INTERBULL BULLETIN NO. 60. 20-21 May 2024, Bled, Slovenia

Next-Level Genomic Selection: Mitigating Inbreeding

***Jan-Thijs van Kaam*[1]*, Christos Dadousis*[2,3,4]*, Francesco Tiezzi*[5]*, Martino Cassandro*[1,6]**

[1]Associazione Nazionale Allevatori della razza Frisona, Bruna e Jersey Italiana (ANAFIBJ), 26100 Cremona, Italy

[2]Department of Veterinary Science, University of Parma, 43126 Parma, Italy

[3]School of Health Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom

[4]Veterinary Health Innovation Engine, School of Veterinary Medicine, University of Surrey, Guildford, United Kingdom

[5]Department of Agriculture, Food, Environment and Forestry, University of Florence, 50144 Florence, Italy

[6]Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, 35020 Legnaro, Italy

Corresponding author: jtkaam@anafi.it

Abstract

Analysis of 88,068 autosomal or X-linked SNPs in the ANAFIBJ Holstein genomic database including males and females showed declining SNP heterozygosity over time. In 1990, average SNP heterozygosity was 0.3620. During the pre-genomics period (1990-2010), the annual decline was -0.0003, reaching 0.3558 in 2010. However, in the genomics period (2010-2024), the average SNP heterozygosity declined to 0.3191 in 2024, with an annual decline of -0.0027, over 7 times higher than before. So far, this trend has been highly linear ($R^2=0.987$), which would extrapolate (in the absence of other sources of genetic variation than selection and inbreeding) to result in a complete loss of genetic variation in ~130 years. We developed measures to estimate genomic expected future inbreeding (Gefi) based on Runs-Of-Homozygosity (ROH). A comparison with CDCB genomic future inbreeding (GFI) based on the genomic relationship matrix (GRM) was done using 38,280 genotyped proven males covering 70 years (1951-2020), which resulted in a Pearson correlation of 0.959. The CDCB GRM **G** was computed as $\mathbf{G} = \mathbf{ZZ}'/\Sigma 2p(1-p)$ using p=0.5. Gefi estimates had a mean of 6.9% with a standard deviation of 2.6%. Minimum Gefi was 0.1% and maximum was 15.3%. GFI estimates had a mean of 7.2% with a standard deviation of 2.6%. Minimum and maximum GFI were -3.1% and 13.5%. The correlation was quite high even though Gefi is an identity-by-descent (IBD) measure in the probability space of [0,1], whereas GFI measures identity-by-state (IBS) in the correlation space of [-1,1]. Comparison of the inbreeding depression across traits showed that the depression is largest on yield traits followed by contents, somatic cell score (SCS) and fertility at around 20% of the depression on yield traits. ANAFIBJ aims to reduce the increase in future inbreeding by giving a premium to male and female animals which are less related to the recent population, while penalizing those that are more related.

Key words: Inbreeding, Holstein, effective population size, genetic diversity, SNP, Runs-Of-Homozygosity, genomic selection

Introduction

If a population has no exchange of genes with other populations, then mutations become the only way for new genetic variation to arise, and they arise slowly. Over time, breeding within such a closed population reduces overall genetic variation unless mutations happen frequently enough. This results in individuals becoming genetically more similar and gradually leads to inbreeding. Choosing just a few breeding parents who are genetically alike speeds up the inbreeding process. This can have serious consequences; animals might become less fertile and less healthy. Modern breeding techniques using genomic information (genomic selection) can worsen this decline, especially due to shorter generation intervals and stronger selection pressure. Unfortunately, low genetic variation makes populations less able to adapt to changing environments, increasing their extinction risk.

The Holstein reigns supreme in the dairy world. However, a hidden threat lurks beneath its success: a remarkably narrow genetic base. Despite the vast number of Holsteins globally, the breed's ancestry is surprisingly limited. Less than 10,000 Friesian animals were imported into North America over 130 years ago, and today, only two male lines effectively remain. This history of restricted gene flow, coupled with multiple genetic bottlenecks, has resulted in a population with low genetic diversity. While this focus has yielded high milk production, it could leave Holsteins vulnerable in the long run.

This paper and the presentation from the 2024 annual Interbull meeting in Bled are a follow-up of the paper and the presentation at the 2023 annual Interbull meeting in Lyon (Van Kaam et al., 2023). Estimates of genomic diversity were updated. Additional work was done to check genomic inbreeding coefficients from imputed data, to estimate the correlation between genomic future inbreeding measures, and to compare the size of inbreeding depression estimates per trait.

