Symposium review: The genomic architecture of inbreeding: How homozygosity affects health and performance*

Christine F. Baes,1,2† Bayode O. Makanjuola,1 Filippo Miglior,1,3 Gabriele Marras,1,4 Jeremy T. Howard,5,6 Allison Fleming,1,3 and Christian Maltecca6

1Centre for the Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, Ontario N1G 2W1, Canada
2Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland
3Canadian Dairy Network, Guelph, Ontario N1K 1E5, Canada
4Semex Alliance, Guelph, Ontario N1H 6J2, Canada
5Department of Animal Science, University of Nebraska at Lincoln, Lincoln 68583-0908
6Department of Animal Science, North Carolina State University, Raleigh 27695-7627

ABSTRACT

Inbreeding depression is a growing concern in livestock because it can detrimentally affect animal fitness, health, and production levels. Genomic information can be used to more effectively capture variance in Mendelian sampling, thereby enabling more accurate estimation of inbreeding, but further progress is still required. The calculation of inbreeding for herd management purposes is largely still done using pedigree information only, although inbreeding coefficients calculated in this manner have been shown to be less accurate than genomic inbreeding measures. Continuous stretches of homozygous genotypes, so called runs of homozygosity, have been shown to provide a better estimate of autozygosity at the genomic level than conventional measures based on inbreeding coefficients calculated through conventional pedigree information or even genomic relationship matrices. For improved and targeted management of genomic inbreeding at the population level, the development of methods that incorporate genomic information in mate selection programs may provide a path forward. In this symposium review article, we describe traditional measures of inbreeding and the recent developments made toward more precise measures of homozygosity using genomic information. The effects of homozygosity resulting from inbreeding on phenotypes, the identification and mapping of detrimental homozygosity haplotypes, management of inbreeding with genomic data, and areas in need of further research are discussed.

Key words: inbreeding, runs of homozygosity, mating programs

INTRODUCTION

An important topic in dairy cattle breeding today is the state of genetic diversity in the population and the implications of its decline. Intensive selection within the dairy cattle population has resulted in remarkable genetic gains in several traits; in particular, those with economic value such as milk production and fat and protein content (Miglior et al., 2017). These gains are a result of management practices, such as improved feeding regimens, rations, housing, and so on, but also of breeding methodologies and computing, reproductive, and genotyping technologies. A notable consequence of genetic gain, however, has been an increasing trend in inbreeding levels. Fitness traits, or traits measuring reproduction (e.g., González-Recio et al., 2007; McParland et al., 2007; VanRaden et al., 2011), health (e.g., Kim et al., 2015), and survival (e.g., Thompson et al., 2000a,b), are particularly sensitive to inbreeding (Falconer and Mackay, 1996). This noted inbreeding depression has also been linked with reduced production performance in dairy cows (e.g., Pryce et al., 2014). As

Received August 8, 2018.
Accepted November 21, 2018.
*Presented as part of the Canadian Society of Animal Science Symposium: Genomic Alterations and Implications on Health—Gut and Beyond at the ADSA Annual Meeting, Knoxville, Tennessee, June 2018.
†Corresponding author: cbaes@uoguelph.ca
technologies emerge and afford new opportunities for genetic improvement of dairy cattle, it is imperative to consider the effect these tools could have on genetic diversity (Howard et al., 2017a; Fleming et al., 2018).

Historically, the introduction of novel tools for animal breeders has altered genetic diversity by not only changing the structure of the breeding population, but also changing how breeding candidates are selected (Dekkers, 2012). Artificial insemination technologies greatly increased the number of progeny produced by top dairy sires and lowered the effective population size (Brotherstone and Goddard, 2005). Superior AI sires currently have thousands of offspring; therefore, few individuals make large contributions to the next generation. Alternatively, at the individual herd level, AI has allowed producers to use multiple sires in their herds, which has potentially increased the within-herd diversity compared with natural mating (Young, 1984). The use of embryo transfer offers the same effect for the female population, but on a much less dramatic scale, giving elite dams the opportunity to produce more offspring than would naturally be possible (Fleming et al., 2018). Cloning, although not widespread, could also have the potential to influence population structure if implemented on a large scale (Butler and Wolf, 2010). Some novel reproductive technologies may also have epigenetic effects (Mc Parland et al., 2007), which are currently not well understood, as high cost still restricts their widespread use. The strategy for incorporating these technologies into routine breeding programs will have varying effects on the rate of inbreeding accumulation in the population, depending on the frequency and strategy with which they are implemented. These technologies may also, however, simultaneously be used as a way to conserve genetic diversity. Cryopreserved semen and embryos harbor historical genetic variability that could be reintroduced in the future with the aid of other reproductive technologies. In this sense, AI, embryo transfer, and cloning can serve to maintain genetic diversity when geared toward maximizing diversity as opposed to trait improvement alone.

