Producer uptake: how might genomic information get translated into industry outcomes

Alison L. Van Eenennaam, PhD; **Bovine Respiratory Disease Complex Coordinated Agricultural Project**[a]
Department of Animal Science, University of California, Davis, CA 95616

Abstract

The Bovine Respiratory Disease Complex Coordinated Agricultural Project (BRD CAP) is a 5-year USDA-funded Coordinated Agricultural Project to develop genetic markers associated with bovine respiratory disease (BRD) to identify cattle that are less susceptible to BRD. Ultimately the aim of this project is to integrate predictive markers for BRD susceptibility into genetic tests and national cattle genetic evaluations. Work is ongoing to identify regions of the genome associated with BRD susceptibility in both dairy and beef cattle. Initial results have identified multiple genomic regions that were significantly associated with BRD susceptibility. Genomic information has typically been integrated into genetic evaluations based on linkage disequilibrium (LD) between single nucleotide polymorphism (SNP) markers and the causative mutations affecting economically relevant traits. This has been successful to develop selection within breeds in the dairy cattle industry, but has been more problematic in the beef cattle industry due to the presence of numerous breeds and the importance of crossbreeding in the commercial cattle population. LD between markers and QTL has not be consistent across breeds, and so markers that were identified in 1 breed were frequently uninformative in other breeds. However, the sequencing of a large number of animals has opened up the possibility of identifying the actual SNP variations that are causing genetic variation and performing sequenced-based genomic selection of cattle. There are several advantages associated with this approach including persistence of the marker affect across generations, and an increased likelihood that causative polymorphisms will be similarly associated with variation across multiple breeds. It is envisioned that by imputing the genotypes of reference animals collected by the BRD CAP up to full sequence and further fine mapping and analyses, that the causative genetic variants associated with BRD susceptibility will be identified, and that inclusion of these markers on genotyping platforms will provide a reliable selection criterion to enable for the selection of both beef and dairy cattle that are less susceptible to BRD.

Key words: bovine, cattle, genomics, BRD

Résumé

La maladie respiratoire bovin coordonné complexe projet agricole (BRD CAP) est une 5-an USDA-financé projet agricole coordonnée pour développer des marqueurs génétiques associés à une maladie respiratoire chez les bovins (BRD) à l'identification des bovins qui sont moins sensibles à BRD. En fin de compte le but de ce projet est d'intégrer les marqueurs prédictifs pour la susceptibilité en tests génétiques BRD et évaluations génétiques bovin national. Des travaux sont en cours pour identifier les régions du génome associées à BRD susceptibilité dans les deux vaches laitières et bovins de boucherie. Les premiers résultats ont identifié plusieurs régions du génome qui sont significativement associées à la BRD susceptibilité. L'information génomique a généralement été intégrées dans les évaluations génétiques basées sur déséquilibre de liaison (DL) entre le polymorphisme de nucléotide simple (SNP) des marqueurs, et les mutations affectant causal économiquement les traits pertinents. Cela a réussi à développer sélection au sein de niche dans l'industrie des bovins laitiers, mais a été plus problématique dans l'industrie des bovins de boucherie en raison de la présence de nombreuses races et l'importance de métissage dans la population bovine commerciale. LD entre les marqueurs et QTL n'a pas être cohérents dans l'ensemble de races, et ainsi de marqueurs qui ont été identifiés dans une race étaient souvent peu éclairantes dans d'autres races. Toutefois, le séquençage d'un grand nombre d'animaux a ouvert la possibilité d'identifier concrètement l'SNP de variations qui se causant la variation génétique et en effectuant la sélection génomiques à base séquencés de bovins. Il y a plusieurs avantages associés à cette approche, y compris la persistance du marqueur affectent à travers les générations, et une probabilité accrue que polymorphismes causal sera également associée à la variation entre plusieurs races. Il est prévu qu'en imputant les génotypes des animaux de référence recueillies par le BRD PAC jusqu'à une séquence complète et plus loin la cartographie fine et les analyses, que la cause des variants génétiques associés à la susceptibilité BRD seront identifiés, et que l'inclusion de ces marqueurs sur les plates-formes de génotypage fournira un critère de sélection fiable à activer pour la sélection de viande de boeuf et de bovins laitiers qui sont moins sensibles à BRD.

