Genomics: What do we need to know?

Jason B. Osterstock, *DVM, PhD*
Zoetis, 333 Portage St, Kalamazoo, MI 49007

Abstract

Use of genetic testing as an adjunct to traditional selection methods is becoming increasingly common. Today, the ratio of female-to-male genotyped Holstein cattle in the USDA Council on Dairy Cattle Breeding dairy genetic evaluation is 5 to 1. Producers are using this information to inform mating, heifer culling, and allocation of advanced reproductive technologies such as sexed semen, embryo transfer, and *in vitro* fertilization. The motivation to do so is not merely interest in new technology, but the impact that genomic data can have on the efficacy of genetic selection strategies, ultimately improving the rate of genetic progress and profitability. Veterinarians have an opportunity to be a part of this discussion, particularly because they actively manage so many of the non-genetic elements that influence the ability to design, implement, and benefit from genomically-enhanced selection. Importantly, the discussion around genomic technologies need not be complex. Sufficient context and understanding can be gained with just a few key concepts including common strategies, the impact of reliability on selection, and the logistics of dairy genetic evaluation systems.

Key words: bovine, genomics, genetics, genotyping

Résumé

L'utilisation des tests génétiques en lien avec les méthodes traditionnelles de sélection devient de plus en plus fréquente. À ce jour, le rapport femelle/mâle chez les bovins laitiers Holstein génotypés dans le programme d'évaluation génétique de l'USDA Council on Dairy Cattle Breeding est de 5 pour 1. Les producteurs utilisent cette information afin d'éclairer leur choix pour l'accouplement, la réforme des génisses et l'utilisation des techniques de reproduction avancées incluant la semence sexée, le transfert embryonnaire et la fertilisation in vitro. La motivation sous-jacente n'est pas seulement l'intérêt pour les nouvelles technologies mais aussi l'impact que les données génomiques pourraient avoir sur l'efficacité des stratégies de sélection génétique visant à améliorer le taux de progrès génétique et la rentabilité. Les vétérinaires ont la chance de faire partie de cette discussion car ils gèrent déjà plusieurs des aspects non-génétiques qui influencent la capacité de planifier, de mettre en œuvre et de rentabiliser la sélection génétiquement améliorée. Il est aussi important de réaliser que la discussion entourant les technologies génomiques n'a pas besoin d'être compliquée. Une mise en situation adéquate et une bonne compréhension nécessitent seulement quelques concepts clés incluant les stratégies communes, l'impact de la fiabilité sur la sélection et la logistique des systèmes d'évaluation génétique des bovins laitiers.

Introduction

Use of genomic technologies is becoming increasingly common in commercial dairy production. This applies to selection of both male and female breeding animals, as well as management of breeding programs, mating, and prioritization of use of other reproductive technologies. The principal benefit of genetic testing in dairy cattle is increased reliability of resulting predictions of genetic merit relative to traditional methods of evaluation. This is derived through genomic verification of pedigree and knowledge of the effects of individual markers on predicted performance. Dairy genetic evaluations that incorporate genomic information do so in a way that complements other sources of information including pedigree, performance, and progeny data. In this way, genomic technologies do not replace traditional evaluation methods, but instead apply another layer of data to our understanding of the genetic potential of animals in the herd, or about to enter the herd. This additional information allows us to make more effective decisions about how best to manage genetic potential, resulting in accelerated genetic progress, improved performance, and improved profitability.

Gathering Data

The process by which genomic predictions are derived is quite complex, in large part because we are seeking to understand something that is inherently complex. For purposes of this discussion, we will keep it relatively simple. There are 4 potential sources of information that may contribute to our ability to predict the genetic merit of an individual animal: pedigree, performance, progeny performance, and genomic information.