Genomic selection (Meuwissen et al., 2001; Schaeffer, 2006) has clearly altered dairy cattle breeding. The most striking change was a considerable reduction in generation interval, which became possible through early and reliably estimated genomic breeding values compared with traditional parent averages (de Roos et al., 2007). Daetwyler et al. (2007) postulated early on that the rate of inbreeding per generation would be reduced through genomic selection, as Mendelian sampling variation would be more accurately captured through reducing co-selection of relatives.

Although this is true (see Figure 1), the much higher increase in the rate of inbreeding per year due to the shorter generation interval (Schaeffer, 2006) seems to negate these possible positive effects on the rate of increase in inbreeding (Forutan et al., 2018). As shown by Miglior and Beavers (2014), the number of bulls siring 50% of the young bulls entering AI has remained relatively constant despite a more diverse set of bulls being screened; therefore, bulls offered to breeders are not more genomically diverse than they were before genomic selection. Overall, genomic selection has dramatically increased the yearly rate of genetic gain, especially for lowly heritable traits (García-Ruiz et al., 2016). Although genomic information has shifted the landscape on how candidates are selected and the speed at which genetic progress is made, Howard et al. (2017a) argued that a similar degree of change has yet to be seen in the utilization of genomic information to manage genetic diversity and unfavorable inbreeding effects in dairy populations.

![Figure 1](image_url)

*Figure 1.* (a) Average increase in inbreeding per year from 1950 to 2018 in Canadian Holsteins. (b) Average increase in inbreeding per generation in Canadian Holsteins. Data provided by the Canadian Dairy Network, 2018.
Before genomic information, pedigree relationships were used to manage inbreeding within a population (VanRaden and Smith, 1999). In contrast to using pedigree information alone, the realized proportion of the genome that 2 individuals share can be estimated using genomic information, either genome-wide, or at specific regions (Hill and Weir, 2011). For this reason, genomic inbreeding coefficients are closer to the true inbreeding value than those calculated using only pedigree information (Forutan et al., 2018). Furthermore, region-specific metrics that characterize the genetic diversity and inbreeding depression could potentially be implemented together with genome-wide metrics (i.e., effective population size) to more precisely control inbreeding. For example, Charlier et al. (2008) narrowed down the probable cause of congenital muscular dystonia 2 in Belgian Blue cattle to a 3.61-Mb interval on chromosome 29 when present in its homozygous form. In addition, specific chromosomal regions that resulted in the reduction of volume and concentration of ejaculate in Brown Swiss population have been identified (Ferenčaković et al., 2014; Curik et al., 2017). In this way, potential matings of a sire to different dams, with similar haplotypes at the genome-wide level, could be further discriminated by maximizing haplootypic diversity, which is expected to affect the genomic load carried by the resulting offspring. With availability and accessibility to genomic information, region-specific metrics would allow more effective management of risks associated with sire or mate selection through the identification and mapping of regions that contribute to inbreeding depression (Kardos et al., 2016; Curik et al., 2017). If such information were available, breeders would be in a better position to evaluate the trade-off between increasing the genetic value of progeny and undesirable side effects associated with inbreeding (Howard et al., 2017a).

As time and technologies progress, new factors will become relevant during sire or mate selection. The inclusion of dominance effects in breeding value estimation (Sun et al., 2013), the exploitation of variation in recombination across the genome (Ma et al., 2015), the increased use of advanced reproductive technologies (Ponsart et al., 2013), and the future introduction of gene editing procedures, among others, will affect the management of genomic diversity, both at the population as well as at the farm level (Cole et al., 2016). Investigating how to manage genetic resources based on all these factors simultaneously, rather than individually, should be a research priority.

In this symposium review, we discuss the current state and future challenges related to inbreeding, future areas of research, and how homozygosity affects health and productivity phenotypes in dairy cattle using a Canadian case study. In the last section, we discuss managing inbreeding using genomic information.

MEASURES OF INBREEDING

Inbreeding is the probability that 2 alleles sampled at the same locus from an individual are identical by descent (Wright, 1922; Malécot, 1948). Inbreeding has traditionally been measured using pedigree information, which relies on parentage relationships and the correctness of the recorded pedigree. The advent of genomic data has presented the opportunity to dissect the genome, and this allows for estimation of realized inbreeding by capturing Mendelian sampling as opposed to the expectations estimated from pedigree data (Franklin, 1977; Hill and Weir, 2011). Various methods have been proposed, developed, and implemented to estimate inbreeding coefficients of millions of animals using both pedigree and genomic data in a computationally efficient manner (Meuwissen and Luo, 1992; Colleau, 2002; Sargolzaei et al., 2005; VanRaden, 2008). Here, we briefly outline different methods commonly used to estimate inbreeding.

