

Scaling of Genomic Inbreeding

By Ezequiel Nicolazzi, Gerald Jansen, Juan Pablo Nani and Paul VanRaden

CDCB routinely distributes pedigree-based and genomic-based inbreeding values. In April 2021, the scaling of genomic inbreeding was revised to account for X chromosome differences between males and females and to better match the scale of genomic- to pedigree-based inbreeding. Also, in April the computational efficiency of the whole procedure was greatly improved to keep up with the massive increase of incoming genotypes.

These changes did not affect the ranking for genomic inbreeding within either sex, as they only impacted the scaling and comparability across sexes. The revisions did not affect animals' PTAs (Predicted Transmitting Abilities), although there were slight adjustments to reliabilities in some cases.

Reasons for the change

Because males were predominantly genotyped in the early phases of genomics, the genomic inbreeding was scaled to match pedigree inbreeding using males. In recent years, genotyping of females increased drastically (see Figure 1). The scaling of genomic inbreeding values included females but did not account for the differences in the X chromosome between sexes.

In cattle (and humans), the inheritance of the X chromosome is different between sexes. Males have one copy of X and one of Y, whereas females have 2 copies of X. The tricky part is, in the files CDCB receives with marker data, males look like they had two identical copies of the same X chromosome (100% homozygous on X). This causes the homozygosity of females to artificially appear 3% less than males.

Key Takeaways

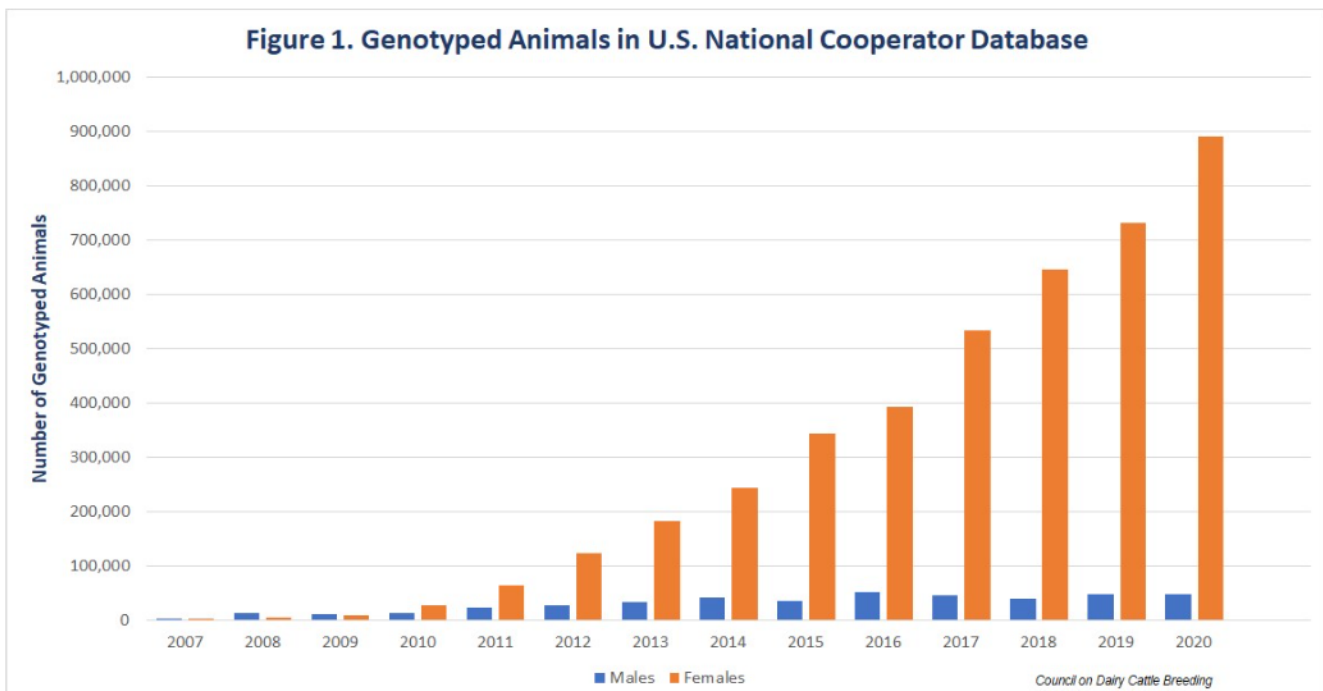
Changes implemented in April evaluations impacted the scaling and comparability across sexes; changes did not affect the ranking for genomic inbreeding within either sex.

Revisions did not affect animals' PTAs (Predicted Transmitting Abilities), except in unique cases. There were slight adjustments to reliabilities for some animals.

Genomic inbreeding values of males and females are now more comparable to each other and to pedigree inbreeding.

Computation has been improved to more rapidly process the larger numbers of genotypes expected in the future.

Figure 1. Genotyped Animals in U.S. National Cooperator Database



While many countries exclude the X chromosome, the U.S. genomic evaluations have always included the X to improve prediction accuracy. Beginning with [December 2018](#) evaluations, many more markers on the X were added. The first commercial chip, the Illumina 50K chip, had only 696 usable markers on the X chromosome (1.6% of the total); now, the 2,397 markers on the X are 3.0% of the 79,060 total markers used in the U.S. evaluations since 2018. These statistics do not count the pseudo-autosomal region that X and Y share which has only 0.2% of the DNA and 275 markers currently used.

The X contains almost 5% of the total DNA (is the 3rd longest chromosome!), meaning it is still underrepresented on commercial chips. Therefore, the contribution of X to inbreeding statistics and to genomic prediction may continue to increase as marker densities become more proportional.

