López de la Torre, R., Fernández, J., García-Ballesteros, S., Morales-González, E., Fernández, A., Meuwissen, T., Caballero, A., Toro, M., Villanueva, B. (2026) Impact of using different coancestry matrices in optimal contribution management on effective population size, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23713: https://doi.org/10.31274/wcgalp.23713
©2026 López de la Torre, et al. CC BY-NC-ND

Impact of using different coancestry matrices in optimal contribution management on effective population size

R. López de la Torre[1,2*], J. Fernández[1], S. García-Ballesteros[1], E. Morales-González[1], A. Fernández[1], T.H.E. Meuwissen[3], A. Caballero[4], M.A. Toro[5], and B. Villanueva[1]

[1]INIA-CSIC, Madrid, Spain; [2]EDUAM, Madrid, Spain. [3]NMBU, Ås, Norway; [4]CIM, Vigo, Spain; [5]UPM, Madrid, Spain; [*]Corresponding author: raul.lopez@inia.csic.es

Introduction

The main objectives of animal conservation programs are to preserve genetic diversity and to prevent inbreeding depression. Optimal Contribution (OC) management (Meuwissen, 1997; Fernández *et al*., 2003) is recognized to be the most effective strategy to achieve these goals by minimizing the global coancestry between potential parents weighted by their contributions, making the coancestry matrix a central component of the methodology.

While early applications of OC relied on pedigree-based coancestry, the availability of dense SNP data has promoted the use of genomic coancestry matrices. Genomic information can improve the accuracy of realized coancestry estimates, capture coancestry from distant ancestors and enables OC management in populations without reliable pedigree information. Accordingly, several genomic estimators of coancestry have been proposed in the literature.

The performance of OC management can be evaluated through the effective population size (N_e), which reflects the extent of genetic drift, inbreeding rate, and loss of genetic variation (Caballero, 2020). N_e is therefore a key parameter for assessing population genetic status (Frankham *et al*., 2010). Despite its importance, few studies have examined how the use of genomic coancestry matrices in OC management affects the N_e achieved. A notable exception is the study of Toro *et al*. (2019), which questioned the classical concept of N_e when parental contributions are optimized using genomic relationships. However, their analysis relied on a single genomic coancestry estimator and treated pedigree-based N_e as the true reference. Rather than assuming pedigree-based N_e as a benchmark, N_e derived from identity-by-descent (IBD) coancestry provides a more accurate definition of true N_e (Caballero *et al*., 2022).

The objective of this study was to evaluate, through stochastic simulations, the evolution of true IBD-based N_e when different genomic coancestry matrices are used within the OC framework, in order to identify the most suitable matrix for effective management of populations under conservation programs.

Materials & Methods

Base population. Management through the OC method was initiated from a base population (t = 0) sampled from an initial population at mutation-drift equilibrium, composed by 25,600 individuals, from which 200 replicates were generated. Each replicate consisted of a randomly sampled base population of N = 100 individuals (50 males and 50 females). On average, approximately 51,500 loci per replicate were segregating and treated as single nucleotide polymorphisms (SNPs) for computing genomic coancestry matrices. Additionally, at t = 0, 10,000 multiallelic loci were simulated, with each locus assigned two unique alleles per individual to create fully informative markers. These loci were tracked across subsequent

generations, allowing the estimation of IBD probabilities, and used to compute true coancestry and inbreeding coefficients.

Population management. Populations were managed for 20 generations using the OC method, which determines the optimal genetic contribution of each candidate to minimize global coancestry. The optimization problem was solved using Lagrange multipliers, following the approach of Meuwissen (1997). Random mating was performed among the selected individuals to produce the next generation. Offspring were assigned sexes at random, ensuring equal numbers of males and females in each generation.