When using pedigree data, the numerator relationship matrix (A) is constructed, which is a covariance matrix that is an estimate of the expected proportions of genes shared by individuals in the pedigree. An algorithm developed by Meuwissen and Luo (1992) allows for the efficient decomposition of the A matrix to estimate inbreeding coefficients for individual animals. With this method, the inbreeding coefficient is usually underestimated when parents are missing, or if the pedigree is incomplete. The formula is

\[ F_{PED_i} = \frac{\sum_{j=1}^{n} L_{ij}^2 D_{ij}}{2} - 1, \]

where \( F_{PED_i} \) is the pedigree inbreeding coefficient of the \( i \)th individual, \( L_{ij} \) is the lower triangular matrix of the numerator relationship between the \( i \)th and \( j \)th individual, and \( D_{ij} \) is the diagonal element of the \( j \)th individual. VanRaden (1992) published a method that accounted for missing or unknown parents by assuming that the relationship between unknown parents and other known parents is twice the average inbreeding coefficient of progeny with known parents. Using this method, Stachowicz et al. (2011) was able to recover information on inbreeding coefficients from missing pedigree.

With genomic data, an approach similar to that used with the A matrix, termed the genomic relationship matrix (G), was adapted from research in human data by VanRaden (2008) to make it amenable to animal breeding. The off-diagonals of the G matrix contain the
actual proportion of alleles shared among individuals, whereas inbreeding coefficients are calculated from the diagonal of $G$ as

$$F_{G_j} = \frac{\sum_{i=1}^{n} z_i z_j'}{\sum_{i=1}^{n} 2p_j (1-p_j)} - 1,$$

where $F_{G_j}$ is the genomic inbreeding of the $j$th individual; $z_i$ is a vector with elements $2 - 2p_j$, $1 - 2p_j$, and $0 - 2p_j$ for the first homozygote, heterozygote, and the second homozygote, respectively; $z_j'$ is the transpose of $z_j$; and $p_j$ is the second allele frequency of the $j$th individual. Inbreeding coefficients estimated from the $G$ matrix are often overestimated due to the inability to discriminate between alleles that truly descended from a common ancestor (i.e., identity by descent, identical by descent, or autozygosity) and alleles that are identical by chance (i.e., identity by state). Inbreeding was explained earlier and is highly correlated with stretches of unbroken homozygous regions within the same individual (McQuillan et al., 2008; MacLeod et al., 2009); these stretches are often referred to as runs of homozygosity (ROH) and are considered good indicators of autozygosity. Additionally, the length of the ROH provides information on whether a ROH segment is the result of recent (long ROH) versus more distant (short ROH) inbreeding (Ferenčakovič et al., 2011; Purfield et al., 2012). Following the application of ROH for inbreeding coefficient estimates in human populations (McQuillan et al., 2008), Ferenčakovič et al. (2011) were the first to report ROH inbreeding estimates in cattle population, where they showed a moderate correlation that ranged from 0.61 to 0.68 between pedigree and genomic inbreeding depending on the length of ROH and depth of the pedigree. The ROH inbreeding can be calculated as

$$F_{ROH_j} = \frac{L_{ROH_j}}{L},$$

where $F_{ROH_j}$ is the genomic inbreeding for the $j$th individual using ROH, $L_{ROH_j}$ is the total length of ROH in the $j$th individual, and $L$ is the total length of the genome covered by the available genomic panel to be used.

**HOMOZYGOUSITY AND ITS EFFECTS ON PHENOTYPES**

The loss of heterozygosity is a consequence of increasing homozygosity and inbreeding, which, in turn, contributes to inbreeding depression. Inbreeding depression is the reduction in the mean value of a phenotype in a given population (Falconer and Mackay, 1996), and is mainly dependent on directional selection and allele frequencies. The degree of inbreeding depression is generally population specific, as drift and selection pressure vary across populations (Howard et al., 2017a). Charlesworth and Willis (2009) stated that the genetic cause of inbreeding depression has 3 possible foundations: (1) increased expression of detrimental recessive alleles (partial dominance); (2) the superiority of heterozygotes compared with homozygotes (over-dominance; Crow and Kimura, 1970); and (3) the increased possibility of favorable gene combinations in heterozygotes (epistasis). In a review by Leroy (2014), which included a meta-analysis of 57 studies comprising 7 livestock species, the authors showed that for every 1% increase in inbreeding there is a corresponding 0.14% reduction in the mean value of a trait. Miglior et al. (1992) observed an inbreeding depression of $-9.84$ kg, $-0.55$ kg, and $-0.0011\%$ for milk, fat, and fat percentage, respectively, per 1% increase in inbreeding coefficient in Canadian Jersey cows. However, inbreeding depression was more pronounced when inbreeding coefficients were greater than 12.5%. Financial economic losses associated with inbreeding depression for Holstein populations were reported by Smith et al. (1998). Those authors reported a $14.79$ reduction in fluid market pricing and a $12.40$ drop in manufacture pricing per 1% increase in inbreeding after adjusting for opportunity cost in Holstein cows with a 7-yr opportunity for herd life. Fitness traits such as survival, reproduction, and disease resistance are generally more affected by inbreeding than other traits because they exhibit a greater degree of directional dominance than traits distantly related to fitness (i.e., type traits or adult body size; Kristensen and Sørensen, 2005). Inbreeding depression is particularly important in dairy cattle because the intensity and speed of current selection programs have caused a rapid increase in relationships between individuals (Rodero-Serrano et al., 2013). Despite the importance of inbreeding depression, the main focus of most breeding programs is still on production traits (Gandini et al., 2014), catalyzing a slow, steady increase in inbreeding levels over time (Sørensen et al., 2005; Stachowicz et al., 2011).

