



Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle¹

T. A. Cooper,² G. R. Wiggans, D. J. Null, J. L. Hutchison, and J. B. Cole

Animal Improvement Programs Laboratory, Agricultural Research Service, US Department of Agriculture (USDA), Beltsville, MD 20705-2350

ABSTRACT

Genomic evaluations of dairy cattle in the United States have been available for Brown Swiss, Holsteins, and Jerseys since 2009. As of January 2013, 1,023 Ayrshires had genotypes in the North American database. Evaluation accuracy was assessed using genomic evaluations based on 646 bulls with 2008 traditional evaluations to predict daughter performance of up to 180 bulls in 2012. Mean gain in reliability over parent average for all traits was 8.2 percentage points. The highest gains were for protein yield (16.9 percentage points), milk yield (16.6 percentage points), and stature (16.2 percentage points). Twelve single nucleotide polymorphisms were useful for Ayrshire breed determination. Fewer breed-determining SNP were available for Ayrshires than for Holsteins, Jerseys, and Brown Swiss because of the similarity of Ayrshires and Holsteins. A haplotype that affects fertility was identified on chromosome 17 and traces back in the genotyped population to the bull Selwood Betty's Commander (born in 1953). The haplotype carrier frequency for genotyped Ayrshires was 26.1%. Sire conception rate was decreased by 4.3 ± 2.5 percentage points for carriers of the haplotype as determined by 618 matings of carrier sire by carrier maternal grandsire. Genomic evaluations for Ayrshires were officially implemented in the United States in April 2013.

Key words: Ayrshire, breed determination, genomic evaluation, fertility haplotype

INTRODUCTION

Official genomic evaluations of dairy cattle in the United States have been available for Brown Swiss, Holsteins, and Jerseys since 2009 (Wiggans et al., 2011).

As of January 2013, over 1,000 Ayrshires had genotypes in the North American database. That amount of data justified investigation of possible benefits of genomic evaluation for the Ayrshire breed. Haplotypes that affect fertility have been identified for Brown Swiss, Holsteins, and Jerseys (VanRaden et al., 2011). Because of the heavy use of popular bulls in the Ayrshire breed, even that small population might have enough matings to uncover haplotypes never observed in the homozygous state. The objectives of this study were to evaluate the accuracy of genomic predictions, investigate SNP that might be useful for breed determination, and identify haplotypes that affect fertility.

MATERIALS AND METHODS

Genotypes determined with the Illumina BovineHD (Illumina Inc., 2010), Illumina BovineSNP50 (Illumina Inc., 2011a), GoldenGate Bovine3K (Illumina Inc., 2011b), Illumina BovineLD (Boichard et al., 2012), GeneSeek Genomic Profiler (Neogen Corporation, 2013a), and GeneSeek Genomic Profiler HD (Neogen Corporation, 2013b) BeadChips were available for 1,023 Ayrshires as of January 2013. Over half (525) of the genotypes were from Canadian Ayrshires that had been genotyped with the BovineHD BeadChip.

Genotypes were imputed to a common set of 45,188 SNP using the findhap.f90 program (VanRaden, 2011). The SNP set was the same as the one used for US genomic evaluation of Brown Swiss, Holsteins, and Jerseys in January 2013; that set had been chosen based on SNP performance criteria such as minor allele frequency, parent-progeny conflict, call rate, and correlation with other SNP (Wiggans et al., 2010). The percentages of SNP that were called after imputation (imputation rate) was determined by the number of called SNP out of the possible 45,188 SNP.

A traditional 4-yr cut-off study, which removed the last 4 yr of information and compared traditional parent averages and parent average reliability with values that included SNP and polygenic effects estimated from genomic information as described by VanRaden et al., (2009), was modified to include animals born after January 1, 2000, in the validation set and used

Received August 27, 2013.

Accepted February 4, 2014.

¹The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the USDA or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

²Corresponding author: Tabatha.Cooper@ars.usda.gov

traditional information from 4 yr ago. Two additional cut-off study methods were considered to determine gains in reliability due to the addition of genomic information. The first method was a traditional 4-yr cut-off study, which removed the last 4 yr of information and compared traditional parent averages and parent average reliability with values that included SNP and polygenic effects estimated from genomic information. The second method used current traditional data and randomly assigned a quarter of the population as the validation set, which was repeated 4 times and averaged across runs (Olson et al., 2011).