Coancestry matrices. Nine management strategies were evaluated, differing only in the coancestry matrix used in OC. These included the true IBD-based matrix ($\boldsymbol{\theta}_{\mathbf{IBD}}$), the pedigree-based matrix ($\boldsymbol{\theta}_{\mathbf{PED}}$), and seven genomic coancestry matrices: those described by Li and Horvitz ($\boldsymbol{\theta}_{\mathbf{LH1}}$: Li and Horvitz, 1953; $\boldsymbol{\theta}_{\mathbf{LH2}}$: Toro *et al*., 2002), VanRaden ($\boldsymbol{\theta}_{\mathbf{VR1}}$: VanRaden, 2008, $\boldsymbol{\theta}_{\mathbf{VR2}}$: Amin *et al*., 2007), Yang and colleagues ($\boldsymbol{\theta}_{\mathbf{YA1}}$: Zhang *et al*., 2022; $\boldsymbol{\theta}_{\mathbf{YA2}}$: Yang *et al*., 2010) and the linkage-based matrix $\boldsymbol{\theta}_{\mathbf{GLA}}$, which combined pedigree and genomic information and was constructed as described by Meuwissen *et al*. (2020). Genomic matrices were computed using allele frequencies at $t = 0$.

Comparison of management strategies. Management strategies were compared based on the reference IBD-based N_e achieved across generations. The reference N_e was calculated from the rate of increase in IBD coancestry (Δf_{IBD}) using the relationship $N_e = 1/(2\Delta f_{IBD})$. The f_{IBD} values were obtained from $\boldsymbol{\theta}_{\mathbf{IBD}}$ matrix.
In addition, strategies were compared in terms of the evolution of different genomic inbreeding coefficients (F): the reference IBD-based inbreeding coefficient and the estimated genomic inbreeding coefficients corresponding to different matrices used in the optimization.

Results

Clear differences were observed depending on the coancestry matrix used in the optimization (Figure 1). Management based on $\boldsymbol{\theta}_{\mathbf{GLA}}$ consistently produced the highest N_e values across generations, closely matching the trajectory obtained under true IBD-based management and reaching values close to 220 in later generations. Management based on $\boldsymbol{\theta}_{\mathbf{PED}}$ resulted in constant N_e values around 200, but still outperformed purely genomic coancestry matrices. Among the remaining genomic matrices, $\boldsymbol{\theta}_{\mathbf{VR2}}$ and $\boldsymbol{\theta}_{\mathbf{YA2}}$ yielded the highest N_e values (~185 across generations), followed by $\boldsymbol{\theta}_{\mathbf{VR1}}$ and $\boldsymbol{\theta}_{\mathbf{YA1}}$ (~170 across generations). In contrast, management based on $\boldsymbol{\theta}_{\mathbf{LH1}}$ and $\boldsymbol{\theta}_{\mathbf{LH2}}$ resulted in substantially lower N_e (~50 and ~25 at t = 1, respectively, and rising to ~125 and ~75 at the end of the simulations).
For populations lacking reliable pedigree information, $\boldsymbol{\theta}_{\mathbf{VR2}}$ and $\boldsymbol{\theta}_{\mathbf{YA2}}$ emerged as the most suitable alternatives among purely molecular matrices, as both produced the highest N_e values within this group. Because true IBD-based coefficients are not available in real populations, monitoring necessarily relies on the same estimator used for management. In the same set of simulations, the corresponding estimated inbreeding coefficient under $\boldsymbol{\theta}_{\mathbf{YA2}}$-based management ($F_{YA2}$) increased steadily across generations, although values remained consistently lower than those obtained for true IBD-based inbreeding (Figure 2a). In contrast, $\boldsymbol{\theta}_{\mathbf{VR2}}$-based management showed a decrease of the inbreeding coefficient F_{VR2} in later generations and a negative correlation with IBD-based inbreeding (–0.57; Figure 2b).

