Genomics in the United States beef industry

Megan M. Rolf a, Jared E. Decker b, Stephanie D. McKay c, Polyana C. Tizioto d, Kimberly A. Branham a, Lynsey K. Whitacre b, Jesse L. Hoff b, Luciana C.A. Regitano e, Jeremy F. Taylor b,*

a Department of Animal Science, Oklahoma State University, Stillwater, OK 74078, USA
b Division of Animal Science, University of Missouri, Columbia, MO 65211, USA
c Department of Animal Science, University of Vermont, Burlington, VT 05405, USA
d Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, SP, Brazil
e Embrapa Pecuária Sudeste, São Carlos, SP, Brazil

ARTICLE INFO

Keywords:
Beef cattle
Bovine
Genomics
Genomic selection
Genetic evaluation

ABSTRACT

While genetic evaluation systems which combine performance records and pedigree data have been utilized in the beef industry for over four decades, the incorporation of genomic information into genetic evaluation, and the effective implementation of genomic tools within the industry is relatively novel. Genomic technologies have been effectively deployed in the dairy, swine, and poultry industries; however, the beef industry possesses unique challenges for technology transfer. In this paper, we discuss the current limitations of genomic technologies and hindrances to the transfer of these technologies to the beef industry, while also considering opportunities for improved genomic and epigenomic tools needed to surmount barriers to technology adoption.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Beef producers have initiated the adoption of genomic technologies. With the seemingly continuous discovery of novel recessive genetic defects in a diverse spectrum of cattle breeds, many producers now utilize available genetic tests to identify carriers. Because carriers of known defects are severely discounted in the registered sector, perhaps overly so, the value of genetic testing is clear to bull breeders. Likewise, breed associations value and appreciate simple genomic technologies such as testing to verify parentage, fostering opportunities for seedstock and commercial producers to benefit from these applications. For example, it is well known that in multiple sire mating programs that some bulls will repeatedly sire significantly more calves than do their pasture mates (Drake et al., 2011). The use of parentage testing can identify which bulls disproportionately affect profitability through siring the greatest number of market calves and replacement heifers and which bulls should be sold without impacting reproductive rates. Establishment of pedigree relationships among the progeny also adds considerable value to the performance records collected in multi-sire mated herds for use in genetic evaluation. It is plausible that the employment of
technologies such as these will continue to increase, becoming as commonplace as vaccination programs.

The use and interpretation of results from genomic testing is often situation-dependent and can vary in different traits and populations. In this article, we will discuss implementation and potential for future applications of genomics in the U.S. beef industry within the context of within-breed application for National Cattle Evaluation (NCE) traits, within breed application for non-NCE traits, and across breed applications for genomics data.

2. Genomic approaches for quantitative traits

There are several examples of single-gene (coat color, horned/polled, genetic defects) or parentage tests which can significantly impact a producer's profit. On the other hand, when causal mutations underlying polygenic traits are not assayed, the cost of genotyping high-density assays, which may have limited predictive power in terms of both traits and populations, creates a challenge to technology transfer, especially for selection and management of commercial cattle. Technology users within the U.S. beef industry covet simple “silver bullet” solutions which genomics can currently provide only for simply-inherited traits or straightforward problems such as parentage verification. For quantitative traits, the solutions are generally complex and suffer from limitations including cost and population specificity. Additional impediments that are specific to the U.S. beef industry are the lack of a profit motive on the part of many small-scale producers who represent approximately 25% of cow ownership (McGrann, 2012) and a lack of transmission of appropriate economic signals from retailers, packers, feedlots and backgrounder/stocker operations to breeders due to the fragmented ownership of cattle during their life cycle. While some beef producers can effectively use genomic technologies to increase profit, technology adoption is currently considerably less than in the competing meat industries (Hayes et al., 2013; Newman, 2013; Fulton, 2012).