Pedigree data allows one to estimate genetic potential based on available estimates of genetic merit in the sire and dam. There are 4 potential weaknesses to relying on pedigree data alone. First, it assumes we have pedigree information, which is not always the case, particularly in commercial dairies where there is often missing data. Second, it assumes we know the genetic merit of the sire and dam. We often encounter missing data here as well, particularly on the dam side of the pedigree in herds that do not register animals and do not record performance data through a Dairy Herd Improvement affiliate. In some situations, we can compensate to some degree by using the maternal grandsire as a proxy, but that is an imperfect solution. The third challenge is that in order for pedigree information to be useful, it needs to be correct. In 2013, more than 50,000 of the animals submitted to the US Department of Agriculture (USDA) – Council on Dairy Cattle Breeding (CDCB) dairy genetic evaluation with genomic data were determined to have an incorrect or missing sire.[3] This figure is approaching 25% of submissions. If the pedigree is recorded incorrectly, then pedigree estimates are quite limited. Finally, we know that even though we may have a good idea of the genetic merit of parents, sometimes offspring over-perform or under-perform relative to expectation. On average and with sufficient numbers of progeny, the progeny of a given sire and dam will perform at or near the average of the parents. However, some animals get more or less than their fair share of the available gene pool. This is a significant limitation to using estimates of genetic potential derived from pedigree data alone.

Performance of animals and their progeny are very important contributors to estimates of genetic merit, and really reflect the gold standard for establishing the genetic potential of cattle. Historically, this has driven the proofs we have used in AI sires and bull dams. By quantifying the actual performance levels for traits evaluated, we can gain a much better appreciation for the genetic potential an individual animal possesses or has transmitted to their progeny. The downside to this data, particularly in females, is the time required to obtain it. By the time we have sufficient progeny data in a typical dairy cow, for example, she is likely long gone, if she had enough progeny to do so at all. Effectively, this data doesn't become available until it is too late to make any meaningful selection decisions.

Genomic information is well-suited to help complement these other sources of information by filling the gaps noted. We can use genetic markers to verify assigned parentage or even discover missing parents, as is the case for AI-sired calves with an unknown sire that has a genotype in the evaluation. We can also use genomic data to more accurately predict genetic potential in animals that lack performance or progeny data, thus enabling selection decisions much earlier in life. The effects of individual genetic markers are estimated from a reference population and then used to estimate a genomic breeding value using complex, but well established, statistical methods. The resulting genomic estimate (e.g., Figure 1) is then blended with available pedigree, performance, and progeny data, where available, to derive a final summary estimate of genetic merit. Genotypes provide insight into the genetic material that was inherited and from whom, providing very specific information about the genetic potential of an individual animal.

Gaining Perspective

As there are several very different sources of information contributing to our estimates of genetic potential, we need some way of quantifying how much we actually know based on the available data. In dairy cattle, we use the term *reliability* to describe how much we know and therefore how confident we are in the information. Reliability is actually a statistical parameter estimated from the correlation between the prediction and the animal's true, but unknown, genetic merit. It varies from 0 to 1, with higher values reflecting animals for which we have a lot of information, generally only achieved in AI sires with thousands of progeny (Table 1).

The important thing to understand about reliability is the practical implications. When making genetic selection decisions, we have to rank or group animals for an outcome we can't see using our best available predictions of that outcome. If we have no information, we can't rank or group the animals for the desired outcome. That's like having a reliability of 0 (Table 1). As we acquire additional sources of information, we can begin to rank or group the animals. If we rank or group the animals in a way that closely matches that outcome of interest (which we still can't see) then our strategies and decisions will be quite effective. Predictions with greater reliability rank animals more closely to that outcome of interest, resulting in more effective selection decisions and accelerated genetic progress.

Reliability is a bit like quantifying your perspective on a matter. Assume we had to choose a basketball team and all of the potential players were in a conference room sitting around a table. We know that height is an important trait associated with being a good basketball player. Not the only trait, but an important one. Let's imagine that in the first attempt to choose the team, the coach had his eyes closed. He had no perspective and would obviously struggle to choose the right players.

Now imagine that we choose again, but this time the coach opens his eyes. Now the coach has better perspective. However, the players are all seated so we have better, but still limited perspective. What if one of the taller players was slouching or had his head on the table? We might miss him. What if one of the shorter players was sitting on his knees or duffle bag? We might incorrectly pick him.

