243. Genetic diversity and inbreeding in US Jersey population and germplasm collection

K. Srikanth[1*], M. Neupane[2], M.A. Jaafar[1], C.W. Wolfe[3], H.J. Huson[1], C.P. Van Tassell[2] and H.D. Blackburn[4]

[1]Department of Animal Science, Cornell University, 507 Tower Rd, Ithaca, NY 14853, USA; [2]Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, 10300 Baltimore Ave, Beltsville, 20705-2350, MD, USA; [3]American Jersey Cattle Association, 6486 E. Main St, Reynoldsburg, OH 43068, USA; [4]National Animal Germplasm Program, USDA, 111 South Mason St, Fort Collins, CO 80521, USA; ks786@cornell.edu

Abstract

Genetic diversity is needed to improve health, production traits and for breed conservation. Germplasm collection conserves diversity and enables population reconstruction or corrective mating as needed. The objective of this study was to evaluate genetic diversity in US Jersey cattle and identify animals for germplasm preservation to capture population diversity. Genotype and pedigree information on 36,179 Jersey bulls were obtained from the Council on Dairy Cattle Breeding (CDCB), including the 782 Jersey bulls preserved by the USDA-ARS National Animal Germplasm Program (NAGP). Genetic diversity indices (*Ho*, *He*, *MAF*) were similar in both the groups. Admixture results suggested some differences in the genetic composition between the groups. The pedigree and genomic inbreeding ranged between 7.68 (F_{PED}) and 15.77 (F_{ROH}) for CDCB animals, and between 6.30 (F_{PED}) and 14.43 (F_{ROH}) in the NAGP collected animals. The average correlation between genomic and pedigree inbreeding was 0.63, while between genomic inbreeding estimates was 0.94.

Introduction

Genetic diversity in cattle is associated with improvement of economically important traits, management decisions, and breed conservation. The excessive accumulation of inbreeding or homozygosity leads to genetic diversity loss. Pedigree-based studies have suggested an increased rate of inbreeding and loss of diversity in Canadian Jersey and Holstein cattle (Stachowicz *et al.*, 2011). Melka *et al.* (2012) found lower levels of diversity in Jersey relative to Holstein, and Brown Swiss bull populations and suggested crossbreeding to recover genetic diversity, if diversity within the breeds continues to get worse. The study of genetic diversity promotes understanding of evolution, breed differentiation, and gene pool within the breeds (Doekes *et al.*, 2020 and Huson *et al.*, 2020; Melka *et al.,* 2012). Cryopreservation of recent and historic germplasm is essential to improve management and production of cattle breeds and maintain breed diversity. The USDA-ARS National Animal Germplasm Program (NAGP) collection helps to maintain and dissipate this diversity and reconstruct it in the future if needed (Blackburn *et al.*, 2019). The objective of this study was to evaluate genetic diversity in the US Jersey breed and to identify additional animals for NAGP germplasm collection that would capture and preserve this diversity in cyropreservation. It also aims to assess distant and recent rates of inbreeding in US Jersey cattle.

Materials & methods

Study population. A total of 36,179 genotyped Jersey bulls born between 1950 and 2020 were used in this study. Genotype and pedigree information on these animals were obtained from the national cooperator database maintained by the Council on Dairy Cattle Breeding (CDCB, https://www.uscdcb.com/). This population included 782 NAGP Jersey sires born between 1953 and 2016. Animals were genotyped using 45 different arrays and were imputed to 79,389 markers using Findhap ver 3 (VanRaden *et al.,* 2011). The SNPs used in this study are routinely used in US dairy genomic evaluations. Bulls missing pedigree and phenotype records were excluded, and genotypes for 36,370 (35,675 CDCB and 695 NAGP) animals were used in the study (Table 1).

R.F. Veerkamp and Y. de Haas (eds)

DOI: 10.3920/978-90-8686-940-4_243, © K. Srikanth *et al.* 2022

Table 1. Distribution of samples used in this study.

Decade	CDCB	NAGP	Total
1951-1980	57	29	86
1981-1990	254	110	364
1991-2000	1,352	243	1,595
2001-2010	2,388	212	2,600
2011-2020	28,892	101	28,993
2021-	2,732	0	2,732
Total	35,675	695	36,370

Genetic relationship and diversity estimates. Basic diversity indices such as observed (*Ho*), expected heterozygosity (*He*) and minor allele frequencies (MAF) were calculated using PLINK ver 1.9. Principal component analysis was performed using GCTA ver 1.94. Admixture composition was estimated using a maximum likelihood model implemented in ADMIXTURE ver 1.23 (Alexander *et al.*, 2009).