Producers of currently commercialized tests for quantitative traits return molecular estimates of breeding value or scores/rankings based on their marker tests to either the producer, the breed association, or, often, both. Some breed associations have also negotiated agreements with the service providers to include the receipt of raw genotypes. The Beef Improvement Federation’s guidelines dictate that the use of genomic information on quantitative traits be within the context of the generation of genomic-enhanced expected progeny differences (EPDs) which integrate performance, pedigree and genomic data. This approach works well for seedstock cattle and for traits which are commonly included in NCE. However, producers are often confused when they receive both genomic-enhanced EPDs from breed associations and molecular estimates of breeding value or scores/rankings from genotyping service providers which suggest different genetic merits for a trait. For this reason, some breed associations only report genomic-enhanced EPDs, and not the molecular breeding values. Additional problems have arisen as breed associations have experimented with different methods for integrating genomic data into their genetic evaluation systems to produce genomic-enhanced EPDs. For example, because prediction equations are currently limited to a single breed, a population of animals from that respective breed must be used in the training set to generate molecular breeding values. Those prediction equations are then utilized in the entire population, including those animals represented in the training population, which results in the double counting of data on some animals. This situation can lead to an artificial increase in accuracy; however, the best method to avoid this problem is not clear. Many training populations are small and consist of the most influential animals within a breed. When the animal’s owners have paid for genotyping, they expect access to genomic-enhanced EPDs on those animals, and excluding their genomic data from the full evaluation to avoid double counting is problematic. The need to retrain prediction equations results in an ever-expanding training population size, which underscores the necessity of developing a solution to this problem. Additionally, it is widely accepted that the degree of relationship between the training and implementation population influences the accuracy of prediction, and there remains a significant opportunity to explore weighting of genomic data in the analysis to account for these pedigree relationships. Lastly, when genotype information is included in genetic evaluation, the data storage requirements, especially with raw genotypes, are greatly increased. Paired with the need to alter National Cattle Evaluation procedures at each respective breed association to accommodate the molecular data, these necessary changes can pose a challenge to the adoption of this technology within the industry.

To this point, no U.S. beef breed association has implemented the procedures adopted by the U.S. dairy industry, which uses genotype information to generate genomic relationships between animals, largely because raw genotypes have only recently become available to the associations. Rather, the U.S. beef associations have adopted one of three approaches to produce genomic-enhanced EPDs: (1) a selection index which weights pedigree-based and molecular EPDs based on the predictive ability of the genomic test, (2) a multivariate analysis in which the molecular EPD is treated as a correlated trait (Kachman, 2008), or (3) an analysis in which the molecular EPD is treated as an external EPD (Quaas and Zhang, 2006; Henderson, 1975). Each approach has advantages and disadvantages; and the selection of method has been primarily predicated on the types of available data (i.e., complete genotype information versus calculated molecular estimates of breeding value) and the flexibility of software used to perform genetic evaluation. Each approach has different advantages including the ability to include an individual accuracy for each genomic test (external EPD), the ability for information to filter through the pedigree to non-genotyped animals to increase the accuracy of genomic-enhanced EPDs for related animals (correlated trait), and possessing readily-available data for retraining that is independent of currently implemented genomic technologies (indexing). However, none of these approaches are optimal, and the ultimate solution to this
problem will require the development of a new methodology that simultaneously incorporates phenotypes, genotypes and pedigree information, not necessarily all recorded on the same animals, to generate genomic-enhanced EPDs.

3. Challenges with genomic evaluations

A multitude of challenges have limited the rate of adoption of molecular technologies in the U.S. beef industry. While some registered producers and breed associations have embraced the concept of genotyping and are driving the implementation of genomic selection within their breeds, the majority of producers are not testing their cattle primarily due to the lack of short-term return on investment. Commercial producers generally do not receive substantial returns on investments to offset the cost of labor and testing of their market calves, with the possible exception of those that retain ownership as opposed to selling their weaned calves. In general, genetic gains achieved by commercial producers are incremental and relatively small annually, but lead to long-term genetic improvement. The adoption of genomic selection may not dramatically change the rate of genetic improvement in commercial herds due to relatively low cow replacement rates, but does require a significant economic investment. However, commercial herds can reduce their risk when buying herd bulls, as genomic-enhanced EPDs are more reliable than traditional EPDs for young bulls which are primarily based on parental average EPDs for all except early growth trait such as birth and weaning weight. This reduction in risk is proportional to the decrease in the possible change value for each EPD due to the increase in accuracy. Genotyping has the opportunity to generate considerable value within the entire beef industry, from the producer to the end product consumer (Bullock et al., 2012); however, most of the expense of genotyping is currently borne by the seedstock sector. Within the seedstock sector, genotyping technologies have largely been used as a marketing tool, and the ability to increase genetic improvement has been a secondary goal. Returns on investment due to increased marketing ability tend to be much higher than in the commercial sector. Nonetheless, for many smaller registered producers who do not receive a bonus due to marketing, the costs of these tests may still be beyond reach. This challenge will likely be solved over time as bull and female buyers become more educated about the value of increased accuracy of EPDs due to genomic testing and as the cost of genotyping declines.