Now imagine that the players all stand up. In this case, the coach now has much better perspective to make the choices. There may still be mistakes in the selection

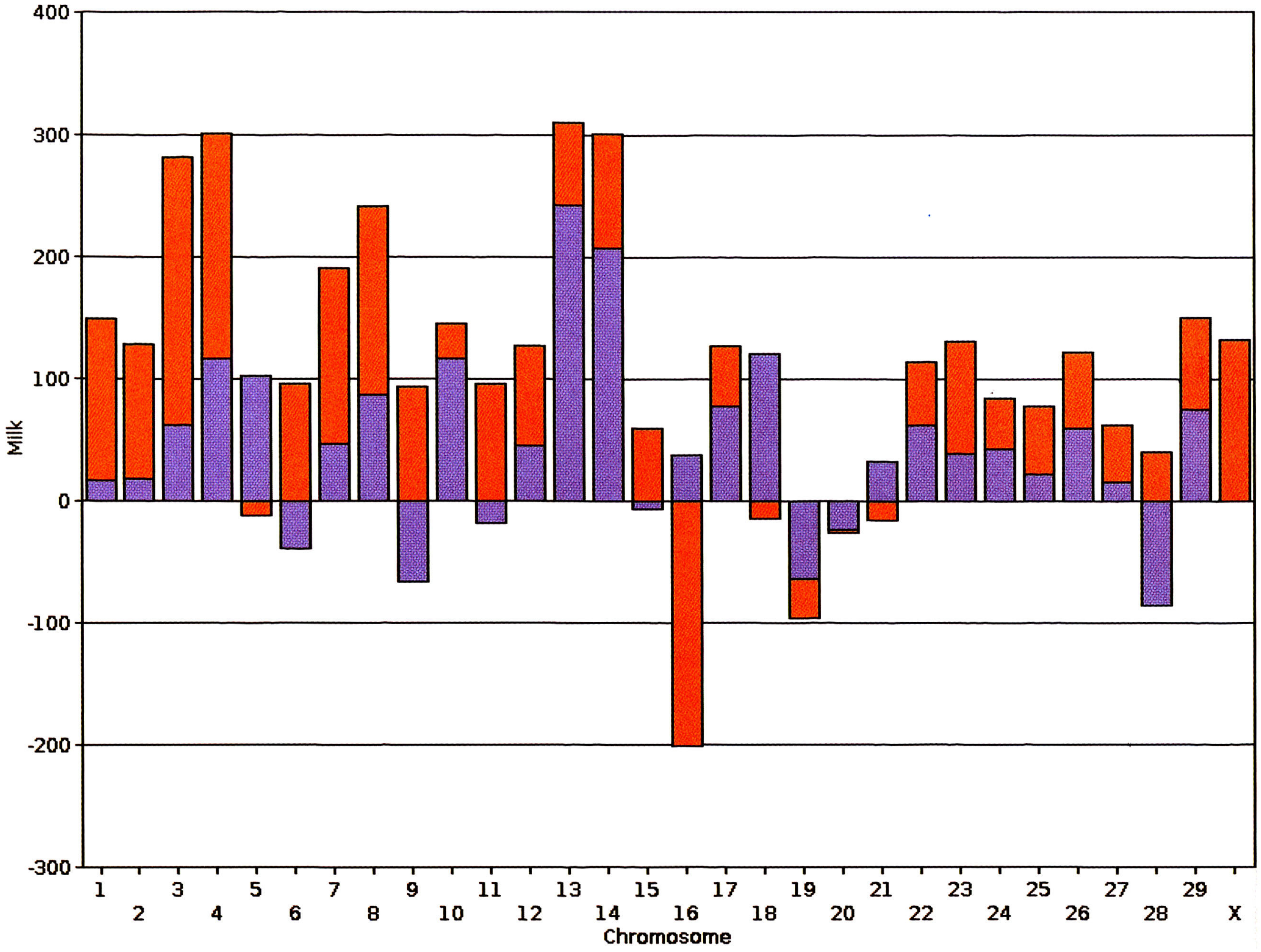

Figure 1. Chromosomal PTA for ENSENADA TABOO PLANET-ET (HOUSA000060597003) for milk. The figure depicts a total genomic transmitting ability of +1976 including contributions from genetic material inherited from the sire (+754, blue) and the dam (+1223, orange). The genomic data is blended with other available data yielding an official predicted transmitting ability (PTA) of +2105. This figure illustrates how genomic data from all of the chromosomes, representing genetic material inherited from each parent, contributes to a single summary estimate of genetic potential. (CDCB, 2014. Available at: https://www.cdcb.us/CF-queries/Bull_Chromosomal_EBV/bull_chromosomal_ebv.cfm? Accessed May 15, 2014.)

process, but with the improved perspective, we will be right more often than not. The same applies to selecting dairy cattle. Reproduction isn't the only thing that makes a cow profitable, but it's a part of it. We can't see an animal's potential to be reproductively sound when it is 4 months old, but if we have genomic information, we can start to sort them out and make better selection decisions.