Pedigree information has been used to investigate inbreeding depression, although this approach is severely limited because inbreeding calculated using pedigree information relies on Mendelian sampling probabilities (and not the realized homozygosity) and because it can only provide genome-wide expected inbreeding estimates. Furthermore, pedigree information may contain errors, which results in incorrect estimations of inbreeding. In a study by Wiggans et al. (2012), par-
entage confirmation with genomic information showed that up to 13.9% of sires were incorrectly recorded. Nevertheless, pedigree information has been used in several studies to calculate inbreeding depression of phenotypes by regressing the phenotype of interest on the inbreeding value and assuming that inbreeding depression is a linear function of an individual’s level of inbreeding (Howard et al., 2017a). Some studies have shown nonlinear inbreeding depression, such that a faster rate of depression is seen at higher levels of inbreeding (Thompson et al., 2000a,b; Wall et al., 2005; Gulisija et al., 2007); however, it is worth noting that the presence of inbreeding depression estimated from a model that includes either an additive by dominance or dominance by dominance interaction may not be accurate when using pedigree information, and more so with traits that are controlled by few loci with large effects (Curik et al., 2001). Moreover, it is difficult to determine whether the observed nonlinearity is due to true epistatic effects or statistical artifacts when dependencies of data showing high inbreeding levels are not accounted for (Lynch and Walsh, 1997; Curik et al., 2001); this is an area which needs further study. Furthermore, as described in Howard et al. (2017a), genome-wide inbreeding metrics do not account for the fact that homozygosity at some regions might be more detrimental than others. Consequently, 2 individuals with the same genome-wide inbreeding value could differ in inbreeding load. This could imply the presence of heterogeneous genetic load among founder ancestors, thus resulting in individuals with similar inbreeding coefficients showing varying levels of inbreeding depression (Miglior et al., 1994; Gulisija et al., 2006). This can be seen as an animal’s deviation from the regression line of its inbreeding level on a phenotype. Therefore, to properly account for heterogeneity among founder ancestors, a shift from a common slope to a slope that includes the interaction of common ancestor is more appropriate (Miglior et al., 1994; Gulisija et al., 2006).

In the following sections, we focus on specific trait areas and provide a short review of literature detailing how homozygosity affects phenotypes in those areas.

**Production Traits**

General consensus exists that high levels of inbreeding negatively affect milk production. In Holsteins, effects of inbreeding on production range from −19.7 to −29.6 kg (Mi et al., 1965; Miglior et al., 1995b; Wiggans et al., 1995; Croquet et al., 2006). Mc Parland et al. (2007) found the effect of inbreeding in Irish Holstein-Friesian cattle was negligible, at $F_{PED} < 6.25\%$, but much more pronounced above that threshold; first lactation Holsteins with $F_{PED} \geq 12.5\%$ were estimated to produce 160 kg less milk, 6.0 kg less fat, and 4.8% less protein than their non-inbred counterparts. More recently, Pryce et al. (2014) estimated inbreeding levels in a subset population of Holstein and Australian Jersey using both pedigree and genomic information. Their results showed that increasing inbreeding by 1% resulted in a negative effect on production traits by a reduction of 0.4 to 0.6%. It should be noted that the negative effects of inbreeding are cumulative and likely nonlinear, and, although the decrease in milk production and content traits may seem small per lactation, the long-term effects of accumulating inbreeding are clearly economically detrimental.

**Reproduction and Fertility Traits**

**Female Fertility.** A large number of studies have investigated the effects of inbreeding on reproduction and fertility traits in dairy cattle. Most have concluded that homozygosity has a small but clear association with reduced female fertility. González-Recio et al. (2007) found significant effects of inbreeding on pregnancy rate (cows $F_{PED} > 6.25\%$ had a pregnancy rate 1.68% lower than lowly or non-inbred animals), but the rate of inbreeding only affected calving ease in cows with $F_{PED} \geq 12.5\%$, and the result was not significant (+0.71%). Mr Parland et al. (2007) found a greater incidence of stillbirths (+0.06% per 1% increase in inbreeding), longer calving intervals (+0.7 d), higher age at first calving (+0.2 d), and lower survival (−0.3%) in inbred animals. Pryce et al. (2014) also found that increasing inbreeding by 1% resulted in an increase in calving interval from 0.02 to 0.05%. Using only pedigree information, Cassell et al. (2003) observed an undesirable but nonsignificant effect of maternal and fetal inbreeding on 70-d nonreturn rate for Holsteins across parities, whereas a significant and undesirable effect of maternal inbreeding for Jerseys was found to cause a 3% reduction in 70-d nonreturn rate per 10% increase in inbreeding.