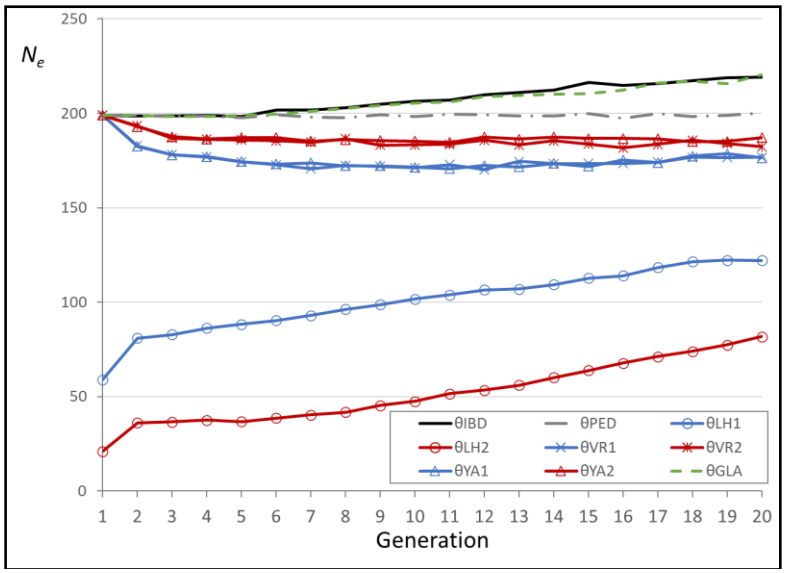

Figure 1. Reference IBD-based N_e across generations, under different management strategies defined by the coancestry matrix used to optimize contributions.

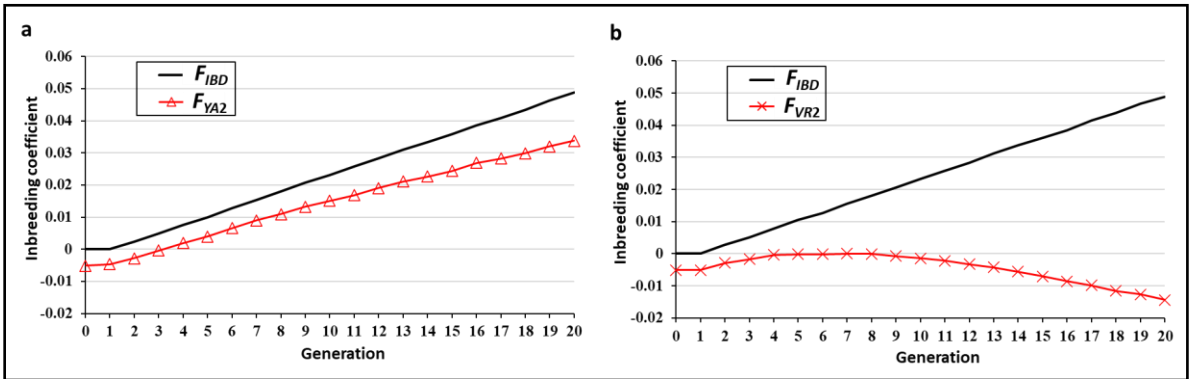

Figure 2. Evolution across generations of the reference IBD-based inbreeding coefficient (F_{IBD}) and the genomic inbreeding coefficient corresponding to the matrix used in the optimization of contributions. Panel (a): F_{IBD} and F_{YA2} when $\boldsymbol{\theta}_{\mathbf{YA2}}$ was used in the optimization. Panel (b): F_{IBD} and F_{VR2} when $\boldsymbol{\theta}_{\mathbf{VR2}}$ was used in the optimization.

Discussion

This study identified $\boldsymbol{\theta}_{\mathbf{GLA}}$ and $\boldsymbol{\theta}_{\mathbf{PED}}$ as the matrices that most closely reflect true IBD relationships. Both $\boldsymbol{\theta}_{\mathbf{GLA}}$ and $\boldsymbol{\theta}_{\mathbf{PED}}$ are derived from pedigree information and therefore intend to reflect IBD relationships. Starting from unrelated founders, $\boldsymbol{\theta}_{\mathbf{PED}}$ remains homogeneous under OC management, which results in equalized parental contributions, making N_e to remain close to $2N$. However, management based on $\boldsymbol{\theta}_{\mathbf{GLA}}$ consistently produced higher N_e, because $\boldsymbol{\theta}_{\mathbf{GLA}}$ combines pedigree and marker information, allowing the differentiation of realized coancestry among individuals sharing identical pedigree relationships. This leads to a reduction in the variance of the Mendelian segregation (Wang and Hill, 2000) and enables N_e to exceed $2N$, approaching the IBD-based trajectory. Similarly, Meuwissen et al. (2020) reported that $\boldsymbol{\theta}_{\mathbf{GLA}}$-based management provides an advantage in maintaining genetic diversity in selected populations.

