

Understanding Genomics and Its Applications on a Commercial Dairy Farm

Dr. Kent Weigel

University of Wisconsin

INTRODUCTION

In the past 2 years, more than 33,000 North American dairy cattle have been genotyped using the Illumina BovinSNP50 BeadChip. This technology, which became possible due to recent sequencing of the bovine genome, was developed in a partnership between Illumina Inc., the USDA Agricultural Research Service, the National Association of Animal Breeders, Merial Inc., and researchers at several universities and institutes.

A key breakthrough provided by this technology is the ability to carry out 54,000 DNA marker tests simultaneously, for a cost of about \$225 per animal. Single nucleotide polymorphism (SNP) markers represent single base changes (A, T, C, or G) within the DNA sequence of a cow or bull – a sequence that consists of approximately 3 billion base pairs distributed over 30 pairs of chromosomes. These SNP markers can be genotyped in an efficient and automated manner, in contrast to the labor-intensive (one at a time) genotyping of microsatellite markers.

Another key breakthrough is the finding that, once a large number of more or less evenly spaced genetic markers (e.g., at least 30,000) become available for an individual animal, it is possible to estimate the breeding value of that animal based on associations between marker genotypes and milk yield, somatic cell score, productive life, daughter pregnancy rate, and other key traits. These associations are estimated using data from the animal's ancestors, in particular the progeny tested bulls

represented in an animal's pedigree. The key animals in this process have been the dairy bulls represented in the Cooperative Dairy DNA Repository (CDDR), which was formed nearly 15 yr ago, when ABS Global, Accelerated Genetics, Alta Genetics, Genex Cooperative, Select Sires, Semex, and Taurus Service began storing semen samples from young bulls entering their progeny testing programs for the purpose of future research.

Although it sounds mysterious, genomic selection is actually rather simple. In the past, all we knew about a young animal's genetic potential was its parent average (PA), which was simply the average predicted transmitting ability (PTA) of its parents; and we had no way to determine whether this young animal got a better than average or poorer than average sample of genes from its parents. We also had no choice but to wait 2 yr until we could measure the animal's performance, in the case of females, or wait 5 yr until we could measure the performance of the animal's progeny, in the case of males.

Now, because the relationships between SNP markers and important functional genes that we observe in an animal's ancestors are maintained for several generations (before recombination breaks down these genetic links), we can glimpse into the crystal ball to see what the future holds for a particular young animal. The genomic PTA became the official genetic evaluation for US Holsteins and Jerseys in January 2009, and

official results for Brown Swiss became available in August 2009.

GENOTYPIC EVALUATIONS

In a recent project by scientists at the USDA-ARS Beltsville Agricultural Research Center, a total of 5,369 Holstein bulls and cows that were born from 1952-1999 were genotyped with the BovinSNP50 BeadChip (VanRaden et al., 2009; Cole et al., 2009). Genotypic and phenotypic data of these bulls were used to estimate the effects of 38,416 SNP markers (after discarding markers with low minor allele frequency and markers that were in complete linkage disequilibrium with adjacent markers) on production, type,

longevity, udder health, and calving ability. Next, the estimated SNP effects were used to compute the genomic PTA of each of 2,035 young Holstein bulls born from 2000-2003 that had no progeny. Finally, the 2009 PTA of each bull in the latter group, which was based on information from its progeny, was compared with the traditional PA and the genomic PTA computed from 2004 data. The same process was repeated in the Jersey breed (1,361 older animals and 388 young bulls) and the Brown Swiss breed (512 older animals and 150 young bulls). Results in Table 1 show the increase in reliability (**REL**) due to genomic information, as compared with the REL from parent average information only.

Table 1. Reliability changes due to the inclusion of genomic data in national genetic evaluations (VanRaden et al., 2009).

Trait	Increase in Reliability Due to Genomics		
	Holstein	Jersey	Brown Swiss
	----- % -----		
Net merit	+24	+8	+9
Milk yield	+26	+6	+17
Fat yield	+32	+11	+10
Protein yield	+24	+2	+14
Fat percentage	+50	+36	+8
Protein percentage	+38	+29	+10
Productive life	+32	+7	+12
Somatic cell score	+23	+3	+17
Daughter pregnancy rate	+28	+7	+18
Final classification score	+20	+2	+5
Udder depth	+37	+20	+8
Foot angle	+25	+11	-1

As shown in Table 1, gains in REL from genomic information were significant for almost all traits and breeds, ranging from -1 % for foot angle in Brown Swiss to +50 % for fat percentage in Holsteins. Gains were largest for traits with previously discovered major genes, such as fat percentage (DGAT1 on chromosome 14; Grisart et al., 2004) and protein percentage (ABCG2 on chromosome 6; Cohen-Zinder et al., 2005). For each trait, we can combine a young animal's PA with information from the BovinSNP50 BeadChip to obtain a genomic PTA of much greater accuracy.