Inbreeding estimation. Pedigree inbreeding (F_{PED}) was estimated using an additive genomic relationship matrix A using only pedigree information to calculate probabilities that gene pairs are identical by descent. Genomic inbreeding was estimated by two approaches: (1) Genomic relationship matrix (F_{GRM}) G (VanRaden *et al.*, 2008). Both pedigree and F_{GRM} were calculated using inbreed1.f90 and inbreed2.f90 implemented in US dairy evaluations. (2) Runs of homozygosity (F_{ROH}) was determined using a sliding window approach following the criteria suggested by (Doekes *et al.* 2019). Inbreeding coefficient was calculated based on the method proposed by McQuillan *et al.* 2008. ROH length based inbreeding was estimated by classifying the identified ROH segments into different length classes: (1) >16 Mb; (2) 8-16 Mb; (3) 4-8 Mb; (4) 2-4 Mb; and (5) 1-2 Mb.

Results & discussion

Genetic diversity between bulls in the CDCB population and NAGP germplasm collection. The average observed heterozygosity, expected heterozygosity and minor allele frequencies in the two populations were similar (Table 2). Principal component analysis revealed considerable variation in the populations, particularly in recent decades (Figure 1). This result could reflect that most of the sires genotyped are from the recent decades (80% of the genotyped bulls were born from 2011 to 2020). ADMIXTURE results (K=2) suggested differences in composition between the animals in the NAGP collection and the wider population (CDCB), with lower levels of admixture in the NAGP collection (Figure 2a). Inclusion of purebred Holstein data in the analysis (K=3) suggested Holstein introgression for CDCB (8.9%) and NAGP (5.4%).

Table 2. Genetic diversity indices observed in the population and germplasm collection.

Group	*Ho*[1] ± Std	*He*[2] ± Std	MAF[3] ± Std
CDCB	0.37±0.12	0.38±0.13	0.28±0.14
NAGP	0.38±0.11	0.39±0.11	0.29±0.12

[1] Ho – observed heterozygosity.
[2] He – expected heterozygosity.
[3] MAF- minor allele frequency.

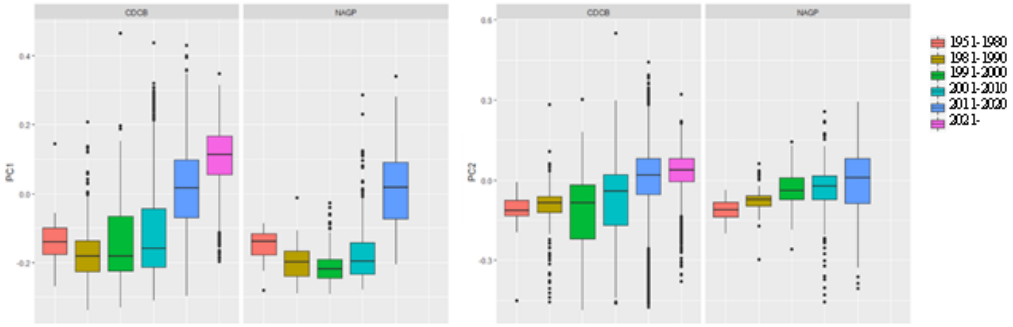

Figure 1. Box plot showing variation captured by Principal components (PC1 and PC2).

Genomic and pedigree inbreeding estimation. The average pedigree inbreeding was 6.3% in the NAGP and 7.68% in the CDCB populations. The genomic inbreeding estimates were 11.12%, 10.16% (F_{GRM}) and 15.77%, 14.43% (F_{ROH}) for CDCB and NAGP collections, respectively. ROH length-based inbreeding ranged between 3.8 and 15.2% for CDCB and 3.94 and 14.43% in NAGP animals (Table 3). The correlation between genomic and pedigree inbreeding ranged between 0.56 to 0.68 with genomic inbreeding estimates ranging between 0.87 to 0.97 (Figure 2b). The average pedigree and genomic inbreeding in NAGP bulls was 10.2 and 6.5 respectively (Table 3). This estimation was similar to previous reports in Jersey (Huson *et al.* 2020). The lack of pedigree completeness, and the random nature of recombination and segregation which pedigree-based measures do not capture, could be a contributing factor for low pedigree inbreeding estimates (Doekes *et al.* 2019). Inbreeding measures based on diagonal elements of the GRM do not distinguish between identity by state (IBS) and identity by descent (IBD) and could lead to inconsistent results (Pryce *et al.* 2014, Villanueva *et al.* 2021). Meuwissen *et al.* 2020, showed that in contrast to classical inbreeding theory, inbreeding measured in terms of drift (F_{DRIFT}) and homozygosity (F_{HOM}) can substantially vary, when inbreeding measures are based on GRM, derived using IBS (F_{GRM}), but not when derived using IBD (F_{ROH}). This could explain the differences in genomic inbreeding measures in this study and based on these reports, F_{ROH} might capture inbreeding more accurately than F_{GRM}. The information generated in this study can be used to improve Jersey diversity, health, and production. It will inform the US Jersey Association of the population status on diversity and inbreeding. Analyses to date suggest the *in situ* population and the NAGP collection have equivalent levels of genetic diversity, but further analysis will better identify gaps in the collection.

