Invited review: Inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of genomic variability

Jeremy T. Howard,*[1] Jennie E. Pryce,†‡ Christine Baes,§ and Christian Maltecca*#
*Department of Animal Science, North Carolina State University, Raleigh 27695-7627
†Agriculture Victoria, Agribio, 5 Ring Road, Bundoora 3083, Australia
‡La Trobe University, Bundoora, Victoria 3086, Australia
§Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, 50 Stone Road E., Guelph N1G 2W1, Canada
#Genetics Program, North Carolina State University, Raleigh 27695-7627

ABSTRACT

Traditionally, pedigree-based relationship coefficients have been used to manage the inbreeding and degree of inbreeding depression that exists within a population. The widespread incorporation of genomic information in dairy cattle genetic evaluations allows for the opportunity to develop and implement methods to manage populations at the genomic level. As a result, the realized proportion of the genome that 2 individuals share can be more accurately estimated instead of using pedigree information to estimate the expected proportion of shared alleles. Furthermore, genomic information allows genome-wide relationship or inbreeding estimates to be augmented to characterize relationships for specific regions of the genome. Region-specific stretches can be used to more effectively manage areas of low genetic diversity or areas that, when homozygous, result in reduced performance across economically important traits. The use of region-specific metrics should allow breeders to more precisely manage the trade-off between the genetic value of the progeny and undesirable side effects associated with inbreeding. Methods tailored toward more effectively identifying regions affected by inbreeding and their associated use to manage the genome at the herd level, however, still need to be developed. We have reviewed topics related to inbreeding, measures of relatedness, genetic diversity and methods to manage populations at the genomic level, and we discuss future challenges related to managing populations through implementing genomic methods at the herd and population levels.

Key words: inbreeding, genomics

INTRODUCTION

Characterizing and efficiently managing inbreeding levels has been, and will continue to be, an important goal in dairy cattle-breeding programs to ensure that populations can adapt to future breeding goals while maintaining genetic diversity and avoiding the accumulation of detrimental effects associated with inbreeding. The use of genomic information as a tool to cost-effectively predict breeding values, referred to as genomic selection (Meuwissen et al., 2001), has dramatically changed dairy cattle breeding (Hayes et al., 2009). Because of genomic selection, an increasing number of selection candidates across both sexes are currently routinely genotyped. For example, the number of genotyped Holstein animals in the United States exceeds 1 million animals (https://www.uscdcb.com/Genotype/cur_ctry.html). The advent of genomic selection has reduced the generation interval and increased the selection intensity, which has dramatically increased the rate of genetic gain, especially for lowly heritable traits (García-Ruiz et al., 2016). Genomic information has shifted the landscape on how candidates are selected, although we argue that a similar degree of change has yet to be seen in the utilization of genomic information to manage dairy populations. Annual rates of genetic gain of US Holstein females assessed using the net merit selection index have nearly doubled for cows born since 2010 (approximately when genomically tested sires became widely used) compared with the previous decade (https://www.uscdcb.com/eval/summary/trend.cfm). As shown by Miglior and Beavers (2014), the introduction of genomic selection has resulted in a more diverse set of bulls sampled for testing, yet the number of bulls siring 50% of the young bulls entering AI has remained relatively constant. Thus, although a more diverse set of bulls are screened, this is not translating into more genomically diverse bulls being offered to breeders.

In general, the majority of methods to manage a population and its associated effective population size are related to the coancestry among the parents used to

obtain the next generation, while taking into account the expected genetic value of the progeny (Kristensen and Sorensen, 2005). Prior to genomic information, pedigree relationships were used to manage inbreeding within a population (VanRaden and Smith, 1999). However, pedigree information is based on the expected proportion of the genome that is identical by descent (**IBD**) between 2 parents and thus does not capture the variation due to Mendelian sampling and linkage during gamete formation (Franklin, 1977; Hill, 1993; Guo, 1996; Hill and Weir, 2011). Genomic information allows an estimate of the realized proportion of the genome that 2 individuals share, either genome-wide, or at specific regions (Hill and Weir, 2011). The implications of this difference will be discussed throughout the paper.

