Genetic and genomic relationships among Canadian Holstein dairy cattle population and international Holstein bulls

B.O. Makanjuola[1], C. Obari[1], G. Condello[1], F. Miglior[1,2], C. Maltecca[1,3], J.B. Cole[1,3,4,5], F.S. Schenkel[1] and C.F. Baes[1]

[1]Centre for Genomic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, Ontario, Canada
[2]Lactanet, Guelph, Ontario, Canada
[3]North Carolina State University, Department of Animal Science and Genetics Program, Raleigh, North Carolina, USA
[4]Council on Dairy Cattle Breeding, Bowie, Maryland, USA
[5]Department of Animal Sciences, Donald Henry Barron Reproductive and Perinatal Biology Research Program, and the Genetics Institute, University of Florida, Gainesville, USA
Corresponding author: bmakanju@uoguelph.ca

Abstract

The increasing level of inbreeding in dairy cattle populations can be a concern for researchers, producers, and artificial insemination (AI) companies. High inbreeding levels can lead to the accumulation of deleterious recessive variants, depression of the mean value of economically important traits, and a reduction in available genetic diversity in the population. Advancements in reproductive technologies and the integration of genomic information into genetic evaluations have contributed to rising inbreeding levels. Given the global interest in inbreeding, there is a need to monitor inbreeding trends and develop strategies to manage its adverse effects while ensuring continued genetic progress. The objective of this study was to measure the current trends in genetic and genomic relationships among Canadian Holstein cows and international Holstein bulls. Pedigree and genotype data were provided by Lactanet Canada (Guelph, ON). Genotype information was available for 168,995 animals, for which a pedigree of 616,258 animals was extracted. Among genotyped animals, 8,491 bulls were born between 2000 and 2023, and 131,139 cows were born between 2010 and 2024. The average pedigree completeness index of all genotyped animals was greater than 99%, with a maximum pedigree depth of 30 generations. Genetic relationship values were estimated using pedigree data (R-value) and genomic data (GR-value). R-value and GR-value represent the expected and realized percentage of shared DNA between an animal and a defined reference population, respectively. This was done by iteratively tracing back the gene contribution of an animal to the reference population. In this study, the reference population was defined as currently active cows and heifers enrolled in milk recording without a documented left-herd date on the latest test day in April 2024. Results show a yearly increase in relationship within the reference and bull populations. The average genetic relationship between bulls and the reference population ranged from 9.3% to 26.5% (R-value) and from 12.9% to 40.8% (GR-value). Among bulls, those with United States registration codes had the highest relationship value with active Canadian cows and heifers, with R-value and GR-value estimates of 20.8% and 30.4%, respectively. Conversely, bulls registered in the Czech Republic had the lowest average relationship values, with R-value and GR-value estimates of 17.1% and 24.3%, respectively. Selecting sires with low average relationship values among defined reference populations as a mating strategy could reduce or maintain inbreeding at acceptable levels while preserving genetic diversity.

Key words: Genetic relationship, inbreeding diversity, Holstein, dairy cattle

Introduction

The availability of genotype information has driven rapid advancements in breeding programs across most intensive dairy producing countries, particularly following the implementation of genomic selection (García-Ruiz et al., 2016; Miglior et al., 2017). Moreover, the international exchange of genotypes among partner countries has been reported to produce higher gains in reliability estimates of breeding values for economically important traits (Schenkel et al., 2009). These exchanges contribute not only to the increase of genetic progress, but also to greater genetic relatedness of animals across and within countries. This increased genetic connectedness leads to rising levels of inbreeding within a given population.

Increased inbreeding can reduce genetic variation, which may limit the response to selection, lower the mean value of economically important traits, and promote the accumulation of deleterious recessive alleles. Ultimately, this can increase the frequency of genetic defects and result in substantial economic losses for producers. Therefore, monitoring relatedness and inbreeding levels within a population is important for sustainable genetic progress. One approach to monitoring genetic relatedness within a population is by estimating genetic relationship values between actively producing cows and currently available sires. Genetic relationship values measure the proportion of DNA an animal shares with a predefined reference population and can be estimated using either pedigree information (R-values) or genomic data (GR-values).