Introduction

The use of DNA testing to predict the genetic merit of dairy and beef cattle has become common place since the introduction of the 50,000 single nucleotide polymorphism (SNP) chip in 2009. The dairy industry rapidly adopted the technology, and as of 2015 the dairy industry is on track to

have run genomic tests on over a million animals. To date, the accurate prediction of genetic merit using genomic information has been limited to within breed predictions. This means that every breed needs to develop its own reference population of at least several hundred phenotyped, genotyped animals to develop accurate genomic prediction equations[9]. Additionally, equations need to be continually updated or refreshed each generation to remain accurate. As a result, the promise of genomic technology has been realized to a much more modest extent in the beef industry, in part due to the presence of a number of different beef breeds, many with limited population sizes, necessitating the development of separate reference populations for each breed.

The traits that can be predicted using genomics are restricted to those for which phenotypes are available in the reference population. As such, all of the traits that breed associations typically collect data on were available to enable the development of genomic predictions for those traits. However, these are the "easy" traits – in that they are often relatively easy to measure, and genetic merit estimates were already available on these traits prior to the development of genomics. There are many valuable and economically-relevant traits for which there are no records and hence no selection criteria, and hence they are not currently included in genetic improvement programs. Typically these are hard or expensive to measure traits, or those that are collected late in an animal's life.

In the dairy industry researchers have variously investigated selection for a number of novel phenotypes including milk fatty acid composition, persistency of lactation, rectal temperature, residual feed intake and feed efficiency, methane production, claw health, immune response, and dairy cattle health. The beef industry is also examining the potential to select for feed efficiency, improved fertility, and cattle health. All of these novel phenotypes are not routinely recorded, and for some of them (e.g., individual feed intake) the cost of routine recording is cost prohibitive.

Bovine respiratory disease (BRD) complex is a trait that falls into this category. Despite the fact that this disease is the leading cause of mortality in both the beef and dairy industries nationally, routine recording of disease incidence is not currently being fed back into the national genetic evaluation systems. However, it is clearly a very valuable trait[5]. In an economic simulation of the relative economic value of selection it was determined that BRD incidence should be weighted approximately 7 times more heavily in a terminal sire, retained ownership selection index than weaning weight, postweaning average daily gain and feed intake, and that these traits should receive 2-3 times more emphasis than marbling score and yield grade.[8]

It was for this reason that a group of geneticists at several large U.S. Universities got together and obtained USDA funding for a project called the Bovine Respiratory Disease Coordinated Agricultural Project (**BRD CAP**) to attempt to address this problem using the tools of genomics.

The premise behind the project was to take DNA from large (> 1,000) cohorts of both Holstein dairy calves[6] and *Bos taurus* feedlot beef cattle[5] that were diagnosed with BRD using the McGuirk standardized scoring system[4], and their immediate neighbor or pen mate that remained healthy. The DNA profiles were then compared between these "cases" and "controls" using a "Genome Wide Association Study" or GWAS. The goal was to find genetic markers in these "reference population" that could be used to predict susceptibility to BRD in selection or tests candidates.

One question that remains to be answered is whether the trait of BRD susceptibility will be associated with several large effect single mutations, or have a more polygenic inheritance pattern (i.e., be associated with many small effect loci). This is called the genetic architecture of the trait. If the former "oligogenic" situation is the case then it is possible that the causative markers can be tracked using a haplotype approach to genomic prediction. A haplotype is a set of DNA variations, or polymorphisms, that tend to be inherited together. A haplotype can refer to a combination of alleles or to a set of single SNPs found on the same chromosome. An important advantage of haplotypes over single SNP markers is their higher ability to identify mutations. In animal breeding studies, SNPs are commonly bi-allelic and even when causative mutations have occurred between 2 SNP markers on a chip it is possible that the allele frequencies remain unaltered. However, when haplotypes are analyzed, mutations in different loci tend to be associated with major changes in the haplotype frequencies. Thus, a marker that is not in complete linkage disequilibrium (LD) with any individual bi-allelic SNP, marker may be in complete LD with a multi-marker haplotype.

The results of the dairy study in pre-weaned Holstein calves[6] showed over 100 genomic regions that were significantly associated with BRD, many of which were associated with biologically meaningful genes. This suggests that in the case of dairy this trait is somewhere between oligogenic and polygenic in trait architecture, and that many genes are associated with susceptibility to this multifactorial disease.