Martikainen et al. (2017) found significant evidence of inbreeding depression on female fertility in Finnish Ayrshire (0.5-d interval from first to last insemination in heifers and first-parity cows, 1.4-d interval from first to last insemination in 2+ parity cows, and 0.3-d interval from calving to first insemination per 1% increase in $F_G$) but only using genomic information (pedigree-based inbreeding coefficients did not show inbreeding depression in any of the traits analyzed).

Less consensus exists when considering the negative effects of inbreeding on embryo viability. Bjelland et al. (2015) compared empirical genomic inbreeding values with expected inbreeding values using simulated data
and a pedigree of 11,484 Holsteins and found only negligible effects of inbreeding on embryo viability. In contrast, Perez et al. (2017) found that increased $F_{PED}$ of the donor significantly reduced the number and the percentage of both viable and high-grade oocytes, as well as the number and percentage of cleaved embryos. Furthermore, increased $F_{PED}$ of the embryo also had significant effects on viability.

**Male Fertility.** In terms of ROH affecting male fertility, Ferencišković et al. (2017) found that inbreeding depression did not affect the percentage of live spermatozoa but did affect the number of spermatozoa. Dorado et al. (2017) found a significantly higher number of highly active but nonprogressive spermatozoa in highly inbred bulls, and also found that cows sired by inbred bulls had a slightly higher calving interval.

Most genes responsible for maleness, and therefore male fertility (including testis development and initiation and maintenance of spermatogenesis in adulthood), in humans are located on the Y chromosome (Colaco and Modi, 2018); it can therefore be postulated that the same is true for cattle. This has 2 main implications for identifying the effects of inbreeding on fertility. (1) Information on sex chromosomes is normally filtered out in most genomic data sets currently available due to its low quality, effectively blocking any in-depth analysis of loci with biological effects on both female and male fertility. (2) Because males are heterogametic, it may be futile to search for ROH in the sex chromosomes in males. Improving the quality of genomic information for sex chromosomes in both sexes could contribute to better understand fertility in livestock.

**Lethal Recessive Mutations**

Inbreeding increases the frequency of homozygous loci and several lethal recessive mutations have been observed in dairy cattle populations as a result. Several major recessive defects affecting embryo or fetal survival have been identified (Shanks and Robinson, 1989; Shuster et al., 1992; Agerholm et al., 2001). Even with very large sets of phenotypic and pedigree data, defects that cause embryo death are difficult to detect without genomic data as a result of the limited number of carrier by carrier matings (VanRaden and Miller, 2006). However, Man et al. (2007) developed an approach that uses pedigree data to estimate the expected frequencies of carriers and affected offspring of an autosomal recessive deleterious mutation when the ancestor transmitting the mutation is known. Using this method, the frequency of Australian Holstein-Friesian cows carrying the deleterious mutation for complex vertebral malformation was estimated to be 2.47 and 4.44% for individuals born in 1992 and 1997 cohorts, respectively (Man et al., 2007). This pedigree approach is suitable for newly recognized autosomal recessive mutations, for which no extant of genotype test exists. The large number of genotyped cattle in North America has resulted in the discovery of several recessive haplotypes and a few causative mutations (VanRaden et al., 2011; Fritz et al., 2013; McClure et al., 2013), most of them affecting fertility due to increased embryonic mortality. This information has permitted the determination of carrier status of genotype animals using haplotypes in place of laboratory tests (Cole et al., 2013). With genomic data, a missense mutation SNP in the PNPLA8 gene present on *Bos taurus* autosome 4 was implicated as the causal mutation responsible for Weaver syndrome in Brown Swiss cattle (Kunz et al., 2016). In addition, the depression in fertility observed in Jersey cattle has been attributed to a nonsense mutation found in the *CWC15* gene (Sonstegard et al., 2013). Those authors found no individuals carrying the *CWC15* mutant allele from 749 genotyped animals, indicating that the presence of homozygous *CWC15* mutant allele resulted in spontaneous abortion. Causative mutations for calf diseases and for the postnatal mortality of calves are less frequently reported (Charlier et al., 2012; Jung et al., 2014; Kipp et al., 2015; Pausch et al., 2015), although such mutations may also play a large role in animal health. Any recessive haplotype that results in the death of calves following their birth and early rearing period is particularly problematic because the economic impact is much greater than that of haplotypes causing early embryonic loss (Cole et al., 2016). The avoidance of mating individuals carrying deleterious recessive haplotypes could potentially save the dairy industry approximately $11,000,000 per annum, which is a culmination of losses from Holsteins, Jerseys, Brown Swiss, and Ayrshires (Cole et al., 2016).

