ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12420

Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows

M. S. Ortega*, A. C. Denicol*[1], J. B. Cole[†], D. J. Null[†] and P. J. Hansen*

*Department of Animal Sciences, D.H. Barron Reproductive and Perinatal Biology Research Program and Genetics Institute, University of Florida, Gainesville, FL, USA. [†]Animal Genomics and Improvement Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD, USA.

Summary

We evaluated 69 SNPs in genes previously related to fertility and production traits for their relationship to daughter pregnancy rate (DPR), cow conception rate (CCR) and heifer conception rate (HCR) in a separate population of Holstein cows grouped according to their predicted transmitting ability (PTA) [≤−1 ($n = 1287$) and ≥1.5 ($n = 1036$)] for DPR. Genotyping was performed using Sequenom MassARRAY®. There were a total of 39 SNPs associated with the three fertility traits. The SNPs that explained the greater proportion of the genetic variation for DPR were *COQ9* (3.2%), *EPAS1* (1.0%), *CAST* (1.0%), *C7H19orf60* (1.0%) and *MRPL48* (1.0%); for CCR were *GOLGA4* (2.4%), *COQ9* (1.8%), *EPAS1* (1.1%) and *MRPL48* (0.8%); and for HCR were *HSD17B7* (1.0%), *AP3B1* (0.8%), *HSD17B12* (0.7%) and *CACNA1D* (0.6%). Inclusion of 39 SNPs previously associated with DPR in the genetic evaluation system increased the reliability of PTA for DPR by 0.20%. Many of the genes represented by SNPs associated with fertility are involved in steroidogenesis or are regulated by steroids. A large proportion of SNPs previously associated with genetic merit for fertility in Holstein bulls maintained their association in a separate population of cows. The inclusion of these genes in genetic evaluation can improve reliabilities of genomic estimates for fertility.

Keywords dairy cow, fertility, genetic evaluation and reproduction, single nucleotide polymorphisms

Introduction

Fertility in cows is a complex trait that is regulated in part by genetics. Breeding values for fertility and production have moved in opposite directions over the past five decades, probably in part because of the negative genetic correlation between fertility and milk production (Boichard & Manfredi 1994; VanRaden *et al.* 2004; Pritchard *et al.* 2013). Genetic improvement of cow fertility using traditional breeding approaches has been slow because of the low heritability of reproductive traits, which ranges from 0.01–0.10 (Averill *et al.* 2004; Pryce *et al.* 2004; VanRaden *et al.* 2004). Similarly, although incorporation of genomic information through genome-wide single nucleotide polymorphism (SNP) arrays has improved the reliability of genetic estimates, it has had a lower impact for traits of low heritability than for traits of higher heritability (Wiggans *et al.* 2011). For example, the gain in reliability from including genomic data was 17% for daughter pregnancy rate (DPR) vs. 30% for milk production (Wiggans *et al.* 2011).

One approach to improving the use of genomic information for estimating genetic merit is to incorporate information on causative mutations affecting phenotypes of interest. Advantages of identifying causal mutations include the facts that the effect of a causal mutation on a trait will be greater than that of a SNP in close linkage disequilibrium; that the relationship with a trait may extend across breeds and not change over time; and that information about the function of the gene, and the physiological control of the trait of interest, will be obtained (Weller & Ron 2011).

Several approaches have been used to identify putative causative mutations affecting fertility. The availability of

Address for correspondence

P. J. Hansen, Department of Animal Sciences, D.H. Barron Reproductive and Perinatal Biology Research Program and Genetics Institute, University of Florida, PO Box 110910, Gainesville, FL 32611-0910, USA.
E-mail: pjhansen@ufl.edu

[1]Present address: Department of Biology, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, USA.

Accepted for publication 24 December 2015

genomic information on large populations of individuals has allowed for the identification of haplotypes negatively affecting fertility in dairy cattle (VanRaden *et al.* 2011), and putative causative mutations associated with some of these haplotypes have been pinpointed in the US Jersey, Holstein and Brown Swiss breeds (Sonstegard *et al.* 2013; McClure *et al.* 2014). Resequencing of regions near quantitative trait loci has led to the identification of causal mutations associated with fetal death in Holsteins (Charlier *et al.* 2012) and embryonic lethality in Nordic Red cattle (Kadri *et al.* 2014). Using analyses of gene networks, candidate genes for cattle reproductive behavior have been identified (Kommadath *et al.* 2011). A candidate gene approach has also been used to identify mutations in genes associated with fertilization and embryonic development during the pre-implantation period (Khatib *et al.* 2009a,b; Cochran *et al.* 2013a).

Cochran *et al.* (2013b) identified 51 SNPs in candidate genes that were associated with one or more fertility traits in Holstein bulls including DPR, cow conception rate (CCR) and heifer conception rate (HCR). Many of these SNPs had neutral associations with milk production, suggesting that it could be possible to select for fertility without selecting against milk yield. However, genetic markers in one study are often not predictive in other studies (Siontis *et al.* 2010; Ioannidis *et al.* 2011). The objective of this study was to evaluate the SNPs in candidate genes previously associated with genetic merit for female fertility in Holstein bulls (Cochran *et al.* 2013b) in a separate population of Holstein cows. In addition, the degree to which incorporation of these SNPs in genomic estimates of DPR would improve reliability was ascertained.

Materials and methods

Selection of animals

Holstein cows were selected based on their predicted transmitting ability (PTA) and reliability for DPR; these values were obtained from the May 2014 US national genetic evaluations (Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD). Cows were selected to have a high (≥1.5) or low regressed PTA for DPR (≤−1.0). The minimum reliability for inclusion was 0.25. The PTA for DPR ranged from −5.4 to −1 for the low DPR group and from 1.5 to 3.7 for the high DPR group. Reliabilities ranged from 0.25 to 0.71 and from 0.25 to 0.73 for the low and high DPR groups respectively. Cows were located in six dairies in Florida and five in California. There were 1036 cows in the high DPR group and 1287 in the low DPR group. All animals had at least one lactation completed at the time of sampling, with a range of one to seven lactations. Summary statistics for cows used in genetic analyses are presented in Table S1.

Genotyping

Samples of whole blood were collected via saphenous vessel venipuncture into tubes coated with ethylenediamine tetracetic acid and shipped on ice to Neogen for DNA isolation and genotype determination. Genotyping was performed using the Sequenom MassARRAY® system (iPLEX GOLD; Sequenom) according to the manufacturer's instructions. The technique is based on the analysis of DNA products using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (Abel *et al.* 2006). The region of DNA containing the SNP was amplified by PCR, a primer extension reaction was performed to generate allele-specific DNA products and the size and amount of each allele-specific product were determined using chip-based mass spectrometry. The average call rate was 95.67%. A random sample of 10 SNPs was determined in duplicate for each animal; agreement between duplicates was 97.02%. When the genotype did not match between samples, both genotypes were deleted and treated as no calls.