Common Applications

Despite all the complexities around how genetic predictions are derived and how the accuracy of the respective types of predictions might be estimated, the application can be quite simple. In fact, there are only 3 possible decisions that can be made that impact genetic improvement:

1. Do I want this animal? – this applies to selection of home-raised replacements, purchase of animals from other herds, or the purchase of semen from a particular bull. The question here is whether we want that individual animal's genetics to contribute in any way to herd performance and profitability.
2. How many progeny do I want from this animal? – this question can have many answers, but simply put, they are many, some, or none. The answer to this question influences whether we might consider a female as a donor in a multiple ovulation and embryo transfer (MOET) or in

Table 1. Examples of typical reliabilities for Net Merit (NM$) in dairy cattle based on available information. (CDCB, 2014, https://www.cdcb.us/cgi-bin/general/Qpublic/do.Q.cgi?qname=getbull&single, accessed May 15, 2014.) (Zoetis, 2014. Data on file.)

Description	Typical traditional NM$ reliability
Registered Holstein bull with more than 30,000 daughter records	0.99
Non-registered Holstein female with 2 lactation records	0.45
Non-registered Holstein female < 12 months of age with genomic data	0.68
Non-registered Holstein female < 12 months of age with verified pedigree, but no genomic data	0.22
Non-registered Holstein female with no pedigree or genomic data	0

vitro fertilization (IVF) program, how we might allocate sexed semen, or perhaps how we might allocate beef semen should that be part of the herd's management strategy.

3. Who should I mate this animal to? – this question is really about planning for the next generation. Each animal has strengths and weaknesses relative to the specific breeding objectives of the dairy. Preferentially choosing mates that accentuate strengths and help correct weaknesses is an important component of the use of genomic data. In addition, mating strategies can help producers breed around deleterious recessive conditions (e.g., fertility haplotypes), prioritize unique genetic composition (e.g., horned/polled), and proactively manage inbreeding.

Given these choices, we next need to consider the basic strategy. At the risk of stating the obvious, we want more of the best, more progeny from the best, and we want everyone mated up correctly. Being able to execute this now depends upon our ability to define who the best are, which is a ranking / grouping exercise. As described previously, reliability has a significant impact on our ability to rank animals correctly and the use of higher reliability predictions, such as genomic-predicted transmitting abilities (GPTA), provides significant benefit. When done correctly, the result of this genomically-enabled decision making is accelerated genetic progress. The animals destined for production, in this generation and the next, possess greater average genetic merit than would have been achieved otherwise using traditional methods.

In non-registered, commercial dairy production, the most common applications include culling strategies and selective allocation of sex-sorted semen. Culling strategies are generally most appropriate for herds with good reproduction and calf health that are not in a position to expand and would like to avoid overcrowding. Historically, all female progeny have been retained as replacements. In herds with strong reproduction and calf health programs, this often means that the number of heifers available to enter the herd exceed the replacement requirements of the dairy. This scenario may lead to overcrowding or more aggressive culling of adult cows to make room for the heifers coming. The challenges associated with the former are detailed elsewhere.[2] The challenges associated with the latter are 2-fold. First, the strategy assumes that the heifers coming in represent a genetic upgrade that exceeds the negative impacts of reduced first-lactation production relative to mature cows. Importantly, this is not always the case as there is generally a substantial proportion of the heifer population with genetic potential well below herd averages (Figure 2). Second, this strategy may increase total cost of production by increasing replacement costs as a percentage of whole-herd cost of production per unit of milk produced.

Heifer culling strategies are best employed with careful consideration of the market to sell the excess. On the surface, the best decision may be to remove them from the heifer inventory as early as possible, thus reducing the dollars spent raising heifers that don't fit the herd's objectives. However, in some situations, available markets to sell excess heifers may favor raising them to breeding age or springers. Regardless of the timing, the genetic selection strategy should remain the same – rank them as accurately as possible, identify the bottom end, and ship them.