Materials and Methods

Issue of declining genetic variation

The ANAFIBJ genomic databank was used to analyze annual trends in genetic variation of Holstein SNP genotypes. After imputation, the annual average SNP heterozygosity of 88,068 autosomal or X-linked SNPs from male and female animals born between 1990 and 2024 was computed, excluding non-genotyped animals and animals without pedigrees. The year 2010 marked the transition from pre-genomic to genomic selection.

The analysis included both males and females. The average inbreeding coefficient per year was computed as (homozygosity this year - homozygosity first year) / (1 - homozygosity first year). Average generation intervals were calculated per year for the four pathways separately and then averaged. The relative year since 1990 was divided by the annual generation interval to estimate the number of generations that had passed since 1990. No Hardy-Weinberg equilibrium was assumed.

Genomic inbreeding coefficients from imputed data

A verification was done to check if genomic inbreeding coefficients from imputed data were reliable. This was done by comparing genomic inbreeding coefficients from imputed and genotyped SNPs. In the verification a set of high-density (139K/778K) genotypes were downgraded to either the GeneSeek Genomic Profiler 3 (26K) or to the Labogena MD (62K) SNP set. The downgraded set of 329 animals was also split into 2 subsets:

- Subset 1 of 266 animals with information on both parental sides i.e., S+D or S+MGS or S+D+MGS (S: Sire, D: Dam, MGS: Maternal grandsire)
- Subset 2 of 63 animals without information on one or both parental sides

Spearman rank, Pearson and concordance correlation coefficients were computed for the full downgraded set as well as the two subsets. Three genomic inbreeding coefficients were used. Fgrm0.5 was based on the GRM using p=0.5. Froh coefficients were based on 27 or 80 SNP segments, which correspond to 0.95 and 2.8 Mb, respectively. In the case of ROH, longer SNP segments indicate more recent inbreeding. The three types of correlation coefficients were computed for each of these three genomic inbreeding coefficients.

Correlation between Gefi and GFI

At Anafibj, a procedure was developed to estimate an Runs-Of-Homozygosity (ROH) style genomic inbreeding coefficient using an output file of our imputation software pedimpute.f90 (Nicolazzi et al., 2013), where haplotype segments are numbered. Using numbered haplotype segments, we could directly compare the segments by their number rather than using time-consuming SNP-by-SNP comparisons. The fraction of autosomal haplotypes with the same number (i.e., identical) on a pair of homologous chromosomes is an estimate of the Froh. A Froh is an identity-by-descent (IBD) type of inbreeding coefficient. We computed the own as well as the future inbreeding coefficients. Our genomic expected future inbreeding coefficient was named Gefi. An advantage of an IBD type of inbreeding measure is that values are within the probability space [0,1], which avoids negative values that are more difficult to understand.

A comparison was undertaken between the Gefi estimated by Anafibj and the genomic future inbreeding coefficients (GFI) estimated by CDCB using the diagonal from the genomic relationship matrix, which is an identity-by-state (IBS) measure with values within the correlation space [-1,1]. This Pearson correlation was computed based on 38,280 genotyped proven bulls covering 70 years (born 1951-2020), which had both estimates available.

Inbreeding depression standardized effect size across traits

In the literature, a number of articles showed results of genomic inbreeding depression using ROH-based inbreeding coefficients (Froh) in Holsteins. In order to compare traits and to understand which traits had a larger inbreeding depression, we standardized the inbreeding depression effects so that they became comparable across traits. We computed the standardized effect size as the inbreeding depression estimate (b) multiplied by the standard deviation of the genomic inbreeding coefficient (SD(F)) used to estimate the inbreeding depression. We then divided this by the observed standard deviation of the trait (SD(y)). Estimates from the following papers were included: Ablondi et al., 2023; Bjelland et al., 2013; Doekes et al., 2019; Makanjuola et al., 2020; Mugambe et al., 2023. Also, new unpublished results from Ablondi et al. were included.

Results & Discussion

Issue of declining genetic variation

Figure 1 shows the annual trend in SNP heterozygosity before and during the genomics era.