As pedigree and genotype information accumulates, genetic relationships among animals in a population can be more accurately estimated. These relationships depend on both the number of descendants an animal contributes to the active population and the number of ancestors that have many descendants in the active population. Consequently, genetic relationships are bound to change continuously as actively producing animals are culled or lost from the population. Increased availability and use of young genomic sires can reduce the heavy reliance on proven sires, which may subsequently lower the average relationship values for proven sires. An analogous way to capture this dynamic is through the expected future inbreeding (EFI), which is approximately half the average genetic relationship between a sire and a random sample of active cows. The objective of this study was to assess the current average genetic relationship trend between active foreign and domestic sires and currently active Canadian Holstein cows using both pedigree and genomic information.

Materials and Methods

Data

Pedigree and genotype data were provided by Lactanet Canada (Guelph, ON, Canada) following the April 2024 evaluation release. In addition, herdbook records for active milk-recorded cows and heifers were provided. The reference population used for analysis was defined as active cows with a recorded test date in April 2024, as well as heifers up to 30 months of age that were registered in the herdbook and had no recorded left-herd date.

In total, there were 616,258 animals in the pedigree, which included all known ancestors for the genotyped animals traced back 20 generations. The pedigree completeness index (PCI) for all animals in the pedigree was estimated going back five generations and only animals with a PCI greater than or equal to 90% were retained for further analyses.

Genotype data were available for a total of 146,698 animals linked to the pedigree. Of these, 8,504 were bulls with birth year between 2000 to 2023 and 138,194 were active cows or heifers that had not been culled from the herd with birth year between 2010 to 2024. All animals had genotypes on the 50K SNP panel (Illumina Inc., San Diego, CA, USA). Quality control was performed to retain autosomal SNP with a call rate greater than 95%, a minor allele frequency (MAF) greater than 5%, and a difference between expected and observed heterozygosity less than 0.15 (Wiggans et al., 2009). In addition, only genotyped animals with a five generation PCI greater than or equal to 90% were included in the final dataset, resulting in 8,491 bulls and 131,139 active cows and heifers retained for further analyses.

Statistical analyses

Genetic relationship values between each animal and the reference group were estimated in accordance with the method developed by Van Doormaal et al. (2003). This was then modified to incorporate genomic information for estimation of GR-values. The first step was iterative estimation of progeny gene contribution of each animal to the reference population for the sire and dam separately.

$$C_{si} = \sum_{i=1}^{m} 1/2C_i\,;\ \ C_{di} = \sum_{i=1}^{m} 1/2C_i \qquad (1)$$

where C_{si} and C_{di} are the i^{th} individual progeny gene contribution to their sire and dam, respectively, C_i is the contribution of the i^{th} individual with an initial starting value of 1 for all active animals and 0 for all other animals, and m is the number of animals in the pedigree.

$$R_i = \begin{cases} C_i & if\ S_i = 0\ and\ D_i = 0 \\ 0.75C_i + 0.5C_{si} & if\ S_i \neq 0\ and\ D_i = 0 \\ 0.75C_i + 0.5C_{di} & if\ S_i = 0\ and\ D_i \neq 0 \\ 0.75C_i + 0.5C_{si} + 0.5C_{di} & if\ S_i \neq 0\ and\ D_i = 0 \end{cases} \qquad (2)$$

where R_i is the total value of the i^{th} individual that reflect the genetic relationship to the reference group, and S_i and D_i are the sire and dam of the i^{th} individual, respectively, and when 0 it indicates unknown parent. The R-value is then defined as:

$$Rval_i = \frac{R_i}{n} \qquad (3)$$

where $Rval_i$ is the percentage genetic relationship value of the i^{th} individual and n is the total number of animals in the reference group.

For the genomic relationship values, equation (1) was modified to use the realized marker-by-marker similarity between progeny and parent, instead of assuming the expected value of 0.5 used for pedigree. In this case, actual SNP genotypes were used to estimate the proportion of identical-by-state (IBS) alleles shared between parent and offspring.

Results & Discussion

The distribution of genotyped active cows and heifers is presented in Table 1. Notably, 45.7% of the reference population represents heifers that have not calved. Including heifers in the reference group is essential to represent the expected future breeding population of active females. Furthermore, approximately 7% of cows were seven years of age or older, suggesting that older animals tend to leave or are culled from the herd. This dynamic turnover contributes to ongoing changes in the genetic relationship values estimated within the population.

Among genotyped bulls, those under the age of four accounted for 30.7% of the total genotyped bulls (Figure 1). This attests to the impact of genomic selection, which allows for early selection of bulls rather than the expected five years for progeny proven time (Schaeffer, 2006). This could contribute to diversification of the pool of available bulls and influence the estimated genetic relationship values within the population.