Preferential allocation of sexed semen to better animals is, in my experience, a pretty common strategy employed by commercial dairy producers using genomic technology, alone or in combination with culling strategies. It has the potential to drive a greater proportion of the next generation of females from those heifers with greater genetic merit for traits of interest. Based on statistical modeling, one can produce about the same number of heifers by using sexed semen only in the top 50% of heifers for 2 or more breedings as would be achieved using sexed semen for a single breeding followed by conventional in all heifers.

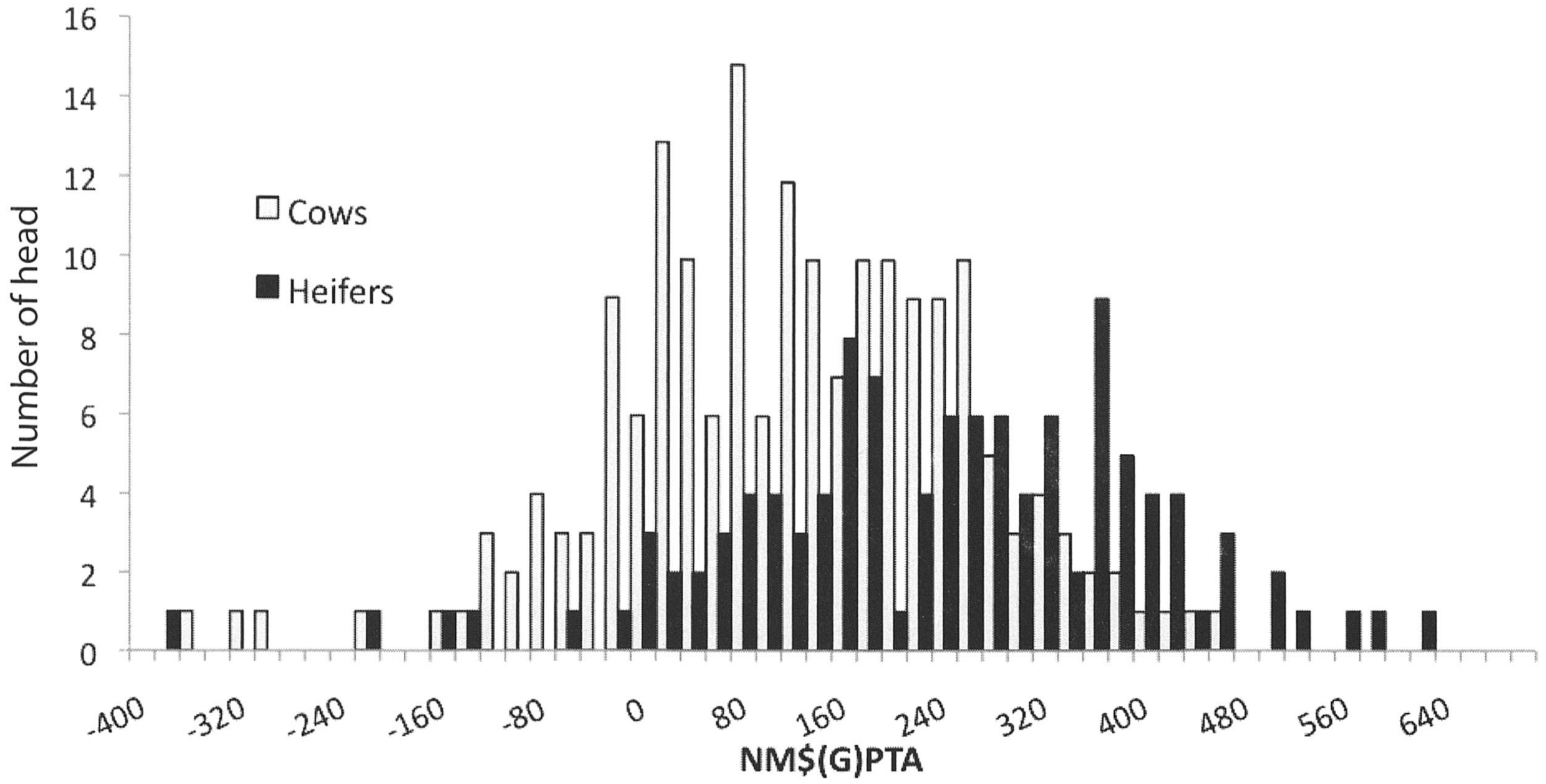

Figure 2. Example distribution of net merit (NM$) from a Holstein herd demonstrating the overlap in genetic merit between heifer and adult populations. Note that while on average the heifers are superior, there are some heifers that fall well below average relative to the cow herd. (Zoetis, 2014. Data on file.)