Animals were genotyped for a total of 69 SNPs previously associated by Cochran *et al.* (2013b) with at least one the following traits: CCR, HCR, productive life (PL), net merit (NM), milk yield (MY), fat yield (FY), protein yield (PY), protein percent (PPCT), fat percent and somatic cell score. Of these, 51 were associated with one or more fertility traits (DPR, CCR and/or HCR), with 39 being previously associated with DPR. The SNPs were from a larger set of 434 genes selected using criteria described by Cochran *et al.* (2013b) for a possible relationship to reproductive function. The 69 SNPs were significantly associated with at least one trait based on a false discovery rate of 0.05.

Details of each SNP are described by Cochran *et al.* (2013b). Most of the 69 SNPs in the present study are missense mutations in the coding sequence of the gene except for *HSP70* (promoter mutation), *FGF2* and *PGR* (intronic mutations) and *ARL6IP1* and *TBC1D24* (nonsense mutations). SNPs causing an amino acid change were selected because they were predicted to change the structure of the encoded protein using an exchangeability matrix (Yampolsky & Stoltzfus 2005).

Statistics

Statistical analysis was performed using the STATISTICAL ANALYSIS SYSTEM (SAS, version 9.4; SAS Institute, Inc.). Minor allele frequency (MAF) was calculated, and the distribution of genotypes for each SNP was tested for deviation from Hardy–Weinberg equilibrium using the ALLELE procedure.

The association of genetic variants for each SNP was estimated individually using the MIXED procedure of SAS. Genotype was considered a categorical variable. The full model was as follows:

$$Y_{ij} = \mu + a_i + g_j + e_{ij}$$

where Y_i is the deregressed value of the trait of interest for the i^{th} cow ($i = 1, 2, \ldots, n$), a_i is the random polygenic effect (including all available pedigree information) of the i^{th} cow, g_j is the fixed effect of SNP genotype j and e_{ij} is the random residual effect. $a_i \sim A\sigma_a^2$ and $\varepsilon i \sim W\sigma_e^2$, where A is the numerator relationship matrix, W is an diagonal matrix of weights, σ_a^2 is the additive genetic variance of the trait of interest and σ_e^2 is the residual error variance. The elements of W are proportional to the number of observations for each cow in the dataset. All of the available pedigree information for each cow was used when modeling the covariance among the polygenic effects.

The additive effect at locus j was estimated as $(g_{j3}\ g_{j1})/2$, and the dominance effect at locus j was estimated as $g_{j2} - (g_{j3} + g_{j1})/2$. The SNPs in which the additive or dominance effect was $P < 0.05$ were noted.

Changes in reliability by including candidate genes in the calculation of the estimate for DPR

An analysis was performed to determine whether the inclusion of genotypes for 39 candidate genes previously related to DPR (Cochran *et al.* 2013b) improved the reliability of genomic estimates of DPR obtained using genome-wide SNP arrays. Using as a reference population the 2323 cows from the present study and the 550 bulls used by Cochran *et al.* (2013b), genotypes for the 39 SNPs of interest were imputed in the rest of U.S. Holstein population that had been subjected to genomic analysis and was available in the national evaluation system. The dataset included 494 050 Holsteins.

The imputation was performed in a two-step process. In the first step, the 39 SNPs of interest were added to genotypes from the commercial genomic chips [Illumina 50K versions 1 and 2, 3K, and high and low density; GeneSeek Genomic Profiler versions 1 and 2, and the GeneSeek Genomic Profiler-HD (Neogen); Zoetis low- and medium-density chips (Zoetis); and the Eurogenomics EuroG10k (Illumina)] and coded as missing values for the animals not genotyped for the 39 SNPs. In the second step, all genotypes were imputed to the 61 013 markers set used for routine genomic evaluation following the procedure of VanRaden *et al.* (2013). Missing genotypes for the 39 novel SNPs were filled during the imputation process.

The gain in reliability from adding the 39 SNPs to the evaluation SNP set was determined using the method described by VanRaden *et al.* (2008), by which the last four years of phenotypic information was removed and parent averages and parent average reliabilities were compared with values that included SNP and polygenic effects estimated using genomic information. The analysis was performed first with the 61 013 SNPs currently used for the US genomic evaluations (USDA-SNP panel), and then the 39 SNPs previously related to DPR (UFL-SNP panel) were added to the official SNP set and the evaluations recalculated (Table S2). This is the approach used in the United States to determine reliability gains when the SNPs used in the evaluations are changed. The analysis was also repeated after removing the SNPs in the USDA-SNP panel that were within 100 or 200 kb of SNPs in the UFL-SNP panel. In addition, the reliability of using only UFL-SNP panel was determined.

Pathway analysis

Genes that were significantly associated with DPR, CCR or HCR were subjected to pathway analysis using QIAGEN's INGENUITY® PATHWAY ANALYSIS (IPA®; www.ingenuity.com). In this analysis, associations were calculated using direct and indirect relationships among genes that were experimentally observed. IPA® software was used to identify (i) canonical pathways in which at least two genes were over-represented, (ii) upstream regulators affecting at least six genes and (iii) biological functions predicted to be regulated by the genes in which a minimum of six genes were acting. For all analyses, only relationships with a P-value ≤ 0.05 are reported.

Results

Allele frequencies

The frequency of each allele is presented in Table S3. In general, the minor allele was the same as previously reported by Cochran *et al.* (2013b). The exceptions were for *BSP3* [minor allele = G (29.8% MAF) vs. A (MAF = 30.3%) in the previous study], *COQ9* [A (MAF = 48.6%) vs. G (MAF = 48.5%)], *MRPL48* [G (MAF = 37.9%) vs. A (MAF = 49.2%)] and *PARM1* [G (MAF = 47.3%) vs. C (MAF = 47.8%)]. Of the 69 SNPs, 17 were not in Hardy–Weinberg equilibrium. For *ACAT2, AP3B1, AVP, C7H19orf60, CCT8, DZIP3, FST, HSP70, OCLN, SLC18A2, TSHB* and *ZP2*, there were more animals homozygous for the major allele than expected. For *HSD17B6, HSD17B7* and *TXN2*, there were fewer heterozygotes than expected, and for *NLRP9* and *PARM1*, there were more heterozygotes than expected (>50%).

Association of SNPs with fertility traits

Results for DPR are shown in Table 1. There were significant relationships with DPR for 29 SNPs, with 22 having an additive effect, four having a dominance effect and three having additive and dominance effects. Of the 29 SNPs, *ARL6IP1* and *TBC1D24* are non sense mutations, *PGR* is a mutation in an intronic region and the remaining 26 are missense mutations. The SNPs that explained the greatest proportion of the genetic variation in DPR were *COQ9*

Table 1 SNPs associated with daughter pregnancy rate.

