	5.68E-05
26	rs43118660	AISC	26874836	−1.7798	0.4138	3.96E-05
29	ss86341011	FTI	27084192	−1.065	0.244	3.01E-05
29	ss86341011	IFLC	27084192	−0.9263	0.2057	1.63E-05

Genome-wide significant associations are in bold.

*FTI, fertility index; AISC, number of inseminations per conception (or culling) (cow); IFLC, days from first to last insemination (cow); and ICF, interval from calving to first insemination. FTI, FT2, and FRT3 indicate veterinary treatments of reproductive disorders in the first, second, and third lactation respectively.

IFLC, FTI, and AISC. The third QTL region on BTA1 was located from 85.3 to 92.4 Mb with evidence of association with FTI and IFLC.

Using the linkage analysis method, Höglund *et al.* (2009) reported QTL segregating between *BMS918* (132.5 cM) and *BMS4043* (142.2 cM) for fertility treatment and ICF, in the same cattle population as in the present study. Additionally, Schulman *et al.* (2008) reported QTL for days open and fertility treatments between markers *BMS599* (125.8 cM) and *BMS4014* (135.5 cM) in Finnish Ayrshire cattle. Considering our data, we found SNPs with significant and suggestive association with ICF, IFLC, FTI, and AISC in these regions. The results also indicate that this region may harbour QTL affecting FERT1 and FERT3. Ben-Jemaa *et al.* (2008) reported a QTL for 90-day NRR segregating in French dairy cattle at 108 Mb. As the position estimates from linkage analysis can have large uncertainties, the QTL reported in French cattle may be the same as the QTL we observed at 92.4 Mb for FTI and IFLC. An increase in ICF and IFLC results in longer intervals before success of insemination, necessitating fertility treatments, increased inseminations/conceptions, and a decreased overall fertility index. Though all of the fertility trait definitions were not identical across the studies, the QTL detected at the same chromosomal regions may be the same loci affecting one or several fertility-related traits.

Four significant SNP-trait associations were identified on BTA3 (Table 1). ss86299093 (at 81.3 Mb) had the smallest *P*-value and was associated with both FTI and IFLC. The QTL region ranged over 81.1–89.2 Mb and also showed a suggestive association with ICF. Interestingly, Ben-Jemaa *et al.* (2008) reported another QTL segregating on the first part of BTA3 for 90-day NRR. However, this QTL position cannot be compared directly, as both positions and marker orders differed from Btau_4.0. Regardless, Druet *et al.* (2008) fine-mapped the same QTL to 19 cM on BTA3. However, the present study could not confirm the presence of this QTL in Danish and Swedish Holstein cattle. There was, however, an indicative QTL at 80 cM in Druet *et al.* (2008), which may be identical to the one we detected at 81.3 Mb.

The QTL region on BTA4 containing the significant SNP rs29020693 is located between 88.1 and 94.0 Mb and affects AISC, FTI, and IFLC. Höglund *et al.* (2009) previously detected a QTL affecting IFLC at 39.6 Mb. A suggestive region associated with AISC, FERT3, IFLC, and FTI was observed in the present study between 46.9 and 51.9 Mb on this chromosome. The present findings together with those of Höglund *et al.* (2009) suggest that this QTL is likely to be legitimate.

We observed the strongest evidence of association between fertility traits and SNPs on BTA5. Two SNPs exceeded the genome-wide threshold, and the QTL region