**Health Traits**

Due to the relative novelty of including health traits in routine breeding evaluations, most literature describing the effect of inbreeding on health traits has focused on SCS. In contrast to investigations on the effect of homozygosity on the traits previously discussed, the effect of inbreeding on SCS is not conclusive. Miglior et al. (1995a) found relatively low levels of inbreeding depression for SCS, but they observed a higher average lactation SCS than non-inbred animals. Thompson et al. (2000b) saw no effect of inbreeding level on SCS, although they did find that survival decreased as the level of inbreeding increased and concluded that inbreeding was likely to have a greater negative effect on the financial health of the dairy farm than directly.
in production losses. Sewalem et al. (2006) noted a trend toward increased risk of culling among inbred animals in Holsteins, Jerseys, and Ayshires, but this effect was more pronounced above inbreeding levels of 12.5%. Sørensen et al. (2006) reported higher levels of clinical mastitis in inbred cows and Mc Parland et al. (2007) reported that inbreeding increased SCS. Kim et al. (2015) concluded that most ROH were not significantly associated with SCS. An interesting area of future research could therefore be calculating the effect of genomic inbreeding on novel health traits that were previously not collectable on a large scale.

Overall, several studies have concluded that an increase in homozygosity in a population can have clearly negative effects on phenotypes, likely owing to both the expression of deleterious homozygous recessive alleles and loss of heterozygosity advantage. Using novel technologies and data, it is becoming possible to localize and partition the negative effects of inbreeding to specific genomic regions to better understand how homozygosity affects phenotypic traits of interest (McLaren et al., 2016; Howard et al., 2017b). In the case of fertility traits, improved quality of genomic information, particularly on the sex chromosomes, will likely contribute greatly to future progress.

**A Case Study: Canada**

**Introduction.** From the literature reviewed in the previous section, it seems clear that inbreeding depression has an effect on many aspects of production, reproduction and fertility, and health of dairy cattle. Furthermore, we have highlighted how inbreeding depression accumulation is heterogeneous across the genome. The challenge remains in the ability to identify specific ROH genotypes associated with reduced performance. Currently no consensus method exists to perform such analyses, given that the definition of ROH and the identification of haplotypes within ROH regions are dependent on the cutoff length chosen and the software used for the analysis.

In the following section, we identify ROH genotypes associated with extreme phenotypes in Canadian dairy cattle and test them for significance using software developed by Howard et al. (2017b). The source code compiled executable files for Linux operating systems and a manual are available at https://github.com/jeremyhoward. We do not aim to discuss these results in detail here, but rather to generally show what types of information on localization and characterization of ROH with an effect on phenotype are available.

**Pedigree and Genotypes.** Pedigree information on 77,951 animals were available for analysis from the Canadian Dairy Network. A total of 14,741 cows were genotyped with a 50K SNP chip panel, and missing genotypes were imputed with FImpute v2.2 (Sargolzaei et al., 2014). After filtering, 43,342 autosomal SNP were analyzed.

**Phenotypes.** Phenotypes on 5 production traits (305-d milk, fat and protein yields, as well as fat and protein percentages), 14 fertility traits (age of the heifer and cow at first service, interval between calving and first insemination for the heifer and cow, 56-d nonreturn rate for heifers and milking cows, the interval between first insemination and conception of heifers and lactating cows, age at calving, calf size, days open, gestation length, number of services, and stillbirth), as well as 2 health-related traits (BHB and SCS) were included in the study.

**Methods.** The software of Howard et al. (2017b) was used to first scan genotypes using a sliding marker window approach of decreasing size and to identify ROH associated with an unfavorable phenotype. Second, the regions identified in the initial scan were tested for their significance using the following linear mixed model (as implemented in Howard et al., 2017b):

\[ y = Xb + Za + Wpe + e, \]

where \( b \) is a vector of fixed effects (environmental classifications effects and the effect of the ROH genotype in a given window), \( a \) is a vector of random additive genetic effects, \( pe \) is a vector of random permanent environmental effects, \( e \) is a vector of random residual effects, \( X, Z, \) and \( W \) are incidence matrices relating \( b, a, \) and \( pe \) respectively, to \( y \). The random additive genetic effect is assumed \( \sim N(0, A\sigma^2_a) \), with \( A \) representing the pedigree additive relationship matrix. The random permanent environmental and residual effects were assumed \( \sim N(0, I\sigma^2_{pe}) \) and \( \sim N(0, I\sigma^2_e) \), respectively, with \( I \) being an identity matrix, and \( \sigma^2_a, \sigma^2_{pe}, \) and \( \sigma^2_e \) are the additive genetic, permanent environmental, and residual variances, respectively. Briefly, variance components were assumed fixed across windows based on the null model of no ROH effect. Given the solutions for each window (obtained using the Cholesky decomposition of the left hand side), contrasts between each ROH genotype versus non-ROH genotypes and associated \( t \)-statistics were obtained. The direction of each one-sided \( t \)-test was dependent on the direction of the unfavorable phenotype; therefore, genotypes not found within an ROH were assumed normal compared with ROH genotypes, aligning with the partial dominance hypothesis. In this manner, ROH identified to have a significant effect on a trait were identified. A summary of significant ROH for each trait and the least squares
Table 1. Significant runs of homozygosity (ROH) identified with the most extreme unfavorable effects in cows and heifers (C and H, respectively)