This was the starting point for the research of Dr. Juan Nani and Dr. Paul VanRaden at USDA AGIL that led to this development ⁽¹⁾, based on recent research of Druet and Legarra⁽²⁾.

Impact of the adjustment

Adjusting for the X chromosome difference of males and females resulted in both sexes having more comparable values in the April evaluations. Although single cases could be affected by many factors, an example may be helpful. The following list shows the genomic inbreeding of the Jersey bull CREDENTIAL and his full-sibs (2 males and 1 female). Two observations can be made in the December 2020 results:

- 1) Male genomic inbreeding values are clearly higher than the female, and;
- 2) There is one of the full sibs that has clearly higher inbreeding than the rest, most likely depending on a more homozygous genotype.

The values in April are clearly more comparable between males and females, and the higher value for the homozygous full sib is still evident.

ID	SEX	GENOMIC INBREEDING	
		Dec. 2020	April 2021
JEUSA000117979857	M	8.4	4.4
JEUSA000117979866	M	12.2	8.2
JEUSA000117979875	M	8.2	4.8
JEUSA000117979884	F	5.7	4.3

A full statistical analysis, including all full sibs in the population, was performed at USDA AGIL to confirm that genomic inbreeding within full sibs now has the same average for male versus female siblings.

Genomic, pedigree inbreeding now on same scale

Another important goal of this research was to have genomic and pedigree inbreeding values on the same scale (e.g., more comparable, although still having a different range). The current results are much more aligned, recognizing that obtaining a perfect match of scales is extremely difficult. Previously, the pedigree inbreeding values used for scaling the genomic inbreeding were approximated with different edits, pedigree depth, and assumptions for missing ancestors. Effective this April, the official rather than approximate pedigree inbreeding is used in the genomic scaling.

In addition, some variability across breeds will be observed, because of how the changes affect the breeds and because of characteristics of the breeds themselves. For example, there are differences in age structure and genetic architecture of the genotyped populations, as well as differences in pedigree depth and completeness. Brown Swiss is a special case, because there is a large proportion of foreign bulls in the genotyped population due to the strategic collaboration with BSW Intergenomics. Regarding pedigree inbreeding, a unique example is that some crossbred animals now have lower inbreeding values because the upward adjustment for missing pedigree information is now only applied when comparing ancestors of the same breed. Note that backcrosses and other more complicated crosses can still have non-zero inbreeding when the same breed is present on both sides of the pedigree.

Need for processing efficiencies

In the recent 12 months (March 2020 to March 2021), CDCB received 1 million new genotypes. We expect to receive the next million in even less time!

With such a rapid increase of data, more efficient methods to obtain those values are needed. The previous methodology required almost one day and a large amount of computer memory to complete. The new procedure performs the same calculations in less than an hour. Computing performance and the scaling investigation were conducted at CDCB by Dr. Gerald Jansen.

Are animal PTAs affected by this change?

The short answer is, no. CDCB uses the pedigree-based future inbreeding (EFI, or Expected Future Inbreeding) to correct PTAs for inbreeding.

None of these changes affect EFI; however, the same procedure that calculates genomic inbreeding also calculates the relationships of animals with the breed reference population. This relationship value is then used to estimate the reliability of the genomic evaluation. Since the PTA is a blend of traditional and genomic evaluation, some animals may have been affected indirectly because their genomic reliability changed. Animals most likely to be affected are those with incomplete or completely missing pedigree, or animals with pedigrees that have little in common with the U.S. reference population.

Except for these unique cases, there were no effects on PTAs for nearly all animals.

Why not use genomic inbreeding to correct PTAs?

The CDCB and AGIL team periodically receives questions about the use of genomic future inbreeding (GFI) instead of pedigree-based inbreeding to correct PTAs. These questions are logical. We have those values, and genomics is expected to be more precise than pedigree-based measures. We see that in the PTAs, so why not use in the inbreeding calculation?

The scaling issue is important because millions of cows in the current and historical database are not genotyped. Also, different ways of measuring and controlling genomic inbreeding can produce different outcomes, like the increase in frequency of recessives in the population. This is definitively not desired, yet these indications are preliminary, and more research is needed. Until we know more, we are continuing the pedigree-based measure to correct PTAs. Consider also that those pedigree-based measures are improved by genotyping, since the pedigree for genotyped animals is verified/corrected using genomics.

The bottom line

Genomic inbreeding values of males and females are now more comparable to each other and to pedigree inbreeding, and computation has been improved to more rapidly process the larger numbers of genotypes expected in the future.

⁽¹⁾ Nani, J.P. and VanRaden, P.M. [*Accounting for X chromosome and allele frequencies in genomic inbreeding estimation.*](#)

⁽²⁾ Druet, T., and Legarra, A.. 2020. *Theoretical and empirical comparisons of expected and realized relationships for the X-chromosome.* *Genet. Sel. Evol.* 52:50. doi:10.1186/s12711-020-00570-6.

Ezequiel Nicolazzi, PhD, is Chief Operating Officer at the Council on Dairy Cattle Breeding.

Gerald Jansen is Technical Advisor at the Council on Dairy Cattle Breeding.

Juan Pablo Nani, PhD, is Geneticist at Animal Genomics and Improvement Laboratory, Agricultural Research Service, United States Department of Agriculture.

Paul VanRaden, PhD, is Senior Research Geneticist at Animal Genomics and Improvement Laboratory, Agricultural Research Service, United States Department of Agriculture.