Gene	SNP id	Chr.	Location	Copies of minor allele[1]			Effect	r^2	*P*–value[2]	
				0	1	2			A	D
ACAT2	rs109967779	9	97478396	0.62 (0.50)	1.59 (0.48)	3.38 (0.74)	1.38	0.0067	0.0006	0.4616
AP3B1	rs133700190	10	9177305	0.84 (0.45)	2.05 (0.50)	3.47 (1.15)	1.32	0.0003	0.0250	0.8881
APBB1	rs41766835	15	47252371	1.99 (0.44)	0.29 (0.52)	−0.58 (1.17)	−1.28	0.0088	0.0298	0.5624
ARL6IP1	rs110541595	25	16544291	0.88 (0.52)	2.08 (0.47)	0.60 (0.73)	−0.14	0.0008	0.7267	0.0143
BCAS1	rs109669573	13	82164839	1.97 (0.52)	1.20 (0.47)	0.03 (0.72)	−0.97	0.0035	0.0144	0.7268
BOLA-DMB	rs109032590	23	7189701	0.52 (0.47)	2.38 (0.49)	3.63 (0.91)	1.55	0.0062	0.0010	0.6255
BSP3	rs110217852	18	51919757	2.14 (0.46)	0.92 (0.50)	−0.35 (0.89)	−1.24	0.0059	0.0061	0.9637
C7H19orf60	rs109332658	7	4533772	2.44 (0.47)	−0.09 (0.52)	0.87 (0.91)	−0.78	0.0100	0.0977	0.0062
CAST	rs137601357	7	98485273	2.51 (0.54)	1.05 (0.47)	−0.70 (0.64)	−1.61	0.0105	<0.0001	0.7907
COQ9	rs109301586	18	25527339	−0.57 (0.57)	1.73 (0.47)	4.17 (0.59)	2.37	0.0319	<0.0001	0.8943
CSNK1E	rs133449166	5	110565337	1.03 (0.50)	1.33 (0.48)	3.10 (0.79)	1.03	0.0040	0.0143	0.1971
DEPDC7	rs110270752	15	63421395	1.58 (0.45)	0.70 (0.51)	3.01 (1.14)	0.71	0.0015	0.2190	0.0253
DNAH11	rs110629231	4	30766895	2.24 (0.49)	0.78 (0.48)	−0.47 (0.87)	−1.36	0.0059	0.0028	0.8630
DSC2	rs109503725	24	27124361	0.14 (0.55)	2.14 (0.47)	1.76 (0.68)	0.81	0.0058	0.0396	0.0276
DZIP3	rs133175991	1	53855896	1.42 (0.44)	1.59 (0.53)	−2.17 (1.52)	−1.80	0.0006	0.0189	0.0278
EPAS1	rs43676052	11	28650973	2.43 (0.46)	−0.43 (0.51)	−1.15 (1.06)	−1.79	0.0200	0.0010	0.1140
FCER1G	rs109137982	3	8308678	1.52 (0.41)	1.06 (0.68)	5.80 (2.32)	2.14	0.0020	0.0648	0.0479
FUT1	rs41893756	18	55831611	1.84 (0.43)	0.75 (0.54)	−2.40 (1.40)	−2.12	0.0078	0.0027	0.2169
HSD17B12	rs109711583	15	74828355	0.10 (0.54)	1.59 (0.46)	2.96 (0.66)	1.43	0.0056	0.0002	0.9055
HSD17B7	rs110828053	3	6630548	1.00 (0.43)	1.90 (0.57)	4.92 (1.34)	1.96	0.0044	0.0038	0.2033
LHCGR	rs41256848	11	30824442	1.68 (0.55)	1.74 (0.47)	0.12 (0.65)	−0.78	0.0023	0.0392	0.1100
MRPL48	rs43703916	15	54119266	0.56 (0.50)	1.40 (0.48)	3.53 (0.74)	1.48	0.0100	0.0002	0.2425
OCLN	rs134264563	20	10167825	0.35 (0.49)	1.58 (0.48)	3.82 (0.86)	1.73	0.0050	0.0001	0.3962
PCCB	rs109813896	1	134130474	0.23 (0.50)	1.99 (0.48)	2.77 (0.78)	1.27	0.0069	0.0023	0.3831
PGR	rs109506766	15	8158458	2.15 (0.54)	1.34 (0.47)	0.42 (0.69)	−0.86	0.0032	0.0265	0.9202
PMM2	rs109629628	25	7716425	0.24 (0.51)	1.63 (0.48)	3.17 (0.70)	1.47	0.0072	0.0001	0.8811
RABEP2	rs133729105	25	26182660	0.96 (0.52)	1.26 (0.48)	2.55 (0.74)	0.79	0.0032	0.0499	0.3694
TBC1D24	rs110660625	25	2007163	1.21 (0.49)	0.92 (0.48)	2.96 (0.81)	0.87	0.0008	0.0429	0.0474
TXN2	rs134031231	5	75266801	0.84 (0.54)	0.95 (0.47)	2.59 (0.68)	0.82	0.0002	0.0311	0.1254

[1]Data are least-squares means (standard error) for the predicted transmitting ability.
[2]A, additive; D, dominance.

(3.2%), *EPAS1* (1.0%), *CAST* (1.0%), *C7H19orf60* (1.0%) and *MRPL48* (1.0%).

There were 23 SNPs that had a significant association with CCR. Of these, 19 had an additive effect and four had a dominance effect (Table 2). The SNPs that explained the greatest proportion of the genetic variation for CCR were *GOLGA4* (2.4%), *COQ9* (1.8%), *EPAS1* (1.1%), *MRPL48* (0.8%) and *FUT1* (0.8%).

There were 14 SNPs with significant associations with HCR, of which nine had an additive effect, three had a dominance effect and two had additive and dominance effects (Table 3). The most explanatory SNPS were *HSD17B7* (1.0%), *AP3B1* (0.8%), *CACNA1D* (0.6%) and *HSD17B12* (0.6%).

Six SNPs were significantly associated with DPR, CCR and HCR (*AP3B1*, *BOLA-DMB*, *DNAH11*, *HSD17B12*, *HSD17B7* and *MRPL48*); 13 were associated with DPR and CCR (*ARL6IP1*, *BSP3*, *C7H19orf60*, *CAST*, *COQ9*, *DSC2*, *DZIP3*, *EPAS1*, *FCER1G*, *FUT1*, *OCLN*, *PCCB* and *PMM2*), one SNP affected both HCR and DPR (*CSNK1E*) and three were associated with CCR and HCR (*CD40*, *FST* and *IBSP*).