Genotyping technologies have rapidly evolved, but in general have increased marker density while maintaining a fairly constant assay price. There are a few examples of low density assays being used for genotype imputation or of targeted panels marketed to commercial producers; however, it is clear that to increase the rate of adoption of genotyping, the cost of assays must continue to decrease without losing predictive ability. Ideally, the assay would include the markers used by breed associations for parentage verification and the cost of the delivered content would not greatly exceed the cost of genotyping for parentage verification. At this point, there would be little impediment to whole-herd genotyping in the registered sector whereby genotypes and genomic-enhanced EPDs for yearling bulls could accompany them at their sale into the commercial sector. This could also lead to the genotyping of cows within commercial herds and perhaps even calves if producers could market their cattle according to their projected carcass yield and eating quality (specifically tenderness and marbling). Genotype service providers are currently working towards this goal and are already offering tests that combine parentage testing and single-gene tests (for genetic defects and/or coat color) or sector-specific tests such as the Certified Angus Beef® GeneMax™ test for feedlot performance (marbling and gain) into “packages” to broaden the appeal of genomic testing at a modest price point to producers. These packages achieve their price point by decreasing the number of markers that are genotyped and consequently may explain less of the variation in the target trait than the commonly used BovineSNP50 (50 K) assay and have little to no predictive power for genetic merit in non-target traits.

Very few causal mutations underlying variation in quantitative traits have been identified, and consequently the majority of tests used to predict molecular breeding values rely upon establishing the phase relationship between SNP and QTL alleles (or chromosomal haplotypes which may contain several QTL) which largely makes the tests breed-specific. Since many of the commercial beef producers in the U.S. utilize composite-breed cows to capitalize on heterosis, the breed-specificity of genomic tests is the primary limitation to the adoption of genotyping in the commercial sector. However, the generation of genotype information for hybrid cattle (i.e., Lim-Flex, Balancer, SimAngus, etc.) for incorporation into National Cattle Evaluation will generate the data that are necessary for the development of multi-breed tests. Simulation studies have indicated that the accuracy of genomic predictions within individual breed-groups were only slightly lower than achieved by within-breed predictions as long as all breeds were included in the training population (Toosi et al., 2010; Kizilkaya et al., 2010). Multi-breed beef cattle studies, comprising either multiple purebred populations (Garrick and Saatchi, 2013) or crossbred multi-breed populations (Weber et al., 2012; Rolf, 2012), support these results. Current results show that only a small amount of genetic variation is explained (Weber et al., 2012); however, an increased training population size will likely be necessary for crossbred and multibreed populations (Toosi et al., 2010) or the identification of important causal variants will be required to increase the efficacy of these tests. Because of these limitations, the development of better tests may be feasible with increased adoption levels (at least for those producers utilizing the technology) for commonly measured traits. Because the U.S. beef industry utilizes more than 80 distinct breeds of beef cattle with at least 12 breeds (http://www.ansi.okstate.edu/breeds/cattle/) being numerically significant, the development of multi-breed genomic evaluations will require the genotyping and phenotyping of significant numbers of purebred cattle from each breed or composites representing all numerically important breeds.
If performance within the seedstock sector is not representative of performance in the commercial sector (Ibanez-Escriche et al., 2009), or is not fully realized due to environmental limitations (Lalman et al., 2013), the development of multi-breed evaluations in commercial cattle may better identify elite seedstock animals based on the performance of their commercial progeny (Zeng et al., 2013). Limitations to the implementation of this approach in the U.S. beef industry have been the cost of genotyping and the limited availability of large samples of animals with available DNA and a broad spectrum of phenotypes for feed intake, growth, reproductive, health, and carcass traits.