doi: 10.1111/j.1365-2052.2010.02064.x

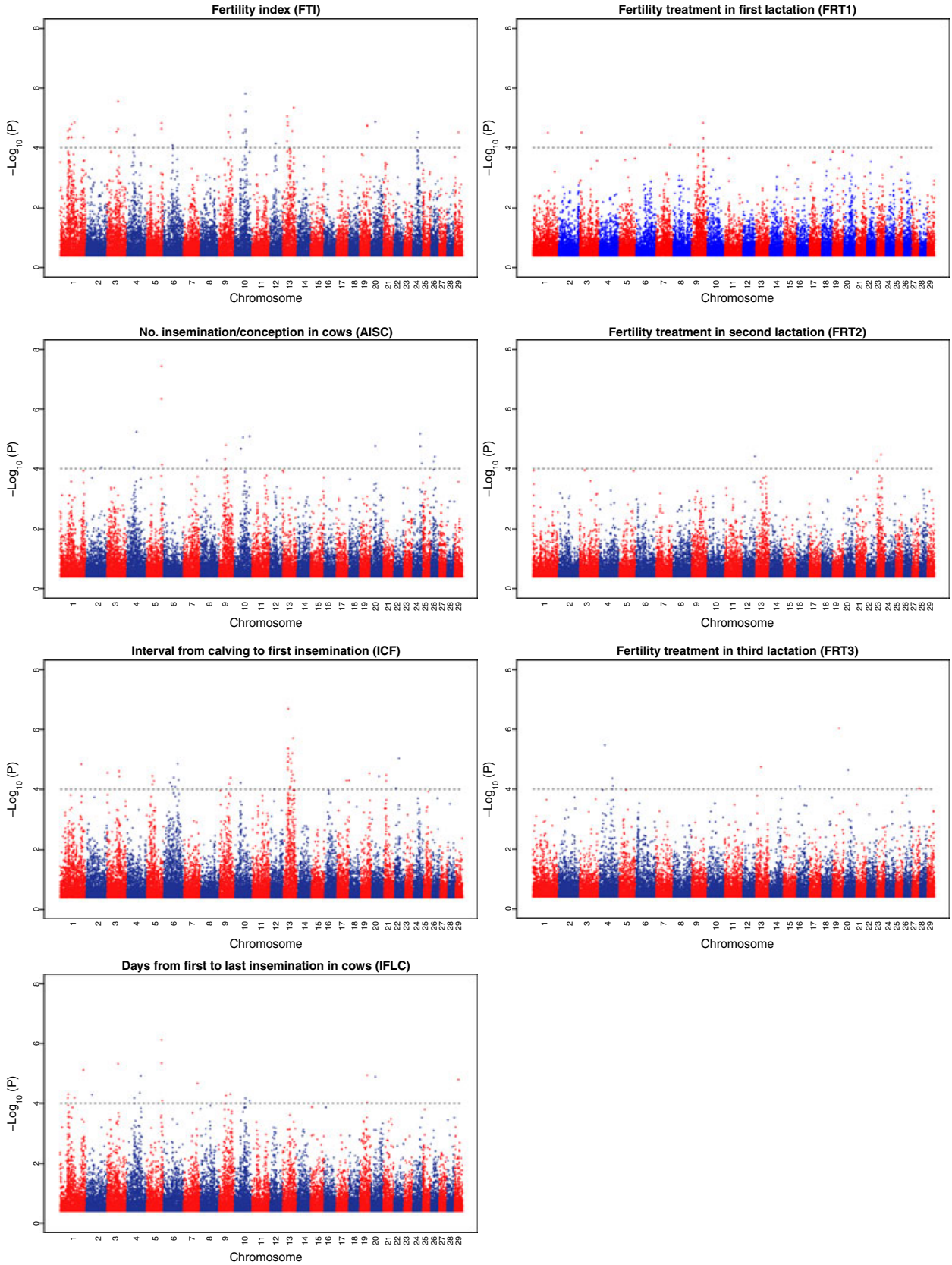

Figure 1 Genome-wide scan for fertility traits: $-\log_{10}$ of the *P*-value analysis for association with SNPs. Chromosomes are shown in alternating colours for clarity. The dotted line represents suggestive association [$-\log_{10}$ (*P*-value) = 4] as considered in the present study.

doi: 10.1111/j.1365-2052.2010.02064.x

Table 2 Most significant SNPs and the boundary of the QTL detected using association mapping for fertility traits.