Table 1: Age distribution of the reference group as of April 2024.

Age (year)	Number of active cows and heifers	Percentage of total
<1	3,394	2.6%
1	28,073	21.4%
2	28,499	21.7%
3	23,775	18.1%
4	18,264	13.9%
5	12,620	9.6%
6	7,947	6.1%
7	4,155	3.2%
8	2,169	1.7%
9	1,214	0.9%
10	527	0.4%
>10	502	0.4%

The average annual genetic relationship per year (Figure 2) trend showed a steady increase over time, mirroring the pattern observed in the estimated increase in pedigree inbreeding reported by Van Doormaal (2024). Although, genomic relationship values were approximately 1.5 times higher than the pedigree relationship values, both increased in parallel from 1990 to 2024. Specifically, R-values increased from 14.6% in 2010 to 21.4% in 2024, while GR-values rose from 20.7% to 31.3% over the same period (Table 2).

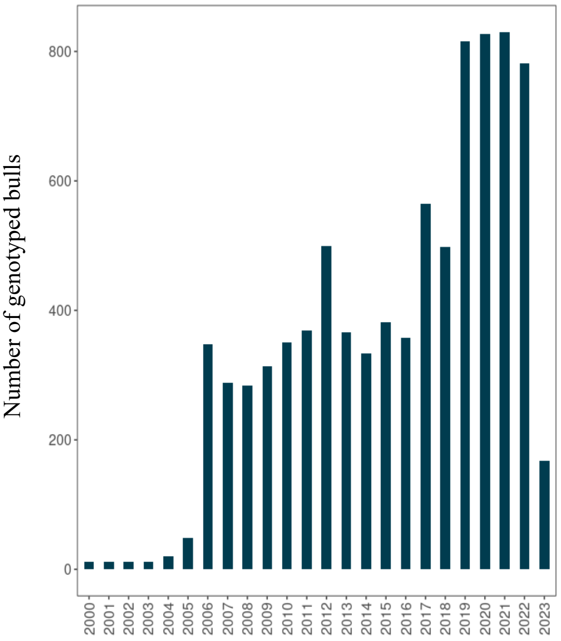

Figure 1. Distribution of the number of genotyped sires across birth year from 2000 to 2023.

Increase in both genomic and pedigree relationships suggest a concurrent increase in inbreeding levels within the population over time. Based on the estimated genetic relationships, the expected pedigree future inbreeding (EFI) and genomic future inbreeding (GFI) was estimated as half the R-values and GR-values, respectively. The EFI increased from 7.0% in 2010 to 10.6% in 2024, while GFI increased from 9.9% to 15.5%. These estimates aligned with the estimated pedigree inbreeding coefficients for the corresponding period and ranged from 6.6% to 11.1%.

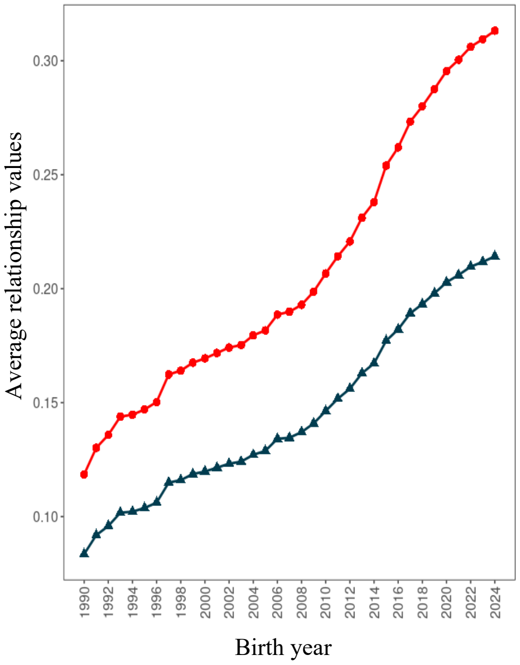

Figure 2. Annual average genetic relationship values per year based on pedigree and genomic information from 1990 to 2024. Blue line with triangles and red line with circles are pedigree and genomic relationship values, respectively.

Additionally, the observed correlations for pedigree inbreeding coefficients (Fped) with R-values and GR-values were moderately high at 0.77 and 0.76, respectively (Table 3). These findings suggest that genetic relationship measures can serve as a sufficient proxy of inbreeding levels and can be effectively used to inform mating strategies and selection decisions with the consideration of reducing inbreeding within the population.