<table>
<thead>
<tr>
<th>Trait</th>
<th>BTA</th>
<th>Start position</th>
<th>End position</th>
<th>Phenotypic mean LSM</th>
<th>SD LSM</th>
<th>SD LSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at calving (H; mo)</td>
<td>19</td>
<td>14,121,650</td>
<td>14,763,072</td>
<td>0.881</td>
<td>0.896</td>
<td>0.144</td>
</tr>
<tr>
<td>Age at first service (H; mo)</td>
<td>23</td>
<td>21,557,504</td>
<td>25,062,520</td>
<td>13.336</td>
<td>13.385</td>
<td>2.047</td>
</tr>
<tr>
<td>BHB (mmol/L)</td>
<td>19</td>
<td>49,398,678</td>
<td>51,317,789</td>
<td>0.365</td>
<td>0.347</td>
<td>0.048</td>
</tr>
<tr>
<td>Calving to first service (C; d)</td>
<td>24</td>
<td>35,022,765</td>
<td>35,982,394</td>
<td>9.290</td>
<td>9.347</td>
<td>1.350</td>
</tr>
<tr>
<td>Calving to first service (H; d)</td>
<td>1</td>
<td>46,370,195</td>
<td>47,296,021</td>
<td>2.203</td>
<td>2.208</td>
<td>0.926</td>
</tr>
<tr>
<td>Calf size (C)</td>
<td>24</td>
<td>13,837,820</td>
<td>15,077,549</td>
<td>0.186</td>
<td>0.180</td>
<td>0.027</td>
</tr>
<tr>
<td>Calf size (H)</td>
<td>24</td>
<td>10,418,163</td>
<td>4,922,801</td>
<td>27.226</td>
<td>27.226</td>
<td>3.066</td>
</tr>
<tr>
<td>Days open (C)</td>
<td>11</td>
<td>59,381,768</td>
<td>60,572,203</td>
<td>−25.954</td>
<td>−32.252</td>
<td>3.539</td>
</tr>
<tr>
<td>Fat (305-d; kg)</td>
<td>4</td>
<td>21,401,937</td>
<td>25,358,659</td>
<td>−0.156</td>
<td>−0.131</td>
<td>0.023</td>
</tr>
<tr>
<td>Fat percentage</td>
<td>14</td>
<td>38,583,008</td>
<td>39,640,732</td>
<td>16.536</td>
<td>16.537</td>
<td>2.479</td>
</tr>
<tr>
<td>Gestation length (C; d)</td>
<td>23</td>
<td>26,727,885</td>
<td>27,545,231</td>
<td>2.101</td>
<td>2.101</td>
<td>0.327</td>
</tr>
<tr>
<td>Gestation length (H; d)</td>
<td>29</td>
<td>10,829,548</td>
<td>11,619,866</td>
<td>2.431</td>
<td>2.107</td>
<td>0.327</td>
</tr>
<tr>
<td>Milk (305-d; kg)</td>
<td>11</td>
<td>60,377,158</td>
<td>62,732,451</td>
<td>−896.693</td>
<td>−917.691</td>
<td>93.186</td>
</tr>
<tr>
<td>Nonreturn rate (C)</td>
<td>28</td>
<td>37,962,946</td>
<td>39,134,584</td>
<td>−0.179</td>
<td>−0.179</td>
<td>0.026</td>
</tr>
<tr>
<td>Nonreturn rate (H)</td>
<td>16</td>
<td>72,878,056</td>
<td>73,425,409</td>
<td>−0.205</td>
<td>−0.205</td>
<td>0.026</td>
</tr>
<tr>
<td>Number of services (C)</td>
<td>11</td>
<td>48,881,609</td>
<td>50,065,266</td>
<td>0.565</td>
<td>0.565</td>
<td>0.081</td>
</tr>
<tr>
<td>Number of services (H)</td>
<td>16</td>
<td>4,329,965</td>
<td>5,253,657</td>
<td>0.457</td>
<td>0.457</td>
<td>0.026</td>
</tr>
<tr>
<td>Protein (305-d; kg)</td>
<td>11</td>
<td>58,524,865</td>
<td>60,578,856</td>
<td>−23.635</td>
<td>−28.525</td>
<td>2.925</td>
</tr>
<tr>
<td>Protein percentage</td>
<td>5</td>
<td>92,074,297</td>
<td>93,920,010</td>
<td>−0.110</td>
<td>−0.071</td>
<td>0.010</td>
</tr>
<tr>
<td>SCS</td>
<td>5</td>
<td>87,104,035</td>
<td>88,105,584</td>
<td>0.696</td>
<td>0.661</td>
<td>0.082</td>
</tr>
</tbody>
</table>

1Chromosome that harbors the ROH with an unfavorable effect.