Chromosome	QTL region#	Lower-boundary		Upper boundary		Most significant SNP			Traits showing association*
		SNP	Position(bp)	SNP	Position(bp)	SNP	Position(bp)	*P*-value	
1	1	ss61508407	85264719	rs29019866	92433977	rs29019866	92433977	1.42E-05	FTI, IFLC
1	2	ss61544878	135219500	ss86336776	142283845	ss86338186	136499200	1.42E-05	ICF, IFLC, FTI, AISC
1	3	ss61476732	148322768	ss86338760	154222249	ss61522035	148765346	7.80E-06	IFLC, FTI, AISC
3	1	rs43041636	81060752	rs43350054	89150703	ss86299093	81303069	2.81E-06	FTI, IFLC, ICF,
4	1	rs42688382	88071518	rs43411018	93978310	rs29020693	90418033	1.22E-05	AISC, FTI, IFLC
5	1	ss86309245	113570138	ss86291377	121206637	ss86299524	116362061	3.69E-08	AISC, FTI, IFLC
6	1	ss86317213	87879378	ss86296213	90008100	ss86337596	89774923	1.39E-05	ICF
7	1	ss117968986	90305805	ss86285499	91300574	ss117968986	90305805	2.18E-05	FERT1, IFLC
9	1	ss61523758	40425232	ss61561345	52271111	rs42466129	51661514	1.79E-05	AISC,IFLC
9	2	rs29027423	63019666	ss86301859	68411525	ss117964048	67688020	2.98E-05	FTI
9	3	ss86293152	82136926	rs43609252	84609287	rs43732223	83133779	8.10E-06	FTI, IFLC, FERT1, ICF
10	1	ss117970005	32331352	rs42525044	41302784	ss61524214	40769608	2.14E-05	AISC, ICF
10	2	ss86336211	47500126	rs43629990	52714675	rs43087086	52660587	8.90E-06	AISC, FTI
10	3	ss105240223	65757089	rs42227324	70287191	ss105241610	69247532	1.55E-06	FTI, IFLC
10	4	ss61471889	90356426	rs43081611	97308023	ss61559544	93016335	8.16E-06	AISC, IFLC
13	1	ss86339552	27708081	ss86284616	38687018	ss86288462	33517748	2.01E-07	FERT3, FTI, ICF
13	2	rs29015694	40802344	ss86312717	52066047	ss61474875	44121804	9.82E-06	ICF
13	3	ss86296563	57904589	ss86336655	76326511	ss86288836	60775960	1.93E-06	FTI, ICF
20	1	ss61520373	20040904	ss46526609	38201199	ss61543090	26416375	1.31E-05	AISC, FERT3, IFLC, FTI
21	1	ss117972978	14946223	ss86298519	16092606	rs29017183	15643297	3.35E-05	FTI,
24	1	rs42609685	29594855	rs42048480	29757516	ss86302405	29730649	4.57E-05	FTI
24	2	rs42050669	35550267	ss86339808	40994513	ss61486559	36746469	2.99E-05	FTI, AISC
24	3	rs29025240	47270718	ss86287710	48799059	ss86301009	48680176	6.62E-06	AISC, FTI, IFLC
26	1	ss86303243	18441296	ss61549738	20994002	ss86336536	19918015	5.68E-05	AISC

*The abbreviations for trait names used are the same as in Table 1. #QTL region within a chromosome

doi: 10.1111/j.1365-2052.2010.02064.x

between 113.4 and 121.2 Mb (strongest associated SNP at 116.3 Mb) was associated with AISC, FTI, and IFLC. This QTL might also affect FERT1 and FERT2 as there was suggestive evidence for association. Using regression-based linkage analysis, Schulman *et al.* (2008) reported QTL for days open and fertility treatments segregating between *BM2830* (113.5 cM) and *BM43* (118.3 cM) on BTA5 in Finnish Ayrshire cattle. Thus, evidence for this QTL segregating in another population increases the confidence for its existence.

SNPs spread over a 1 Mb chromosomal region centered at 90.3 on BTA7 showed association with FERT1 and IFLC and also showed suggestive association with FERT2, FTI, ICF, and AISC. Höglund *et al.* (2009) detected a QTL for heat strength between *BMS2258* (77.2 cM) and *AE129* (95.9 cM) and for AISH between *DIK2895* (103.1 cM) and *MB057* (116.6 cM); the QTL we found at 90.3 Mb may be the same as reported for heat strength (Höglund *et al.* 2009). Boichard *et al.* (2003) detected a QTL for success of insemination in daughters at 103.1 Mb on this chromosome in the French Holstein population, which could be the same one reported by Höglund *et al.* (2009) for AISH. However, our results could not confirm this QTL.

