

# Simulating genetic progress for traits with expensive phenotyping

J. Jenko<sup>1</sup>, A. Ehsani<sup>2</sup> and G. Gorjanc<sup>2</sup>

<sup>1</sup> Geno SA, Storhamargata 44, 2317 Hamar, Norway

<sup>2</sup> The Roslin Institute and Royal (Dick) School of Veterinary Medicine, The University of Edinburgh, Edinburgh, UK

Corresponding author: [janez.jenko@geno.no](mailto:janez.jenko@geno.no)

---

## Abstract

Phenotyping costs in dairy cattle breeding exhibit significant variability across traits. While milk production is recorded routinely at only low costs, traits such as feed efficiency and methane emissions pose challenges due to their expensive measurement requirements. This study leveraged the real-size digital twin of the Geno breeding program for the Norwegian Red dairy cattle breed to simulate genetic progress following ten years of selective breeding, particularly targeting traits demanding costly phenotyping. Multiple scenarios were simulated, varying in the number of phenotypes recorded, economic weight, and genetic correlation between the trait and total merit index. Our results highlight the importance of genetic correlation in achieving progress for traits with expensive phenotypes recorded at a limited scale. Increasing economic weight and the number of phenotypes increased genetic progress. Thus, there is an indirect indication that traits with low phenotyping costs and high correlation to expensive phenotypes should be prioritized when selecting for genetic improvement of a trait with expensive phenotypes. However, precise phenotypes are required for accurately estimating genetic correlations between traits with expensive phenotypes and traits with cheap phenotyping.

**Key words:** genomic selection, phenotyping, Norwegian Red Dairy Cattle, breeding program, digital twin, future genetic progress

---

## Introduction

Since the introduction of genomic selection in the Norwegian Red Dairy Cattle (NR) breeding program in 2016, single step genomic prediction approach was used (Nordbø et al., 2019). This method integrates pedigree and genotype information into a single relationship matrix, allowing the inclusion of all the individuals with phenotype and genotype information in the reference (Christensen and Lund, 2010). Consequently, the reference population comprises progeny tested bulls and phenotyped cows, enhancing the accuracy of predicted breeding values (Legarra et al., 2014).

As of 2024, the reference population of the NR breeding program includes approximately 100 000 animals for production traits and 47 000 animals for conformation traits. While

phenotype data for production traits are collected routinely at low costs, the recording of type traits has a long history and incurs intermediate cost. However, recording for enteric methane emission and feed efficiency began only recently, resulting in smaller reference populations for these traits (Heringstad and Bakke, 2023). Due to high cost recording these traits, it will take significantly more time to establish a reference population sufficient for predicting highly accurate genomic breeding values.

This study aims to demonstrate the potential of genomic selection in the NR breeding program for traits with expensive phenotyping by utilizing the digital twin of Geno's breeding program (Ehsani et al., 2022). We analysed the effect of the reference population size, different economic weights,

and the correlation between the selection index and the expensive trait.

### Materials and Methods

A real-size breeding program was simulated beginning with fifty years of historical breeding (from 1971 to 2020), followed by ten years of alternative future breeding scenarios. The future scenarios differed in the number of phenotypes collected annually (1000, 2000, or 3000), the economic weight of an expensive trait in the future selection index (0.2 or 0.5), and the genetic correlation between the index trait and the expensive trait.

reflecting the estimated heritability of enteric methane emission and dry matter intake in NR (Heringstad and Bakke, 2023).

For the first 45 years of historical breeding, breeding values were predicted using only pedigree data. From 2016 onward, pedigree information was combined with genotype information into a single step genomic evaluation approach. Genomic breeding values were calculated using the singular value decomposition method (Ødegård et al., 2018). In this method, chromosome specific principal components explained 98% of genetic variance among the 20 000 core individuals. This core group included genotypes from all progeny