For a heifer calf, REL of the genomic PTA is equivalent to the information we could obtain by measuring several lactation records on the animal and its daughters. For a young cow, information from the BeadChip can be combined with lactation records to obtain a genomic PTA that is significantly more informative than its traditional PTA.

For a bull calf, REL of the genomic PTA is equivalent to what we could obtain by measuring performance on 25 or 30 progeny test daughters. Improvements in accuracy can even be obtained for bulls that have completed progeny testing, although the gain in information for a bull that already has performance data from 80 to 100 daughters is much smaller. Gains in REL for Jerseys and Brown Swiss have not been as large as for Holsteins. However, this poor performance is largely due to the fact that fewer progeny tested bulls have been genotyped, and results for these breeds could be improved by combining information from North American sires with that of key populations internationally.

IMPACT OF GENOMICS

What has been the impact of genomics on the AI industry? The AI studs are already in the midst of tremendous change because of this technology. Virtually every young bull entering a North American AI center today is tested with the BovinSNP50 BeadChip, and DNA testing of potential bull dams is rapidly becoming the norm. The genomic PTA for a genotyped young bull typically has REL in the range of 60-80 %, as opposed to only 30-40 % or so for its traditional PA. The success rate (i.e., graduate rate) in progeny testing programs, which is currently about 1 in 10, will increase significantly in the future, because we will know prior to entry into the AI stud that each young bull has received a favorable sample of genes from its parents.

Progeny testing has been the cornerstone of the dairy cattle breeding industry for nearly a half century, and anything that competes with progeny testing in terms of accuracy will have an enormous impact. Within the next year, it is likely that parentage verification via DNA testing will become much more widespread, so the accuracy of a traditional PTA based on progeny test daughters will increase due to fewer misidentified daughters. The PTA of a bull that already has first-crop progeny test daughters will change slightly based on genomic information, but the largest impact will be for a young bull that doesn't yet have any progeny.

The North American AI centers are now marketing semen from hundreds of young bulls that have genomic PTAs, but no daughters of their own. These genome tested bulls tend to replace older proven bulls that were at the lower end of the sire line-up, and some young bulls with

outstanding genomic evaluations are being used for contract matings.

Over time, as AI centers and producers become more comfortable with this technology, we will see a decline in progeny testing, because its purpose is the same as that of the BeadChip – to see which young bull got the best sample of genes from its parents. Eventually, some of the additional genotyping costs will be offset by a lowered investment in progeny testing (e.g., bull housing, semen distribution, incentive payments, etc.); but because this technology has been available for only 24 months, most studs aren't ready to pull the plug on progeny testing just yet.

What will be the impact of genomics on pedigree breeders who are merchandising breeding stock? Breeders who are selling young bulls to AI are already seeing an impact, because the AI studs are asking to genotype the young bulls, and potentially their dams as well. Based on this initial genomic screening, many bull dams and young bulls are rejected. Conversely, the price paid for young bulls that pass this initial genomic screening is often higher. Furthermore, because sire analysts now have the ability to distinguish between sets of full brothers that all have the same PA, the premium for securing first choice from a flush is much greater. For these reasons, young bull contracts have been affected.

The impact on the female side, whether selling embryos or live heifers, will be similar. Potential buyers of embryos may want to know the genomic PTA of the dam, and buyers of live calves or heifers may want to genotype an animal before completing the purchase. Genomic information may be desired at consignment sales as well, and as buyers begin to understand this information they will pay a premium for young animals with favorable genotypes. Lastly, buyers of first or second

choice from a flush may pay more at an auction, knowing that they can genotype the resulting calves before making a decision.

What will be the impact of this project on commercial producers? These producers will not see a huge impact immediately, but several things are likely to happen. Producers now see semen that is marketed based on genomic breeding values. These may be yearling bulls that have no progeny, or they may be three-year-old bulls that have calving ease information but have no milking daughters. These bulls have attractive pedigrees (because they're younger than the current crop of proven bulls); but their reliabilities are lower, mostly in the 60-80 % range rather than the 80-90 % range typically seen for a first-crop proven bull. When buying semen from genome tested bulls, producers should avoid heavy use of 1 or 2 top bulls and instead spread out their risk by using a larger group of bulls.

Cooperator herds may see a decline in progeny testing activity, with a corresponding reduction in progeny test semen and incentive payments. This won't happen immediately, but over time it is likely that studs will test fewer bulls, and the semen of young bulls with genomic information will be sold at a much higher price than the semen of young bulls offered today.