Three QTL regions that were associated with fertility traits on BTA9 are presented in Table 2, but the demarcation between QTL regions is not distinct. SNPs with significant and suggestive associations with fertility traits were spread across the entire chromosome. Of these, four SNP-Trait combinations showed significant association involving AISC, FTI, and FERT1. Höglund *et al.* (2009) detected a highly significant QTL for IFLC between *UWCA9* (49.9 cM) and *DIK4912* (51.9 cM) and IFLH near *BMS2151* (4.9 cM). Schrooten *et al.* (2000) and Holmberg & Andersson-Eklund (2006) reported a QTL for NRR close to the same microsatellite marker, *TGLA73*, located at 76.7 Mb on BTA9. Holmberg *et al.* (2007) fine-mapped the QTL between *BMS1724* (80.1 Mb) and *BM7209* (81.2 Mb) on BTA9, and we additionally observed significant association with AISC, FERT1, and FTI. The NRR and ICF are strongly related to the number of inseminations/conceptions, fertility treatment, and thereby, the overall FTI. Therefore, all of these studies, including the present one, provide evidence for fertility trait QTL segregating on BTA9. However, the number of QTL and their possible positions vary among these studies, making it difficult to select regions in which to search for candidate genes underlying the QTL.

The present study yielded strong evidence indicating that more than one QTL segregates on BTA10. Table 2 presents four probable chromosomal regions that may harbour fertility-related QTL. The strongest evidence of SNPs associated with AISC, FTI, and IFLC was in the 65.8–70.3 Mb region, and the most significant SNP (ss105241610) was located at 69.2 Mb. The other regions on BTA10 showing significant associations were at 40.7, 52.7, and 93 Mb. Clusters of SNPs suggestive of QTL were also observed at 40.7 and 93 Mb, distinct from other QTL regions. Meanwhile, the QTL at 52.7 and 69.2 Mb were not clearly separated, as there are SNPs with suggestive association located between them. Höglund *et al.* (2009) reported a QTL for IFLC between *BMS2641* (87.5 cM) and *bms614* (100.0 cM) that might also affect AISC. We mapped this QTL, which we found to affect AISC and IFLC (with possible additional effects on FTI and FERT3), at 93.0 Mb. Boichard *et al.* (2003) reported a QTL for insemination success at 95.2 Mb, close to our QTL location. Further, Schulman *et al.* (2008) reported QTL on BTA10 for NRR between *ILSTS053* (37.9 cM) and *ILSTS070* (72.3 cM) and fertility treatment at 95.2 Mb. The QTL we observed at 69.2 Mb may be the same as the NRR QTL reported by Schulman *et al.* (2008).

About one-third of all significant trait-SNP associations in the whole genome scan were on BTA13. Out of these 21 significant trait-SNP combinations, 15 involved ICF. The association between ss86288462 and ICF was significant genome-wide, and the region delineated around this SNP spreads between 27.7 and 38.7 Mb (Table 2). The other two probable QTL regions on BTA13 were located in the 40.8–52.1 Mb region and in the 57.9–76.3 Mb region, with most significant SNPs at 44.1 and 60.8 Mb respectively. Höglund *et al.* (2009) observed a QTL for ICF between 71.9 and 77.6 Mb on BTA13, and while we detected several SNPs significantly associated with ICF, the one between 57.9 and 76.3 Mb could be the same QTL in the two studies.

A SNP located at 41.7 Mb on BTA19 showed genome-wide significant association with FERT3. There were also three more SNPs, two of which were located at 45.8 Mb and the other at 54.1 Mb, that showed significant association with fertility traits. Schulman *et al.* (2008) reported a QTL affecting fertility treatment located on BTA19 near marker *ETH3* (81.5 cM). The SNP with genome-wide significance that we observed also affected fertility treatment, but the positions of the QTL observed in these two studies were far apart.