Table 2: Average pedigree and genomic relationship values, pedigree and genomic future inbreeding estimated from relationship values and inbreeding coefficients across birth year from 2010 to 2024.

Birth Year	R-value (%)	GR-value (%)	EFI (%)	GFI (%)	Fped (%)
2010	14.6	20.7	7.0	9.9	6.6
2011	15.2	21.4	7.3	10.3	6.8
2012	15.6	22.1	7.6	10.7	7.1
2013	16.3	23.1	7.8	11.0	7.4
2014	16.7	23.8	8.1	11.6	7.5
2015	17.7	25.4	8.4	11.9	7.8
2016	18.2	26.2	8.9	12.7	8.1
2017	18.9	27.3	9.1	13.1	8.6
2018	19.3	28.0	9.5	13.7	8.8
2019	19.8	28.8	9.7	14.0	9.3
2020	20.3	29.6	9.9	14.4	9.7
2021	20.6	30.0	10.1	14.8	10.1
2022	21.0	30.6	10.3	15.0	10.4
2023	21.2	30.9	10.5	15.3	10.8
2024	21.4	31.3	10.6	15.5	11.1

This analysis identified the degree of relationship between each countries' bulls and the active Canadian cows and heifers. Results showed that bulls registered to the United States, Belgium, and the Netherlands had the highest average genetic relationship with the Canadian reference population, with R-values of 20.8%, 20.7%, and 20.2% and corresponding GR-values of 30.4%, 30.1%, and 29.3%, respectively (Table 4).

Table 3: Correlation coefficients between pedigree and genomic relationship values and pedigree inbreeding.

	R-value	GR-value	Fped
R-value	1		
GR-value	0.99	1	
Fped	0.77	0.76	1

These findings indicate frequent exchange of genetic materials between these countries and Canada. In contrast, two decades ago, bulls from Spain (11.4%), Japan (9.4%), and Italy (9.4%) had the highest R-values with active Canadian cows in 2004, which may indicate a shift in bull selection (Van Doormaal et al., 2005). Moreover, in 2004, the average Canadian bull R-values with active Canadian cows was estimated to be 11.7% and currently in 2024 it has increased to 19.4%, indicating an increase in genetic relatedness and, by extension, inbreeding. Additionally, a wider range of R-values (4.2% to 11.4.%) was observed between international bulls and active Canadian cows in 2004, which had substantially narrowed by 2024 (17.1% to 20.8%). This pattern indicates that the Holstein breed across countries has become increasingly similar over time, which is likely due to the frequent international exchange of genomic material and the widespread adoption of genomic selection.

Table 4: Properties of countries with genotype bulls that are connected with the active Canadian cows and heifers.

Country	Number of Bulls	Number of Daughters	R-value (%)	GR-value (%)	EFI (%)	GFI (%)
CAN	3,022	50,107	19.4	28.1	9.7	14.0
AUS	14	902	19.1	27.3	9.6	13.7
BEL	10	8	20.7	30.1	10.3	15.1
CHE	50	361	20.0	29.1	10.0	14.6
CZE	10	0	17.1	24.3	8.5	12.2
DEU	121	461	19.6	28.4	9.8	14.2
DNK	21	57	17.5	25.3	8.8	12.6
ESP	5	50	19.4	28.0	9.7	14.0
FRA	82	108	19.1	27.7	9.6	13.9
GBR	49	415	19.4	27.9	9.7	13.9
HUN	19	3	17.4	24.9	8.7	12.4
ITA	111	571	19.0	27.5	9.5	13.7
NLD	225	3,944	20.2	29.3	10.1	14.7
USA	4,745	57,120	20.8	30.4	10.4	15.2

The average genetic relationship values varied across bulls, ranging from approximately 10% to 25% for R-values and 14% to 36% for GR-values (Figure 3). This variation highlights the opportunity for strategic selection of bulls that are less related to the reference population to manage and minimize inbreeding levels. Additionally, AI companies can support this effort by diversifying bull selection pools and offering sires that are less genetically related to specific herds or producers. Figure 4 presents the genetic relationship values of the top 100 lifetime performance index (LPI) bulls, with R-values ranging from 19% to 25% and GR-values from 28% to 36%, reflecting a considerable genetic contribution to the population. This is expected because bulls with high LPI tend to have more daughters and shared ancestry, which ultimately increases their genetic relationship with the active cow and heifer population.