Comparison of significant SNPs with results of a previous candidate gene study

The SNPs that were significantly associated with DPR, CCR and HCR were compared to SNPs previously associated with fertility traits by Cochran *et al.* (2013b) (Table 4). For DPR, 19 of 39 SNPs found by Cochran *et al.* (2013b) to be associated with DPR were associated with DPR in the present experiment. For 15 of these SNPs (*ACAT2*, *AP3B1*, *ARL6IP1*, *C7H19orf60*, *CAST*, *COQ9*, *CSNK1E*, *DEPDC7*, *FUT1*, *HSD17B12*, *HSD17B7*, *OCLN*, *PCCB*, *PMM2* and *TBC1D24*), the allele associated with a higher value for DPR was the same as previously reported. For four of the SNPs (*APBB1*, *BSP3*, *DSC2* and *RABEP2*), the allele positively associated with DPR was the opposite of that found by Cochran *et al.* (2013b). In addition, 10 significant SNPs found in the present study were not associated with DPR in Cochran *et al.*'s (2013b) study. Of these 10 SNPs, four were previously associated with fertility, including CCR (*BCAS1*), HCR (*DZIP3*), PL (*LHCGR*) and NM (*EPAS1*), and six were previously associated with production traits including MY (*BOLA-DMB*), FY (*DNAH11* and *MRPL48*),

Table 2 SNPs associated with cow conception rate.

				Copies of minor allele[1]					*P*–value[2]	
Gene	SNP id	Chr.	Loc.	0	1	2	Effect	r^2	A	D
ACAT2	rs109967779	9	97478396	0.81 (0.95)	2.28 (0.91)	5.48 (1.40)	2.34	0.0051	0.0023	0.4127
AP3B1	rs133700190	10	9177305	0.98 (0.85)	2.82 (0.96)	5.94 (2.22)	2.48	0.0023	0.0287	0.6471
ARL6IP1	rs110541595	25	16544291	0.84 (0.99)	3.14 (0.90)	0.46 (1.40)	−0.19	0.0008	0.8093	0.0183
BOLA-DMB	rs109032590	23	7189701	0.35 (0.89)	3.48 (0.94)	5.51 (1.76)	2.58	0.0057	0.0046	0.6438
BSP3	rs110217852	18	51919757	2.89 (0.88)	1.27 (0.95)	−2.00 (1.71)	−2.45	0.0048	0.0055	0.4825
C7H19orf60	rs109332658	7	4533772	3.63 (0.89)	−0.48 (0.98)	0.70 (1.73)	−1.47	0.0077	0.1042	0.0300
CAST	rs137601357	7	98485273	3.99 (1.03)	0.58 (0.88)	−0.50 (1.22)	−2.25	0.0074	0.0017	0.2581
CD40	rs41711496	13	75567844	3.02 (1.09)	1.38 (0.88)	0.16 (1.20)	−1.43	0.0002	0.0472	0.8393
COQ9	rs109301586	18	25527339	−1.12 (1.08)	2.26 (0.89)	6.18 (1.12)	3.65	0.0187	<0.0001	0.7845
DNAH11	rs110629231	4	30766895	3.54 (0.92)	0.78 (0.91)	−0.63 (1.65)	−2.08	0.0047	0.0161	0.5497
DSC2	rs109503725	24	27124361	0.01 (1.04)	3.35 (0.89)	1.78 (1.30)	0.89	0.0033	0.2368	0.0176
DZIP3	rs133175991	1	53855896	1.51 (0.82)	2.95 (1.01)	−3.40 (2.92)	−2.45	0.0011	0.0966	0.0241
EPAS1	rs43676052	11	28650973	3.14 (0.88)	−0.95 (0.97)	−1.09 (2.06)	−2.11	0.0114	0.0461	0.1343
FST	rs109247499	20	25589648	0.92 (1.06)	1.53 (0.87)	4.49 (1.28)	1.78	0.0011	0.0162	0.2461
FUT1	rs41893756	18	55831611	2.55 (0.79)	0.88 (1.03)	−4.21 (2.83)	−3.38	0.0080	0.0185	0.3169
GOLGA4	rs42339105	22	10887536	0.55 (0.70)	7.67 (1.50)	15.10 (6.39)	7.28	0.0244	0.0232	0.9639
HSD17B12	rs109711583	15	74828355	−0.21 (1.03)	2.11 (0.88)	4.38 (1.26)	2.29	0.0073	0.0017	0.9829
HSD17B7	rs110828053	3	6630548	1.27 (0.81)	3.36 (1.08)	8.49 (2.55)	3.62	0.0047	0.0050	0.3429
IBSP	rs110789098	6	3809790	3.28 (0.98)	1.22 (0.89)	0.06 (1.42)	−1.16	0.0027	0.0412	0.6746
MRPL48	rs43703916	15	54119266	0.40 (0.96)	1.81 (0.91)	5.58 (1.41)	2.59	0.0083	0.0007	0.2653
OCLN	rs134264563	20	10167825	0.29 (0.92)	2.57 (0.90)	5.67 (1.65)	2.69	0.0036	0.0021	0.7160
PCCB	rs109813896	1	134130474	−0.27 (0.95)	2.93 (0.90)	4.44 (1.48)	2.35	0.0064	0.0031	0.4339
PMM2	rs109629628	25	7716425	−0.13 (0.96)	2.53 (0.91)	3.92 (1.33)	2.03	0.0041	0.0053	0.5335

[1]Data are least-squares means (standard error) for the predicted transmitting ability.
[2]A, additive; D, dominance.

Table 3 SNPs associated with heifer conception rate.

				Copies of minor allele[1]					Effect *P*–value[2]	
Gene	SNP id	Chr.	Loc.	0	1	2	Effect	r^2	A	D
AP3B1	rs133700190	10	9177305	3.79 (1.02)	6.48 (1.18)	16.95 (2.92)	6.58	0.0076	<0.0001	0.0347
BOLA-DMB	rs109032590	23	7189701	4.11 (1.08)	7.78 (1.16)	8.92 (2.30)	2.40	0.0031	0.0471	0.4243
C17H22orf25	rs133455683	17	74976374	7.31 (1.30)	5.23 (1.12)	3.38 (1.64)	−1.96	0.0005	0.0450	0.9325
CACNA1D	rs135744058	22	47726446	4.34 (1.12)	5.61 (1.14)	12.18 (2.20)	3.92	0.0061	0.0008	0.0799
CD40	rs41711496	13	75567844	7.33 (1.35)	5.37 (1.06)	3.56 (1.51)	−1.89	0.0005	0.0459	0.9508
CSNK1E	rs133449166	5	110565337	6.82 (1.17)	4.27 (1.11)	7.82 (1.92)	0.50	0.0008	0.6395	0.0349
DNAH11	rs110629231	4	30766895	7.07 (1.12)	4.75 (1.10)	2.10 (2.15)	−2.48	0.0020	0.0322	0.9148
FAM5C	rs135071345	16	16039905	2.57 (0.90)	5.91 (1.57)	25.75 (6.40)	10.24	0.0027	0.0014	0.0069
FST	rs109247499	20	25589648	4.64 (1.33)	5.49 (1.05)	8.77 (1.63)	2.07	0.0015	0.0362	0.3659
GCNT3	rs109830880	10	50709147	6.66 (0.96)	3.73 (1.44)	14.16 (4.85)	3.75	0.0017	0.1247	0.0166
HSD17B12	rs109711583	15	74828355	3.30 (1.29)	6.31 (1.08)	9.14 (1.61)	2.92	0.0059	0.0027	0.9483
HSD17B7	rs110828053	3	6630548	3.49 (0.96)	9.18 (1.34)	15.33 (3.35)	5.92	0.0098	0.0005	0.9111
IBSP	rs110789098	6	38309790	7.98 (1.19)	3.33 (1.07)	5.11 (1.81)	−1.44	0.0040	0.1689	0.0243
MRPL48	rs43703916	15	54119266	4.43 (1.17)	5.58 (1.10)	8.95 (1.81)	2.26	0.0025	0.0283	0.4367