The SNP ss61543090, located at 26.4 Mb on BTA20, showed significant associations with FTI, AISC, IFLC, and FERT3. This SNP was flanked by two regions, 20.0 and 35.1 Mb, where there is suggestive evidence of association with several SNPs. We also found an additional region (60.8–73.8 Mb) on BTA20 displaying suggestive association with fertility traits. Höglund *et al.* (2009) reported a QTL for FERT2 at 49.1–65.5 Mb. Combining our results with those from Höglund *et al.* (2009), it seems that one QTL is located at 35.1 cM, and the map position of ss1543090 may be not correct (or it could be a false positive). Further, we found evidence suggestive of a fertility treatment QTL at 60.8–67.1 Mb, which may be identical to the fertility treatment QTL reported by Höglund *et al.* (2009). The suggestive QTL in the region of 60.8–73.8 Mb was consistent with the QTL for insemination success in a French cattle population (Boichard *et al.* 2003).

doi: 10.1111/j.1365-2052.2010.02064.x

Four SNPs showed significant association with fertility traits on BTA24, and there were three QTL regions on this chromosome. The QTL region at 29.7 Mb was very narrow (less than 0.2 Mb), affected FTI, and was suggestive for ICF. The second QTL region, with the most significant SNP at 36.7 Mb, extended from 35.6 to 41.0 Mb. The third QTL region contained the most significant BTA24 SNP and extended from 47.3 to 48.8 Mb. Further, there were two additional clusters of SNPs showing suggestive associations with fertility traits at 55.4 and 60.6 Mb. Höglund *et al.* (2009) reported a QTL affecting ICF between *ILSTS065* (27.4 cM) and *BMS1862* (35.5 cM), which is likely the same QTL we observed at 29.7 cM. However, they also reported a QTL for NRRH at 4.2 Mb, and our evidence for this location was very weak. Thus, this QTL could not be confirmed.

There were two SNPs on BTA26 that showed significant association with AISC. The QTL region, with centre at 19.9 Mb, extended across 2.6 Mb (18.4–21.0 Mb). Previously, Höglund *et al.* (2009) found QTL for NRRH and IFLC at 34.8–41.0 and 41.2–46.0 Mb respectively.

The present study did not detect the QTL for NRR segregating on five chromosomes described by Höglund *et al.* (2009). The q-values for these QTL were very high (four of them had q-values higher than 0.39), indicating that about half of them are likely false positives. Association analyses with NRR did not reveal any significant association of NRR with SNPs. The heritability of NRR for both cows and heifers was very low (0.011 and 0.007 respectively, Guosheng Su, *pers. comm.*), indicating that the genetic contribution to the total variance in NRR is negligible.

QTL supported by multiple traits

Most of the identified QTL showed evidence of association with multiple fertility traits, and some SNPs had significant association with more than one trait. The highest number of overlapping QTL regions were observed for FTI, ICF, IFLC, and AISC. The correlation between breeding values of FTI with IFLC was 0.95, followed by AISC (0.84) and ICF (0.62). This was evident because many of the identified QTL regions are common to these four traits. Six SNPs showed significant association both with FTI and IFLC, while three were significant for FTI and AISC, and FTI and ICF. AISC and IFLC measure the same biological phenomena, and therefore it was expected that most of the QTL for these two traits would be common. Similarly, FTI is a combination of AIS, ICF, and IFLC. Therefore, QTL showing association with any of these three traits are generally also associated with FTI. Many of these QTL also showed indications of having an effect on one of the fertility treatment traits but did not cross the significance threshold. For instance, a delay in conception may result in a fertility treatment. Thus, it is reasonable that genes that affect fertility are also associated with fertility treatment. We did not definitively detect any SNPs significantly associated with 56 day NRR or with any of the heifer fertility traits. However, there were indications that some of the identified QTL affect these traits.

QTL position estimates

Genome-wide association studies generally map QTL in narrower chromosomal regions than earlier reports using linkage analysis. Thus, this work represents a marked improvement from prior data reported earlier by Höglund *et al.* (2009) using linkage analysis to study the same population. The smallest QTL region was 0.16 Mb for QTL-1 on BTA24, which is indicative of the power of GWA to tightly define such loci. Additionally, eight QTL regions were mapped in less than 5 Mb regions.