Individual SNP effects were computed using a training set of 646 bulls with traditional evaluations as of August 2008. The validation set of bulls was animals born on or after January 1, 2000, and had a traditional evaluation by December 2012. The number of bulls in the validation set varied by trait and ranged from 147 to 180. Gains in reliability were calculated as the difference between genomic reliability, which included SNP and polygenic effects estimated from the August 2008 predictor population (genotyped animals with traditional evaluations) and August 2008 traditional parent averages, and parent average reliability. Additional statistics included coefficient of determination (R^2) between August 2008 evaluations and December 2012 daughter deviations deregressed from traditional evaluations, coefficient for regression (b) of December 2012 daughter deviations on August 2008 genomic evaluations, and bias in genomic evaluation between December 2012 daughter deviation and August 2008 genomic evaluation.

Determination of SNP that could be used for breed identification was found by comparing allele frequencies by SNP for 869 Ayrshires, 7,728 Brown Swiss, 258,678 Holsteins, and 31,417 Jerseys. To qualify as an Ayrshire identifier SNP, the allele frequency had to be ≥ 0.90 monomorphic for Ayrshires and ≤ 0.30 monomorphic for Brown Swiss, Holsteins, and Jerseys for the same allele.

After imputation, haplotype segments of 600, 200, and 75 markers were examined to find those segments with the highest frequency in the Ayrshire population that were never homozygous (VanRaden et al., 2011). Based on 611 actual mating patterns that created the genotyped individuals in the population, the expected number of carrier animals was determined as the number of matings for carrier sire by carrier maternal grandsire (**MGS**) divided by 4 for the region.

A total of 22,253 Ayrshire breeding records were analyzed using December 2012 data for official USDA phenotypic evaluations for sire conception rate (Kuhn and Hutchison, 2008; Norman et al., 2008). The overall conception rate for the Ayrshire breed in December

2012 was $37 \pm 4.8\%$. An interaction for sire carrier status with MGS carrier status was added to the official model (Norman et al., 2012). To evaluate haplotype carrier effects on sire conception rate, 618 matings of carrier sire by carrier MGS were compared with 6,733 matings from noncarrier sires.

RESULTS AND DISCUSSION

Mean percentage of SNP that were called after imputation (Table 1) ranged from 78.9 to 99.7 depending on the genotype chip used. Animals genotyped with the Illumina BovineSNP50 BeadChip had the highest imputation rate because they had the greatest number of called SNP in common with the 45,188 SNP set used for genomic evaluation. Imputed dams had the lowest imputation rate because their genotypes were derived completely from parent and progeny genotypes. Imputation rates for Ayrshire genotypes were similar to those for simulated Holstein genotypes (VanRaden et al., 2013) for all chip types except imputed dams, which were lower for Ayrshires most likely because of fewer genotyped progeny and other relatives. All 45,188 SNP were used in imputation and SNP effect calculation to maintain consistency with the SNP set used for routine evaluation of Brown Swiss, Holstein, and Jersey even though differences in minor allele frequencies exist.

The 2 additional cut-off study methods considered to determine gains in reliability due to the addition of genomic information were not used. The first method, a traditional 4-yr cut-off study, did not include enough animals in the validation group. The second method, which used current traditional data and randomly assigned a quarter of the population as the validation set, used overlapping data, which caused the contribution from genomics to be overvalued.

Table 2 shows that the mean gain in reliability over parent average for all traits was 8.2 percentage points. The largest reliability gains were for protein yield (16.9 percentage points), milk yield (16.6 percentage points), and stature (16.2 percentage points). Coefficients of determination for prediction of December 2012 daughter deviations (deregressed from traditional evaluations) by August 2008 evaluations ranged from 0.05 to 0.29 for parent average and from 0.10 to 0.42 for genomic evaluation. Coefficients for regression of December 2012 daughter deviations on August 2008 genomic evaluations ranged from 0.84 to 1.31; none were greatly different from the expected regression coefficient of 1.00 when a representative sample of genotyped bulls is assumed (Mäntysaari et al., 2010). Bias, which was determined by subtracting August 2008 genomic evaluations from December 2012 daughter deviations, was close to 0 for all traits except yield and net merit. The negative

Table 1. Percentages of SNP that were called for Ayrshire animals after imputation to the 45,188 SNP used for US genomic evaluations in January 2013 by genotyping chip

Item	Animals (no.)	Mean (%)	SD (%)	Minimum (%)	Maximum (%)
Genotyping chip					
Illumina BovineSNP50 ¹	386	99.7	0.0	99.4	99.9
GoldenGate Bovine3K ¹	3	92.0	0.0	92.0	92.0
Illumina BovineHD ¹	525	99.5	0.0	99.0	99.5
Illumina BovineLD ¹	33	95.4	1.9	94.6	99.8
GeneSeek Genomic Profiler ²	4	95.0	0.0	95.0	95.0
GeneSeek Genomic Profiler HD ²	72	98.1	0.0	98.1	98.1
Imputed dams ³	19	78.9	4.7	71.0	87.1

¹Illumina Inc. (San Diego, CA).²Neogen Corp. (Lansing, MI).³Imputed from genotypes of parents and progeny.

bias, especially milk yield and net merit, indicated that genomic predictions overestimated daughter deviations, which is similar to what was observed in the Brown Swiss population when genomic evaluations were first introduced for the small population of genotyped animals (Wiggans et al., 2011).