Means difference between animals with a significant ROH versus those animals without ROH in the same window is presented in Table 1.

**Results and Discussion.** In this small case study using Canadian data, ROH with significant detrimental effects on various traits were identified for all traits. Table 1 shows the difference in phenotypic mean of animals with ROH compared with those without the extreme ROH. Traits were not standardized and were given in phenotypic units of measurement (e.g., animals with the extreme ROH haplotype on chromosome 23 affecting age at first service in heifers are likely to be inseminated an average of 13 d later than those without the ROH). The observed effect of the ROH on the phenotype was, in some cases, quite strong. These associations should be investigated further to better understand the underlying biological mechanisms causing the difference in phenotype.

The average length of extreme ROH haplotypes was 1.17 Mb, with a maximum length of 2.35 Mb and a minimum length of 0.55 Mb. Within a region, the smallest significant ROH haplotype was used to tag animals with variable ROH lengths due to recombination occurring at different locations; as a result, the overall ROH region may be much larger across animals. Using more dense chip information would provide more precise data and may result in longer ROH. This information can then be further analyzed to more precisely localize significant regions of homozygosity, to investigate the biological mechanisms of specific genes in these regions, and to better understand the implications of homozygosity throughout the genome.

As discussed in depth by Howard et al. (2017b), the unfavorable effect of an ROH was likely due to a single unique ROH genotype within a given region, with some ROH having an unfavorable effect or a favorable effect. Thus, identifying unique ROH genotypes associated with an unfavorable phenotype may provide researchers with a starting point for uncovering the biological implications of homozygosity at specific loci. The primary outcome of the software algorithm was a list of unfavorable haplotypes, which can be incorporated into one of the many algorithms that already exist to manage unfavorable mutations or haplotypes within breeding programs (e.g., Kinghorn, 2011; Cole, 2015). A further use for this type of information is to better understand how and why ROH may affect traits of interest.

**MANAGING INBREEDING USING GENOMIC INFORMATION**

Currently, several strategies to control inbreeding accumulation are available, each with varying levels of complexity. The simplest strategy is simply not mating genetically related animals (Weigel and Lin, 2000); however, more complex methods of managing inbreeding at the population level, such as restricting the rate of inbreeding at either the animal or the population level (Fernández et al., 2000), have been proposed. Alternatively, optimal contribution selection, which maximizes genetic gain under constrained levels of accepted inbreeding, can be implemented (Wray and Goddard, 1994; Meuwissen, 1997; Grundy et al., 1998). Overall, the majority of these strategies rely on either pedigree...
information or genomic relationships averaged over the genome (for a detailed review of inbreeding management strategies, please refer to Howard et al., 2017a).

The use of region-specific metrics that characterize the genetic diversity (i.e., effective population size) and mapping inbreeding depression at specific regions could be used alongside commonly estimated genome-wide metrics (Kardos et al., 2016; Curik et al., 2017). These regions can be identified and mapped when only affected individuals have the homozygous haplotypes carrying the causal mutation and also when only heterozygous haplotypes are identified, suggesting the lethality of the homozygous haplotypes (VanRaden et al., 2011; Kardos et al., 2016). In this manner, matings between individuals with similar co-ancestry values at the genome-wide level might be further discriminated on the basis of haplotypic diversity, which is expected to have an effect on the genomic load carried by the resulting offspring (Howard et al., 2017a). The use of region-specific metrics should allow breeders to more effectively manage the risks associated with mate selection, allowing better evaluation of trade-offs between the genetic value of the progeny and the undesirable side effects associated with inbreeding. Lastly, as time progresses, a larger number of novel 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 the future introduction of gene-editing procedures, among others, will affect the management of genomic diversity, both at the population as well as the farm level (Howard et al., 2017a). Research on how to manage genetic resources based on all these factors simultaneously, rather than individually, should become a priority.

CONCLUSIONS

In this symposium review, we discussed the current state and future challenges related to inbreeding, and how homozygosity affects health and productivity phenotypes in dairy cattle using a Canadian case study, as well managing inbreeding using genomic information at the individual herd level. Future areas of research could include improved partitioning of genomic areas that are detrimental when homozygous and those that are beneficial when homozygous, better characterization, genotyping, and understanding of sex chromosomes and their association with fertility traits, as well as more in-depth analysis of the effect of genomic inbreeding on novel health traits. Overall, characterization and management of inbreeding at both the herd and population levels could be improved using genomic information.

ACKNOWLEDGMENTS

American authors (J. T. Howard and C. Maltecca) gratefully acknowledge funding from the American Holstein Association (Brattleboro, VT). Canadian authors (C. Baes, B. Makanjuola, A. Fleming, F. Miglior, and G. Marras) gratefully acknowledge funding by the DairyGen Council of the Canadian Dairy Network (Guelph, Ontario, Canada) and the Natural Sciences and Engineering Research Council of Canada (Ottawa, Ontario, Canada). The authors thank two anonymous reviewers who contributed to the quality of our manuscript.

REFERENCES