Some of the QTL regions revealed in this study extended across large intervals. For example, QTL-3 on BTA13 extended across a long interval (57.9–76.3 Mb) that consists of two smaller regions, one spanning 60.6–61.0 Mb and another 67.3–69.8 Mb. High-resolution mapping typically shows that single QTL fractionate into multiple closely linked QTL (Flint & Mackay 2009). Thus, the broad QTL regions observed in this study likely harbour more than one QTL. In the future, these regions can be further enriched with SNPs to narrow down the chromosomal regions for candidate gene and candidate polymorphism studies.

The LD in cattle is spread over a wide region due to low effective population size and strong selection. This LD breadth places limitations on mapping QTL to regions suitable for searching for candidate genes. However, joint analysis of several cattle populations may help to narrow the QTL confidence interval and increase the confidence that a particular QTL is a true positive (Sahana *et al.* 2008).

Conclusion

The present GWA study presented strong and highly consistent evidence for association of SNPs with fertility traits. To our knowledge, no GWA study for these fertility traits in cattle has been published thus far. On the contrary, there are a few studies using linkage analysis to map QTL for fertility traits, and the majority of them came from Nordic countries. However, the QTL mapping regions reported in these studies were large and not suitable for further candidate gene analysis. Our study was able to map QTL in narrow regions, some of them less than 1 Mb wide, and these results can be used for identifying candidate causal genes by enrichment for SNPs in the identified intervals. It will also be possible to determine the genetic merit of individual animals by combining effects estimated for the SNPs.

Acknowledgements

This work was performed in the projects 'Genomic Selection – from function to efficient utilization in cattle breeding'

doi: 10.1111/j.1365-2052.2010.02064.x

(grant no. 3412-08-02253), funded by the Danish Directorate for Food, Fisheries and Agri Business, VikingGenetics, Nordic Genetic Evaluation, and Aarhus University, and 'DNA based selection to improve disease resistance, fertility, calf survival and production in Danish cattle' (grant no. 3401-04-00853), funded by the Danish Directorate for Food, Fisheries and Agri Business. We thank the Nordic Cattle Genetic Evaluation for providing phenotypic data. VikingGenetics is acknowledged for providing blood and semen samples. Peter Sørensen and Ole Fredslund Christensen, Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, Aarhus University are gratefully acknowledged for help in mixed model analyses and valuable discussion.

References

Ancker S., Norberg E. & Callesen H.. (2006) Videnssyntese reproduction 113. In: Dansk Landbrugsrådgivning. ISSN0906-1398.

Ben-Jemaa S., Fritz S., Guillaume F., Druet T., Denis C., Eggen A. & Gautier M. (2008) Detection of quantitative trait loci affecting non-return rate in French dairy cattle. *Journal of Animal Breeding and Genetics* **125,** 280–8.

Boichard D., Grohs C., Bourgeois F., Cerqueira F., Faugeras R., Neau A., Rupp R., Amigues Y., Boscher M.Y. & Levéziel H. (2003) Detection of genes influencing economic traits in three French dairy cattle breeds. *Genetics Selection Evolution* **35,** 77–101.

Collins F.S., Brooks L.D. & Chakravarti A. (1998) A DNA polymorphism discovery resource for research on human genetic variation. *Genome Research* **8,** 1229–31.

Dematawewa C.M.B. & Berger P.J. (1998) Genetic and phenotypic parameters for 305-day yield, fertility, and survival in Holsteins. *Journal of Dairy Science* **81,** 2700–9.

Devlin B. & Risch N. (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29,** 311–22.

Druet T., Fritz S., Boussaha M. *et al.* (2008) Fine mapping of quantitative trait loci affecting female fertility in dairy cattle on BTA03 using a dense single-nucleotide polymorphism map. *Genetics* **178,** 2227–35.

Flint J. & Mackay T.F.C. (2009) Genetic architecture of quantitative traits in flies, mice and humans. *Genome Research* **19,** 723–33.