Twelve SNP (Table 3) can be used for Ayrshire breed determination because they are almost monomorphic

(minor allele frequency of <10%) for Ayrshires, but <30% of Holstein, Jersey, and Brown Swiss cattle are monomorphic for those alleles. Fewer SNP qualify for breed determination of Ayrshires compared with Holsteins, Jerseys, and Brown Swiss, which have approximately 230 unique SNP each on the Illumina BovineSNP50 BeadChip. The smaller number of qualifying SNP results primarily from the small number of unique

Table 2. Reliabilities (REL) for August 2008 traditional parent averages and genomic evaluations¹ of young bulls without daughter information, coefficients of determination (R²) of August 2008 evaluations predicting December 2012 daughter deviations deregressed from traditional evaluations, coefficients for regression (*b*) of December 2012 daughter deviations on August 2008 genomic evaluations, and bias in genomic evaluation by trait

Trait	Validation bulls (no.) ²	REL (%)			R ²			
		Parent average	Genomic evaluation	Gain ⁴	Parent average	Genomic evaluation	<i>b</i>	Bias ³
Milk (kg)	180	35.0	51.6	16.6	0.15	0.28	1.05 ± 0.13	-87.8 ± 39.1
Fat (kg)	180	35.0	49.0	14.0	0.22	0.32	1.01 ± 0.11	-2.8 ± 1.5
Protein (kg)	180	35.0	51.8	16.9	0.29	0.42	1.17 ± 0.10	-3.7 ± 1.0
Fat (%)	180	35.0	47.0	11.9	0.25	0.35	1.30 ± 0.13	0.0 ± 0.0
Protein (%)	180	35.0	46.3	10.7	0.19	0.28	1.02 ± 0.12	0.0 ± 0.0
Net merit (\$)	180	31.5	45.1	13.6	0.14	0.23	1.26 ± 0.17	-50.3 ± 33.2
Productive life (mo)	171	26.1	30.3	4.2	0.07	0.10	1.16 ± 0.27	-0.3 ± 0.4
SCS	180	30.5	33.5	3.0	0.20	0.22	1.01 ± 0.14	0.0 ± 0.0
Daughter pregnancy rate (%)	147	27.9	28.0	0.1	0.15	0.15	1.01 ± 0.20	0.0 ± 0.2
Final score	175	26.2	26.1	-0.1	0.13	0.13	1.00 ± 0.19	-0.1 ± 0.1
Stature	177	31.3	47.5	16.2	0.23	0.36	1.03 ± 0.10	-0.3 ± 0.1
Strength	173	28.2	36.4	8.2	0.05	0.10	0.87 ± 0.20	-0.3 ± 0.1
Dairy form	172	25.9	34.0	8.0	0.06	0.12	0.92 ± 0.21	0.0 ± 0.1
Foot angle	166	24.2	24.1	-0.1	0.08	0.10	1.12 ± 0.29	-0.3 ± 0.1
Rear legs (side view)	174	28.1	37.5	9.4	0.06	0.13	0.87 ± 0.18	0.0 ± 0.1
Rump angle	176	30.6	37.2	6.6	0.21	0.26	1.06 ± 0.14	0.0 ± 0.1
Rump width	173	28.9	40.4	11.5	0.07	0.15	0.84 ± 0.15	-0.1 ± 0.1
Fore udder attachment	172	28.1	35.5	7.3	0.16	0.21	1.13 ± 0.17	0.0 ± 0.1
Rear udder height	176	27.9	32.9	5.0	0.14	0.17	0.88 ± 0.15	0.0 ± 0.1
Udder depth	174	30.2	35.0	4.8	0.22	0.26	1.03 ± 0.13	-0.1 ± 0.1
Udder cleft	172	25.8	33.3	7.5	0.09	0.12	1.09 ± 0.22	0.1 ± 0.1
Front teat placement	175	28.5	37.2	8.7	0.19	0.25	1.31 ± 0.17	0.0 ± 0.1
Teat length	175	29.2	34.3	5.1	0.06	0.10	1.03 ± 0.24	0.0 ± 0.1

¹Includes SNP and polygenic effects estimated from the August 2008 predictor population (genotyped animals with traditional evaluations) and August 2008 traditional parent averages.²Received traditional evaluation by December 2012.³December 2012 daughter deviation - August 2008 genomic evaluation.⁴Genomic REL - parent average REL.