[1]Data are least-squares means (standard error) for the predicted transmitting ability.
[2]A, additive; D, dominance.

PY (*PGR* and *TXN2*) and PPCT (*FCER1G*) (Cochran *et al.* 2013b).

For CCR, nine of the 33 SNPs examined by Cochran *et al.* (2013b) were associated with the trait (*ACAT2, AP3B1, ARL6IP1, CAST, COQ9, FUT1, HSD17B7, OCLN* and *PMM2*). In each case, the beneficial allele was the same in both studies. There were an additional 14 SNPs associated with CCR, and of these, 11 were associated with another fertility trait or trait involving fertility in the Cochran *et al.* (2013b) study including DPR (*BSP3, C7H19orf60,CD40, DSC2, HSD17B12* and *PCCB*), HCR (*DZIP3* and *GOLGA4*), PL (*IBSP*) and NM (*EPAS1* and *FST*).

Table 4 Comparison of the number of SNPs previously associated with daughter pregnancy rate (DPR), cow conception rate (CCR) and heifer conception rate (HCR).

Description of the SNP	Number of SNPs per trait		
	DPR	CCR	HCR
Previously associated with the trait[1]	39	33	22
Associated with each trait in the present study			
Same direction	15	9	4
Opposite direction	4	0	0
New significant	10	14	10
Total significant	29	23	14

[1]Associated with each trait by Cochran *et al.* (2013b).

Table 5 Changes in reliability by including UFL-SNP in the calculation of genomic estimate for daughter pregnancy rate.

SNP array	Reliability[1] (%)			Gain vs. USDA-SNP (%)
	PA	PA + Gen	Gain	
USDA-SNP	30.44	60.79	30.35	–
USDA-SNP + UFL-SNP	30.44	60.99	30.55	0.20
USDA-SNP (−100 kb) + UFL-SNP[2]	30.44	61.12	30.68	0.33
USDA-SNP (−200 Kb) + UFL-SNP[3]	30.44	61.17	30.73	0.38
UFL-SNP	30.44	33.20	2.76	–

[1]PA, parental average; Gen, genomic information; Gain, difference between the parental average and the parental average + genomic information.
[2]USDA-SNP panel without markers within 100 kb of UFL-SNP markers.
[3]USDA-SNP panel without markers within 200 kb of UFL-SNP markers.

Three were previously associated with production traits, including MY (*BOLA-DMB*), FY (*DNAH11*) and PY (*MRPL48*).

For HCR, four of the 22 SNPs found by Cochran *et al.* (2013b) were significantly associated with HCR (*AP3B1*, *CACNA1D*, *CSNK1E* and *HSD17B7*). For each gene, the allele associated with higher HCR was the same in both studies. In addition, 10 genes that were associated with HCR were previously associated with other traits including DPR (*C17H22orf25*, also known as *TANGO2; CD40;* and *HSD17B12*), PL (*IBSP*), NM (*FST*), MY (*BOLA-DMB* and *FAM5C*), FY (*DNAH11* and *MRPL48*) and PY (*GCNT3*) by Cochran *et al.* (2013b).

Reliability gain for DPR by including candidate gene SNPs in genomic estimates

The results for reliability gain calculation are presented in Table 5. For DPR, the reliability using parental averages was 30.44%. Inclusion of genomic information based on the 61 013 SNPs currently used for the US genomic evaluations (USDA-SNP) in the genetic evaluation increased reliability to 60.79%, a gain of 30.35%. When SNPS from the UFL-SNP panel were also added, the reliability gain was increased by a further 0.20% to 60.99%. Removal of markers in the USDA-SNP panel near the markers in the UFL-SNP panel (within both 100 and 200 kb) increased the reliability gain from adding SNPs from the UFL-SNP panel as compared to using the USDA-SNP panel alone by 0.33% (100 kb) 0.38% (200 kb). The increase in reliability of DPR through incorporation of the UFL-SNP panel alone was 2.76%, from 30.44 to 33.20%.

Pathway analysis

There were three significant canonical pathways in which at least two genes associated with DPR, CCR or HCR were represented. These were the cholesterol biosynthesis pathway (*ACAT2* and *HSD17B7*), estrogen biosynthesis pathway (*HSD17B7* and *HSD17B12*) and autoimmune thyroid disease signaling pathway (*CD40* and *FCER1G*). The biological functions represented by the largest numbers of genes were cellular growth and proliferation/cell morphology/cell death (*AP3B1*, *APBB1*, *ARL6IP1*, *CACNA1D*, *CAST*, *CD40*, *CSNK1E*, *EPAS1*, *FAM5C*, *FCER1G*, *FST*, *IBSP*, *LHCGR*, *OCLN*, *PGR*, *TBC1D24* and *TXN2*) and reproductive system development and function (*CACNA1D*, *CAST*, *CD40*, *DSC2*, *EPAS1*, *FAM5C*, *FCER1G*, *FST*, *HSD17B7*, *LHCGR*, *OCLN* and *PGR*).

Several upstream regulators of genes significantly associated with fertility traits were identified, including estradiol, which is an upstream regulator for 13 genes (*APBB1*, *BCAS1*, *CAST*, *CD40*, *FST*, *HSD17B12*, *HSD17B7*, *LHCGR*, *PMM2*, *OCLN*, *PGR*, *RABEP2* and *TXN2*); tumor necrosis factor, regulating nine genes (*CD40*, *DSC2*, *FCER1G*, *FST*, *HSD17B7*, *LHCGR*, *OCLN*, *RABEP2* and *TXN2*); as well as chorionic gonadotropin (*DNAH11*, *EPAS1*, *FST*, *HSD17B7*, *LHCGR* and *PGR*), progesterone (*CD40*, *EPAS1*, *FST*, *LHCGR*, *PGR* and *OCLN*) and transforming growth factor β1 (*CD40*, *DSC2*, *DZIP3*, *FCER1G*, *LHCGR* and *PCCB*), which each regulate six genes.

