Basics of genetic variation and the use of genomic technologies in cattle

Jason B. Osterstock, DVM, PhD
Zoetis, Kalamazoo, MI 49007; email: jason.osterstock@zoetis.com

Abstract

Genetic variation is an important factor in determining phenotypic variation in cattle. This variation arises in large part from differences in the genetic composition among animals. Genomic technologies can help to describe this variation and expected phenotypic impacts in part associated with differences in coding regions of the genome affecting protein structure and function, and by characterizing similarities among animals at the molecular level that can help inform prediction of genetic merit among genomically similar animals. When applied in this manner, genomic technologies help to provide more accurate estimation of genetic merit for more traits than can often be effectively measured phenotypically much earlier in life. This improves and accelerates decision making, thus optimizing production efficiency.

Key words: genomics, genetic variation, genetic evaluation, SNP

Résumé

La variation génétique est un important facteur dans la détermination de la variation phénotypique chez les bovins. Cette variation provient en grande partie de différences dans la composition génétique d'un animal à l'autre. Les technologies génomiques peuvent aider à décrire cette variation et l'impact sur les phénotypes attendus qui sont en partie associés à des différences dans les régions codantes du génome qui affectent la fonction et la structure des protéines. Ces technologies peuvent aussi caractériser les similarités entre les animaux au niveau moléculaire afin d'aider à raffiner les prédictions sur le mérite génétique d'un ensemble d'animaux génétiquement similaires. Appliquées de cette manière, les technologies génomiques peuvent aider à faire des estimations plus précises du mérite génétique pour plus de traits qui peuvent souvent être mesurés adéquatement au niveau phénotypique beaucoup plus tôt dans la vie. Cela améliore et accélère la prise de décision optimisant ainsi l'efficacité de la production.

Introduction

Differences in the genetic composition of cattle represent an important component of variation in economically relevant outcomes. To a significant degree, livestock production is all about managing this variation, making understanding and management of genetic variation a powerful lever in optimizing production. While our appreciation of the contribution of genetics to variation in performance, reproduction, and health is not new, the emergence of genomic technologies over the last decade has created tremendous opportunity.

From a biological perspective, the contribution of genetic variation to differences in phenotype can be a reasonably linear exercise. Differences in DNA sequence and subsequent mRNA transcripts leads to varying combinations of amino acids and, therefore, variants in protein structure and function that are expressed as phenotypic variants by virtue of their contribution to metabolism and structure. Pretty simple. But that only applies to differences in DNA sequence within regions of the genome that code for proteins (i.e., genes), and differences in DNA sequence that actually manifest as a difference in amino acids. The reality is that most economically relevant traits are influenced by many genes, and variation in DNA sequence occurs within coding and non-coding regions, both of which may alter phenotypic expression. So that tidy linear pathway derived from basic molecular genetics begins to feel quite inadequate.

Classically, animal breeding and quantitative genetics take a different view. Rather than worrying about the underlying biology, quantitative genetics focuses instead on understanding how phenotypes are expressed in related animals. In its most basic form, this process uses statistical methods to estimate the genetic merit of 1 individual based on what has already been observed from parents and other relatives, who in turn have their genetic merit estimated from their parents and relatives, and so on. Intrinsic to this model is knowledge of the pedigree, collection of phenotypes, and mathematical models to reduce the impact of non-genetic sources of variation in phenotype. We can even estimate the prediction error variance and derive an estimate of confidence (i.e., accuracy) in the estimated genetic merit. Making progress. But pedigree is sometimes incorrect, and most would readily admit that assuming progeny are the average of parents is not terribly accurate, and by the time all the phenotypes are measured it's too late.

Contemporary genetics leverages genomic technologies to help create a bridge between the gaps in classical methods and the need to accurately assess the genetic merit of individual animals. Genotypes are used to verify recorded pedigree, discover otherwise undocumented genetic similarities among individual animals, and consider variants that are directly or indirectly tied to how genotype translates into biological function. Over the last 10 years, genomic technologies have evolved and are now routinely applied across the breeding layer of beef cattle production. Genotypes consisting of tens of thousands of single nucleotide polymorphisms (SNP) are commonplace, providing high-resolution characterization of genetic similarity that delivers substantial improvement in accuracy and decision making.

One of the primary challenges for veterinarians in understanding contemporary genomic technologies is what I might refer to as "biology bias." Veterinary and animal science training provide exposure to key concepts like modes of inheritance (i.e., dominant, recessive, codominant, etc.) and heritability (the percent of phenotypic variance attributed to genetic effects). The aforementioned ties between molecular biology and phenotype work well in tying structure to function. And the use of molecular methods in diagnostic testing for pathogen testing provides an apparent parallel. Veterinarians often try to draw from these concepts to understand genomic technologies. Therein lies the bias… trying to explain the biology to explain the value of understanding the genotype. In practice, application of genomic technologies is more strongly tied to statistics (ugh), and big data concepts (double-ugh). The key to understanding how the technologies work is largely tied to divorcing oneself from understanding the biology and thinking about the genotype as another source of data that, when combined with other data points like pedigree, phenotype, phenotype of relatives, etc, provides a more accurate assessment of the likely outcome. Part of that value is absolutely drawn from the biological significance of the genotype. Much more of that value is drawn from identifying patterns of DNA that were actually inherited from the parents and recognizing how other animals with similar DNA patterns performed.

The benefits of genomic technologies in cattle are several fold. First, as referenced above, genomic data increases the accuracy of estimates of genetic merit. This is a result of addressing potential errors in recorded pedigree and understanding the genomic content that the offspring actually inherited rather than assuming they are a balanced blend of the parents' best and worst qualities. Improved accuracy imparts an improved understanding of the animal's true genetic merit, thus increasing the effectiveness of any selection decision based upon this information. The second benefit of genomic technologies is that the data can be obtained very early in life, well in advance of any observation of phenotype for economically relevant traits. Earlier decision-making helps to drive efficiency into the selection process and allows advanced reproductive technologies to be applied in elite animals much sooner than would have been prudent using traditional methods. Finally, genomic technologies can provide insights for traits that are otherwise difficult or expensive to measure phenotypically.

These benefits notwithstanding, it is important to note that the use of genomic technologies is dependent upon the availability of phenotypes, and specifically, phenotypes that can be linked to genotyped animals through either pedigree of genomic relationships. Sustained collection of phenotypic data is critical to maintaining the validity of the genomic prediction.

Conclusions

Genetic variation makes significant contributions to differences in phenotypic performance of cattle across all economically relevant traits. When understood, this variation can be effectively managed to optimize production efficiency. Genomic technologies represent an important tool by which we can understand the extent of genetic variation within and among populations.

Acknowledgement

The author is employed by Zoetis and serves as the global head of the Genetics business unit.