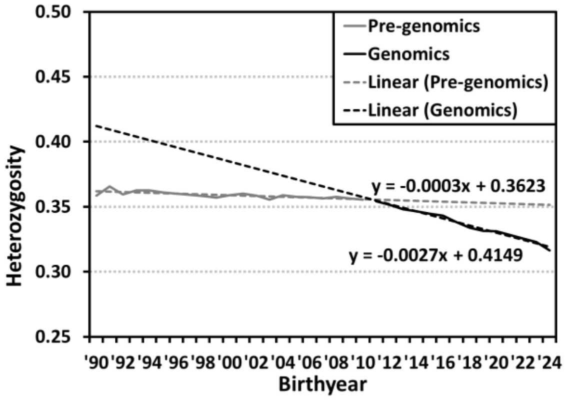

Figure 1. Pre- and post-genomic trends of SNP heterozygosity by birth year

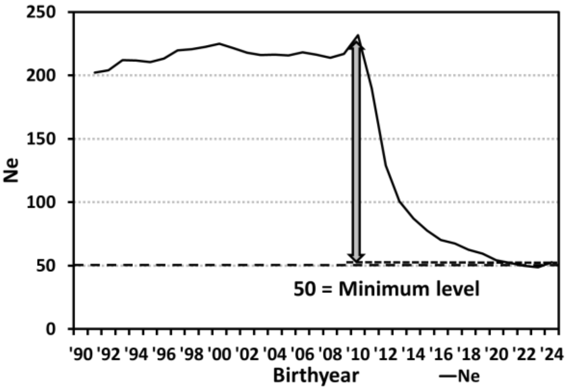

Figure 2. Decline of effective population size (Ne) during the most recent 16 birth years

Genomic inbreeding coefficients from imputed data

Table 1 shows different types of correlation coefficients between the genomic inbreeding coefficients of the samples in the imputed downgraded set with their fully genotyped (i.e. not downgraded) samples.

In the same manner Table 2 and 3 show the correlation coefficients for the subsets 1 and 2.

Table 1. Correlations of inbreeding coefficients entire downgraded set (N=329)

Downgrade	Correlation	GRM	ROH 27	ROH 80
GGP3 (26K)	Spearman	0.85	0.85	**0.90**
	Pearson	*0.43*	*0.48*	0.89
	Concordance	*0.26*	*0.39*	0.87
Labogena MD (62K)	Spearman	**0.93**	**0.96**	**0.96**
	Pearson	0.65	**0.97**	**0.98**
	Concordance	*0.55*	**0.97**	**0.96**

Table 2. Correlations of inbreeding coefficients downgraded set (N=266) with info on both parental sides

Downgrade	Correlation	GRM	ROH 27	ROH 80
GGP3 (26K)	Spearman	**0.93**	**0.96**	**0.96**
	Pearson	0.89	**0.96**	**0.96**
	Concordance	0.82	**0.96**	**0.94**
Labogena MD (62K)	Spearman	**0.98**	**0.99**	**0.98**
	Pearson	**0.97**	**0.99**	**0.98**
	Concordance	**0.95**	**0.99**	**0.97**

Table 3. Correlations of inbreeding coefficients downgraded set (N=63) without info on one or both parental sides

Downgrade	Correlation	GRM	ROH 27	ROH 80
GGP3 (26K)	Spearman	0.68	*0.54*	0.67
	Pearson	0.65	*0.34*	0.65
	Concordance	*0.14*	*0.10*	*0.55*
Labogena MD (62K)	Spearman	0.83	0.77	0.83
	Pearson	0.72	0.82	**0.90**
	Concordance	*0.27*	0.79	0.84

Correlation between Gefi and GFI

Summary statistics from Gefi and GFI are presented in Table 4 based on 38.280 proven bulls. The Pearson correlation between Gefi and GFI was 0.959.

Table 4: Summary statistics of Gefi and GFI

	Gefi	GFI
Average	6.9	7.2
Standard deviation	2.6	2.6
Maximum	15.3	13.5
Minimum	0.1	*-3.1*

Inbreeding depression standardized effect size across traits

Table 5 shows the average standardized effect size of the inbreeding depression per trait based on the estimates in literature. For all traits the effect was in the undesirable direction.