Common Challenges

As with any technology, genetic testing does present some challenges. The biggest is death by data. While perhaps not an ancient proverb, it does hold very true. The USDA-CDCB dairy genetic evaluation produces a wealth of information on a variety of traits. There are production traits, calving traits, health and longevity, reproduction, linear type traits, milk components, parentage data, and a host of selection indexes and composite traits. This information can be overwhelming for dairy producers, and in some cases paralyzes them from executing the planned decisions. There are a couple of key approaches that should be employed up front to prevent this analysis-paralysis phenomenon. First, establish a plan before testing so that we know what predictions will influence selection decisions and which will be ignored or only used for mating. Second, focus on the available indexes including Total Performance Index (TPITM) from Holstein Association USA or Net Merit (NM$). Indexes are developed based on careful assessment of the contribution of individual component traits to comprehensive genetic improvement and profitability. To that end, there is considerable science behind the process, thus providing confidence in using these as principal selection criteria.

Another significant challenge is implementation without upsetting the routine work and standard operating procedures on the dairy. With respect to sample collection, this can be avoided by identifying an opportunity for sampling that fits well within other management decisions and actions in advance of the timing of the selection decision. A more difficult logistical issue to overcome is related to animal identification. Recall that the dairy genetic evaluations have both genomic data and production records coming in. It is imperative that the animal identification numbers used for both are the same for a given individual in the evaluation, and unique for every animal in the evaluation. To that end, one cannot request a genomic evaluation as a heifer calf and then assign a new number later in life that is used when reporting milk records. This undermines the entire system. For dairies that utilize permanent, official identification from birth, there is no problem. For dairies that do not utilize official identification systems in heifer calves, plans must be made to modify identification strategies to ensure that problems with animal identification do not interfere with successfully obtaining the genomic results or threaten the integrity of the evaluation.

There are limits to the current evaluations. An important one is breed, whereby only purebred Holstein, Jersey, Brown Swiss, and Ayrshire animals are supported. Crossbred animals cannot be tested and in most instances, an animal needs to be at least 7/8 of a supported breed to qualify. There are also important caveats when testing males and in herds not participating in herd recording, including additional fees that are in addition to any genotyping costs.[1]

Use of genomic testing in dairy cattle production has increased considerably in the past 2 to 3 years (Figure 3). Drivers of that adoption have largely been the technology's power to provide timely and accurate data to dairy producers faced with a variety of selection decisions, and the opportunity for some producers to make strategic decisions regarding heifer inventory management and use of complementary reproductive technologies.

Conclusions

For practicing veterinarians, there is little need to become intimately familiar with the nuances of dairy genetic evaluation and quantitative genetics unless one seeks to play an active, consultative role for client dairies. However, a basic understanding of the process, including the types of decisions that can be made regarding genetic improvement and the principal benefits of genetic testing, are important in helping producers understand the opportunities to leverage genomic technology. In many instances, veterinarians are in a unique position to understand and affect key drivers of heifer inventory, reproductive success, and performance. All of these elements can help shape genomic strategies on commercial dairies. In addition, a veterinary advisor may be able to help appropriately position genomic technologies as part of a comprehensive strategy to improve herd performance and profitability.

References

1. CDCB, Fee schedule for genomic evaluation, 2013. Available at: https://www.cdcb.us/News/CDCB_Fee_Schedule_03_22_13_FINAL.pdf. Accessed May 19, 2014.
2. Schefers JM, Weigel KA, Rawson CL, Zwald NR, Cook NB. Management practices associated with conception rate and service rate of lactating Holstein cows in large, commercial dairy herds. *J Dairy Sci* 2010;93:1459-1467.
3. VanRaden P. Impact of genomics on genetic improvement, in *Advancing Dairy Cattle Genetics: Genomics and Beyond*, Phoenix, AZ, February 19, 2014.