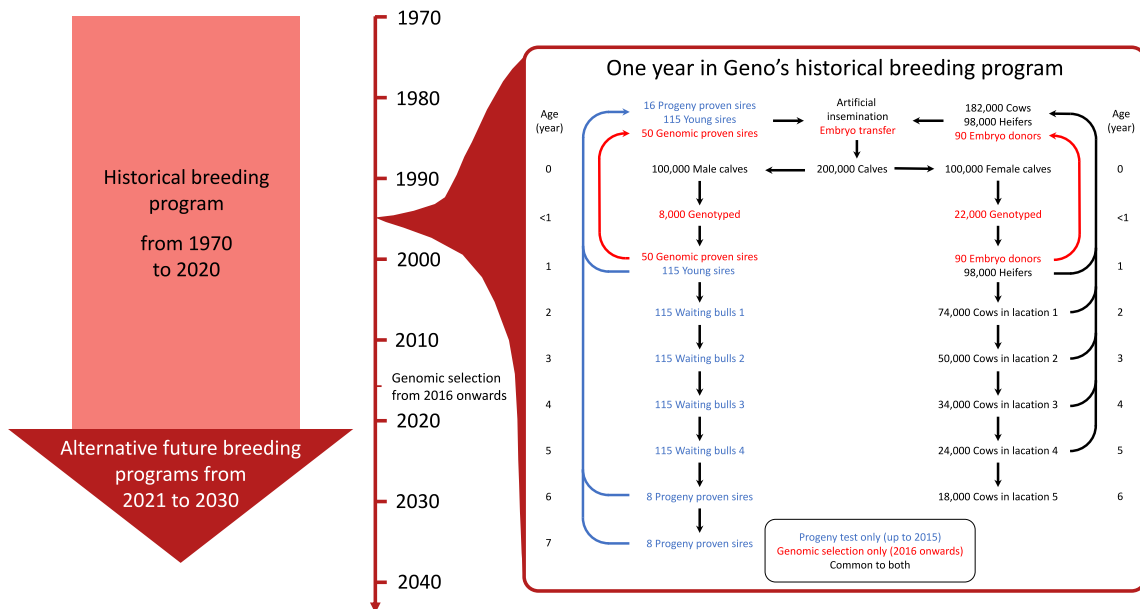


Figure 1. Fifty years of historical breeding program of Norwegian Red Dairy Cattle breed followed by ten years of future breeding as integrated in digital twin of Geno breeding program

During the historical breeding period, animals were selected based on the predicted breeding values for the index trait. For simplicity, this trait was represented milk yield ( $h^2=0.192$ ), one of the most important traits in the history of NR breeding program. In the future breeding scenarios, animals were selected based on the future selection index, which combined the breeding values of the index trait and the expensive trait according to the specified economic weights. The heritability of the expensive trait was set at 0.3,

proven bulls while the rest were genotypes from the cows with available phenotype data.

The estimated breeding values (rEBV) from the final year of future breeding were standardised (EBV) so that:  $EBV = m + k * (rEBV - mEBV) / sEBV$ , where  $m = 100$ ,  $k = 12$ ,  $mEBV$  is the mean breeding value of females born between 2023 and 2028, and  $sEBV$  is the standard deviation of the bulls breeding values born between April 2011 and March 2016. Each scenario was run in ten replicates, and the mean genetic progress was

calculated across the ten replicates for each year of selection. The different scenarios were compared based on the achieved genetic progress after then years of future breeding.

## Results & Discussion

We present the achieved genetic progress in the final year of future breeding separately for the index trait and the expensive trait in each simulated scenario, as shown in Table 1. We will begin by analysing the effect of increasing the number of phenotypes. Next, we will examine the impact of assigning a higher economic weight to the expensive trait. Following this, we will explore the effect of varying the correlation between the index trait and the expensive trait. Lastly, we will compare the changes when two or three parameters are simultaneously adjusted

Table 1. Genetic progress ( $\Delta G$ ) in ten years of future breeding for index trait (IT) and expensive trait (ET) with different number of phenotypes (N) for ET, different economic weights (EW) for ET and IT in the future selection index and different genetic correlations between IT and ET ( $r_g$ )

N for ET	EW for ET*	$r_g$	$\Delta G$ for IT	$\Delta G$ for ET
1000	0.2	0	64.1	2.7
2000	0.2	0	63.9	3.7
3000	0.2	0	63.4	6.2
1000	0.5	0	58.7	11.6
2000	0.5	0	57.2	15.1
3000	0.5	0	56.0	17.0
1000	0.2	0.3	64.4	20.9
1000	0.2	0.6	64.2	38.6
1000	0.5	0.3	61.9	26.4
1000	0.5	0.6	63.3	40.7
2000	0.2	0.3	64.5	21.9
2000	0.2	0.6	64.4	39.5
2000	0.5	0.3	60.8	28.8
2000	0.5	0.6	62.6	42.6
3000	0.2	0.3	64.3	22.1
3000	0.2	0.6	64.4	39.9
3000	0.5	0.3	59.8	30.3
3000	0.5	0.6	62.1	43.4

\*EW for IT is: 1 – EW for ET

### *Effect of increasing phenotype numbers*

Collecting a higher number of phenotypes for the expensive trait slightly decreased the

genetic gain for the index trait while increasing the gain for the expensive trait. When 1000 phenotypes were collected each year during the future breeding period, with the economic weight for the expensive trait set at 0.2 and no correlation between the traits, the genetic progress achieved was 64.1 for the index trait and 2.7 for the expensive trait. As the number of phenotypes increased to 2000 per year, the genetic progress for the index trait dropped slightly to 63.9, while the gain for the expensive trait rose to 3.7. With 3000 phenotypes per year, the genetic progress further declined to 63.4 for the index trait, but it increased to 6.2 for the expensive trait. This indicates that doubling the number of phenotypes for the expensive trait does not result in a proportional increase in its genetic progress.

### *Impact of economic weight*

An increase in the economic weight of the expensive trait in the future selection index reduced the genetic gain of the index trait but enhanced the gain for the expensive trait. Specifically, when the economic weight of the expensive trait was raised from 0.2 to 0.5 in the future selection index, the genetic progress over ten years of future breeding fell from 64.1 to 58.7 for the index trait, while it rose significantly from 2.7 to 11.6 for the expensive trait. This indicates that increasing the economic weight by two and a half times leads to more than a fourfold increase in the genetic progress of the expensive trait, while the genetic progress for the index trait decreases by only 8.4%.