Table 5. Standardized effect size of the inbreeding depression

Trait	Standardized effect size
Milk kg	-.068%
Fat kg	-.063%
Protein kg	-.084%
Fat %	-.015%
Protein %	-.013%
SCS*	.020%
SCS5-150	.006%
SCS151-400	.017%
Age at 1st calving	.007%
Heifer interval 1st-last insemination	.022%
Heifer NR56	-.009%
Heifer conception rate	-.022%
Cow interval 1st-last insemination	.016%
Cow NR56	-.019%
Cow conception rate	-.028%
Interval calving to 1st insemination	.008%
Days open	.060%
Calving interval	.018%

*SCS: Somatic Cell Score

A further condensed overview of the standardized effect size of the inbreeding depression is given in Table 6. Here the average value per trait group is given.

Table 6. Average across trait standardized effect size of the inbreeding depression per trait group

Trait	Standardized effect size
Yields	-.072%
Contents	-.014%
SCS*	.014%
Fertility	-.019%

*SCS: Somatic Cell Score

Conclusions

Regarding genomic inbreeding coefficients from imputed data, we can conclude:

- The 62K chip outperforms the 26K chip.
- Froh exhibits stronger correlations than Fgrm0.5.
- Longer Froh segments demonstrate higher correlations than shorter segments.
- Spearman rank correlations > Pearson correlations > Concordance correlations
- Results are satisfactory when both parental sides have genotypes.
- In 2023, >97.3% of animals have genotypes on both parental sides, so for recent animals the results based on imputed data should be fine.

The high correlation between Gefi and GFI shows that they are both measuring future inbreeding producing very similar results.

For the inbreeding depression per trait group, we can conclude that by far the largest impact is on the yield traits. Fertility, contents and SCS all have an undesirable inbreeding depression of around 20% from the inbreeding depression on yields.

Overall take-home messages are:

- There has been a rapid increase in inbreeding since the advent of genomic selection.
- Inbreeding detrimentally affects nearly all traits, with the most pronounced impact seen in yield traits.
- Anafibj intends to introduce a premium/penalty for expected future inbreeding later this year.
- We will use genomic estimates when possible and otherwise pedigree-based estimates on a comparable scale.
- It is important to give a signal regarding the impact of inbreeding.

Acknowledgments

This study was supported by “Latteco2 project, sottomisura 10.2 of the PSRN-Biodiversity 2020–2023” (MIPAAF. D.M. no. 465907 del 24/09/2021, project unique code J12C21004080005).

References

Ablondi, M., Summer, A., Stocco, G., Finocchiaro, R., Van Kaam, J.-T., Cassandro, M., Dadousis, C., Sabbioni, A., and Cipolat-Gotet, C., 2023. The role of inbreeding depression on productive performance in the Italian Holstein breed. *J. Anim. Sci.* 101:1–10. https://doi.org/10.1093/jas/skad382.

Bjelland, D.W., Weigel, K.A., Vukasinovic, N., and Nkrumah, J.D., 2013. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *J. Dairy Sci.* 96:4697–4706. https://doi.org/10.3168/jds.2012-6435.

Doekes, H.P., Veerkamp, R.F., Bijma, P., De Jong, G., Hiemstra, S.J., and Windig, J.J., 2019. Inbreeding depression due to recent and ancient inbreeding in Dutch Holstein–Friesian dairy cattle. *Genet. Sel. Evol.* 51:54. https://doi.org/10.1186/s12711-019-0497-z.

Makanjuola, B.O., Maltecca, C., Miglior, F., Schenkel, F.S., and Baes, C.F., 2020. Effect of recent and ancient inbreeding on production and fertility traits in Canadian Holsteins. *BMC Genomics* 21:605. https://doi.org/10.1186/s12864-020-07031-w.

Mugambe, J., Ahmed, R., Thaller, G., and Schmidtmann C., 2024. Impact of inbreeding on production, fertility, and health traits in German Holstein dairy cattle utilizing various inbreeding estimators. *J. Dairy Sci.* TBC. https://doi.org/10.3168/jds.2023-23728.

Nicolazzi, E.L., Biffani, S., and Jansen, G., 2013. Short communication: Imputing genotypes using PedImpute fast algorithm combining pedigree and population information. *J. Dairy Sci.* 96:2649–2653. https://doi.org/10.3168/jds.2012-6062.

Van Kaam, J-T., Ablondi, M., Maltecca, C., and Cassandro, M., 2023. Inbreeding Becomes a Serious Issue. *Interbull Bulletin.* 59:101-104.