COMPARISON OF GENOMIC AND TRADITIONAL EVALUATION

Recently, we took an early peek at how well genomic evaluations are working in practice, based on results from the first batch of bulls that had genomic data in January 2009 and at least 50 milking daughters in August 2009. In routine genetic evaluations, genomic data are combined with pedigree data when computing the official PTA that are released to the industry. Therefore,

Table 2. Comparison of January 2009 parent averages (PA) and genomic predicted transmitting abilities (GPTA) for milk, fat, protein, somatic cell score (SCS), and daughter pregnancy rate (DPR) with August 2009 daughter yield deviations (DYD) for bulls whose first-crop daughters calved between January and August.

	Milk	Fat	Prot	SCS	DPR
No. bulls	238	238	238	237	60
Reliability (Jan '09 PA), %	42	42	42	39	36
Reliability (Jan '09 GPTA), %	72	72	72	67	62
No. daughters (Aug '09 DYD)	71	71	71	71	62
Reliability (Aug '09 DYD), %	84	84	84	67	52
Correlation (Jan '09 PA, Aug '09 DYD)	.444	.540	.476	.376	.213
Correlation (Jan '09 GPTA, Aug '09 DYD)	.624	.695	.632	.531	.341

genomic information still has a considerable impact on the official PTA of a bull or cow, even after that animal has records or progeny of its own. Instead of using the official August 2009 genetic evaluations for bulls in this study, we used unofficial, *traditional* August 2009 genetic evaluations that did not contain any genomic information. In fact, we actually used daughter yield deviations (**DYD**) of these bulls, so that we could also escape the influence of pedigree information.

As shown in Table 2, a total of 238 Holstein bulls had official PTA in January 2009 that were based only on genomic information, as well as daughter yield deviations in August 2009 that included at least 50 milking daughters. Note that only 60 bulls had at least 50 daughters in their genetic evaluations for daughter pregnancy rate, and because none of these bulls had productive life data from at least 50 daughters in August, that trait could not be considered in our study. Average January 2009 REL based on PA information was 42 % for yield traits, 39 % for somatic cell score, and 26 % for DPR; whereas REL of

the genomic PTA, which include both pedigree and genomic information, averaged 72, 67, and 62 %, respectively. Data from an average of 71 daughters per bull (62 for DPR) provided an average REL of 84 % for yield traits, 67 % for SCS, and 62 % for DPR in August 2009.

It is interesting to note that the average REL of January 2009 GPTA for SCS was equivalent to that of August 2009 evaluations based on 71 milking daughters; whereas average REL of the August 2009 evaluations for DPR was still considerably lower than REL of the corresponding GPTA in January. This illustrates the challenge in improving lowly heritable health and fertility traits through genetic selection – a situation that may or may not be improved by genomics, as discussed later. It is also very important to note that August 2009 DYD of the bulls in this study are somewhat preliminary, as significant changes can occur when more records from first- and second-crop daughters become available. The correlations between August 2009 DYD resulting from progeny testing and January 2009 PA and GPTA for each trait are also

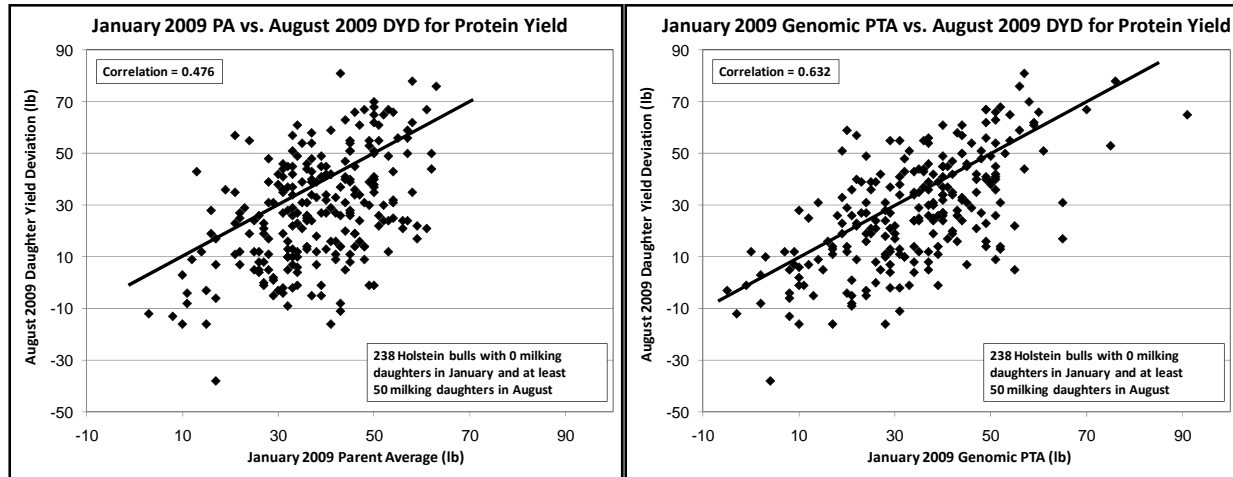


Figure 1. Scatter plot of January 2009 parent average (PA) based on pedigree information (graph on left) and genomic predicted transmitting ability (GPTA) based on DNA testing (graph on right) with August 2009 daughter yield deviation (DYD) for protein yield.