### *Effect of genetic correlation*

A higher positive genetic correlation between the index trait and the expensive trait had no significant effect on the genetic improvement of the index trait, but it strongly enhanced the genetic progress of the expensive trait. When the genetic correlation between the two traits increased from 0 to 0.3 and then to 0.6, the genetic progress of the index trait remained

relatively stable, at 64.1, 64.4, and 64.2. In contrast, the genetic progress for the expensive trait saw substantial increases: from 2.7 with no correlation to 20.9 at a correlation of 0.3, and further rising to 38.6 when the correlation was 0.6.

### ***Combined parameter adjustments***

Increasing the number of phenotypes for the expensive trait, raising its economic weight, and having a higher genetic correlation between the index and expensive traits positively impacted the realized genetic gain of the expensive trait. The highest genetic gain of 43.4 for the expensive trait occurred in the scenario where 3000 phenotypes were collected, the economic weight was 0.5, and the genetic correlation with the index trait was 0.6. This gain is more than sixteen times greater than the baseline scenario, where only 1000 phenotypes were collected per year, the economic weight was 0.2, and there was no genetic correlation between the traits.

Among the three parameters analysed, the genetic correlation between the index trait and the expensive trait had the greatest impact on the genetic gain of the expensive trait. Therefore, identifying a phenotype with lower recording costs but a strong genetic correlation to the expensive trait could be a viable strategy for future trait improvement. Since building a reference population for the expensive trait requires many years, continuing to collect precise phenotypes is crucial. This allows for the accurate estimation of genetic correlations between traits with costly phenotyping and those with cheaper phenotyping. Thus, even on a smaller scale, ongoing phenotyping for the expensive traits is justified.

### **Conclusions**

Our findings underscore the critical role of genetic correlation in enhancing genetic progress for traits with expensive phenotypes, especially when phenotyping is limited. Increasing the economic weight assigned to

these traits in the selection index, along with the number of phenotypes collected, significantly boosts genetic gains. This suggests that traits with lower phenotyping costs but strong genetic correlations to expensive traits should be prioritized in breeding programs to achieve indirect genetic improvements in costly traits.

Moreover, the study highlights the necessity of obtaining precise phenotypes to accurately estimate genetic correlations between expensive and inexpensive traits. This precision is essential for developing cost-effective strategies in future breeding efforts aimed at enhancing traits with prohibitive phenotyping costs.

### **Acknowledgments**

A. Ehsani and G. Gorjanc would like to acknowledge the contribution from: HORIZON2020 through the TRAIN@Ed programme and funded under EXCELLENT SCIENCE – Marie Skłodowska-Curie Actions; and the BBSRC ISP grant to The Roslin Institute (BBS/E/D/30002275, BBS/E/RL/230001A, BBS/E/RL/230001C); and The University of Edinburgh.

J. Jenko would like to thank Norwegian Research Council through the project 330747, «Data Driven Breeding of Resource Efficient Cattle» under the SkatteFUNN scheme and the project 346741, «NrfTwin: Advancing Sustainable Breeding of Norwegian Red Dairy Cattle through Digital Twin» funded by the Research Council of Norway.

### **References**

- Christensen, O.F. and Lund, M.S. 2010. Genomic prediction when some animals are not genotyped. *Genet. Sel. Evol.* 42 (2). <https://doi.org/10.1186/1297-9686-42-2>
- Ehsani, A., Jenko, J., Storlien, H., Tajet, H.M., Pocrnic, I., Obsteter, J. and Gorjanc, G. 2022. Full-scale simulation of the Norwegian red dairy cattle breeding

- programme to enable testing future breeding directions, *Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP)*. Wageningen Academic Publishers, 1934–1937. [https://doi.org/10.3920/978-90-8686-940-4\\_465](https://doi.org/10.3920/978-90-8686-940-4_465)
- Heringstad, B. and Bakke, K. 2023. Heritability of methane emission in young Norwegian Red bulls estimated from GreenFeed measures at the test station. *Interbull Bull.* 59, 69–73. <https://journal.interbull.org/index.php/ib/article/view/1881/1909>
- Legarra, A., Christensen, O.F., Aguilar, I. and Misztal, I. 2014. Single Step, a general approach for genomic selection. *Livest. Sci.* 166, 54–65. <https://doi.org/10.1016/j.livsci.2014.04.029>
- Nordbø, Ø., Gjuvsland, A.B., Eikje, L.S. and Meuwissen, T. 2019. Level-biases in estimated breeding values due to the use of different SNP panels over time in ssGBLUP. *Genet. Sel. Evol.* 51 (76). <https://doi.org/10.1186/s12711-019-0517-z>
- Ødegård, J., Indahl, U., Strandén, I. and Meuwissen, T.H.E. 2018. Large-scale genomic prediction using singular value decomposition of the genotype matrix. *Genet. Sel. Evol.* 50 (6). <https://doi.org/10.1186/s12711-018-0373-2>