Purely molecular matrices showed poorer overall performance, as they attempt to infer IBD correcting by base population allele frequencies but may fail to adequately distinguish IBD from identity-by-state. Drift-based genomic matrices ($\boldsymbol{\theta}_{\mathbf{VR1}}$, $\boldsymbol{\theta}_{\mathbf{VR2}}$, $\boldsymbol{\theta}_{\mathbf{YA1}}$ and $\boldsymbol{\theta}_{\mathbf{YA2}}$) generally led to higher N_e than homozygosity-based matrices ($\boldsymbol{\theta}_{\mathbf{LH1}}$ and $\boldsymbol{\theta}_{\mathbf{LH2}}$), as the former group tends to

maintain a larger number of contributors and reduce the loss of rare alleles. In contrast, homozygosity-based matrices strongly reduced the contributor´s pool in early generations, leading to a sharp initial decline in N_e. In later generations, N_e gradually increased (Figure 1) as allele frequencies became more balanced, allowing a larger proportion of individuals to contribute.

The discrepancies observed in the evolution of estimated inbreeding against the true IBD inbreeding (Figure 2) indicate that $\boldsymbol{\theta}_{\mathbf{YA2}}$ provides more coherent monitoring than $\boldsymbol{\theta}_{\mathbf{VR2}}$. Although both estimators underestimated inbreeding relative to IBD-based values, under $\boldsymbol{\theta}_{\mathbf{YA2}}$ management the estimated inbreeding increased steadily over generations. In contrast, under $\boldsymbol{\theta}_{\mathbf{VR2}}$ management, estimated inbreeding followed a decreasing trajectory that is inconsistent with true IBD-based inbreeding.

Conclusions

These results indicated that pedigree-based OC remains an effective strategy when accurate pedigree information is available, but that its performance can be further improved by incorporating genomic data. Management based on $\boldsymbol{\theta}_{\mathbf{GLA}}$ consistently produced the highest N_e and closely followed the trajectory obtained under $\boldsymbol{\theta}_{\mathbf{IBD}}$-based management, making it the most powerful strategy for maximizing N_e in conservation programs. For populations lacking reliable pedigree information, $\boldsymbol{\theta}_{\mathbf{YA2}}$ emerged as the most suitable alternative.

Acknowledgements

This work was funded by MCIN/AEI/10.13039/501100011033, Spain (grants FPI PRE2021-096997 and PID2020-114426GB-C2). We thank the 'Centro de Supercomputación de Galicia' (Cesga) for their computing support.

References

Amin N., van Duijn C.M., and Aulchenko Y.S. (2007) PLOS One 2:e1274. doi.org/10.1371/journal.pone.0001274

Caballero A. (2020) Quantitative genetics. Cambridge: Cambridge University Press.

Caballero A., Fernández A., Villanueva B., *et al*. (2022) Genet. Sel. Evol. 54:82. doi.org/10.1186/s12711-022-00772-0

Fernández J., Toro M.A., and Caballero A. (2003) Genetics 165(2):885-894. doi.org/10.1093/genetics/165.2.885

Frankham R., Ballou J.D., and Briscoe D.A. (2010) Introduction to Conservation Genetics. Cambridge University Press, Cambridge.

Li C.C., and Horvitz D.G. (1953) Am. J. Hum. Genet. 5(2):107–117. PMID: 13065259

Meuwissen T.H.E. (1997) J. Anim. Sci. 75(4): 934–940. doi.org/10.2527/1997.754934x

Meuwissen T.H.E., Sonesson A.K., Gebregiwergis G., *et al*. (2020) Front. Genet. 11:880. doi.org/10.3389/fgene.2020.00880

Morales-González E., Fernández J., Pong-Wong R., *et al*. (2021) Genes 12(5):673. doi.org/10.3390/genes12050673

Wang J., and Hill W.G. (2000) Genetics 154(1):475–489. https://doi.org/10.1093/genetics/154.1.475

Toro M.A., Villanueva B., and Fernández J. (2019) J. Anim. Breed. Genet. 137(4):345-355. doi.org/10.1111/jbg.12455

VanRaden P.M. (2008) J. Dairy Sci. 91(11):4414-4423. doi.org/10.3168/jds.2007-0980

Yang J., Benyamin B., McEvoy B.P., *et al*. (2010) Nat. Genet. 88:76-82. doi.org/10.1038/ng.608

Zhang Q., Goudet J., and Weir B.S. (2022) Heredity. 128:1–10. doi.org/10.1038/s41437-021-00471-4