A second limitation of these approaches is the difference in application of data for traits that are included in NCE (i.e., growth and carcass), and those that are novel (i.e., feed efficiency and disease resistance). Within the NCE system, adoption of genotyping by producers and the sustained recording of phenotypes ensure a dynamic training population for future iterations of prediction equation development. For non-NCE traits, research populations are utilized to develop prediction models, but because these phenotypes are not commonly recorded during NCE, there is no clear path to obtaining the data that will be required to retrain prediction models in the future. The use of novel traits that exhibit a genetic correlation with other commonly recorded production traits within a multi-trait NCE model is one possible way to integrate these data into NCE. Some examples that might be useful include the incorporation of beef tenderness or fatty acid profile data with carcass data and feed intake or enteric methane emission data within the evaluation for growth traits. The inclusion of these traits with appropriate weights within the context of a selection index would make this information easy for producers to effectively utilize. As methodologies progress and the ability to identify causal mutations increases, these issues may be solved through development of new tools to address these selection goals.

Another consequence of the use of SNPs that are associated with phenotype rather than the causal mutations themselves is that the accuracy of genomic predictions decays with the number of meioses that separate the training population from the implementation population (Taylor, 2014). This results in losses in the accuracy of genomic predictions between populations from the same breed (e.g., in Australian or Argentine Angus when training occurred in U.S. Angus) and in advanced generations of selectively bred cattle if their genotypes and phenotypes are not used to retrain the genomic predictions. The effect of the magnitude of genetic distance between the training and implementation populations on the accuracy of genomic predictions has been well-documented using simulated data (i.e., Habier et al., 2007), but fewer studies have been performed in actual beef cattle populations (Rolf, 2012). Furthermore, and perhaps most importantly, the development of genomic predictions for novel phenotypes such as disease resistance or feed efficiency will be limited to only those breeds, or widely-used families within breeds, in which the phenotypes were collected and not to the industry as a whole, which was the original hope for genomic selection. Consequently, the implementation of genomic selection within the beef industry will require the on-going collection of both genotypes and phenotypes within a multitude of breeds to enable the retraining of genomic prediction models in order to maintain high prediction accuracies.

4. Challenges and opportunities for indicine-based breeds

In the United States, most of the cattle production is centered on the use of taurine-based cattle breeds, and it is primarily in the southern tier of the U.S. that the use of Brahman-based breeds occurs. While Brahman is predominantly used in the U.S. to form composites with taurine breeds, the majority of available genomic tools have been developed in taurine cattle primarily because of the greater ability of the taurine breeds to capture large numbers of genotypes and phenotypes upon which to develop prediction models. Unfortunately, this has led to a lag in technology transfer for many of the largest cattle herds in the world which are located in tropical climates. Bos taurus indicus based breeds are particularly well adapted to tropical environmental conditions (Hanotte et al., 2000) and the overwhelming majority of the cattle in the world’s tropical regions have indicine ancestry.

Bos t. indicus breeds are well-suited to crossbreeding programs that involve Bos t. taurus breeds to exploit hybrid vigor due to the considerable genotypic divergence between the subspecies (The Bovine HapMap Consortium, 2009; Decker et al., 2014). Despite the production advantages of Bos t. indicus composite breeds in tropical and subtropical environments due to increased resistance to heat, disease, internal and external parasites (Crouse et al., 1989), their inferior beef quality, particularly meat tenderness, is a substantial drawback and impacts demand and price of indicine sourced beef (O’Connor et al., 1997; Johnson et al., 1990; Whipple et al., 1990).

While the existing high-density SNP genotyping assays were designed to primarily capture common variation within the genomes of taurine breeds (Matukumalli et al., 2009), the marker density, particularly of the Illumina BovineHD BeadChip, is sufficient to enable genomic selection within indicine or indicine × taurine breeds (Bennewitz et al., 2009; Tizioto et al., 2013; Espigolan et al., 2013). The success of genomic prediction depends on genotyping enough markers to identify the spectrum of haplotypes which exist within a breed and in genotyping enough animals that the substitution effect of each haplotype can be reasonably well estimated (Wray et al., 2013). The models that are fit to estimate SNP allele substitution effects for genomic prediction can simultaneously be used in Genome Wide Association Studies (GWAS) to identify QTLs and their candidate genes in an effort to better understand the biological mechanisms which underlie production traits. Marker density should be sufficient to ensure that QTLs of large effect are not (\(\Delta LD\)) is considerably less than in taurine populations (The Bovine HapMap Consortium, 2009). While high-resolution LD maps and descriptions of haplotype block configurations are in development for many breeds,
the fine-scale description of the structure of LD at the breed/population level is crucial for understanding the associations between genes and phenotypic traits (Jakobsson et al., 2008).