Discussion

In the present study, the effects of a large proportion of the SNPs previously associated with genetic merit for fertility traits were validated in a separate population of animals. Additionally, inclusion of these SNPs in genetic evaluations for DPR improved reliability. Many of the genes represented by the SNPs in this study are involved in steroidogenesis or are regulated by steroids. This latter supports the idea that ovarian steroidogenesis is important for cow fertility.

One characteristic of many GWAS and candidate gene studies is that associations between genotype and phenotype are not repeatable (Ioannidis *et al.* 2011). For example, Littlejohn *et al.* (2012) were not able to replicate the effect of any of 138 markers previously associated with residual feed

intake in dairy cattle. In the present study, however, many SNPs previously related to fertility traits (Cochran *et al.* 2013b) had similar relationships in a separate set of animals. Of the 39 SNPs previously found to be related to DPR by Cochran *et al.* (2013b), 19 were significantly associated with DPR in the present study. For 15 of the 19 genes, the beneficial allele was the same as that found by Cochran *et al.* (2013b). This result suggests that many of the candidate gene SNPs found by Cochran *et al.* (2013b) are likely to represent true causal variants.

Failure of many of the SNPs found by Cochran *et al.* (2013b) to have a significant effect on DPR in the present study may represent a combination of circumstances, including false positives as well as the fact that reliabilities of the cow population used in the present study are lower than the reliabilities for the bull population used by Cochran *et al.* (2013b). As anticipated, because CCR and HCR have lower reliabilities than does DPR (Cooper *et al.* 2014), fewer of the significant SNPs for the former two traits found by Cochran *et al.* (2013b) were significant in the present study.

A further indication of the utility of the candidate gene SNPs for DPR was the finding that these SNPs by themselves increased the reliability of estimates of DPR in the national herd by 2.8%. The increase in reliability may have been greater if more genotyped animals were available for imputation. In addition, the reliability of estimates from the use of the USDA-SNP panel could be improved by 0.20% with the inclusion of information from the UFL-SNP panel. This is a noteworthy gain when compared to the 0.4% gain for DPR reported by VanRaden *et al.* (2013) by increasing the number of random genetic markers by 300 000. In another study of Jerseys and Holsteins, there was no improvement in the accuracy of the genetic prediction for MY, FY and PY when a higher density SNP panel (800K) was compared to the 50K panel (Erbe *et al.* 2012; Cuyabano *et al.* 2014). The 0.2% improvement in reliability means that the UFL-SNP panel can provide useful information for estimating fertility above that achieved with the current USDA-SNP panel. This is despite the fact that some of the variation explained by the UFL-SNP panel is already explained by the USDA-SNP panel. This latter conclusion is based on the fact that the increase in reliability from the addition of the UFL-SNP panel to the USDA-SNP panel was 0.38% if SNPs from the USDA-SNP panel located 200 kb from a UFL-SNP were removed from the analysis (vs. 0.20% if the SNPs were not removed).

All of the SNPs have the same variance in the Bayes A model used in the US genomic evaluation system (Wiggans *et al.* 2011), and the shrinkage in such a model could prevent the 39 putative causal SNPs from contributing substantially to the reliabilities of the breeding values. One modeling strategy to avoid that problem is to fit the causal SNPs as fixed effects and to compute the direct genomic values as the sum of the random effects of the 61 013 markers currently used in the evaluation and the fixed effects of the 39 causal SNPs. Although this is computationally tractable in the current study, it is not feasible to scale that to approximately 30 traits for each of the six breeds that receive genomic evaluations because that would require the use of different sets of SNPs for each breed and trait.

Many of the genes related to fertility in the present study are involved in actions of steroid hormones, particularly estradiol and, to a lesser extent, progesterone. Two of the canonical pathways in which genes related to DPR, CCR or HCR were over-represented were involved in steroidogenesis–the cholesterol biosynthesis pathway (*ACAT2* and *HSD17B7*) and estrogen biosynthesis pathway (*HSD17B7* and *HSD17B12*). Furthermore, a total of 13 genes related to fertility traits are targets of estrogen regulation (*APBB1, BCAS1, CAST, CD40, FST, HSD17B12, HSD17B7, LHCGR, PMM2, OCLN, PGR, RABEP2* and *TXN2*), and six genes are regulated by progesterone (*CD40*, *EPAS1, FST, LHCGR, PGR* and *OCLN*). The lactating cow, which is less fertile than the non-lactating heifer (Pursley *et al.* 1997), experiences low circulating concentrations of estradiol at estrus and low progesterone concentrations during the luteal phase of the estrous cycle (Sartori *et al.* 2002). At least in part, reduced circulating concentrations of steroids reflects increased liver metabolism during lactation (Wiltbank *et al.* 2006). Administration of estradiol cypionate around the time of ovulation has been reported to increase pregnancy rate (Cerri *et al.* 2004; Emadi *et al.* 2014). In contrast, benefits of supplementation of cows with progesterone after insemination are small (Wiltbank *et al.* 2014). Taken together, these results implicate steroid hormones as critical determinants of fertility.

Genes for several SNPs related to fertility in this study have previously been implicated in variations in reproductive function. *CAST* has been previously linked to genetic merit for fertility in dairy cattle (Garcia *et al.* 2006). *DNAH11* and *FAM5C* are differentially regulated in the brain of animals exhibiting strong vs. weak estrus (Kommadath *et al.* 2011). *BOLA-DMB*, *CD40*, *EPAS1* and *FCER1G* are differentially regulated in the endometrium of pregnant vs. non-pregnant cows, whereas *BCAS1, C7H19orf60, MRPL48* and *RABEP2* are differentially regulated in the endometrium of lactating vs. non-lactating cows (Cerri *et al.* 2012). *PCCB* is differentially regulated in the liver during the transition period in the cow (Graber *et al.* 2010). *GOLGA4* and *FUT1* are differentially regulated in embryos produced from superovulated vs. unstimulated dams (Gad *et al.* 2011), *EPAS1* is upregulated in caruncular and chorioallantoic tissues of somatic cell nuclear transfer pregnancies (Hoffert-Goeres *et al.* 2007) and *CSNK1E* has been associated with regulation of β-catenin (Kim *et al.* 2010), which in turn is involved in the WNT signaling pathway and is regulated by progesterone in the endometrium (Satterfield *et al.* 2008). Some genes are linked to reproductive tract diseases, such as *CACNA1D*, involved in proliferation of endometrial cancer cells (Hao *et al.* 2015),

and *HSD17B7*, whose expression is higher in ovarian endometriosis lesions (Huhtinen *et al.* 2012).