Ideally mutations that directly impact BRD susceptibility will be unambiguously identified through fine mapping, and sequencing of genomic regions associated with BRD susceptibility. This would eliminate the concerns associated with LD, as selection could be on the causative SNP itself. Such an objective is no small feat, as up until this point only a few causative mutations have been identified for any trait, let alone a complex trait like disease resistance. It is possible that this hunt will be facilitated by imputation of the genotypes of animals in the reference population using data from the 1000 Bulls project[1]. As suggested by its name, this large project is sequencing the entire 3 billion base pair sequence of over 1000 bulls of many different breeds. This information can be used to fill in the missing information (a process known as imputation) the full genome sequence of animals that have been genotyped using a lower density genotyping panel (Figure 1).

It is envisioned that performing a GWAS on this will dataset will help to precisely locate and enable the identification of the causative mutations for BRD susceptibility. An added advantage of this is that causative mutations are likely to be more persistent in their ability to predict disease susceptibility across generations.

Ultimately, the identified markers will need to be tested to assess how much of the genetic variation in BRD susceptibility they are associated with in an independent validation population. The purpose of this validation step is to use phenotypes available on an independent set of genotyped individuals to those used in the reference population to produce an estimate of the reliability of the genomic prediction for BRD susceptibility. The accuracy of genomic selection is known to rapidly decrease with increasing genetic distance between the reference population and the animal being tested, thus the loss of reliability with each generation removed from the reference population, although this is likely to be less of a problem when using causative mutations. The individuals sampled to form the validation population should be representative of the likely selection candidates. Ultimately it is envisioned that selection candidates will be genotyped and in addition to other traits the genomic information will provide a genetic prediction of BRD susceptibility. Given there is currently no selection criterion for this trait, even if only a handful of mutations are reliably associated with genetic variation in BRD susceptibility, this would be extremely valuable information as it would provide a selection criterion for this trait where now there are none.

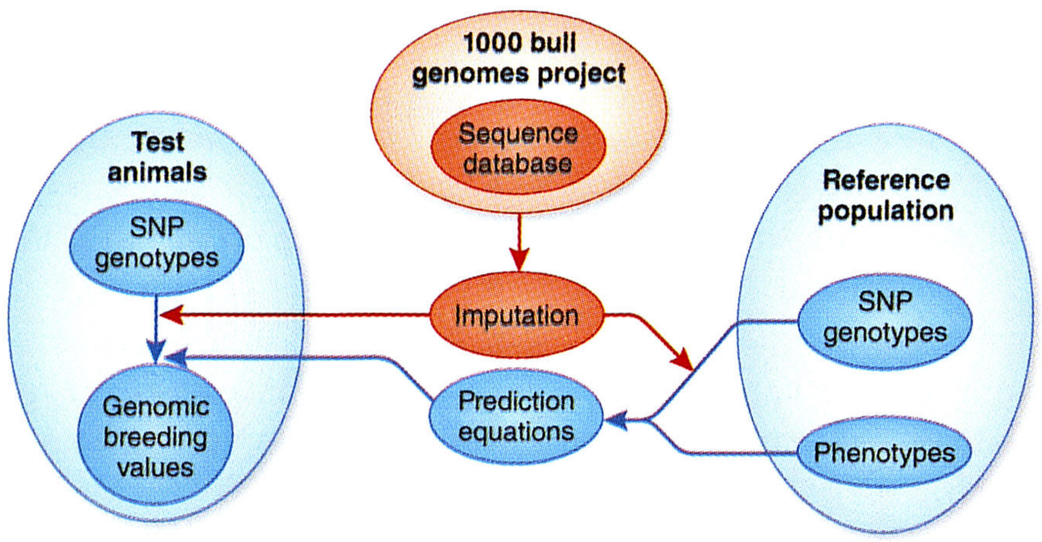

Figure 1. In blue, the classical process of genomic selection generating prediction equations from the joint analysis of 10,000 TO 50,000 SNP genotypes and phenotypes recorded in a large (>20,000 individuals) reference population is shown. These equations can then be used to predict genomic breeding values for test animals from their SNP genotypes alone. The red area shows that the sequence database of the 1000 bull genomes project allows for imputation of genotypes for millions of additional DNA variants for both reference and test animals to generate more robust prediction equations and genomic breeding values. Reprinted by permission from Macmillan Publishers Ltd: Nature Genetics, Copyright (2014).

Complex diseases, such as BRD, involve the influence of many genes and as a result, disease outcomes are difficult to predict. However, incorporating genomic information into the calculation of genetic merit estimates for mastitis in dairy cattle has demonstrated the ability to predict and select for disease risk. Preliminary data show BRD susceptibility has moderate heritability (0.21) [6], supporting the development of genetic predictions of BRD to enable selection for this important trait complex.