The most recently developed SNP chips including the BovineHD and Affymetrix Axiom® BOS 1, utilized sequence data generated from indicine animals in the assay design to reduce the allele frequency ascertainment bias that favored taurine cattle in the design of the BovineSNP50 assay, thereby producing a tool that was better suited to both taurine and indicine breeds. The BovineHD assay has been utilized in several studies of Zebu and Nelore populations and has demonstrated that different genomic regions influence meat quality traits in indicine cattle (Tizioto et al., 2013) compared to taurine cattle (McClure et al., 2012). These results may reflect differences between breeds for allele frequencies at the causal mutations, the extent of LD (Bolormaa et al., 2013), or the presence of epistasis which could influence the magnitude of QTL effects within different genetic backgrounds and result in different marker-trait associations. Differences in the patterns of variation within genes in pathways regulating metabolism and production traits in indicine and taurine cattle due to the extent of the divergence between indicine and taurine cattle (250 Ky, The Bovine HapMap Consortium, 2009; Decker et al., 2014) could also explain the incongruity between the genes but not the pathways that are associated with production traits in the two subspecies (Tizioto et al., 2013). These findings reinforce the need for the development of population-specific genomic prediction training sets involving Bos t. indicus breeds particularly in view of the possibility of genotype-by-environment interactions caused by differences in performance recording (phenotype and EPD definitions) and production systems (grazing vs. concentrate feeding) between locations or the performance potential of indicine cattle under adverse conditions.

5. Influence of genetics on beef sustainability and adaptability

From an economic perspective, the beef industry has made great strides in advancing sustainability; however, much work remains to address the environmental and social aspects of sustainable beef production. Increasing competition for the natural resources needed to raise cattle (corn production used for ethanol and competition for water resources, for example) will drive the need for further research. Furthermore, because sustainability and adaptability traits are not routinely collected in most National Cattle Evaluations, genomic tools will likely emerge as being the most useful for generating improvement in these traits.

How U.S. consumers perceive the well-being of cattle has increasingly substantial effects on the social aspects of sustainability by influencing and establishing the demands for alternative beef products (i.e., grass-fed beef). The need to tackle consumer concerns that have major consequences for beef production systems, paired with the need to increase the environmental sustainability of cow/calf production operations, will increase focus on the efficiency of nutrient utilization and methane output in grazing cows and will require the development of new research protocols for phenotype collection. Methane emissions in cattle are driven largely by the intake of fermentable carbohydrates (Johnson and Johnson, 1995; Ellis et al., 2007). Enteric methane emissions per head have increased in recent times; however, the beef industry has concomitantly decreased the amount of methane produced per unit of beef product by decreasing the cattle inventory while increasing beef production through various production efficiencies (Capper, 2011; Rotz et al., 2013). Further reductions in emissions could be economically beneficial for cattle producers, because methane production primarily represents an inefficiency in the cycle of converting feed energy to growth, milk and other desirable production traits (Hayes et al., 2013). Because selection for increased feed efficiency has been shown to decrease methane emissions in some studies (Hegarty et al., 2007), it is possible that continued progress in beef cattle production efficiencies will also result in an improvement in environmental sustainability.

Feed efficiency and other production efficiencies must be tied to any discussion of sustainability. Increasing feed efficiency has the potential to profoundly impact beef profitability and food security. Many studies have assessed the extent of genetic control of feed efficiency in conventional feedlot settings (i.e., Rolf et al., 2012; Barendse et al., 2007; Sherman et al., 2009); however, the relationship between feed efficiency of growing calves and stocker cattle or breeding age cows grazing forage has received limited attention due to the inability to collect accurate data on large numbers of animals. While it is common practice to assume that animals will rank similarly for their feed efficiency on pasture as they do on concentrate-based rations (Arthur and Herd, 2005), this relationship is not known.