shown in Table 2. In all cases, correlations were much higher with the inclusion of genomic information, which seems to indicate better overall agreement when genomic information is used.

Scatter plots of the January 2009 PA and GPTA for protein yield are shown in Figure 1, along with the corresponding August 2009 DYD for these traits. As you can see, there is generally good agreement between genomic predictions and progeny test results, as we would expect by correlations of 0.632 for protein. Bulls that lie above the line in Figure 1 were under-predicted in January 2009, such that their August 2009 DYD were better than expected, and bulls that lie below the line were over-predicted. On the other hand, agreement between PA information and DYD from progeny testing was poorer, as one would expect based on a correlation of 0.476.

A large proportion of the bulls have evaluations within a few pounds of their genomic predictions, but some wide misses

occur as well. This phenomenon is not unique to genomic evaluations, as genetic predictions based on limited information have occasionally missed the mark for decades, regardless of the method or approach that was used.

The important thing for producers is to recognize that, although the GPTA for a young animal will usually be much more accurate than its traditional PA, it will generally be less accurate than information that comes later via progeny testing. Therefore, producers should limit the amount of semen that is used from an individual bull with high-ranking GPTA, and instead focus on choosing a larger group of bulls (perhaps 3 or 4 times as many as they've chosen in the past) to control the risks associated with lower REL bulls.

What will be the impact of genomics on academic researchers? As one would expect, the aforementioned developments have led to an explosion in dairy genomics research. Many new and useful tools, some

of which cannot yet be imagined, will be developed in the next decade based on genome-related technologies. A topic of keen interest is the development of a low-cost system that can be used for widespread genotyping of animals on commercial farms (e.g., replacement heifers). Such a product could also be used for screening of outcross families to find elite males or females that can offer unique genetic contributions to the breed.

Another topic of interest is the development of mating programs that are based on the results of DNA testing, rather than pedigree or conformation data. Lastly, the development of genotype-guided management programs is an interesting long-term possibility. Genomics has created a great deal of excitement in the dairy cattle breeding industry, but both producers and AI studs are still relatively low on the learning curve with respect to understanding this technology and maximizing its impact.

CONCLUSIONS

In summary, it is clear that genomic information enhances the accuracy of genetic evaluations for yield traits and, to a lesser extent, for health and fertility traits. As of August 2009, North American breeding companies were marketing 365 young Holstein bulls based solely on their genomic information, in contrast to 697 proven Holstein bulls with data from 10 or more milking daughters. However, when we focus on bulls with PTA for Lifetime Net Merit of \$450 or greater, the list is dominated by genome-tested bulls – a total of 306 *G status* (genomic) bulls exceed \$450 for Net Merit, whereas only 138 *A status* (active) bulls met this criterion. So, using

young, genome-tested bulls comes with increased risk, but ignoring these bulls comes with a heavy opportunity cost.

As with most good things, moderation is the key. Producers who supplement their traditional sire selections with a group of outstanding genome-tested bulls (each used in moderation) will achieve the greatest genetic progress in their herds.

LITERATURE CITED

- Cole, J. B., P. M. VanRaden, J. R. O'Connell, C. P. Van Tassell, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and G. R. Wiggans. 2009. Distribution and location of genetic effects for dairy traits. *J. Dairy Sci.* 92:2931-2946.
- Cohen-Zinder, M., E. Seroussi, D. M. Larkin, J. J. Loor, A. Everts-van der Wind, J. H. Lee, J. K. Drackley, M. R. Band, A. G. Hernandez, M. Shani, H. A. Lewin, J. I. Weller, and M. Ron. 2005. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res.* 15:936-944.
- Grisart B., F. Farnir, L. Karim, N. Cambisano, J. J. Kim, A. Kvasz, M. Mni, P. Simon, J. M. Frere, W. Coppieters, and M. Georges. 2004. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc. Natl. Acad. Sci. U.S.A.* 101:2398-2403.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. Schenkel. 2009. Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92:16-24.

*The High Plains Dairy Conference does not support one product over another
and any mention herein is meant as an example, not an endorsement.*