Translation in Industry

The incorporation of a diagnostic marker set for BRD susceptibility is perhaps easiest to envision in the dairy industry. This is partly due to the fact that there is a single predominant Holstein breed that was used in the reference population and the widespread use of a single selection tool by the US dairy industry; Lifetime Net Merit selection index ($NM). Selection indices provide a way to economically weight multiple traits into an a single value that differentiates animals based on expected profit. The $NM has evolved over the years, starting off originally focused only on milk and fat production in 1971, and maturing in 2015 to a balanced index that considers 12 traits associated with production, reproduction, type traits, and health traits (Table 1).

Table 1. Year that genetic rankings began and emphasis placed on dairy traits in 2014 US national dairy selection indexes

Trait	Year begun	2014 $NM Emphasis (%)
1. Milk	1935	-1
2. Milk fat	1935	22
3. Milk protein	1977	20
4. Productive life	1994	19
5. Somatic cell score (mastitis indicator)	1994	-7
6. Udder composite	2000	8
7. Feet and leg composite	2000	3
8. Body size composite	2000	-5
9. Daughter pregnancy rate	2003	7
10. Calving ability	2006	2
11. Cow conception rate	2104	1
12. Heifer conception rate	2014	5

The marker set that is shown to be associated with BRD susceptibility could be directly included on the genotyping chip that is currently being used for dairy genomic predictions. This development will require the incorporation of BRD susceptibility into the Lifetime Net Merit Index at an appropriate emphasis. Such calculations will depend upon how accurate the markers are at predicting BRD susceptibility in the selection population (i.e. reliability), and the economic value of the trait. Even if the markers predict only 20% of the genetic variation for this trait, this is likely to be valuable information given the significant economic costs associated with BRD.

Translation to the beef industry is likely to be a little more problematic due both to the larger number of breeds and breed associations involved in beef cattle genetic evaluation, and the structure of the beef industry. We do not yet know whether the trait of BRD susceptibility in beef cattle will be associated with several large effect single mutations, or have a more polygenic inheritance pattern (i.e. be associated with many small effect loci). In the former case we can likely genotype all breeds for these large effect causative mutations and get accurate predictions that will be robust over generations. This assumes that the same causative mutations are segregating across breeds. If BRD susceptibility is more polygenic in nature, this will make it more difficult to identify these small effect causative mutations. Each breed association will need to assess how accurate the markers are at predicting BRD susceptibility in their breed

Even if the technical challenges can be overcome, the real task may be getting the industry to incorporate the trait of disease susceptibility into beef cattle selection indices and selection decisions. One issue is that not all beef breed associations have economic selection indices, which would leave the value determination of BRD susceptibility up to the individual breeder. Additionally, selection against BRD susceptibility would obviously have great value to the feedlot sector, but breeders will need some incentive to include it in their selection criteria, especially given most producers do not retain ownership of their cattle through the feedyard. There needs to be some value transfer of the benefits derived from procuring cattle that remain free from disease in the feedlot, back to the producers who are providing those cattle. Such value transfer might be analogous to a backgrounding premium, but in this case the premium would be associated with including a cattle health trait in their breeding program.

The Importance of Recording Health Traits

Genomics has the potential to accelerate the rate of genetic improvement in low heritability, hard-to-measure traits such as disease status. Several studies show that the use of direct health observations is an effective way to incorporate heath traits into breeding programs. Such observations require a standardized system to record diagnoses to ensure phenotypes are comparable between farms. Consistent recording of health data is more difficult than for other traits due to subjectivity of diagnosis and reporting. Several studies have shown that for use in genetic evaluations, common health disorders recorded by farmers are of a similar quality as those documented by veterinarians[2]. A recent study showed that genetic selection for health traits (including cystic ovaries, displaced abomasum, ketosis, lameness, mastitis, metritis, and retained placenta) using producer-recorded health data collected from on-farm computer systems is feasible in the United States (Table 2).[7]

To be successful, there needs to be a balance between the effort required to collect these health data and subsequent benefits. Electronic systems that make such data capture easy and automated are likely key to the long-term success. The authors conclude that "The development of genomic selection methodologies, with accompanying substantial gains in reliability for low-heritability traits, may dramatically improve the feasibility of genetic improvement of dairy cow health."[7]

Table 2. Mean reliabilities of sire PTA computed with pedigree information and genomic information for health traits based on producer records of health events in U.S. dairy cattle. The right column shows how genomics can improve the overall gain.[7]