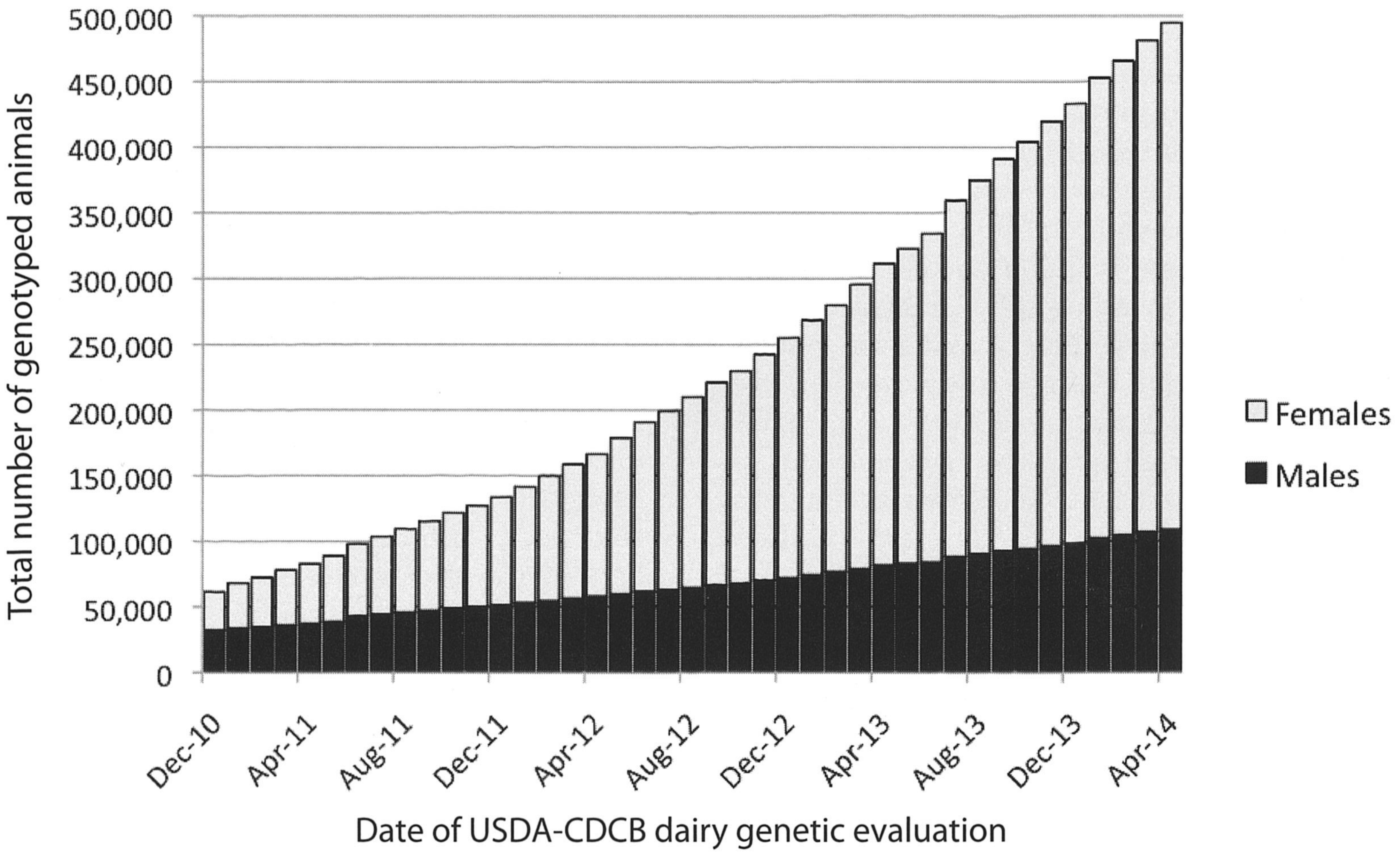

Figure 3. Cumulative number of genotyped animals by sex in USDA-CDCB dairy genetic evaluation over time since introduction of low-density genomic testing platforms in September 2010. (CDCB, 2014. Genotype counts: Total (as of 2014-04-28). Available at: https://www.cdcb.us/Genotype/counts.html. Accessed May 19, 2014.)