Table 3. Single nucleotide polymorphisms used for Ayrshire breed determination by *Bos taurus* chromosomes and their homozygous genotype frequencies for Ayrshires, Brown Swiss, Holsteins, and Jerseys

SNP	Chromosome	Homozygous genotype frequency (%)							
		Ayrshire		Brown Swiss		Holstein		Jersey	
		AA	BB	AA	BB	AA	BB	AA	BB
rs42351867 (BTB-01196862)	1	0	94	21	28	21	28	49	10
rs42354762 (BTB-01197254)	2	0	100	39	14	58	5	84	1
rs110894651 (ARS-BFGL-NGS-119271)	5	0	94	45	11	39	14	27	22
rs29009927 (Hapmap54948)	5	98	0	9	50	13	40	11	44
rs41869408 (BTB-00707438)	6	0	93	73	2	29	21	60	5
rs42498573 (BTB-01377157)	8	91	0	25	25	22	28	14	39
rs41656027 (Hapmap43070-BTA-83139)	9	0	93	53	8	29	22	28	22
rs111012814 (ARS-BFGL-NGS-104500)	12	0	92	27	23	42	13	56	6
rs109713098 (ARS-BFGL-BAC-14435)	13	93	0	5	60	13	40	5	58
rs110914539 (ARS-BFGL-NGS-106716)	13	0	93	58	6	40	13	57	6
rs109436345 (ARS-BFGL-NGS-111777)	13	100	0	5	65	17	34	4	67
rs109646517 (ARS-BFGL-NGS-70175)	17	92	0	19	30	29	21	17	33

SNP found for Ayrshires compared with Holsteins. That lack of uniqueness may result from the common use of Dutch or Flemish cattle during the 1750s for the improvement of both the Ayrshire and Holstein breeds (Oklahoma State University, 2001).

A haplotype that affects Ayrshire fertility was discovered on *Bos taurus* chromosome 17 in the range of 65.9 to 66.2 Mbp. That haplotype, which was designated as AH1, was expected to have 14 homozygous animals (56 matings of carrier sire by carrier MGS divided by 4). However, none were observed in the genotyped population. In the genotyped Ayrshire population, the oldest ancestor with the AH1 haplotype was Selwood Betty's Commander (born in 1953), the most heavily used bull in the breed (VanRaden and Smith, 1999); however, the original mutation most likely did not occur in Selwood Betty's Commander. The carrier frequency of AH1 for genotyped Ayrshires is 26.1%, which has remained steady since the 1970s and is the most prevalent recessive haplotype discovered using this method to date (VanRaden et al., 2011). The haplotype was able to persist in the population with such high frequency because of the popularity of Selwood Betty's Commander and his descendants.

Sire conception rate was 4.3 ± 2.5 percentage points lower (Table 4) for AH1 carriers as determined by

matings of carrier sire with carrier MGS compared with noncarrier sire matings. The effect on stillbirth rate could not be determined because of limited calving data. Carrier status for AH1 can be integrated into a mating scheme to reduce the likelihood of carrier-to-carrier matings and decrease the loss in overall fertility.

CONCLUSIONS

Genomic evaluations for Ayrshire have been official in the United States since April 2013. Those evaluations have improved prediction accuracy over parent average, and evaluation reliability has increased by an average of 8.2 percentage points over all traits. Because the genomic evaluation methods were based on the North American population, they may not be suitable for all red dairy cattle because of the differences in linkage disequilibrium in red dairy populations. Additional information extracted from the genotypes of the Ayrshire population identified a group of SNP unique to the breed that can be used in breed identification. Because of the high prevalence of the deleterious AH1 haplotype in the Ayrshire population, knowledge of carrier animals can be used in breeding decisions to decrease fertility losses.

Table 4. Differences¹ (percentage points) between mating groups based on carrier status for haplotype AH1 for sire conception rate (SCR)

Mating group	Matings (no.)	Difference (%)
Carrier sire × carrier MGS	618	-4.3 ± 2.5
Noncarrier sire × noncarrier MGS or MGS with unknown carrier status	6,733	0.0 ± 2.4
Unknown carrier status for both sire and MGS	14,902	-0.4 ± 1.2

¹Mating group value minus matings with a noncarrier sire and noncarrier maternal grandsire (MGS) or MGS with unknown carrier status.