With an increasing focus now being placed on animal welfare, it is reasonable to assume that research to identify QTL for important behavioral traits in cattle will increase. In particular, temperament is important to both human safety and animal well-being and has been implicated in variation in meat quality and average daily gain (Voisinet et al., 1997). Because of the relationship between behavioral traits, meat quality, and other performance traits, this affords a unique opportunity for cattle producers to improve genetic merit in a variety of economically important traits in conjunction with enhanced consumer acceptability of beef products and production practices.

Morbidity and mortality as a result of disease and injury, and the corresponding lower performance of affected cattle results in substantial economic losses within the beef industry. Due to the economic importance and heritable basis of disease resistance, there is considerable interest in studying the mechanisms underlying disease resistance in beef cattle (Morris, 2007). Again, the lack of consistent and widespread phenotyping is the largest impediment to the development of genetic evaluations for disease resistance; however, selection could be made possible through the use of genomic tools. Several diseases and disorders have recently been studied, including infectious bovine keratoconjunctivitis (Snowder et al.,
of longevity, calculated either using either linear models or survival analysis (Caraviello et al., 2004). While missing or censored data and environmental interactions with production traits that lead to involuntary culling are an issue with fertility and longevity analyses in the absence of whole-herd reporting (Essl, 1998), these tools provide a foundation for genetic progress in these traits.

Despite its considerable economic importance, negative genetic trends in fertility have been observed in livestock (Rauw et al., 1998; Decker et al., 2012) and genotypic tests for fertility have a high value proposition (Van Eenennaam et al., 2011; Fritz et al., 2013; Sonstegard et al., 2013). As the breed associations now have genotypes for more than 1000 individuals which is considered by many to be the de facto entry point for developing genomic predictions, although the information content present in this sample of animals can vary widely. These animals provide a substantial nucleus of data for training genomic prediction equations, but the cost of genotyping animals has been an impediment to technology adoption. As in other livestock sectors (Cleveland and Hickey, 2013), genotyping strategies that include imputation offer an opportunity to cost effectively genotype large numbers of individuals by acquiring most genotypes at low density (Boichard et al., 2012). When high-density genotype data (≥ 50 K) are available on individuals, the use of low density assays to genotype large numbers of animals is likely sufficient to allow accurate imputation to high density for the accurate estimation of genomic-enhanced EPDs (Cleveland et al., 2011). Genotyping a large number of individuals is important for maximizing selection intensity in genomic breeding, so the use of a dual density strategy with imputation, such as that implemented by the American Gelbvieh Association, is clearly sensible. However, there may be limitations to the accuracy of imputation for animals distantly related to those with high-density genotypes, which underscores the importance of strategic genotyping of individuals according to their pedigree (Druet et al., 2014). Since beef breeds typically have higher effective population sizes (especially breeds with open herd books; The Bovine HapMap Consortium, 2009) and smaller numbers of high-density genotyped individuals than do the U.S. dairy breeds, the improvement in accuracy of genomic-enhanced EPDs is critically linked to the number of collected high-density genotypes. As cost structures evolve for the different assays, the marginal returns to imputation will likely change. However, preliminary evidence in dairy cattle suggests that genomic predictions based on assays larger than 50 K SNPs offer only a small improvement in the accuracy of genomic-enhanced EPDs and the cost of genotyping with the BovineHD or BOS 1 assays is not justified for this purpose (VanRaden et al., 2013). A strategy based on the use of 50 K and 10 K assays with the opportunity for the inclusion of causal mutations as they are detected is the most cost effective strategy for performing genomic selection in beef cattle.