Other genes associated with fertility play important roles in immune function, including *AP3B1*, which is involved in antigen processing and presentation (Sasai *et al.* 2010); *GCNT3*, which is necessary for antigen expression in mammalian cells (Nonaka *et al.* 2014); and *CD40* and *FCER1G*, which are members of the TNFR family that mediates actions of TNF in inflammation, immunity and cellular differentiation and death (Brenner *et al.* 2015). Immune function is likely to be closely connected to fertility. Indeed, cows that experience more than one postpartum disease are less fertile than are cows that have experienced zero or one disease (Santos *et al.* 2011).

In conclusion, a large proportion of SNPs previously associated with genetic estimates of fertility in Holstein bulls maintained their association in a separate population of cows. The inclusion of these genes in genetic evaluations can provide moderate improvement in the reliability of genomic estimates for fertility. Moreover, these genes point out the importance of steroids for optimal reproduction in dairy cattle.

Acknowledgements

Research was supported by Agriculture and Food Research Initiative Competitive Grant no. 2013-68004-20365 from the USDA National Institute of Food and Agriculture and a grant from the Southeast Milk Inc. Milk Check-off Program. Cole and Null were supported by appropriated project 1265-31000-096-00, 'Improving Genetic Predictions in Dairy Animals Using Phenotypic and Genomic Information', of the Agricultural Research Service of the United States Department of Agriculture.

The authors thank the dairies that formed part of the study: Davie 1 & 2 (Okeechobee, FL), University of Florida Dairy (Hague, FL), Genasci Brothers Dairy (Modesto, CA), Larson Dairy Barn 5 (Okeechobee, FL), Maddox and Ruann Dairies (Riverview, CA), North Florida Holsteins (Bell, FL), PH Dairies 1 & 2 Dairies (Winton, CA) and Shenandoah Dairy (Live Oak, FL). The authors also thank Geneseek for outstanding cooperation in completion of the study. Thanks are also extended to the following for assistance with obtaining samples and data: Dr. Alvaro Magalhaes, Lander Veterinary Clinic (Turlock, CA); and Kelly Buchanan, Serdal Dikmen, Kyle Dobbs, Javier Juarez, James Moss, Antonio Ruiz de King and Paula Tribulo from the University of Florida. The Council on Dairy Cattle Breeding (Reynoldsburg, OH) and the Cooperative Dairy DNA Repository (Columbia, MO) are acknowledged for providing data used in this study. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The USDA is an equal opportunity provider and employer. Authors declare no conflict of interest.

References

Abel K., Reneland R., Kammerer S., Mah S., Hoyal C., Cantor C.R., Nelson M.R. & Braun A. (2006) Genome-wide SNP association: identification of susceptibility alleles for osteoarthritis. *Autoimmunity Reviews* **5**, 258–63.

Averill T.A., Rekaya R. & Weigel K. (2004) Genetic analysis of male and female fertility using longitudinal binary data. *Journal of Dairy Science* **87**, 3947–52.

Boichard D. & Manfredi E. (1994) Genetic analysis of conception rate in French Holstein cattle. *Acta Agriculturae Scandinavica. Section A, Animal Science* **44**, 138–45.

Brenner D., Blaser H. & Mak T.W. (2015) Regulation of tumor necrosis factor signalling: live or let die. *Nature Reviews Immunology* **15**, 362–74.

Cerri R.L.A., Santos J.E.P., Juchem S.O., Galvão K.N. & Chebel R.C. (2004) Timed artificial insemination with estradiol cypionate or insemination at estrus in high-producing dairy cows. *Journal of Dairy Science* **87**, 3704–15.

Cerri R.L.A., Thompson I.M., Kim I.H., Ealy A.D., Hansen P.J., Staples C.R., Santos J.E.P. & Thatcher W.W. (2012) Effects of lactation and pregnancy on gene expression of endometrium of Holstein cows at day 17 of the estrous cycle or pregnancy. *Journal of Dairy Science* **95**, 5657–75.

Charlier C., Agerholm J.S., Coppieters W. *et al.* (2012) A deletion in the bovine *FANCI* gene compromises fertility by causing fetal death and brachyspina. *PLoS ONE* **7**, e43085.

Cochran S.D., Cole J.B., Null D.J. & Hansen P.J. (2013a) Single nucleotide polymorphisms in candidate genes associated with fertilizing ability of sperm and subsequent embryonic development in cattle. *Biology of Reproduction* **89**, 1–7.

Cochran S.D., Cole J.B., Null D.J. & Hansen P.J. (2013b) Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genetics* **14**, 49.

Cooper T.A., Wiggans G.R. & VanRaden P.M. (2014) Including cow information in genomic prediction of Holstein dairy cattle in the US. *Proceedings of 10th World Congress of Genetics Applied to Livestock Production*, Vancouver, Canada.

Cuyabano B.C., Su G. & Lund M.S. (2014) Genomic prediction of genetic merit using LD-based haplotypes in the Nordic Holstein population. *BMC Genomics* **15**, 1171.

Emadi S.R., Rezaei A., Bolourchi M., Hovareshti P. & Akbarinejad V. (2014) Administration of estradiol benzoate before insemination could skew secondary sex ratio toward males in Holstein dairy cows. *Domestic Animal Endocrinology* **48**, 110–8.

Erbe M., Hayes B.J., Matukumalli L.K., Goswami S., Bowman P.J., Reich C.M., Mason B.A. & Goddard M.E. (2012) Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *Journal of Dairy Science* **95**, 4114–29.

Gad A., Besenfelder U., Rings F. *et al.* (2011) Effect of reproductive tract environment following controlled ovarian hyperstimulation treatment on embryo development and global transcriptome profile of blastocysts: implications for animal breeding and human assisted reproduction. *Human Reproduction* **26**, 1693–707.

Garcia M.D., Michal J.J., Gaskins C.T., Reeves J.J., Ott T.L., Liu Y. & Jiang Z. (2006) Significant association of the *calpastatin* gene with fertility and longevity in dairy cattle. *Animal Genetics* **37**, 304–5.

Graber M., Kohler S., Kaufmann T., Doherr M.G., Bruckmaier R.M. & Van Dorland H.A. (2010) A field study on characteristics and diversity of gene expression in the liver of dairy cows during the transition period. *Journal of Dairy Science* **93**, 5200–15.

Hao J., Bao X., Jin B., Wang X., Mao Z., Li X., Wei L., Shen D. & Wang J. (2015) Ca2+ channel subunit α 1D promotes proliferation and migration of endometrial cancer cells mediated by 17β-estradiol via the G protein-coupled estrogen receptor. *The FASEB Journal*, **29**, 2883–93.

Hoffert-Goeres K.A., Batchelder C.A., Bertolini M., Moyer A.L., Famula T.R. & Anderson G.B. (2007) Angiogenesis in day-30 bovine pregnancies derived from nuclear transfer. *Cloning and Stem Cells* **9**, 595–607.

Huhtinen K., Desai R., Ståhle M., Salminen A., Handelsman D.J., Perheentupa A. & Poutanen M. (2012) Endometrial and endometriotic concentrations of estrone and estradiol are determined by local metabolism rather than circulating levels. *The Journal of Clinical Endocrinology & Metabolism* **97**, 4228–35.