Han B., Kang H.M. & Eskin E.. (2009) Rapid and accurate multiple testing correction and power estimation for millions of correlated markers. *PloS Genetics* **5,** e1000456. doi:10.1371/journal.pgen.1000456.

Hästbacka J., de la Chapelle A., Kaitila I., Sistonen P., Weaver A. & Lander E. (1992) Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nature Genetics* **2,** 204–11.

Höglund J.K., Guldbrandtsen B., Su G., Thomsen B. & Lund M.S. (2009) Genome scan detects quantitative trait loci affecting female fertility traits in Danish and Swedish Holstein cattle. *Journal of Dairy Science* **92,** 2136–43.

Holmberg M. & Andersson-Eklund L. (2006) Quantitative trait loci affecting fertility and calving traits in Swedish dairy cattle. *Journal of Dairy Science* **89,** 3664–71.

Holmberg M., Sahana G. & Andersson-Eklund L. (2007) Fine mapping of a quantitative trait locus on chromosome 9 affecting non-return rate in Swedish dairy cattle. *Journal of Animal Breeding and Genetics* **124,** 257–63.

Liu Y., Qin X., Song X.Z. *et al.* (2009) *Bos taurus* genome assembly. *BMC Genomics* **10,** 180. doi:10.1186/1471-2164-10-180.

Matukumalli L.K., Lawley C.T., Schnabel R.D. *et al.* (2009) Development and characterization of a high density SNP genotyping assay for cattle. *PLoS ONE* **4,** e5350. doi:10.1371/journal.pone.0005350.

Meuwissen T.H.E., Hayes B.J. & Goddard M.E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157,** 1819–29.

Pedersen J., Sørensen M.K., Toivonen M., Eriksson J. & Aamand G.P.. (2008) Report on economic basis for a Nordic total merit index. http://www.nordicebv.info/Publications/English/publicatuions+ny.htm.

Pryce J.E. & Veerkamp R.F. (2001) The incorporation of fertility indices in genetic improvement programmes. *BSAS occasional meeting* '*Fertility in high producing dairy cow*'. 10–22 September 1999, Galway. BSAS Publication 26 (I), 237–49.

Pryce J.E., Veerkamp R.F., Thompson R., Hill W.G. & Simm G. (1997) Genetic aspects of common health disorders and measures of fertility in Holstein Friesian dairy cattle. *Animal Science* **65,** 353–60.

Roxström A., Strandberg E., Berglund B., Emanuelson U. & Philipsson J. (2001) Genetic and environmental correlations among female fertility traits and milk production in different parities of Swedish Red and White dairy cattle. *Acta Agriculturae Scandinavica A* **51,** 7–14.

Sahana G., Lund M.S., Andersson-Eklund L. *et al.* (2008) Fine-mapping QTL for mastitis resistance on BTA9 in three Nordic red cattle breeds. *Animal Genetics* **39,** 354–62.

Schaeffer L.R. (2006) Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics* **123,** 218–23.

Schrooten C., Bovenhuis H., Coppieters W. & Van Arendonk J.A. (2000) Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. *Journal of Dairy Science* **83,** 795–806.

Schulman N.F., Sahana G., Lund M.S., Viitala S.M. & Vilkki J.H. (2008) Quantitative trait loci for fertility traits in Finnish Ayrshire cattle. *Genetics Selection Evolution* **40,** 195–214.

Sørensen A.C., Lawlor T. & Ruiz F.. (2007) *A survey on fertility in the Holstein populations of the world. Proc. Fertility in dairy cows – bridging the gaps*. Liverpool Hope University, Liverpool, UK. Pp. 17.

Teo Y.Y., Inouye M., Small K.S., Gwilliam R., Deloukas P., Kwiatkowski D.P. & Clark T.G. (2007) A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics* **23,** 2741–6. doi:10.1093/bioinformatics/btm443

The Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447,** 661–78.

Veerkamp R.F. & Beerda B. (2007) Genetics and genomics to improve fertility in high producing dairy cows. *Theriogenology* **68S,** S266–73.

Yu J., Pressoir G., Briggs W.H. *et al.* (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* **38,** 203–8.

doi: 10.1111/j.1365-2052.2010.02064.x