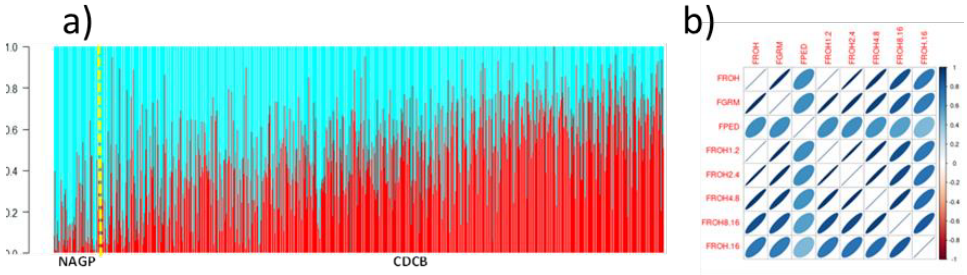

Figure 2. (a) Admixture analysis showing compositional difference between NAGP collection and CDCB population at K=2. (b) Pearson correlation between different inbreeding measures.

Table 3. Comparisons between different inbreeding measures.

Inbreed estimator	CDCB ± SD	NAGP ± SD
F_{PED}	7.68±2.95	6.30±3.16
F_{GRM}	11.12±4.12	10.16±4.15
F_{ROH}	15.77±3.63	14.43±3.73
$F_{ROH\ 1\text{-}2}$	15.23±3.41	14.64±3.02
$F_{ROH\ 2\text{-}4}$	12.56±3.70	12.38±3.83
$F_{ROH\ 4\text{-}8}$	9.34±3.66	9.44±3.73
$F_{ROH\ 8\text{-}16}$	7.16±3.33	7.06±3.41
$F_{ROH\ >16}$	3.80±2.47	3.94±2.59

References

Alexander D.H., Novembre J., and Lange K. (2009) Genome Research 19(9):1655-1664. https://doi.org/10.1101/gr.094052.109.

Blackburn H.D., Wilson C.S., and Krehbiel B. (2019) Diversity 11(12):244. https://doi.org/10.3390/d11120244.

Doekes H.P., Bijma, P., Veerkamp, R.F., de Jong, G., *et al.* (2020) Genetics Selection Evolution 52(1):1-18. https://doi.org/10.1186/s12711-020-00583-1.

Huson H.J., Sonstegard T.S., Godfrey J., Hambrook D., Wolfe C., *et al.* (2020) Frontiers in genetics 11:366. https://doi.org/10.3389/fgene.2020.00366.

McQuillan R., Leutenegger A.L., Abdel-Rahman R., Franklin C.S., Pericic, M., *et al.* (2008) The American Journal of Human Genetics 83(3):359-372. https://doi.org/10.1016/j.ajhg.2008.08.007

Melka M.G., and Schenkel F.S. (2012) BMC Research Notes 5(1): 1-8. https://doi.org/10.1186/1756-0500-5-161.

Meuwissen T. H., Sonesson A. K., Gebregiwergis, G., and Woolliams, J. A. (2020) Frontiers in genetics 11:880. https://doi.org/10.3389/fgene.2020.00880.

Pryce J. E., Haile-Mariam M., Goddard M. E., and Hayes, B. J. (2014). Genetics Selection Evolution, 46(1): 1-14. https://doi.org/10.1186/s12711-014-0071-7.

Stachowicz K., Sargolzaei M., Miglior F., and Schenkel F.S. (2011) Journal of dairy science 94(10):5160-5175. https://doi.org/10.3168/jds.2010-3308.

VanRaden PM., (2008) Journal of Dairy Science 91:4414–4423. https://doi.org/10.3168/jds.2007-0980.

Villanueva B., Fernández A., Saura, M., Caballero, A., Fernández, J., *et al.* (2021) Genetics Selection Evolution 53(1): 1-17.