7. Opportunities to use genomics to fill critical knowledge gaps in the beef industry

Genotyping-by-sequencing (GBS) is an emerging technology that is rapidly becoming an option for the whole genome genotyping of cattle for research purposes and of targeted loci for commercial purposes. The whole-genome genotyping technology was originally developed for SNP discovery in cattle based on the sequencing of pools of restriction enzyme digested genomic fragments (Van Tassell et al., 2009) and has since been extended for genotyping individuals using multiplexes of bar-coded samples (Elshire et al., 2011; De Donato et al., 2013). The approach is unbiased with regards to the minor allele frequency of detected loci – with the exception that very low frequency variants and sequencing errors become confounded, and is flexible because it does not require the development of locus-specific probes for the analyzed variants.
De Donato et al. (2013) generated 40,601 SNPs assigned to chromosomes with a 90% call rate in 47 cattle from seven different breeds for $30–40 per sample, a sufficient marker density for genomic selection, GWAS, and population genetic analyses at a cost of about one half that of the available 50–70 K chip-based genotyping technologies.

Drawbacks of this whole-genome approach are higher missing data and genotype error rates, primarily due to the depth of sequence coverage at each called variant, which may limit its cost-effectiveness and utility in commercial applications. Whereas average sample call rates (proportion of SNPs with genotypes and SNP call rates (proportion of samples with genotypes at each SNP) are near 99%, for SNP assays, both are lower than 90% for GBS. Furthermore, if the depth of sequence coverage is only 2–3 × per individual at each SNP, heterozygous genotypes will not all be accurately called. This necessitates an imputation step in most GBS pipelines. Furthermore, there is a wide variation in inter-marker distances with GBS relative to the tight distributions achieved with designed high-density SNP assays. Finally, with GBS there is a bias towards more SNPs on the smaller chromosomes due to the lower content of repetitive sequence on these chromosomes.

A recent study in yeast has shown that a finite number of quantitative trait variants explain all, or most of, the genetic variance in complex traits (Bloom et al., 2013). The decreasing cost of next-generation sequencing technologies has made it possible to produce genome-wide genotypes on individuals although the cost remains an order of magnitude greater than the cost of genotyping with the highest-density BovineHD or Axiom BOS 1 assays. Nevertheless, it is becoming clear that the imputation of high-density genotype data used in GWAS to sequence-level variation will lead to the simultaneous identification of the functional variants that underlie QTL of large effect (Sellner et al., 2007). Consequently, instead of genomic predictions based on ever increasing numbers of DNA markers, in the future, genomic predictions will be based on smaller numbers of markers with increasingly more markers having functionally significant effects. This will alleviate two shortcomings of genomic selection in the U.S. beef industry. Firstly, assays should become less expensive and have broader application across breeds and sectors of the industry. Secondly, there should be smaller decreases in the accuracy of genomic-enhanced EPDs as tested animals become further removed from training populations by either genetic divergence or time.

Sequencing will clearly have a role to play in the development of low-cost and high-throughput genotyping assays which target specific SNP loci. Thallman et al. (2013) used Next Generation Genotyping to produce genotypes for 95 SNP on 1080 cattle samples via target locus amplification using barcoded primers and the next generation sequencing of the pooled amplicons. The produced genotypes were 99.1% concordant with those produced using an independent commercial assay and the genotype call rate was an acceptable 96.1%. The advantages of this genotyping strategy include cost, provided a sufficient number of animals will be genotyped to amortize the cost of primer synthesis and also potentially flexibility since, at least in theory, loci can be substituted within the target locus multiplex. The greatest limitation at present is the small number of loci that can be targeted for genotyping, although Thallman et al. (2013) indicate that work is currently underway to develop a panel with from 1000–3000 SNPs.

8. Targeting non-additive genetic effects

Traditional selection practices are based solely on selection for additive genetic merit of economically relevant traits. However, much of the commercial beef industry capitalizes on non-additive genetic merit through heterosis achieved by crossbreeding. While the molecular basis for heterosis and its obverse, inbreeding depression, are not well understood the existence of large numbers of non-lethal loss of function (LOF) alleles within breeds may very well explain both phenomena. The explanation postulates the existence of overlapping sets of LOF alleles between breeds, with the extent of the intersection set being determined by the genetic distance between breeds. If these LOF alleles are phenotypically detrimental (i.e., they are functionally responsible for QTL within breeds), inbreeding within each breed will increase the proportion of individuals that are homozygous for LOF alleles at each locus and phenotype will be depressed. When breeds are crossed, heterozygosity at loci with LOF alleles will increase and will be maximized when the breeds carry LOF alleles at completely distinct loci (the intersection set is null). Under this model we would expect to see heterosis maximized for crosses between genetically distant breeds where evolution would be expected to result in a small overlap in the number of loci sharing LOF alleles (distinct LOF mutations within the same gene). This is an intriguing and testable hypothesis which suggests an approach to simultaneously identify QTL which have large non-additive effects within breeds and which contribute to heterosis in breed crosses.