ACKNOWLEDGMENTS

The cooperation of the Council on Dairy Cattle Breeding (Reynoldsburg, OH) in supplying pedigree, performance, and genotypic data is acknowledged.

REFERENCES

- Boichard, D., H. Chung, R. Dasonneville, X. David, A. Eggen, S. Fritz, K. J. Gietzen, B. J. Hayes, C. T. Lawley, T. S. Sonstegard, C. P. Van Tassell, P. M. VanRaden, K. A. Viaud-Martinez, and G. R. Wiggans. 2012. Design of a bovine low-density SNP array optimized for imputation. *PLoS ONE* 7:e34130.
- Illumina Inc. 2010. BovineHD Genotyping BeadChip. Accessed Nov. 25, 2013. http://www.illumina.com/Documents/products/datasheets/datasheet_bovineHD.pdf.
- Illumina Inc. 2011a. BovineSNP50 Genotyping BeadChip. Accessed Nov. 25, 2013. http://www.illumina.com/Documents/products/datasheets/datasheet_bovine_snp50.pdf.
- Illumina Inc. 2011b. GoldenGate Bovine3K Genotyping BeadChip. Accessed Nov. 25, 2012. http://www.illumina.com/Documents/products/datasheets/datasheet_bovine3K.pdf.
- Kuhn, M. T., and J. L. Hutchison. 2008. Prediction of dairy bull fertility from field data: Use of multiple services and identification and utilization of factors affecting bull fertility. *J. Dairy Sci.* 91:2481–2492.
- Mäntysaari, E., Z. Liu, and P. VanRaden. 2010. Interbull validation test for genomic evaluations. *Interbull Bull.* 41:17–22.
- Neogen Corporation. 2013a. GeneSeek Genomic Profiler for dairy cattle. Accessed Dec. 4, 2013. http://www.neogen.com/Agrigenomics/pdf/Slicks/GGP-LD_Dairy.pdf.
- Neogen Corporation. 2013b. GeneSeek Genomic Profiler HD for dairy cattle. Accessed Dec. 4, 2013. http://www.neogen.com/Agrigenomics/pdf/Slicks/GGP_HD_Dairy.pdf.
- Norman, H. D., J. L. Hutchison, and J. R. Wright. 2008. Sire conception rate: New national AI bull fertility evaluation. *AIPL Res. Rep. SCR1(7-08)*. Accessed Nov. 18, 2013. <http://aipl.arsusda.gov/reference/arr-scr1.htm>.
- Norman, H. D., R. H. Miller, J. R. Wright, J. L. Hutchison, and K. M. Olson. 2012. Factors associated with frequency of abortions recorded through Dairy Herd Improvement test plans. *J. Dairy Sci.* 95:4074–4084.
- Oklahoma State University. 2001. Breeds of livestock, Ayrshire. Accessed Nov. 26, 2013. <http://www.ansi.okstate.edu/breeds/cattle/ayrshire/index.htm>.
- Olson, K. M., P. M. VanRaden, M. E. Tooker, and T. A. Cooper. 2011. Differences among methods to validate genomic evaluations for dairy cattle. *J. Dairy Sci.* 94:2613–2620.
- VanRaden, P. 2011. findhap.f90. Accessed Nov. 26, 2013. <http://aipl.arsusda.gov/software/findhap/>.
- VanRaden, P. M., D. J. Null, M. Sargolzaei, G. R. Wiggans, M. E. Tooker, J. B. Cole, T. S. Sonstegard, E. E. Connor, M. Winters, J. B. C. H. van Kaam, A. Valentini, B. J. Van Doormaal, M. A. Faust, and G. A. Doak. 2013. Genomic imputation and evaluation using high-density Holstein genotypes. *J. Dairy Sci.* 96:668–678.
- VanRaden, P. M., K. M. Olson, D. J. Null, and J. L. Hutchison. 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J. Dairy Sci.* 94:6153–6161.
- VanRaden, P. M., and L. A. Smith. 1999. Selection and mating considering expected inbreeding of future progeny. *J. Dairy Sci.* 82:2771–2778.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92:16–24.
- Wiggans, G. R., P. M. VanRaden, L. R. Bacheller, M. E. Tooker, J. L. Hutchison, T. A. Cooper, and T. S. Sonstegard. 2010. Selection and management of DNA markers for use in genomic evaluation. *J. Dairy Sci.* 93:2287–2292.
- Wiggans, G. R., P. M. VanRaden, and T. A. Cooper. 2011. The genomic evaluation system in the United States: Past, present, future. *J. Dairy Sci.* 94:3202–3211.