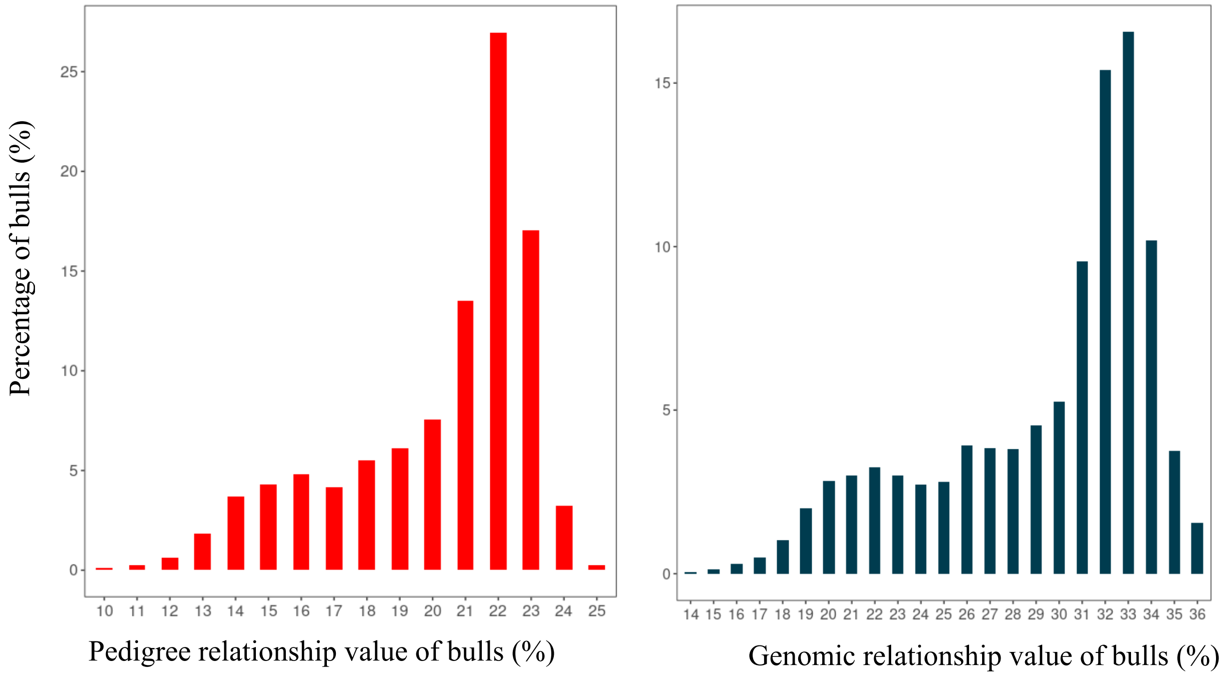

Figure 3. Distribution of relationship values of genotyped bulls with the active cows and heifers in the Canadian Holstein population using pedigree and genomic information.