Ioannidis J.P.A., Tarone R. & McLaughlin J.K. (2011) The false–positive to false–negative ratio in epidemiologic studies. *Epidemiology* **22**, 450–6.

Kadri N.K., Sahana G., Charlier C. *et al.* (2014) A 660-kb deletion with antagonistic effects on fertility and milk production segregates at high frequency in Nordic red cattle: additional evidence for the common occurrence of balancing selection in livestock. *PLoS Genetics* **10**, e1004049.

Khatib H., Huang W., Mikheil D., Schutzkus V. & Monson R.L. (2009a) Effects of signal transducer and activator of transcription (STAT) genes *STAT1* and *STAT3* genotypic combinations on fertilization and embryonic survival rates in Holstein cattle. *Journal of Dairy Science* **92**, 6186–91.

Khatib H., Huang W., Wang X., Tran A.H., Bindrim A.B., Schutzkus V., Monson R.L. & Yandell B.S. (2009b) Single gene and gene interaction effects on fertilization and embryonic survival rates in cattle. *Journal of Dairy Science* **92**, 2238–47.

Kim S.Y., Dunn I.F., Firestein R. *et al.* (2010) CK1ε is required for breast cancers dependent on β-catenin activity. *PLoS One* **5**, e8979.

Kommadath A., Woelders H., Beerda B., Mulder H.A., de Wit A.A., Veerkamp R.F., te Pas M. & Smits M. (2011) Gene expression patterns in four brain areas associate with quantitative measure of estrous behavior in dairy cows. *BMC Genomics* **12**, 200.

Littlejohn M., Grala T., Sanders K., Walker C., Waghorn G., Macdonald K., Spelman R., Davis S. & Snell R. (2012) Non-replication of genome-wide-based associations of efficient food conversion in dairy cows. *Animal Genetics* **43**, 781–4.

McClure M.C., Bickhart D., Null D. *et al.* (2014) Bovine exome sequence analysis and targeted SNP genotyping of recessive fertility defects BH1, HH2, and HH3 reveal a putative causative mutation in *SMC2* for HH3. *PLoS One* **9**, e92769.

Nonaka M., Fukuda M.N., Gao C., Li Z., Zhang H., Greene M.I., Peehl D., Feizi T. & Fukuda M. (2014) Determination of carbohydrate structure recognized by prostate-specific F77 monoclonal antibody through expression analysis of glycosyltransferase genes. *Journal of Biological Chemistry* **289**, 16478–86.

Pritchard T., Coffey M., Mrode R. & Wall E. (2013) Genetic parameters for production, health, fertility and longevity traits in dairy cows. *Animal* **7**, 34–46.

Pryce J.E., Royal M.D., Garnsworthy P.C. & Mao I.L. (2004) Fertility in the high-producing dairy cow. *Livestock Production Science* **86**, 125–35.

Pursley J.R., Wiltbank M.C., Stevenson J.S., Ottobre J.S., Garverick H.A. & Anderson L.L. (1997) Pregnancy rates per artificial insemination for cows and heifers inseminated at a synchronized ovulation or synchronized estrus. *Journal of Dairy Science* **80**, 295–300.

Santos J.E.P., Bisinotto R.S., Ribeiro E.S., Lima F.S., Greco L.F., Staples C.R. & Thatcher W.W. (2011) Applying nutrition and physiology to improve reproduction in dairy cattle. *Reproduction in Domestic Ruminants VII*, 387–403.

Sartori R., Rosa G.J.M. & Wiltbank M.C. (2002) Ovarian structures and circulating steroids in heifers and lactating cows in summer and lactating and dry cows in winter. *Journal of Dairy Science* **85**, 2813–22.

Sasai M., Linehan M.M. & Iwasaki A. (2010) Bifurcation of Toll-like receptor 9 signaling by adaptor protein 3. *Science*, **329**, 1530–4.

Satterfield M.C., Song G., Hayashi K., Bazer F.W. & Spencer T.E. (2008) Progesterone regulation of the endometrial WNT system in the ovine uterus. *Reproduction, Fertility and Development* **20**, 935–46.

Siontis K.C.M., Patsopoulos N.A. & Ioannidis J.P.A. (2010) Replication of past candidate loci for common diseases and phenotypes in 100 genome-wide association studies. *European Journal of Human Genetics* **18**, 832–7.

Sonstegard T.S., Cole J.B., VanRaden P.M., Van Tassell C.P., Null D.J., Schroeder S.G., Bickhart D. & McClure M.C. (2013) Identification of a nonsense mutation in *CWC15* associated with decreased reproductive efficiency in Jersey cattle. *PLoS One* **8**, e54872.

VanRaden P.M., Sanders A.H., Tooker M.E., Miller R.H., Norman H.D., Kuhn M.T. & Wiggans G.R. (2004) Development of a national genetic evaluation for cow fertility. *Journal of Dairy Science* **87**, 2285–92.

VanRaden P.M., Van Tassell C.P., Wiggans G.R., Sonstegard T.S., Schnabel R.D., Taylor J.F. & Schenkel F.S. (2008) Invited review: Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science* **92**, 16–24.

VanRaden P.M., Olson K.M., Null D.J. & Hutchison J.L. (2011) Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *Journal of Dairy Science* **94**, 6153–61.

VanRaden P.M., Null D.J., Sargolzaei M. *et al.* (2013) Genomic imputation and evaluation using high-density Holstein genotypes. *Journal of Dairy Science* **96**, 668–78.

Weller J.I. & Ron M. (2011) Invited review: quantitative trait nucleotide determination in the era of genomic selection. *Journal of Dairy Science* **94**, 1082–90.

Wiggans G.R., VanRaden P.M. & Cooper T.A. (2011) The genomic evaluation system in the United States: past, present, future. *Journal of Dairy Science* **94**, 3202–11.

Wiltbank M., Lopez H., Sartori R., Sangsritavong S. & Gümen A. (2006) Changes in reproductive physiology of lactating dairy cows due to elevated steroid metabolism. *Theriogenology* **65**, 17–29.

Wiltbank M.C., Souza A.H., Carvalho P.D., Cunha A.P., Giordano J.O., Fricke P.M., Baez G.M. & Diskin M.G. (2014) Physiological and practical effects of progesterone on reproduction in dairy cattle.

Animal: an international journal of animal bioscience **8**(Suppl 1), 70–81.

Yampolsky L.Y. & Stoltzfus A. (2005) The exchangeability of amino acids in proteins. *Genetics* **170**, 1459–72.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Summary statistics of PTA values for DPR, HCR and CCR in high and low DPR groups.

Table S2 Genes included in the UFL-SNP array.

Table S3 Allele frequencies of 69 SNPs tested for association with fertility traits.