	Pedigree information alone			Blended pedigree and genomic information			
Health event	**Overall mean**	**Unproven sires***	**Proven sires†**	**Overall mean**	**Unproven sires**	**Proven sires**	**Overall gain‡**
Displaced abomasum	0.44	0.22	0.65	0.55	0.38	0.71	0.11
Ketosis	0.35	0.18	0.52	0.48	0.35	0.61	0.13
Lameness	0.24	0.15	0.32	0.39	0.31	0.47	0.15
Mastitis	0.39	0.26	0.52	0.51	0.40	0.612	0.12
Metritis	0.35	0.24	0.46	0.48	0.38	0.57	0.13
Retained placenta	0.55	0.42	0.67	0.64	0.54	0.73	0.09

*Unproven sires considered sires with less than 10 daughters.
†Proven sires considered sires with at least 10 daughters.
‡The increase in mean reliability calculated as the difference in overall mean reliability between the blended model and the traditional (pedigree data only) model.

Conclusion

The USDA, through the Agricultural and Food Research Initiative (AFRI) competitive grants program, is investing in several other similar grants focused on using DNA-based technologies to make genetic progress in other traits that have proven difficult to improve using traditional selection schemes. These include projects focused on the development of genomic approaches to improve feed efficiency and fertility. None of these traits are the "low hanging fruit" of genetic improvement. They are typically traits that are measured late in life, are expensive to measure or are not routinely measured at all, and frequently have low heritability making it difficult to differentiate the genetic component of phenotype from the environmental influences. However, a successful adoption resulting in a 1% improvement in feed efficiency, fertility, or reduced BRD disease incidence would translate into a huge costs savings for the US cattle industry.

Endnote

[a]The Bovine Respiratory Disease Complex Coordinated Agricultural Project Team: Sharif Aly, University of California Davis; Noah Cohen, Texas A&M University; Alan Dabney, Texas A&M University; Scott Dindot, Texas A&M University; Mark Enns, Colorado State University; Laurel Gershwin, University of California Davis; Robert Hagevoort, New Mexico State University; Terry Lehenbauer, University of California Davis; Holly Neibergs, Washington State University; Shannon Neibergs, Washington State University; Tim Ross, New Mexico State University; Christopher Seabury, Texas A&M University; Loren Skow, Texas A&M University; Jeremy Taylor, University of Missouri Columbia; Milton Thomas, Colorado State University; Cassandra Tucker, University of California Davis; Curtis Van Tassell, USDA ARS; Adroaldo Zanella, University of São Paulo, Brazil; and J.E. Womack, PD, Texas A&M University.

Acknowledgements

The Bovine Respiratory Disease Consortium Coordinated Agricultural Project (BRD CAP) was funded by the USDA Agriculture and Food Research Initiative (AFRI Grant no. 2011-68004-30367; J. E. Womack, Texas A&M, PD). More information about the BRD CAP can be found at http://www.brdcomplex.org/.

References

1. Daetwyler HD, Capitan A, Pausch H, et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat Genet* 2014;46:858-865.
2. Egger-Danner C, Cole JB, Pryce JE, et al. Invited review: overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. *Animal* 2015;9:191-207.
3. Georges M. Towards sequence-based genomic selection of cattle. *Nat Genet* 2014;46:807-809.
4. McGuirk SM. Disease management of dairy calves and heifers. *Vet Clin North Am Food Anim Pract* 2008;24:139-153.
5. Neibergs HL, Neibergs JS, Wojtowicz AJ, et al. Economic benefits of using genetic selection to reduce the prevalence of bovine respiratory disease complex in beef feedlot cattle. *Proceedings*. 46th Annual Research Symposium and Annual Meeting Lincoln, Nebraska, 2014.
6. Neibergs HL, Seabury CM, Wojtowicz AJ, et al. Susceptibility loci revealed for bovine respiratory disease complex in pre-weaned Holstein calves. *BMC Genomics* 2014;15:1164.
7. Parker Gaddis KL, Cole JB, Clay JS, et al. Genomic selection for producer-recorded health event data in US dairy cattle. *J Dairy Sci* 2014;97:3190-3199.
8. Van Eenennaam AL. The potential value of DNA-based tests for bovine respiratory disease (BRD) resistance to the beef cattle supply chain. *Proceedings.* 45th Annu Conf Am Assoc Bov Pract 2012; 60-65.
9. Van Eenennaam AL, Weigel KA, Young AE, et al. Applied Animal Genomics: Results from the Field. *Annu Rev Anim Biosciences* 2014;2:105-139.