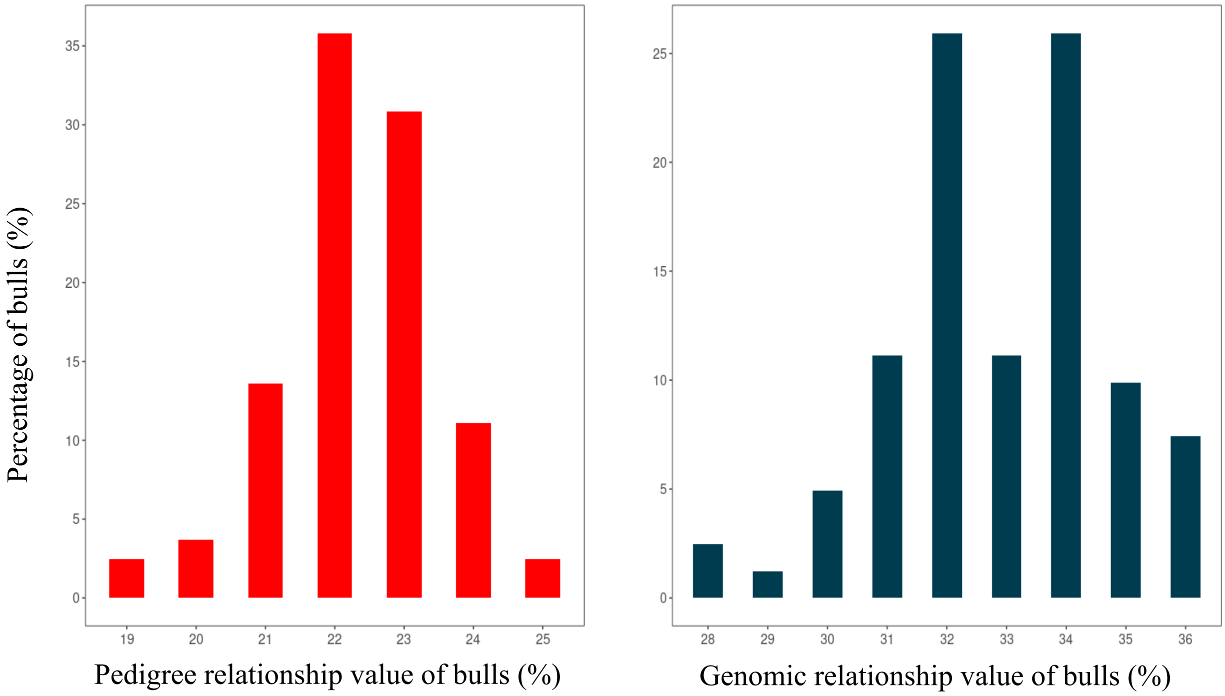

Figure 4. Distribution of relationship values for the top 100 LPI bulls in April 2024 with the active cows and heifers in the Canadian Holstein population using pedigree and genomic information.

Conclusions

Genetic relationships within Canadian Holsteins have steadily increased over time, corresponding with the observed rise in inbreeding levels. Additionally, genetic relationships between international bulls and active Canadian cows and heifers have increased, with differences between countries narrowed. The United States remains the major contributor of bulls to the active Canadian population. Among available bulls, variation exists in the genetic relationships with the active cows and heifers. This variation affords the opportunity to select less related bulls that could help manage and minimize inbreeding at the population level without trading-off the desired genetic gain. Finally, preventing the continued rise in inbreeding will require collaborative efforts from academia, AI companies, and producers.

Acknowledgments

We gratefully acknowledge all funding and support for the Resilient Dairy Genome Project (www.resilientdairy.ca/funders-and-partners/), as administered by Genome Canada, Genome Alberta, Ontario Genomics, Genome Quebec, and Genome British Columbia, as well as from the Net Zero Dairy Genome Project (//www.ndgp.ca/), as administered by Genome Canada, Ontario Genomics, Genome Alberta and Genome Quebec. Christine Baes also gratefully acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC, Ottawa, ON, Canada) Canada Research Chair and Discovery programs. As per the research agreement, researchers maintain independence in conducting their studies, own their data, and report the outcomes regardless of the results. The decision to publish the findings rests solely with the researchers.

References

García-Ruiz, A., Cole, J.B., VanRaden, P.M., Wiggans, G.R., Ruiz-López, F.J. and Van Tassell, C.P. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci.* doi:10.1073/pnas.1519061113.

Miglior, F., Fleming, A., Malchiodi, F., Brito, L.F., Martin, P. and Baes, C.F. 2017. A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. *J. Dairy Sci.*100:10251-10271. doi:10.3168/jds.2017-12968.

Schaeffer, L.R. 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218–223. doi:10.1111/j.1439-0388.2006.00595.x.

Schenkel, F.S., Sargolzaei, M., Kistemaker, G., Jansen, G.B., Sullivan, P., Van Doormaal, B.J., VanRaden, P.M. and Wiggans, G.R. 2009. Reliability of genomic evaluation of Holstein cattle in Canada. Proccedings Interbull Int. Work. - genomic Inf. Genet. Eval. Uppsala, Sweden 51–60.

Wiggans, G.R., Sonstegard, T.S., VanRaden, P.M., Matukumalli, L.K., Schnabel, R.D., Taylor, J.F., Schenkel, F.S. and Van Tassell C.P. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. *J. Dairy Sci.* 92:3431–3436. doi:10.3168/jds.2008-1758.

Van Doormaal, B.J., Kistemaker, G., Brand, P. and Beaumont, P. 2003. Development of "RValues" to represent an animal's genetic relationship to the active female population. Genetic Evaluation Board of Canadian Dairy Network.

Van Doormaal, B.J. 2024. Inbreeding Update - August 2024. Accessed June 8, 2025. https://lactanet.ca/en/inbreeding-update/.

Van Doormaal, B.J., Miglior, F., Kistemaker, G. and Brand, P. 2005. Genetic diversification of the Holstein breed in Canada and internationally. *Interbull Bull.* 93–97.