The use of region-specific metrics that characterize the genetic diversity (i.e., effective population size) and inbreeding depression at specific regions can be used along with commonly estimated genome-wide metrics. Thus, the matings of a sire with potential dams with similar coancestry values at the genome-wide level might be further discriminated on the basis of haplotypic diversity within low-diversity regions. This is expected to have an effect on the genomic load carried by the resulting offspring. The use of region-specific metrics should allow breeders to more effectively manage the risks associated with sire or mate selection, allowing one to better evaluate the trade-off between the genetic value of the progeny and undesirable side effects associated with inbreeding. Lastly, as time progresses, a larger number of new factors will come into play during sire or mate selection. The inclusion of dominance effects (Sun et al., 2013), the exploitation of variation in recombination across the genome (Ma et al., 2015), the increase in use of advanced reproductive technologies (Ponsart et al., 2013), and introduction in the future of gene-editing procedures, among others, will affect the management of genomic diversity, both at the population as well as the farm level. Research on how to manage genetic resources based on all these factors simultaneously, rather than individually, should be a top priority.

In this paper, we review the current state and future challenges related to inbreeding, measures of relatedness, genetic diversity, and methods to manage populations at the genomic level. In the last section, we discuss the challenges of managing inbreeding at the individual herd level and future areas of research.

INBREEDING

The mating of related individuals is unavoidable in populations of finite sizes, as the number of ancestors increases exponentially per generation. The inbreeding coefficient of an individual is defined as the probability of 2 randomly chosen alleles, at the same locus from 2 gametes, which unite to be IBD from a common ancestor (Crow and Kimura, 1970). Another consequence of finite populations is the random fluctuation of gene frequencies (i.e., genetic drift) and the eventual loss of alleles due to the sampling process during gamete formation (Falconer and Mackay, 1996). Inbreeding leads to changes in the distribution of genetic variance, leading to allelic fixation and, in principle, reducing the additive genetic variance within a population as well as decreasing the rate of response to selection for a trait under selection and other traits (Kristensen and Sorensen, 2005). The nonadditive (i.e., dominance and epistatic) proportion of the genetic variance is highest at intermediate frequencies and, similar to the additive genetic variation, as alleles become fixed due to inbreeding the nonadditive variance also decreases. It should be noted though that, in principle, the genetic variance is expected to decrease across the long-term horizon, whereas within the short-term horizon it is possible for the additive genetic variance to increase as a result of nonadditive effects being converted to additive effects (Kristensen and Sorensen, 2005). Furthermore, increased levels of inbreeding result in a reduction in the mean value of a trait, referred to as inbreeding depression, which has been documented across all livestock species (reviewed by Leroy, 2014). Within dairy populations, the accumulation of inbreeding is unavoidable due to several factors. A few of these include multiple generations of intense directional selection (Robertson, 2007), high variance of reproductive success across individuals due to the use of advanced reproductive technologies (Nicholas and Smith, 1983), and the use of BLUP-based genetic evaluations in combination with truncation selection (Verrier et al., 1993). These processes lead to the widespread use of related individuals as parents of the next generation. The extent of inbreeding depression depends on allele frequencies and the strength of directional selection; therefore, it is expected to be population specific, as drift and selection pressure vary across populations. Lastly, fitness traits, such as survival, reproduction, and disease resistance, are more affected by inbreeding. It is thought that fitness traits exhibit a greater degree of directional dominance than traits distantly related to fitness (i.e., type traits or adult body size; Kristensen and Sorensen, 2005).

The Genetic Basis of Inbreeding Depression

Changes due to inbreeding on the phenotypic mean of a trait at n biallelic loci, and assuming no epistasis, can be estimated as