9. Epigenetics

Epigenetic modifications have the ability to regulate gene transcription, and subsequently influence variation in economically important phenotypes of agricultural species. Consequently, epigenetic modifications can cause one genotype to produce alternative phenotypes. One of the most thoroughly studied epigenetic modifications is DNA methylation which involves the addition of a methyl group to the 5 position of cytosines which are found in CpG dinucleotides. When 5-methylcytosine is found in the most thoroughly studied epigenetic modifications is DNA methylation which involves the addition of a methyl group to the 5 position of cytosines which are found in CpG dinucleotides. When 5-methylcytosine is found in the 5’ promoter region of a gene, it can prevent the transcriptional machinery from binding to its target site, resulting in the silencing of gene transcription (Portela and Esteller, 2010).

To understand the impact of DNA methylation on environmental (transient and non-heritable modifications) and additive genetic variation (modifications are stably transmitted across several generations), first requires the ability to establish the distribution of DNA methylation within the genome of tissue-specific cells. Recent studies have generated atlases of the porcine adipose and muscle
tissue methylomes to enable the investigation of the association between methylation and obesity (Li et al., 2012). The identified differentially methylated regions (DMRs) were found to contain ~80% of the known or candidate obesity-related genes in human and 72% were in QTL regions that affect fatness and pork quality. Natt et al. (2012) found 145 genes to be heritably differentially methylated between White Leghorn chickens and Red Jungle Fowl, with 79% being hypermethylated in the thalamus/hypothalamus of White Leghorn. The over representation of DMRs in selective sweep regions associated with domestication suggests that novel heritable methylation patterns may be sufficiently stable to allow strong recurrent selection for epialleles during the domestication of White Leghorn.

Epigenetic modifications may be stably transmitted over several generations (Johannes et al., 2008; Natt et al., 2012) and because beef cattle populations have only a relatively small number of generations of individuals with phenotypes that are used in NCE, the effects of these loci will be incorporated into the additive genetic component of variation and thus into EPDs. However, the long-term transmission stability of epialleles which influence economically important phenotypes is unknown and if these have large effects on phenotypes, they will not be detected by the usual forms of DNA sequence based analysis. If the transmission of epigenetic mutations is unstable and rapidly dissipates, their effects would be captured in progeny tested young bulls but would bias their genetic evaluations relative to the performance of their grand-progeny or great-grand-progeny. Consequently, the identification of epi-QTLs and the quantification of their stability of transmission may have consequences for the accuracy of EPDs and long-term response to selection. Of even greater interest will be coming to an understanding of the mechanisms which induce stable epigenetic modifications to DNA which have large phenotypic effects and how these may be harnessed to enhance food production. If epigenetic modifications evolved as a mechanism to allow organisms to respond more rapidly to environmental changes then is possible from selection on hard-wired DNA variants, we might expect the effects of epialleles to be somewhat larger than the effects of quantitative trait nucleotides, which primarily approach the infinitesimal model (Cole et al., 2011; Tizioto et al., 2013).

10. Conclusions

Progress towards the implementation of genomic selection in the U.S. beef industry has accelerated rapidly over the last five years, and the rate of change promises to increase as new technologies are developed, the cost of current technologies are decreased, and novel mechanisms underlying heritable genetic variation are elucidated. The decreasing costs of next generation sequencing will facilitate new approaches to the identification of variation underlying economically important traits and will lead to increasingly inexpensive genotyping which will facilitate mate selection to minimize homozygosity of loss of function alleles, enable genomic selection within all of the breeds present within the U.S. beef industry and facilitate the marker-assisted management of calves as they proceed through the production chain.

Acknowledgments

The authors appreciate the support of National Research Initiative grants numbers 2011-68004-30214, 2011-68004-30367 and 2013-68004-20364 from the USDA National Institute of Food and Agriculture.

References:


